Topics in Number Theory

taught by Manjul Bhargava and Jerry Wang

February 26, 2016

Introduction

Notes from Manjul Bhargava and Jerry Wang's class "Topics in Number Theory" at Princeton in Spring 2016.

Chapter 1

Overview

We'll talk about statistics for the number of rational points on curves, usually algebraic curves over $\mathbb{Q}$. (Later on we may change to another global field.)

The main question is the following:

- Given a "family" of algebraic curves over $\mathbb{Q}$ (ordered in some way), what "proportion" of those curves have a rational point? How many rational points do we expect on those curves?

The families will usually be given in terms of explicit equations with varying coefficients in $\mathbb{Z}$, giving these curves in projective space $\mathbb{P}^n$.

We'll order these curves by the maximum of absolute values of coefficients in $\mathbb{Z}$, called the **height** of the equation(s)/curve. (If it's a weighted projective space, then we have to take those weights into account.) We'll abuse langauge and refer to the equations as the curve (they're actually a model of the curve).

- When the curves in a family are ordered by height, what is the distribution of the number of rational points?

There are three types of behavior for rational points on curve.

Trichotomy:

1 Genus 0

Genus 0: A genus 0 curve over $\mathbb{Q}$ has either no rational points or its rational points are in bijection with $\mathbb{P}^1(\mathbb{Q})$. The key principle that makes it easier is that the **local-global principle** holds. A genus zero curve has rational iff it has $\mathbb{Q}_p$ poitns for all p and $\mathbb{R}$-points. Checking for $\mathbb{Q}_p, \mathbb{R}$ points is much easier. We just have to understand which curves locally have points everywhere; those curves globally have points.

Example 1.1.1 (Circles)**:** Consider $C_n : x^2 + y^2 = nz^2$ in $\mathbb{P}^2$. C_n has a rational point iff n is a sum of 2 squares iff the factorization of n has only even exponents for primes $\equiv 3 \pmod 4$.

The natural density of such n is 0. The average number of prime factors is infinite. The chance that none of the odd exponent primes are $\equiv 3 \pmod 4$ goes to 0.

This is not a good way of producing rational points.

We're not asking the right question. How 0 is 0—i.e. how fast is it approaching 0? The more precise answer is

$$|\{n < X : n \text{ is sum of 2 squares}\}| \sim \frac{cX}{\sqrt{\ln X}}.$$

This says how fast the density approaches 0.

It's easy to get unsolved problems, for example, if we allow arbitrary coefficients in front of x.

Example 1.1.2 (Diagonal conics)**:** Consider diagonal conics $C_{a,b,c} : ax^2 + by^2 + cz^2 = 0$ in $\mathbb{P}^2$. The density of $C_{a,b,c}$ having a rational point is 0. Asymptotics are different depending on your family.

Proof. We handle genus 0 curves with the local-global principle. By Chinese remainder theorem we can multiply a finite number of primes. To prove the proportion is 0, it suffices to show that when we take more and more primes, the density goes to 0. We're not using any infinite version of the CRT.

Take $p > 2$. How can we tell if it has a point over $\mathbb{F}_p$? A smooth conic always has a rational point. If How do we tell if the conic is smooth? Look at the determinant, it's smooth iff it's nonzero. The determinant is abc. If it's smooth it has a point over $\mathbb{F}_p$, by Hensel's lemma it lifts to a $\mathbb{Z}_p$ point.

The only way it has no points is if it's not smooth. If the conic is not smooth, it breaks up into 2 points. If the 2 lines are over $\mathbb{F}_p$, it still has smooth points, which lift to $\mathbb{Q}_p$ points. The only way it can not have points is if it is a product of two lines not defined over $\mathbb{F}_p$. There are no $\mathbb{F}_p$-points so no $\mathbb{Q}_p$-points.

- The probability that the determinant is a multiple of p ($p \mid abc$) is $1 - (1 - \frac{1}{p})^3 \approx \frac{3}{p} + O(p^2)$.
- Then the probability that the singular conic breaks up into two $\mathbb{F}_p$-conjugate lines over $\mathbb{F}_p$, i.e., the 2 lines are not defined over $\mathbb{F}_p$, is $\frac{1}{2}$. For example, if c is a multiple of p, we get $ax^2 + by^2 = 0$. The probability this doesn't factor is $\approx \frac{1}{2}$, the probability $\frac{b}{a}$ is not a square.

The density of $C_{a,b,c}$ having a point over $\mathbb{Q}_p$ is hence

$$\approx 1 - \underbrace{\frac{3}{p}}_{\mathbb{P}(p|abc)} \cdot \frac{1}{2} + O\left(\frac{1}{p^2}\right).$$

The density of cubics hving a $\mathbb{Q}_p$-point for all $p < Y$ is

$$\prod_{p<Y} \left(1 - \frac{3}{2p} + O\left(\frac{1}{p^2}\right)\right) \to 0.$$

As $Y \to \infty$ this goes to 0. □

(We did the calculations intuitively; it's worth going through this rigorously.)

(Alternatively we can try to fix a, b and let $c \to \infty$ for each a, b. This is what people have tried to do to understand the decay.)

What is the rate of decay to 0?

$$|\{|a|, |b|, |c| < X : C_{a,b,c} \text{ has a rational point}\} \ll \frac{X^3}{(\ln X)^{\frac{3}{2}}}.$$

The upper bound is due to Serre. He conjectured this should be a lower bound. This wasn't shown until later by Hooley (Representation of 0 by ternary quadratic forms). The analytic number theory gets technical, and he loses track of the constants quickly. Are the constants equal? Is it asymptotic with a fixed constant, or is it oscillating? This is unknown.

These are simple analytic number theory problems that we don't know the answer to.

The upper bound is straightforward with the right tool (large sieve). The lower bound is hard.

The real question for genus 0 curves is general conics.

Example 1.1.3 (General conic)**:** Consider

$$ax^2 + bxy + cxz + dy^2 + eyz + fz^2.$$

The density of such curves with a rational points is 0.

Proof. To get the rate of decay do the same proof. The probability that such a conic is singular over $\mathbb{F}_p$ is different. Take the determinant of the matrix $\begin{pmatrix} \frac{a}{2} & b & c \\ b & \frac{d}{2} & e \\ c & e & \frac{f}{2} \end{pmatrix}$. The determinant is equidistributed; the probability it's 0 is $\frac{1}{p}$. (This is different from the diagonal case. You can't just this by a diagonal conic.) Are they defined over $\mathbb{F}_p$? Again, it doesn't iff its singular and doesn't factor over $\mathbb{F}_p$, so the probability is $\frac{1}{2}$.

The probability that such a conic over $\mathbb{Z}$ has a $\mathbb{Q}_p$ point for all $p < Y$ is approximately

$$\prod_{p<Y} (1 - \frac{1}{2p}) \to 0$$

as $Y \to \infty$.

We expect

$$|\{a, b, c, d, e, f : \text{height} < X \text{ with } ax^2 + \cdots + fz^2 \text{ has a rational point}\}| \ll \frac{X^6}{\sqrt{X}}.$$

This was proved by Serre (it's what he really developed the large sieve for). It was open for 20 years what the lower bound is. Again it was done ($\gg \frac{X^6}{\sqrt{X}}$) by Hooley (Representation of zero by ternary quadratic forms II) in the last few years. Again, the precise asymptotics are unknown. □

The next frontier is finding the constant. Lower bounds are not a science yet; they tend to be ad hoc; there are common themes like the circle method. Upper bound sieves tend to be uniform.

The group action is implicit in Hooley's paper ("effecting linear changes of variable"). Hooley writes old-style. Perhaps if you translate it into modern group theory language it will be clearer.

Can you say anything about the height of the rational point (ex. to find it)? You have to use effective version of the local-global principle, which Cassel has bounds for. They're not very strong though.

There are many things in genus 0 we don't quite understand yet.

2 Genus 1

The main issue is that the local-global principle does not always hold. You need global arguments as well.

The famous example is

$$3X^3 + 4Y^3 + 5Z^3 = 0$$

due to Selmer. This has points over $\mathbb{Q}_p$ for all p and over $\mathbb{R}$ but not over $\mathbb{Q}$. This has an another level of algebraic difficulty.

The first simplification is to suppose that our curve locally has points everywhere, so much so that it has a global point. Take a family of curves where you already have a rational global point; then there are no local obstructions to having points.

Example 1.2.1: Thus, the first natural case is a genus one curve coming with a marked rational point, i.e., an **elliptic curve**.

Any elliptic curve over $\mathbb{Q}$ can be expressed as $y^2 = x^3 + Ax + B$, $A, B \in \mathbb{Z}$. We embedded the curve using the divisor $3(P)$ where P is the marked point, and then changed variable. The rational points of $E = E_{A,B}$ form an abelian group where the marked point is the identity, denoted $E(\mathbb{Q})$. (The group is the same as the divisor class group.)

Mordell showed that $E(\mathbb{Q})$ is finitely generated. Once you have a marked point, you have a group of rational points. It's abelian and finitely generated. We can ask about statistics of the group.

The discriminant is $-4a^3 - 27b^2$. In terms of modular forms these are the G_4, G_6 of the corresponding lattice; G_4^3, G_6^2 are comparable. It's natural to define the height

$$H(E_{A,B}) = \max\{|4A^3|, |27B^2|\}.$$

How are the groups distributed in the space of finitely generated abelian groups?

That it is finitely generated means

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times T$$

where r is the **rank** and T is finite (the torsion group). It turns out that $|T| < 16$ by Mazur. It's easy to show that 100% of the time $|T| = 1$. To get $|T| = c > 1$, certain algebraic things happen to a, b; they have to be on a subvariety. There are finitely many of these subvarieties, so there's probability 0 it's on one of them.

For $r = 0$ there are finitely many rational points. For $r \geq 1$ there are infinitely many. The rank quantifies how "infinitely many" you are.

The statistics of the number of rational points is not so interesting. We can distinguish them by the rank.

The natural questions are not about the number of rational points but about the distribution of the rank r. In particular, do they have finitely many points or infinitely many points more often? When they have infinitely many points, what does the rank tend to be?

Do elliptic curves tend to have finitely many or infinitely many rational points? People guessed both, and they were both wrong.

Goldfeld looked at families of twists of elliptic curves. Katz-Sarnak used a the big book of random matrix philosophy.

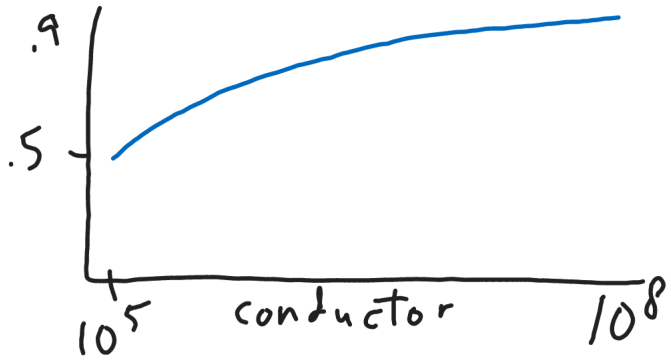
Conjecture 1.2.2 (Goldfeld, Katz-Sarnak)**.** *The average rank of elliptic curves over* $\mathbb{Q}$ *is* $\frac{1}{2}$*. More precisely,* 50% *have rank 0 and* 50% *have rank 1.*

There are infinite families of curves with rank 2, but they are very sparse. As long as you order by something not correlated with rank (ex. conductor, discriminant, height), we think it should be true (no formal equivalence is known though). The number of curves of conductor $< X$ should be $X^{\frac{5}{6}}$. A goes up to $X^{\frac{1}{3}}$ and B goes to $X^{\frac{1}{2}}$. We don't know this yet—it's something we need to answer first.

The best bounds are due to Duke, who shows a $o(X)$ upper bound. There are many ingredients. One has to use class group bounds, modularity, etc. You can get a lower bound of $X^{\frac{5}{6}}$.

If you solve this problem, presumably a lot of theorems we prove for height can be adapted to if we order by conductor.

Katz-Sarnak were busy writing their book. They had their manuscript floating around. People (Brumer-Mebuinness, Bektemerov-Stein-Watkins) were doing these computations of ranks. There was no correlation between what the computations were saying and what the conjecture was saying. We made a graph.



Moreover, the number of rank 2 curves went up to $\approx 20\%$. Katz-Sarnak though the graph would turn around. But they believed in their random matrix philosophy. That's what got me interested in the problem as a grad student.

At that time, one couldn't even prove that the average rank is greater than 0 or less than ∞. One of the goals I set out to do was to prove this.

What was known?

- Based off an idea of Goldfeld, Brumer and Heath-Brown gave the first theoretical evidence that the average rank is bounded. The average rank is going up; people on the data side weren't even sure the average rank is finite.

  If you assume BSD and GRH for L-functions of elliptic curves (2 million dollars of conjectures). Then the average rank is ≤ 2.3 (Brumer) and ≤ 2 (Heath-Brown).

  Heath-Brown tried to get < 2, because that would imply a positive proportion have rank 0 or 1.

- Young showed with the same assumptions that the average rank is ≤ 1.7.

Can we show this by leaving the analytic and L-function world and using algebraic techniques? Can we combine these two techniques?

Theorem 1.2.3 (Bhargava, Shankar)**.** *The average rank is bounded and in fact is* $< .885$.

We'll start talking about the proof of this next week and Shankar will finish it the fourth week. The final ideas involved root numbers.

Theorem 1.2.4 (Bhargava, Skinner, Wei Zhang)**.** *The average rank is strictly positive, and in fact* $> .2$.

This is tricky because you have to construct rational points. The point constructions are done using work of Skinner and Urban. The expected is 0.5—this looks good for the conjecture.

The graph goes past .9 now (after another order of magnitude), so it has to turn around! When A, B are small, it's easier for coincidences to happen. There's still around 20% of rank 2 curves, and it's still not going down. It looks like we'll have to compute quite a bit further.

The results are proved using 2-Selmer and 5-Selmer groups. If these methods continue to work for n-Selmer groups for higher n, then the bounds will approach 0.5. This gives another way to conjecture that the true average rank is 0.5.

Their conjectures work over other global fields; we get a universal bound < 1.05 over any global fields of characteristic $\neq 2$. The techniques indicate that the conjecture should be true over other global fields. More is not known for function fields. The lower bound does not work because we need to apply Gross-Zagier (totally real fields probably work). Jerry will talk about generalizing the bounds to any number field.

Corollary 1.2.5 (of proof)**.** *A positive proportion ($> 20\%$) of elliptic curves over $\mathbb{Q}$ have rank 0.*

Corollary 1.2.6 (of proof)**.** *A positive proportion ($> 20\%$) of elliptic curves over $\mathbb{Q}$ have rank 1.*

(Arul will explain how to get this with root numbers.)

Corollary 1.2.7 (together with work of Wei Zhang)**.** *A positive proportion ($> 20\%$) of elliptic curves over $\mathbb{Q}$ have analytic rank 0, 1.*

Theorem 1.2.8. *A positive proportion ($> 66\%$) of elliptic curves satisfy the Birch and Swinnerton Dyer rank conjecture.*

What if any obstacle is there to taking 5-Selmer to higher Selmer groups? We think it's a question of algebraic geometry. The algebraic geometry we used was classical stuff known in the 18th century. Presumably we can do better.

Example 1.2.9: A genus 1 curve with a degree 2 divisor, i.e., a degree 2 map to $\mathbb{P}^1$. (In Example 1.2.1 we had a degree 1 divisor.) By Riemann-Hurwitz, it has 4 ramification poitns, so such a curve can be expressed as

$$z^2 = f(x, y)$$

where f is a binary quartic form.

This is not an elliptic curve because it may not necessarily have rational points. We may get 2 points at ∞ that may be conjugate to one another. (The usual way you see this is $z^2 = f(x)$ where f is a quartic polynomial.)

We can ask the same question: varying the 5 coefficients, how are the rational points distributed?

All the theorem use each other!

Theorem 1.2.10. *If all binary quartic forms $ax^4 + bx^3y + cx^2y^2 + dxy^3 + dxy^3 + ey^4$ $(a, b, c, d, e \in \mathbb{Z})$ are ordered by height,*

1. *A positive proportion ($> 75\%$) of $z^2 = f(x, y)$ have a point over $\mathbb{Q}_p$ for all p and over $\mathbb{R}$.*

   This is very different from the genus 0 case. (We will show a sieve technique, for any genus 1 or higher curve, to get lower bounds.)

2. *A positive proportion ($< 75\%$) of $z^2 = f(x, y)$ have a point everywhere locally but no point over $\mathbb{Q}_p$ (they fail Hasse).*

3. *A positive proportion of $z^2 = f(x, y)$ hve a rational point.*

All 4 possibilites (local points everywhere? global points?) occur a positive proportion of the time.

Example 1.2.11: Genus 1 curves in $\mathbb{P}^2$ (corresponding to a genus 1 curve with a degree 3 divisor), i.e., a smooth plane cubic, i.e., a ternary cubic form, which has 10 coefficients:

$$ax^3 + bx^2y + \cdots + jz^3 = 0.$$

This is the most common cubic that people talk about: smooth cubics in the plane.

Theorem 1.2.12. *The same theorem holds for genus 1 curves in $\mathbb{P}^2$.*

Example 1.2.13: Genus 1 curves in $\mathbb{P}^3$ (corresponding to a genus 1 curve with a degree4 divisor), i.e., two quadrics in $\mathbb{P}^3$, i.e., two quaternary quadratic forms. They are given by 2 4×4 matrices with coefficiens $a, b, c, d, e, f, g, h, i, j; k, l, m, n, o, p, q, r, s, t,\ a, \ldots, t \in \mathbb{Z}$ (20 coefficients).

This is a complete intersection: any genus 1curve in $\mathbb{P}^3$ can be represented this way, and conversely, given two quadrics, generically they will intersect in genus 1 curve. This is a full description of genus 1 curves in $\mathbb{P}^3$.

Theorem 1.2.14. *The same theorem holds for genus 1 curves in $\mathbb{P}^3$.*

These are all classical objects in 18th century algebraic geometry.

There's one more example.

Example 1.2.15 (Example 5)**:** Genus one curves in $\mathbb{P}^4$. This is NOT a complete intersection. Generically, 5 quadrics intersect in nothing. Many people gave up; some persisted. You can completely describe them as well.

This is a quintuple of 5×5 skew symmetric matrices A, B, C, D, E.

How to make a genus 1 curve using 5×5 skew symmetric matrices? Given A, B, C, D, E, you can make a single 5×5 matrix of linear forms

$$Ax + By + Cz + Dt + Eu.$$

We can takes its determinant? The determinant is 0 because it's an odd skew-symmetric matrix.

What if we take the determinant of a 4×4 skew-symmetric matrix? It's a square. It's a degree 4 form in 5 variables. Taking its square root we get a quadric. There are 5 primary (on-diagonal) 4×4 matrices.

The square root of the determinant is called the **Pfaffian**. The Pfaffian of any primary 4×4 block of $Ax + By + Cz + Dt + Eu$ gives a quadric in $\mathbb{P}^4$.

We get 5 quadrics $Q_1, \ldots, Q_5$ in $\mathbb{P}^4$. Classical algebraic geometers noted that

$$Q_1 \cap \cdots \cap Q_5$$

is a genus 1 curve in $\mathbb{P}^4$. Conversely any genus 1 curve embedded by a complete linear system is given uniquely (up to change in basis) as $Q_1 \cap \cdots \cap Q_5$ where $Q_1, \ldots, Q_5$ come from a unique (A, B, C, D, E), a quintuple of 5×5 skew-symmetric matrices.

(Old reference: Cayley. (Old algebraic geometry) Modern reference: Buchsbaum-Eisenbud. (Modern commutative algebra) It's not clear which is easier. We do a hybrid. See also Fisher; he writes concetely.)

These case $\mathbb{P}^1$ helps us understand 2-Selmer, up to $\mathbb{P}^4$ helping us understand 5-Selmer. This is the last case we know how to solve.

The problem in general is to describe how to use this to understand Selmer. The open question is how can we study genus 1 curves in higher projective spaces? Can you describe, even over $\mathbb{C}$, the genus 1 curves in higher dimensional projective space?

(What if we do something that is an approximate parametrization? The problem is that not all 5 quadrics intersect in a genus 1 curve; we get an overcount.) Even understanding any infinite sequence of n will do! Note we haven't used modern algebraic geometry.

3 Some basic algebraic geometry

2-12: missed class today. Notes are sketchy.

Let k be a field of characteristic $\neq 2, 3$. Let C/k be a "nice" curve of genus g. $\mathrm{Pic}^0(C)$, the Jacobian, is an abelian variety of dimension g, defined over k. We have that $C \to \mathrm{Pic}^\wedge(C)$ is injective if $g \geq 1$, an isomorphism if $g = 1$. For $D \in \mathrm{Pic}^n(C)(k)$, indices a map

$$\begin{aligned} \pi_D : C &\to \mathrm{Pic}^0(C) \\ P &\mapsto n(P) - D. \end{aligned}$$

For the case $g = 1$, $C \cong \mathrm{Pic}^1(C)$ is an E-torsor. For $D \in \mathrm{Pic}^n(C)(k)$,

$$\pi_P(\underbrace{c}_{\in C} + \underbrace{e}_{\in E}) = \pi_P(c) + ne.$$

Definition 1.3.1: An **n-cover of** E is a torsor C of E along with a morphism $\pi : C \to E$ defined over k such that for all $c \in C(\overline{k}), e \in E(\overline{k})$,

$$\pi(c + e) = \pi(c) + ne.$$

Note that π is determined by (the divisor class of $E = \mathrm{Pic}^0(C)$) $\pi(P)$ for any fixed $P \in C(\overline{k})$. If $\pi : C \to E = \mathrm{Pic}^0(C)$ is an n-cover, then the divisor class $D = n(P) - \pi(P) \in Pic^n(C)(\overline{k})$ doesn't depend on $P \in C(\overline{k})$ and is thus Galois invariant, i.e., $n(P) - \pi(P) \in \mathrm{Pic}^n(C)(k)$. Then $\pi = \pi_0$. So we have the correspondence between divisors $D \in \mathrm{Pic}^n(C)(k)$ and n-covers $C \to \mathrm{Jac}(C)$.

Example 1.3.2: The trivial n-cover is multiplication by n, $[n] : E \to E$. For any $d \in E(k)$,

$$\begin{aligned} [n]_d : E &\to E \\ P &\mapsto nP + d \end{aligned}$$

is a n-cover equivalent to $[n]$ if $d \in nE(k)$.

Definition 1.3.3: Two n-covers $\pi_1 : C_1 \to E, \pi_2 : C_2 \to E$ are equivalent if there exists an isometry $\alpha : C_1 \xrightarrow{\cong} C_2$ of n-torsors ($\alpha(c_1 + e) = \alpha(c_1) + e$) such that the following commutes:

$$\begin{array}{ccc} C_1 & \xrightarrow{\alpha\cong} & C_2 \\ & {\scriptstyle \pi_1}\searrow \quad \swarrow{\scriptstyle \pi_2} & \\ & E & \end{array}$$

Example 1.3.4: $[n]_{d_1}$ and $[n]_{d_2}$ are equivalent iff $d_1 \equiv d_2 \pmod{n}E(k)$. We obtain an inclusion

$$E(k)/nE(k) \hookrightarrow \{n\text{-covers of } E\}/\text{equivalence}$$

Example 1.3.5: Suppose $\pi : C \to E$ is an n-cover. Then π is equivalent to $[n]$ iff there exists $P \in C(k)$ with $\pi(P) = O$.

Example 1.3.6: Say π is **soluble** if $C(k) \neq \phi$, so $C \cong E$ as E-torsors (the isomorphism sending P to O). Then π is equivalent to $[n]_d : E \to E$ for some $d \in E(k)/nE(k)$, $d = \pi(P)$. We have the bijection

$$E(k)/nE(k) \hookrightarrow \{\text{soluble } n\text{-covers of } E\}/\text{equivalence}$$

Example 1.3.7: If k is algebraically closed, every n-cover is soluble and $E(k)/nE(k) = 0$. Any n-cover $\pi : C \to E$ is equivalent to $[n]$. Thus, over arbitrary k, n-covers modulo equivalence are twists of $[n] : E \to E$.

$$\begin{aligned} \{n\text{-covers}\}/\text{equivalence} &\cong H^1(k, \mathrm{Aut}([n] : E \to E)) \\ &= H^1(k, E[n]). \end{aligned}$$

Definition 1.3.8: Let k be a global field. A n-cover $\pi : C \to E$ is **locally soluble** if for every place v of k, $C(k_v) \neq \phi$.

We have

$$\begin{aligned}\mathrm{Sel}_n(E/k) :&= \{\text{locally soluble } n\text{-covers of } E\}/\text{equivalence}\\ &= \{\text{elements of } H^1(k, E[n]) \text{ lying locally in the image of } E(k_v)/nE(k_v)\}.\end{aligned}$$

There is a map

$$E(k)/nE(k) \hookrightarrow \mathrm{Sel}_n(E/k)$$

so

$$|\mathrm{Sel}_n(E/k)| \geq |E(k)/nE(k)| \geq n^{\mathrm{rank}\, E(k)}.$$

This means we can bound average ranks in terms of average sizes of Selmer groups.

Our goal is to write down equations for locally soluble n-covers and count solutions.

A locally soluble n-cover corresponds to $D \in \mathrm{Pic}^n(C)(k)$. (The n-cover can be represented by a rational divisor since the n-cover is locally soluble.) This corresponds to a map $C \to \mathbb{P}^{n-1}$.

Example 1.3.9: For $n = 2$, the map $C \to \mathbb{P}^1$ corresponds to writing $C : z^2 = f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$ where f is unique up to PGL_2-equivalence. ($C \subset \mathbb{P}(1, 1, 2)$).

For example, $D = (1, 0, \sqrt{a}) + (1, 0, -\sqrt{a})$ on $E : y^2 = x^3 - \frac{I}{3}x - \frac{J}{27}$ where

$$\begin{aligned}I &= 12ae - 3bd + c^2\\ J &= 72ace + 9bcd - 27ad^3 - 27eb^3 - 2c^3\end{aligned}$$

are the PGL_2-invariants of binary quartic forms. $\pi : C \to E$ is trivial iff $f(x, y)$ has a linear factor over K. Call a 2-cover **irreducible** if f has no linear factor over k.

(Note that $[2] : E \to E$ corresponds to $f(x, y) = x^3y - \frac{I}{3}xy^3 - \frac{J}{27}y^4$.)

The group $G_2 := \mathrm{PGL}_2$ acts on $V_2 = \mathrm{Sym}^4(2)$, the set of binary quartic forms. The action is given by

$$\gamma f(x, y) = \frac{1}{(\det \gamma)^2} f((x, y)\gamma).$$

The invariants $[V_2]^{\sigma_2}$ are the free polynomial ring generated by I, J.

Example 1.3.10: For $n = 3$, $C \to \mathbb{P}^2$ given by D, write

$$C : ax^3 + bx^2y^2 + cx^2z + \cdots + jz^3 = 0.$$

The terms I, J in the equation for E are defined using the Hessian.

$\pi : C \to E$ is trivial iff C has a rational flex point (triple tangent). Otherwise, call π **irreducible**.

The group $G_3 := \mathrm{PGL}_3$ acts on $V_3 = \mathrm{Sym}^3(3)$ by

$$\gamma f(x, y, z) = \frac{1}{\det \gamma} f((x, y, z)\gamma).$$

Example 1.3.11: For $n = 4$, $C \to \mathbb{P}^3$, $C : Q_1(t_1, \dots, t_4) = Q_2(t_1, \dots, t_4) = 0$, $C' : z^2 = \det(xQ_1 - yQ_2)$ is a 2-cover of $\mathrm{Jac}(C)$. $E : y^2 = x^3 - \frac{I}{3} - \frac{J}{27}$. For fixed x, y, $xQ_1 - yQ_2$ is a quadric surface in $\mathbb{P}^3$ and a solution z to $z^2 = \det(xQ_1 - yQ_2)$ corresponds to a ruling of the surface.

For $p \in C = ((Q_1 = 0) \cap (Q_2 = 0))$, there is a unique (x, y) such that the line T_pC lies in $xQ_1 - yQ_2 = 0$. The ruling containing this line corresponds to a point $(x, y, z) \in C^1$. We obtain a map $C \to C^1$. We call a 4-cover $C \to E$ irreducible iff $\det(xQ_1 - yQ_2)$ has no linear factors. $[4] : E \to E$ corresponds to

$$Q_1 = \begin{bmatrix} & & & 1 \\ 0 & 0 & 0 & 0 \\ & & 1 & \\ 1 & & & \end{bmatrix} \qquad Q_2 = \begin{bmatrix} & & -1 & \\ & -1 & & -\frac{I}{6} \\ -1 & & & \\ & & -\frac{I}{6} & -\frac{J}{27} \end{bmatrix}$$

The group

$$G_4 = \{(g_2, g_4) \in \mathrm{GL}_2 \times \mathrm{GL}_4 : \det(g_2)\det(g_4) = 1\} / \left\{(\lambda^{-2}, \lambda) : \lambda \in \mathbb{G}_m\right\}$$

acts on $V_4 = 2 \otimes \mathrm{Sym}^2(4)$.

Example 1.3.12: For $n = 5$,

$$G_5 = \left\{(g_1, g_2) \in \mathrm{GL}_5 \times \mathrm{GL}_5 : (\det g_1)^2 \det g_2 = 1\right\} / \{(\lambda, \lambda^{-2})\}$$

acts on $V_5 = (\Lambda^2 5) \otimes 5$.

We don't know what happens for $n \geq 6$.

Theorem 1.3.13. *Let k be a field of characteristic $\neq 2, 3$. Let $n = 2, 3, 4, 5$, $E : y^2 = x^3 - \frac{I}{3}x - \frac{J}{27}$.*

1. *There is a bijection*

$$E(k)/nE(k) \leftrightarrow \textit{soluble } G_n(k)\textit{-orbits in } V_n(k) \textit{ with invariants } I, J$$

2. *For $v \in V_n(k)$ with invariants I, J there is an isomorphism*

$$\mathrm{Stab}_{G_n}(V) \cong E[n]$$

as group schemes. (cf. Ho-Bhargava: regular representations of genus 1 curves; Birch-and-Swinnerton-Dyer: notes on elliptic curves 1, Cremona-Fischer-Stoll: minimalization and reduction $n = 2, 3, 4$, explicit n-descent for elliptic curves I)

3. *Suppose now that K is a global field of characteristic $\neq 2, 3$. Then there is a bijection*

$$\mathrm{Sel}_n(E/k) \leftrightarrow \textit{locally soluble } G_n(K)\textit{-orbit of } V_n(K) \textit{ with invariants } I, J,$$

Theorem 1.3.14. *Let K be a global field of characteristic $\neq 2, 3$. When all elliptic curves over K are ordered by height, the average size of $\mathrm{Sel}_n(E/K)$ for $n = 2, 3, 4, 5$ is 3, 4, 7, 6. (These are $\sigma(n) = \sum_{d|n} d$.) The average number of irreducible locally soluble n-covers mod equivalence is n for $n = 2, 3, 4, 5$.*

An application: Using $5^r \geq 20r - 15$ for $r \in \mathbb{Z}_{\geq 0}$, the average rank is $\leq \frac{21}{20} = 1.05$.

3.1 Heights over global fields

Let K be a number field, and M_∞ be the set of archimedean places.

Let K is the function field of a nice curve $X/\mathbb{F}_q$. Then $M_\infty = S_0$ is the set of valuations at any fixed finite nonempty set of points on X. Let $\mathcal{O}_K$ be the functions which are regular on $X \backslash S_0$.

Two elliptic curves $E_{A,B}, E_{A',B'}$ are isomorphic iff $A' = c^4 A$, $B' = c^6 B$ for some $c \in K^\times$. We can think of $E_{A,B}$ as a point $(A, B) \in \mathbb{P}(4,6)(K) = \mathbb{A}^2(K)/G_m(K)$.

Given $(A, B) \in S(K)$, define $\widetilde{I} = \{\alpha \in K : \alpha(A, B) \in S(\mathcal{O}_K)\}$ and

$$H(A,B) = N\widetilde{I} \prod_{v \in M_\infty} \max\{|A|_v^{\frac{1}{4}}, |B|_v^{\frac{1}{6}}\}.$$

The strategy of the proof of the theorem is as follows.

1. Find a "convenient" fundamental domain for the action of $G_m(K)$ on $S(K)$, i.e., $\mathcal{E} \subset S(\mathcal{O}_K)$. $\mathcal{E}$ is the set of elliptic curves up to isomorphism. (Homework)
2. Compute $\#\mathcal{E}$. (Simple case of step 3)
3. Compute the number of locally soluble orbits with invariants in $\mathcal{E}$. (Find integral representations in rational orbits.)
4. Divide, look for cancellation, and profit. (Product formula)

Question: start with $v \in V_n(K)$ locally soluble with integral invariants $(I, J) \in \mathcal{E}$. Does there exist $v' \in V_n(\mathcal{O}_K)$ that is $G_n(K)$-equivalent to V?

Theorem 1.3.15 (Local minimization)**.** *Suppose V_φ is a nonarchimedean local field with ring of integers $\mathcal{O}_\varphi$. Let $v \in V_n(K_\varphi)$ be soluble with invariants $(I, J) \in 6S(\mathcal{O}_K)$. Then $G_n(K_\varphi)v \cap V_n(\mathcal{O}_\varphi) \neq \phi$.*

Proof. We have correspondences between

1. soluble cubics,
2. $d \in E(K_\varphi)/nE(K_\varphi)$,
3. $[n]_d : E \to E$,
4. $E \to \mathbb{P}^{n-1}$ via the divisor $(n-1)\infty + (P)$ where P is in the class of d.

The idea is to find a good representative P for d. Write down $E \to \mathbb{P}^{n-1}$. Observe they are integral.

If $d = 0$, $x^3 y - \frac{I}{3} y^3 - \frac{J}{27} y^4 \in \mathrm{Sym}^4(2)$.

If $d \neq 0$, let $P = (a, b)$ be an arbitrary representative. If $a, b \in \mathcal{O}_\varphi$, then the explicit embedding is integral (do the change of variables $(a, b) \leftrightarrow (0, 0)$ and write down sections). Otherwise, $P \in nE(K_\varphi)$.

For $n = 2$, $L = K_\varphi[x]/(x^3 + Ax + b) = K_\varphi[\theta]$. We have a map

$$\begin{aligned} E(K_\varphi)/2E(K_\varphi) &\hookrightarrow (L^\times/L^{\times 2})_{\text{Norm}=1} \\ (a,b) &\mapsto (a-\theta) \in L^{\times 2} \end{aligned}$$

since $(a,b) \in 2E(K_\varphi)$. □

3.2 Patching

Suppose $v \in U_n(K)$ is locally soluble with invariants in Σ. Then there exist $g_\varphi \in \sigma_n(K_\varphi)$ such that $g_\varphi v \in V_n(\mathcal{O}_\varphi)$ for all $\varphi \notin M_\infty$. Consider the adele $(g_\varphi)_\varphi \in \sigma_n(\mathbb{A}_f)$.

$$d_{\sigma_n,K} := \left(\prod_{\varphi \notin M_\infty} G_n(\mathcal{O}_\varphi) \right) \backslash G_n(\mathbb{A}_f)/G_n(K)$$

is always finite (and $= 1$ for $K = \mathbb{Q}$), by Borel, Prasad, and Conrad.

Fix lifts of $cl_{G_{n,K}}$ in $G_n(\mathbb{A}_f)$. Then there exist $(g'_\varphi) \in \prod_{\varphi \notin M_\infty} G_n(\mathcal{O}_\varphi)$, $\beta \in cl_{G_{n,\mathbb{A}_f}}$ one of the chosen lifts, $h \in G_n(K)$, such that $(g_\varphi) = (g'_\varphi)\beta h$. Then $hv \in V_n(K) \cap \beta^{-1}\left(\prod_{\varphi \notin M_\infty V_n(\mathcal{O}_\varphi)}\right) = V_{n,\beta}$ where $V_{n,\beta}$ is commensurable with $V_{n,1} = V_(\mathcal{O}_K)$. The subgroup

$$G_{n,\beta} = G_n(K) \cap \beta^{-1} \left(\prod_{\varphi \notin M^\infty} G_n(\mathcal{O}_\varphi) \right) \beta$$

is commensurable with $G_{n,1} = G_n(\mathcal{O}_\varphi)$ and acts on $V_{n,\beta}$. Count $\#G_n(K)\backslash V_n^{\text{loc. sol.}}(K)$ by counting $\#G_{n,\beta}\backslash V_{n,\beta}$.

Some problems:

1. One v can correspond to multiple β.
2. One $G_n(K)$-orbit in $V_{n,\beta}$ might break up into multiple $G_{n,\beta}$-orbits.
3. Elements of $V_{n,\beta}$ might not be locally soluble.

The solution is to count lattice points with a weight function.

Chapter 2

Counting forms

2-19

1 Counting binary forms

This problem goes back several hundred years (especially for quadratic forms). Counting them has important consequencs in terms of class number, Selmer groups, etc. More generally, the problem is the following.

Let G be an algebraic group and V a representation over $\mathbb{Z}$. We would like to

1. understand the meaning of the "integer orbits" $G(\mathbb{Z})\backslash V(\mathbb{Z})$ (two points are equivalent if you can get from one to the other by an element of $G(\mathbb{Z})$). For nice representations, like those with a free ring of invariants, they always have some meaning. It's usually something fundamental in arithmetic.

2. count the number of elements of $G(\mathbb{Z})\backslash V(\mathbb{Z})$ having bounded invariants. The invariants are polynomial.

Definition 2.1.1: (G, V) is called **coregular** if the action of $G^{ss}(\mathbb{C})$ (ss means semisimple) on $V(\mathbb{C})$ has a **ring of polynomial invariants** that is freely generated, isomorphic to $\mathbb{C}[t_1, \ldots, t_k]$ for some k (there are no relations among generators).

We want each invariant generator to be bounded.

Theorem 2.1.2 (Hilbert's fundamental theorem of invariant theory)**.** *The ring of invariants of a reductive algebraic group G is finitely generated.*

Theorem 2.1.3 (Borel, Harish-Chandra in Annals)**.** *Assume G is a reductive algebraic group.*

The number of elements in $G(\mathbb{Z})\backslash V(\mathbb{Z})$ with fixed values for all invariants is finite.

The argument is explicit and uses reduction theory. They bring every element into a finite volume region, which has finitely many integer points. As the invariants grow, we want to

know how the number of orbits grows; we'll actually be counting something of arithmetic interest.

(You can also do S-integral orbits or specify conditions at ∞.)

Example 2.1.4 (Binary quadratic forms)**:** Consider $\{ax^2 + bxy + cy^2\} = V$. This was first studied by Gauss. (Gauss considered $ax^2 + 2bxy + cy^2$, but this is a minor difference.) We have that $\mathrm{SL}_2(\mathbb{C})$ on $V(\mathbb{C})$ has one invariant, the discriminant

$$D = b^2 - 4ac.$$

The **class number** is

$$h(D) = \#[G(\mathbb{Z})\backslash V(\mathbb{Z})]_D.$$

Theorem 2.1.5 (Gauss, Mertens, Siegel)**.** *The following hold*

1. $\sum_{-X<D<0} h(D) \sim \frac{\pi}{18} X^{\frac{3}{2}}$.
2. $\sum_{0<D<X} h(D) \ln \varepsilon_D \sim \frac{\pi^2}{18} X^{\frac{3}{2}}$. *Here* $\ln \varepsilon_D$ *is the regulator.*

The problem for positive discriminant is unsolved. It's conjectured that

$$\sum_{0<D<X} h(D) \sim X^{1+\varepsilon}.$$

Proof of (1). First, we construct a a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $V(\mathbb{R})$ (a discrete group acting on a real vector space). The domain is $|b| \leq a < c$.

Recall tht $\mathrm{SL}_2(\mathbb{Z})$ acts on the upper half plane. Every element of the upper half plane can be brought into the region $|z| \geq 1$, $\Re z \in (-\frac{1}{2}, \frac{1}{2}]$. $\mathrm{SL}_2(\mathbb{Z})$ acts on the root by a linear fractional transformaiton. It has 1 root in the upper half-plane. Just make sure the root is mapped into the fundamental domain of $\mathcal{H}$.

To prove (a) we need to count lattice points in $\mathbb{R}^3 = V(\mathbb{R})$ such that

$$|b| < a < c \text{ and } -X < b^2 - 4ac < 0.$$

This region in $\mathbb{R}^3$ has volume $\frac{\pi}{18} X^{\frac{3}{2}}$. Gauss concludes that the number of lattice points is $\sim \frac{\pi}{18} X^{\frac{3}{2}}$. (The numbers are usually the same, but not always.)

What is the complication? The region looks something like this.

Even though it has finite volume, it has a cusp going off to ∞. This is a big problem in the geometry of numbers: the region may not be bounded. The tiny volume going off to ∞ may have few or many lattice points. You have to argue that if you go far enough, you don't pick up as many lattice points as you expect. (This is a good problem to try.) □

The modern interpretation of this theorem is that $h(D)$ is the class number of $\mathbb{Z}\left[\frac{D+\sqrt{D}}{2}\right]$, the unique quadratic ring with discriminant D.

It tells us how fast orders in quadratic fields grow. The number of ideal classes is about $\sqrt{D}$. In the positive discriminant case, $h(D)\ln\varepsilon_D$ is the thing that grows like $\sqrt{D}$.

(There is another subtlety. Usually people just count the number of invertible ideal classes. For general orders there may be noninvertible ideal classes. We count those too.)

The sieve in this case if elementary—you can do it with your bare hands. In general it's hard and unsolved. If you're only interested in invertible classes you need another sieve (add certain congruence conditions, i.e., a, b, c are relatively prime, to get invertible classes; require d squarefree to count only the ring of integers of quadratic fields).

Example 2.1.6 (Binary cubic forms)**:** Let $V = \{ax^3 + bx^2y + cxy^2 + dy^3\}$, $G = \mathrm{GL}_2$. The action of SL_2 on V over $\mathbb{C}$ again has a unique invariant, the discriminant

$$D = b^2c^2 + 18abcd - 4ac^3 - 4b^3d - 27a^2d^2.$$

(When $a = 1, b = 0$ this is $-4c^3 - 27d^2$.)

Cubics are a lot harder. It's in 4 dimensions and the surface is more complicated. Writing down all the boundaries explicitly and working by hand as Gauss did, Davenport found the following. You can't count all orbits; you have to restrict the ones you count.

Let $h(D)$ be the number of $\mathrm{GL}_2(\mathbb{Z})$-orbits on $V(\mathbb{Z})$ that are irreducible over $\mathbb{Q}$ (don't factor over $\mathbb{Q}$ or equivalently $\mathbb{Z}$ by Gauss's Lemma).

Theorem 2.1.7 (Davenport, 1960s)**.** *The following hold.*

1. $\sum_{-X<D<0} h(D) \sim \frac{\pi^2}{72}X$.
2. $\sum_{0<D<X} h(D) \sim \frac{\pi^2}{24}X$.

The averge size is a constant, different from the quadratic case where the number was growing.

The proof is the same but more complicated because the cusp contains tons of $a = 0$ points of integer discrimant. They can be ignored because they are reducible. Ignoring those, when $a \neq 0$, 100% of the points you count are irreducible.

We'll prove a harder case from which these will be easy.

There are 2 arithmetic interpretations, to do with classification of cubic rings.

2 Classification of cubic rings

Definition 2.2.1: A **n-ic ring** (ring of rank n) is a ring that is free of rank n as a $\mathbb{Z}$-module, $\cong \mathbb{Z}^n$ as a $\mathbb{Z}$-module. Ex. an order in a degree n number field is a n-ic ring.

The **discriminant** of a n-ic ring R is defined as follows. For $\alpha \in \mathbb{R}$, let x_α be multiplication-by-α. Define $\det(x\alpha) =: N(\alpha)$ and $\mathrm{Tr}(x\alpha) =: \mathrm{Tr}(\alpha)$. There is a pairing $\langle \alpha, \beta \rangle = \mathrm{Tr}(\alpha\beta)$. Then $\mathrm{Disc}(R) = \det(\langle , \rangle)$. (This definition works more generally than the definition for a field.)

What are the n-ic rings for...

1. $n = 1$: $\mathbb{Z}$
2. $n = 2$: $\left\{\mathbb{Z}\left[\frac{D+\sqrt{D}}{2}\right] : D \equiv 0, 1 \pmod 4\right\}$, where D is the discriminant. (When $D = 0$, you adjoin a square to 0, i.e., get $\mathbb{Z}[X]/(X^2)$. When $D = 1$ we get $\mathbb{Z} \times \mathbb{Z}$. We're modding out by the formal polynomial relation $\frac{D+\sqrt{D}}{2}$ that it would satisfy.

   (To show these are all the quadratic rings, write τ as a linear combination of $1, \tau$. Write as $\mathbb{Z}[\tau]/f(\tau)$, f quadratic. Translate τ by in integer. D characterizes quadratic rings.
3. $n = 3$: The ring is not specified any more by the discriminant. There can be many cubic rings of the same discriminant, but at least 1 for every discriminant, $Z \times \mathbb{Z}\left[\frac{D+\sqrt{D}}{2}\right]$.

   A cubic ring always has a basis over $\mathbb{Z}$, $1, \omega, \theta$. To specify the ring, we need to know how ω, θ multiply. We have

$$\begin{aligned}\omega^2 &= ? + ?\omega + ?\theta \\ \theta^2 &= ? + ?\omega + ?\theta \\ w\theta &= ? + ?\omega + ?\theta.\end{aligned}$$

   We can't assign random things because we have to impose associativity. Simplification: we can translate ω by any constant. It only changes the cofficient of θ of $w\theta$; we can make it disappear. Similarly, we can translate θ to make $?\omega$ disappear.

   A cubic ring always has a **normal basis** over $\mathbb{Z}$, with

$$\begin{aligned}\omega^2 &= l + b\omega + a\theta \\ \theta^2 &= m + d\omega - c\theta \\ \omega\theta &= n.\end{aligned}$$

   Exercise: write the multiplication table and use the associative law. Find that l, m, n uniquely are determined by a, b, c, d. Use $\omega^2\theta = \omega\theta \cdot \omega$ and $\omega \cdot \theta^2 = \omega\theta \cdot \theta$—these imply associativity of the entire ring. We get 6 equations (that are dependent) having a unique solution,

$$\begin{aligned}\omega^2 &= -ac + b\omega + a\theta \\ \theta^2 &= -bd + d\omega - c\theta \\ \omega\theta &= -ad.\end{aligned}$$

   Giving a cubic ring with a normal basis is equivalent to giving a quadruple (a, b, c, d).

   How do we get rid of the normal basis? A cubic ring can have many normal bases, but not that many. Do a $\mathrm{GL}_2(\mathbb{Z})$-change of basis on ω, θ. There's a unique way to normalize to get a normal basis. There's a $\mathrm{GL}_2(\mathbb{Z})$-set of normal bases.

Thus cubic rings are in bijection with quadruples modulo whatever the action of $\mathrm{GL}_2(\mathbb{Z})$ defined above is. We know the representations of $\mathrm{SL}_2(\mathbb{Z})$, and (a,b,c,d) gives a 4-D representation. We can guess what it is: Sym^3, binary cubics (check this).

In summary, $\mathrm{GL}_2(\mathbb{Z})$ acts on normal bases $\langle 1, \omega, \theta\rangle$, therefore acts on (a,b,c,d), in fact just as $\mathrm{GL}_2(\mathbb{Z})$ acts on the binary cubic $ax^3 + bx^2y + cxy^2 + dy^3$.

We've sketched the proof of the following.

Theorem 2.2.2 (Levi, Delone-Faddoev, Gan-Gross-Savin)**.** *There is a bijection*

$$\{\text{cubic rings}\}/\sim \leftrightarrow G(\mathbb{Z})\backslash V(\mathbb{Z}).$$

The discriminants of the cubic ring is the discriminant of the corresponding binary cubic form.

(Both disciminants are degree 4 and $\mathrm{GL}_2(\mathbb{Z})$-invariant, so must be the same up to a constant; you can check the constant is 1.)

Things are as nice as can be.

This was done very coordinate-full, but you can do it coordinate-free.

The polynomial should have some meaning in terms of the ring. We need this for when we want to do over other base rings/schemes. We discover things by explicit calculations, but then we want to know what's going on.

There is a natural meaning. Cubic form on rank 2 module... (RECORDING) $R/\mathbb{Z} \to \Lambda^3 R$. The map is

$$x \mapsto 1 \wedge x \wedge x^2.$$

Now that we have a nice interpretation, the theorem of Davenport immediately have meaning: they count cubic rings of negative/positive discriminant. There are $\frac{\pi^2}{72}$ cubic rings (that are integral domains, i.e., order in cubic field) per negative discriminant on average, and $\frac{\pi^2}{24}$ orders per positive discriminant.

In the quadratic case the form is determined by the discriminant. In the cubic case it is not determined, but there are a constant number on average.

A irreducible form corresponds to an irreducible ring, i.e., integral domain. So we are counting something very natural.

The sieve to maximal orders is done by Davenport and Heilbronn. We want to restrict to squarefree D; this is already nontrivial for binary cubic forms. In general, this is unsolved to impose squarefree condition.

What fails to generalize to n-ic rings?

We can generalize to quartic or quintic rings, or forms. They diverge at $n = 4$. We'll focus on forms.

3 Binary quartic forms

What changes? They don't classify quartic rings.

Now there are 2 independent invariants for the action of SL_2. (The representation is coregular over $\mathbb{C}$. It is not free over $\mathbb{Z}$.) For $ax^4 + \cdots + ey^4$, the invariants are generated by I, J where

$$\begin{aligned} I &= 12ac - 3bd + c^2 \\ J &= 72ace + 7bcd - 27ad^2 - 27b^2e - 2c^3 \\ D &= \frac{4I^3 - J^2}{27}. \end{aligned}$$

(The discriminant is not generated by I, J over $\mathbb{Z}$; it is over $\mathbb{Z}[\frac{1}{3}]$.) This is classic, in Hilbert's book on invariant theory. The way to find this: SL_2 acts on Sym^4 of the standard representation. Decompose

$$\mathrm{Sym}^2(\mathrm{Sym}^4(\mathrm{std})) = \mathrm{Sym}^8(\mathrm{std}) \oplus \underbrace{\mathrm{Sym}^4(\mathrm{std})}_{\text{Hessian}} \oplus \underbrace{\mathrm{Sym}^0(\mathrm{std})}_{I}.$$

We can count binary quartic forms with bounded invariants, but now we have to say both I, J are bounded. We define a height in terms of both I, J.

$$H(f) := H(I, J) := \max\left\{|I|^3, \frac{J^2}{4}\right\}$$

(They have different degrees; we want to scale correctly. Constants don't really matter, and are for convenience.)

Question: Let $G = \mathrm{GL}_2$. What is $\#[G(\mathbb{Z})\backslash V(\mathbb{Z})]_{H(I,J)<X}$?

For a given height, there are a given number of I, J. How many are there for given values of I, J? To answer this we get total counts and divide by number of I, J.

Birch and Merriman show that if you only fix tje discriminant, only finitely many of that discriminant. This is special for binary quartic forms.

If the answer is $cX^{\frac{5}{6}}$ for $c > 0$, then this means there exists a positive constant number of binary quartics up to equivalence per (I, J), on average. The best known bounds classically look like cX, pretty far off.

The complication is that the reduction theory is harder: it's in five dimensions. The degree of the boundaries have exploded. There are 2 invariants. You have to loosely estimate everything.

Theorem 2.3.1 (Yang, Ph.D. thesis (2005))**.** *For every $\varepsilon > 0$,*

$$\#[G(\mathbb{Z})\backslash V(\mathbb{Z})]_{H(I,J)<X} \ll X^{\frac{5}{6}+\varepsilon}.$$

In the binary and cubic case we can do things by hand. That strategy we have to abandon at some point. Very soon the boundaries cannot be written in the history of the universe; they gets big very quickly. The spaces have very commplex invariants. We want a coordinate-free description, carrying over to the reduction theory, etc. There must be something deeper.

If you put in the irreducibility condition, . we get rid of ε and specify the constant. We need the constant to average Selmer sizes.

Binary quadratics have 2 roots: 2 real or 2 complex. Binary cubics have 3 roots: 3 real or 1 real and 2 complex. That's why there were 2 parts. Now there are 3 parts: 4 real, 2 real and 2 complex, 4 complex.

Theorem 2.3.2 (Bhargava, Shankar, Binary quartics)**.** *Let $h^{(i)}$ denote the number of forms with i pairs of complex roots.*

1. $\sum_{H(I,J)<X} h^{(0)}(I,J) = \frac{4}{135}\zeta(2)X^{\frac{5}{6}} + O(X^{\frac{3}{4}+\varepsilon})$.
2. $\sum_{H(I,J)<X} h^{(1)}(I,J) = \frac{32}{135}\zeta(2)X^{\frac{5}{6}} + O(X^{\frac{3}{4}+\varepsilon})$
3. $\sum_{H(I,J)<X} h^{(2)}(I,J) = \frac{8}{135}\zeta(2)X^{\frac{5}{6}} + O(X^{\frac{3}{4}+\varepsilon})$

(Note it's possible for forms to have the same I, J but have different number of real roots.)

The analogue of $D \equiv 0, 1 \pmod 4$ for binary quadratic forms is the following.

Theorem 2.3.3. *I, J satisfy one of the following 4 congruences.*

1. $I \equiv 0 \pmod 3$ *and* $J \equiv 0 \pmod{27}$.
2. $I \equiv 1 \pmod 9$ *and* $J \equiv \pm 2 \pmod{27}$.
3. $I \equiv 4 \pmod 9$ *and* $J \equiv \pm 16 \pmod{27}$.
4. $I \equiv 7 \pmod 9$ *and* $J \equiv \pm 7 \pmod{27}$.

Such (I, J) are called ***eligible****.*

How many I, J of bounded height are there satisfying these conditions? From these congruence conditions we can work out the proportion.

Corollary 2.3.4. *The number of binary quartics per eligible (I, J) with $0, 1, 2$ pairs of complex roots is $\frac{\zeta(2)}{2}$, $\zeta(2)$, $\frac{\zeta(2)}{2}$, respectively.*

We will prove this. We want to visualize the regions without writing explicit equations. This can improve binary quadratic and cubics as well. The equations for binary quartics were too complicated to handle classically.

We won't make 1 fundamental domain and count. We use the fact that no matter which you pick, the number of lattice points is the same. Let's count in the union of fundamental domains and divide by the number of domains. Their union will be a nicer region. In fact, we'll define a continuous sweep of fundamental domains!

There are many maps going from one n-ic form to a m-ic form. A common, classic one is the Hessiam. Given a binary n-ic form, obtain a binary $2(n-2)$-ic form

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{vmatrix}.$$

This sends binary cubics to binary quadratics. For $n = 4$, it sends a binary quartic to binary quartics.

People have searched for geometric meaning. Fulton-Harris talk about geometric plethisms.

Binary cubic means 3 points in $\mathbb{P}^1$; binary quadratic means 2 points in $\mathbb{P}^2$. Take the linear transformation which cycles them. It has 2 fixed points. That's the binary quadratic!

Can you similarly interpret the map from binary quartics to binary quartices? I don't know.

Here's something people don't know the answer to.

$$\mathrm{Sym}^2(\mathrm{Sym}^{2n+1}(\mathrm{std})) \overset{\text{unique}}{\supset} \mathrm{Sym}^2(\mathrm{std})$$

using weights (representation theory). There's always a way to go from an odd number of points in $\mathbb{P}^1$, to 2 points. We saw the geometric way of constructing that for $n = 3$. It's unsolved to find a geometric description.

4 Counting binary cubic forms

2-26:

We want to count the average number of $\mathrm{GL}_2(\mathbb{Z})$-classes of integral binary n-ic forms having integral invariants.

Binary quadratic forms $ax^2 + bxy + cy^2$. The unique invariant is the discriminant $D = b^2 - 4ac$. Let $h(D)$ be the number of classes of binary quadratic forms having discriminant D.

It's conjectured that for every $\varepsilon > 0$, $\mathrm{avg}(h(D)) = D^\varepsilon$ if $D > 0$. This is unsolved. We don't know how to prove an exponent better than $\frac{1}{2}$. It's conjectured that for infinitely many D, it's 1.

We know $\mathrm{avg}(h(D)) = c|D|^{\frac{1}{2}}$ if $D < 0$ (Gauss/Mertens/Siegel). We understand negative discriminants but not positive discriminants because there is a nontrivial stabilizer for $D > 0$.

For binary cubic forms $ax^3 + bx^2y + cxy^2 + dy^3$, $D = b^2c^2 - 27a^2d^2 + 18abcd - 4ac^3 - 4b^3d$. We will prove that $\mathrm{avg}(h(D)) = \begin{cases} \frac{\pi^2}{72}, & \text{if } D > 0, \\ \frac{\pi^2}{24}, & \text{if } D < 0, \end{cases}$ both due to Davenport. Note we only count irreducible forms over $\mathbb{Q}$.[1] This says that the average number of classes per discriminant is finite.

The method of proof is to count the number of binary cubic forms of discriminant of absolute value $< X$ in a fundamental domain for the action of $\mathrm{GL}_2(\mathbb{Z})$ on $V(\mathbb{R})$. Here V is the space of binary cubic forms. Davenport's method does not work for higher degree forms. We need to have a conceptual reason for why we can count points in these regions in general.

We give Davenport's fundamental domain when $D > 0$. Recall Gauss's fundamental domain for binary quadratic forms $ax^2 + bxy + cy^2$ with $D < 0$: it is $|b| < a < c$.

[1] Otherwise the answers would be bigger; the total number of lattice points would be greater than the volume.

The Hessian of a binary cubic form that has $D > 0$ gives a binary quadratic form with $D < 0$. The Hessian of $ax^3 + bx^2y + cxy^2 + dy^3$ is[2]

$$\text{Hess}(x^3 + bx^2y + cxy^2 + dy^3) = (b^2 - 3ac)x^2 + (bc - 9ad)xy + (c^2 - 3bd)y^2.$$

Call f **reduced** if $\text{Hess}(f)$ is reduced in the sense of Gauss. Recall that SL_2 acting on the standard representtion; we have $\text{Sym}^3(s+d)$ is the set of binary cubics, and $\text{Sym}^2(\text{Sym}^3(\text{std})) = \text{Sym}^6(\text{std}) \oplus \text{Sym}^2(\text{std})$.

There is no symmetric power of binary quartics that has Sym^2 in it. (This is a nice exercise using the representation theory of SL_2. Use weights.) Thus this reduction of reduction does not work for binary quartics. (It also doesn't work with binary cubics with $D < 0$. Then we get a binary quadratic with $D > 0$; we don't know how to reduce those.)

Counting average size of class number was nicer because you have linear inequalities. Use the Principle of Lipschitz. If the region is rounded enough, the number of lattice points grows as the volume. This is basically the theory of Riemann integration. Letting $x \to \infty$, you're looking at a homogeneously expanding region. By hand there are few points of interest in the cusps which has small volume. Try this; this is the only principle you need for binary quadratics.

The principle of Lipschitz isn't enough for binary cubics. In the cusp there are more points than the volume would indicate. He wrote a paper "On a paper of Lipschitz" that says we need a better way to count, and to give an explicit error term depending on the region.

Theorem 2.4.1 (Davenport, "On a principle of Lipschitz" with erratum)**.** [3] *Let R be a bounded semi-algebraic region in $\mathbb{R}^n$ defined by $\leq k$ polynomial inequalities of degree $\leq l$. The number of lattice points in R is*

$$Vol(R) + O_{k,l}(\max\{ Vol(\overline{R}, 1)\}),$$

where $\overline{R}$ ranges over all projections of R onto smaller dimensional coordinate hyperplanes.

When the region is round and big, we expect number of lattice points to be the volume. The volume is the maximum of the projections to smaller dimensional hyperplanes. If the region is round, it will be 1 order of magnitude less. When it is wrong is when the projection is thin. Then there is a projection where the error term is the same order of magnitude.

Take a line through the region R. Get a union of integrals. It will be bounded in terms of the number of polynomial inequalities. The degrees of those inequalities. (A line can intersect at most degree many times. Induct dimension by dimension. In each dimension count on a union of unit intervals. Add an error with the number of intevals.

If you change the lattice do you do better?

[2]The 2 fixed points of the transformations cycling the roots are the 2 roots of the Hessian. Suppose they're defined over $\mathbb{Q}$ together. Take the linear transformation that cycles them. Each should be defined over a quadratic field. Is it the resolvent quadratic field? That doesn't happen. You get a quadratic field of disciminant -3 times the original discriminant, $\text{Disc}(\text{Hess}(f)) = -3\text{Disc}(f)$.

[3]Lang made a big fuss for many years about the error. Finally, Lang published an erratum to fix it.

Theorem 2.4.2 (cf. Theorem 2.1.7)**.**

$$\sum_{-X<D<0} h(D) = \frac{\pi^2}{72}X + O(X^{\frac{15}{16}})$$
$$\sum_{0<D<X} h(D) = \frac{\pi^2}{24}X + O(X^{\frac{15}{16}}).$$

Davenport proved this usign explicit fundamental domain. Davenport uses clever relations between the inequalities. (When D is negative, it has a unique root in the upper half-plane. Do the Gauss trick there. He works the inequalities on a, b, c, d explicitly; they are degree 2 inequalities.)

In general we don't want to write down explicit inequalities. We want to know that are the features that allow us to prove bounds, without knowing explicitly knowing the inequalities.

The way to do this is to use fundamental domains in the group. Here's a way to make fundamental domains for $G(\mathbb{Z})$ on $V(\mathbb{R})$, for $G(\mathbb{Z}) = \mathrm{GL}_2(\mathbb{Z})$ and $V(\mathbb{R})$ the space of binary cubics.

Let $\mathcal{F}$ denote a fundamental domain for the action of $G(\mathbb{Z})$ on $G(\mathbb{R})$. Let $v_0 \in V(\mathbb{R})$ have positive discriminant. Then $\mathcal{F}v_0$, considered as a multiset $\{gv_0 : g \in \mathcal{F}\}$, is a union (as multisets) of 6 fundamental domains for the action of $G(\mathbb{Z})$ on $V^+(\mathbb{R})$.

Proof. Any positive discriminant binary cubic form is equivalent to any other by an element of $G(\mathbb{R})$ because there's only 1 invariant. Binary cubic form is 3 points on $\mathbb{P}^1$. By linear fractional transformation you can take them to any other 3 points. Any can be brought to any other by an element of $G(\mathbb{R})$.

Intuitively,

$$[G(\mathbb{Z})\backslash G(\mathbb{R})] \cdot \left[G(\mathbb{R})\backslash V^+(\mathbb{R})\right] = G(\mathbb{Z})\backslash V^+(\mathbb{R}).$$

Where does the 6 come from?

Over $\mathbb{R}$ there are stabilizers of size 6! There are further transformations that fix v_0, 6 of them! A fundamental domain for this action is $\frac{1}{6}$ of the point v_0. The right statement is

$$\underbrace{[G(\mathbb{Z})\backslash G(\mathbb{R})]}_{\mathcal{F}} \cdot \underbrace{\left[G(\mathbb{R})\backslash V^+(\mathbb{R})/\operatorname{Stab}_{G(\mathbb{R})}(v_0)\right]}_{\frac{1}{6}v_0} = G(\mathbb{Z})\backslash V^+(\mathbb{R}).$$

□

We can describe $\mathcal{F}$ explicitly.

For $\mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$, we can describe $\mathcal{F}$ as the set of $g \in G(\mathbb{R})$ taking i into the guillotine region—that's a fundamental region for the $G(\mathbb{Z})$ action on $G(\mathbb{R})$. The upper half-plane is a homogeneous space for $G(\mathbb{R})$.

Gauss's fundamental domain for binary quadratic forms is

$$\mathcal{F} \cdot (x^2 + y^2) = \left\{ax^2 + bxy + cy^2 : |b| < a < c\right\}.$$

The reason this is a set not a multiset is that everything is weighted by SO_2 so everything has the same weight. You can hit with the orthogonal group and then go everywhere.

We can think of Davenport's fundamental domain for $D > 0$ this way as well, even though he didn't. His fundamental domain is

$$\mathcal{F} \cdot f = \left\{(a,b,c,d) : |bc - 9ad| < b^2 - 3ac < c^2 - 3bd\right\}.$$

where f is a binary cubic such tht the Hessian is a multiple of $x^2 + y^2$ (e.g., $f = x^3 - 3xy^2$).

For Davenport's fundamental domain for binary cubic forms with $D > 0$, there's an explicit description of $\mathcal{F}$ using the **Iwasawa decomposition** (ex. for the general linear group N' is lower triangular matrices, A' is the diagonal matrices, K is the orthogonal group)

$$\mathcal{F} = N'A'K'\Lambda$$

where

$$\begin{aligned}
N' &= \left\{\begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} : |n| \le \frac{1}{2}\right\} \\
N'(t) &= \left\{\begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} : n \in \nu(t) \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right]\right\} \\
A' &= \left\{\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} : |t| \le \sqrt{\frac{\sqrt{3}}{2}}\right\} \\
K &= \mathrm{SO}_2 \\
\Lambda &= \left\{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} : \lambda > 0\right\} = \{\lambda > 0\}.
\end{aligned}$$

What are all elements take i back into the fundamental domain? Hitting with λ does nothing. We can hit i with all $K = \mathrm{SO}_2$. Now you can move it around. Use matrices $\left(\begin{smallmatrix} t^{-1} & 0 \\ 0 & t \end{smallmatrix}\right)$ to move it around. The diagonal matrices move it up and down on the vertical line $\Re z = 0$. The bound on t makes sure it doesn't go below $\Im < \frac{1}{2}$. Now we can hit it with lower triangular matrices which moves it left and right. We get a little more than the fundamental domain (we get a segment of the circle).

SO_2 and $N'(t)$ are compact. The tentacle-ness is A'. (Λ is bounded because the discriminant is bounded.) We've isolated the problem! $\mathcal{F}$ is nonbounded because of the A. That's the part we have to deal with when we hit v_0.

The point is that $\mathcal{F}v$ would give a fundamental domain no matter which v you take. We take advantage of this. Davenport, etc. took a clever v and tried to count points there. We can avoid this and take any v. We can count points in many fundamental domains. Even better, take a whole ball of points, and divide by the ball. When you have a ball of v's, the tentacles swirl about. Divide by how much you swirled.

Let B be a K-invariant compact region in $V^+(\mathbb{R})$. For any $v \in B$ let

$$\begin{aligned}\mathcal{R}_X(v) &= \mathcal{F} \cdot v \cap \{|\operatorname{Disc}| < X\} \\ &= N'A'K\Lambda' v \qquad \text{where } \Lambda' = \left\{ \lambda < \left(\frac{X}{|\operatorname{Disc}(v)|} \right)^{\frac{1}{4}} \right\}.\end{aligned}$$

Lemma 2.4.3. *Let $v \in B$. The number of reducible integral binary cubic forms in $\mathcal{R}_X(v)$ with $a \neq 0$ is $O_\varepsilon(X^{\frac{3}{4}+\varepsilon})$.*

Note $a = 0$ means reducible, and there are tons of $a = 0$ forms; that's where the cusp is. As $t \to \infty$, c, d become large and a, b become small. t^{-3} makes a really small.

Think of the action

$$N'A'K\Lambda v.$$

1. Λ: coordinates up to $X^{\frac{1}{4}}$.
2. K: compact, so size is still $O(X^{\frac{1}{4}})$.
3. A': a, b, c, d are multiplied by t^{-3}, t^{-1}, t, t^3, respectively.
4. N', as a lower triangular matrix, adds the small stuff (a, b) to the large stuff (c, d).

c, d can get really big; that's where the infinite part is.

Proof. If $d = 0$, the number of choices for a, b, c is $O(X^{\frac{3}{4}+\varepsilon})$. (But it's possible $b, c = 0$. But if $b = 0$, we can bound ac and do better, etc.)

(Important in this argument: when small coordinates are nonzero, you have a negative 2-weight which you can use to bound things by multiplying.)

If $d \neq 0$, and a form (a, b, c, d) is reducible, then it has a linear factor $px + qy$. (Then c is determined.) Fix a, b, d; then $abd = O(X^{\frac{3}{4}})$ so thare are $O(X^{\frac{3}{4}+\varepsilon})$ choices. Then $p \mid a, q \mid d$, implying that p, q are determined up to X^ε choices. (N has at most N^ε factors.) □

If $a = 0$ then we have $O(X)$ instead.

Note one can use a p-dic argument (reducible means reducible mod p) to get $o(X)$. But the global argument gives the right order of magnitude.

4.1 Averaging

Let

$$N^+(X) = \sum_{0<D<X} h(D)$$

be the number of irreducible lattice points with $a \neq 0$ in $\mathcal{R}_X(v)$ for any $v \in B$. We can sum this up over all $v \in B$. B is an infinite set, so integrate; use the volume. This is (the numerator is a constant)

$$= \frac{\int_{v \in B} \#\left\{x \in \mathcal{F}v \cap V^{\text{irr}}(\mathbb{Z}) : |\operatorname{Disc}(x)| < X\right\} |\operatorname{Disc}(v)|^{-1}\, dv}{6 \int_{v \in B} |\operatorname{Disc}(v)|^{-1}\, dv}$$

The nice measure to take here is $|\operatorname{Disc}(v)|^{-1}$ because it's an invariant measure—it doesn't change under $\mathrm{SL}_2(\mathbb{Z})$. When you scale v by 2, the discriminant goes up by 16. But the Euclidean measure goes up by 16 as well. (dv is not a nice measure. The group likes the measure above.) The degree of the derivative is the dimension of the space.

This is a tautology. Now we change the order of integration—we make the integral over the group, so we can do integrals inside $\mathrm{GL}_2(\mathbb{Z})$. (The denominator is a constant C.) Here dg is the Haar measure on $G(\mathbb{R})$.

$$= \frac{1}{C} \int_{g \in \mathcal{F}} \#\left\{x \in yB \cap V^{\mathrm{irr}} \cap V^{\mathrm{irr}} : |\operatorname{Disc}(x)| < X\right\} \, dg$$

The Haar measure has a nice description in terms of the Iwasawa decomposition.

$$dg = t^{-2} \, dn \, d^\times t \, dk \, d^\times \lambda.$$

Here $d^\times t = \frac{dt}{t}$ is the multiplicatively invariant measure. Now

$$= \frac{1}{C} \int_{\lambda=0}^{X^{\frac{1}{4}}} \int_K \int_{t=C''}^{\infty} \int_{n \in \nu(t)} \#\left\{x \in n \begin{pmatrix} t & \\ & t \end{pmatrix}^{-1} k\lambda B \cap V^{\mathrm{irr}}(\mathbb{Z}) : |\operatorname{Disc}| < X\right\} t^{-2} \, dn \, d^\times t \, dk \, d^\times \lambda$$

$$= \frac{1}{C} \int_{\lambda=0}^{X^{\frac{1}{4}}} \int_{t=C''}^{\infty} \int_{n \in \nu(t)} \#\left\{x \in n \begin{pmatrix} t & \\ & t \end{pmatrix}^{-1} \lambda B \cap V^{\mathrm{irr}}(\mathbb{Z}) : |\operatorname{Disc}| < X\right\} t^{-2} \, dn \, d^\times t \, d^\times \lambda$$

B is a fat ball in 4-D space. Hitting it with λ; it homogeneously expands; thus we can remove K above. Instead of counting points in an unbounded region, we're not counting in a simple region.

Let

$$B(n, t, \lambda, X) := n \begin{pmatrix} t & \\ & t \end{pmatrix}^{-1} \lambda B,$$

B stretched using these parameters.

Lemma 2.4.4. *Let C''' be the maximum of the absolute values of coefficients of $v \in B$. (B is compact, so this exists.) Then the number of lattice points in $B(n, t, \lambda, X)$ with $a \neq 0$ is*

$$\begin{cases} 0, & \text{if } \frac{c'''\lambda}{t^3} < 1 \\ \mathit{Vol}(B(n, t, \lambda, v)) + O(\max\{t^3\lambda^3, 1\}), & \text{otherwise.} \end{cases}$$

If we squeezed the box too much, the only thing left is $a = 0$. Once that smallest side (range of a) is big enough, the volume is a good approximation of the number of points.

Proof. If $\frac{C'''\lambda}{t^3} < 1$, then the only integral value of a is $a = 0$.

Otherwise, the number of lattice points with $a \neq 0$ is the volume with error given by Davenport's lemma 2.4.1.

Note the 1 in $\max\{t^3\lambda^3, 1\}$ is a major source of errors in the geometry of numbers? □

Then (note we cut off the integral before)

$$N^+(X) = \frac{1}{6}\mathrm{Vol}(\mathcal{R}_X(v)) + \underbrace{O\left(\int_\lambda \int_{t=\lambda^{\frac{1}{3}}} + \iiint \max\{t^3\lambda^3, 1\}\, dg\right)}_{O(X^{\frac{5}{6}})}$$

If we put the volume everywhere we get the volume back. But we didn't, so we have to include errors.

These are standard integrals. (Exercise.) They give the same thing, $O(X^{\frac{5}{6}})$. Our result is better than Davenport's. If you work harder, you can show $X^{\frac{5}{6}}$ is the second-order term.

The method did better, and we never wrote down inequalities for the fundamental domain! We just need to know what the features were, the $\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$ which characterizes the infinite part.

Look in the slice $a = 1$, if you count the number of points you get $X^{\frac{5}{6}}$. If you look across all classes, which have a transformation to make the leading term 1 (which represent 1)? Presumably the $a = 1$ slice are 100% of the cubics representing 1 and contribute the entire $O(X^{\frac{5}{6}})$ error term, but we don't know how to prove this.

Note there's no 3rd order term, it's $O(X^{\varepsilon})$! See Bhargava, Shankar, Tsimerman.

(The Shintani zeta function only have poles at $1, \frac{5}{6}$; residues give coefficients. It encodes counts of cubic rings locally, Tamagawa numbers.)

Index