SHIMURA CURVES LECTURE NOTES 11: INTEGRAL STRUCTURES, GENERA AND CLASS NUMBERS

PETE L. CLARK

1. Formulae for genera, volumes, class numbers

Throughout these notes, we shall use D to denote the product of an even number (possibly zero) of distinct primes, and D' the product of an odd number of distinct primes. Thus D (resp. D') is the discriminant of a unique indefinite (resp. definite) rational quaternion algebra $B/\mathbb{Q}$ of discriminant D (resp. D'). Recall that the discriminant map induces a bijection from the set of isomorphism classes of rational quaternion algebras over $\mathbb{Q}$ to the set of squarefree positive integers.

We let N denote a *squarefree* positive integer which is prime to D.

We shall be concerned with the Shimura curves $X_0^D(N)$, defined, at present, as nonsingular, complete geometrically connected algebraic curves over $\mathbb{Q}$. (We are taking $X_0^1(N) = X_0(N)$ to be the usual – compactified – elliptic modular curve.)

In this lecture, we will define integral models on $X_0^D(N)$ and use these integral structures to prove some basic formulae on genera, volumes and class numbers.

Definition: For a squarefree positive integer m, let $d(-m)$ be the discriminant of $\mathbb{Q}(\sqrt{-m})$, i.e., $-m$ if $m \equiv 3 \pmod 4$ and $-4m$ otherwise.

Definition: For coprime squarefree positive integers A and N and a squarefree positive integer m, define

$$e_m(A,N) = \prod_{p \mid A} 1 - \left(\frac{-d(m)}{p}\right) \prod_{q \mid N} 1 + \left(\frac{-d(m)}{q}\right). \tag{1}$$

Exercise 1:
a) Show that $e_1(D,N)$ is the number of $\mathbb{Z}[\sqrt{-1}]$-CM points on $X_0^D(N)$, and also the number of order 2 elliptic points on $X_0^D(N)$.[1]
b) Show that $e_3(D,N)$ is the number of $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$-CM points on $X_0^D(N)$, and also the number of elliptic points of order 3.

Since $\Gamma_0^D(N)$ is an arithmetic Fuchsian group with trace field $\mathbb{Q}$, the only possible orders of elliptic points are 2 and 3.

[1]It is thus unfortunate that we cannot denote this quantity by e_2. We appear to have painted ourselves into a corner notationwise.

We write $e_\infty(D,N)$ for the number of cusps on $X_0^D(N)$; recall that this is $2^{\#q|N}$ if $D=1$ and 0 if $D>1$.

Theorem 1. *The genus of $X_0^D(N)$ is*

$$g(D,N) = 1 + \frac{1}{12}\prod_{p \mid D}(p-1)\prod_{q \mid N}(q+1) - \frac{e_1(D,N)}{4} - \frac{e_3(D,N)}{3} - \frac{e_\infty(D,N)}{2}.$$

Theorem 2. *(Shimizu) The volume of a fundamental region for $X_0^D(N)$ is*

$$\frac{\pi}{3}\prod_{p \mid D}(p-1)\prod_{q \mid N}(q+1).$$

Remark: For the reader's convenience, we record also a more general volume formula for a Shimura curve corresponding to a quaternion algebra B of type $(1, g-1)$ and discriminant $\mathcal{D}$ over a totally real field F (and no level structure), namely

$$\mathrm{Vol}(\Gamma^{\mathcal{D}}(1)) = \frac{1}{\pi}\zeta_F(2)\delta^{\frac{3}{2}}(4\pi^2)^{1-[F:\mathbb{Q}]}\prod_{v \mid D}(Nv-1).$$

Here ζ_F is the Dedekind zeta function and δ is the discriminant of $F/\mathbb{Q}$. (See Vignéras' book for a proof.) As we discuss briefly at the end, our method allows us to derive these formulas for certain other totally real fields F, but not in general; however, the method fails for very interesting reasons.

For any Fuchsian group Γ of the first kind, write $V(\Gamma) = \frac{-1}{2\pi}\mathrm{Vol}(\Gamma)$. Then we have the basic formula

$$2 - 2g(X) = V(\Gamma) + \sum_{n \geq 1}\left(\frac{n-1}{n}\right)e_n, \tag{2}$$

where e_n is the number of elliptic points of order n and e_∞ is the number of cusps.[2] Thus, in light of our formulas (1) and (2) for the ramification indices, Theorems 1 and 2 are equivalent. We shall prove Theorem 1 and deduce Theorem 2; note that this is the reverse of the usual state of affairs.

Let us denote by $h(D',N)$ the class number of a level N Eichler order in the definite rational quaternion algebra of discriminant D'.[3]

Theorem 3. *(Eichler)*

$$h(D',N) = \frac{1}{12}\prod_{p \mid D'}(p-1)\prod_{q \mid N}(q+1) + \frac{e_2(D',N)}{4} + \frac{e_3(D',N)}{3}.$$

Note the striking similarity between the formulas for $g(D,N)$ and $h(D',N)$!

As one might imagine, Theorem 3 admits a purely algebraic proof (Eichler's proof is purely algebraic) and Theorem 2 admits an analytic proof. Of course, working in such different contexts one does not gain insight into why the formulas of Theorems 1 and 3 have such a similar form, a phenomenon which is just as interesting (and

[2]We again apologize for the notational dissonance concerning elliptic points of order 2.

[3]Our convention here is to write D for the discriminant of an indefinite rational quaternion algebra and D' for the discriminant of a definite rational quaternion algebra.

perhaps more interesting) than the formulae themselves.

The method we present here proves Theorems 1 and 3 simultaneously. Indeed we induct on the number of primes dividing the quaternionic discriminant, hence prove cases of Theorem 3 and Theorem 1 in alternate steps.

I would hesitate to call the proof presented here "new," although I doubt that it has been written down before. At least the spirit of the proof is very well-known: it is an exploitation of what we shall call the **Jacquet-Langlands phenomenon**.[4] This is a many-faceted thing: an adequate discussion would involve deep ideas in the following areas: (i) theory of automorphic forms on $B^\times$ (for all rational quaternion algebras $B/\mathbb{Q}$); (ii) classical modular forms on $X_0^D(N)$ and $X_0(DN)$; (iii) morphisms between the Jacobians of these modular curves; and (iv) relations between the bad reductions (at distinct primes!) of the integral models of these curves; and (v) monodromy pairings.

The importance of the Jacquet-Langlands phenomenon in modern arithmetic geometry cannot be overestimated. Among other spectacular applications, it lies at the heart of Ribet's level-lowering theorem (and hence, of the proof of Fermat's Last Theorem). One may also say that it ensures the relevance of the arithmetic of Shimura curves to the arithmetic of elliptic curves. Nowadays, if $E_{/\mathbb{Q}}$ is an elliptic curve of conductor ND (in the above notation), then we know to consider not only the modular parameterization $X_0(ND) \to E$ but also the Shimura curve parameterization $X_0^D(N) \to E$.

Sadly, we are at the end of our course and will not get to describe these weighty matters in any detail. Thus we offer up our simultaneous proof of Theorems 1 and 3 (and the entailing theory of good and bad reduction of Shimura curves) as a tasty morsel of the Jacquet-Langlands phenomenon. We hope the reader will be inspired to take a bigger bite.

Exercise 2: (Shimura curves of low genus): Let D and N be squarefree corprime positive integers, with $D > 1$.
a) Show that the Shimura curves $X_0^D(N)$ (with N squarefree!) are precisely the following:

$$(N, D) = (6, 1),\ (6, 7),\ (10, 1),\ (22, 1).$$

b) Show that $X_0^D(N)$ has genus one precisely for

$$(N, D) = (6, 13),\ (10, 7),\ (14, 1),\ (15, 1), (21, 1),\ (33, 1),\ (34, 1),\ (46, 1).$$

c) All of the above Shimura curves have points rational over $\mathbb{Q}(\sqrt{-1})$ or $\mathbb{Q}(\sqrt{-3})$. Why?

2. Introducing the Integral Canonical Model

Consider the functor $F(D, N)$ on the category of all schemes, which associates to any scheme S the set of isomorphism classes of structures (A, ι, Q_N), where $A_{/S}$ is an abelian scheme of relative dimension 2, $\iota : \mathcal{O}_D \hookrightarrow \mathrm{End}_S(A)$ is a QM structure (here $\mathcal{O}_D$ is a maximal order), and $Q_N \subset A[N]$ is a rank N^2 subgroup scheme

[4]But it would be best to attach at least the following other names: Shimizu, Eichler, Shimura, Ribet, Faltings, Helm.

which is, fppf locally on S, cyclic as an $\mathcal{O}_D$-module. Moreover, we require the QM structure to satisfy the following condition (which is automatic in characteristic 0): for all $b \in \mathcal{O}_D$, the trace of $\iota(m)$ acting on the Lie algebra of A coincides with the reduced trace $tr(m)$.

Note that upon restriction of $\mathbb{Q}$-schemes, this precisely the functor which was coarsely represented by $X_0^D(N)_{/\mathbb{Q}}$.

Theorem 4. *(Morita) The functor $F(D,N)$ is coarsely represented by a $\mathbb{Z}$-scheme which is flat, quasi-projective and of relative dimension one. It is proper iff $D > 1$.*

In other words, the solution of the moduli problem can be viewed as a curve over $\mathbb{Z}$ (or, an "arithmetic surface") whose generic fiber is canonically isomorphic to the canonical model $X_0^D(N)_{/\mathbb{Q}}$. We will denote the scheme by $X_0^D(N)_{/\mathbb{Z}}$, the **integral canonical model**.

As usual, although demonstrating the representability of the functor is nontrivial, much more interesting is an explicit description of the corresponding scheme, or at least of its special fibers $X_0^D(N)_{/\mathbb{F}_p}$. Rather unsurprisingly, we get qualitatively different behavior depending upon whether p is prime to ND, whether p divides N or whether p divides D.

3. Special Fibers I: smooth reduction

Theorem 5. *Let p be prime to ND. Then $X_0^D(N)_{/\mathbb{F}_p}$ is smooth.*

In this (best) case, one can develop a theory which looks quite similar to the $D = 1$ case of good reduction of $X_0(N)$. In particular, one can describe the locus $X_0^D(N)(\mathbb{F}_q)$ of $\mathbb{F}_q$-rational points and can compute the zeta function. Here we will concentrate on the dichotomy between ordinary points and supersingular points.

Recall that a g-dimensional abelian variety over an algebraically closed field k of characteristic $p > 0$ is **ordinary** if $\#A(k)[p] = p^g$ (the largest possible value, in view of the self-Cartier duality of $A[p]$). There are various equivalent definitions for an abelian variety to be **supersingular**, the most elementary being that it is **isogenous** to E^g, where E is a supersingular elliptic curve. When $g \leq 2$, it is equivalent to require $\#A(k)[p] = 1$, but – beware – this is no longer the case when $g \geq 3$. (The "correct" definition is in terms of the slopes of the Dieudonné module of the p-divisible group of A, which we do not rehearse here.) Note well that in dimension $g \geq 2$ an abelian variety need not be either ordinary or supersingular; e.g. consider the product of an ordinary elliptic curve with a supersingular elliptic curve. (This is not the only counterexample in dimension 2.)

For the initiates, we also recall the definition of **superspecial**, namely an abelian variety which is *isomorphic* to the product of supersingular elliptic curves. Clearly there is no distinction between supersingular and superspecial in dimension 1; however, in all dimensions g, there are only finitely many principally polarized superspecial abelian varieties – indeed, each has field of moduli in $\mathbb{F}_{p^2}$ – whereas the supersingular locus is a closed subvariety of dimension $\lfloor \frac{g^2}{4} \rfloor$.

Put $\mathbb{F} = \overline{\mathbb{F}_p}$.

Note that in the next result we do *not* need to make any assumptions about p.

Proposition 6. *Let (A, ι) be a QM surface defined over $\mathbb{F}$. Then $A \sim E^2$ for some elliptic curve $E_{/\mathbb{F}}$.*

In particular, any QM surface defined over $\mathbb{F}$ *is* either ordinary or supersingular. This is one of many analogies between QM surfaces and elliptic curves.

Proof: This follows from the classification of endomorphism algebras of abelian surfaces over finite fields (Honda-Tate theory), which has a quite different character from the characteristic zero endomorphism algebra. Indeed, whereas the maximal endomorphism algebras in characteristic zero are those of CM type, over $\mathbb{F}$ CM is the *minimal* endomorphism algebra. (Moreover, an abelian variety over $\mathbb{F}$ is ordinary if and only if its endomorphism algebra is of CM type). We give only the following (easy) consequence of the classification: namely, the only possible division algebras are a quartic CM field and a quaternion algebra D with center an imaginary quadratic field K, such that $p\mathfrak{o}_K = \mathfrak{p}_1\mathfrak{p}_2$, such that D is ramified precisely at $\mathfrak{p}_1$ and $\mathfrak{p}_2$. Obviously a number field does not admit B as a subalgebra. To rule out the second algebra, we consider two cases.
Case 1: p does not divide D. Then there does not exist a $\mathbb{Q}$-algebra embedding $B \hookrightarrow D$, for if so, we tensor up to $\mathbb{Q}_p$, getting $M_2(\mathbb{Q}_p) \hookrightarrow D_{\mathfrak{p}_1} \oplus D_{\mathfrak{p}_2}$. Since the former algebra has nilpotent elements and the latter does not, this is a contradiction.
Case 2: $p \mid D$. We claim that every QM surface (A, ι) is supersingular in this case. Indeed, an abelian surface, if not supersingular, has a nonzero (naive, aka étale) p-adic Tate module: $T_p(A) \cong \mathbb{Z}_p^i$ with $0 < i \leq 2$. Then we get an action of $B \otimes \mathbb{Q}_p = B_p$ on $V_p(A)$ which (as usual!) takes 1 to 1 so is injective, giving an embedding $B_p \hookrightarrow M_i(\mathbb{Q}_p)$. Since $i \leq 2$, this is a contradiction.

Remark: This result does not hold for fields of positive transcendence degree. Indeed, fix D and let k be the algebraic closure of $\mathbb{F}_p(X^D)$. Let $A_{/k}$ be the generic QM surface, i.e., the one corresponding to the generic point of the moduli space. Then it is not hard to see that $\mathrm{End}^0(A) = B$.

Proposition 7. *Suppose $(p, D) = 1$. Then every supersingular QM abelian surface (A, ι) is superspecial.*

Proof: It is known that a g-dimensional abelian variety is superspecial if and only if Oort's invariant $a(A) = \dim_{\mathbb{F}} \mathrm{Hom}(\alpha_p, A)$ is equal to g, and also that any supersingular abelian variety has $a(A) \geq 1$. Thus we need to show that $a(p) \neq 1$ for a supersingular QM surface. But $\mathrm{Hom}(\alpha_p, A)$ is naturally a right $\mathcal{O}_B \otimes \mathbb{F}$-module, i.e., a $M_2(\mathbb{F})$-module. But as we have seen before ("Morita equivalence"), this implies that the underlying $\mathbb{F}$-module has even dimension.

3.1. **A description of the ordinary locus via the Eichler-Selberg trace formula.** Omitted for now. All that we would have liked to say can be found in the papers *Descent on certain Shimura curves*, by Skorobogatov and Yafaev; and *Points on Shimura varieties modulo p* by Milne.

3.2. **A description of the supersingular locus.** We assume that $(p, ND) = 1$. We say that a point on $X_0^D(N)(\mathbb{F})$ is supersingular if its image in $X^D(\mathbb{F})$ is supersingular.

Lemma 8. *There exists a supersingular point $P \in X^D(\mathbb{F})$.*

Proof: We saw this at the beginning of the course when $D = 1$; a very similar argument applies here. Namely, choose an imaginary quadratic field K such that (i) K splits D and (ii) p is nonsplit in K: for instance we could take $K = \mathbb{Q}(\sqrt{-Dp})$. Then there exists at least one point $P \in X^D(K)$ with CM by (any preassigned order in) K. If $D = 1$, we use the fact that CM elliptic curves have potentially good reduction; for $D > 1$, the curve X^D is complete, so there is a reduction map $X^D(\mathbb{Q}_p) = X^D(\mathbb{Z}_p) \to X^D(\mathbb{F}_p)$. The mod p reduction of P is supersingular since p is inert in K.

Clearly then there are supersingular points on $X_0^D(N)(\mathbb{F})$.

Theorem 9. *Fix a supersingular point $P_0 \in X_0^D(N)(\mathbb{F})$. Then there are canonical bijections between the following sets:*
(a) The set of supersingular points on $X_0^D(N)(\mathbb{F})$.
(b) The set of classes of (left or right) ideals for an Eichler order $\mathcal{O}$ of level N in the definite rational quaternion algebra B' of discriminant Dp.
(c) The adelic double coset space $B'^{\times}(\mathbb{Q})\backslash B'^{\times}(A)/\mathcal{O}(A)^{\times}$.
In particular all three sets have cardinality $h(Dp, N)$.

Sktech proof: A key point is that if (A, ι) is a supersingular QM abelian surface, then $\mathrm{End}_{QM}(A)$ is an order in the commutant of B in $M_2(B_{p,\infty})$. It is a general fact that if $A_1 \subset B$ is an inclusion of central simple algebras over a field K, and A_2 is the commutant of A_1 (i.e., the subalgebra of elements commuting with every element of A_1) then $\dim B = \dim A_1 \dim A_2$ and $[B] = [A_1] + [A_2]$, where $[A]$ denotes the class of a central simple algebra in the Brauer group of K. Applying this in the current case, we get that the QM-equivariant endomorphism algebra is B', the definite rational quaternion algebra of discriminant $D' = Dp$. The setup is now identical to the $D = 1$ case: an "enhanced" (i.e., endowed with a QM $\Gamma_0(N)$-level structure Q_N) supersingular QM surface (A, ι, C_N) has QM endomorphism ring a level N Eichler order $\mathcal{O}'$ in B'. Then any other supersingular point is QM-equivariantly isogenous to (A, ι), and the set of QM-equivariant isogenies $(A, \iota, C_N) \to (A', \iota', C'_N)$ is a left $\mathcal{O}'$-module, in fact locally free of rank 1. This defines the map from (a) to (b). To see that it is a bijection, recall that our adelic theory gave a natural equivalence between (b) and (c). What is needed, then, is to interpret the supersingular locus on $X_0^D(N)_{/\mathbb{F}_p}$ via the same double coset construction. One finds in Milne's paper *Points on Shimura varieties modulo p* an adelic description of every isogeny class of QM surfaces mod p, including the supersingular class. As the construction uses the Dieudonné module attached to a p-divisible group, we do not give it here. Note however that the double coset space is exactly what we would have gotten had we tried to associate a Shimura variety to the definite quaternion algebra B': since $SL_1(B')(\mathbb{R})$ is *compact*, we get a complex manifold of dimension 0.

The following result is FYI; it is not used in the proofs of Theorems 1 and 3.

Proposition 10. *a) Every supersingular point on $X_0^D(N)_{/\mathbb{F}_p}$ is rational over $\mathbb{F}_{p^2}$.*
b) There exists an $\mathbb{F}_p$-rational supersingular point on $X_0^D(N)$ iff $e_p(Dp, N) > 0$.

Proof: a) This follows from the identification of the action of Frobenius on each isogeny class in terms of the above adelic construction. See Milne's paper for details.

b) Under the categorical isomorphism between supersingular QM surfaces and ideal clases, the Frobenius map corresponds to the Atkin-Lehner involution w_p. Thus are reduced to information about the traces of Eichler-Brandt matrices, which can be found in Vignéras' book.[5]

For $n \geq 1$, let $\mathbb{Q}_{p^n}$ denote the unique degree n unramified extension of $\mathbb{Q}_p$ and let $\mathbb{Z}_{p^n}$ denote its ring of integers.

Corollary 11. *Assume $(p, DN) = 1$. Then $X_0^D(N)(\mathbb{Q}_{p^r}) \neq \emptyset$ if:*
a) r is even.
b) $e_p(Dp, N) > 0$.
c) r is sufficiently large with respect to D, N, p.

Proof: By Hensel's Lemma, we reduce to the corresponding statements about non-emptiness of the set of $\mathbb{F}_{p^r}$-rational points. By the previous result, there are always $\mathbb{F}_{p^2}$-rational points, namely the supersingular points; *a fortiori* there are $\mathbb{F}_{p^r}$-rational points for every even r. Similarly part b) follows from the condition for there to be an $\mathbb{F}_p$-rational supersingular point. Finally, if $X_{/\mathbb{F}_p}$ is any smooth algebraic curve, then $\#X(\mathbb{F}_{p^r}) \to \infty$ with r (and in particular is positive for sufficiently large r), as follows from the Weil bounds for rational points on algebraic curves.

Exercise 3: Using the fact that an algebraic curve of genus g over $\mathbb{F}_q$ has at least $q + 1 - 2g\sqrt{q}$ $\mathbb{F}_q$-rational points, use Theorem 1 to give an explicit bound for r in part c).

Remark: Note that we have omitted the discussion of how to calculate how many $\mathbb{F}_q$-rational *ordinary* points there are on $X_0^D(N)$. In any given case, this number can readily be calculated by the Eichler-Selberg trace formula; however, the term "formula" here is somewhat misleading. It is not a closed-form expression and it is far from clear (at least to me, and I have never seen anyone else do better) how to extract qualitative information from the formula: e.g., there is no known explicit criterion on D, N and q for there to exist an $\mathbb{F}_q$-rational ordinary point on $X_0^D(N)$.

Problem 1: Fix a positive integer M. Prove (or disprove) that there exist D, N and p such that the least r such that $X_0^D(N)$ has $\mathbb{F}_{p^r}$-rational *ordinary* points is at least M.

Exercise 4:
a) Suppose that $X_0^D(N)$ has genus 0 or 1 (compare Exercise 2). Show that for all p prime to ND, $X_0^D(N)(\mathbb{Q}_p) \neq \emptyset$. (Hint: This is really a statement about smooth curves of genus 0 or 1 over finite fields.)
b) Recall that there is a canonical (and functorial) bijection between smooth genus zero curves over a field K (say, of characteristic not equal to 2) and quaternion algebras over K, (Try to prove this, if you haven't seen it before.) under which $\mathbb{P}^1$ corresponds to $M_2(K)$. Deduce that if K is a number field, a genus zero curve $V_{/K}$ has K-rational points if and only if it has points rational over every completion of K. Moreover, the number of places of K for which V fails to have K-rational

[5]From our perspective, this is cheating: in a later draft, I will explain how to prove this geometrically.

points is even.
c) Consider now X^D with $D = 6,\ 10,\ 22$ (so of genus 0). What can be said about $X^D(\mathbb{Q}_p)$ based upon parts a) and b) and the fact that $X^D(\mathbb{R}) = \emptyset$? Show that there exists exactly one prime $p \mid D$ such that $X^D() \neq \emptyset$.
d) Use part c) to show that $\mathbb{Q}(X^D)$ does not split B when X^D has genus zero.

Remark: By Jordan's Theorem, this implies that the generic point on the Shimura curve X^D is an abelian surface which cannot be defined over its field of moduli, at least when D is 6, 10 or 22. The work of Jordan and Livne shows that there exist infinitely many discriminants for which $X^D(\mathbb{Q}_p) \neq \emptyset$ for some $p \mid D$, hence infinitely many more examples of this. However, this can only happen when D is even. As far as I know, whether or not there exists a discriminant $D > 1$ such that $\mathbb{Q}(X^D)$ splits B (equivalently, whether or not there exists some D such that every QM abelian surface can be defined over its field of moduli) is an open question.[6]

Final remark: We did *not* use the hypothesis that N is squarefree in this section.

4. Special fibers II: Deligne-Rapoport reduction

let k be an algebraically closed field. We say that a curve $X_{/k}$ is **semistable** if it is reduced, connected but possibly reducible, and the only singularities are ordinary double points (aka nodes, aka points of transverse intersection). A bit more informally, all semistable curves can be constructed as follows: we start with a curve C which is the disjoint union of a finite set $C_1, \ldots, C_N$ of smooth curves defined over k. Then we iterate the following construction: choose any $1 \leq i,\ j \leq N$ ($i = j$ is allowed), pick a point $P_i \in C_i(k)$ and a point $Q_j \in C_j(k)$ and glue P_i to Q_j so as to make a transverse intersection (in other words, so that the completed local ring at the intersection point is isomorphic to $k[[x,y]]/(xy)$). We repeat this process any finite number of times, the only proviso being that once we have glued two points together we don't glue either of them to any other point. We will say that a curve is **degenerate semistable** if each $C_i \cong \mathbb{P}^1$.

If now k is an arbitrary (say perfect) field, then a curve $C_{/k}$ is semistable (resp. degenerate semistable) if its basechange to $\overline{k}$ is. We will call a curve k-split if all the singular points are k-rational.

We are now ready to describe the special fiber of $X_0^D(N)_{/\mathbb{F}_p}$, where $p \mid N$.

Theorem 12. *Suppose $p \mid N$. Then $X_0^D(N)_{/\mathbb{Z}_p}$ has the following structure:*
a) The special fiber $X_0^D(N)_{/\mathbb{F}_p}$ has two irreducible components, each isomorphic to the smooth curve $X_0^D(\frac{N}{p})_{/\mathbb{F}_p}$.
b) The two irreducible components intersect transversely at the supersingular points, a corresponding point on the first copy of $X_0^D(\frac{N}{p})_{/\mathbb{F}_p}$ being glued to its image under the quadratic Frobenius map.
c) At each supersingular point $z \in X_0^D(N)_{/\mathbb{F}_{p^2}}$, the complete local ring is isomorphic to $\mathbb{Z}_{p^2}[[X,Y]]/(XY - p^{a_z})$ for some positive integer a_z.

[6]However, this is very much in the spirit of work of Shimura, who showed, e.g., that the generic principally polarized abelian variety of dimension g can be defined over its field of moduli if and only if g is odd. Perhaps the answer to this question appears in one of his papers.

Remark: The hypotheses "N squarefree, $p \mid N$" can be weakened to $p \mid\mid N$.

Proof: The analogous statement for $D = 1$ (i.e., classical modular curves) is due to Deligne and Rapoport. In that same work it was pointed out that their result would continue to hold in the quaternionic context. Apparently the first careful treatment of the quaternionic case is due to Buzzard , who worked however under the assumption of some additional rigidifying level structure. The case of $X_0^D(N)$ (which is not a fine moduli scheme) was worked out (independently) in the theses of the author and of David Helm.

Proposition 13. *For any prime $p \mid N$, $m_p(X_0^D(N)) \leq 4$.*

Proof: Consider $M_0^D(N)_{\mathbb{F}_{p^2}}$. Let z be any point on this curve coming from a supersingular point of $M_0^D(\frac{N}{p})(\mathbb{F}_{p^2})$. The completed local ring at z is isomorphic to $\mathbb{Z}_{p^2}[[x,y]]/(xy - p^a)$ for some integer $a \geq 1$.

Suppose first that $a > 1$. Then in order to get the minimal regular model, one must blow up the point z $(a-1)$ times, getting a chain of $a-1$ rational curves defined over $\mathbb{F}_{p^2}$. Each of these curves has $p^2 + 1 - 2$ smooth $\mathbb{F}_{p^2}$-rational points, which lift to give points on $X_0^D(N)$ rational over $\mathbb{Q}_{p^2}$, the unramified quadratic extension.

For a possibly singular projective curve $C_{/k}$, let $\chi(C) = 1 - g$ be its Euler characteristic, where g is its arithmetic genus.

If $C_{/k}$ is a semistable curve, then there is a simple recipe to compute its Euler characteristic (hence its arithmetic genus) from the Euler characteristics of its irreducible components. Namely, $\chi(C) = \sum_i \chi_(C_i) - \#C_i \cap C_j$, i.e., we add the individual Euler characteristics and then subtract the number of intersection points.

Thus we get:

Corollary 14. *Let N be a positive integer and p prime to DN. Then*

$$\chi(X_0^D(Np)) = 2\chi(X_0^D(N)) - h(Dp, N).$$

Unpacking this, we get:

$$h(Dp, N) = g(D, Np) - 2g(D, N) + 1. \tag{3}$$

In the special case $D = N = 1$, this says that the class number of $B_{p,\infty}$ – i.e., the number of supersingular elliptic curves in characteristic p, is equal to one more than the genus of $X_0(p)$. This was shown in G. Tornaria's lecture by other (but closely related!) methods.

Note that we could now prove Theorems 1 and 3 in the $D = 1$ case; however, we shall press on to give the last piece of the puzzle.

5. Special fibers III: Cerednik-Drinfeld reduction

In this section all notation is as above, and we work with a prime $p \mid D$. The fiber $X_0^D(N)_{/\mathbb{F}_p}$ again has semistable bad reduction, but this time it is totally degenerate and $\mathbb{F}_{p^2}$-split. This is a corollary of the fact that $X_0^D(N)_{/\mathbb{Z}_p}$ is the $\mathbb{Z}_{p^2}/\mathbb{Z}_p$-twist of a **Mumford curve** $C(D,N)_{/\mathbb{Z}_p}$. These are the curves which are uniformized by

cocompact discrete subgroups of automorphisms of the *p-adic* upper halfplane $\mathcal{H}_p$. The theory of p-adically uniformized curves is rich enough so that we will attempt to give it on the fly.[7] Let us record the following properties of any Mumford curve $X_{/\mathbb{Z}_p}$:

(MC1) The special fiber $X_{\mathbb{F}_p}$ is semistable, totally degenerate and $\mathbb{F}_p$-split.

Thus we may attach the **dual graph**, whose vertices are the irreducible components C_i (all isomorphic to $\mathbb{P}^1$) and whose edges are the distinct intersection points of C_i and C_j. [8]

(MC2) The degree of each vertex in the dual graph is at most $p+1$.[9]

We can nicely package all this data by saying that $X_0^D(N)_{/\mathbb{F}}$ is a totally degenerate semistable curve, whose dual graph is endowed with an action of $\mathrm{Gal}(\mathbb{F}_{p^2}/\mathbb{F}_p)$.

Thus, the heart of the matter is to discuss the structure of this graph – what is the vertex set, what is the edge set, and what is the Galois action? – but even before we get into this we can make the following observation.

Proposition 15. *For any $p \mid D$, $X_0^D(N)(\mathbb{Q}_{p^2}) \neq \emptyset$.*

Proof: All that is needed here is that, over $\mathbb{Z}_{p^2}$, $X_0^D(N)$ is equal to the base extension to $\mathbb{Z}_{p^2}$ of a Mumford curve $M(D,N)_{\mathbb{Z}_p}$. Indeed, by Hensel's Lemma, any *smooth* $\mathbb{F}_{p^2}$-valued point on $M(D,N)$ will lift to a $\mathbb{Z}_{p^2}$-valued point. Now look at any component C of $M(D,N)_{/\mathbb{F}_{p^2}}$: it is isomorphic to $\mathbb{P}^1_{\mathbb{F}_{p^2}}$, thus it has p^2+1 $\mathbb{F}_{p^2}$-rational points. Some of these points may be singular; indeed, the singular points correspond to the edges in the dual graph with vertex corresponding to C. But since the degree is at most $p+1$, there are at most $p+1$ singular points, hence at least $p^2+1-(p+1) = p^2-p > 0$ smooth $\mathbb{F}_{p^2}$-rational points. This completes the proof.

Now, for the structure of the dual graph:

Theorem 16. *(Ribet) Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dual graph of the special fiber of $X_0^D(N)$. Let $\mathcal{O}_N$ (resp. $\mathcal{O}_{Np}$) be a level N (resp. level Np) Eichler order in the rational quaternion algebra of discriminant D/p.*
a) There exists a bijection from $\mathcal{G}$ to the disjoint union of two copies of $\mathrm{Pic}_r(\mathcal{O}_N)$.
b) There exists a bijection form $\mathcal{E}$ to $\mathrm{Pic}_r(\mathcal{O}_{Np})$.
c) The $\mathbb{F}_{p^2}/\mathbb{F}_p$-Frobenius element acts on $\mathcal{E}$ by the Atkin-Lehner involution w_p.

See Ribet's 1990 *Inventiones* paper for the proof. Let us at least note that there is a natural map from $\mathrm{Pic}_r(\mathcal{O}_{Np}) \to \{1,2\} \times \mathrm{Pic}_r(\mathcal{O}_N)$: we may choose $\mathcal{O}_{Np} = \mathcal{O}_N \cap w_p \mathcal{O}_N w_p^{-1}$ to be the intersection of two conjugate level N Eichler orders,

[7] I regret the lack of time to discuss this important topic. Were I to teach such a course again, I would omit some other topic – perhaps part of the discussion of arithmetic Fuchsian groups – in order to make room. Sorry!

[8] The dual graph is, roughly speaking, the quotient of the Bruhat-Tits tree – i.e., the homogeneous tree of degree $p+1$ whose vertex set can be viewed as $PGL_2(\mathbb{Q}_p)/PGL_2(\mathbb{Z}_p)$ – by the corresponding uniformizing subgroup Γ.

[9] This follows from the previous footnote.

and these two inclusions define two "pushforward" maps $\mathrm{Pic}_r(\mathcal{O}_{Np}) \to \mathrm{Pic}_r(\mathcal{O}_N)$. Moreover, w_p is obtained by conjugating the $\mathcal{O}_{Np}$-structure by the element w_p, that is, by an integral element of B' of norm p.

Exercise 5: Let $X_{/k}$ be a totally degenerate semistable curve. show that the Euler characteristic $\chi(X)$ in the sense of algebraic geometry coincides with the Euler characteristic of the dual graph of X in the sense of algebraic topology, so in particular with $\#\mathcal{V} - \#\mathcal{E}$.

We conclude:

Corollary 17.

$$\chi(X_0^D(N)) = 2h(\frac{D}{p}, N) - h(\frac{D}{p}, Np).$$

Thus:

$$g(D, N) = h(\frac{D}{p}, Np) - 2h(\frac{D}{p}, N) + 1. \tag{4}$$

For complete results on when $X_0^D(N)$ has points over a p-adic field, where $p \mid D$, see *Local Diophantine properties of Shimura curves* by B. Jordan and R. Livne and *Mauvaise réduction des courbes de Shimura* by A. Ogg.

6. Proof of Theorems 1 and 3

We now have all the information necessary to prove Theorems 1 and 3. Before doing this, note the remarkable similarity between Equations 3 and 4. In fact, we can consolidate them into a single formula, by introducing the formal notation suggested above:

$$g(D', N) = h(D', N) - 1.$$

Thus for any coprime squarefree integers A and N, we have attached a meaning to $g(A, N)$, and we have the following formula:

$$g(Ap, N) = g(A, Np) - 2g(A, N). \tag{5}$$

We also have a (less pretty) formula relating $g(D, Np)$ to $g(D, N)$ using, essentially, the Riemann-Hurwitz formula. More precisely, if we write down Equation (1) for the groups $\Gamma_0^D(Np)$ and Γ^D and use the fact that

$$V(D, Np) = (p+1)V(D, N),$$

then we get

$$2 - 2g(D, Np) = \tag{6}$$

$$(p+1)\left(2 - 2g(D, p) - \frac{e_1(D, p)}{2} - \frac{2e_3(D, p)}{3} - \frac{e_\infty(D, p)}{2}\right) + \frac{e_1(D, Np)}{2} + \frac{2e_3(D, Np)}{3} + \frac{e_\infty(D, Np)}{2}.$$

This completes the proof of Theorems 1 and 3, or at least reduces it to a routine calculation: for instance, it is clear that (5) and (6) have a unique solution with $g(1, 1) = 0$, so we can verify that our claimed formula for $g(A, N)$ satisfies both equations.

Exercise 6: We derived (6) in the case of an indefinite quaternionic discriminant D. Show that it holds for D' a definite discriminant as well.

Let us close with some remarks about possible generalizations of the above material.

I. As indicated above, some of the argument generalizes to the non-squarefree case. Rather, the argument succeeds "at p" for any prime p which divides N exactly. However, if N is not squarefree, our inductive argument does not yield genus / class number formulas in this case, because we cannot lower the level all the way down to $D = N = 1$.

II. One can ask what goes wrong with the geometry in case, say, $p^2 \mid N$. The answer is that in this case the curve $X_0^D(N)$ does *not* have semistable reduction. (Although we have not explained why, this is essentially the same phenomenon as an abelian variety $A_{/\mathbb{Q}}$ having semistable reduction at p if and only if p divides exactly the conductor of A.) Obviously the geometry of non-semistable curves can be considerably more intricate. From the Diophantine perspective, the advantage of semistable reduction is that it is easy to obtain the *minimal regular model* of $X_0^D(N)$ over from its integral canonical model: as we saw above, one needs only to blow up the singularities. In contrast, computing the minimal regular model (or the **stable model**, which is easily obtained in terms of the minimal model and conversely) of $X_0(p^n)$ is a very difficult problem: the case of $X_0(p^2)$ is due to Edixhoven circa 1990, and the case of $X_0(p^3)$ is due to Coleman and McMurdy (as yet unpublished).[10] Similarly, the special fiber of $X_1^D(N)$ modulo a prime dividing N (even $p \mid\mid N$) is nonreduced, again reflecting the fact that $J_1^D(N)$ does not have semistable reduction.

II'. For more information on the bad reduction of the curves $X_0^D(N)$ and $X_1^D(N)$ in the $D = 1$ case, consult Katz and Mazur, *Arithmetic moduli of modular curves*. Note that this only treats the case $D = 1$! However, if one restricts attention to primes dividing N, then it has long been known that one can, essentially, go through the book adding a superscripted D to each instance of $X_\bullet(N)$ (and ignoring all discussions of cusps when $D > 1$).

III. In a very influential 1987 paper, Gross defines a curve $X_0^{D'}(N)$, i.e., meant to be the analogue to $X_0^D(N)$ when D' is the discriminant of a definite rational quaternion algebra. The motives behind this definition our in accord with our motives here: i.e., to give a geometric meaning to the arithmetic of definite quaternion algebras. It must be said that, from a purely geometric standpoint, Gross' definition is not completely satisfying. Namely he defines $X_0^{D'}(N)$ as a curve over $\mathbb{Q}$ which is a *disjoint* union of $h(D', N)$ curves of genus 0 (which are not in general isomorphic to $\mathbb{P}^1$). The arithmetic genus of this curve is then $1 - h(D', N)$, i.e., a nonpositive integer. Thus we find, curiously, that the genera of Gross' curves and our "formal" genera are equal in absolute value but opposite in sign.

IV. Here is another argument to establish the base case of the theorem, i.e., that the genus of $X(1)$ equals zero. Namely, note that $X(1)_{/\mathbb{Q}}$ is a projective curve with everywhere smooth reduction (or, if you like, it has a smooth $\mathbb{Z}$-model). Take now

[10] According to Ken McMurdy's web page, he is now working on $X_0(81)$.

$J(1)_{/\mathbb{Q}}$, the Jacobian of $X(1)$. This is now an abelian variety over $\mathbb{Q}$ with everywhere good reduction (equivalently, with conductor 1). By a famous (and quite deep) theorem of Fontaine, there are no abelian varieties over $\mathbb{Q}$ with everywhere good reduction, except of course for $A = 0$. Since the dimension of the Jacobian is the genus of the curve, we conclude that $X(1)$ genus zero, and then (by the Weil estimates and the Hasse principle for quadrics) that it is isomorphic to $\mathbb{P}^1$.

V. It is natural to ask whether the argument can be extended to the case of quaternion algebras over a totally real field F. The answer is yes, as far as it goes: let $\mathcal{D}$ and $\mathcal{N}$ be coprime squarefree integral ideals of F, such that the number of primes dividing $\mathcal{D}$ has the opposite parity to $[F : \mathbb{Q}]$, so there exists a unique quaternion algebra B/F ramified at all but one infinite place of F, at the primes dividing $\mathcal{D}$, and nowhere else. For simplicity, let us also assume that F has narrow class number one, so that the canonical model of $X = X^{\mathcal{D}_0(\mathcal{N})}$ is defined over F itself, and not merely some extension of F unramified at every finite place. Then there exists a canonical model for X over $\mathfrak{o}_F[\frac{1}{|\delta|}]$, which has smooth reduction at finite places v of F not divididing $\mathcal{D}\mathcal{N}$. Moreover one can define a notion of supersingular points on the special fiber modulo v, which are finite in number and in bijection with $\mathrm{Pic}_r(\mathcal{O}_{\mathcal{N}})$, where $\mathcal{O}_{\mathcal{N}}$ is a level N Eichler order in the totally definite quaternion algebra over F of discriminant $\mathcal{D} \cdot v$. Moreover the Deligne-Rapoport and Cerednik-Drinfeld reduction theories still go through. Thus one gets formulas relating all genera of Shimura curves of type $(1, g-1)$ over F, all class numbers of Eichler orders in totally definite quaternion algebras over F, and all volumes of fundamental domains of Shimura curves over F. However, (especially when g is even) we may not have a "base case," i.e., a Shimura curve in this family which evidently has genus zero. When $[F : \mathbb{Q}]$ is odd, the simplest looking Shimura curve is the one with $\mathcal{D} = \mathfrak{o}_F$, i.e., unramified at every finite prime. Thus this construction gives algebraic curves $X^1_{/F}$ (or, in general, over the Hilbert class field of F) which have smooth reduction at *every* prime of F. These curves will, in general, have positive genus (e.g., I believe it is true that in the (countable) family of Shimura curves attached to totally real fields of given degree g, there are only finitely many of genus less than or equal to any given positive integer), so are *a priori* interesting arithmetic-geometric objects.

Problem: Let $p \equiv -1 \pmod 4$ be a prime number, and $F_p := \mathbb{Q}(\zeta_p + \zeta_p^{-1})$, a totally real number field of odd degree $\frac{p-1}{2}$. Let $X^1_{/F_p}$ be the Shimura curve as described above. Explore the arithmetic of this family of curves.

Exercise 7: Take $F = \mathbb{Q}(\zeta_7)^+ = \mathbb{Q}(\zeta_7 + \zeta_7^{-1})$. Let B/F be the quaternion algebra over F of type $(1, g-1)$ and unramified at every finite prime. Show that the corresponding Fuchsian group Γ^1 is the triangle group $\Delta(2, 3, 7)$, hence this "base" Shimura curve has genus zero. Derive genus formulas and class number formulas for all quaternion algebras over F.

7. Shimura curves violating the Hasse Principle

We will end by applying some of the earlier results on the structure of local points on $X_0^D(N)$ to prove the following result.[11]

Theorem 18. *(Clark) There exists an absolute constant C such that if $D \cdot N > C$, there exists a number field K such that $X_0^D(N)_{/K}$ has points rational over every completion of K but not over K itself. Moreover:*
a) If $N = 1$, K can be taken to be quadratic imaginary.
b) In general, K can be taken to be a quartic field.

Remarks:
a) The constant C can be effectively computed, although I have not done so in the general case. When $N = 1$, then (using some results of V. Rotger) one can take any $D > 546$.
b) The theorem also holds for $X_1^D(N)$, although in this case it can be shown that the degree of the number field required goes to infinity with N (uniformly in D).

Definition: For an algebraic variety V defined over a field K, let $m(V)$ be the least degree of a field extension L/K such that $V(L) \neq \emptyset$. If K is a number field, we define $m_{\text{loc}}(V)$ to be the least common multiple of $m(V_{/K_v})$ as v ranges over all places of K. (It is easy to see that V has K_v-rational points for almost every v.)

Exercise 8: Show that if $m = m_{\text{loc}}$, then there exist infinitely many number fields L/K with $[L : K] = m$ and such that V has points over every completion of L. (Hint: use Krasner's Lemma and the standard (weak) approximation theorem for a finite set of valuations on a field.)

The considerations of the previous sections give us the following result:

Proposition 19. *We have $m_{\text{loc}}(X_0^D(N)_{/\mathbb{Q}})$ is 2 or 4. When $D = 1$, it is always 2.*

Proof: Since $X_0^D(N)_{/\mathbb{R}} = \emptyset$, the m-invariant at the real place is 2, so m_{loc} is even. For p not dividing DN, we have $m_p \leq 2$ by Corollary 11. For $p \mid N$, we have $m_p \mid 4$ by Proposition 13. For $p \mid D$, we have $m_p \leq 2$ by Proposition 15.

Remark: I believe that there should in fact be examples where $m_{\text{loc}}(X_0^D(N)) = 4$, but I have not done the computations.

Definition: For an algebraic curve X defined over a field K, we define the K-gonality to be the least degree of a K-morphism $\varphi : X \to \mathbb{P}^1$. For instance, only $\mathbb{P}^1$ has K-gonality 1, and the curves with K-gonality 2 are hyperelliptic over K.

Proposition 20. *(Ogg) The $\mathbb{Q}$-gonality of $X_0^D(N)$ approaches ∞ with $\min(D, N)$.*

Sketch Proof: We may suppose that the genus of $X_0^D(N)$ is positive (because from our genus formulas there are only finitely many curves of genus zero). Then it is known that if $\varphi : X \to C$ is a finite morphism of curves defined over $\mathbb{Q}_p$, and X

[11]This section is an exposition of some of the results in my paper *On the Hasse principle for Shimura curves*, which can be found online at www.arxiv.org. Our presentation here is shorter and less formal; on the other hand, the reader might prefer to see more of the details spelled out.

has smooth reduction, then so does C and the morphism φ can itself be reduced to give a map of smooth curves over $\mathbb{F}_p$. Suppose that this morphism has degree M. Then for all $r \geq 1$, $\#X(\mathbb{F}_{p^r}) \leq M(p^r + 1)$. Let us apply this with $X = X_0^D(N)$, any prime p with $(p, DN) = 1$ and $r = 2$, we get that if this curve has $\mathbb{Q}$-gonality at most M, then $\#X_0^D(N)(\mathbb{F}_{p^2}) \leq M(p^2 + 1)$. On the other hand, because all the supersingular points are defined over $\mathbb{F}_{p^2}$, we know that the number of $\mathbb{F}_{p^2}$-rational points is at least $h(Dp, N)$. We leave it to the reader to use the class number formulas to check that for any fixed M, the set of all pairs (D, N) such that for all primes p not dividing DN, $h(Dp, N) \leq M(p^2 + 1)$ is finite.

Remark: A different argument appears in my paper. Namely, Abramovich has shown that if $\Gamma \subset \Gamma^D(1)$ is any congruence subgroup, then the $\mathbb{C}$-gonality of the Shimura curve $X^D(\Gamma)$ is at least $\frac{21}{200}(g(X^D(\Gamma)) - 1)$. This is a stronger result for three reasons: (i) the $\mathbb{C}$-gonality is evidently less than or equal to the $\mathbb{Q}$-gonality (because any map to $\mathbb{P}^1$ which is defined over $\mathbb{Q}$ can be defined over $\mathbb{C}$, but not conversely); (ii) the bound it gives is better than what can be obtained using Ogg's method (Ogg's method would give a bound of order square root of the genus, rather than linear in the genus), and (iii) it applies also to curves like $X_1^D(N)$, for which I prove a similar Hasse principle violation result. On the other hand, Ogg's method is beautifully thematic and quite influential: it shows again how the supersingular points influence the geometry of the curves, and has been used in many different contexts.[12]

Why do we care about the $\mathbb{Q}$-gonality of $X_0^D(N)$? Well, if $\varphi_{\mathbb{Q}} : X \to \mathbb{P}^1$ is a degree M map, then clearly X has infinitely many points of degree at most M. The following result is a sort of converse.

Theorem 21. *Let $X_{/K}$ be a curve defined over a number field, and for $n \in \mathbb{Z}^+$, let $\mathcal{S}_n(X)$ be the set of points $P \in X(\overline{K})$ of degree dividing n. If $\mathcal{S}_n(X)$ is infinite, then the K-gonality of X is at most $2n$.*

We will not discuss this result here, except to say that it is by far the deepest result that goes into the proof. This theorem appears for the first time in my paper. However, it is a very small variation of a theorem of Frey, which proves the same reuslt for points of degree *at most* n rather than dividing n under the additional assumption that $X(K) \neq \emptyset$ (which we certainly do not want to make here). Frey's argument is clever but still quite elementary. It uses, however, Faltings' proof of the Mordell-Lang Conjecture, which is probably the single deepest general result concerning rational points algebraic varieties.

Proof of Theorem 18: We need only put together the pieces: for $\min(D, N)$ sufficiently large, the $\mathbb{Q}$-gonality of $X_0^D(N)$ is greater than 8, so that by the preceding theorem, there are only finitely many points of $X_0^D(N)(\overline{\mathbb{Q}})$ which are rational over fields of degree 1, 2 or 4. Thus, letting M be the field which is generated by the coordinates of all the points rational over fields of degree dividing 4, we know that M is a number field. On the other hand, there are no $\mathbb{Q}$-rational points at all (since there are no $\mathbb{R}$-rational points), and it follows that if $L/\mathbb{Q}$ is a field extension of

[12]Also Abramovich's argument uses some deep results of *differential* geometry that I do not pretend to understand.

degree 4 which is linearly disjoint from M, then $X_0^D(L) = \emptyset$. But it is easy to produce a number field L of degree 4 with the following local behavior: (i) it is totally imaginary; (ii) at every finite place p such that $X_0^D(N)(\mathbb{Q}_p) = \emptyset$, the completion M_p is a quartic extension of $\mathbb{Q}_p$ such that $X_0^D(N)(L_p) \neq \emptyset$, and (iii) L is totally ramified at some prime p which is unramified in in M, so that M and L are linearly disjoint over $\mathbb{Q}$. We conclude that $X_0^D(N)_{/L}$ violates the Hasse principle. In fact infinitely many such L exist of degree 4, and L can be taken to be of degree 2 when $N = 1$.

Final remark: Obviously Theorem 18 is false if we do not exclude the pairs (D, N) for which $X_0^D(N)$ has genus zero. If instead of Theorem 20 we use a theorem of Harris and Silverman, we get that X^D violates the Hasse principle over some quadratic field if it is neither hyperelliptic nor bielliptic (i.e., admits a degree 2 morphism to an elliptic curve). It is an exercise to apply Ogg's method to show that this condition excludes only finitely many discriminants. More precise work of Rotger shows that the condition holds for all $D > 546$ (whereas $D = 546$ is bielliptic). On the other hand, the theorem may well be true for all D's for which $X_0^D(N)$ has positive genus. In the same paper, I made the following (rather ambitious)

Conjecture 22. *Let K be a number field, and $C_{/K}$ a (smooth, projective, geometrically irreducible) algebraic curve without K-rational points, and of positive genus. Then there exists a finite field extension L/K such that $C_{/L}$ violates the Hasse principle.*

In other words, we claim that Theorem 18 is not a special property of Shimura curves; rather, it is just that we know enough about the arithmetic of Shimura curves to *prove* the conjecture for almost all of them. In the same paper I mention that the conjecture can in fact be established for curves with everywhere semistable reduction satifying some modest technical assumptions. The general case seems quite forbidding – let me know if you have any ideas!