THREE LECTURES ABOUT THE ARITHMETIC OF ELLIPTIC CURVES.

(– These are very rough, unedited, and preliminary notes – B. M.)

Lecture I. Introduction to ABC Problems.

I.1 What is Number Theory? And why does it turn out to be so directly tied to geometry? to the representation theory of groups? or, nowadays, to physics? How difficult it is, to gauge the importance, the centrality, of a question posed about numbers! Which are the questions that will turn out to be frivolous dotings on mere surface phenomena? And which questions lead to an understanding of deeper structure? Or perhaps you might want to phrase these wonderings with the opposite tonality: Which are the questions that will merely lead the question-asker to remoter structures, further and further from the basic object? And which of the questions will keep faithful focus on the texture of numbers and their interrelations?

The subject has a bounty of famous ancient problems: direct queries, which can be asked in not too technical, almost pre-mathematical, language: questions about the placement of prime numbers among all numbers (the Goldbach conjecture, the twin prime conjecture, the"Schinzel hypothesis" predicting when there are an infinite number of prime number values of a given polynomial, etc.), and also questions about the behavior of the sets of "perfect powers" under simple arithmetic operations.

It is this second type of question that I want to talk about this hour. A *perfect power* is the n-th power of an integer where n is some natural number > 1. These have attracted attention from the earliest times, beginning with perfect squares, which arise in reflections concerning the relationship given in the Pythagorean Theorem, applied to right-angle triangles all three of whose sides are integral multiples of a given unit. Of course, perfect squares arise in other ways as well; consider Fibonacci's reflection in his treatise"Liber Quadratorum" in 1225 on perfect squares:

"I thought about the origin of all square numbers and discovered that they arise out of the increasing sequence of odd numbers; for the unity is a square. namely 1; to this unity is added 3. making the second square, namely 4, with root 2; if the sum is added to the third odd number, namely, 5,..."

There is no end of famous problems regarding the most simple-seeming questions of placement of perfect powers, and sums of them, on the number line, (**Fermat**: for $n > 2$ the sum of two n-th powers is never an n-th power; **Catalan** (1844): the numbers 8 and 9 are the only consecutive perfect powers; **Waring** etc. Also, there is the problem (at first glance, it is somewhat curious to single this problem out!) of finding for any fixed integer k, all integral solutions to the "**Mordell Equation**"

$$Y^2 - X^3 = k.$$

As for Fermat's Last Theorem, we now have a proof, thanks to the celebrated efforts of Wiles and Taylor (1995).

In the direction of the Catalan Problem, we know, thanks to Tijdeman (1976) –who used Baker's theory of lower bounds for nonvanishing linear forms in logarithms– that there are only a *finite* set of (pairs of) consecutive perfect powers. An upper bound (for a perfect power whose successor is also a perfect power) can be computed from Tijdeman's proof (work of Langevin) to be

$$e^{e^{e^{e^{730}}}}.$$

As for the Mordell Equation, a general theorem of Siegel (1929) guarantees that for a given nonzero integer k it has only a finite number of integral solutions (as does any affine curve of genus > 0 over the ring of integers). Moreover, much explicit work has been done (both unconditionally and dependent upon standard conjectures) to finding concretely the solutions for given values of $k < 100,000$ (cf [ ]).

If one views each of the problems above as "Diophantine", i.e., as the problem of finding integral solutions to specific algebraic equations, one is struck by how specific, indeed, these equations are. To nudge oneself towards a more flexible type of problem which still carries much of the flavor of the problems we have reviewed, let us generalize somewhat the idea of a "perfect power"– and deal, instead, with numbers possessing comparatively large perfect power divisors. My main reason for considering this kind of generalization is that it is a liesurely way of getting some intuition for, and appreciation of, the recent ABC-conjecture due to Masser and Oesterlé. A second reason is to produce a source of problems which can be stated in relatively nontechnical language, which might ever-so-slightly remind one of the constellation of "Manin Conjectures" (these being much more precise predictions concerning the asymptotics of rational points of bounded height in varieties with ample anti-canonical bundle) currently being framed and studied by Batyrev, Peyre, Tschinkel, and Manin.

To prepare for all this, recall that the *radical* of a non-zero number N, denoted rad(N), is the product of all the prime divisors of N; so rad$(-1) = 1$, rad$(24) = 6$, etc.

Definition. If $N \neq 0, \pm 1$ is an integer, by the **power** of N, denoted $P(N)$, we mean the real number

$$P(N) = \frac{\log |N|}{\log \operatorname{rad}(N)}.$$

It is reasonable to simply convene $P(\pm 1) := \infty$ so that the power function is defined for all non-zero integers. We have that $P(N) \geq 1$ and (for $N > 1$) $P(N) = 1$ if and only if N is "square-free", i.e., if and only if N is not divisible by any perfect square (> 1). If N is a perfect n-th power, we have that $P(N) \geq n$.

For $a > 1$ a real number, by an **a-powered number** let us mean a non-zero number N with $P(N) \geq a$.We will be wanting to study the properties of the set of a-powered numbers– the "placement" of these sets among all integers, the behavior of these sets under simple arithmetic operations.

As a way of introduction, let us first answer the question of how "many" numbers N are there with $P(N) = 1$? More exactly, for a positive real number X let *Sq.free*(X)

denote the number of square-free numbers N in the interval $1 < N < X$; how fast does $Sq.\mathit{free}(X)$ tend to ∞ with X?

The answer, which involves a small piece of "sieve theory" (legacy of Eratosthenes), has been known for quite a while and I will review the basic idea behind it. The first step in setting up our "sieve" is to array "in a line" all the integers N in the range $1 < N < X$. There are roughly X of them. Next, so as not to get confused by too many numbers appearing in our calculation, let us rename the prime numbers as $p_1, p_2, \cdots$ in increasing order, so that, in fact, p_1 is the prime number 2;p_2 is, in fact, the prime number 3 , etc.). Now, cross off all integers N divisible by the square of the first prime number, i.e., divisible by p_1^2 $(= 4)$ for surely none of these N's are squarefree. After having crossed these off, we are left with roughly $(1 - \frac{1}{p_1^2}) \cdot X$ remaining numbers in our line. We now want to cross off all integers N divisible by the square of the second prime number, i.e., divisible by p_2^2 $(= 9)$ for, again, surely none of these N's are squarefree. But we must be careful to make the count of what remains. Thinking about this, you find that at this stage, after numbers divisible by p_1^2 and p_2^2 are crossed off, is roughly $(1 - \frac{1}{p_1^2} - \frac{1}{p_2^2} + \frac{1}{(p_1p_2)^2}) \cdot X$ numbers N. The point here is that if we had not included the third term $+\frac{1}{(p_1p_2)^2}$, one would have (erroneously) counted *twice* "as removed" all the numbers N which are divisible by $(p_1p_2)^2$, the first time because N is divisible by p_1^2 and a second time because N is divisible by p_2^2. Since

$$(1 - \frac{1}{p_1^2} - \frac{1}{p_2^2} + \frac{1}{(p_1p_2)^2}) \cdot X = (1 - \frac{1}{p_1^2})(1 - \frac{1}{p_2^2}) \cdot X,$$

I hope you see the pattern that is emerging. Indeed, with some work, one can (control the error terms, cf. [***] and) show the asymptotics to be:

$$Sq.\mathit{free}(X) = \prod_{p=2,3,\cdots} (1 - \frac{1}{p^2}) \cdot X + o(X),$$

as X tends to ∞. Since

$$\prod_{p=2,3,\cdots} (1 - \frac{1}{p^2}) = \frac{1}{\zeta(2)},$$

where $\zeta(s)$ is Riemann's ζ-function, and since

$$\zeta(2) = \sum_{n=1}^{\infty} 1/n^2 = \frac{\pi^2}{6},$$

we get:

$$Sq.\mathit{free}(X) = \frac{6}{\pi^2} \cdot X + o(X),$$

or, speaking loosely, the probability that a given number is squarefree is a bit under two-thirds.

Question for computer scientists: The analogous "power function" P define on the ring of polynomials in one variable over a field is very rapidly calculated by applying the

Euclidean algorithm to find the g.c.d's of the polynomial and its derivatives. Is there an algorithm for computing $P(N)$ which is significant faster than factorizing N?

For real numbers $a \geq 1$ and X, let $S(a;X)$ denote the number of integers $1 \geq N < X$ such that $P(N) \geq a$, i.e., the number of a-powered numbers less than X. Aa Andrew Granville explained to me, an easy argument gives that for fixed $a \geq 1$, and for any $\epsilon > 0$ we have:

$$X^{1/a-\epsilon} < S(a;X) < X^{1/a+\epsilon}$$

as X tends to ∞. We will abbreviate this type of estimate as $S(a;X) = X^{1/a+o(1)}$ emphasizing that $S(a;X)$ grows very roughly like $X^{1/a}$. This is a good thing for us insofar as $X^{1/a}$ is also the rate of growth of "perfect a-th powers" ; i.e., by generalizing from "perfect a-th powers" to "a-powered numbers" we haven't, at least, changed the rough asymptotics.

We are now ready to raise a question which has at least a remote connection to each of the problems in our illustrative list. Fix real numbers $a, b, c \geq 1$, and another real number X. Consider the set $\mathcal{S}(a,b,c;X)$ of triples (A,B,C) of nonzero integers which sum to zero, which are relatively prime, each of which is in absolute value $< X$, and such that

$$P(A) \geq a; \quad P(B) \geq b; \quad P(C) \geq c,$$

i.e., A is a-powered, B is b-powered, and C is c-powered,

Question. How fast can we expect the cardinality of the set $\mathcal{S}(a,b,c;X)$ to grow,if at all, for fixed $a,b,c \geq 1$ and X tending to ∞?

Here is the typical "secret calculation" that is popular to make, to come up with an "expected rate of growth" in this circumstance, but as you will see, it is unlikely that one could come up with a *proof* that these asymptotics are correct just by pursuing the argument that we will give!

Ignoring for the moment the requirement that A, B, C be relatively prime and that they sum to 0, and remembering that the A's are chosen from a set of (roughly) $X^{1/a}$ elements, the B's from a set of (roughly) $X^{1/b}$ elements, and similarly for the C's , we have (roughly) $X^{1/a+1/b+1/c}$ triples (A,B,C) with the requisite lower bounds on their "power functions". The requirement that A, B, C be relatively prime shouldn't change the asymptotics, but the requirement that they sum to 0 should. The absolute value of the sum $|A+B+C|$ is bounded by a constant (3, in fact) times X and so the "chances" that the sum be zero (provided that no other mitigating large effect has been ignored– an important proviso) is inversely proportional to X; call it X^{-1}. Feeding all this into our calculation, our line of reasoning might then lead us to "expect" that the cardinality of $\mathcal{S}(a,b,c;X)$ is comparable to $X^{1/a+1/b+1/c-1}$.

But how to interpret this "expectation"?

Let us refer to $d := 1/a + 1/b + 1/c - 1$ as the **basic exponent** of our problem.

(sub-)Question If the "basic exponent" d is positive, i.e., if

$$1/a + 1/b + 1/c > 1,$$

does the cardinality of $\mathcal{S}(a, b, c; X)$ tend to ∞ as X grows, and with the asymptotics

$$\text{card } \mathcal{S}(a, b, c; X) = X^{d+o(1)}?$$

Remark 1. I asked Trevor Wooley about this question, and he sketched an argument using the circle method that proves these asymptotics when a, b, c are bounded above by 6/5. As he pointed out, the circle method does extraordinarily well in handling ternary additive problem involving square-free numbers ("because" wrote Wooley, "one has unexpectedly strong control over the relevant exponential sums on minor arcs– one does a lot better than square-root cancellation").

The basic idea, here, is to note that an a-powered integer N, with,say, $a < 6/5$ can be written as xy with x square-free and y "square-full" and y is small compared to x. Also, when we are estimating a-powered integers, the bulk of our count will consist in integers which are essentially a-powered and not much more highly poered, for the more higly powered numbers are relatively sparse. Now write $A = xy$, $B = uv$, and $C = zw$, with x, u, z square-free and y, v, w "square-full" and x, u, z the dominating variables; that is we re studying the equation

$$xy + uv + zw = 0$$

with with x, u, z square-free and y, v, w "square-full". The circle method then gives an asymptotic formula for the number of solutions x, u, z for fixed y, v, w with uniform error term. Summing over the possible triples y, v, w, noting that there aren't too many of them, gives a more precise asympototic statement than is formulated in the above "sub-question" in the range $a, b, c \leq \kappa$ with $\kappa = 6/5$. Specifically, in this range,

$$\text{card } \mathcal{S}(a, b, c; X) \sim \gamma \cdot X^d$$

for an appropriate constant $\gamma = \gamma(a, b, c)$, and he suggests that with more work, the upper bound $\kappa = 6/5$ might be improved to 10/7. It is clear, however, that one cannot expect to improve this too much further. We cannot have the same shape of asymptotic formula for the region given by $a, b, c \leq \kappa = 2$ for example; for Pythagorean triples *alone* will give rise to a contribution to card $\mathcal{S}(2, 2, 2; X)$ on the order of $X^{1/2} \cdot \log(X)$.

Remark 2. There is the following natural extension of this problem to m integers where $m \geq 3$. Consider the following "region" D in m-space (a kind of "sconce")

$$D = \{(a_1, a_2, \ldots, a_m) \in \mathbf{R}^m \mid \ a_j \geq 1, \ \ j = 1, \ldots, m \text{ and } d = \sum_{j+1}^{j=m} \frac{1}{a_j} - 1 \ > \ 0\}.$$

Let $S(a_1, a_2, \ldots, a_m; X)$ denote the number of m-tuples of integers $A_1, A_2, \ldots, A_m$ which are pair-wise relatively prime, which sum to zero, are of absolute value $< X$, and such that A_j is an a_j-power number for $j = 1, \ldots, m$. Might one expect, using the analogous rough calculation as above, that for $(a_1, a_2, \ldots, a_m) \in D$ (or at least in some large sub-region in D), we have

$$S(a_1, a_2, \ldots, a_m; X) = X^{d+o(1)}?$$

As Wooley pointed out, the circle method should work all the better the larger m is. For example, can one get a "respectably large" explicitly described sub-region $C \subset D$ (the letter C is for "circle method") such that for $(a_1, a_2, \ldots, a_m) \in C$,the sharper statement that $S(a_1, a_2, \ldots, a_m; X)$ is asymptotic to X^d is true (and is provable by the circle method)?

Returning to our original question, if the "basic exponent" d is negative, i.e., if

$$1/a + 1/b + 1/c < 1,$$

the rough calculation above might suggest the following:

Conjecture 1. *If $1/a + 1/b + 1/c < 1$, then there are, in total, only a finite number of triples A, B, C none zero, which sum to zero, which are relatively prime, and such that*

$$P(A) \geq a; \quad P(B) \geq b; \quad P(C) \geq c.$$

If we are at the boundary, i.e. if $d = 1/a + 1/b + 1/c - 1$ is zero, well, our rough calculation would certainly suggest that for any positive ϵ, we might hope for an upper bound card $\mathcal{S}(a, b, c; X) \leq\leq X^{\epsilon}$, but on the basis of that calculation we wouldn't have any grounds for entertaining a prejudice regarding whether or not card $\mathcal{S}(a, b, c; X)$ tends to ∞ as X grows.

As in the remark above, the "heuristic" we have outlined (to get "expected asymptotics") can be altered to fit a number of other related problems (but of course this "heuristic" never provides any logical justification for the answers it comes up with!).

For example, here is a mild variant of our original problem, which is related to the Catalan problem. Fix real numbers a, b, and a non-zero integer k. For X a real number, consider the set $\mathcal{S}_k(a, b; X)$ of b-powered numbers $N < X$ which are translates by k of an a-powered number. That is, $N < X$ is a b-powered number, and $N - k$ is an a-powered number. The "basic exponent" for this problem is $d = 1/a + 1/b - 1$, as you can easily calculate in analogy with the calculation we made above.

In analogy with Conjecture 1 above, then, an optimist might make the following conjecture concerning this problem in the case of negative "basic exponent":

Conjecture 2. *If k is a nonzero integer, and $a, b > 1$ real numbers such that $1/a + 1/b < 1$, then there are, in total, only a finite number of pairs A, B such that $B - A = k$ and*

$$P(A) \geq a; \quad P(B) \geq b.$$

Aside: Both Conjectures 1 and 2 have the current status (a happy or unhappy status, depending upon your attitude!) of having been verified in NO case. E.G., consider the special case of Conjecture 2 where we fix $k = 1$ and $a = 3$, $b = 2$. The set of pairs (A, B) satisfying the hypotheses in this instance of Conjecture 2 consists of pairs of consecutive numbers $(B = A + 1)$ with A "3-powered" and B "2-powered (which, for example, would include the pair 8 and 9 encountered in our discussion of the Catalan problem). We do not, at present, seem to have the techniques to prove that this set is finite.

The two conjectures we have just displayed are along the lines of Masser-Oesterlé's ABC-Conjecture and will, I believe, provide some motivation for the eventual formulation of the ABC-Conjecture in this lecture. Nevertheless Conjectures 1 and 2, strong as they are, are a good deal weaker than ABC. To see, though, that Conjectures 1 and 2 are already quite strong, you can perform the (easy!) exercise of showing that Conjecture 1 (for *any* fixed choice of a, b, c with negative basic exponent) implies Fermat's Last Theorem for large enough degree, and also implies that there are only a finite number of solutions to the Catalan problem.

I.2 The ABC-conjecture. By an **ABC-solution** let us mean a triple of nonzero integers (A, B, C) which are relatively prime, and which sum to zero. Define the **power** $P(A, B, C)$ of an ABC-solution (A, B, C) to be

$$P(A, B, C) := \frac{\log \max(|A|, |B|.|C|)}{\log \operatorname{rad}(ABC)}.$$

Conjecture (Masser-Oesterlé's ABC): Given any real number $a > 1$ there are only a finite number of ABC-solutions of "power" $\geq a$.

You can do the exercise of seeing that the ABC-conjecture implies Conjectures 1 and 2.

Numerical Examples. Elkies and Kanapka have tabulated all ABC-solutions with $\log \max(|A|, |B|.|C|) < 2^{32}$ and with power > 1.2; see the display. The four "top" ABC-solutions (in this range) are:

$$2 + 3^{10} \cdot 109 + (-23^5) = 0$$

(discovered by Reyssat; its power is 1.629912...)

$$11^2 + 3^2 5^6 7^3 + (-2^{21} 23 = 0$$

(discovered by de Weger; its power is 1.625991...)

$$283 + 5^{11} 13^2 + (-2^8 3^8 17^3 = 0$$

(discovered by Browkin-Brzezinski; its power is 1.580756...)

$$1 + 23^7 + (-5^4 7) = 0$$

(discovered by de Weger; its power is 1.567887...)

Lecture II. The equivalence between *ABC* and general conjectures in the arithmetic of curves; relations between *ABC* and the arithmetic of elliptic curves. (sketch)

II. 1 (Introductory Remarks about (Effective) "*ABC*" implying (Effective) "Mordell" (Elkies); Elkies' "near-misses" coming from rational points of elliptic curves; as a lead in to:) Elliptic curves and their "classifying invariants".

I will assume some familiarity with the very basic definitions, and some properties, of the theory of elliptic curves over **Q**. To give such, you can do it in "Weierstrass form" just by giving two rational numbers (which I will denote c_4 and c_6) and writing the cubic equation

$$E: \; y^2 = x^3 +$$

with the understanding that E is the *projective* plane curve defined by that equation, with the single point at infinity included:– the point at infinity being taken to be the origin of the celebrated group law in E. or you can be more particular about E and choose a "Tate-Weierstrass" model for E over the ring of integers **Z** by giving an equation for it in form,

$$E: \; y^2 + = x^3 +$$

for $a_1, ..., a_6 \in \mathbf{Z}$. You can reproduce an explicit "Weierstrass equation" for E over **Q** from the above "Tate-Weierstrass" model by putting

$$c_4 =; \quad c_6 = .$$

By the **discriminant** $\Delta = \Delta_E$ (of the equation defining E) we mean the integer given by

One says that the Tate-Weierstrass model for E has "good reduction" modulo a prime number p if the above equation, interpreted over $\mathbf{F}_p$, defines a smooth projective curve (and hence an elliptic curve) over $\mathbf{F}_p$. This happens if and only if the prime number p does not divide Δ_E.

I will be interested, this hour, only in *semistable* elliptic curves over **Q**; that is, I want E to be given over **Z** in such a way that its reduction over $\mathbf{F}_p$ is "good" or else is a curve of genus zero with a nodal singularity over $\mathbf{F}_p$. It is equivalent to ask, simply, that c_4 and Δ be relatively prime. If E is a semi-stable elliptic curve over **Q**, we may (cheaply) define the **conductor** N_E of E by the formula:

$$N_E := \prod_{p|\Delta} p.$$

By Wiles-Taylor, any semi-stable elliptic curve E over $\mathbf{Q}$ is *modular* in the sense that there is a nonconstant mapping defined over $\mathbf{Q}$, a *modular parametrization*

$$\phi : X_0(N) \to E,$$

where $X_0(N)$ is the modular curve whose associated Riemann surfce is the compactification of the quotient of the upper half-plane by the action of the group $\Gamma_0(N) \subset \mathrm{PSL}_2(\mathbf{Z})$ which are represented by matrices

$$\begin{pmatrix} a\, b \\ c, d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$$

with $c \equiv 0 \mod N$.

I want to focus the discussion, this hour and the next, on the following list of classifying invariants of semi-stable E's:

- **the discriminant (of a Tate-Weierstrass equation for** E **)** $= \Delta_E$,
- **the conductor** $N_E := \mathrm{rad}(\Delta_E)$,
- **the modular degree** δ_E: = the minimal degree of all modular parametrizations

$$\phi : X_0(N_E) \to E,$$

- **the (Faltings) height** $= h_E$ (to be defined in the next section of today's lecture),
- **the order** $|\mathrm{Sha}_E|$ **of the Shafarevich-Tate group of** E **(to be defined in tomorrow's lecture)**.

For the rest of this lecture I propose that we discuss, and muse about the curious fact that the ABC-conjecture is *equivalent* to knowledge of specific upper bounds for the rate of growth (relative to the conductor N_E) of *any one* of the disparate invariants $|\Delta_E|$, δ_E, $H_E := \exp(h_E)$; and also $|\mathrm{Sha}_E|$ (this last being conditional on the Birch-Swinnerton-Dyer Conjecture). As an aside, it seems to me that it might pay to think about what general profile this phenomenon might have. Are there, for example, other classes of modular or automorphic forms attached to whcih there is a list of similar disparate invariants, where the upper bound asymptotics for any one of these invariants is "deep", but knowledge of the upper bound asymptotics for one, is provably equivalent to knowledge for all?

But to return to the study of our list of invariants, let us say that a real-valued invariant $\Phi(E)$ of isomorphism classes of a given class of elliptic curves over $\mathbf{Q}$ has **minimal upper bound exponent** α if for any E in that class,

$$\Phi(E) << N_E^{\alpha+\epsilon}$$

for any positive ϵ, and α is the smallest possible exponent for which this is true (equivalently, the stated inequality holds, and for any $\epsilon > 0$ there are elliptic curves E in the given class, of arbitrary high conductor, such that $\Phi(E) > N_E^{\alpha-\epsilon}$). We also use the phrase **maximal lower bound exponent**, defined analogously.

The following theorem collects a number of known results (due to ***):

Theorem. These are equivalent (subject, in the last instance, to BST= the Birch-Swinnerton-Dyer conjecture):

- The ABC-conjecture,
- $|\Delta_E| << N_E^{6+\epsilon}$ for $|\Delta|$ ranging over semi-stable elliptic curves,
- $\delta_E << N_E^{2+\epsilon}$ for $|\delta|$ ranging over semi-stable elliptic curves,
- $H_E := \exp(h_E) << N_E^{1/2+\epsilon}$ for h ranging over semi-stable elliptic curves,
- [the following is equivalent to the above, subject to BST]

$|\mathrm{Sha}_E| << N_E^{1/2+\epsilon}$ for $|\mathrm{Sha}_E|$ ranging over all elliptic curves.

Note. In the first three instances above, the exponents given, $6, 2$ and $1/2$ are the minimal lower bound exponents.The corresponding maximal upper bound exponents are $1, 7/6$ and $1/6$ respectively (for these invariant ranging over the class of semi-stable elliptic curve over $\mathbf{Q}$. If the Birch-Swinnerton-Dyer conjecture and the Riemann hypothesis for certain Rankin-Selberg zeta functions are both true, then $1/2$ is the minimal upper bound exponent for $|\mathrm{Sha}|$ (ranging over quadratic twists of semi-stable elliptic curves). I would guess that there are semi-stable elliptic curves over $\mathbf{Q}$ of arbitrary large conductor, with trivial Shafarevich-Tate group (and if that were true, 0 would then be the corresponding maximal lower bound exponent for $|\mathrm{Sha}|$).

But now we must go more slowly, and define height.

II.2. The (Faltings) height of an elliptic curve over Q.

Consider semi-stable elliptic curves E over $\mathbf{Q}$ and as above let $c_4(E)$, $c_6(E)$ and $\Delta(E)$ stand for the correspondingly named invariants of a minimal Tate-Weierstrass equation for E over $\mathbf{Z}$, with $NN_E = \text{rad } \Delta$. Let ω_E denote a Néron differential for E, i.e., ω_E is a regular differential that reduces to a regular differential modulo p for all prime numbers p. This property determines ω_E up to sign, and ω_E can be read off from the Tate-Weierstrass equation, as follows:

The ("Faltings") height of $E_{/\mathbf{Q}}$ is given by:

$$h(E/\mathbf{Q}) := -1/2 \log \max \int_{E(\mathbf{C})} \omega_E \wedge \bar{\omega}_E.$$

To shorten the statement of some inequalities, let us adopt the following convention. Given two functions $A(E)$ and $B(E)$ defined on a class of elliptic curves E, say that

$$A(E) \asymp B(E)$$

if

$$(1-\epsilon)\cdot A(E) - \mathrm{O}_\epsilon(1) < B(E) < (1+\epsilon)\cdot A(E) + \mathrm{O}_\epsilon(1)$$

for all $\epsilon > 0$.

We have (cf. Silverman)

Proposition 1

$$h(E/\mathbf{Q}) \asymp 1/12 \log \max\{|c_4|^3, |c_6|^2\}.$$

Corollary. For all $\epsilon > 0$,

$$h(E/\mathbf{Q}) \geq (1/12 - \epsilon) \log N - O_\epsilon(1).$$

Proof. Since

$$c_4^3 - c_6^2 = 1728 \cdot \Delta,$$

we have

$$\log N \leq \log \Delta \leq \log \max\{|c_4|^3, |c_6|^2\}.$$

Aside: The height of elliptic curves of prime conductor.

It is expected, but not known, that there is an infinite number of (non-isomorphic) elliptic curves over $\mathbf{Q}$ of prime discriminant (and hence also prime conductor). This is equivalent to the Schinzel-type conjecture that the rational polynomial $(X^3 - Y^2)/1728$ takes on an infinite number of integer values of the form $\pm p$ where p is a prime number as X and Y run through integer values of the form specific (Hardy-Littlewood) conjecture that there are an infinite number of prime numbers p of the form $u^2 + 64$ for $u \in \mathbf{Z}$: for each prime number p of this form there is a pair of elliptic curves of conductor p (called "Neumann-Setzer" curves) given over $\mathbf{Z}[1/2]$ by the formulas:

$$\begin{aligned} E_1 &: y^2 = x^3 - 2ux^2 + px \\ E_2 &: y^2 = x^3 + ux^2 - 16x, \end{aligned}$$

where the sign of u is taken so that $u \equiv 1 \bmod 4$.

The elliptic curves E_1 and E_2 are 2-isogenous, one to another, and have minimal discriminants $-p^2$ and p respectively. This family of (Neumann-Setzer) curves contains all elliptic curves of prime conductor which possess a $\mathbf{Q}$-rational point of order 2, with the exception of a pair of elliptic curves of conductor 17. The curves E_1 of the Neumann-Setzer family are the only elliptic curves E of conductor a prime number p, and such that the minimal discriminant of E is distinct from $\pm p$, with the exception of five elliptic curves, of conductors $11, 17, 17, 19$, and 37.

A straight calculation gives that if c_4 and c_6 are the conventionally named invariants of the minimal Weierstrass equation for E_1, we have

$$\max\{|c_4|^3, |c_6|^2\} = (p - 2^6)(p + 2^9)^2$$

if $p \neq 73$, while if if c_4 and c_6 are the invariants for E_2 we have

$$\max\{|c_4|^3, |c_6|^2\} = (p - 16)^3$$

Corollary 2. Let E " run through" all (optimal) Neumann-Setzer elliptic curves (assuming that there are an infinity of these!). Then

$$h(E/\mathbf{Q}) \asymp 1/4 \text{ log } N.$$

Remarks. 1. One can get infinite sequences of semi-stable elliptic curves (with nonprime conductors) such that $h(E/\mathbf{Q}) \geq 1/2$ log N, as we shall see in section 3 below.

2. Although this remark is neither about elliptic curves of prime conductor, nor even about semi-stable elliptic curves, let me translate into the language of "heights of elliptic curves" Zagier's comment [Z] that if one fixes a given elliptic curve E_1 over $\mathbf{Q}$ and lets E run through all quadratic twists E_d of E_1 (where E_d means the "twist" of E by the quadratic character of discriminant d) then one also has the asymptotic estimate $h(E/\mathbf{Q}) \asymp 1/4$ log N.

——————-

To be removed from the final draft. By Taylor-Wiles, the Neumann-Setzer curves are modular. If p is the conductor of a Neumann-Setzer curve, then clearly $p \equiv 1$ mod 8. If $p \not\equiv 1$ mod 16 then (compare [M] Ch III Prop. 7.5) E_1 is equal to the "2"-Eisenstein factor of $J_0(p)$. In particular, E_1 is an optimal factor of $J_0(p)$. Is E_1 an optimal factor even when $p \equiv 1$ mod 16?

——————-

II.3. Frey (-Hellegouarch) curves.

Recall from our first lecture that by an ABC-**solution** (A, B, C) we mean a triple of nonzero integers A, B, C which sum to zero and are relatively prime. By the **power**, $P(A, B, C)$ of an ABC-**solution** (A, B, C) we mean the real number

$$P(A, B, C) := \text{log max } (|A|, |B|, |C|)/\text{log rad } (A, B, C,).$$

By the (Hellegouarch ?-) Frey curve $E = E_{A,B,C}$ associated to an ABC- solution (A, B, C), we mean the elliptic curve

$$y^2 = x(x - A)(x + B).$$

See Oesterlé's article [O] concerning this. In the case of ABC- solutions (A, B, C) satisfying the congruences $A \equiv -1 \bmod 4$ and $B \equiv 0 \bmod 16$, we have that $E_{A,B,C}$ is semi-stable. A minimal equation for $E_{A,B,C}$ can be taken to be

$$y^2 + xy = x^3 + (B - A - 1)/4 \cdot x^2 - AB/16 \cdot x.$$

One calculates

$$c_4 = -(AB + AC + BC),$$

$$c_6 = (B - A)(C - B)(A - C)/2,$$

and

$$\Delta = (ABC/16)^2 \quad ; \quad N = \text{rad } ABC.$$

Using the above data one calculates (compare [O], [Mu]) that

$$h(E/\mathbf{Q}) \asymp 1/2 \ P(A, B, C) \cdot \log N,$$

where E ranges through all (semi-stable) Frey curves attached to ABC-solutions satisfying the above congruence conditions on B and A.

Example. Take the Frey curve corresponding to the "top" of the list of ABC-solutions we gave at the end of Lecture I:

$$2 + 3^{10} \cdot 109 + (-23^5) = 0.$$

As is reported in de Weger's [de W] the corresponding elliptic curve,

$$y^2 = x^3 - 6436339x^2 - 12872682x,$$

of conductor $N = 240672$ has Mordell-Weil rank zero, and (anticipating our discussion of "Sha" in Lecture III) its Shafarevich-Tate group, Sha_E is a product of two cyclic groups of order 19.

Corollary. There is an infinity of semi-stable (Frey) curves with

$$h(E/\mathbf{Q}) \geq 1/2 \cdot \log N.$$

Proof. Form the Frey curves associated to Elkies' sequence of ABC-"near-misses" (cf. [E]: In Elkies ' notation we must take an infinite family with N(r) a fixed positive multiple of H(r) and then throw out a finite number of members of this family to get the inequality in the Corollary).

4. The ABC-conjecture and heights of semi-stable elliptic curves.

Recall that the ABC-**conjecture** asserts that for any number $\eta > 1$, there are only a finite number of ABC-solutions with $P(A,B,C) \geq \eta$.

By the **congruence** ABC-conjecture with modulus m we mean the ABC-conjecture as above, but restricted to ABC-solutions (A,B,C) such that B is divisible by m.

Proposition 2. (Oesterlé, Szpiro, Hindry,...) These statements are equivalent.

1. There exists a prime power $m = \ell^n$ such that the congruence ABC-conjecture with modulus m is true.

2. The ABC-conjecture is true.

3. $h(E_{/\mathbf{Q}}) \leq (1/2 + \epsilon) \cdot \log N \ + \ O_\epsilon(1)$ for all semi-stable elliptic curves E.

4. $h(E_{/\mathbf{Q}}) \leq (1/2 + \epsilon) \cdot \log N \ + \ O_\epsilon(1)$ for all semi-stable Frey curves E.

Remarks. The ideas behind these equivalences are (using Proposition 1) due to Oesterlé, Szpiro, and Hindry. Jordan Ellenberg showed me a simple proof of the equivalence of **1.** and **2.**. To see that **2** implies **3**, we must show that

$$\log \max \{|c_4|^3, |c_6|^2\} \leq (6 + \epsilon) \cdot \log N \ + \ O_\epsilon(1)$$

when E ranges through all semi-stable elliptic curves. For this apply the ABC-conjecture to the three-term equation

$$c_4^3 - c_6^2 - 1728\Delta = 0,$$

where A, B, and C are obtained from the three terms in the equation by removing, if necessary, common factors of 2 and 3 (cf. [O]). Clearly **3** implies **4**.

Since $h(E/\mathbf{Q}) \asymp 1/2 \ P(A,B,C) \cdot \log N$, we have that **4** implies **1** with modulus $m = 16$.

II.3 Relationship between the ABC-conjecture and modular degree.

Let E be an optimal elliptic factor of conductor N and let f_E denote the normalized newform on $\Gamma_0(N)$ attached to E. Among other things this means that if $\psi : X_0(N) \to E$ is the modular parametrization, the pullback (to the upper half-plane) of a holomorphic differential on E (via the composition of the projection of the upper half-plane onto $X_0(N)$

with ψ) yields a multiple of the differential form $f_E(\tau)d\tau$. We may use this to normalize the differential on the Riemann surface $E(\mathbf{C})$ and an identification,

$$E(\mathbf{C}) \cong \mathbf{C}/\Lambda_E$$

where $\Lambda_E \subset \mathbf{C}$ is a lattice. We may find such a lattice, and identification, such that the holomorphic differential dz on $\mathbf{C}/\Lambda_E$ is identified with the Néron differential ω_E (well-defined up to sign) on $E(\mathbf{C})$. Pulling dz back to the upper half plane (via ψ) we obtain a differential form on the upper half plane which is equal to

$$c_E \cdot 2\pi i f_E(\tau)d\tau,$$

where c_E ("Manin's constant") is a nonzero rational number, and $f_E = \sum_{n=1} a_n q^n$ is the newform associated to E ("newform" includes the requirement that f_E be normalized so that $a_1 = 1$). By appropriate choice of sign of the Néron differential, we can guarantee that c_E is positive; c_E is proven to be an integer for all conductors N (not only those that are square-free) and is conjectured always to be equal to 1. We have:

***Proposition 3** Let N be square-free. The Manin constant c_E is either 1 or 2 (and the latter possibility can not happen unless E has good, ordinary reduction in characteristic 2).

Proof. This follows by combining [M ?] and [Ra]. For further results due to Stephens and Edixhoven when N is not square-free, see [E].

Proposition 4.

$$4\pi^2 c_E \cdot (f_E, f_E) = \delta_E \cdot \int_{E(\mathbf{C})} \omega_E \wedge \bar{\omega}_E.$$

Proof. Here, (,) denotes the Petersson inner product, and by $\mathrm{Vol}(\Lambda_E)$ we mean rather the volume of a fundamental domain of the lattice $\Lambda_E \subset \mathbf{C}$. To prove this formula, just consider

$$4\pi^2 c_e \cdot (f_E, f_E) = 2\pi^2 i \int_{X_0(N)(\mathbf{C})} f_E(\tau)d\tau \wedge \overline{f_E(\tau)d\tau}$$

$$= i/2 \int_{X_0(N)(\mathbf{C})} (2\pi i f_E(\tau))d\tau \wedge \overline{2\pi i f_E(\tau)d\tau}$$

$$= i/2 \int_{X_0(N)(\mathbf{C})} {\psi_E}^*(dz) \wedge \overline{{\psi_E}^*(dz)}$$

$$= i/2 \cdot \delta_E \cdot \int_{\mathbf{C}/\Lambda_E} dz \wedge d\bar{z}$$

$$= \delta_E \cdot \int_{E(\mathbf{C})} \omega_E \wedge \bar{\omega}_E.$$

Writing the integral on the right in terms of the height of the elliptic curve, $h(E/\mathbf{Q}) := -1/2\log \int_{E(\mathbf{C})} \omega_E \wedge \bar{\omega}_E$, and noting that for semi-stable modular elliptic curves E of conductor N we have

$$(f_E, f_E) = N \cdot L(\mathrm{Symm}^2(f_E), 2)/288\ \pi^3$$

where $L(\mathrm{Symm}^2(f_E), s)$ is the L-function of the "symmetric square" of f_E, (cf.) we get:

***Proposition 5**

$$\log\ \delta_E = \log\ \ N + 2h(E/\mathbf{Q}) + \log\ \ L(\mathrm{Symm}^2(f_E), 2) + \log\ (c_E/72\pi)$$

We wish to get some estimate for the size of δ_E. By Proposition 3, we have $\log\ (c_E/72\pi) =$ O(1). We need:

***Proposition 6** (Mai-Murty ?? Hoffstein-Lockhart??): For any $\epsilon > 0$ there is a positive constant κ_ϵ such that

$$\kappa_\epsilon \cdot N^{-\epsilon} \leq L(\mathrm{Symm}^2(f_E), 2) \leq O(\log\ N).$$

NOTE:* Include fuller description of this result with commentary about implied constants. Can we get, for example, an explicit lower bound for $L(\mathrm{Symm}^2(f_E), 2)$ in terms of N, so as to get a lower bound for $\log\ \delta_E$? Compare with Yau's bound; see below. *****

Corollary 1. As E runs through all semi-stable elliptic curves, we have

$$\log\ \delta_E \asymp \log\ \ N + 2\ h(E/\mathbf{Q}),$$

(or, equivalently, in view of Proposition 1 of section 1.

$$\log\ \delta_E \asymp \log N + 1/6 \log\ \max\{|c_4|^3, |c_6|^2\}.$$

Corollary 2. For every postive ϵ we have

$$\log\ \delta_E \geq (7/6 - \epsilon)\ \log\ N + O_\epsilon(1).$$

Proof. This follows from the Corollary to Proposition 1 in section 1.

6. Explicit lower bounds for δ_E. There are various approaches to obtaining lower bounds for log δ_E in terms of N (e.g., one can try to by work out the constants, for example, in the above Corollary). There is also the approach given by Yau [Y] and Li-Yau [L-Y] (see also [Ab]) which has the advantage that it gives an explicit lower bound for the **gonality** of any congruence modular curve.

Definition. If X is a Riemann surface, its **gonality** is the minimum degree of any (nonconstant) meromorphic function on X.

In [Y], Yau shows that if X is a a congruence modular curve of genus g and having ν cusps, the gonality $\gamma(X)$ admits the following lower bound:

$$\gamma(X) \geq 3(2g-2+\nu)/64.$$

Since δ_E is visibly $\geq \gamma(X_0(N))/2$, we have:

Proposition 7. (Yau)

$$\delta_E \geq 3(2g-2+\nu)/128,$$

where g is the genus of $X_0(N)$ and ν is the number of cusps of $X_0(N)$.

Since

$$2g-2+\nu \geq 1/6\ N \prod_{p|N}(1+p) - 1/2 \prod_{p|N}(1+\left(\frac{-1}{p}\right)) - 2/3 \prod_{p|N}(1+\left(\frac{-3}{p}\right)),$$

Yau's bound gives a fairly usable estimate (e.g., it can be used to show that 131 is the largest conductor of an elliptic curve whose modular degree is 2, as was asserted in section 3 of the main text of this article).

Corollary. As $E = E_{(A,B,C)}$ runs through all semi-stable Frey elliptic curves

$$\log\,\delta_E \asymp (1+P(A,B,C)) \cdot \log\,\text{radical}(ABC).$$

Corollary 4. The four equivalent conjectures formulated in the Proposition 2 of section

$$\log\,\delta_E\ \leq (2+\epsilon)/\log\,N + O_\epsilon(1).$$

Lecture III Relationship between the ABC-conjecture and the Shafarevich-Tate group.

III.1 The Shafarevich-Tate group.

Introduction. An element in the Shafarevich-Tate group, Sha_E, of an elliptic curve E over $\mathbf{Q}$ is given by an isomorphism class of pairs $(\Sigma,\ \iota)$ where Σ is a curve over $\mathbf{Q}$ which is of genus 1 possessing a $\mathbf{Q}_\ell$-rational point for every prime number ℓ as well as an $\mathbf{R}$-rational point, and where $\iota : E \to \mathrm{jac}\ \Sigma$ is an isomorphism over $\mathbf{Q}$ between E and the jacobian of Σ. It is known that Sha_E is a torsion abelian group, that for each prime number p, the pontrjagin dual of its ("ind"-)p-primary component is a module of finite type over $\mathbf{Z}_p$. From the very definition we have given of it, the "size" of the Shafarevich-Tate group is a measure of how badly the "local-to-global principle" (also known as the "Hasse principle") fails. For example, it is clear from our definition that to say that Sha_E vanishes is equivalent to saying that any curve of genus 1 over $\mathbf{Q}$ whose jacobian is isomorphic to E has a $\mathbf{Q}$-rational point if and only if it has a $\mathbf{Q}_p$ rational point for all prime numbers p and has a real point.

Recall that the Shafarevich-Tate Conjecture is that Sha_E is a *finite* group.

If Sha_E is finite, it supports a natural alternating non-degenerate self-pairing, and therefore its order is a perfect square. Finiteness of Sha_E is now known for a class of elliptic curves, thanks to the work of Kolyvagin and Rubin. What can we say about the size of Sha_E? Goldfeld and Szpiro conjecture (for all elliptic curves, semi-stable or not; cf [G-S]) that

Conjecture (Goldfeld-Szpiro).

$$|\mathrm{Sha}_E| << N_E^{1/2+\epsilon},$$

for any positive ϵ.

They prove [G-S] that (given the Birch- Swinnerton-Dyer Conjecture) that the ABC-Conjecture is equivalent to this upper bound restricted to semi-stable elliptic curves E.

As for lower bounds, I would expect that there are semi-stable elliptic curves of arbitrarily high conductor, with trivial Sha_E. If that is so, the natural "next question" in the lower bound direction (a question which was in effect posed by de Weger, in his preprint [de W]) is to determine some positive exponent d such that there is an infinity of elliptic curves with $N_E^{d-\epsilon} << \mathrm{Sha}_E$. de Weger has shown that (at least if one broadens the class of elliptic curves ever-so-slightly more generally than semi-stable) that some "standard" analytic conjectures (Birch-Swinnerton-Dyer, and the Riemann hypothesis for Rankin-Selberg zeta functions associated to certain weight 3/2 modular forms) imply that you can take $d = 1/2$; i.e., you cannot improve on the exponent 1/2 occurring in the Goldfeld-Szpiro conjecture.

III.2 Visualizing the Shafarevich-Tate group. How well can we ”see” the curves of genus 1 classified by the Shafarevich-Tate group of elliptic curves over $\mathbf{Q}$?

Here is one instance where we can do this satisfactorily: Consider the smooth plane cubic curve $E : x^3 + y^3 + 60z^3 = 0$. By work of Kolyvagin and Rubin, one knows that

$$\mathrm{Sha}_E \cong \mathrm{Z}/3\mathrm{Z} \times \mathrm{Z}/3\mathrm{Z}.$$

Noting that change of sign of ι corresponds to multiplying by -1 in Sha_E we see that there are, in toto, four “nontrivial” Σ's to find here and after some minor computations (cf [M]) one finds all of them again as smooth plane cubics:

$$\Sigma_1 : 3x^3 + 4y^3 + 5z^3 = 0$$
$$\Sigma_2 : 12x^3 + y^3 + 5z^3 = 0$$
$$\Sigma_3 : 15x^3 + 4y^3 + z^3 = 0$$
$$\Sigma_4 : 3x^3 + 20y^3 + z^3 = 0.$$

Of course, the only elements of any Sha_E that could be realized as plane cubic curves are elements of order ≤ 3, so a pre-condition for “seeing” all of Sha_E in anything like the above format is that Sha_E be annihilated by 6. This suggests that, in general, we should look to varieties other than $\mathbf{P}^2$ as candidate “ambient spaces” within which we might hope to find the curves of genus 1 representing the elements of Sha_E.

Assume that E is an elliptic curve over $\mathbf{Q}$ of square-free conductor N. The elliptic curve E is semi-stable, and consequently (by [W] and [T-W]) modular. We shall assume even more, that E is an “optimal factor” of the jacobian $J = J_0(N)$ of the modular curve $X_0(N)$ over $\mathbf{Q}$ and therefore (see section 2 below) there is an imbedding (unique up to multiplication by ±1) of E into J. In a word, we can “find”, uniquely, the elliptic curve E as an abelian sub-variety of $J = J_0(N)$. The idea suggests itself that, since we have “found” the elliptic curve E in the ambient space J, perhaps we should look for elements of Sha in the same ambient space.

How much of Sha_E can we find represented as curves in J? If a given element x of Sha_E can be represented as a curve, defined over $\mathbf{Q}$ in J, we shall say (see the technical definition below) that x is **visible**.

As we will discuss in more detail below, the visible part Sha_E° of Sha_E is simply the kernel of the homomorphism $\mathrm{Sha}_E \to \mathrm{Sha}_J$ induced from the injection $E \to J = J_0(N)$. Loic Merel has been independently investigating (along with his student Amod Agoshe) the order of the Shafarevich-Tate group of the winding quotients of $J_0(N)$ for N prime (see XXXX). They find that ********

There is a companion to these questions which, to focus on, I'll phrase very narrowly. Let E over $\mathbf{Q}$ be an elliptic curve optimal factor of $J_0(N)$ (as above, N is the conductor of E, assumed square-free). Suppose that $\mathrm{Sha}_E \cong \mathrm{Z}/p\mathrm{Z} \times \mathrm{Z}/p\mathrm{Z}$, for p a prime number and suppose further that the Mordell-Weil group of E is finite, of order prime to p. Can we find an optimal factor A of $J_0(N)$, i.e., of the same conductor as E, with the following properties?

i. The ring of endomorphisms of A is an order $\mathcal{O} \subset K$ in a number field K.

ii. There is a maximal ideal $m \subset \mathcal{O}$ of residual characteristic p such that if $A[m]$ denotes the kernel of the ideal m in A, $A[m]$ is an $\mathbf{F}_p$-vector space of dimension 2.

iii. There is a "p-congruence between E and A"; more explicitly, there is a Galois (i.e., Gal $(\bar{\mathbf{Q}}/\mathbf{Q})$-) equivariant isomorphism of $\mathbf{F}_p$-vector spaces of dimension 2,

$$E[p] \cong A[m].$$

iv. The isomorphism above induces an isomorphism on the m-torsion in the corresponding Selmer groups.

v. If the subscript m denotes m-adic completion, the m-adic completion of the Mordell-Weil group of A (i.e. $A(\mathbf{Q})_m = A(\mathbf{Q}) \otimes_{\mathcal{O}} \mathcal{O}_m$) is free of rank two over $\mathcal{O}_m$. (This, combined with **iv.** implies that the m-part of Sha_A vanishes.)

When such an optimal factor A can be found, we shall say (but see the more technical definition below which applies more widely) that Sha_E is **explained by jumps in the rank of Mordell-Weil**. The motivation for this terminology is simply that (assuming **i-v**) although the p-congruence allows us to identify p-torsion in the Selmer group of E with that of A, this "same" Selmer group accounts for rank of the Mordell-Weil group of A while it accounts for the p-torsion in the Shafarevich-Tate group of E (whose Mordell-Weil rank is 0 and therefore needs no "accounting for"). When Sha_E is explained by jumps in the rank of Mordell-Weil, Sha_E is also "visible" (in the sense described above and defined precisely below).

If you come across an element of Sha_E of order p it might, at least at first glance, seem quite unlikely that you could explain it away by "jumps in the rank of Mordell-Weil" as described above. But consider the first two instances of nontrivial Shafarevich-Tate group for optimal semi-stable elliptic curves (i.e. the lowest conductor N for which this occurs) are for the curves $571A$ and $681B$ as given in [Cr]. Both $571A$ and $681B$ have trivial Mordell-Weil group and their Shafarevich-Tate groups are isomorphic to $\mathrm{Z}/2\mathrm{Z} \times \mathrm{Z}/2\mathrm{Z}$ and to $\mathrm{Z}/3\mathrm{Z} \times \mathrm{Z}/3\mathrm{Z}$, respectively. Checking Cremona's book [Cr] one immediately finds the happy "accident" that $571A$ seems to admit a 2-congruence with the optimal elliptic curve factor $571B$, whose Mordell-Weil rank is 2 and whose 2-part of Sha is trivial. As for $681B$, a similar "accident" happens: $681B$ seems to admit a 3-congruence with the optimal elliptic curve factor $681C$, whose Mordell-Weil rank is 2 and whose 3-part of Sha is trivial. By the phrase "seems to admit a p-congruence" what I mean here is that one checks that the required congruence mod p hold for all the traces of Frobenius tabulated in [Cr] for the elliptic curves involved. It would take further work (using, say, effective Chebotarev) to actually prove that the p-congruence holds, and this I have not done. If one does this further work in both instances, one will have shown that the Shafarevich-Tate groups are "explained by jumps in the rank of Mordell-Weil" in the sense formulated above. In particular, the Shafarevich-Tate groups involved are all "visible". To investigate numerically how much longer these "accidents" occur, Adam Logan made computations amplifying Cremona's data, and compiled the table which we reproduce in section 9 below. The surprise for us is that, restricting to the odd part of the Shafarevich-Tate groups for the moment, all of the (odd part of) the Shafarevich-Tate group of optimal elliptic curves E of (squarefree) conductor $N < 2849$ "seem to be" explained by jumps in the rank of Mordell-Weil, and in all these cases, the optimal factor A that does the explaining can be taken to be an elliptic curve (these are the curves F in the table). In particular, all odd Sha for semi-stable optimal elliptic curves of conductor < 2849 is visible. Indeed, the situation is pretty nearly the same if we include 2-primary components but the prevalence of rational 2-torsion in these elliptic curves requires that we make a more elaborate discussion of these cases. As our numerical investigations now stand, the only **guaranteed** invisible elements in Sha_E that we have found for elliptic curves E of square-free conductor is the single instance of $2849A$ where Sha_{2849A} is of order 9 (assuming Birch and Swinnerton-Dyer) and is entirely invisible. Consulting the table in section 9 one finds a few unresolved cases among which one might expect a few more instances of invisibility.

I feel that these issues deserve to be investigated further. Is the prevalence of "visibility" a phenomenon occurring only in this modest range of conductors? Is most of Sha invisible? Or is most of Sha visible? In the vocabulary of the previous lecture, what is the minimal upper bound exponent for Sha°_E as E ranges through all semi-stable optimal elliptic curves over $\mathbf{Q}$? That is,

Query! What is the minimal exponent $\alpha \geq 0$ for which the subgroup of visible elements in Sha satisfies the upper bound

$$|\mathrm{Sha}^{\circ}_E| << N_E^{\alpha+\epsilon}$$

as E ranges through all semi-stable optimal elliptic curves over $\mathbf{Q}$?

Given the ABC- conjecture and [BST] we would have

$$0 \leq \alpha \leq 1/2.$$

But at present we haven't even a guess about the size of α beyond these inequalities.

As we shall see below, the subgroup of visible elements in Sha_E is annihilated by δ_E, the modular degree of E so it is particularly relevant, in the search for invisible Sha, to find prime divisors of the order of Sha_E which do not divide δ_E.

At present writing we can prove almost nothing of a general nature. The rest of this article is devoted to making precise some of the definitions we have alluded to in this introduction, proving some related basic lemmas (all of which are either well known, or should have been) and reproducing the table of Adam Logan (based to a large part on computations of Cremona).

1. Optimal factors of the jacobians of modular curves

The varieties that appear in this paragraph will be varieties defined over $\mathbf{Q}$, unless otherwise indicated. Let N be a square-free positive integer, and let $\pi : J_0(N) \to A$ be an *optimal factor.* That is, π is a surjective homomorphism of abelian varieties over $\mathbf{Q}$ with the following properties:

1. The homomorphism π factors through the new part, $J_0(N) \to J_0(N)^{new}$.

2. The homomorphism π has connected kernel.

3. The abelian variety A is $\mathbf{Q}$-simple.

Let $\mathbf{T}_A$ be the image in the totally real field $\mathbf{K}_A := End_{\mathbf{Q}-a.v.}(A) \otimes \mathbf{Q}$ of the algebra generated by (*all*) Hecke operators T_p for primes p not dividing N and all Atkin-Lehner operators U_q for primes $q|N$. We refer to $\mathbf{T}_A$ as *the* Hecke algebra of the optimal factor A.

If B is an abelian variety, let $\hat{B}$ denote the dual abelian variety. Since $J_0(N)$ is the jacobian of a curve, the classical Poincaré divisor gives us a canonical identification

$$\iota : \hat{J}_0(N) \cong J_0(N).$$

If A is an optimal factor, and $\hat{A}$ is its dual, the Hecke algebra $\mathbf{T}_A$ acts naturally on $\hat{A}$ and the isogeny ι commutes with the action of $\mathbf{T}_A$ on its domain and range.

Dualizing the homomorphism π we get a homomorphism

$$\hat{\pi} : \hat{A} \to \hat{J}_0(N).$$

Compose the morphisms $\hat{A} \to \hat{J}_0(N) \cong J_0(N) \to A$ to give us a homomorphism

$$\iota_A := \pi\iota\hat{\pi} : \hat{A} \to A.$$

Proposition 8. the homomorphism $\iota_A : \hat{A} \to A$ is an isogeny.

Proof. Since A and $\hat{A}$ are $\mathbf{Q}$-simple and since ι_A is defined over $\mathbf{Q}$, it follows that ι_A is either an isogeny or else it is 0. But if it were 0, it would then follow that the factorization of $J_0(M)$ (up to isogeny) as product of $\mathbf{Q}$-simple abelian varieties would have at least *two* factors isogenous to A, contradicting the fact that A occurs in this factorization with multiplicity one.

Proposition 9. the homomorphism $\hat{\pi} : \hat{A} \to \hat{J}_0(N)$ is injective.

Proof. By Proposition 1, we see that the kernel of $\hat{\pi}$ is a finite subgroup(scheme) of $\hat{A}$; call it $\Phi \subset \hat{A}$. Now factor $\hat{\pi}$ as the composition of the isogeny $\hat{A} \rightarrow \hat{A}/\Phi$ and an injection $\hat{A}/\Phi \rightarrow \hat{J}_0(N)$. Dualize this composition, and use the canonical isomorphism between the double dual and the identity functor to get that π factors through an isogeny $A' \rightarrow A$ which is dual to $\hat{A} \rightarrow \hat{A}/\Phi$. Since π is optimal, this isogeny $A' \rightarrow A$ is an isomorphism. So, $\Phi = \{0\}$, and $\hat{\pi} : \hat{A} \rightarrow \hat{J}_0(N)$ is injective.

Data-gathering Problem. Cremona has tabulated the degrees of the isogenies ι_A for optimal elliptic curve factors A of conductor $N < 5,000$. It would be interesting to get some similar data for optimal abelian variety factors A (of dimensions > 1 as well). How often is it the case that the kernel of ι_A in $\hat{A}$ is the kernel of a locally principal ideal $\mathcal{I}_A \subset \mathbf{T}_A$; i.e.,

$$ker(\iota_A) = \hat{A}[\mathcal{I}_A]?$$

2. Optimal factors which are elliptic curves

Let A be an optimal factor of dimension one, i.e., an elliptic curve. Then $\mathbf{T}_A = \mathbf{Z}$. To emphasize that A is an elliptic curve, we will change letters $A = E$. We have a canonical identification of an elliptic curve with its dual, and by means of this, we identify $\hat{A}$ with E. The injective homomorphism $\hat{\pi}$ then identifies E with a sub-elliptic curve in $J_0(N)$; so that E may be considered as both sub-elliptic curve and quotient elliptic curve of $J_0(N)$, the identifications given by the injection $\hat{\pi}$ and the projection π. By the "multiplicity one" theorem, the factorization of the abelian variety $J_0(N)^{\mathrm{new}}$ into simple abelian varieties over $\mathbf{C}$ (up to isogeny) contains the isogeny class of E only once, and hence this is also true for the factorization of $J_0(N)$. Since it is also true that the automorphism group of E is $\{\pm 1\}$, it then follows that the injection $E \hookrightarrow J_0(N)$ is unique (up to multiplication by $\{\pm 1\}$; that is, we may identify E with a sub-abelian variety of $J_0(N)$ in a unique manner, up to sign.

The isogeny ι_A is the composition of π and $\hat{\pi}$; ι_A then may be identified with a $\mathbf{Q}$-rational endomorphism

$$\iota_E : E \rightarrow E$$

which is of necessity multiplication by a scalar.

Recall the definition of the "modular degree" of the (semi-stable, optimal) elliptic curve E, denoted δ_E, as introduced in the previous lecture. One way of thinking of δ_E is to consider the nautral mapping $X_0(N) \rightarrow J_0(N)$ given by sending the cusp ∞ to the origin. Let $A = E$ be our optimal elliptic curve factor of $J_0(N)$. Composing the modular parametrization $\pi_E : J_0(N) \rightarrow E$ with the "natural" mapping $X_0(N) \rightarrow J_0(N)$ (which we normalize by requiring that the cusp $\infty \in \mathcal{X}_{\prime}(\mathcal{N})$ be sent to the origin in $J_0(N)$) we get a surjective mapping of curves,

$$\psi = \psi_E : X_0(N) \rightarrow E,$$

whose degree is, by definition, δ_E.

Proposition 10. The endomorphism $\iota_E : E \to E$ is multiplication by the modular degree, δ_E.

Proof of Proposition 10. The mapping $\psi_E : X_0(N) \to E$ induces the homomorphism $\iota_E : J_0(N) \to E$ on "jacobians". The adjoint ${\psi_E}^* : Pic^o(E) \to Pic^o(X_0(N))$ is identified with $\hat{\iota}_E : \hat{E} \to \hat{J}_0(N)$. After making the natural identifications, we get

$$\delta = \psi_E \cdot {\psi_E}^* : E \to E.$$

But for any surjective mapping $\psi : X \to E$ (of any curve X onto E) $\psi \cdot \psi^*$ is multiplication by the degree of ψ. The proposition follows.

Proposition 11. Let $B = \ker(\pi) \subset J_0(N)$. Then B is an abelian subvariety of $J_0(N)$ and the intersection of the abelian subvarieties E and B in $J_0(N)$ is given by the kernel of multiplication by δ_E in E:

$$E \cap B = E[\delta_E] \subset E.$$

Proof. This is evident. We may interpret it as saying that δ_E is the largest "modulus of congruence" connecting the newform attached to E to the complementary space of modular forms of weight two on $\Gamma_0(N)$.

The importance of the modular degree was clear already from examples (and computations of congruences satisfied by modular eigenforms) due to Doi, and Shimura. The notion of modular degree plays a prominent role in the work of Hida, Ribet, Tilouine, and others. In particular, Ribet's "level-raising" and "level-lowering" theories both pivot on the "modular degree".

3. Examples. Low modular degrees.

For a more complete discussion of the modular degree, see the appendix below. To say that the modular degree is 1 is to say that $X_0(N)$ is of genus 1, and that $E = X_0(N)$ and that's all there is to say about it. This happens for $N = 11$, 14, 15, 17, 19, 20, 21, 23, 27, 32, 36, 40, 48, 49, and 64. To say that the modular degree is 2 is to say that there is an involution of the modular curve $X_0(N)$ whose quotient is the elliptic curve E. The elliptic curve E then has the property that there is an isomorphism of its 2-torsion subgroup $E[2]$ with a subgroup of the complementary abelian variety $B \subset J_0(N)$. This happens for $N = 26$, 26, 30, 34, 35, 37, 37, 38, 39, 43, 44, 45, 50, 50, 51, 53, 54, 55, 56, 61, 62, 65, 69, 79, 83, 89, 92, 94, 101, and 131. Particularly amusing are the cases $N = 26, 37$. The curve $X_0(37)$ is of genus 2, and is therefore hyperelliptic. Let $u : X_0(37) \to X_0(37)$ denote the hyperelliptic involution, and let $w : X_0(37) \to X_0(37)$ be the Atkin-Lehner involution $w = w_{37}$. These involutions can be seen to be *distinct.* Since the hyperelliptic involution of a hyperelliptic curve of genus ≥ 2 is in the center of the automorphism group, these involutions u and w commute. They generate the full group of automorphisms of the Riemann surface $X_0(37)$, which is of order four. The quotient of $X_0(37)$ by the hyperelliptic involution u is of genus zero (this characterizes the hyperelliptic involution of a hyperelliptic curve of genus ≥ 2). Denoting by E and F the quotients of $X_0(37)$ by w and by uw, one sees easily that E and F are elliptic curves, both optimal quotients of $J_0(37)$, both clearly having modular degree equal to 2, and, viewing them as sub-abelian varieties of $J_0(37)$ they are complementary sub-abelian varieties, whose intersection is precisely given by the 2-torsion subgroup of each of them:

$$E[2] = E \cap F = F[2].$$

In particular, if $e = \sum a_n(e) \cdot q^n$ and $f = \sum a_n(f) \cdot q^n$ denote the newforms attached to E and F, we have that their Fourier coefficients admit a congruence modulo 2, i.e.,

$$a_n(e) \equiv a_n(f)$$

for all $n \geq 1$; they do not admit a congruence modulo any larger integer.

4. Visible elements in the Shafarevich-Tate group.

Let E be a semi-stable, elliptic curve over $\mathbf{Q}$ of conductor N which is an "optimal" factor of $J_0(N)$. Let $J = J_0(N)$, and recall that given an optimal modular parametrization $J \to E$, we may dualize this mapping (and use the canonical self-dulality of both domain and range) to get an injection $E \to J$. In this paragraph, then we view E as a sub-elliptic curve of J. Let Σ denote an E-torsor; that is, Σ is a curve of genus 1 over $\mathbf{Q}$ given together with an isomorphism $E \cong \text{Jac}\,(\Sigma)$. Equivalently, Σ represents an element of the continuous cohomology group $H^1(G_{\mathbf{Q}}, E)$.

DEFINITION 1. Say that the E-torsor Σ is **visible** (i.e., "visible in $J_0(N)$") if the curve Σ is isomorphic over $\mathbf{Q}$ to a subvariety (a curve, of course) in the variety J.

Proposition 12. Let Σ be an E-torsor (over $\mathbf{Q}$). These are equivalent:

1. Σ is visible.

2. Σ is isomorphic (over $\mathbf{Q}$) to a subvariety (defined over $\mathbf{Q}$) of J which is a translate of $E \subset J$.

3. The pushout of the E-torsor Σ with respect to the homomorphism $E \to J$ is the trivial J-torsor; equivalently, the image of Σ under $H^1(G_{\mathbf{Q}}, E) \to H^1(G_{\mathbf{Q}}, J)$ vanishes.

Proof. Suppose **1** . Let Σ be visible, and let us identify Σ with a curve in J. Since N is squarefree, and equal to the conductor of E, the quotient abelian variety J/E (even over $\mathbf{C}$) contains no factor isogenous to E, and therefore the image of Σ in J/E is a point, which gives **2** . Now suppose **2** and note that the imbedding of Σ in J (as a translate of E) enables us to define an isomorphism from the "trivial" J-torsor (i.e., J) to $\mathcal{J} :=$ the pushout of Σ with respect ot the inclusion $E \subset J$. This gives **3** . But **3** clearly implies **1.**

In a word, the visible E-torsors are represented by the kernel of the homomorphism

$$H^1(G_{\mathbf{Q}}, E) \to H^1(G_{\mathbf{Q}}, J).$$

If $B_{/\mathbf{Q}}$ is any abelian variety, let Sha_B denote its Shafarevich-Tate group (over $\mathbf{Q}$).

Let E be an optimal elliptic curve factor, of squarefree conductor, as before. Consider the exact sequence of abelian varieties

$$0 \ \to \ E \ \to \ J \ \to \ B \ \to \ 0,$$

where $B = J/E$ is the quotient .

NOTATION. Let Sha_E° denote the **visible part** of Sha_E. That is, an element of Sha_E is in Sha_E° if and only if it is represented by a visible E-torsor.

Proposition 13.

1. The visible part of Sha_E is a subgroup and we have an exact sequence

$$0 \ \to \ Sha_E^\circ \ \to \ Sha_E \ \to Sha_J.$$

2. Let $\mathcal{B}(\mathbf{Q}) \subset B(\mathbf{Q})$ denote the subgroup of elements $\beta \in B(\mathbf{Q})$ with the property that if $\beta_v \in B(\mathbf{Q}_v)$ denotes the image of β under the mapping induced by completion $\mathbf{Q} \subset \mathbf{Q}_v$, then for all primes v (the archimedean prime included) β_v is in the image of the mapping $J(\mathbf{Q}_v) \to B(\mathbf{Q}_v)$. Clearly $\mathcal{B}(\mathbf{Q})$ contains the image of $J(\mathbf{Q})$. The long exact sequence of $G_{\mathbf{Q}}$ cohomology applied to the displayed exact sequence of abelian varieties above yields

$$Sha_E^\circ \cong \mathcal{B}(\mathbf{Q}) \ / \ \text{image } J(\mathbf{Q}).$$

Corollary. The subgroup Sha_E° is annihilated by the modular degree δ_E.

(for further discussion of modular degree, see the appendix.)

5. A "finite type" version of the Selmer group .

We have the standard exact sequence,

$$0 \to B(\mathbf{Q}) \otimes \mathbf{Q}/\mathbf{Z} \to \hat{S}(B) \to \text{Sha}_B \to 0,$$

where Sha_B is the Shafarevich-Tate group of B, and $\hat{S}(B)$ is the (pro-finite) Selmer group of B. The conjecture of Shafarevich-Tate implies that Sha_B is finite. *We assume that* Sha_B *is finite* (for *all* B) in what follows.

Now consider the natural injection

$$\iota : \mathbf{Q}/\mathbf{Z} \hookrightarrow \mathbf{R}/\mathbf{Z}$$

which induces an injection $1 \otimes \iota : B(\mathbf{Q}) \otimes \mathbf{Q}/\mathbf{Z} \hookrightarrow B(\mathbf{Q}) \otimes \mathbf{R}/\mathbf{Z}$. Let us "push out" the displayed exact sequence with respect to the mapping $1 \otimes \iota$ to get an exact sequence:

$$0 \to B(\mathbf{Q}) \otimes \mathbf{R}/\mathbf{Z} \to S(B) \to \mathrm{Sha}_B \to O,$$

where S(B) is a compact abelian topological group, given our assumption that Sha_B is finite. Let B^* be the (Néron model over $\mathbf{Z}$ of) the dual abelian variety to B. Let Selmer (B) denote the pontrjagin dual of the compact topological group $S(B^*)$. We have slipped in the dual here to keep the functor Selmer$(-)$ covariant. The abelian group Selmer(B) (after our hypothesis) is finitely generated abelian. We shall refer to it as the *finite type* **Selmer group** of B. The rank of the Mordell-Weil group of B is the rank of the finitely generated abelian group Selmer(B) and we can "retrieve" (the dual of) Sha_{B^*} as the torsion subgroup in Selmer(B). If $B_{/\mathbf{Q}}$ admits a commutative ring of endomorphisms R (i.e., defined over $\mathbf{Q}$) then $\hat{B}$, $S(B)$ and Selmer(B) inherits an R-action, as does Sha_B .

6. The Selmer group of the jacobian of modular curves, and of optimal factors. The example that interests us primarily is when $B = J = J_0(N)$, the jacobian of the modular curve $X_0(N)$ for N a squarefree positive integer. We have a natural action of the (full) Hecke algebra $\mathbf{T}$ on J, and therefore on Selmer(J).

Continuing with this example, consider now an (optimal) projection $\pi : J_0(N) \to A$ where A is an optimal factor, as in section 1. In particular, we have $\mathbf{T}_A \subset K_A$ where $\mathbf{T}_A$ denotes the image of the Hecke algebra in the endomorphism ring of A, and $K_A = \mathbf{T}_A \otimes \mathbf{Q}$ is a totally real field. Recall the diagram

$$(*) \qquad A^* \hookrightarrow J \to A,$$

which induces a diagram of homomorphisms of ind-quasi-finite flat groups $\hat{A}^* \to \hat{J} \to \hat{A}$, and therefore it also induces homomorphisms $S(A^*) \to S(J) \to S(A)$ and

$$\mathrm{Selmer}(A^*) \to \mathrm{Selmer}(J) \to \mathrm{Selmer}(A).$$

The homomorphism Selmer$(J) \to$ Selmer(A) is equivariant with respect to the action of $\mathbf{T}$ (and $\mathbf{T}_A$), and therefore we have an induced homomorphism of $\mathbf{T}_A$-modules:

$$\lambda_A : \mathrm{Selmer}(J) \otimes_{\mathbf{T}} \mathbf{T}_A \to \mathrm{Selmer}(A).$$

Let $J_0(N) \to A$ be an optimal factor, and $m \subset \mathbf{T}_A$ a maximal ideal. We view m as a point of $\mathrm{Spec}\mathbf{T}_A \subset \mathrm{Spec}\mathbf{T}$. Let the subscript m attached to a $\mathbf{T}$-algebra denote "completion at m". In particular, $\mathbf{T}_m \to \mathbf{T}_{A,m}$ denotes the homomorphism induced from $\mathbf{T} \to \mathbf{T}_A$ on completions with respect to m. Let the subscripts m attached to J, A or A^* denote associated m-divisible subgroup-schemes, i.e. the p-divisible groups, $J_m = \bigcup_{\nu=1}^{\infty} J[m^\nu]$, and $A_m = \bigcup_{\nu=1}^{\infty} A[m^\nu]$, etc. , where p is the residual characteristic of m. Passing to m-divisible group schemes attached to the diagram **(*)** gives a diagram

$$(**) \qquad A^*_m \hookrightarrow J_m \to A_m,$$

where A_m^* may be taken to be either the m-divisible group scheme attached to the dual abelian variety A^* to A, or also the "Cartier" dual of the m-divisible group scheme A_m.

Proposition 14. Let $m \in \mathrm{Spec}\mathbf{T}_A$ be a maximal ideal whose associated residual representation is absolutely irreducible and is not of residual characteristic dividing 2N. Then the induced homomorphism

$$\lambda_{A,m} : \mathrm{Selmer}(J)_m \otimes_{\mathbf{T}_m} \mathbf{T}_{A,m} \to \mathrm{Selmer}(A)_m$$

is an isomorphism.

Proof. Since the associated residual representation of m is absolutely irreducible, and is not of characteristic 2, we may identify

$$S(J)_m \cong H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, \mathcal{J}_m)$$

and

$$\{Sha_E\}_m \cong H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, \mathcal{J})_m,$$

where in the displayed formulas above, $\mathcal{J}$ is the Néron model of $J_0(N)$ over Spec $\mathbf{Z}$; for this see [M]. One needn't make use of the flat topology here, and one can perfectly well follow through (what would essentially be a very close paraphrase of) the proof we are about to give using the more traditional definitions via Galois cohomology. We leave this proof in more elementary langauge as an exercise to the reader, and simply record the more rapid proof below using sheaves for the flat topology.

Consider the functor $H^1 = H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, -)$ (i.e., flat one-dimensional cohomology) on the category of ind-quasi-finite flat group schemes isomorphic to subquotients of J_m. H^0 of any such subquotient is zero, and H^1 is left-exact on this category. Moreover, $H^1_{\text{finite flat}}(\text{Spec } \mathbf{Z}, {J_m}^o) \to H^1_{\text{finite flat}}(\text{Spec } \mathbf{Z}, J_m)$ is an isomorphism and (since m is of residual characteristic different from 2) we may identify the Pontrjagin dual of either domain or range with the m-adic completion of Selmer(J). One computes that for any ideal $\mathcal{A} \subset \mathbf{T}$,

$$(***) \quad H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, J_m[\mathcal{A}]) \cong H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, J_m)[\mathcal{A}].$$

Now let $\mathcal{A} \subset \mathbf{T}$ be the kernel of $\mathbf{T} \to \mathbf{T}_A$ and note that by [R] Th. 5.2 (b) we have that the Pontrjagin dual of $J_m(\bar{\mathbf{Q}})$ is free of rank two over $\mathbf{T}_m$ and therefore so is the Pontriagin dual of $J_m(\bar{\mathbf{Q}})[\mathcal{A}]$ free of rank two over $\mathbf{T}_{A,m}$. It follows by an easy argument then, that $A_m^* = J_m[\mathcal{A}]$. Consequently **(***)** may be written as follows:

$$H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, A_m^*) \cong H^1_{\text{flat}}(\text{Spec } \mathbf{Z}, J_m)[\mathcal{A}].$$

Passing, now, to Pontrjagin duals, we have

$$\mathrm{Selmer}(J) \otimes_{\mathbf{T}} \mathbf{T_m}/\mathcal{A} \cong \mathrm{Selmer}(A) \otimes_{\mathbf{T}_A} \mathbf{T}_{A,m}$$

which is what we want.

DEFINITION Let us say that the m-primary component of Sha_A is **"explained by" jumps in the rank of Mordell-Weil** if:

a. The abelian group $\mathrm{Selmer}(J_0(N))_m$ is torsionfree,
and

b. the induced homomorphism

$$\lambda_{A,m} : \mathrm{Selmer}(J)_m \otimes_{\mathbf{T}_m} \mathbf{T}_{A,m} \to \mathrm{Selmer}(A)_m$$

is an isomorphism.

Note that since $\mathbf{T}_A$ is not flat over $\mathbf{T}$ (unless $J = A$) and $\mathrm{Selmer}(J)$ need not be locally free over $\mathbf{T}$, it is perfectly conceivable that for some values of N, the abelian group $\mathrm{Selmer}(J)$ is torsionfree (equivalently, Sha_J is trivial) , and the homomorphisms λ_A are isomorphisms (for optimal factors A of $J_0(N)$) and yet some Sha_A's do not vanish. Indeed, the numerical examples below suggest that phenomena of this sort happen more than one might at first expect.

Proposition 15. If E is an optimal factor of $J_0(N)$ and if p is a prime number for which the p-primary component of Sha_E is "explained by" jumps in the rank of Mordell-Weil, then the p-primary component of Sha_E is visible (and is annihilated by δ_E).

Proof. This is evident.

7. Primes of fusion, and Mordell-Weil "jumps". If A is an optimal factor of $J_0(N)$, by the **Mordell-Weil rank** of A we mean

$$\rho(A) := \dim_{K_A}(A(\mathbf{Q}) \otimes \mathbf{Q}).$$

By a **fusion prime** $m \in \mathrm{Spec}\ \mathbf{T} = \mathrm{Spec}\ \mathbf{T}_0(N)$ let us mean a (maximal) prime ideal contained in $\mathrm{Spec}\ \mathbf{T}_{A_1}$ and $\mathrm{Spec}\ \mathbf{T}_{A_2}$ where A_1 and A_2 are *distinct* optimal factors of $J_0(N)$. If A_1 and A_2 are the *only* optimal factors A of $J_0(N)$ such that $m \in \mathrm{Spec}\ \mathbf{T}_A$ and if the residual characteristic of m is different from 2 let us say that m is an **ordinary** fusion prime. If m is an ordinary fusion prime, by the **Mordell-Weil jump** at m we mean the non-negative integer

$$j(m) := |\rho(A_1) - \rho(A_2)|.$$

8. Remarks on the Numerical Examples. Cremona's book [Cr] lists all optimal elliptic curve factors (referred to there as "strong Weil curves") of conductor < 1000 and among these there are *only four instances* where the (Birch Swinnerton-Dyer conjectured) value of the order of the Shafarevich-Tate group is > 1. In three of these instances the (conjectured) value is 4. If we want *odd* prime divisors of the order of Sha, we are left with precisely one example: namely it is given by the curve $E :=$ **681B1** which has $\rho(E) = 0$ and $|Sha_E| = 9$. The prime 3 seems to be a prime of fusion for the optimal factor E because there is another optimal elliptic curve factor $F :=$ **681B1** such that the Hecke eigenvalues for newforms corresponding to E and F are congruent modulo 3, at least for the Hecke operators T_p for $p \leq 97$ (according to Cremona's table on page 272 of [C]). Since the Galois representations on $E[3]$ and $F[3]$ are irreducible (which one also deduces from the facts given in [C]) it is natural to imagine that they are indeed isomorphic (this could be determined rigorously by some further computation). The curve F has the property that Selmer(F) is a free abelian group on two generators. In particular, the jump at the prime m of fusion in $\mathbf{T}_0(681)$ corresponding to the prime 3 in $\mathbf{T}_E = \mathbf{Z}$ and $\mathbf{T}_F = \mathbf{Z}$ is equal to 2.

Remark. For m a prime of fusion, I sometimes find it convenient to "sketch" Spec$\mathbf{T}_0(N)_m$ labelling each irreducible component by the letter denoting the optimal factor of conductor N corresponding to that irreducible component, and also "decorating" each irreducible component by the corresponding Mordell-Weil rank "ρ". For instance in the case of the prime of fusion m in $\mathbf{T}_0(681)$ of residual characteristic 3 which fuses the two elliptic curves $E = 681B$ and $F = 681C$ we would just have the simple diagram

9. Adam Logan's data.

Adam Logan studied all instances of nontrivial Shafarevich-Tate groups for elliptic curves E of square-free conductor < 3000. Indeed his calculations extended to 5077, and we include this data below, but the tables are still relatively incomplete in the higher ranges. As it turns out, all of the (semi-stable) elliptic curves with nontrivial Shafarevich-Tate group with conductor < 5077 have Mordell-Weil rank 0. The nontrivial Shafarevich-Tate groups in this range are either of order p^2 for $p = 2, 3,$ or 5 or else of order 16. In discussing the data, it is useful to distinguish between instances where the Shafarevich-Tate group is of odd order or of order a power of two (these being the only cases that arise in the range tabulated).

the odd order cases: Logan finds only one case where an odd nontrivial p-primary component of Sha_E is provably "unexplained" among all the instances he investigated. The one provably "unexplained case" is given by the curve $E = 2849A$ which has Sha_E of order 9, but modular degree not divisible by 3. In all other cases where an odd prime number p divides the order of Sha, Logan found much the same pattern as was exhibited by the example given above, of conductor 681. Namely, for optimal elliptic curve factors E (other than $2849A$) for which an odd prime p divides the order of Sha_E, Logan finds (by means of the data available, thanks to John Cremona, via anonymous ftp) another optimal *elliptic curve* factor F of the same conductor as E such that p "seems to be" a prime of fusion, fusing E to F and F seems to provide the "explanation" for the nontrivial Sha of E, in that the order of Sha_E is p^2 and the Mordell-Weil rank of E is trivial, while F has trivial Sha but Mordell-Weil rank 2. In one case (conductor 2534) three optimal elliptic curve factors seem to be fused together at the prime $p = 3$. Two of these elliptic curves have Mordell-Weil rank zero and Sha of order 9 and the third seems to be the "explanatory" optimal factor: it has trivial Sha but Mordell-Weil rank equal to 2.

the "invisible" example: $2849A$. Since this is our first square-free example, it may be worth looking a bit more closely at it. Both Loic Merel and Richard Taylor have suggested that one test to see if its Shafarevich-Tate group becomes visible in $J_1(2849A)$. This is a natural suggestion in view of Glenn Stevens ideas regarding J_1-optimality (cf. [St]) but we have not yet made this test.

examples where Sha is of even order. As for the cases where the order of Sha is even, a similar pattern is found. In all but one of these cases, either the Sha_E is "explained" by the existence of a congruence modulo 2 between E and another elliptic curve F which has Mordell-Weil rank 2 (or in one instance, Mordell-Weil rank one), or else E has a point of order 2. In the remaining instance, $E = 3017A$, the order of Sha_E is 4, and E admits no congruence to any of the other elliptic curves of its conductor. The modular degree δ_E is equal to 1944 and therefore (in contrast to the case $2849A$) it is conceivable that Sha_E is "explained" by a congruence modulo two to an optimal abelian variety factor A of $J_0(3017)$ such that dim $A > 1$, with Mordell-Weil of rank two (over its Hecke algebra); we have not yet investigated whether this is the case. One feature peculiar to the prime $p = 2$ is that it is possible for two optimal factors of $J_0(N)$ to admit a congruence modulo $p = 2$ and have the property that they have different sign in their functional equations. This happens in the table below exactly once ($2886A$ has Sha of order 4 and is congruent mod 2 to an elliptic curve with Mordell-Weil rank one).

In the table below, the data compiled by Logan is reproduced. The elliptic curves E of squarefree conductor with nontrivial Sha_E are listed, together with the corresponding elliptic curve F of positive Mordell-Weil rank which "explains" Sha_E (except in the cases where it doesn't exist). If there is no indication to the contrary, the " congruence prime" (or "prime of fusion") linking E and F is is $\sqrt{|\mathrm{Sha}|}$. The modular degrees δ_E are tabulated, as are δ_F when they are known.

E	$\sqrt{\lvert\mathrm{Sha}_E\rvert}$	δ_E	**F**	δ_F	Remarks
571A	2	$2^3 \cdot 3 \cdot 5$	**571B**	$2^4 \cdot 3$	
681B	3	$3 \cdot 5^5$	**681C**	$2^5 \cdot 3$	
1105A	2	$2^5 \cdot 5^5$	**none**	–	2-torsion
1246B	5	$2^6 \cdot 3^4 \cdot 5$	**1246C**	$2^6 \cdot 5$	
1309A	4	$2^7 \cdot 3 \cdot 7 \cdot 17$	**1309B**	2^8	
1365F	2	$2 \cdot 3^4$	**none**	–	2-torsion

1443D	2	$2^4 \cdot 3^3 \cdot 5$	**1443C**	$2^4 \cdot 7$	2-torsion
1613B	2	$2^4 \cdot 19$	**1613A**	$2^4 \cdot 5$	
1717A	2	$2^3 \cdot 41$	**1717B**	$2^3 \cdot 13$	
1738B	2	$2^11 \cdot 3^3 \cdot 7$	**1738A**	2^8	
1785G	2	2^7	**none**	–	2-torsion
1785H	2	$2^7 \cdot 3$	**none**	–	2-torsion
1913B	3	$3 \cdot 103$	**1913A**	$2^2 \cdot 3 \cdot 5^2$	
2006E	3	$2^6 \cdot 3 \cdot 5 \cdot 7 \cdot 23$	**2006D**	$2^7 \cdot 3$	
2035A	2	$2^4 \cdot 3 \cdot 5$	**2035C**	$2^4 \cdot 3 \cdot 11$	
2089D	2	$2^5 \cdot 3 \cdot 5$	**2089E**	$2^5 \cdot 11$	
2145D	2	$2^8 \cdot 3^7$	**none**	–	
2145G	2	$2^6 \cdot 3$	**none**	–	cong. mod 4 to 2145D
2265A	2	$2^5 \cdot 3^2 \cdot 5^2 \cdot 7$	**2265B**	$2^5 \cdot 5 \cdot 7$	cong. mod 4
2310B	2	$2^9 \cdot 7^2$	**none**	–	2-torsion
2405B	4	$2^3 \cdot 3 \cdot 5^2 \cdot 7 \cdot 17$	**2406C**	$2^4 \cdot 7 \cdot 13$	
2405D	2	$2^5 \cdot 19$	**none**	–	4-torsion
2409B	2	$2^9 \cdot 5^2$	**2409D**	$2^5 \cdot 7^2$	
2429B	3	$2 \cdot 3 \cdot 73$	**2429D**	$2^3 \cdot 3 \cdot 13$	
2534E	3	$2^2 \cdot 3^2 \cdot 5^3 \cdot 11$	**2534G**	$2^5 \cdot 3^2 \cdot 13$	
2534F	3	$2^2 \cdot 3^2 \cdot 5 \cdot 7$	**2534G**	$2^5 \cdot 3^2 \cdot 13$	
2554B	2	$2^5 \cdot 13$	**2554C**	$2^4 \cdot 3^2 \cdot 7$	
2563C	2	$2^6 \cdot 3 \cdot 7$	**2563D**	$2^4 \cdot 3 \cdot 5$	
2665A	2	$2^6 \cdot 5$	**none**	–	2-torsion
2674B	3	$2^4 \cdot 3^3 \cdot 13$	**2674A**	$2^4 \cdot 3^2$	
2678A	2	$2^9 \cdot 3^2 \cdot 23$	**2678B**	$2^7 \cdot 3$	cong. mod 4
2678A	2	$2^9 \cdot 3^2 \cdot 23$	**2678I**	$2^5 \cdot 3 \cdot 11$	cong. mod 2
2710C	3	$2^5 \cdot 3^3 \cdot 7$	**2710B**	$2^5 \cdot 3^2$	
2710A	2	$2^5 \cdot 3 \cdot 5^2$	**2710B**	$2^5 \cdot 3^2$	
2710A	2	$2^5 \cdot 3 \cdot 5^2$	**2710D**	$2^5 \cdot 5 \cdot 11$	
2742B	4	$2^6 \cdot 5 \cdot 17 \cdot 23$	**2742C**	$2^6 \cdot 5$	rat'l 2-torsion; cong. mod 4
2834D	5	$2^2 \cdot 3^5 \cdot 5 \cdot 109$	**2834C**	$2^6 \cdot 3^2 \cdot 5$	
2849A	3	$2^5 \cdot 5 \cdot 61$	**NONE**	–	
2886A	2	$2^9 \cdot 3^2$	**2886B**	$2^8 \cdot 3$	2886B has 2-torsion, rank 1
2955B	3	$2^3 \cdot 3^5 \cdot 5$	**2955C**	$2^6 \cdot 3^3$	
3017A	2	$2^3 \cdot 5^2 \cdot 11$	**none**	–	
3054A	5	$2^3 \cdot 5^2 \cdot 11$	**3054C**	$2^4 \cdot 3 \cdot 5 \cdot 7$	
3306B	3	?	**none**	–	
3370D	2	?	**3370E**	?	
3742A	2	?	**3742B**	?	
3774G	2	?	**3774D**	?	cong. mod 4
3686D	2	?	**3686E**	?	cong. mod 4
3883B	2	?	**3883A**	?	
3886B	2	?	**3886G**	?	
3954C	3	?	**3954D**	?	

3995A	2	?	**none**	–
4229A	3	$2^5 \cdot 3 \cdot 5^2$	**none**	–
4334B	3	?	**none**	–
4630D	2	?	**4630B,4630C**	?
4749A	2	?	**4749B**	?
5073D	3	?	**none**	–