On the computation of coefficients of a modular form

Bas Edixhoven[*], in collaboration with
Jean-Marc Couveignes[†], Robin de Jong[‡], Franz Merkl[§] and Johan Bosman[¶]

May 11, 2006

Contents

[*]Mathematisch Instituut, Universiteit Leiden
[†]Groupe de Recherche en Informatique et Mathématiques du Mirail, Université de Toulouse II, Le Mirail
[‡]Mathematisch Instituut, Universiteit Leiden
[§]Mathematisches Institut, Ludwig-Maximilians-Universität München
[¶]Mathematisch Instituut, Universiteit Leiden

1 Introduction

This is a progress report, the aim of which is to describe in more detail and to place in a wider context than is possible in a research article the project that is presently being carried out by Bas Edixhoven, Jean-Marc Couveignes, Robin de Jong and Johan Bosman. The project concerns the complexity issues of the computation of coefficients of modular forms, and, more importantly, of the Galois representations with coefficients in finite fields that are associated to eigenforms. In the Netherlands, this project is supported financially by NWO, the Dutch organisation for scientific research (Nederlandse organisatie voor Wetenschappelijk Onderzoek); see Edixhoven's home page for the grant proposal for this project.

The main results are the following (for more details see Theorems 23.1 and 22.1):

> *There exists a probabilistic algorithm that on input a prime number p gives $\tau(p)$, in expected running time polynomial in $\log p$.*

and:

> *There exists a probabilistic algorithm that computes the mod l Galois representation associated to Δ in time polynomial in l.*

A concise research article on this subject will shortly be extracted from these notes by Couveignes, Edixhoven and de Jong, and will be submitted as soon as possible. The article is aimed at experts in arithmetic geometry. This report, written in the context of a contract[1] between the University of Leiden and the French CELAR (Centre Électronique de l'Armement), is intended to be accessible to a wider audience, including people with an interest in cryptology and with some background in computational algebraic geometry and number theory. I hope that this text will serve as a kind of overview where it concerns material that can be found in the literature, providing sufficiently many references to it, as well as a detailed proof where new material is concerned. As the contract stipulates, this report will be published on Edixhoven's home page, accessible to everyone. Comments are welcome, and, if appropriate and if possible, will be taken into account.

1.1 Personal historical account

This project started when René Schoof asked me (Bas Edixhoven), in the Fall of 1995, at the end of a short workshop in Münster on Néron models, the question if the value of Ramanujan's τ-function at a prime number p could be computed in time polynomial in $\log p$, in analogy of

[1] Contrat d'Études 04.42.217

his algorithm that computes in polynomial time the number of points on the reduction mod p of an elliptic curve over $\mathbb{Q}$. I found this question immediately very attractive. It was clear at once that one could try to first compute the mod l Galois representation V_l associated to the modular form Δ, and from that compute $\tau(p)$ mod l as the trace of the Frobenius element at p. But then the computation of V_l should be done in time polynomial in l.

The problem that one runs into is that Deligne's construction of V_l uses higher étale cohomology, of which we do not know how to treat it algorithmically (to be precise, it is cohomology in degree 11 with constant coefficients of a variety of dimension 11). A standard step to get rid of the higher degree étale cohomology was then to consider, still following Deligne, the dual of V_l in the first degree cohomology of a non-constant sheaf on the j-line (the 10th symmetric power of the local system given by the l-torsion of the universal elliptic curve). I invested quite some time in studying this first degree cohomology group. Its elements are torsors over the base curve, and therefore are given by certain covers of the base curve. I showed that these torsors can be described efficiently by algebraic equations, and that a system of equations for the coefficients of those equations can be computed efficiently. However, it seems that there is no method known for solving systems of polynomial equations efficiently, and that even there is no real hope that such a method will ever be found (for example, the satisfiability problem for Boolean equations, known as SAT, is known to be NP-complete). Therefore, working directly with the first degree cohomology of the non-constant sheaf was judged too difficult.

A standard technique in étale cohomology is to make locally constant sheaves constant by passing to a cover. In our case, this boils down to the fact that V_l occurs in the l-torsion of the Jacobian $J_1(l)$ of the modular curve $X_1(l)$. Using this, one is back in the familiar situation of torsion points on Abelian varieties. The main problem here, however, is that the dimension of $J_1(l)$ is not bounded, so that one cannot apply the results of Jonathan Pila in [88]. In fact, the dimension of $J_1(l)$ grows quadratically in l. This means that our 2-dimensional $\mathbb{F}_l$-vector space V_l is embedded in the very large space consisting of the l-torsion of $J_1(l)$. It is now easy to write down equations for V_l, but standard methods from computer algebra for solving such equations usually take an amount of time that is exponential in the dimension, hence in l^2.

In February of 1999, I discussed this problem with Jean-Marc Couveignes when he visited Rennes as a speaker in the algebraic geometry seminar. He suggested another method. He said that as the goal was to compute a number field (namely, the field of definition of the elements of V_l), one should try to construct a generator of that field, and approximate it, numerically, with a precision so high that the exact minimal polynomial of the generator is determined by the approximation. The precision that is required is determined by the height of the generator. A way to construct a generator would be to take a function on $J_1(l)$, defined over $\mathbb{Q}$, and evaluate it at the points of V_l. However, it was not clear what function one could take that would give

values of small enough height, and that could be approximated easily enough. From February 1999 on Couveignes and I collaborated on this project, and we corresponded regularly but not often, without making much progress.

Finally, in October 2000 I had an idea that allows to carry out Couveignes approach. The idea was simply to consider a suitable function on the curve $X_1(l)$ instead of on its Jacobian $J_1(l)$, and use the description of points on $J_1(l)$ in terms of divisors on $X_1(l)$; the function can then be used to push divisors to the projective line, and get equations for their image. There are then two problems to be dealt with. The first problem is to give a construction of a generator of the field to be computed and show that the height of that generator is polynomial in l when l varies. It was decided that I would study this problem, using Arakelov theory. The second problem is to show that the necessary approximations can be done in time polynomial in l. This seems reasonable, using the complex uniformisation of $X_1(l)$ and of $J_1(l)$, however, to really *prove* that it works is a very different matter. Couveignes would try to solve the approximation problem. At first sight, it is perhaps not so easy to see why working on the curve $X_1(l)$ is so much different from working on its Jacobian $J_1(l)$. As this is an important point, easy to appreciate and not so hard to understand, we invite the reader to muse a bit on this and try to find an answer before reading the one in the footnote[2] below.

In December 2000, I gave, for the first time, a lecture on this subject, during the workshop "Computational Arithmetic Geometry" held at the MSRI in Berkeley; see [35]. By that time I had already worked out the strategy in sufficient detail in order to give a lecture. In particular, I had already shown that in the function field analog, under the assumption of good reduction everywhere, a polynomial upper bound for the required precision could be obtained using intersection theory and the theorem of Grothendieck-Riemann-Roch. The next step was then to try to make the same arguments work in the number field case, using Arakelov intersection theory and Faltings's arithmetic Grothendieck-Riemann-Roch theorem. On the one hand, it was clear that there was a very nice and promising project to be carried out. But on the other hand it was also clear that there was really a lot of work to be done.

Curiously enough, I did not work on this project for three years. Instead, I worked on other things: some cases of the André-Oort conjecture, the typesetting of SGA 1, computational aspects of modular forms of weight one, to name the most important ones. Also, Compositio Mathematica started taking a lot of my time (and it still does!) after I started working as co-managing editor for it in January 2003. In the Summer of 2002 I moved from Rennes to Leiden,

[2] In the complex analytic setting, a rational function on $J_1(l)$ is usually locally given as a power series in $g_l = \mathrm{genus}(X_1(l))$ variables, and therefore hard to approximate because the number of terms needed is too large. On the other hand, the map from $X_1(l)^{g_l}$ to $J_1(l)$ given by a divisor of degree g_l is given by integration of g_l power series in one variable, and therefore easy to approximate.

and wrote a very detailed research proposal for NWO (the Dutch organisation for Scientific Research) for the project of polynomial time computation of $\tau(p)$ (almost identical to the proposal that is on my home page). Unfortunately, this proposal was not funded at that moment (January 2003). Nevertheless, I started working on the project again in March 2003, a good reason being that I was going to give a lecture on it during the AIM workshop "Future directions in algorithmic number theory" in Palo Alto.

In April 2003 I gave a lecture in the Dutch Intercity Number Theory seminar, where Robin de Jong was among the audience. He was then a PhD student in Amsterdam, working on Arakelov theory, under the direction of Gerard van der Geer. In the three months that followed, we succeeded together in applying Arakelov theory and the arithmetic Riemann-Roch theorem to get precise expression for an upper bound on the required precision. To be precise, we proved what is now Theorem 15.2.5. With this result proved, it was clear what quantities in the Arakelov theory of modular curves $X_1(l)$ needed to be bounded polynomially in l: the Faltings height of $X_1(l)$, the log of the sup-norm of a certain theta function, and, finally, the sup of the Arakelov-Green functions associated to single points on $X_1(l)$. Robin and I collaborated regularly on the Arakelov aspect of the project since then. Most of Sections 14–19 have been written by us together, most often first in scratch by me, and then in a more acceptable form by Robin, and then in final form again by me. (Of course, this is not true for Section 18.1 which was provided by Franz Merkl.)

Sinnou David, who visited Leiden for a workshop during the Summer of 2003, provided a suitable upper bound for the sup-norm of the theta function. His arguments, based on results by him and Philippon in [17] and [18] are not the ones that are now in this report, as we (Robin and I) found more direct arguments for it, using ingredients that we use also for bounding the other two quantities. The result is Theorem 17.1.

Getting a bound for the height of $X_1(l)$ was not so much of a problem, as it would suffice to generalise the method used in [10] in order to estimate the height of $X_0(l)$. This part of the work just had to wait until I took the time for it. The result is Theorem 16.7.

The last main hurdle to overcome was to get suitable upper bounds for the Arakelov-Green functions. As I am not knowledgeable enough in this matter, I started shopping around for help. At the end of October 2003 I asked Peter Sarnak what he knew about it. He replied that he would transmit the question to Jay Jorgenson. In the academic year 2003-2004, Franz Merkl had a position in Leiden. His office was close to mine, and I saw him often at work in the evenings or during weekends. Even though he worked in the statistics group, I had seen his thesis and knew he was a person I could ask about the so much desired upper bounds. This happened in December 2003, and turned out to work marvellously. Franz said that bounding Green functions on Riemann surfaces was something that he had already done for his thesis but that had not

gotten in it. By the beginning of February of 2004 he gave me what is now (after some editing) Section 18.1. Meanwhile, after Peter Sarnak had done what he had promised to do, Jay Jorgenson and Jürg Kramer started working on the same question at the beginning of January 2004. I told them that Franz had already solved the problem, but also that it would be preferable to have two proofs, and that, due to their approach, their result would probably be more precise, hence could give a better exponent for the final complexity of the algorithm for computing $\tau(p)$. At the end of March 2004 they had a more or less final version of their result, which is now the article [59].

In May 2004 I had a 3 day meeting with Robin de Jong and Jean-Marc Couveignes in Rennes, where we discussed the non-trivial problem of how to fit all the pieces of our project together. There it was decided to work not with the modular curves $X_1(l)$ but with the $X_1(5l)$ instead, as I had found (in September 2003) that on these last curves it was not hard to find cuspidal divisors as in Theorem 12.8.

In the meantime, Jean-Marc had advanced on the approximation problem. In July 2003 he had suggested the possibility of doing the computations over a p-adic field with p small. He had written a first version of his preprint [13] on computations over the complex numbers already in January of 2004. In that preprint he shows (Theorems 1 and 2) that addition and subtraction in the complex Jacobians $J_0(l)(\mathbb{C})$ can be done deterministically in time polynomial in l and the required precision, and ditto for the inversion of the Jacobi map, i.e., for representing points of $J_0(l)(\mathbb{C})$ by divisors on $X_0(l)(\mathbb{C})$. The notion of precision here is the required number of correct digits, i.e., a bound for minus the logarithm of the error in the approximation. His methods extend almost certainly to the case of the Jacobians $J_1(n)$ for all $n \geq 1$. But as that work has not yet been done, we will use, in this text, another approach to the approximation part.

A nice byproduct of our estimates from Arakelov theory is an upper bound for the number of prime numbers where the geometry of our methods (uniqueness of divisors on X_l, poles of certain functions on X_l) does not specialise well; for the technical statement, see Theorem 19.5. As a consequence, it is then possible to do all necessary computations over finite fields. At our meeting in Rennes in May 2004, it was decided that Jean-Marc would work out this finite field approach. His results are described, very briefly, in Section 20, and the details are in Section 21, written by himself. In fact, already in August 2003 Jean-Marc had suggested to use a mod p or p-adic method, for p a suitable small prime, but at that time we did not yet have an upper bound for the number of p that cannot be used.

In 2004, I resubmitted my research proposal to NWO, and this time I was more lucky: it got funded, starting January 2005. On January 1, 2005, Robin was appointed as post-doc in Leiden on those funds (he had defended his thesis in Amsterdam in December 2004). We then started writing our work up more properly, filling the gaps that were still there. Also Johan Bosman, who started working as a PhD-student with me in June 2004, was from then on paid by the

project. Johan's task was to try to really compute the representations V_l for as many l as possible. In March 2005, he found the polynomials of degrees 14 and 168 whose splitting fields give the mod 13 representations (to $\mathrm{PGL}_2(\mathbb{F}_{13})$ and $\mathrm{GL}_2(\mathbb{F}_{13})$, respectively). Since then he has also found such polynomials for $l = 17$ and $l = 19$, but the amount of time needed for his computations grows fast with l. His results are described in Section 24, written by himself.

In Summer of 2005, at the Oberwolfach conference "Explicit methods in number theory" our results were presented in a series of 5 lectures by me, Robin, Jean-Marc and Johan. After that conference, I started writing the more technical sections of this report. Now, at the beginning of May 2006, this report is ready to be published on arxiv.

A few words about the future. In September 2005, Fabio Mainardi was appointed in the project, with the goal of generalising the whole approach to other situations, such as Hilbert modular forms. In the Summer of 2006, two PhD students will start working in the project in Leiden. One will try to get the complexity of our algorithms as small as possible; actually, the reader will note that although we prove the existence of polynomial time algorithms in this report, we do not give any exponent in the final results. The other will work on the function field analog of the whole situation. In the function field case, more precise results can be expected, because ordinary intersection theory is easier to deal with than Arakelov intersection theory.

1.2 Short overview of this report

Mostly, the titles of the sections speak for themselves. The final results are in sections 22 and 23, all the rest is introduction (sections 2–10 and Section 14), or preparation (sections 11–13), or proofs of intermediate results (sections 15–21), or examples (Section 24).

It is fair to say that the technical level of the sections varies quite a lot. The introductory ones are aimed at non-specialists, but the ones where the real work is done are certainly more difficult. Nevertheless, I hope that after reading the introductory ones, the reader has an idea of what goes on in the more difficult sections.

I did not write all in this report. Jean-Marc Couveignes wrote Section 21. Johan Bosman wrote Section 24. Franz Merkl provided Section 18.1. The rest of sections 14–19 were written in collaboration with Robin de Jong, as explained in the personal historical account above.

1.3 Acknowledgements

During this project that started with Schoof's question in the Fall of 1995, I have been in contact with many people. This is an appropriate place to thank at least some of them. First of all, I must name those with who I have worked together most on this project: Jean-Marc Couveignes and

Robin de Jong. Very important to me was the help I got from Franz Merkl, and Jay Jorgenson and Jürg Kramer, concerning upper bounds for Green functions, and the help of Sinnou David concerning bounds on theta functions. I spoke several times with Bjorn Poonen. He told me about SAT, and at some point after a lecture at Oberwolfach (I do not remember which year) he told me that now he was convinced of my project, which was nice to hear. When I was still in Rennes, I discussed matters concerning Arakelov theory with Laurent Moret-Bailly. In Leiden, the computational number theory environment (Hendrik Lenstra, Bart de Smit, Peter Stevenhagen) has certainly been beneficial. I am very happy with Johan Bosman's examples. His work has had influence at several points of this report: in Section 22.6, and in Section 17 to name two of them. Peter Bruin's master thesis (in preparation) has helped me to understand the details of Franz Merkl's work on Green functions.

Thanks are due to NWO for funding my project from 2005 on, for 5 years. With the funds of NWO, I could appoint Robin de Jong and Johan Bosman, and the two PhD students who are about to start. And there are some post-doc years left to fill.

I am very grateful to the CELAR (David Lubicz and Reynald Lercier, in particular), for proposing me the contract under which this report has been written, at a time when I had no money for my own research in Leiden. One might think that institutions such as the CELAR are all about secrecy, but, in fact, this contract is a good example of open access to information. Moreover, this information would have been available only later, or not at all in this form, without that contract.

The Mathematical Institute in Leiden has been a good place for me to work on the project, although I have many duties there. The Mathematics Department in Rennes has played an important role for me, as the project started there, and after my departure to Leiden, I could still meet there regularly with Jean-Marc without having to ask for travel money at Leiden. I should acknowledge the Research Training Network of the European Union "Arithmetic Algebraic Geometry", of which I am a member attached to the node in Rennes, for their facilities.

2 Historical context: Schoof's algorithm

One way to understand the project described here is to view it in relation to Schoof's method to count points on elliptic curves over finite fields, see [96] and [97]. Jan Nekovar's DEA lecture notes [86] constitute a good reference for the theory of elliptic curves. René Schoof gave an algorithm to compute, for E an elliptic curve over a finite field $\mathbb{F}_q$, the number $\#E(\mathbb{F}_q)$ of $\mathbb{F}_q$-rational points in a time $O((\log q)^{5+\varepsilon})$. His algorithm works as follows.

The elliptic curve is embedded, as usual, in the projective plane $\mathbb{P}^2_{\mathbb{F}_q}$ as the zero locus of a

Weierstrass equation, which, in inhomogeneous coordinates, is of the form:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

with the a_i in $\mathbb{F}_q$. We let $\mathbb{F}_q \to \overline{\mathbb{F}}_q$ be an algebraic closure. We let $\mathrm{F}_q\colon E \to E$ denote the so-called q-Frobenius. It is the endomorphism of E with the property that for all (a, b) in the affine part of $E(\overline{\mathbb{F}}_q)$ given by the Weierstrass equation above we have $\mathrm{F}_q((a, b)) = (a^q, b^q)$. The theory of elliptic curves over finite fields, see for example Algebraic Theory, Chapter III of [86], says:

1. there is a unique integer a, called the *trace* of F_q, such that $\mathrm{F}_q^2 - a\mathrm{F}_q + q = 0$ (an identity in the endomorphism ring of E);
2. $\#E(\mathbb{F}_q) = 1 - a + q$;
3. $|a| \leq 2q^{1/2}$.

So, computing $\#E(\mathbb{F}_q)$ is equivalent to computing this integer a. Schoof's idea is now to compute a modulo l for small prime numbers l. If the product of the prime numbers l exceeds $4q^{1/2}$, the length of the interval in which we know a to lie, then the congruences modulo these l determine a uniquely. Analytic number theory tells us that it will be sufficient to take all primes l up to approximately $(\log q)/2$.

Then the question is how one computes a modulo l. This should be done in time polynomial in $\log q$ and l. The idea is to use the elements of order dividing l in $E(\overline{\mathbb{F}}_q)$. We assume now that l does not divide q, i.e., we avoid the characteristic of $\mathbb{F}_q$. For each l, the kernel $E(\overline{\mathbb{F}}_q)[l]$ of multiplication by l on $E(\overline{\mathbb{F}}_q)$ is a two-dimensional vector space over $\mathbb{F}_l$. The map F_q gives an endomorphism of $E(\overline{\mathbb{F}}_q)[l]$, and it follows that the image of a in $\mathbb{F}_l$ is the unique element of $\mathbb{F}_l$, also denoted a, such that for each v in $E(\overline{\mathbb{F}}_q)[l]$ we have $a\mathrm{F}_q(v) = \mathrm{F}_q^2 v + qv$. We remark that the image of a in $\mathbb{F}_l$ is the trace of the endomorphism of $E(\overline{\mathbb{F}}_q)[l]$ given by F_q, but this is not really used at this point.

To find this element a of $\mathbb{F}_l$, one proceeds as follows. We suppose that $l \neq 2$. There is a unique monic element ψ_l of $\mathbb{F}_q[x]$ of degree $(l^2 - 1)/2$, whose roots in $\overline{\mathbb{F}}_q$ are precisely the x-coordinates of the $l^2 - 1$ non-zero elements in $E(\overline{\mathbb{F}}_q)[l]$ (the rational function x on E is a degree two map to $\mathbb{P}^1_{\mathbb{F}_q}$, which as such is the quotient for the multiplication by -1 map on E). One then lets A_l be the $\mathbb{F}_q$-algebra obtained as:

$$A_l := \mathbb{F}_q[x, y]/(y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6, \psi_l(x)).$$

The dimension of A_l as $\mathbb{F}_q$-vector space is $l^2 - 1$. An equivalent description of A_l is to say that it is the affine coordinate ring of the subscheme of points of order l of E. By construction

of A_l, there is a tautological A_l-valued point v in $E(A_l)$ (its coordinates are the images of x and y in A_l). Now to find the element a of $\mathbb{F}_l$ that we are looking for one then tries one by one the elements i in $0, \pm 1, \ldots, \pm(l-1)/2$ until $i\mathrm{F}_q(v) = \mathrm{F}_q^2 v + qv$; then $i = a \mod l$.

It is easy to see that all required computations can be done in time $O((\log q)^{5+\varepsilon})$ (using fast arithmetic for the elementary operations, e.g., a multiplication in A_l costs about $(l^2(\log q))^{1+\varepsilon}$ time; $l^2(\log q)$ is the number of bits needed to store one element of A_l).

For the sake of completeness, let us mention that shortly after the appearance of Schoof's algorithm, Atkin and Elkies have added some improvements to it, making it possible in certain cases to reduce the dimension of the $\mathbb{F}_q$-algebra from l^2-1 to linear in $l+1$ or $l-1$. This improvement, called the Schoof-Atkin-Elkies (SEA) algorithm, is important mainly for implementations. Its (average) complexity is $O((\log q)^{4+\varepsilon})$; for details, the reader is referred to [97].

3 Schoof's algorithm described in terms of étale cohomology

In order to describe Schoof's algorithm in the previous section, we referred to the theory of elliptic curves over finite fields. But there is a more general framework for getting information on the number of rational points of algebraic varieties over finite fields: cohomology, and Lefschetz's trace formula. Cohomology exists in many versions. The version directly related to Schoof's algorithm is étale cohomology with coefficients in $\mathbb{F}_l$. Standard references for étale cohomology are [51], [25], [58], [81], [44]. The reader is referred to these references for the notions that we will use below. We also recommend Appendix C of [53].

For the sake of precision, let us say that we define the notion of *algebraic variety over a field* k to mean *k-scheme that is separated and of finite type*. Associated to an algebraic variety X over a field k there are *étale cohomology groups with compact supports* $\mathrm{H}^i_{\mathrm{c}}(X_{\mathrm{et}}, \mathbb{F}_l)$, for all $i \geq 0$ and for all prime numbers l. Actually, the coefficients $\mathbb{F}_l$ can be replaced by more general objects, sheaves of Abelian groups on the étale site X_{et} of X, but we do not need this right now. If X is a proper k-scheme, then the $\mathrm{H}^i_{\mathrm{c}}(X_{\mathrm{et}}, \mathbb{F}_l)$ are equal to the étale cohomology groups $\mathrm{H}^i(X_{\mathrm{et}}, \mathbb{F}_l)$ without condition on supports.

If k is separably closed then the $\mathrm{H}^i_{\mathrm{c}}(X_{\mathrm{et}}, \mathbb{F}_l)$ are finite dimensional $\mathbb{F}_l$-vector spaces, zero for $i > 2\dim(X)$. In that case, they are the analog of the more easily defined cohomology groups $\mathrm{H}^i_c(X, \mathcal{F})$ for complex analytic varieties: the derived functors of the functor that associates to a sheaf $\mathcal{F}$ of $\mathbb{Z}$-modules on X equipped with its Archimedean topology its $\mathbb{Z}$-module of global sections whose support is compact.

The construction of the $\mathrm{H}^i_{\mathrm{c}}(X_{\mathrm{et}}, \mathbb{F}_l)$ is functorial for proper morphisms: a proper morphism $f\colon X \to Y$ of algebraic varieties over k induces a pullback morphism f^* from $\mathrm{H}^i_{\mathrm{c}}(Y_{\mathrm{et}}, \mathbb{F}_l)$ to

$\mathrm{H}^i_{\mathrm{c}}(X_{\mathrm{et}}, \mathbb{F}_l)$.

Let now X be an algebraic variety over $\mathbb{F}_q$. Then we have the q-Frobenius morphism F_q from X to itself, and, by extending the base field from $\mathbb{F}_q$ to $\overline{\mathbb{F}}_q$, from $X_{\overline{\mathbb{F}}_q}$ to itself. This morphism F_q is proper, hence induces maps:

$$\mathrm{F}_q^*\colon \mathrm{H}^i_{\mathrm{c}}(X_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l) \longrightarrow \mathrm{H}^i_{\mathrm{c}}(X_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l).$$

Hence, for each i in $\mathbb{Z}$, the trace $\mathrm{trace}(\mathrm{F}_q^*, \mathrm{H}^i_{\mathrm{c}}(X_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l))$ of the map above is defined, and it is zero for $i < 0$ and $i > 2\dim(X)$. The set of fixed points of F_q on $X(\overline{\mathbb{F}}_q)$ is precisely the subset $X(\mathbb{F}_q)$. The Lefschetz trace formula then gives the following identity in $\mathbb{F}_l$:

$$\#X(\mathbb{F}_q) = \sum_i (-1)^i \mathrm{trace}(\mathrm{F}_q^*, \mathrm{H}^i_{\mathrm{c}}(X_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)). \tag{3.1}$$

We can now say how Schoof's algorithm is related to étale cohomology. We consider again an elliptic curve E over a finite field $\mathbb{F}_q$. We assume that l does not divide q. Then, as for any smooth proper geometrically connected curve, $\mathrm{H}^0(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l) = \mathbb{F}_l$ and F_q^* acts on it as the identity, and $\mathrm{H}^2(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)$ is one-dimensional and F_q^* acts on it by multiplication by q, the degree of F_q. According to the trace formula (3.1), we have:

$$\#E(\mathbb{F}_q) = 1 - \mathrm{trace}(\mathrm{F}_q^*, \mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)) + q.$$

It follows that for the integer a of the previous section, the trace of Frobenius, we have, for all l not dividing p the identity in $\mathbb{F}_l$:

$$a = \mathrm{trace}(\mathrm{F}_q^*, \mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)).$$

This identity is explained by the fact that there is a natural isomorphism, compatible with the action of F_q:

$$\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l) = E(\overline{\mathbb{F}}_q)[l].$$

Let us describe how one constructs this isomorphism. On E_{et} we have the short exact sequence of sheaves, called the Kummer sequence:

$$0 \longrightarrow \mu_l \longrightarrow \mathbb{G}_{\mathrm{m}} \longrightarrow \mathbb{G}_{\mathrm{m}} \longrightarrow 0$$

where the map on $\mathbb{G}_{\mathrm{m}}$ is multiplication by l in the group law of $\mathbb{G}_{\mathrm{m}}$, i.e., taking lth powers. This short exact sequence gives an exact sequence of cohomology groups after pullback to $E_{\overline{\mathbb{F}}_q,\mathrm{et}}$:

$$\mu_l(\overline{\mathbb{F}}_q) \hookrightarrow \overline{\mathbb{F}}_q^{\times} \to \overline{\mathbb{F}}_q^{\times} \to \mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mu_l) \to \mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{G}_{\mathrm{m}}) \to \mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{G}_{\mathrm{m}}) \to \cdots$$

Just as for any scheme, one has:

$$\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{G}_\mathrm{m}) = \mathrm{Pic}(E_{\overline{\mathbb{F}}_q})$$

It follows that

$$\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mu_l) = \mathrm{Pic}(E_{\overline{\mathbb{F}}_q})[l].$$

Finally, using the exact sequence:

$$0 \longrightarrow \mathrm{Pic}^0(E_{\overline{\mathbb{F}}_q}) \longrightarrow \mathrm{Pic}(E_{\overline{\mathbb{F}}_q}) \xrightarrow{\deg} \mathbb{Z} \longrightarrow 0$$

and the fact that E is its own Jacobian variety, i.e., $\mathrm{Pic}^0(E_{\overline{\mathbb{F}}_q}) = E(\overline{\mathbb{F}}_q)$, we obtain:

$$\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mu_l) = \mathrm{Pic}(E_{\overline{\mathbb{F}}_q})[l] = \mathrm{Pic}^0(E_{\overline{\mathbb{F}}_q})[l] = E(\overline{\mathbb{F}}_q)[l].$$

The choice of an isomorphism between $\mu_l(\overline{\mathbb{F}}_q)$ and $\mathbb{F}_l$ gives us the desired isomorphism between $\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l))$ and $E(\overline{\mathbb{F}}_q)[l]$. In fact, we note that by using the Weil pairing $E(\overline{\mathbb{F}}_q)[l] \times E(\overline{\mathbb{F}}_q)[l] \to \mu_l(\overline{\mathbb{F}}_q)$, we get an isomorphism:

$$\mathrm{H}^1(E_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)) = E(\overline{\mathbb{F}}_q)[l]^\vee$$

that is more natural than the one used above; in particular, it does not depend on the choice of an isomorphism $\mathbb{F}_l \to \mu_l(\overline{\mathbb{F}}_q)$.

4 Some natural new directions

So, indeed, we have seen that the two-dimensional $\mathbb{F}_l$-vector spaces that are used in Schoof's algorithm for elliptic curves can also be seen as étale cohomology groups. A natural question that arises is then the following.

> *Are there other interesting cases where étale cohomology groups can be used to construct polynomial time algorithms for counting rational points of varieties over finite fields?*

4.1 Curves of higher genus

The first step in the direction of this question was taken by Jonathan Pila. In [88] he considered principally polarised Abelian varieties of a fixed dimension, and curves of a fixed genus, and showed that in those cases polynomial time algorithms for computing the number of rational points over finite fields exist. In these cases, the only relevant cohomology groups are in degree

one, i.e., they are of the form $\mathrm{H}^1(X_{\overline{\mathbb{F}}_q,\mathrm{et}}, \mathbb{F}_l)$ with X a smooth proper curve, or an Abelian variety, over the field $\mathbb{F}_q$. As in Schoof's algorithm, the way to deal with these cohomology groups is to view them as $J(\overline{\mathbb{F}}_q)[l]$, the kernel of multiplication by l on the Abelian variety J. In the case where X is a curve, one lets J be the Jacobian variety of X.

As Pila makes use of explicit systems of equations for Abelian varieties, his algorithm has a running time that is at least exponential in the dimension of the Abelian variety, and hence, in the case of curves, as a function of the genus of the curve.

The current state of affairs concerning the question of counting the rational points of curves over finite fields seems still to be the same: algorithms have a running time that is exponential in the genus. As an illustration, let us mention that in [2] Adleman and Huang give an algorithm that computes $\#X(\mathbb{F}_q)$ in time $(\log q)^{O(g^2 \log g)}$, where X is a hyperelliptic curve over $\mathbb{F}_q$, and where g is the genus of X.

Recent progress in the case where the characteristic of the finite fields $\mathbb{F}_q$ is fixed, using so-called p-adic methods, will be discussed in Section 6 below. In that case, there are algorithms whose running time is polynomial in g and $\log q$.

4.2 Higher degree cohomology, modular forms

Another direction in which one can try to generalise Schoof's algorithm is to varieties of higher dimension, where non-trivial cohomology groups of degree higher than one are needed. In this context, we would call the degree 2 cohomology group of a curve trivial, because the trace of F_q on it is q.

More generally speaking, cohomology groups, but now with l-adic coefficients, that are of dimension one are expected to have the property that the trace of F_q can only be of the form $q^n\zeta$, with n an integer greater than or equal to zero, and ζ a root of unity. This means that one-dimensional cohomology groups are not so challenging. Indeed, it is the fact that for elliptic curves over $\mathbb{F}_p$ all integers in the Hasse interval $[p+1-2p^{1/2}, p+1+2p^{1/2}]$ can occur that makes the problem of point counting very different from point counting on non-singular quadric surfaces in $\mathbb{P}^3_{\mathbb{F}_q}$, for example, where the outcome can only be q^2+2q+1 or q^2+1.

It follows that the simplest case to consider is cohomology groups of dimension two, in degree at least two, on which the action of F_q is not given by a simple rule as in the one-dimensional case. Such cohomology groups are provided by modular forms, as we will explain later in Section 8. Let us just say for the moment, that there is a direct relation with elliptic curves, via the concept of *modularity* of elliptic curves over $\mathbb{Q}$, that we will now sketch.

Let E be an elliptic curve over $\mathbb{Q}$, given by some Weierstrass equation. Such a Weierstrass equation can be chosen to have its coefficients in $\mathbb{Z}$. A Weierstrass equation for E with coeffi-

cients in $\mathbb{Z}$ is called *minimal* if its *discriminant* is minimal among all Weierstrass equations for E with coefficients in $\mathbb{Z}$; this discriminant then only depends on E and will be denoted $\mathrm{discr}(E)$. In fact, two minimal Weierstrass equations define isomorphic curves in $\mathbb{P}^2_{\mathbb{Z}}$, the projective plane over $\mathbb{Z}$. In other words, E has a Weierstrass minimal model over $\mathbb{Z}$, that will be denoted by $E_{\mathbb{Z}}$. For each prime number p, we let $E_{\mathbb{F}_p}$ denote the curve over $\mathbb{F}_p$ given by reducing a minimal Weierstrass equation modulo p; it is the fibre of $E_{\mathbb{Z}}$ over $\mathbb{F}_p$. The curve $E_{\mathbb{F}_p}$ is smooth if and only if p does not divide $\mathrm{discr}(E)$. The possible singular fibres have exactly one singular point: an ordinary double point with rational tangents, or with conjugate tangents, or an ordinary cusp. The three types of reduction are called split multiplicative, non-split multiplicative and additive, respectively, after the type of group law that one gets on the complement of the singular point. For each p we then get an integer a_p by requiring the following identity:

$$p + 1 - a_p = \#E(\mathbb{F}_p).$$

This means that for all p, a_p is the trace of F_p on the degree one étale cohomology of $E_{\overline{\mathbb{F}}_p}$, with coefficients in $\mathbb{F}_l$, or in $\mathbb{Z}/l^n\mathbb{Z}$ or in the l-adic numbers $\mathbb{Z}_l$. For p not dividing $\mathrm{discr}(E)$ we know that $|a_p| \leq 2p^{1/2}$. If $E_{\mathbb{F}_p}$ is multiplicative, then $a_p = 1$ or -1 in the split and non-split case. If $E_{\mathbb{F}_p}$ is additive, then $a_p = 0$. We also define, for each p an element $\varepsilon(p)$ in $\{0, 1\}$ by setting $\varepsilon(p) = 1$ for p not dividing $\mathrm{discr}(E)$ and setting $\varepsilon(p) = 0$ for p dividing $\mathrm{discr}(E)$. The *Hasse-Weil L-function* of E is then defined as:

$$L_E(s) = \prod_p L_{E,p}(s), \qquad L_{E,p}(s) = (1 - a_p p^{-s} + \varepsilon(p) p p^{-2s})^{-1},$$

for s in $\mathbb{C}$ with $\Re(s) > 3/2$ (indeed, the fact that $|a_p| \leq 2p^{1/2}$ implies that the product converges for such s). To explain this function more conceptually, we note that for all p and for all $l \neq p$ we have the identity:

$$1 - a_p t + \varepsilon(p) t^2 = \det(1 - t\mathrm{F}_p^*, \mathrm{H}^1(E_{\overline{\mathbb{F}},\mathrm{et}}, \mathbb{Q}_l))$$

The reader should notice that now we use étale cohomology with coefficients in $\mathbb{Q}_l$, the field of l-adic numbers, and not in $\mathbb{F}_l$. The reason for this is that we want the last identity above to be an identity between polynomials with integer coefficients, and not with coefficients in $\mathbb{F}_l$.

The function L_E was conjectured to have a holomorphic continuation over all of $\mathbb{C}$, and to satisfy a certain precisely given functional equation relating the values at s and $2 - s$. In that functional equation appears a certain positive integer N_E called the *conductor* of E, composed of the primes p dividing $\mathrm{discr}(E)$ with exponents that depend on the behaviour of E at p, i.e., on $E_{\mathbb{Z}_p}$. This conjecture on continuation and functional equation was proved for semistable E

(i.e., E such that there is no p where E has additive reduction) by Wiles and Taylor-Wiles, and in the general case by Breuil, Conrad, Diamond and Taylor; see [34] for an overview of this. In fact, the continuation and functional equation are direct consequences of the modularity of E that was proved by Wiles, Taylor-Wiles etc. (see below). The weak Birch and Swinnerton-Dyer conjecture says that the dimension of the $\mathbb{Q}$-vector space $\mathbb{Q} \otimes E(\mathbb{Q})$ is equal to the order of vanishing of L_E at 1; see [86], Section 3.5, Part II for more information. Anyway, the function L_E gives us integers a_n for all $n \geq 1$ as follows:

$$L_E(s) = \sum_{n \geq 1} a_n n^{-s}, \quad \text{for } \Re(s) > 3/2.$$

From these a_n one can then consider the following function:

$$f_E \colon \mathbb{H} = \{\tau \in \mathbb{C} \,|\, \Im(\tau) > 0\} \to \mathbb{C}, \quad \tau \mapsto \sum_{n \geq 1} a_n e^{2\pi i n \tau}.$$

Equivalently, we have:

$$f_E = \sum_{n \geq 1} a_n q^n, \quad \text{with} \quad q \colon \mathbb{H} \to \mathbb{C}, \quad \tau \mapsto e^{2\pi i \tau}.$$

A more conceptual way to state the relation between L_E and f_E is to say that L_E is obtained, up to elementary factors, as the *Mellin transform* of f_E:

$$\int_0^\infty f_E(it) t^s \frac{dt}{t} = (2\pi)^{-s} \Gamma(s) L_E(s), \quad \text{for } \Re(s) > 3/2.$$

After all these preparations, we can finally state what the modularity of E means:

f_E is a modular form of weight two for the congruence subgroup $\Gamma_0(N_E)$ of $\mathrm{SL}_2(\mathbb{Z})$.

For some more details on the concept of modular forms we refer to Section 8. At this moment, we just want to say that the last statement means that f_E has, as Mazur says in the Singh's BBC documentary on Wiles's proof of Fermat's Last Theorem, an enormous amount of symmetry. This symmetry is with respect to the action of $\mathrm{GL}_2(\mathbb{Q})^+$, the group of invertible 2 by 2 matrices with coefficients in $\mathbb{Q}$ whose determinant is positive, on the upper half plane $\mathbb{H}$. This symmetry gives, by Mellin transformation, the functional equation of L_E. Conversely, it had been proved in [110] by Weil that if sufficiently many twists of L_E by Dirichlet characters satisfy the conjectured holomorphic continuation and functional equation, then f_E is a modular form of the type mentioned.

We now remark that Schoof's algorithm implies that, for p prime, the coefficient a_p in the q-expansion of $f_E = \sum_{n \geq 1} a_n q^n$ can be computed in time polynomial in $\log p$. One of the aims of the research project described in this report is to generalise this last fact to certain modular forms of higher weight. Before we give precise definitions in Section 8, we will discuss a typical case in the next section.

5 More historical context: congruences for Ramanujan's τ-function

References for this section are the articles [98], [106] and [19] by Serre, Swinnerton-Dyer and Deligne.

A typical example of a modular form of weight higher than two is the *discriminant* modular form, usually denoted Δ. One way to view Δ is as the holomorphic function on the upper half plane $\mathbb{H}$ given by:

$$\Delta = q\prod_{n\geq 1}(1-q^n)^{24}, \tag{5.1}$$

where q is the function from $\mathbb{H}$ to $\mathbb{C}$ given by $z \mapsto \exp(2\pi i z)$. The coefficients in the power series expansion:

$$\Delta = \sum_{n\geq 1}\tau(n)q^n \tag{5.2}$$

define the famous *Ramanujan τ-function*.

To say that Δ is a modular form of weight 12 for the group $\mathrm{SL}_2(\mathbb{Z})$ means that for all elements $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ of $\mathrm{SL}_2(\mathbb{Z})$ the following identity holds for all z in $\mathbb{H}$:

$$\Delta\left(\frac{az+b}{cz+d}\right) = (cz+d)^{12}\Delta(z), \tag{5.3}$$

which is equivalent to saying that the multi-differential form $\Delta(z)(dz)^{\otimes 6}$ is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$. As $\mathrm{SL}_2(\mathbb{Z})$ is generated by the elements $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0\end{smallmatrix}\right)$, it suffices to check the identity in (5.3) for these two elements. The fact that Δ is q times a power series in q means that Δ is a *cusp form*: it vanishes at "$q = 0$". It is a fact that Δ is the first example of a non-zero cusp form for $\mathrm{SL}_2(\mathbb{Z})$: there is no non-zero cusp form for $\mathrm{SL}_2(\mathbb{Z})$ of weight smaller than 12, i.e., there are no non-zero holomorphic functions on $\mathbb{H}$ satisfying (5.3) with the exponent 12 replaced by a smaller integer, whose Laurent series expansion in q is q times a power series. Moreover, the $\mathbb{C}$-vector space of such functions of weight 12 is one-dimensional, and hence Δ is a basis of it.

The one-dimensionality of this space has as a consequence that Δ is an eigenform for certain operators on this space, called *Hecke operators*, that arise from the action on $\mathbb{H}$ of $\mathrm{GL}_2(\mathbb{Q})^+$, the subgroup of $\mathrm{GL}_2(\mathbb{Q})$ of elements whose determinant is positive. This fact explains that the coefficients $\tau(n)$ satisfy certain relations which are summarised by the following identity of Dirichlet series (converging for $\Re(s) \gg 0$, for the moment, or just formal series, if one prefers

that):

$$L_\Delta(s) := \sum_{n\geq 1} \tau(n) n^{-s} = \prod_p (1 - \tau(p)p^{-s} + p^{11}p^{-2s})^{-1}. \tag{5.4}$$

These relations:

$$\begin{aligned} &\tau(mn) = \tau(m)\tau(n) && \text{if } m \text{ and } n \text{ are relatively prime} \\ &\tau(p^n) = \tau(p^{n-1})\tau(p) - p^{11}\tau(p^{n-2}) && \text{if } p \text{ is prime and } n \geq 2 \end{aligned}$$

were conjectured by Ramanujan, and proved by Mordell. Using these identities, $\tau(n)$ can be expressed in terms of the $\tau(p)$ for p dividing n.

As L_Δ is the Mellin transform of Δ, L_Δ is holomorphic on $\mathbb{C}$, and satisfies the functional equation (Hecke):

$$(2\pi)^{-(12-s)}\Gamma(12-s)L_\Delta(12-s) = (2\pi)^{-s}\Gamma(s)L_\Delta(s).$$

The famous *Ramanujan conjecture* states that for all primes p one has the inequality:

$$|\tau(p)| < 2p^{11/2}, \tag{5.5}$$

or, equivalently, that the complex roots of the polynomial $x^2 - \tau(p)x + p^{11}$ are complex conjugates of each other, and hence are of absolute value $p^{11/2}$. This conjecture was proved by Deligne as a consequence of his article [19] and his proof of the analog of the Riemann hypothesis in the Weil conjectures in [20].

Finally, Ramanujan conjectured congruences for the integers $\tau(p)$ with p prime, modulo certain powers of certain small prime numbers. In order to state these congruences we define, for $n \geq 1$ and $r \geq 0$:

$$\sigma_r(n) := \sum_{1\leq d|n} d^r,$$

i.e., $\sigma_r(n)$ is the sum of the rth powers of the positive divisors of n. We will now list the congruences that are given in the first pages of [106]:

$$\begin{aligned} &\tau(n) \equiv \sigma_{11}(n) && \bmod 2^{11} && \text{if } n \equiv 1 \mod 8 \\ &\tau(n) \equiv 1217\sigma_{11}(n) && \bmod 2^{13} && \text{if } n \equiv 3 \mod 8 \\ &\tau(n) \equiv 1537\sigma_{11}(n) && \bmod 2^{12} && \text{if } n \equiv 5 \mod 8 \\ &\tau(n) \equiv 705\sigma_{11}(n) && \bmod 2^{14} && \text{if } n \equiv 7 \mod 8 \end{aligned}$$

$$\tau(n) \equiv n^{-610}\sigma_{1231}(n) \qquad \mod 3^6 \qquad \text{if } n \equiv 1 \mod 3$$
$$\tau(n) \equiv n^{-610}\sigma_{1231}(n) \qquad \mod 3^7 \qquad \text{if } n \equiv 2 \mod 3$$

$$\tau(n) \equiv n^{-30}\sigma_{71}(n) \qquad \mod 5^3 \qquad \text{if } n \text{ is prime to } 5$$

$$\tau(n) \equiv n\sigma_9(n) \qquad \mod 7 \qquad \text{if } n \equiv 0, 1, 2 \text{ or } 4 \mod 7$$
$$\tau(n) \equiv n\sigma_9(n) \qquad \mod 7^2 \qquad \text{if } n \equiv 3, 5 \text{ or } 6 \mod 7$$

$$\tau(p) \equiv 0 \qquad \mod 23 \qquad \text{if } p \text{ is prime and not a square mod } 23$$
$$\tau(p) \equiv 2 \qquad \mod 23 \qquad \text{if } p \neq 23 \text{ is a prime of the form } u^2 + 23v^2$$
$$\tau(p) \equiv -1 \qquad \mod 23 \qquad \text{for other primes } p \neq 23$$

$$\tau(n) \equiv \sigma_{11}(n) \mod 691$$

The reader is referred to [106] for the origin and for proofs of these congruences. There, Swinnerton-Dyer remarks that the proofs do little explain why such congruences occur. Serre conjectured an explanation in [98]. First of all, Serre conjectured the existence, for each prime number l, of a continuous representation:

$$\rho_l \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{Aut}(V_l), \tag{5.6}$$

with V_l a two-dimensional $\mathbb{Q}_l$-vector space, such that ρ_l is unramified at all primes $p \neq l$, and such that for all $p \neq l$ the characteristic polynomial of $\rho_l(\mathrm{Frob}_p)$ is given by:

$$\det(1 - x\mathrm{Frob}_p, V_l) = 1 - \tau(p)x + p^{11}x^2. \tag{5.7}$$

To help the reader, let us explain what unramified at p means, and what the Frobenius elements Frob_p are. For p prime, we let $\mathbb{Q}_p$ denote the topological field of p-adic numbers, and $\mathbb{Q}_p \to \overline{\mathbb{Q}}_p$ an algebraic closure. The action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the set $\mathrm{Hom}(\overline{\mathbb{Q}}, \overline{\mathbb{Q}}_p)$ of embeddings of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_p$ is transitive, and each embedding induces an injection from $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ into $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$,

the image of which is called a decomposition group of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ at p. The injections from $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ into $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and the corresponding decomposition groups at p obtained like this are all conjugated by the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. In order to go further we need to say a bit about the structure of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. We let $\mathbb{Q}_p^{\mathrm{unr}}$ be the maximal unramified extension of $\mathbb{Q}_p$ in $\overline{\mathbb{Q}}_p$, i.e., the composite of all finite extension K of $\mathbb{Q}_p$ in $\overline{\mathbb{Q}}_p$ such that p is a uniformiser for the integral closure O_K of $\mathbb{Z}_p$ in K. We let $\mathbb{Z}_p^{\mathrm{unr}}$ be the integral closure of $\mathbb{Z}_p$ in $\mathbb{Q}_p^{\mathrm{unr}}$; it is a local ring, and its residue field is an algebraic closure $\overline{\mathbb{F}}_p$ of $\mathbb{F}_p$. The sub-extension $\mathbb{Q}_p^{\mathrm{unr}}$ gives a short exact sequence:

$$I_p \hookrightarrow \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \twoheadrightarrow \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p). \tag{5.8}$$

The subgroup I_p of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ is called the inertia subgroup. The quotient $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$ is canonically isomorphic to $\hat{\mathbb{Z}}$, the profinite completion of $\mathbb{Z}$, by demanding that the element 1 of $\hat{\mathbb{Z}}$ corresponds to the Frobenius element Frob_p of $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$ that sends x to x^p for each x in $\overline{\mathbb{F}}_p$.

Let now ρ_l be a continuous representation from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathrm{GL}(V_l)$ with V_l a finite dimensional $\mathbb{Q}_l$-vector space. Each embedding of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_p$ then gives a representation of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ on V_l. Different embeddings give isomorphic representations because they are conjugated by an element in the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ under ρ_l. We now choose one embedding, and call the representation $\rho_{l,p}$ of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ on V_l obtained like this the local representation at p associated to ρ_l. This being defined, ρ_l is then said to be unramified at a prime p if $\rho_{l,p}$ factors through the quotient $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \to \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$, i.e., if I_p acts trivially on V_l. If ρ_l is unramified at p, then we get an element $\rho_l(\mathrm{Frob}_p)$ in $\mathrm{GL}(V_l)$. This element depends on our chosen embedding of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_p$, but its conjugacy class under $\rho_l(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$ does not. In particular, we get a well-defined conjugacy class in $\mathrm{GL}(V_l)$, and so the characteristic polynomial of $\rho_l(\mathrm{Frob}_p)$ is now defined if ρ_l is unramified at p.

Continuous representations such as ρ_l can be reduced modulo powers of l as follows. The compactness of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ implies that with respect to a suitable basis of V_l the representation ρ_l lands in $\mathrm{GL}_2(\mathbb{Z}_l)$, and hence gives representations to $\mathrm{GL}_2(\mathbb{Z}/l^n\mathbb{Z})$ for all $n \geq 0$. This reduction of ρ_l modulo powers of l is not unique, but the semi-simplification of the reduction modulo l is well-defined, i.e., two reductions lead to the same Jordan-Hölder constituents. According to Serre, the congruences above would then be explained by properties of the image of ρ_l.

For example, if the image of the reduction modulo l of ρ_l is reducible, say an extension of two characters α and β from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathbb{F}_l^*$, then one has the identity in $\mathbb{F}_l$, for all $p \neq l$:

$$\tau(p) \equiv \alpha(\mathrm{Frob}_p) + \beta(\mathrm{Frob}_p). \tag{5.9}$$

The characters α and β are unramified outside l. By the Kronecker-Weber theorem, the maximal Abelian subextension of $\mathbb{Q} \to \overline{\mathbb{Q}}$ that is unramified outside l is the cyclotomic extension gener-

ated by all l-power roots of unity, with Galois group $\mathbb{Z}_l^*$. It then follows that $\alpha = \chi_l^n$ and $\beta = \chi_l^m$ for suitable n and m, where χ_l is the character giving the action on the lth roots of unity in $\overline{\mathbb{Q}}$: for all σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and for all ζ in $\overline{\mathbb{Q}}^*$ with $\zeta^l = 1$ one has $\sigma(\zeta) = \zeta^{\chi_l(\sigma)}$. The identity (5.9) in $\mathbb{F}_l$ above then takes the form:

$$\tau(p) = p^n + p^m \mod l, \quad \text{for all } p \neq l, \tag{5.10}$$

which indeed is of the same form as the congruences mod l for $\tau(p)$ listed above. For example, the congruence mod 691 corresponds to the statement that the reduction modulo l of ρ_l contains the two characters 1 and χ_l^{11}.

Deligne, in [19], proved the existence of the ρ_l, as conjectured by Serre, by showing that they occur in the degree one l-adic étale cohomology of certain sheaves on certain curves, and in the degree 11 étale cohomology with $\mathbb{Q}_l$-coefficients of a variety of dimension 11. This last variety is, loosely speaking, the 10-fold fibred product of the universal elliptic curve. Deligne's constructions will be discussed in detail in Sections 8 and 9. It should be said that Shimura had already shown how to construct Galois representations in the case of modular forms of weight two; in that case one does not need étale cohomology, but torsion points of Jacobians of modular curves suffice, see [102].

In [106], Swinnerton-Dyer gives results, partly resulting from his correspondence with Serre, in which the consequences of the existence of the ρ_l for congruences of $\tau(p)$ modulo l are explored. A natural question to ask is if there are primes l other than 2, 3, 5, 7, 23 and 691 modulo which there are similar congruences for $\tau(p)$.

For each $p \neq l$, $\tau(p)$ is the trace of $\rho_l(\mathrm{Frob}_p)$, and the determinant of $\rho_l(\mathrm{Frob}_p)$ equals p^{11}. Hence, a polynomial relation between $\tau(p)$ and p^{11}, valid modulo some l^n for all $p \neq l$, is a relation between the determinant and the trace of all $\rho_l(\mathrm{Frob}_p)$ in $\mathrm{GL}_2(\mathbb{Z}/l^n\mathbb{Z})$. But Chebotarev's theorem (see [65], or [9], for example) implies that every element of the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}_2(\mathbb{Z}/l^n\mathbb{Z})$ is of the form Frob_p for infinitely many p. Hence, such a polynomial relation is then valid for all elements in the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}_2(\mathbb{Z}/l^n\mathbb{Z})$. For this reason, the existence of non-trivial congruences modulo l^n as above for $\tau(p)$ depends on this image.

The image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathbb{Z}_l^*$ under $\det \circ \rho_l$ is equal to the subgroup of 11th powers in $\mathbb{Z}_l^*$. To explain this, we note that $\det \circ \rho_l$ is a continuous character from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathbb{Z}_l^*$, unramified outside l, and such that Frob_p is mapped to p^{11} for all $p \neq l$; this implies that $\det \circ \rho_l$ is the l-adic cyclotomic character, giving the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the l-power roots of unity.

In order to state the results in [106], one calls a prime number l *exceptional (for Δ)* if the image of ρ_l, taking values in $\mathrm{GL}_2(\mathbb{Z}_l)$, does *not* contain $\mathrm{SL}_2(\mathbb{Z}_l)$. For l not exceptional, i.e., such that the image of ρ_l contains $\mathrm{SL}_2(\mathbb{Z}_l)$, the image of ρ_l in $\mathrm{GL}_2(\mathbb{Z}_l)$ is the subgroup of elements x such that $\det(x)$ is an 11th power in $\mathbb{Z}_l^*$. For x in this subgroup, the trace of x modulo l is not

determined by its determinant modulo l, hence there can be no congruence for $\tau(p)$ modulo l as above.

The Corollary to Theorem 4 in [106] states, among others, that the list of primes that are exceptional for Δ is $\{2, 3, 5, 7, 23, 691\}$. The main tool that is used and that we have not discussed is the theory of modular forms modulo l, or, equivalently, the theory of congruences modulo l between modular forms. As a consequence, there are no similar congruences for $\tau(p)$ modulo primes other than the ones listed above. The special form of the congruences modulo 23 is explained by the fact that in that case the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}_2(\mathbb{F}_{23})$ is dihedral; in the other cases the residual representation, i.e., the representation to $\mathrm{GL}_2(\mathbb{F}_l)$, is reducible. In the case $l = 2$, Swinnerton-Dyer has determined the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}_2(\mathbb{Z}_2)$ exactly: see the appendix in [106].

The direction in which we generalise Schoof's algorithm is to give an algorithm that computes for prime numbers l that are not exceptional for Δ the field extension $\mathbb{Q} \to K_l$ that corresponds to the representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathrm{GL}_2(\mathbb{F}_l)$ that comes from Δ. The field K_l is given in the form $\mathbb{Q}[x]/(f_l)$. The computation has a running time that is polynomial in l. It is fair to say that this algorithm makes the mod l Galois representations associated to Δ accessible to computation, at least theoretically. As the field extensions that are involved are non-solvable, this should be seen as a step beyond computational class field theory, and beyond the case of elliptic curves, in the direction to make the results of Langlands' program accessible to computations.

As a consequence, one can compute $\tau(p)$ mod l in time polynomial in $\log p$ and l, by factoring f_l as above mod p and some more computations that will be described later. By doing this for sufficiently many l, just as in Schoof's algorithm, one then gets an algorithm that computes $\tau(p)$ in time polynomial in $\log p$.

In principle, the method used here applies to more general modular forms that are eigenforms for the Hecke operators. In particular, if $f = \sum_{n\geq 1} a_n(f)q^n$ is a such a form, normalised by the condition that $a_1(f) = 1$, the coefficients $a_p(f)$ with p prime can be computed in time polynomial in $\log p$. We should mention, however, that the exponent of $\log p$ depends on the degree of the number field $E(f)$ generated by the $a_n(f)$. The reason for that is that for each prime number l the l-adic Galois representation associated to f takes values in $\mathrm{GL}_2(\mathbb{Q}_l \otimes E(f))$, which leads to residual (mod l) Galois representations with images in $\mathrm{GL}_2(\mathbb{F}_{l^i})$ with i depending on how l decomposes in E. Therefore, the Galois extensions of $\mathbb{Q}$ to be computed are of degree about l^{4d}, where $d = \dim_{\mathbb{Q}}(E(f))$, in the worst case, where l is inert in E. The possibility of using only the residual representations taking place in various $\mathrm{GL}_2(\mathbb{F}_l)$ should be considered. Effective Chebotarev theorems (see [99]) should then be used to get lower bounds for the number of maximal ideals of degree one of the ring of integers $O_{E(f)}$ of $E(f)$.

6 Comparison with p-adic methods

Before we start seriously with the theory of modular forms and the Galois representations associated to them in the next sections, we make a comparison between our generalisation of Schoof's algorithm and the so-called p-adic methods that have been developed recently by Satoh [94], Kedlaya [62] (see also [36], and [55]), Wan and Lauder [70], [71], [68] and [69], Fouquet, Gaudry, Gürel and Harley [43], [48], Denef and Vercauteren [27], Mestre, Lercier and Lubicz [75], Carls (thesis, Groningen and Leiden) and Gerkmann (thesis, Essen). Actually, we should notice that such a method was already introduced in [60] in 1982, but that this article seems to have been forgotten (I thank Fre Vercauteren for having drawn my attention to this article).

In all these methods, one works with fields of small characteristic p, hence of the form $\mathbb{F}_q$ with $q = p^m$ and p fixed. All articles cited in the previous paragraph have the common property that they give algorithms for computing the number of $\mathbb{F}_q$-rational points on certain varieties X over $\mathbb{F}_q$, using, sometimes indirectly, cohomology groups with p-adic coefficients, whence the terminology "p-adic methods".

For example, Satoh [94] uses the canonical lift of ordinary elliptic curves and the action of the lifted Frobenius endomorphism on the tangent space, which can be interpreted in terms of the algebraic de Rham cohomology of the lifted curve. Kedlaya [62] uses Monsky-Washnitzer cohomology of certain affine pieces of hyperelliptic curves. In fact, all cohomology groups used here are de Rham type cohomology groups, given by complexes of differential forms on certain p-adic lifts of the varieties in question. Just as an example, let us mention that Kedlaya [62] gives an algorithm that for fixed $p \neq 2$ computes the zeta functions of hyperelliptic curves given by equations:

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

where f has arbitrary degree, in time $m^3 \deg(f)^4$. The running times of the other algorithms are all similar, but all have in common that the running time grows at least linearly in p, hence exponentially in $\log p$. The explanation for this is that somehow in each case non-sparse polynomials of degree at least linear in p have to be manipulated.

Summarising this recent progress, one can say that, at least from a theoretical point of view, the problem of counting the solutions of systems of polynomial equations over finite fields of a fixed characteristic p and in a fixed number of variables has been solved. However, if p is not bounded, then almost nothing is known about the existence of polynomial time algorithms.

A very important difference between the project described here, using étale cohomology with coefficients in $\mathbb{F}_l$, and the p-adic methods, is that the Galois representations on $\mathbb{F}_l$-vector spaces that we obtain are *global* in the sense that they are representations of the absolute Galois group

of the global field $\mathbb{Q}$. The field extensions such as the $K_l = \mathbb{Q}[x]/(f_l)$ arising from Δ discussed in the previous section have the advantage that one can choose to do the required computations over the complex numbers, approximating f_l, or p-adically at some suitable prime p, or in $\mathbb{F}_p$ for sufficiently many small p. Also, as we have said already, being able to compute such field extensions K_l, that give mod l information on the Frobenius elements at all primes $p \neq l$, is very interesting. On the other hand, the p-adic methods force one to compute with p-adic numbers, or, actually, modulo some sufficiently high power of p, and it gives information only on the Frobenius at p. The main drawback of the étale cohomology with $\mathbb{F}_l$-coefficients seems to be that the degree of the field extensions as K_l to be dealt with grows exponentially in the dimension of the cohomology groups; for that reason, we do not know how to use étale cohomology to compute $\#X(\mathbb{F}_q)$ for X a curve of arbitrary genus in a time polynomial in $\log q$ and the genus of X.

7 Modular curves

As a good reference for getting an overview of the theory of modular curves and modular forms we recommend the article [28] by Fred Diamond and John Im. This reference is quite complete as results are concerned, and gives good references for the proofs of those results. Moreover, it is one of the few references that treats the various approaches to the theory of modular forms, from the classical analytic theory on the upper half plane to the more modern representation theory of adelic groups. Another good first introduction could be the book [29]. An on-line reference for some of the theory is [33]. Let us also mention that there is a forthcoming book "Modular forms and the Ramanujan conjecture" by Brian Conrad, and also the information in the *wikipedia* is getting more and more detailed, see:

`http://en.wikipedia.org/wiki/Modular_form`.

In this section our aim is just to give the necessary definitions and results for what we need later (and we need at least to fix our notation). Readers who want more details, or more conceptual explanations are encouraged to consult [28] and the wikipedia.

7.1 Definition For n an integer greater than or equal to one we let $\Gamma(n)$ be the kernel of the surjective morphism of groups $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ given by reduction of the coefficients modulo n, and we let $\Gamma_1(n)$ be the inverse image of the subgroup of $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ that fixes the element $(1,0)$ of $(\mathbb{Z}/n\mathbb{Z})^2$. Similarly, we let $\Gamma_0(n)$ be the inverse image of the subgroup of $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ that fixes the subgroup $\mathbb{Z}/n\mathbb{Z}\cdot(1,0)$ of $(\mathbb{Z}/n\mathbb{Z})^2$. Hence the elements of $\Gamma_0(n)$ are the $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ of $\mathrm{SL}_2(\mathbb{Z})$ such that $c \equiv 0 \mod n$, those of $\Gamma_1(n)$ are the ones that satisfy the extra

conditions $a \equiv 1 \mod n$ and $d \equiv 1 \mod n$ and those of $\Gamma(n)$ are the ones that satisfy the extra condition $b \equiv 0 \mod n$.

The group $\mathrm{SL}_2(\mathbb{R})$ acts on the upper half plane $\mathbb{H}$ by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

The subgroup $\mathrm{SL}_2(\mathbb{Z})$ of $\mathrm{SL}_2(\mathbb{R})$ acts discontinuously in the sense that for each z in $\mathbb{H}$ the stabiliser $\mathrm{SL}_2(\mathbb{Z})_z$ is finite and there is an open neighbourhood U of z such that each translate γU with γ in $\mathrm{SL}_2(\mathbb{Z})$ contains exactly one element of the orbit $\mathrm{SL}_2(\mathbb{Z})\cdot z$ and any two translates γU and $\gamma' U$ with γ and γ' in $\mathrm{SL}_2(\mathbb{Z})$ are either equal or disjoint. This property implies that the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, equipped with the quotient topology and with, on each open subset U, the $\mathrm{SL}_2(\mathbb{Z})$-invariant holomorphic functions on the inverse image of U, is a complex analytic manifold of dimension one, i.e., each point of the quotient has an open neighbourhood that is isomorphic to the complex unit disk. Globally, the well-known j-function from $\mathbb{H}$ to $\mathbb{C}$ *is* in fact the quotient map for this action. One way to see this is to associate to each z in $\mathbb{H}$ the elliptic curve $E_z := \mathbb{C}/(\mathbb{Z} + \mathbb{Z}z)$, and to note that for z and z' in $\mathbb{H}$ the elliptic curves E_z and $E_{z'}$ are isomorphic if and only if z and z' are in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit, and to use the fact that two complex elliptic curves are isomorphic if and only if their j-invariants are equal.

The quotient $\Gamma(n)\backslash\mathbb{H}$ can be identified with the set of isomorphism classes of pairs (E, ϕ) where E is a complex elliptic curve and $\phi\colon (\mathbb{Z}/n\mathbb{Z})^2 \to E[n]$ is an isomorphism of groups, compatible with the Weil pairing $E[n] \times E[n] \to \mu_n(\mathbb{C})$ and the $\mu_n(\mathbb{C})$-valued pairing on $(\mathbb{Z}/n\mathbb{Z})^2$ that sends $((a_1, a_2), (b_1, b_2))$ to $\zeta_n^{a_1 b_2 - a_2 b_1}$, where $\zeta_n = e^{2\pi i/n}$.

The quotient $\Gamma_0(n)\backslash\mathbb{H}$ is then identified with the set of pairs (E, G) where E is a complex elliptic curve, and $G \subset E$ a subgroup that is isomorphic to $\mathbb{Z}/n\mathbb{Z}$. Equivalently, we may view $\Gamma_0(n)\backslash\mathbb{H}$ as the set of isomorphism classes of $E_1 \xrightarrow{\phi} E_2$, where ϕ is a morphism of complex elliptic curves, and $\ker(\phi)$ is isomorphic to $\mathbb{Z}/n\mathbb{Z}$.

Finally, the quotient $\Gamma_1(n)\backslash\mathbb{H}$ is then identified with the set of pairs (E, P) where E is a complex elliptic curve, and P is a point of order n of E. Explicitly: to each z in $\mathbb{H}$ corresponds the pair $(\mathbb{C}/(\mathbb{Z}z + \mathbb{Z}), [1/n])$, where $[1/n]$ denotes the image of $1/n$ in $\mathbb{C}/(\mathbb{Z}z + \mathbb{Z})$.

In order to understand that the quotients considered above are in fact the complex analytic varieties associated to affine complex algebraic curves, it is necessary (and sufficient!) to show that these quotients can be compactified to compact Riemann surfaces by adding a finite number of points, called *the cusps*. As the quotient by $\mathrm{SL}_2(\mathbb{Z})$ is given by $j\colon \mathbb{H} \to \mathbb{C}$, it can be compactified easily by embedding $\mathbb{C}$ into $\mathbb{P}^1(\mathbb{C})$; the point ∞ of $\mathbb{P}^1(\mathbb{C})$ is called the cusp. Another way to view this is to note that the equivalence relation on $\mathbb{H}$ given by the action of $\mathrm{SL}_2(\mathbb{Z})$ identifies

two elements z and z' with $\Im(z) > 1$ and $\Im(z') > 1$ if and only if $z' = z + n$ for some n in $\mathbb{Z}$; this follows from the identity:

$$\Im\left(\frac{az+b}{cz+d}\right) = \frac{\Im(z)}{|cz+d|^2}, \quad \text{for all } \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \text{ in } \mathrm{SL}_2(\mathbb{R}) \text{ and } z \text{ in } \mathbb{H}. \tag{7.2}$$

Indeed, if moreover $c \neq 0$, then:

$$\frac{\Im(z)}{|cz+d|^2} \leq \frac{\Im(z)}{(\Im(cz))^2} = \frac{1}{c^2\Im(z)}. \tag{7.3}$$

Hence on the part "$\Im(z) > 1$" of $\mathbb{H}$ the equivalence relation given by $\mathrm{SL}_2(\mathbb{Z})$ is given by the action of $\mathbb{Z}$ by translation. As the quotient for that action is given by the map $q\colon \mathbb{H} \to D(0, e^{-2\pi})^*$, $z \mapsto \exp(2\pi i z)$, where $D(0, e^{-2\pi})^*$ is the open disk of radius $e^{-2\pi}$, centred at 0, and with 0 removed, we get an open immersion of $D(0, e^{-2\pi})^*$ into $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. The compactification is then obtained by replacing $D(0, e^{-2\pi})^*$ with $D(0, e^{-2\pi})$, i.e., by adding the centre back into the punctured disk.

Let us now consider the problem of compactifying the other quotients above. Let Γ be one of the groups considered above, or, in fact, any subgroup of finite index in $\mathrm{SL}_2(\mathbb{Z})$. We consider the morphism $f\colon \Gamma\backslash\mathbb{H} \to \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} = \mathbb{C}$, and our compactification $\mathbb{P}^1(\mathbb{C})$ of $\mathbb{C}$. By construction, f is proper (i.e., the inverse image of a compact subset of $\mathbb{C}$ is compact). Also, we know that ramification can only occur at points with j-invariant 0 or 1728. Let D^* be the punctured disk described above. Then $f\colon f^{-1}D^* \to D^*$ is an unramified covering of degree $\#\mathrm{SL}_2(\mathbb{Z})/\Gamma$ if Γ does not contain -1, and of degree $(\#\mathrm{SL}_2(\mathbb{Z})/\Gamma)/2$ if -1 is in Γ. Up to isomorphism, the only connected unramified covering of degree n, with $n \geq 1$, of D^* is the map $D_n^* \to D^*$, with $D_n^* = \{z \in \mathbb{C} \,|\, 0 < |z| < e^{-2\pi/n}\}$, sending $z \mapsto z^n$. It follows that $f^{-1}D^*$ is, as a covering of D^*, a disjoint union of copies of such $D_n^* \to D^*$. Each D_n^* has the natural compactification $D_n := \{z \in \mathbb{C} \,|\, |z| < e^{-2\pi/n}\}$. We compactify $\Gamma\backslash\mathbb{H}$ by adding the origin to each punctured disk in $f^{-1}D^*$. The points that we have added are called the cusps. By construction, the morphism $f\colon \Gamma\backslash\mathbb{H} \to \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ extends to the compactifications. It is a fact that a compact Riemann surface can be embedded into some projective space, using the theorem of Riemann-Roch, and that the image of such an embedding is a complex algebraic curve. This means that our quotients are, canonically, the Riemann surfaces associated to smooth complex algebraic curves.

7.4 Definition For $n \geq 1$ we define $X(n)$, $X_1(n)$ and $X_0(n)$ to be the proper smooth complex algebraic curves obtained via the compactifications of $\Gamma(n)\backslash\mathbb{H}$, $\Gamma_1(n)\backslash\mathbb{H}$, and $\Gamma_0(n)\backslash\mathbb{H}$, respectively. The affine parts obtained by removing the cusps are denoted $Y(n)$, $Y_1(n)$ and $Y_0(n)$.

The next step in the theory is to show that these complex algebraic curves are naturally defined over certain number fields. Let us start with the $X_0(n)$ and $X_1(n)$, which are defined over $\mathbb{Q}$.

A simple way to produce a model of $X_0(n)$ over $\mathbb{Q}$, i.e., an algebraic curve $X_0(n)_{\mathbb{Q}}$ over $\mathbb{Q}$ that gives $X_0(n)$ via extension of scalars via $\mathbb{Q} \to \mathbb{C}$, is to use the map:

$$(j, j') \colon \mathbb{H} \longrightarrow \mathbb{C} \times \mathbb{C}, \quad z \mapsto (j(z), j(nz)).$$

This map factors through the action of $\Gamma_0(n)$, and induces a map from $X_0(n)$ to $\mathbb{P}^1 \times \mathbb{P}^1$ that is birational to its image. This image is a curve in $\mathbb{P}^1 \times \mathbb{P}^1$, hence the zero locus of a bi-homogeneous polynomial often denoted Φ_n, the minimal polynomial of j' over $\mathbb{C}(j)$. One can then check, using some properties of the j-function, that Φ_n has integer coefficients. The normalisation of the curve in $\mathbb{P}^1_{\mathbb{Q}} \times \mathbb{P}^1_{\mathbb{Q}}$ defined by Φ_n is then the desired curve $X_0(n)_{\mathbb{Q}}$. As Φ_n has coefficients in $\mathbb{Z}$, it even defines a curve in $\mathbb{P}^1_{\mathbb{Z}} \times \mathbb{P}^1_{\mathbb{Z}}$ (here, one has to work with schemes), whose normalisation $X_0(n)_{\mathbb{Z}}$ can be characterised as a so-called *coarse moduli space*. For this notion, and for the necessary proofs, the reader is referred to [28, II.8], to [21] and to [61], and to [33]. One consequence of this statement is that for any algebraically closed field k in which n is invertible, the k-points of $Y_0(n)_{\mathbb{Z}}$ (the complement of the cusps) correspond bijectively to isomorphism classes of $E_1 \stackrel{\phi}{\to} E_2$ where ϕ is a morphism of elliptic curves over k of which the kernel is cyclic of order n.

The notion of moduli space also gives natural models over $\mathbb{Z}[1/n]$ of $X_1(n)$ and $Y_1(n)$. For $n \geq 4$ the defining property of $Y_1(n)_{\mathbb{Z}[1/n]}$ is not hard to state. There is an elliptic curve $\mathbb{E}$ over $Y_1(n)_{\mathbb{Z}[1/n]}$ with a point $\mathbb{P}$ in $E(Y_1(n)_{\mathbb{Z}[1/n]})$ that has order n in every fibre, such that *any* pair $(E/S, P)$ with S a $\mathbb{Z}[1/n]$-scheme and P in $E(S)$ of order n in all fibres arises by a *unique* base change:

$$\begin{array}{ccc} E & \to & \mathbb{E} \\ \downarrow & & \downarrow \\ S & \to & Y_1(n)_{\mathbb{Z}[1/n]} \end{array}$$

that is compatible with the sections P and $\mathbb{P}$. The pair $(\mathbb{E}/Y_1(n)_{\mathbb{Z}[1/n]}, \mathbb{P})$ is therefore called *universal*.

The moduli interpretation of $X(n)$ is a bit more complicated, because of the occurrence of the Weil pairing on $E[n]$ that we have seen above. The curve $X(n)$ has a natural model $X(n)_{\mathbb{Z}[1/n,\zeta_n]}$ over $\mathbb{Z}[1/n, \zeta_n]$. The complement of the cusps $Y(n)_{\mathbb{Z}[1/n,\zeta_n]}$ then has an elliptic curve $\mathbb{E}$ over it, and an isomorphism ϕ between the constant group scheme $(\mathbb{Z}/n\mathbb{Z})^2$ and $\mathbb{E}[n]$ that respects the pairings on each side. The pair $(\mathbb{E}/Y(n)_{\mathbb{Z}[1/n,\zeta_n]}, \phi)$ is universal in the same sense as above. We warn the reader that the notation $X(n)$ is also used sometimes for the moduli scheme for pairs (E, ϕ) where ϕ does not necessarily respect the pairings on the two sides.

8 Modular forms

Let us now turn our attention to modular forms. It will be enough for us to work with modular forms for the congruence subgroups $\Gamma_1(n)$. Therefore, we restrict ourselves to that case.

8.1 Definition Let $n \geq 1$ and k an integer. A (holomorphic) *modular form* for $\Gamma_1(n)$ is a holomorphic function $f\colon \mathbb{H} \to \mathbb{C}$ that satisfies the following properties:

1. for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\Gamma_1(n)$ and for all z in $\mathbb{H}$: $f((az+b)/(cz+d)) = (cz+d)^k f(z)$;
2. f is holomorphic at the cusps (see below for an explanation).

A modular form is called a *cuspform* if it vanishes at the cusps.

We still need to explain the condition that f is holomorphic at the cusps. In order to do that, we first explain what this means at the cusp ∞. That cusp is the point that was added to the punctured disk obtained by taking the quotient of $\mathbb{H}$ by the unipotent subgroup $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$, which acts on $\mathbb{H}$ by translations by integers. The coordinate of that disk is q, the map that sends z to $\exp(2\pi i z)$. Therefore, f admits a Laurent series expansion in q:

$$f = \sum_{n \in \mathbb{Z}} a_n(f) q^n, \quad \text{called the } q\text{-expansion at } \infty. \tag{8.2}$$

With this notation, f is called holomorphic at ∞ if $a_n(f)$ is zero for all $n < 0$, and f is said to vanish at ∞ if $a_n(f)$ is zero for all $n \leq 0$.

To state this condition at the other cusps, we need some description of the set of cusps. First, we note that $\mathbb{P}^1(\mathbb{C}) - \mathbb{P}^1(\mathbb{R})$ is the same as $\mathbb{C} - \mathbb{R}$, and therefore the disjoint union of $\mathbb{H}$ and its complex conjugate (which explains, by the way, that $\mathrm{GL}_2(\mathbb{R})^+$ acts by fractional linear transformations on $\mathbb{H}$). We can then consider $\mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$ inside $\mathbb{P}^1(\mathbb{C})$, with the $\mathrm{SL}_2(\mathbb{Z})$ action on it. Then the subgroup $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ stabilises the point $\infty = (1:0)$ of $\mathbb{P}^1(\mathbb{Q})$, and ∞ can be naturally identified with the origin that we added to the disk D^* above, because ∞ is the unique element of $\mathbb{P}^1(\mathbb{Q})$ that lies in the closure of the inverse image "$\Im(z) > 1$" of D^* in $\mathbb{H}$. Then, the images of the region "$\Im(z) > 1$" under the action of elements of $\mathrm{SL}_2(\mathbb{Z})$ correspond bijectively to the elements of $\mathbb{P}^1(\mathbb{Q})$ (note that $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q}) = \mathbb{P}^1(\mathbb{Z})$), and also to the maximal unipotent subgroups of $\mathrm{SL}_2(\mathbb{Z})$ (i.e., the subgroups that consist of elements whose eigenvalues are 1). It follows that we can identify the set of cusps of $X_1(n)$ with $\Gamma_1(n)\backslash\mathbb{P}^1(\mathbb{Q})$, and that the images of the region "$\Im(z > 1)$" under $\mathrm{SL}_2(\mathbb{Z})$ give us punctured disks around the other cusps. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be an element of $\mathrm{SL}_2(\mathbb{Z})$. The conditions of holomorphy and vanishing at the cusp $\gamma\infty = (a:b)$ are then given in terms of the q-expansion of $z \mapsto (cz+d)^{-k} f(\gamma z)$ at ∞. The group $\gamma^{-1}\Gamma_1(n)\gamma$ contains the group $\left(\begin{smallmatrix} 1 & n* \\ 0 & 1 \end{smallmatrix}\right)$ (indeed, $\Gamma_1(n)$ contains $\Gamma(n)$ and that

one is normal in $\mathrm{SL}_2(\mathbb{Z})$). Therefore, putting $q_n\colon \mathbb{H} \to \mathbb{C}$, $z \mapsto \exp(2\pi i z/n)$, the function $z \mapsto (cz+d)^{-k} f(\gamma z)$ then has a Laurent series expansion in q_n, and one asks that this Laurent series is a power series (for holomorphy) or a power series with constant term zero (for vanishing).

8.3 Example Some simple examples of modular forms for $\mathrm{SL}_2(\mathbb{Z})$ are given by Eisenstein series. For each even $k \geq 4$ one has:

$$E_k = 1 - \frac{2k}{B_k} \sum_{n \geq 1} \sigma_{k-1}(n) q^n,$$

where the B_k are the Bernoulli numbers defined by:

$$\frac{te^t}{e^t - 1} = \sum_{k \geq 0} B_k \frac{t^k}{k!},$$

and where, as before, $\sigma_r(n)$ denotes the sum of the rth powers of the positive divisors of n. In particular, one has the formulas:

$$E_4 = 1 + 240 \sum_{n \geq 1} \sigma_3(n) q^n, \quad E_6 = 1 - 504 \sum_{n \geq 1} \sigma_5(n) q^n, \quad \text{and} \quad \Delta = \frac{E_4^3 - E_6^2}{1728}.$$

8.4 Remark Let us note that, from a computational point of view, the coefficients of q^p with p prime of E_k are very easy to compute, namely, up to a constant they are $1 + p^{k-1}$, but that computing the $\sigma_{k-1}(n)$ for composite n is equivalent to factoring n. This is a strong indication that, for computing coefficients $a_n(f)$ of a modular form f, there is a real difference between the case where n is prime and the case where n is composite.

The space of cuspforms of weight k on $\Gamma_1(n)$ will be denoted $S_k(\Gamma_1(n))$. This space can be interpreted as the space of sections of some holomorphic line bundle $\underline{\omega}^{\otimes k}(-\mathrm{Cusps})$ on $X_1(n)$, if $n \geq 5$ (for $n < 4$ there are non-trivial stabilisers of the action of $\Gamma_1(n)$ on $\mathbb{H}$ that cause a problem, and for $n = 4$ there is a problem at one of the cusps):

$$S_k(\Gamma_1(n)) = \mathrm{H}^0(X_1(n), \underline{\omega}^{\otimes k}(-\mathrm{Cusps})) \quad \text{if } n \geq 5. \tag{8.5}$$

This implies that the spaces $S_k(\Gamma_1(n))$ are finite dimensional, and in fact zero if $k \leq 0$ because the line bundle in question then has negative degree. The restriction to $Y_1(n)$ of the line bundle giving the weight k forms is given by dividing out the action of $\Gamma_1(n)$ on $\mathbb{C} \times \mathbb{H}$ given by:

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\colon (x, z) \mapsto \left((cz+d)^k x, \frac{az+b}{cz+d}\right). \tag{8.6}$$

The extension of this line bundle over the cusps is then given by decreeing that, at the cusp ∞, the constant section 1 (which is indeed invariant under the translations $z \mapsto z+n$) is a generator for the bundle of holomorphic forms, and q times 1 is a generator for the bundle of cusp forms.

The moduli interpretation for $Y_1(n)$ can be extended to the holomorphic line bundles giving the modular forms as follows. Recall that a point on $Y_1(n)$ is an isomorphism class of a pair (E,P) with E a complex elliptic curve and P a point of order n of E. The complex line at (E,P) of the bundle of forms of weight k is then $\underline{\omega}_E^{\otimes k}$, the kth tensor power of the dual of the tangent space at 0 of E. In this way, a modular form f of weight k for $\Gamma_1(n)$ can be described as follows: it is a function that assigns to each (E,P) an element $f(E,P)$ of $\underline{\omega}_E^{\otimes k}$, varying holomorphically with (E,P), and such that it has the right property at the cusps (being holomorphic or vanishing). The function f has to be compatible with isomorphisms: if $\phi\colon E \to E'$ is an isomorphism, and $\phi(P) = P'$, then $f(E,P)$ has to be equal to $(\phi^*)^{\otimes k} f(E',P')$. In what follows we will simply write ϕ^* for $(\phi^*)^{\otimes k}$.

The fact that f should be holomorphic can be stated by evaluating it on the family of elliptic curves that we have over $\mathbb{H}$. Recall that to z in $\mathbb{H}$ we associated the pair $(\mathbb{C}/(\mathbb{Z}z+\mathbb{Z}),[1/n])$. Let us denote x the coordinate of $\mathbb{C}$, then dx is a generator of the cotangent space at 0 of this elliptic curve. Then for f a function as above, we can write:

$$f((\mathbb{C}/(\mathbb{Z}z+\mathbb{Z}),[1/n])) = F_f(z)(dx)^{\otimes k}, \quad F_f\colon \mathbb{H}\to\mathbb{C}. \tag{8.7}$$

The function F_f is then required to be holomorphic. The requirement that f is compatible with isomorphisms means precisely that F_f transforms under $\Gamma_1(n)$ as in Definition 8.1 above. The requirement that f vanishes at the cusps is equivalent to the statement that the Laurent expansions in $q^{1/n}\colon z\mapsto \exp(2\pi i z/n)$ obtained by evaluating f on all pairs $(\mathbb{C}/(\mathbb{Z}z+\mathbb{Z}),(az+b)/n)$, with a and b in $\mathbb{Z}$ such that $(az+b)/n$ is of order n are in fact power series with constant term zero.

The spaces $S_k(\Gamma_1(n))$ are equipped with certain operators, called Hecke operators and diamond operators. These operators arise from the fact that for every element γ of $\mathrm{GL}_2(\mathbb{Q})^+$ the subgroups $\Gamma_1(n)$ and $\gamma\Gamma_1(n)\gamma^{-1}$ are commensurable, i.e., their intersection has finite index in each of them. The diamond operators are then the simplest to describe. For each a in $(\mathbb{Z}/n\mathbb{Z})^\times$, $Y_1(n)$ has the automorphism $\langle a\rangle$ given by the property that it sends (E,P) to (E,aP). This action is then extended on modular forms by:

$$(\langle a\rangle f)(E,P) = f(E,aP). \tag{8.8}$$

Similarly, there are Hecke operators T_m on $S_k(\Gamma_1(n))$ for all $m\geq 1$, defined by:

$$(T_m f)(E,P) = \frac{1}{m}\sum_{\phi} \phi^* f(E_\phi, \phi(P)), \tag{8.9}$$

where the sum runs over all quotients $\phi\colon E \to E_\phi$ of degree m such that $\phi(P)$ is of order n. Intuitively, the operator T_m is to be understood as a kind of averaging operator over all possible isogenies of degree m, however, the normalising factor $1/m$ is not equal to the inverse of the number of such isogenies.

Of course, each element f of $S_k(\Gamma_1(n))$ is determined by its q-expansion $\sum_{m\geq 1} a_m(f)q^m$ at the cusp ∞. The action of the Hecke operators can be expressed in terms of these q-expansions (see [28, (12.4.1)]):

$$a_m(T_r f) = \sum_{\substack{0<d|(r,m)\\(d,n)=1}} d^{k-1} a_{rm/d^2}(\langle d\rangle f), \tag{8.10}$$

for f in $S_k(\Gamma_1(n))$, r and m positive integers.

From this formula, a lot can be deduced. It can be seen that the T_r commute with each other (but there are better ways to understand this). The $\mathbb{Z}$-algebra generated by the T_m for $m \geq 1$ and the $\langle a\rangle$ for a in $(\mathbb{Z}/n\mathbb{Z})^\times$ is in fact generated by the T_m with $m \geq 1$, i.e., one does not need the diamond operators, and also by the T_p for p prime and the $\langle a\rangle$ with a in $(\mathbb{Z}/n\mathbb{Z})^\times$ (see [28, §3.5]). The multiplication rules for the T_m acting on $S_k(\Gamma_1(n))$ can be read off from the formal identity ([28, §3.4]):

$$\sum_{m\geq 1} T_m m^{-s} = \prod_p (1 - T_p p^{-s} + p^{k-1}\langle p\rangle p^{-2s})^{-1}, \tag{8.11}$$

where $\langle p\rangle$ is to be interpreted as zero when p divides n. The fact that the Hecke and diamond operators commute means that they have common eigenspaces. Taking $m = 1$ in (8.10) gives:

$$a_1(T_r f) = a_r(f). \tag{8.12}$$

It follows that if f is a non-zero eigenvector for all T_r, then $a_1(f) \neq 0$, so that we can assume that $a_1(f) = 1$. Then, for all $r \geq 1$, $a_r(f)$ is the eigenvalue for T_r. In particular, this means that the common eigenspaces for the T_r are one-dimensional, and automatically eigenspaces for the diamond operators. Eigenforms with $a_1(f) = 1$ are called *normalised eigenforms*.

From (8.11) above it follows that for a normalised eigenform f one has:

$$L_f(s) := \sum_{n\geq 1} a_n(f) n^{-s} = \prod_p (1 - a_p(f)p^{-s} + p^{k-1}\varepsilon_f(p)p^{-2s})^{-1}, \tag{8.13}$$

where $\varepsilon_f\colon (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^*$ is the character via which the diamond operators act on f, with the convention that $\varepsilon_f(p) = 0$ if p divides n. In particular, the L-function of a modular form has such an Euler product expansion if and only if the modular form is an eigenform for all Hecke operators.

An element of $S_k(\Gamma_1(n))$ that is a normalised eigenform for all Hecke operators is called a *newform* if the system of eigenvalues $a_p(f)$, with p not dividing n, does not occur in a level strictly smaller than n, i.e., in some $S_k(\Gamma_1(m))$ with $m < n$ (actually, we will see in a moment that one only needs to consider the m's dividing n). The set of newforms in $S_k(\Gamma_1(n))$ will be denoted $S_k(\Gamma_1(n))^{\mathrm{new}}$.

Now we want to recall briefly how one obtains a basis of $S_k(\Gamma_1(n))$ in terms of the sets of newforms $S_k(\Gamma_1(m))^{\mathrm{new}}$ for m dividing n. For details and references to proofs, see [28, I.6]. First of all, for each n, $S_k(\Gamma_1(n))^{\mathrm{new}}$ is a linearly independent subset of $S_k(\Gamma_1(n))$, hence finite. For m dividing n and for d dividing n/m, we have a map $B_{n,m,d}\colon X_1(n) \to X_1(m)$, with moduli interpretation $(E, P) \mapsto (E/\langle (n/d)P\rangle, d'P)$, where $dd' = n/m$. For example, this means:

$$B_{n,m,d}\colon (\mathbb{C}/(\mathbb{Z}z + \mathbb{Z}), 1/n) \mapsto (\mathbb{C}/(\mathbb{Z}zd + \mathbb{Z}), 1/m), \tag{8.14}$$

which means that the cusp ∞ of $X_1(n)$ is mapped to the cusp ∞ of $X_1(m)$. Each such map $B_{n,m,d}$ induces by pullback a map:

$$B^*_{n,m,d}\colon S_k(\Gamma_1(m)) \to S_k(\Gamma_1(n)). \tag{8.15}$$

In terms of q-expansions at the cusp ∞ we have, for f in $S_k(\Gamma_1(m))$:

$$B^*_{n,m,d} f = \sum_{r\geq 1} a_r(f) q^{dr}, \tag{8.16}$$

i.e., the effect is just substitution of q by q^d. With these definitions, we can describe a basis for $S_k(\Gamma_1(n))$:

$$\coprod_{m|n} \coprod_{d|(n/m)} B^*_{n,m,d} S_k(\Gamma_1(m))^{\mathrm{new}} \quad \text{is a basis of } S_k(\Gamma_1(n)). \tag{8.17}$$

In the case where $\Gamma_1(n)$ is replaced by $\Gamma_0(n)$, this kind of basis is due to Atkin and Lehner.

In the sequel, we will also make use of a (hermitian) inner product on the $S_k(\Gamma_1(n))$: the Petersson scalar product. It is defined as follows. For f and g in $S_k(\Gamma_1(n))$, viewed as functions on $\mathbb{H}$ as in Definition 8.1 one has:

$$\langle f, g\rangle = \int_{\Gamma_1(n)\backslash\mathbb{H}} f(z)\overline{g(z)}\, y^k \, \frac{dxdy}{y^2}, \tag{8.18}$$

where the integral over $\Gamma_1(n)\backslash\mathbb{H}$ means that one can perform it over any fundamental domain. Indeed, formula 7.2 shows that the function $z \mapsto f(z)\overline{g(z)}y^k$ is invariant under $\Gamma_1(n)$.

We also want to explain the definition of $\langle f, g\rangle$ in terms of the moduli interpretation of $S_k(\Gamma_1(n))$, if $k \geq 2$. For simplicity, let us suppose $n \geq 5$ now. Then $S_k(\Gamma_1(n))$ is the space

of global sections of $\underline{\omega}^{\otimes k}(-\mathrm{Cusps})$ on $X_1(n)$. Now we let $\Omega^1 := \Omega^1_{X_1(n)}$ denote the line bundle of holomorphic differentials on $X_1(n)$. Then there is an isomorphism, named after Kodaira and Spencer:

$$\mathrm{KS}\colon \underline{\omega}^{\otimes 2}(-\mathrm{Cusps}) \xrightarrow{\sim} \Omega^1, \quad \text{Kodaira-Spencer isomorphism.} \tag{8.19}$$

Explicitly, for f in $S_2(\Gamma_1(n))$, viewed as a $\Gamma_1(n)$-invariant section of $\underline{\omega}^{\otimes 2}$ for the family of elliptic curves over $\mathbb{H}$ whose fibre at z is $\mathbb{C}/(\mathbb{Z}z + \mathbb{Z})$ we have:

$$\mathrm{KS}\colon f(dx)^{\otimes 2} \mapsto (2\pi i)^{-2} f\,\frac{dq}{q}. \tag{8.20}$$

Equivalently, for this family of elliptic curves, the Kodaira-Spencer isomorphism sends $(dx)^{\otimes 2}$ to $(2\pi i)^{-2}(dq)/q$. Note that indeed $(dx)^{\otimes 2}$ and $(dq)/q$ transform in the same way under the action of $\mathrm{SL}_2(\mathbb{R})$. We note that without f being required to vanish at the cusps, $\mathrm{KS}(f)$ could have poles of order one at the cusps. The factor $(2\pi i)^{-2}$ is to make the isomorphism compatible with the coordinates $t = \exp(2\pi i x)$ on $\mathbb{C}^\times/\langle \exp(2\pi i z)\rangle$ (which is another way to write $\mathbb{C}/(\mathbb{Z}z + \mathbb{Z})$), and the coordinate $q = \exp(2\pi i z)$ on the unit disk. In those coordinates, that have a meaning "over $\mathbb{Z}$", which means that formulas relating them are power series (or Laurent series) with integer coefficients, KS sends $((dt)/t)^{\otimes 2}$ to $(dq)/q$.

For every complex elliptic curve, the one dimensional complex vector space $\underline{\omega}_E$ has the inner product given by:

$$\langle \alpha, \beta\rangle = \frac{i}{2}\int_E \alpha\,\overline{\beta}, \tag{8.21}$$

where we interpret α and β as translation invariant differential forms on E. The factor $i/2$ comes from the fact that, for $z = x + iy$, one has $dx\,dy = (i/2)dz\,d\overline{z}$. Applying this to the family of elliptic curves $\mathbb{C}/(\mathbb{Z}z + \mathbb{Z})$ over $\mathbb{H}$ gives an inner product on the line bundle $\underline{\omega}$ on $\mathbb{H}$, and also on the line bundle $\underline{\omega}$ on $Y_1(n)$ (recall that we are supposing that $n \geq 5$). Taking tensor powers and duals, this induces inner products on $\underline{\omega}^{\otimes k}$ for all k. The Kodaira-Spencer isomorphism (8.20) gives isomorphisms:

$$\mathrm{KS}\colon \underline{\omega}^{\otimes k}(-\mathrm{Cusps}) \xrightarrow{\sim} \Omega^1 \otimes \underline{\omega}^{\otimes(k-2)} \tag{8.22}$$

For f and g in $S_k(\Gamma_1(n))$, now viewed as sections on $X_1(n)$ of $\underline{\omega}^{\otimes k}(-\mathrm{Cusps})$, one has:

$$\langle f, g\rangle = \frac{i}{2}\int_{X_1(n)} \langle \mathrm{KS}(f), \mathrm{KS}(g)\rangle, \tag{8.23}$$

where the inner product on the left hand side is the Petersson scalar product (8.18), and where for two local sections $\omega_1 \otimes \alpha_1^{\otimes(k-2)}$ and $\omega_2 \otimes \alpha_2^{\otimes(k-2)}$ of $\Omega^1 \otimes \underline{\omega}^{\otimes(k-2)}$ we have defined:

$$\langle \omega_1 \otimes \alpha_1^{\otimes(k-2)}, \omega_2 \otimes \alpha_2^{\otimes(k-2)}\rangle = \langle \alpha_1, \alpha_2\rangle^{k-2}\omega_1\overline{\omega_2} \tag{8.24}$$

The operators T_m on $S_k(\Gamma_1(n))$ with m relatively prime to n are normal: they commute with their adjoint. As a consequence, distinct newforms in $S_k(\Gamma_1(n))$ are orthogonal to each other. On the other hand, the basis (8.17) above of $S_k(\Gamma_1(n))$ is not orthogonal if it consists of more than only newforms.

9 Galois representations associated to eigenforms

The aim of this section is to describe the construction of the Galois representations associated to modular forms, as used in the case of Δ in Section 5. Before giving the construction, let us state the result, which is due, for $k = 2$, to Eichler and Shimura [102], to Deligne [19] for $k > 2$, and to Deligne and Serre [22] for $k = 1$. See Section 12.5 in [28]. A long account of the construction in the case $k \geq 2$ is given in the book [11] (now in preparation) by Brian Conrad.

9.1 Theorem *Let f be a normalised newform, let n be its level, let k be its weight, and let $\varepsilon\colon (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^*$ be its character. Then the subfield K of $\mathbb{C}$ generated over $\mathbb{Q}$ by the $a_n(f)$, $n \geq 1$, and the image of ε is finite over $\mathbb{Q}$. For every prime number l and for any embedding λ of K into $\overline{\mathbb{Q}}_l$, there is a continuous two-dimensional representation V_λ over $\overline{\mathbb{Q}}_l$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ that is unramified outside nl and such that for each prime number p not dividing nl the characteristic polynomial of the Frobenius at p acting on V_λ equals:*

$$\det(1 - x\mathrm{Frob}_p, V_\lambda) = 1 - a_p(f)x + \varepsilon(p)p^{k-1}x^2.$$

For $k \geq 2$ the representations V_λ can be found in the l-adic étale cohomology in degree $k-1$ of some variety of dimension $k-1$, or in the cohomology in degree one of some sheaf on a curve, as we will describe below. The determinant of the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on V_λ is easily described. We let $\chi_l\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Z}_l^\times$ be the l-adic cyclotomic character defined by $\sigma(z) = z^{\chi_l(\sigma)}$, for all σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and all z in $\overline{\mathbb{Q}}^\times$ of l-power order. We let $\varepsilon\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to K^\times$ be the composition of the character $\varepsilon\colon (\mathbb{Z}/n\mathbb{Z})^\times \to K^\times$ with the mod n cyclotomic character $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ given by the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\mu_n(\overline{\mathbb{Q}})$. With these definitions, the determinant of the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on V_λ is given by the character $\varepsilon\chi_l^{k-1}$. As the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ under the determinant of V_λ is infinite, its image in $\mathrm{GL}(V_\lambda)$ is infinite. On the other hand, for $k = 1$, the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{GL}(V_\lambda)$ is finite, and in fact all these representations when λ varies can be realised over some fixed finite extension of $\mathbb{Q}$. The proof of Theorem 9.1 by Deligne and Serre in the case $k = 1$, which uses the result in the case $k \geq 2$, is quite different from the case $k \geq 2$. In particular, no direct construction of the Galois representations for forms of weight one is known. We remark that in the case $k = 2$ the V_λ occur in the first degree étale cohomology

with constant coefficients $\mathbb{Q}_l$ of modular curves, hence can be constructed from l-power torsion points of Jacobians of modular curves (in fact, of the modular curve $X_1(n)$).

The representation V_λ is irreducible by a theorem of Ribet, see Theorem 2.3 of [89], and hence it is characterised by its trace. As the Frobenius conjugacy classes at the primes not dividing nl are dense by Chebotarev's theorem, the representation V_λ is unique up to isomorphism. Non-cuspidal eigenforms lead to Galois representations that are reducible; as our interest lies in going beyond class field theory, we do not discuss this case.

Let us now start the description of the construction, by Deligne, of the representation V_λ as in Theorem 9.1 above in the case where $k \geq 2$. First, if $n < 5$, we replace n by say $5n$ and f by a normalised Hecke eigenform in the 2-dimensional $\mathbb{C}$-vector space generated by $f(q)$ and $f(q^5)$. Then f is no longer a newform, but it is an eigenform, which will be good enough, and as $n \geq 5$ we can view it as a section of the line bundle $\underline{\omega}^{\otimes k}(-\mathrm{Cusps})$ on the smooth complex projective curve $X_1(n)$. The eigenvalues at primes other than 5 have not been changed by this operation. As one can compute from the formulas in the previous section, the two possible eigenvalues for T_5 on the space generated by $f(q)$ and $f(q^5)$ are the two roots of the polynomial $x^2 - a_5(f)x + \varepsilon(5)5^{k-1}$, i.e., the two eigenvalues of the Frobenius element at 5 associated to f if λ does not divide 5. For a detailed computation for this, see Section 4 of [10]; that article also explains why one should expect the two eigenvalues always to be distinct, and that this is a theorem if $k = 2$.

On $Y_1(n)$, we have a universal family $(\mathbb{E}/Y_1(n), \mathbb{P})$ of elliptic curves with a given point of order n. Taking fibre wise the cohomology $\mathrm{H}^1(\mathbb{E}_s, \mathbb{Z})$ gives us a locally constant sheaf on $Y_1(n)$, denoted $\mathrm{R}^1 p_* \mathbb{Z}_\mathbb{E}$ because it is the first higher direct image of the constant sheaf $\mathbb{Z}_\mathbb{E}$ on $\mathbb{E}$ via the morphism $p\colon \mathbb{E} \to Y_1(n)$. We define:

$$\mathcal{F}_k := \mathrm{Sym}^{k-2}(\mathrm{R}^1 p_* \mathbb{Z}_\mathbb{E}), \tag{9.2}$$

where Sym^{k-2} denotes the operation of taking the $(k-2)$th symmetric power. The stalks of the locally constant sheaf $\mathrm{R}^1 p_* \mathbb{Z}_\mathbb{E}$ on $Y_1(n)$ are free $\mathbb{Z}$-modules of rank 2. More concretely, the sheaf $\mathrm{R}^1 p_* \mathbb{Z}_\mathbb{E}$ is obtained from the constant sheaf $\mathbb{Z}^2$ on $\mathbb{H}$ by dividing out the $\Gamma_1(n)$-action given by:

$$\gamma \cdot \left(\begin{pmatrix} n \\ m \end{pmatrix}, \tau \right) = \left(\gamma \cdot \begin{pmatrix} n \\ m \end{pmatrix}, \gamma \cdot \tau \right) = \left(\begin{pmatrix} an + bm \\ cn + dm \end{pmatrix}, \frac{a\tau + b}{c\tau + d} \right), \tag{9.3}$$

where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. The sheaf $\mathcal{F}_k$ is then obtained by dividing out the $\Gamma_1(n)$-action on the constant sheaf $\mathrm{Sym}^{k-2}(\mathbb{Z}^2)$ on $\mathbb{H}$. It is useful to view $\mathbb{Z}^2$ as the $\mathbb{Z}$-submodule $\mathbb{Z}x \oplus \mathbb{Z}y$ of the polynomial ring $\mathbb{Z}[x, y]$. The grading $\mathbb{Z}[x, y] = \oplus_i \mathbb{Z}[x, y]_i$ by the degree then gives the symmetric powers of $\mathbb{Z}x \oplus \mathbb{Z}y$:

$$\mathrm{Sym}^{k-2}(\mathbb{Z}^2) = \mathbb{Z}[x, y]_{k-2} = \bigoplus_{i+j=k-2} \mathbb{Z}x^i y^j. \tag{9.4}$$

We extend the sheaf $\mathcal{F}_k$ to $X_1(n)$ by taking the direct image via the open immersion $j\colon Y_1(n) \to X_1(n)$; this gives us $j_*\mathcal{F}_k$ on $X_1(n)$, again denoted $\mathcal{F}_k$. Outside the cusps, $\mathcal{F}_k$ is locally constant, with stalks free of rank $k-1$ as $\mathbb{Z}$-modules. At the cusps, the stalks of $\mathcal{F}_k$ are free of rank one. At the cusp ∞ this follows from the fact that the subring of invariants of $\mathbb{Z}[x,y]$ for the action of $\left(\begin{smallmatrix}1 & *\\ 0 & 1\end{smallmatrix}\right)$ is $\mathbb{Z}[x]$. At the other cusps it then follows by conjugating with a suitable element of $\mathrm{SL}_2(\mathbb{Z})$. We note that for $k=2$ the sheaf $\mathcal{F}_k$ is the constant sheaf $\mathbb{Z}$ on $X_1(n)$.

The *Eichler-Shimura isomorphism* gives a relation between modular forms and the cohomology of $\mathcal{F}_k$. One way to view this, due to Deligne, is as a Hodge structure. More precisely, the $\mathbb{C}$-vector space $\mathbb{C}\otimes \mathrm{H}^1(X_1(n),\mathcal{F}_k)$ carries a Hodge decomposition:

$$\mathbb{C}\otimes \mathrm{H}^1(X_1(n),\mathcal{F}_k)\tilde{\longrightarrow} S_k(\Gamma_1(n))\oplus \overline{S_k(\Gamma_1(n))}, \tag{9.5}$$

where the two terms on the right are of type $(k-1,0)$ and $(0,k-1)$, respectively. The complex conjugation on the second term on the right comes from the complex conjugation on the tensor factor $\mathbb{C}$ on the left. A good reference for this decomposition and its properties is [4]; we will not go into details here. For an account using group cohomology we refer to Section 12.2 of [28]. For $k=2$ all of this is quite easy. Via the Kodaira-Spencer isomorphism (8.20) it then is the decomposition:

$$\mathrm{H}^1(X_1(n),\mathbb{C}) = \mathrm{H}^0(X_1(n),\Omega^1)\oplus \overline{\mathrm{H}^0(X_1(n),\Omega^1)}. \tag{9.6}$$

We should mention that instead of working with the sheaf $\mathcal{F}_k$ on the curve $X_1(n)$, one can also work with a constant sheaf on a $k-1$-dimensional variety. As before, we let $(\mathbb{E},\mathbb{P})$ denote the universal object over $Y_1(n)$. Then we let $\mathbb{E}^{k-2}$ denote the $k-2$-fold fibre power of $\mathbb{E}$ over $Y_1(n)$; these are the simplest cases of so-called Kuga-Sato varieties. The Künneth formula gives, for s in $Y_1(n)$:

$$\mathrm{H}^{k-2}(\mathbb{E}_s^{k-2},\mathbb{Z}) = \mathrm{H}^1(\mathbb{E}_s,\mathbb{Z})^{\otimes(k-2)} \twoheadrightarrow \mathrm{Sym}^{k-2}(\mathrm{H}^1(\mathbb{E}_s,\mathbb{Z})) = \mathcal{F}_{k,s}. \tag{9.7}$$

In view of the Leray spectral sequence for the cohomology $\mathrm{H}(\mathbb{E}^{k-2},\mathbb{Z})$ of $\mathbb{E}^{k-2}$ in terms of the cohomology of the higher derived direct images $\mathrm{H}(Y_1(n),\mathrm{R}p_*\mathbb{Z}_{\mathbb{E}})$ it is then not so surprising that $S_k(\Gamma_1(n))$ can be identified with a piece of $\mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}},\mathbb{C})$, where $\overline{\mathbb{E}^{k-2}}$ is a certain smooth projective model of $\mathbb{E}^{k-2}$ over $X_1(n)$. Some details for this can be found in [19], and more of them in [95], and still more in [11]. A very explicit way to describe this identification is the map:

$$S_k(\Gamma_1(n)) \longrightarrow \mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}},\mathbb{C}),\quad f\mapsto (2\pi i)^{k-1} f\,d\tau\,dz_1\cdots dz_{k-2}, \tag{9.8}$$

where τ is the coordinate on $\mathbb{H}$, and the z_j are the coordinates on the copies of $\mathbb{C}$ using $\mathbb{E}_\tau = \mathbb{C}/(\mathbb{Z}\tau+\mathbb{Z})$. It is indeed easy to verify that the differential form on the right is invariant

under the actions of $\mathbb{Z}^{2(k-2)}$ and $\mathrm{SL}_2(\mathbb{Z})$, precisely because f is a modular form of weight k for $\Gamma_1(n)$. The claim (proved in the references above) is that it extends without poles over $\overline{\mathbb{E}^{k-2}}$. As it is a holomorphic form of top-degree, it is automatically closed, and hence defines a class in the de Rham cohomology of $\overline{\mathbb{E}^{k-2}}$, hence in $\mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}}, \mathbb{C})$.

There are natural Hecke correspondences on $\mathbb{C} \otimes \mathrm{H}^1(X_1(n), \mathcal{F}_k)$ and on $\mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}}, \mathbb{C})$, and the identification of $S_k(\Gamma_1(n))$ as a piece of these cohomology groups is compatible with these correspondences. Let now f be our eigenform in $S_k(\Gamma_1(n))$ as above. Then the Hecke eigenspace in $\mathbb{C} \otimes \mathrm{H}^1(X_1(n), \mathcal{F}_k)$ with the eigenvalues $a_m(f)$ for T_m is two-dimensional: the sum of the one-dimensional subspace $\mathbb{C}f$ in $S_k(\Gamma_1(n))$ and the one-dimensional subspace $\mathbb{C}\overline{f'}$ in $\overline{S_k(\Gamma_1(n))}$, where $f' = \sum_{m\geq 1} \overline{a_m(f)} q^m$, the Galois conjugate of f obtained by letting complex conjugation act on the coefficients of f. This element f' has eigenvalue $\overline{a_m(f)}$ for T_m, hence $\overline{f'}$ has eigenvalue $a_m(f)$ again. The $(k-1)$-form corresponding to $\overline{f'}$ is $\overline{f'} d\overline{\tau}\, d\overline{z_1} \cdots d\overline{z_{k-2}}$, indeed a form of type $(0, k-1)$.

We let $\mathbb{T}$ denote the $\mathbb{Z}$-algebra in $\mathrm{End}_{\mathbb{C}}(S_k(\Gamma_1(n)))$ generated by the T_m ($m \geq 1$) and the $\langle a \rangle$ (a in $(\mathbb{Z}/n\mathbb{Z})^\times$). The fact that the Hecke correspondences act on both sides of the Eichler-Shimura isomorphism (9.5) implies that the image of $\mathrm{H}^1(X_1(n), \mathcal{F}_k)$ in $\mathbb{C} \otimes \mathrm{H}^1(X_1(n), \mathcal{F}_k)$ is a faithful $\mathbb{T}$-module. As this image is free of finite rank as $\mathbb{Z}$-module, $\mathbb{T}$ is free of finite rank as $\mathbb{Z}$-module.

For A a subring of $\mathbb{C}$, we let $S_k(\Gamma_1(n), A)$ be the sub-A-module of $S_k(\Gamma_1(n))$ consisting of elements g such that $a_m(g) \in A$ for all $m \geq 1$. In particular, $S_k(\Gamma_1(n), \mathbb{Z})$ is the submodule of forms whose q-expansion has all its coefficients in $\mathbb{Z}$. For example, Δ belongs to $S_{12}(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Z})$. The $S_k(\Gamma_1(n), A)$ are $\mathbb{T}$-submodules of $S_k(\Gamma_1(n))$; see Propositions 12.3.11 and 12.4.1 of [28].

We have the following pairing between $\mathbb{T}$ and $S_k(\Gamma_1(n), \mathbb{Z})$:

$$S_k(\Gamma_1(n), \mathbb{Z}) \times \mathbb{T} \longrightarrow \mathbb{Z}, \quad (g, t) \mapsto a_1(tg). \tag{9.9}$$

This pairing is perfect, in the sense that it identifies each side with the $\mathbb{Z}$-linear dual of the other; this follows easily from the identity (8.12). It follows that the $\mathbb{Z}$-dual $S_k(\Gamma_1(n), \mathbb{Z})^\vee$ of $S_k(\Gamma_1(n), \mathbb{Z})$ is free of rank one as $\mathbb{T}$-module. See [28, 12.4.13]. For any $\mathbb{Z}$-algebra A we let $\mathbb{T}_A$ denote $A \otimes \mathbb{T}$, and $\mathbb{T}_A^\vee$ will denote the A-linear dual of $\mathbb{T}_A$. It can be proved that $\mathbb{T}_{\mathbb{Q}}^\vee$ is free of rank one as $\mathbb{T}_{\mathbb{Q}}$-module, i.e., that $\mathbb{T}_{\mathbb{Q}}$ is *Gorenstein*. One proof is by explicit computation, see Theorem 3.5 and Corollary 3.6 of [87]. Another, more conceptual proof, uses the Petersson inner product, and a so-called Atkin-Lehner pseudo-involution w_{ζ_n}, to show that $S_k(\Gamma_1(n))^\vee$ is isomorphic as $\mathbb{T}_{\mathbb{C}}$-module to $S_k(\Gamma_1(n))$ itself; see [28, 12.4.14]. It follows that $S_k(\Gamma_1(n))$ is free of rank one as $\mathbb{T}_{\mathbb{C}}$-module, and that $\mathbb{Q} \otimes \mathrm{H}^1(X_1(n), \mathcal{F}_k)$ and its dual $\mathbb{Q} \otimes \mathrm{H}^1(X_1(n), \mathcal{F}_k)^\vee$ are free of rank two as $\mathbb{T}_{\mathbb{Q}}$-module. It is this freeness result that will lead to the fact that the Galois representations we get are two-dimensional.

The step from the cohomological interpretation of modular forms, given, over the complex numbers, by the Eichler-Shimura isomorphism (9.5), to two-dimensional l-adic Galois representations is made by comparing the cohomology groups above to their l-adic counterparts for the étale topology, and noting that $p\colon \mathbb{E} \to X_1(n)$ is naturally defined over $\mathbb{Z}[1/n]$ as we have seen at the end of Section 7. From now on we will denote by $X_1(n)$ this model over $\mathbb{Z}[1/n]$, and by $X_1(n)(\mathbb{C})$ the Riemann surface given by $X_1(n)$. For any $\mathbb{Z}[1/n]$-algebra A, $X_1(n)_A$ will denote the A-scheme obtained from $X_1(n)$ by extending scalars via $\mathbb{Z}[1/n] \to A$.

We let $\mathcal{F}_{k,l}$ denote the sheaf of $\mathbb{Q}_l$-vector spaces $\mathbb{Q}_l \otimes \mathcal{F}_k$ on $X_1(n)$. Then we have a canonical isomorphism:

$$\mathrm{H}^1(X_1(n)(\mathbb{C}), \mathcal{F}_{k,l}) = \mathbb{Q}_l \otimes \mathrm{H}^1(X_1(n)(\mathbb{C}), \mathcal{F}_k). \tag{9.10}$$

The sheaves $\mathcal{F}_{k,l}$ can also be constructed on the étale site $X_1(n)_{\mathrm{et}}$, by taking the first derived direct image of the constant sheaf $\mathbb{Q}_l$ on $\mathbb{E}_{\mathrm{et}}$ under $p\colon \mathbb{E} \to Y_1(n)$, then the $(k-2)$th symmetric power of that and finally the pushforward from $Y_1(n)$ to $X_1(n)$.

The usual comparison theorems (comparing cohomology for étale and Archimedean topology, and étale cohomology over various algebraically closed fields) give:

$$\mathrm{H}^1(X_1(n)(\mathbb{C}), \mathcal{F}_{k,l}) = \mathrm{H}^1(X_1(n)_{\mathbb{C},\mathrm{et}}, \mathcal{F}_{k,l}) = \mathrm{H}^1(X_1(n)_{\overline{\mathbb{Q}},\mathrm{et}}, \mathcal{F}_{k,l}). \tag{9.11}$$

We put:

$$W_l := \mathrm{H}^1(X_1(n)_{\overline{\mathbb{Q}},\mathrm{et}}, \mathcal{F}_{k,l})^\vee. \tag{9.12}$$

By the results and the comparisons above, W_l is, as $\mathbb{T}_{\mathbb{Q}_l}$-module, free of rank 2, and $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts continuously on it. To be precise: an element σ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts as $((\mathrm{id} \times \mathrm{Spec}(\sigma^{-1}))^*)^\vee$, which is indeed covariant in σ. The fact that the Hecke correspondences exist over $\mathbb{Q}$ makes that the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-action on W_l commutes with the Hecke operators. The choice of a $\mathbb{T}_{\mathbb{Q}_l}$-basis of W_l gives us a representation:

$$\rho_l\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{T}_{\mathbb{Q}_l}). \tag{9.13}$$

Recall that we have fixed an eigenform f in $S(\Gamma_1(n), \mathbb{C})$. Sending a Hecke operator to its eigenvalue for f then gives us a morphism of rings:

$$\phi_f\colon \mathbb{T} \longrightarrow \mathbb{C}. \tag{9.14}$$

We let $K(f)$ be the image of $\mathbb{T}_{\mathbb{Q}}$ under ϕ_f; it is the finite extension of $\mathbb{Q}$ obtained by adjoining all coefficients $a_m(f)$ of the q-expansion of f. We now view ϕ_f as a morphism from $\mathbb{T}$ to $K(f)$. The

tensor product $\mathbb{Q}_l \otimes K(f)$ is the product of the completions $K(f)_\lambda$, with λ ranging through the finite places of $K(f)$ that divide l. For each such λ we then get a morphism $\phi_{f,\lambda}\colon \mathbb{T}_{\mathbb{Q}_l} \to K(f)_\lambda$, and a representation:

$$\rho_{f,\lambda}\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(K(f)_\lambda). \tag{9.15}$$

These are the representations mentioned in Theorem 9.1. It may be useful to note that the space on which the representation is realised is:

$$V_{f,\lambda} := K(f)_\lambda \otimes_{\mathbb{T}_{\mathbb{Q}_l}} W_l. \tag{9.16}$$

The representations $\rho_{f,\lambda}$ are continuous by construction. The sheaves $\mathcal{F}_{k,l}$ on $X_1(n)_{\mathbb{Z}[1/nl]}$ are "lisse" away from the cusps, and tamely ramified at the cusps, hence, by Proposition 2.1.9 of [26, XIII, §2], $\rho_{f,\lambda}$ is unramified at all p not dividing nl.

In the case where $k = 2$ the construction of the $\rho_{f,\lambda}$ is much simpler, because then the sheaf $\mathcal{F}_k$ is the constant sheaf $\mathbb{Z}$ on $X_1(n)(\mathbb{C})$. The use of étale cohomology can then be replaced by Tate modules of the Jacobian variety of $X_1(n)$. We let $J := J_1(n)$ be this Jacobian variety, actually an Abelian scheme over $\mathbb{Z}[1/n]$. Then we have:

$$W_l = \mathbb{Q} \otimes \varprojlim_m J(\overline{\mathbb{Q}})[l^m]. \tag{9.17}$$

The fact that for p a prime not dividing nl the characteristic polynomial of $\rho_{f,\lambda}(\mathrm{Frob}_p)$ is as stated in Theorem 9.1 is obtained by studying the reduction modulo p of the Hecke correspondence T_p, i.e., as a correspondence on $X_1(nl)_{\mathbb{F}_p}$, compatibly with the sheaf $\mathcal{F}_{k,l}$. For details we refer to Conrad's book [11] and Deligne's article [19]. In the case $k = 2$ this result is known as the *Eichler-Shimura congruence relation*, expressing the endomorphism T_p of $J_{\mathbb{F}_p}$ as $F + \langle p\rangle V$, where F denotes the Frobenius endomorphism, and V its dual, i.e, the endomorphism satisfying $FV = p = VF$ in $\mathrm{End}(J_{\mathbb{F}_p})$. For details in the case $k = 2$ we refer to Section 12.5 of [28].

Now that we have sketched the construction of the l-adic Galois representations associated to modular forms, we mention some more of their properties, that are not mentioned in Theorem 9.1 and in the remarks directly following that theorem.

The fact that Deligne proved the Riemann hypothesis part of Weil's conjectures in [20] implies very precise bounds on the coefficients of modular forms. The reason for that is that the roots of the equation $x^2 - a_p(f)x + \varepsilon_f(p)p^{k-1}$ are eigenvalues of the Frobenius at p on the space $\mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}}_{\overline{\mathbb{F}}_p,\mathrm{et}}, \mathbb{Q}_l)$. We state these bounds, called *Ramanujan bounds*, in a theorem, due to Deligne in the case $k \geq 2$, and to Deligne-Serre ([22]) in the case $k = 1$.

9.18 Theorem *Let f be a normalised newform, let n be its level and k its weight. Then for p not dividing n, we have:*

$$|a_p(f)| \leq 2{\cdot}p^{(k-1)/2}. \tag{9.19}$$

A slightly weaker result than in the theorem above, stating that, for a given f as above, $|a_m(f)| = O(m^{k/2})$, can be obtained in a very elementary way; see [82, Cor. 2.1.6] (the idea is to use that the function $z \mapsto |f(z)|(\Im z)^k$ is bounded on $\mathbb{H}$ and to view $a_m(f)$ as a residue).

Theorem 9.1 gives us information on the restriction $\rho_{f,\lambda,p}$ of $\rho_{f,\lambda}$ to decomposition groups $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ for p not dividing nl. Namely, the theorem says that such restrictions $\rho_{f,\lambda,p}$ are unramified, and it gives the eigenvalues of $\rho_{f,\lambda,p}(\mathrm{Frob}_p)$. Unfortunately, it is not known if $\rho_{f,\lambda,p}(\mathrm{Frob}_p)$ is semi-simple; see [10] for information on this.

We should note that also in the case that p divides nl almost everything is known about $\rho_{f,\lambda,p}$. For p not dividing l, this is the very general statement that the "Frobenius semi-simplification" of $\rho_{f,\lambda,p}$ corresponds, via a suitably normalised local Langlands correspondence, to a certain representation $\pi_{f,p}$ of $\mathrm{GL}_2(\mathbb{Q}_p)$ associated to f. This result is due, in increasing order of generality, to Langlands, Deligne, and Carayol. For details on this the reader is referred to [8], which gives this result in the more general context of Hilbert modular forms (i.e., $\mathbb{Q}$ is replaced by a totally real number field). The result for $p = l$ uses Fontaine's p-adic Hodge theory, and is due to Saito ([92] for the case of modular forms, and [93] for the case of Hilbert modular forms).

10 Galois representations over finite fields, and reduction to torsion in Jacobians

We start this section by explaining how to pass from l-adic Galois representations to Galois representations over finite fields.

Let $f = \sum a_m q^m$ be a (complex) normalised cuspidal eigenform for all Hecke operators T_m, $m \geq 1$, of some level $n \geq 1$ and of some weight $k \geq 2$. As in Theorem 9.1 we have the Galois representations $\rho_{f,\lambda}$, from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathrm{GL}_2(\overline{\mathbb{Q}}_l)$. It follows from the construction of those representations that there is a finite subextension $\mathbb{Q}_l \to E$ of $\mathbb{Q}_l \to \overline{\mathbb{Q}}_l$ such that $\rho_{f,\lambda}$ takes its values in $\mathrm{GL}_2(E)$. (Actually, this can also be deduced from the continuity alone; see the proof of Corollary 5 in [23] for an argument.) The question as to what the smallest possible E is can be easily answered. Such an E must contain the traces $a_p(f)$ of the $\rho_{f,\lambda}(\mathrm{Frob}_p)$ for all p not dividing nl. So let K be the extension of $\mathbb{Q}$ generated by the $a_p(f)$ with p not dividing n, i.e., K is the field of definition of the newform corresponding to f. Then E can be taken to be K_λ, the l-adic completion of K specified by the embedding λ of K into $\overline{\mathbb{Q}}_l$ (see Section 12.5 in [28]).

Let now $\rho_{f,\lambda}\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(E)$ be a realisation of $\rho_{f,\lambda}$ over E as above. As $\rho_{f,\lambda}$ is semisimple (it is even irreducible), such a realisation is unique up to isomorphism (because it is determined by the traces). Let O_E be the ring of integers in E, i.e., the integral closure of $\mathbb{Z}_l$ in E. As $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is compact, it stabilises some O_E lattice in E^2 (in the set of lattices,

the orbits under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ are finite, take the intersection of the lattices in one orbit). This means that, after suitable conjugation (choose such a lattice, and an O_E-basis of it), $\rho_{f,\lambda}$ takes values in $\mathrm{GL}_2(O_E)$. We let $O_E \to \overline{\mathbb{F}}_l$ denote the morphism induced by the given embedding of E into $\overline{\mathbb{Q}}_l$ (we view $\overline{\mathbb{F}}_l$ as the residue field of the subring of integers $\overline{\mathbb{Z}}_l$ of $\overline{\mathbb{Q}}_l$). We can then define the *residual* Galois representation $\overline{\rho}_{f,\lambda}$ to be the *semi-simplification* of the composed representation $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(O_E) \to \mathrm{GL}_2(\overline{\mathbb{F}}_l)$. Another choice of E or of lattice or basis leads to an isomorphic $\overline{\rho}_{f,\lambda}$, but we note that without the operation of semi-simplification this would not be true (see Chapter III of [100]).

Given f, all but finitely many of the $\overline{\rho}_{f,\lambda}$ are irreducible. This was proved for f of level one and with coefficients in $\mathbb{Z}$ in Theorem 4 of [106]. The general case follows easily from Theorem 2.3 of [41], which says that if $\overline{\rho}_{f,\lambda}$ is reducible with $l > k$ not dividing n, then $\overline{\rho}_{f,\lambda} = \alpha \oplus \beta\overline{\chi}_l^{k-1}$ with α and β unramified outside n, and $\overline{\chi}_l \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{F}_l^\times$ the mod l cyclotomic character. Moreover, the proof shows that the set of l such that some $\overline{\rho}_{f,\lambda}$ is reducible can be bounded explicitly.

The next question that we want to answer is the following: what is the smallest subfield of $\overline{\mathbb{F}}_l$ over which $\overline{\rho}_{f,\lambda}$ can be realised? Just as for $\rho_{f,\lambda}$ itself, that subfield must contain the traces of the $\overline{\rho}_{f,\lambda}(\mathrm{Frob}_p)$ for all p not dividing nl. That condition turns out to be sufficient, as we will now show. So we let $\mathbb{F}$ be the subfield of $\overline{\mathbb{F}}_l$ that is generated by the images $\overline{a_p(f)}$ in $\overline{\mathbb{F}}_l$ of the $a_p(f)$ in $\overline{\mathbb{Z}}_l$. Then for any σ in $\mathrm{Gal}(\overline{\mathbb{F}}_l/\mathbb{F})$ the conjugate $\overline{\rho}_{f,\lambda}^\sigma$ of $\overline{\rho}_{f,\lambda}$ and $\overline{\rho}_{f,\lambda}$ itself are both semisimple and give the same characteristic polynomials as functions on $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Therefore, by a theorem of Brauer-Nesbitt (see Theorem 30.16 of [16]), $\overline{\rho}_{f,\lambda}$ is isomorphic to all its conjugates over $\mathbb{F}$. (A more general statement of this kind is given in Exercise 1 of Section 18.2 of [100].) The fact that $\mathrm{Gal}(\overline{\mathbb{F}}_l/\mathbb{F})$ is equal to $\hat{\mathbb{Z}}$ then implies that $\overline{\rho}_{f,\lambda}$ can be realised over $\mathbb{F}$. Let us give an argument for that in terms of matrices, although a much more conceptual argument would be to say that a "gerbe over a finite field is trivial". Let σ be the Frobenius element of $\mathrm{Gal}(\overline{\mathbb{F}}_l/\mathbb{F})$, and let s be an element of $\mathrm{GL}_2(\overline{\mathbb{F}}_l)$ such that for all g in the image of $\overline{\rho}_{f,\lambda}$ we have $\sigma(g) = sgs^{-1}$. Then take a t in $\mathrm{GL}_2(\overline{\mathbb{F}}_l)$ such that $s = \sigma(t)^{-1}t$. Then all tgt^{-1} are in $\mathrm{GL}_2(\mathbb{F})$. By Brauer-Nesbitt, the realisation over $\mathbb{F}$ is unique.

For a discussion on possible images of $\overline{\rho}_{f,\lambda}$ we refer the reader to the introduction of [63] (we note however that for f a "CM-form" infinitely many of the $\overline{\rho}_{f,\lambda}$ can have dihedral image in $\mathrm{PGL}_2(\overline{\mathbb{F}}_l)$). In particular, Theorem 2.1 of [90] states that for f not a CM-form only finitely many of the images of the $\overline{\rho}_{f,\lambda}$ are exceptional in the sense that they are of order prime to l. For $l > 3$ such that $\overline{\rho}_{f,\lambda}$ is irreducible and not exceptional, a result of Dickson, see [24], or [91], says that the image of $\overline{\rho}_{f,\lambda}$ is, after suitable conjugation, equal to $\mathrm{PGL}_2(\mathbb{F})$ or $\mathrm{SL}_2(\mathbb{F})/\{1,-1\}$ for some finite extension $\mathbb{F}$ of $\mathbb{F}_l$. We note that this field $\mathbb{F}$ be smaller that the extension $\mathbb{F}_l$ generated by the traces of $\overline{\rho}_{f,\lambda}$ (indeed, twisting does not change the projective image, but it can make the field

generated by the traces bigger).

We will use the easily proved statement that if $\overline{\rho}_{f,\lambda}$ takes values in $\mathrm{GL}_2(\mathbb{F}_l)$, and is irreducible and not exceptional, then $\mathrm{im}\overline{\rho}_{f,\lambda}$ contains $\mathrm{SL}_2(\mathbb{F}_l)$, and therefore is the subgroup of elements of $\mathrm{GL}_2(\mathbb{F}_l)$ whose determinant is in the image of the character $\overline{\varepsilon_f\chi_l}^{k-1}$.

Let us suppose now that $\overline{\rho}_{f,\lambda}$ is irreducible. The construction of $\rho_{f,\lambda}$ that we recalled in Section 9 implies that the dual of $\overline{\rho}_{f,\lambda}$ occurs in $\mathrm{H}^{k-1}(\overline{\mathbb{E}^{k-2}}_{\overline{\mathbb{Q}},\mathrm{et}},\mathbb{F}_l)$, as well as in $\mathrm{H}^1(X_1(n)_{\overline{\mathbb{Q}},\mathrm{et}},\overline{\mathcal{F}}_{k,l})$, where $\overline{\mathcal{F}}_{k,l}$ is defined as $\mathcal{F}_{k,l}$ but with $\mathbb{Q}_l$ replaced by $\mathbb{F}_l$. Let us now assume that $k>2$. Then both these realisations are difficult to deal with computationally. In the first representation the difficulty arises from the degree $k-1$ étale cohomology; it seems to be unknown how to deal explicitly with elements of such cohomology groups. In the second representation, the elements of the cohomology group are isomorphism classes of $\overline{\mathcal{F}}_{k,l}$-torsors, on $X_1(n)_{\overline{\mathbb{Q}},\mathrm{et}}$. Such torsors can be described explicitly, as certain covers of $X_1(n)_{\overline{\mathbb{Q}}}$ with certain extra data. The set of such torsors can probably be described by a system of polynomial equations that can be written down in time polynomial in nl (think of the variables as coefficients of certain equations for the torsors). But the problem is that, apparently, there are no good methods known to solve these systems of equations (the number of variables grows too fast with l and the equations are not linear). In fact, the *satisfiability problem* SAT, which is known to be NP-complete (Cook's theorem, see for example [83], or the wikipedia), is a special case of the problem of deciding whether or not a polynomial system of equations over $\mathbb{F}_2$ has a solution over $\mathbb{F}_2$. We note that the description of the set of torsors by system of polynomial equations should also work over suitable finite extensions of finite fields $\mathbb{F}_p$, in time polynomial in $l\log p$.

Another place where $\overline{\rho}_{f,\lambda}$ occurs is in $J_1(nl)(\overline{\mathbb{Q}})[l]$, i.e., in the l-torsion of the Jacobian of the modular curve with level nl, if $l+1\geq k$ and $l\not|n$. This means that at the cost of increasing the level by a factor l, we are reduced to dealing with torsion points on Abelian varieties. Of course, the l-adic representation $\rho_{f,\lambda}$ does not occur in the Jacobian of any curve, simply because the Frobenius eigenvalues are Weil numbers of the wrong weight. What happens here for $\overline{\rho}_{f,\lambda}$ is a "mod l phenomenon" having to do with "congruences" between modular forms. Before we give a detailed statement, let us explain why this happens (such explanations date back to the 1960's; Shimura, Igusa, Serre,...).

For simplicity, and only during this explanation, we assume that $n\geq 5$. Then we have a universal elliptic curve with a given point of order n over $\mathbb{Z}[1/nl]$-schemes: $(\mathbb{E}/Y_1(n),\mathbb{P})$. We let $p\colon\mathbb{E}\to Y_1(n)$ denote the structure morphism. By definition, we have:

$$\overline{\mathcal{F}}_{k,l}=\mathrm{Sym}^{k-2}\mathrm{R}^1p_*\mathbb{F}_l. \tag{10.1}$$

As explained at the end of Section 3, we have a natural isomorphism:

$$\mathrm{R}^1p_*\mathbb{F}_l=\mathbb{E}[l]^\vee. \tag{10.2}$$

And by the definition of $Y_1(nl)$, and the Weil pairing, we have an exact sequence on $Y_1(nl)_{\mathrm{et}}$:

$$0 \longrightarrow \mathbb{F}_l \longrightarrow \mathbb{E}[l] \longrightarrow \mu_l \longrightarrow 0, \tag{10.3}$$

where $\mathbb{F}_l$ and μ_l denote the corresponding constant sheaves. It follows that the pullback of $\mathrm{R}^1 p_* \mathbb{F}_l$ to $Y_1(nl)_{\mathrm{et}}$ has a 2-step filtration with successive quotients $\mathbb{F}_l$ and $\mu_l^\vee$. Therefore, $\overline{\mathcal{F}}_{k,l}$ has a filtration in $k-1$ steps, with successive quotients $\mathbb{F}_l^{\otimes i} \otimes (\mu_l^\vee)^{\otimes j} = \mu_l^{\otimes -j}$, with $i+j = k-2$, $i \geq 0$, $j \geq 0$. In particular, we get a map:

(10.4)

$$\mathrm{H}^1(X_1(n)_{\overline{\mathbb{Q}},\mathrm{et}}, \overline{\mathcal{F}}_{k,l}) \longrightarrow \mathrm{H}^1(X_1(nl)_{\overline{\mathbb{Q}},\mathrm{et}}, \overline{\mathcal{F}}_{k,l}) \longrightarrow \mathrm{H}^1(X_1(nl)_{\overline{\mathbb{Q}},\mathrm{et}}, \mathbb{F}_l) = J_1(nl)(\overline{\mathbb{Q}})[l]^\vee.$$

This map explains that our representation $\overline{\rho}_{f,\lambda}$ is likely to occur in $J_1(nl)(\overline{\mathbb{Q}})[l]$. A better way to analyse this map is in fact by studying the direct image of the constant sheaf $\mathbb{F}_l$ via the map $X_1(nl) \to X_1(n)$. A recent detailed treatment of this method, and precise results can be found in [112].

Another way to show that $\overline{\rho}_{f,\lambda}$ occurs in $J_1(nl)(\overline{\mathbb{Q}})[l]$ is to study modular forms mod l of level nl and of weight 2. This is more complicated than the modular forms that we have seen before, as it uses the study of the reduction mod l of the modular curve $X_1(nl)$, which is not smooth. The study of these reductions has its roots in Kronecker's congruence relation. The most complete modern accounts of such material are given in the article [21] by Deligne and Rapoport and in the book [61] by Katz and Mazur. A construction of $\overline{\rho}_{f,\lambda}$ in $J_1(nl)(\overline{\mathbb{Q}})[l]$, following suggestions from Serre, was given by Gross in [50].

We are now in a position to state the following theorem, that, combining Gross's result with a so-called multiplicity one theorem, gives us a useful realisation of $\overline{\rho}_{f,\lambda}$. The theorem is case 1 of Theorem 9.2 of [32], to which we refer for the proof. As it is nowadays customary to say, it is a result due to "many people" (mainly Mazur, Ribet, Gross (and Edixhoven for the multiplicity one part)).

10.5 Theorem *Let $f = \sum_{m \geq 1} a_m q^m$ be a complex normalised cuspidal eigenform of weight $k > 2$ and level n. Let $l > k$ be a prime number and let λ be an embedding of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_l$. Assume that the residual representation $\overline{\rho}_{f,\lambda}$ from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathrm{GL}_2(\overline{\mathbb{F}}_l)$ is irreducible. Let $\mathbb{T}$ be the Hecke algebra of $J_1(nl)$, i.e., the subring of $\mathrm{End}(J_1(nl))$ generated by all T_m with $m \geq 1$, and all $\langle a \rangle$ with a in $(\mathbb{Z}/nl\mathbb{Z})^\times$. Then there is a unique morphism of rings $\mathbb{T} \to \overline{\mathbb{F}}_l$ such that for all $m \geq 1$ the operator T_m is mapped to $\overline{a_m}$. The image of $\mathbb{T}$ is the subextension $\mathbb{F}$ of $\overline{\mathbb{F}}_l$ generated by the $\overline{a_m}$. Let $m_{f,\lambda}$ be the kernel of $\mathbb{T} \to \overline{\mathbb{F}}_l$, and let $V_{f,\lambda} \subset J_1(nl)(\overline{\mathbb{Q}})$ denote the kernel of $m_{f,\lambda}$, i.e., the $\mathbb{F}$-vector space of elements x in $J_1(nl)(\overline{\mathbb{Q}})$ such that $tx = 0$ for all t in $m_{f,\lambda}$. Then $V_{f,\lambda}$ is 2-dimensional, and realises $\overline{\rho}_{f,\lambda}$ over $\mathbb{F}$. Under the morphism $\mathbb{T} \to \overline{\mathbb{F}}_l$, $\langle a \rangle$*

is sent to $\overline{\varepsilon}(a)a^{k-2}$, *where* $\overline{\varepsilon}\colon (\mathbb{Z}/n\mathbb{Z})^\times \to \overline{\mathbb{F}}_l^\times$ *is the character* ε *from* $(\mathbb{Z}/n\mathbb{Z})^\times$ *to* $\overline{\mathbb{Z}}^\times$ *associated to* f, *composed with the reduction map* $\overline{\mathbb{Z}} \to \overline{\mathbb{F}}_l$, *and where* a^{k-2} *denotes the* $k-2$*nd power of the image of* a *under* $\mathbb{Z}/nl\mathbb{Z} \to \mathbb{F}_l$.

As we want to describe $V_{f,\lambda}$ explicitly, we will need a bound on the amount of Hecke operators needed to describe $\mathbb{T}$ and its ideal $m_{f,\lambda}$. We start by quoting a result of Jacob Sturm (see [105]).

10.6 Theorem (Sturm) *Let* $N \geq 1$ *be an integer,* Γ *a subgroup of* $\mathrm{SL}_2(\mathbb{Z})$ *containing* $\Gamma(N)$, f *and* g *modular forms on* Γ *of weight* k, *with coefficients in a discrete valuation ring* R *contained in* $\mathbb{C}$. *Let* F *be the residue field of* R, *and suppose that the image* $\sum a_n q^n$ *in* $F[[q^{1/N}]]$ *of the* q*-expansion of* $f-g$ *has* $a_n = 0$ *for all* $n \leq k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]/12$. *Then* $a_n = 0$ *for all* n, *i.e.,* f *and* g *are congruent modulo the maximal ideal of* R.

This result of Sturm gives as a direct consequence a bound for up to where one has to take T_i so that one gets a system of generators of the Hecke algebra as $\mathbb{Z}$-module, for a given level and weight. This has been written down for the Hecke algebras associated to modular forms for the groups $\Gamma_0(N)$ by Agashe and Stein, in an appendix to [67]; see also William Stein's home page. We want to use it in the case of the groups $\Gamma_1(N)$, and therefore give the result here.

10.7 Theorem *Let* $N \geq 1$ *and* $k \geq 1$ *be integers, and let* $\mathbb{T}$ *be the Hecke algebra associated to* $S_k(\Gamma_1(N))$, *i.e.,* $\mathbb{T}$ *is the* $\mathbb{Z}$*-submodule of* $\mathrm{End}_{\mathbb{C}}(S_k(\Gamma_1(N)))$ *generated by all* T_n, $n \geq 1$, *and all* $\langle a\rangle$, a *in* $(\mathbb{Z}/N\mathbb{Z})^\times$. *Then* $\mathbb{T}$ *is generated, as* $\mathbb{Z}$*-module, by the* T_i *with* $1 \leq i \leq k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)]/12$.

Proof The argument is the same as the one used by Agashe and Stein. Let $S := S_k(\Gamma_1(N), \mathbb{Z})$. We have: $S = \mathbb{T}^\vee$, and $\mathbb{T} = S^\vee$ (see 9.9). Then Sturm's result says that for each prime number p, the elements T_i, $1 \leq i \leq k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(n)]/12$, generate $\mathbb{F}_p \otimes S^\vee$, and hence they generate $\mathbb{F}_p \otimes \mathbb{T}$. So, indeed, these T_i generate $\mathbb{T}$ as a $\mathbb{Z}$-module. □

We can now state a complement to Theorem 10.5.

10.8 Proposition *In the situation of Theorem 10.5, the Hecke algebra* $\mathbb{T}$ *is generated, as* $\mathbb{Z}$*-module, by the* T_i *with* $1 \leq i \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(nl)]/6$.

We remark that, still in the same situation, giving generators of $m_{f,\lambda}$ is then a matter of simple linear algebra over $\mathbb{F}_l$ in a vector space of suitably bounded dimension. In the particular case of the modular form Δ we have proved the following result (we state it here because it gives the subspace $V_{\Delta,l}$ so explicitly).

10.9 Theorem *Let $l > 12$ be prime with $l \notin \{23, 691\}$. Let:*

$$V_{\Delta,l} := \bigcap_{1 \le i \le (l^2-1)/6} \ker(T_i - \tau(i), J_1(l)(\overline{\mathbb{Q}})[l]).$$

Then $V_{\Delta,l}$ is a 2-dimensional $\mathbb{F}_l$-vector space realising the irreducible Galois representation $\overline{\rho}_{\Delta,l}$. For $p \neq l$ prime, $\langle p \rangle$ acts on $V_{\Delta,l}$ as multiplication by p^{10}.

We also state the following definition and theorem here, because the result, to be used later, is directly related to Theorem 10.5. The theorem is due, again, to "many people", just as Theorem 10.5 itself.

10.10 Definition Let $N \geq 1$, and let $\mathbb{Z}[\zeta_N]$ be the subring of $\mathbb{C}$ generated by a root of unity of order N. To a pair $(E/S/\mathbb{Z}[1/N, \zeta_N], P)$ consisting of an elliptic curve E over a $\mathbb{Z}[1/N, \zeta_N]$-scheme S, together with a point P in $E(S)$ that is of order N everywhere on S, we associate another such pair $(E'/S'/\mathbb{Z}[1/N, \zeta_N], P')$ as follows. Let $\beta\colon E \to E'$ be the isogeny whose kernel is the subgroup of E generated by P. Let $\beta^\vee\colon E' \to E$ be the dual of β (see Section 2.5 of [61]). Let P' be the unique element of $\ker(\beta^\vee)(S)$ such that $e_\beta(P, P') = \zeta_N$, where e_β is the perfect μ_N-valued pairing between $\ker(\beta)$ and $\ker(\beta)$ as described in Section 2.8 of [61]. This construction induces an automorphism of the modular curve $X_1(N)_{\mathbb{Z}[1/N,\zeta_N]}$, called an "Atkin-Lehner pseudo-involution".

10.11 Theorem *In the situation of Theorem 10.5 the completion $\mathbb{T}_{m_{f,\lambda}}$ of $\mathbb{T}$ at $m_{f,\lambda}$ is* Gorenstein, *i.e., the $\mathbb{Z}_l$-linear dual of $\mathbb{T}_{m_{f,\lambda}}$ is free of rank one as $\mathbb{T}_{m_{f,\lambda}}$-module. For all $r \geq 1$, the $(\mathbb{Z}/l^r\mathbb{Z}) \otimes \mathbb{T}_{m_{f,\lambda}}$-module $J_1(nl)(\overline{\mathbb{Q}})[l^r]_{m_{f,\lambda}}$ is free of rank 2.*

For any t in $\mathbb{T}$ we have $t^\vee = wtw^{-1}$, where $t^\vee$ is the dual of t as endomorphism of the self-dual Abelian variety $J_1(nl)_{\mathbb{Q}(\zeta_{nl})}$, and where w is the endomorphism of $J_1(nl)_{\mathbb{Q}(\zeta_{nl})}$ induced via Picard functoriality by the automorphism $w_{\zeta_{nl}}$ of $X_1(nl)_{\mathbb{Q}(\zeta_{nl})}$.

For $r \geq 0$, let $(\cdot,\cdot)_r$ denote the Weil pairing on $J_1(nl)(\overline{\mathbb{Q}})[l^r]$, and let $\langle \cdot, \cdot \rangle_r$ denote the pairing defined by:

$$\langle x, y \rangle_r = (x, w(y))_r.$$

Then $\langle \cdot, \cdot \rangle_r$ is a perfect pairing on $J_1(nl)(\overline{\mathbb{Q}})[l^r]$ for which the action of $\mathbb{T}$ is self-adjoint. In particular, $\langle \cdot, \cdot \rangle_r$ induces a perfect pairing on $J_1(nl)(\overline{\mathbb{Q}})[l^r]_{m_{f,\lambda}}$.

Proof See Sections 6.4 and 6.8 of [32]. □

11 Strategy for computing residual Galois representations

Let $f = \sum_{m\geq 1} a_m q^m$ be a complex normalised cuspidal newform of weight $k > 2$ and level n. (We could also treat the easier case of weight 2, by the same method, but that would lead to more notation.) We let $E \subset \overline{\mathbb{Q}}$ be the subfield generated by all a_m, $m \geq 1$. For each finite place λ of E we let $\mathbb{F}_\lambda$ denote the finite extension of $\mathbb{F}_l$ (where l is the prime of $\mathbb{Q}$ "under λ") that is generated by the reductions $\overline{a_m}$, $m \geq 1$. We then have the residual Galois representations as explained in the previous section:

$$\overline{\rho}_\lambda \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{F}_\lambda)$$

We assume in this section, from now on, that f is not a CM-form, and we exclude the finitely many exceptional λ, i.e., the λ such that $\overline{\rho}_\lambda$ is reducible or such that the image $\mathrm{im}\overline{\rho}_\lambda$ of $\overline{\rho}_\lambda$ is of order prime to l. In particular, if $\mathbb{F}_\lambda = \mathbb{F}_l$ then $\mathrm{im}\overline{\rho}_\lambda$ contains $\mathrm{SL}_2(\mathbb{F}_l)$.

We let $K_\lambda \subset \overline{\mathbb{Q}}$ be the field "cut out by $\overline{\rho}_\lambda$", i.e., the finite Galois extension of $\mathbb{Q}$ contained in $\overline{\mathbb{Q}}$ consisting of the elements of $\overline{\mathbb{Q}}$ that are fixed by all elements in $\ker(\overline{\rho}_\lambda)$. Then we have, by definition, the following factorisation of $\overline{\rho}_\lambda$:

$$\overline{\rho}_\lambda \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(K_\lambda/\mathbb{Q}) \hookrightarrow \mathrm{GL}_2(\mathbb{F}_\lambda).$$

Our aim is then to compute these residual representations $\overline{\rho}_\lambda$, in time polynomial in $\#\mathbb{F}_\lambda$. We recall that f is now fixed and that λ varies. By computing $\overline{\rho}_\lambda$ we mean giving a monic polynomial in $\mathbb{Q}[T]$ that is the minimal polynomial of some generator t of K_λ, and giving the elements σ of $\mathrm{Gal}(K_\lambda/\mathbb{Q})$ by giving their matrices with respect to the $\mathbb{Q}$-basis of K_λ consisting of the first so many powers of t, together with the element $\overline{\rho}_\lambda(\sigma)$ of $\mathrm{GL}_2(\mathbb{F}_\lambda)$.

Theorem 10.5 tells us that $\overline{\rho}_\lambda$ is realised by the 2-dimensional $\mathbb{F}_\lambda$-vector space V_λ contained in $J_1(nl)(\overline{\mathbb{Q}})[l]$ as the set of elements annihilated by all elements of the kernel m_λ of the morphism $\mathbb{T} \twoheadrightarrow \mathbb{F}_\lambda$ that sends T_m to $\overline{a_m}$. Therefore, the field K_λ is the compositum of the fields of definition of the elements of V_λ. The crucial step in computing $\overline{\rho}_\lambda$ is to compute the extension K_λ by giving the minimal polynomial of some suitable generator. Instead of giving such a polynomial for the whole Galois extension K_λ, we can also give one for a suitable subextension, from which K_λ can then be computed as the splitting field. In order to describe such a subfield it is good to use the modern version of Galois theory that says that the functor $A \mapsto \mathrm{Hom}_{\mathbb{Q}}(A, \overline{\mathbb{Q}})$ is an anti-equivalence from the category of finite separable $\mathbb{Q}$-algebras to that of finite discrete (continuous) $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-sets. An inverse is given by the map that sends X to $\mathrm{Hom}_{\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})}(X, \overline{\mathbb{Q}})$, the $\mathbb{Q}$-algebra of functions f from X to $\overline{\mathbb{Q}}$ such that $f(gx) = g(f(x))$ for all g in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and all x in X. Under this correspondence, fields correspond to transitive $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-sets, and the field K_λ corresponds to an orbit in the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-set $\mathrm{Isom}(\mathbb{F}_\lambda^2, V_\lambda)$ of $\mathbb{F}_\lambda$-bases of V_λ.

Instead of looking at such a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-set, we look at $V_\lambda - \{0\}$ and let K'_λ denote the $\mathbb{Q}$-algebra that corresponds to it; it is the product of the fields that correspond to the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-orbits in $V_\lambda - \{0\}$. As $K'_\lambda \otimes K'_\lambda$ corresponds to $(V_\lambda - \{0\})^2$, which contains the set of $\mathbb{F}_\lambda$-bases of V_λ, K_λ is one of the factors of the decomposition of $K'_\lambda \otimes K'_\lambda$ as a product of fields. If the minimal polynomial of a generator of K'_λ is known, then we get the field K_λ using factoring algorithms. See [72], [73] and [64] for the fact that such factoring can be done in polynomial time. We note that if $\mathbb{F}_\lambda = \mathbb{F}_l$ then $V_\lambda - \{0\}$ is a transitive $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-set, and hence K'_λ is a field.

We wish to produce a generator of K'_λ, and its minimal polynomial over $\mathbb{Q}$. This means that we must produce a $\overline{\mathbb{Q}}$-valued function k on $V_\lambda - \{0\}$ such that $k(gx) = g(k(x))$ for all x in $V_\lambda - \{0\}$ and all g in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Such a function is a generator of K'_λ if and only if it does not arise from a strictly smaller quotient of $V_\lambda - \{0\}$ (such quotients correspond to subalgebras), hence, equivalently, if and only if k is injective. The minimal polynomial over $\mathbb{Q}$ of such a generator k is given as follows:

$$P_\lambda(T) = \prod_{x \in V_\lambda - \{0\}} (T - k(x)). \tag{11.1}$$

The question is now how we are going to produce such a generator. A direct way would be to compute the elements of V_λ in $J_1(nl)(\overline{\mathbb{Q}})$, by writing down polynomial equations in a suitable coordinate system that is defined over $\mathbb{Q}$, and solving them, using computer algebra. This is essentially how Schoof's algorithm deals with elliptic curves. However, the dimension of $J_1(nl)$ is quadratic in l. Writing down equations in polynomial time still seems possible. But we do not know of a way of solving the equations in a time that is not exponential in the dimension.

It is at this point that *Jean-Marc Couveignes* suggested a decisive idea: to use numerical computations to approximate the coefficients of a minimal polynomial P_λ as above, in combination with a bound on the *height* of those coefficients. We recall that the (standard, logarithmic) height of a rational number a/b, with a and b integers that are relatively prime, is $\log \max\{|a|, |b|\}$ (a variant would be $\log(a^2 + b^2)$). This rational number $x = a/b$ is known if we know an upper bound h for its height, and an approximation y of it (in $\mathbb{R}$, say), with $|x - y| < e^{-2h}/2$. Indeed, if $x' = a'/b'$ also has height at most h, and $x' \neq x$, then:

$$|x - x'| = \left|\frac{a}{b} - \frac{a'}{b'}\right| = \left|\frac{ab' - ba'}{bb'}\right| \geq \frac{1}{|bb'|} \geq e^{-2h}. \tag{11.2}$$

We also note that there are good algorithms to deduce x from such a pair of an approximation y and a bound h (continued fractions).

The question is now: how we are going to implement this suggestion of Couveignes? The basic idea in doing this is to not work on the Abelian variety $J_1(nl)$ but rather on the product

$X_1(nl)^{g_l}$ of copies of $X_1(nl)$, where g_l is the genus of $X_1(nl)$. To compare the two, we first choose an effective divisor $D = P_1 + \cdots + P_{g_l}$ on $X_1(nl)_{\mathbb{Q}}$, and we consider the well-known map:

$$X_1(nl)^{g_l} \longrightarrow J_1(nl), \quad (Q_1, \ldots, Q_{g_l}) \mapsto [Q_1 + \cdots + Q_{g_l} - D]. \tag{11.3}$$

To understand the definition of this map rigorously, one must use the interpretation of $X_1(nl)$ as its functor of points with values in $\mathbb{Z}[1/nl]$-schemes, and that of $J_1(nl)$ as the degree zero part of the relative Picard functor $\mathrm{Pic}^0_{X_1(nl)/\mathbb{Z}[1/nl]}$. For the necessary background on this, see Chapters 8 and 9 of [6]. The divisor D is rational over $\mathbb{Q}$, and it extends uniquely over $\mathbb{Z}[1/nl]$ to an effective relative Cartier divisor of degree g_l on $X_1(nl)$. The points P_i of which D is the sum need not be rational over $\mathbb{Q}$.

The inverse image of a point x in $J_1(nl)(\overline{\mathbb{Q}})$ under the map (11.3) can be described as follows. Let $\mathcal{L}_x$ denote a line bundle of degree zero on $X_1(nl)_{\overline{\mathbb{Q}}}$ that corresponds to x (x is an isomorphism class of such line bundles). Then the inverse image of x is the set of $(Q_1, \ldots, Q_{g_l})$ such that $\mathcal{L}_x$ has a rational section whose divisor is $Q_1 + \cdots + Q_{g_l} - D$, or, equivalently, the set of $(Q_1, \ldots, Q_{g_l})$ such that there is a nonzero section of $\mathcal{L}_x(D)$ with divisor $Q_1 + \cdots + Q_{g_l}$.

When x runs through $J_1(nl)(\overline{\mathbb{Q}})$, the classes of the $\mathcal{L}_x(D)$ run through $\mathrm{Pic}^{g_l}(X_1(nl)_{\overline{\mathbb{Q}}})$. The function on $J_1(nl)(\overline{\mathbb{Q}})$ that assigns to x the dimension $h^0(\mathcal{L}_x(D))$ of the space of global sections of $\mathcal{L}_x(D)$ is semi-continuous in the sense that for each i the locus of x where $h^0(\mathcal{L}_x(D)) \geq i$ is closed (the condition $h^0(\mathcal{L}_x(D)) \leq i$ need not be closed). On a non-empty open subset of $J_1(nl)(\overline{\mathbb{Q}})$ this value is one, as can be seen using the theorem of Riemann-Roch, and Serre duality. This means that for x outside a proper closed subset of $J_1(nl)(\overline{\mathbb{Q}})$, the inverse image in $X_1(nl)^{g_l}(\overline{\mathbb{Q}})$ of x consists of the g_l-tuples obtained by permutation of coordinates of a single $(Q_1, \ldots, Q_{g_l})$. Another way to express this is to say that the map (11.3) above factors through the symmetric product $X_1(nl)^{(g_l)}$ and that the map from $X_1(nl)^{(g_l)}$ to $J_1(nl)$ is birational (i.e., an isomorphism on suitable non-empty open parts).

It is then reasonable to assume that we can take D such that for all x in $V_\lambda - \{0\}$ there is, up to permutation of the coordinates, a unique $Q = (Q_1, \ldots, Q_{g_l})$ in $X_1(nl)^{g_l}(\overline{\mathbb{Q}})$ that is mapped to x via the map (11.3). On the other hand, on a curve of high genus such as $X_1(nl)$ it is not clear how to make a large supply of inequivalent effective divisors D on $X_1(nl)_{\mathbb{Q}}$. We will see in the next section that we can indeed find a suitable divisor, supported on the cusps, and defined over $\mathbb{Q}(\zeta_l)$, on the $X_1(5l)$, which will suffice for treating the modular form Δ. In situations where such a cuspidal divisor cannot be found, one could try at random $P_1, \ldots, P_g$ in $X_1(nl)(K)$, corresponding to elliptic curves lying in one isogeny class, with complex multiplications, for example by $\mathbb{Q}(i)$. Then K is a solvable Galois extension of $\mathbb{Q}$, which makes it sufficiently disjoint from the extension K_λ so that K_λ can be reconstructed from the compositum KK_λ. If

one chooses the P_i reasonably, the degree of K and the logarithm of the discriminant of K are polynomial in l.

Let us now assume that we have a divisor D as described above. Then we choose a non-constant function:

$$f\colon X_1(nl)_{\mathbb{Q}} \twoheadrightarrow \mathbb{P}^1_{\mathbb{Q}}, \tag{11.4}$$

that will have to satisfy some conditions that will be given in a moment.

With these two choices, D and f, and a choice of $d \in \{1, \ldots, g_l\}$, we get an element $k_{D,f,d}$ of the $\mathbb{Q}$-algebra K'_λ corresponding to $V_\lambda - \{0\}$ as follows. For x in $V_\lambda - \{0\}$ we let $D'_x = Q_{x,1} + \cdots + Q_{x,g_l}$ be the unique effective divisor of degree g_l such that:

$$x = [D'_x - D], \tag{11.5}$$

and we put:

$$k_{D,f,d}\colon V_\lambda - \{0\} \longrightarrow \overline{\mathbb{Q}}, \quad x \mapsto \Sigma_d(f(Q_{x,1}), \ldots, f(Q_{x,g_l})), \tag{11.6}$$

where Σ_d is the elementary symmetric polynomial of degree d in g_l variables. (If the complexity of evaluating the Σ_d becomes a problem, we will replace them by the power sums of the same degree.) Obviously, one condition that f must verify is that all the D'_x are disjoint from the locus of poles of f. This condition will not be guaranteed to hold later when we treat Δ, but then it will be possible to omit the $Q_{x,i}$ at which f has a pole from the sum in (11.6) (f will have its poles at certain cusps). For the moment, let us just assume that this condition is satisfied. Then the $f_*D'_x$ are effective divisors of degree g_l on $\mathbb{A}^1_{\overline{\mathbb{Q}}}$.

We will choose f in such a way that the $f_*D'_x$ are distinct (one way to achieve this is to make sure that this holds after reduction at a suitable prime number p); we assume now that this is so. The g_lth symmetric product of $\mathbb{A}^1_{\overline{\mathbb{Q}}}$, that parametrises effective divisors of degree g_l on $\mathbb{A}^1_{\overline{\mathbb{Q}}}$, is, as a quotient of the product $(\mathbb{A}^1_{\overline{\mathbb{Q}}})^{g_l}$, given by the map:

$$\Sigma\colon (\mathbb{A}^1_{\overline{\mathbb{Q}}})^{g_l} \to \mathbb{A}^{g_l}_{\overline{\mathbb{Q}}}, \quad (y_1, \ldots, y_{g_l}) \mapsto (\Sigma_1(y_1, \ldots, y_{g_l}), \ldots, \Sigma_{g_l}(y_1, \ldots, y_{g_l})). \tag{11.7}$$

The $\overline{\mathbb{Q}}$-point of $\mathbb{A}^{g_l}$ corresponding to $f_*D'_x$ is the point $k_{D,f}(x) \in \overline{\mathbb{Q}}^{g_l}$ whose coordinates are the $k_{D,f,d}(x)$, $1 \leq d \leq g_l$. This means that these $k_{D,f,d}$ together generate K'_λ. A suitable linear combination:

$$k := \sum_d a_d k_{D,f,d} \tag{11.8}$$

with "small" a_d in $\mathbb{Z}$ is then a generator of K'_λ.

Finally, we want to have control on the heights of the coefficients of the minimal polynomial of k, because these heights determine the required precision of the approximations of those coefficients that we must compute. The whole strategy depends on the possibility to choose, for all λ, a divisor D and a function f, such that those heights grow at most polynomially in $\#\mathbb{F}_\lambda$. Using a great deal of machinery from Arakelov theory, we will show (at least in the case of Δ) that any reasonable choices of D and f will lead to an at most polynomial growth of those heights. Intuitively, and completely non-rigorously, one can believe that this should work, because of the following argument. Our x are torsion points, so that their Néron-Tate height is zero. As x and D determine D'_x, the height of D'_x should be not much bigger than the height of D. As we choose D ourselves, it should have small height. Finally, the height of $k_{D,f,d}$ should be not much bigger than the sum of those of f and the D'_x. Turning these optimistic arguments into rigorous statements implies a lot of work (that will in fact occupy most of the rest of this text). An important problem here is that in Arakelov theory many results are available that deal with a single curve over $\mathbb{Q}$, but in our situation we are dealing with the infinitely many curves $X_1(nl)$ as l varies.

A few words about the numerical computations involved. What we need is that these can be done in a time that is polynomial in $\#\mathbb{F}_\lambda$ and the number of significant digits that one wants for the coefficients of P_λ. It is not at all obvious that this can be done, as the genus of $X_1(nl)$ and hence the dimension of $J_1(nl)$ are quadratic in l.

One way to do the computations is to use the complex uniformisations of $X_1(nl)(\mathbb{C})$ and $J_1(nl)(\mathbb{C})$. The Riemann surface $X_1(nl)(\mathbb{C})$ can be obtained by adding finitely many cusps (the set $\Gamma_1(nl)\backslash\mathbb{P}^1(\mathbb{Z})$) to the quotient $\Gamma_1(nl)\backslash\mathbb{H}$ (see Section 7). This means that $X_1(nl)(\mathbb{C})$ is covered by disks around the cusps, which are well suited for computations (functions have q-expansions, for example). In order to describe $J_1(nl)(\mathbb{C})$ as $\mathbb{C}^{g_l}$ modulo a lattice, we need a basis of the space of holomorphic differential forms $\mathrm{H}^0(X_1(nl)(\mathbb{C}),\Omega^1)$. The basis that we work with is the one provided by Atkin-Lehner theory, as given in (8.17); we write it as $\omega=(\omega_1,\dots,\omega_{g_l})$. Then we have the following complex description of the map (11.3):

$$X_1(nl)(\mathbb{C})^{g_l} \longrightarrow\!\!\!\!\!\!\rightarrow J_1(nl)(\mathbb{C}) = \mathbb{C}^{g_l}/\Lambda \tag{11.9}$$

$$(Q_1,\dots,Q_{g_l}) \longmapsto [Q_1+\cdots+Q_{g_l}-P_1-\cdots-P_{g_l}] = \sum_{i=1}^{g_l}\int_{P_i}^{Q_i}(\omega_1,\dots,\omega_{g_l}),$$

where Λ is the period lattice of this basis, i.e., the image of $\mathrm{H}_1(X_1(nl)(\mathbb{C}),\mathbb{Z})$ under integration of the ω_i. This map can be computed up to any desired precision by formal integration of power series on the disks mentioned above. The coefficients needed from the power series expansions

of the ω_i can be computed using the method of modular symbols, as has been implemented by William Stein in Magma (see William Stein's home page). We note that modular symbols algorithms can be used very well to locate V_λ inside $l^{-1}\Lambda/\Lambda$, hence in $J_1(nl)(\mathbb{C})$. A strategy to approximate a point $Q_x = (Q_{x,1}, \ldots, Q_{x,g_l})$ as above for a nonzero x in V_λ is to lift the straight line that one can draw in $\mathbb{C}^{g_l}/\Lambda$ from 0 to x (within a suitable fundamental domain for Λ) to a path in $X_1(nl)(\mathbb{C})^{g_l}$ starting at $(P_1, \ldots, P_{g_l})$. In practice this seems to work reasonably well, see Section 24. A theoretical difficulty with this approach is that one needs to bound from below the distance to the ramification locus of $X_1(nl)^{(g_l)} \to J_1(nl)$.

We note that Couveignes has another approach to the computational issue over the complex numbers; see [13]. Unfortunately, we cannot use the results of [13] in our situation, as those results are concerned with the Jacobians of the modular curves $X_0(p)$ for p prime.

Another way to do the "approximation" is to compute the minimal polynomial P_λ of k modulo many small primes p. Indeed, the map (11.3) can be reduced mod p. In this case one has no analytic description of the curve and its Jacobian, but one can make random points in $J_1(nl)(\mathbb{F}_q)$ for a suitable finite extension $\mathbb{F}_p \to \mathbb{F}_q$. Such random points can then be projected, using Hecke operators, into V_λ. Elements of $J_1(nl)(\mathbb{F}_q)$ can be represented by divisors on $X_1(nl)_{\mathbb{F}_q}$, and all necessary operations can be done in polynomial time. This approach is the one taken in [14]; we will explain it in detail in Section 21.

Finally, in case one has a natural rigid analytic uniformisation at some prime p, one may want to use that. For the modular curves that we are dealing with this is not the case, but the closely related Shimura curves associated to quaternion algebras over $\mathbb{Q}$ do admit such uniformisations at the primes where the quaternion algebra is ramified (as was proved by Cerednik, Drinfeld, see [5]).

12 Construction of a suitable cuspidal divisor on $X_1(5l)$

In this section we put ourselves in the situation where we want to compute the residual representation $\overline{\rho}_l$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to $\mathrm{GL}_2(\mathbb{F}_l)$ associated to the cuspform Δ. We assume that l is not exceptional, so that the image of $\overline{\rho}_l$ contains $\mathrm{SL}_2(\mathbb{F}_l)$. We let V_l denote the representation $\overline{\rho}_l$ viewed as a 2-dimensional $\mathbb{F}_l$-vector space with $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-action. By Theorem 10.9 V_l is naturally embedded in $J_1(l)(\overline{\mathbb{Q}})$.

As explained in the previous section, we would like to have an effective divisor D on $X_1(l)_{\mathbb{Q}}$ of degree the genus of $X_1(l)$ such that for all non-zero x in the submodule V_l of $J_1(l)(\overline{\mathbb{Q}})$ we have $h^0(\mathcal{L}_x(D)) = 1$. It would be nice to have a cuspidal divisor (i.e., a divisor supported on the cusps) with this property. The first complication is that the cusps are not all rational over $\mathbb{Q}$: half

of them have the maximal real subfield of $\mathbb{Q}(\zeta_l)$ as field of definition. Moreover, even working with all the cusps, we have not succeeded to find a suitable cuspidal divisor D. On the other hand, below we will give explicitly a cuspidal divisor D on the curve $X_1(5l)_{\mathbb{Q}(\zeta_l)}$ that has the property that $h^0(\mathcal{L}_x(D)) = 1$ for each x in $J_1(5l)(\overline{\mathbb{Q}})$ that specialises to 0 at some place of $\overline{\mathbb{Q}}$ over l. In particular, D has the required property for V_l embedded in $J_1(5l)(\overline{\mathbb{Q}})$ in an arbitrary way (this will be shown in the next section). We have chosen to work with $X_1(5l)$, but the same method will work for modular curves corresponding to some level structure if the prime to l part of the level structure is fine, and of genus zero.

For the rest of this section, our assumptions are the following: l is a prime number, not equal to 5. We let $X := X_1(5l)_{\mathbb{Q}(\zeta_l)}$ over $\mathbb{Q}(\zeta_l)$. References for facts about X that we use can be found in [50], and also in [32]; they are derived from results in [21] and in [61].

The curve $X_0(5)_{\mathbb{Q}}$ has 2 cusps, both $\mathbb{Q}$-rational, called 0 and ∞ (after the points of $\mathbb{P}^1(\mathbb{Q})$ of which they come). The cusp ∞ has as moduli interpretation the degenerate elliptic curve (generalised elliptic curve in the terminology of [21]): the 1-gon, equipped with the unique subgroup of order 5 of $\mathbb{G}_{\mathrm{m}}$. The cusp 0 corresponds to the 5-gon, equipped with a subgroup of order 5 that meets all 5 components. The group $\mathbb{F}_5^\times$ acts (as diamond operators) on $X_1(5)$, with quotient $X_0(5)$; the subgroup $\{\langle\pm 1\rangle\}$ acts trivially, and the quotient $\mathbb{F}_5^\times$ by this subgroup acts faithfully. The inverse images of 0 and ∞ both consist of two cusps. Those over 0 are $\mathbb{Q}$-rational (the subgroup of order 5 of the 5-gon is the constant groupscheme $\mathbb{Z}/5\mathbb{Z}$), whereas those over ∞ are conjugated over $\mathbb{Q}(\sqrt{5})$. We fix one $\mathbb{Q}$-rational cusp c of $X_1(5)$.

The group $\mathbb{F}_l^\times$ acts faithfully, and in fact, even freely, on X. The set of cusps of X over the cusp c of $X_1(5)$ form two $\mathbb{F}_l^\times$-orbits, corresponding to the type of degenerate elliptic curve that they correspond to: 5-gon or $5l$-gon. The orbit corresponding to the 5-gon consists of points rational over $\mathbb{Q}(\zeta_l)$, all conjugates of each other. The orbit corresponding to the $5l$-gon consists of $\mathbb{Q}$-points.

We let J denote the Jacobian of X. What we want is an effective divisor D of degree g on X (with g the genus of X), supported on the cusps over c, such that for all x in $J(\overline{\mathbb{Q}})$ that specialise to 0 at some place of $\overline{\mathbb{Q}}$ over l we have $h^0(X_{\overline{\mathbb{Q}}}, \mathcal{L}_x(D)) = 1$. For the notion of specialisation we use Néron models; the reader is referred to [6] for this notion. For x in $J(K)$ with $K \subset \overline{\mathbb{Q}}$ a finite extension of $\mathbb{Q}(\zeta_l)$, and λ a place of $\overline{\mathbb{Q}}$ over l, we say that x specialises to 0 at λ if x, viewed as element of $J_{O_K}(O_K)$, with J_{O_K} the Néron model of J over O_K, specialises to 0 at the place of K given by λ. For $K \subset K'$ a finite extension we have $J(K) \subset J(K')$, hence we can also view x as element of $J(K')$. The notion of x specialising to zero is the same for K and K', because $J_{\mathbb{Z}[\zeta_l]}$ is semi-stable at l.

The moduli interpretation of X gives a semi-stable model $X_{\mathbb{Z}[\zeta_l, 1/5]}$ over $\mathbb{Z}[\zeta_l, 1/5]$, described in [50] for example. A result of Raynaud identifies the connected component of the Néron model

J_{O_K} with the connected component of the Picard scheme of X_{O_K} (see Section 9.5 of [6]). This means that for x in $J(K)$ specialising to 0 at λ the line bundle $\mathcal{L}_x$ on X_K associated to x can be extended uniquely over the local ring $O_{K,\lambda}$ to a line bundle $\mathcal{L}_x$ on $X_{O_{K,\lambda}}$ such that the restriction $\overline{\mathcal{L}_x}$ of $\mathcal{L}_x$ to the special fibre $X_{\mathbb{F}_\lambda}$ is trivial. The divisor D on X_K extends, by taking the Zariski closure, to an effective Cartier divisor on $X_{O_{K,\lambda}}$.

We now note that $h^0(X_K, \mathcal{L}_x(D))$ is at least one, by Riemann-Roch, and that $h^0(X_K, \mathcal{L}_x(D))$ is at most $h^0(X_{\mathbb{F}_\lambda}, \overline{\mathcal{L}_x}(\overline{D}))$ by semi-continuity of cohomology of coherent sheaves. As $\overline{\mathcal{L}_x}(\overline{D})) = \mathcal{O}(\overline{D})$, it now suffices to take D such that $h^0(X_{\overline{\mathbb{F}}_l}, \mathcal{O}(\overline{D})) = 1$. We do this by looking at the geometry of $X_{\overline{\mathbb{F}}_l}$. As $\mathbb{Z}[\zeta_l]$ has a unique morphism to $\overline{\mathbb{F}}_l$, the curve $X_{\overline{\mathbb{F}}_l}$ does not depend on K. The scheme of cusps of $X_{\mathbb{Z}_l[\zeta_l]}$ is finite étale over $\mathbb{Z}_l[\zeta_l]$, hence the cusps lying over c specialise injectively to $X_{\mathbb{F}_l}$.

The curve $X_{\overline{\mathbb{F}}_l}$ is the union of two irreducible components, X_1 and X_2, say, both isomorphic to the Igusa curve of level l over $X_1(5)$ over $\overline{\mathbb{F}}_l$, that meet transversally in the set Σ of supersingular points. We will take $\overline{D} = D_1 + D_2$, with D_1 on X_1 and D_2 on X_2; note that the cusps are disjoint from Σ, so $\overline{D}$ lies in the smooth locus of $X_{\overline{\mathbb{F}}_l}$. In order to simplify, we let X and D denote $X_{\overline{\mathbb{F}}_l}$ and $\overline{D}$, from now on.

We let Ω_X be the dualising sheaf on X (see Section 8 of [50], or [78]). By Riemann-Roch, what we want is that $h^1(X, \mathcal{O}(D)) = 0$, and hence, by Serre duality, that $h^0(X, \Omega_X(-D)) = 0$. In other words, a regular differential on X that vanishes on D must be zero. Restriction to X_1 gives a short exact sequence:

$$0 \to \mathrm{H}^0(X_2, \Omega^1_{X_2/\overline{\mathbb{F}}_l}(-D_2)) \to \mathrm{H}^0(X, \Omega_X(-D)) \to \mathrm{H}^0(X_1, \Omega^1_{X_1/\overline{\mathbb{F}}_l}(\Sigma - D_1)) \to 0. \tag{12.1}$$

Hence, it suffices to take D_1 such that $\mathrm{H}^0(X_1, \Omega^1_{X_1/\overline{\mathbb{F}}_l}(\Sigma - D_1)) = 0$ and D_2 such that $\mathrm{H}^0(X_2, \Omega^1_{X_2/\overline{\mathbb{F}}_l}(-D_2)) = 0$. Let us now first see of which degrees d_1 and d_2 we want to take D_1 and D_2. Let g_1 and g_2 be the genera of X_1 and X_2 (note: they are equal). Then we have that $g = g_1 + g_2 + \#\Sigma - 1$, and $g = \deg D = d_1 + d_2$. As the stack $[\Gamma_1(5)]$ has degree 24 over [Ell], the degree of the sheaf $\underline{\omega}$ on it is one (see Corollary 10.13.12 of [61]). Therefore, the Hasse invariant has exactly $l - 1$ zeros on $X_1(5)$ over $\overline{\mathbb{F}}_l$ and therefore we have:

$$\#\Sigma = l - 1. \tag{12.2}$$

Applying Hurwitz's formula to the covering $X_1 \to X_1(5)_{\overline{\mathbb{F}}_l}$, which is totally ramified over Σ and unramified outside it, gives:

$$2g_1 - 2 = -2(l-1) + (l-1)(l-2), \qquad g_1 = \frac{1}{2}(l-2)(l-3). \tag{12.3}$$

This implies that we want to take:

$$d_1 = g_1 + \#\Sigma - 1 = \frac{1}{2}(l-1)(l-2), \quad d_2 = g_2 = \frac{1}{2}(l-2)(l-3). \tag{12.4}$$

Now we use equations to compute. We choose a coordinate x on $X_1(5)_{\overline{\mathbb{F}}_l}$ such that Σ does not contain 0 or ∞ and such that 0 is the (rational) cusp 0 of $X_1(5)_{\overline{\mathbb{F}}_l}$. Let f be the monic polynomial in x whose roots are the elements of Σ, with multiplicity one. Then X_1 and X_2 are both isomorphic to the cover of $X_1(5)_{\overline{\mathbb{F}}_l}$ given by the equation $y^{l-1} = f$. Here we use that the action of $\mathbb{F}_l^*$ on the tangent spaces of X_1 at the elements of Σ are all given by the same character.

We compute a basis of $\mathrm{H}^0(X_1, \Omega^1_{X_1/\overline{\mathbb{F}}_l}(\Sigma))$. On X_1 we have:

$$-y^{l-2}dy = f'dx. \tag{12.5}$$

Hence $(dx)/y^{l-2} = -(dy)/f'$ is a generating section of $\Omega^1_{X_1/\overline{\mathbb{F}}_l}$ on the affine part given by our equation. Hence $(dx)/y^{l-1}$ is generating section of $\Omega^1_{X_1/\overline{\mathbb{F}}_l}(\Sigma)$ on the affine part, and it is $\mathbb{F}_l^*$-invariant. At each point of X_1 over the point where x has its pole, both x and y have a simple pole, and $(dx)/y^{l-1}$ has order $-2 + l - 1 = l - 3$. So we have a basis:

$$\mathrm{H}^0(X_1, \Omega^1_{X_1/\overline{\mathbb{F}}_l}(\Sigma)) = \bigoplus_{i+j \leq l-3} \overline{\mathbb{F}}_l x^i y^j \cdot (dx)/y^{l-1}. \tag{12.6}$$

Note that this agrees with the fact that $d_1 = \frac{1}{2}(l-1)(l-2)$.

We can now say how to choose D_1. At each of the $l - 1$ points where x has a zero we must give a multiplicity. Here is how we do it: just distribute the multiplicities $(0, 1, \ldots, l-2)$ over these points. Then one directly sees that any linear combination of our basis elements that vanishes on D_1 is zero.

Now we do D_2. A basis is the following:

$$\mathrm{H}^0(X_2, \Omega^1_{X_2/\overline{\mathbb{F}}_l}) = \bigoplus_{i+j \leq l-4} \overline{\mathbb{F}}_l x^i y^j \cdot (dx)/y^{l-2}. \tag{12.7}$$

We note that this agrees with $g_2 = (l-3)(l-2)/2$. So, for D_2, just distribute the multiplicities $(0, 0, 1, \ldots, l-3)$ over the points where x has a zero.

We summarise our results. As the action of $\mathbb{F}_5^\times$ permutes the two $\mathbb{Q}$-rational cusps of $X_1(5)$, our arguments above work for both of them.

12.8 Theorem *Let l be a prime number not equal to 5. Let c be one of the two $\mathbb{Q}$-rational cusps of $X_1(5)$. Then the cusps of $X_1(5l)$ over c are $\mathbb{Q}(\zeta_l)$-rational, and consist of two $\mathbb{F}_l^\times$-orbits, on which $\mathbb{F}_l^\times$ acts freely. Let D_1 be a divisor on $X_1(5l)_{\mathbb{Q}(\zeta_l)}$ obtained by distributing the multiplicities $(0, 1, \ldots, l-2)$ over one of these two orbits. Let D_2 be the divisor obtained by distributing the multiplicities $(0, 0, 1, \ldots, l-3)$ over the other orbit. Then $D := D_1 + D_2$ has degree equal to the genus of $X_1(5l)$ and has the property that for any $\overline{\mathbb{Q}}$-point x of the Jacobian of $X_1(5l)$ that specialises to 0 at some place over l we have $h^0(X_1(5l)_{\overline{\mathbb{Q}}}, \mathcal{L}_x(D)) = 1$.*

13 The exact setup for computing Ramanujan's τ-function

In Section 11 we described our strategy for computing the residual Galois representations associated to a fixed newform. That strategy depends on properties of divisors D and functions f to be chosen, on modular curves of varying level. These D and f must satisfy a number of conditions. In general we do not know how to choose divisors D of which we can prove, without a computer computation, that they have the required property. This is the main reason for which we will now restrict ourselves to just the case of the discriminant newform Δ of weight 12. The aim of this section is to describe exactly our strategy for computing the residual representations V_l in this case. We assume from now on that $l \geq 13$ and that l is not exceptional for Δ (i.e., $l \neq 23$ and $l \neq 691$).

Theorem 10.9 tells us that V_l occurs in $J_1(l)(\overline{\mathbb{Q}})[l]$, and describes it in terms of Hecke operators. Theorem 12.8 gives us a divisor D on $X_1(5l)_{\mathbb{Q}(\zeta_l)}$ that we want to use. Therefore, we want to embed V_l into $J_1(5l)(\overline{\mathbb{Q}})[l]$.

Let $\pi\colon X_1(5l) \to X_1(l)$ be the standard map (i.e., the one that forgets the 5-part of the level structure, the one denoted $B_{5l,l,1}$ in Section 8). Then the degree of π is $5^2 - 1 = 24$, which is prime to l. This implies that $\pi_*\pi^*$ is multiplication by 24 on $J_1(l)$, and that π^* is injective on $J_1(l)(\overline{\mathbb{Q}})[l]$. We have a projector:

$$\frac{1}{24}\pi^*\pi_*\colon J_1(5l)(\overline{\mathbb{Q}})[l] \twoheadrightarrow \pi^*J_1(l)(\overline{\mathbb{Q}})[l] \subset J_1(5l)(\overline{\mathbb{Q}})[l]. \tag{13.1}$$

We will consider V_l embedded in $J_1(5l)(\overline{\mathbb{Q}})[l]$ via its embedding into $J_1(l)(\overline{\mathbb{Q}})[l]$, followed by π^*.

13.2 Proposition *Let D be a divisor on $X_1(5l)_{\mathbb{Q}(\zeta_l)}$ as given in Theorem 12.8. Then for every x in V_l we have $h^0(X_1(5l)_{\overline{\mathbb{Q}}}, \mathcal{L}_x(D)) = 1$.*

Proof In view of Theorem 12.8, it suffices to show that for each x in V_l there is a place of $\overline{\mathbb{Q}}$ over l at which x specialises to 0. The notion of specialisation is explained in Section 12. Let $J_{\mathbb{Z}[\zeta_l]}$ denote the Néron model of $J := J_1(5l)$ over $\mathbb{Z}[\zeta_l]$. Then V_l is the group of $\overline{\mathbb{Q}}$-points of a groupscheme (even a $\mathbb{F}_l$-vector space scheme) $\mathcal{V}_{\mathbb{Q}(\zeta_l)}$ in $J_{\mathbb{Q}(\zeta_l)}$. Let $\mathcal{V}$ be the Zariski closure of $\mathcal{V}_{\mathbb{Q}(\zeta_l)}$ in $J_{\mathbb{Z}[\zeta_l]}$. Then it is shown in Section 12 of [50] and in Section 6 of [32] that $\mathcal{V}_{\mathbb{Z}_l[\zeta_l]}$ is finite locally free over $\mathbb{Z}_l[\zeta_l]$, and that the rank of the local part of $\mathcal{V}_{\mathbb{Z}_l[\zeta_l]}$ is l if $l \not| \tau(l)$ and l^2 if $l|\tau(l)$. (We note that it does not matter if we take Zariski closure in $J_1(l)$ or in $J_1(5l)$, as π^* gives a closed immersion of the l-torsion of $J_1(l)$ over $\mathbb{Z}_l$ into that of $J_1(5l)$.) This means that at each place of $\overline{\mathbb{Q}}$ over l there is a non-zero x in V_l that specialises to 0. Under our assumptions, the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$ acting on V_l is $\mathrm{SL}(V_l)$. Hence $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$ acts transitively on $V_l - \{0\}$. Hence for each x in $V_l - \{0\}$ there are places of $\overline{\mathbb{Q}}$ over l where x specialises to 0. $\square$

The fact that our divisor D lives on $X_1(5l)_{\mathbb{Q}(\zeta_l)}$, and not on $X_1(5l)_{\mathbb{Q}}$, forces us to work over $\mathbb{Q}(\zeta_l)$, and not over $\mathbb{Q}$, as in Section 11.

We let X_l denote $X_1(5l)_{\mathbb{Q}}$ and g_l its genus, and we let $K'_{l,\mathbb{Q}(\zeta_l)}$ denote the extension of $\mathbb{Q}(\zeta_l)$ that corresponds to the transitive $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$-set $V_l - \{0\}$.

For each x in $V_l - \{0\}$ there is a unique effective divisor $D'_x = \sum_i Q_{x,i}$ of degree g_l on $X_{l,\overline{\mathbb{Q}}}$ such that $x = [D'_x - D]$. The uniqueness of D'_x implies that:

$$D'_{gx} = gD'_x, \quad \text{for all } g \text{ in } \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l)). \tag{13.3}$$

We write each D'_x as:

$$D'_x = D''_x + (D'_x)^{\mathrm{rest}}, \tag{13.4}$$

where $(D'_x)^{\mathrm{rest}}$ is supported on the cusps of $X_{l,\overline{\mathbb{Q}}}$ and where D''_x is disjoint from the cusps. As the cusps of $X_{l,\overline{\mathbb{Q}}}$ form a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-stable subset of $X_{l,\overline{\mathbb{Q}}}$ we have:

$$D''_{gx} = gD''_x, \quad \text{for all } g \text{ in } \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l)). \tag{13.5}$$

In particular, the D''_x all have the same degree, d_l, say, and we write $D''_x = \sum_{i \le d_l} Q_{x,i}$.

Any choice of $f\colon X_{l,\mathbb{Q}(\zeta_l)} \twoheadrightarrow \mathbb{P}^1_{\mathbb{Q}(\zeta_l)}$ whose poles are supported on the cusps then gives us elements $k_{D,f,d}$ of $K'_{l,\mathbb{Q}(\zeta_l)}$, with $1 \le d \le d_l$, defined as:

$$k_{D,f,d}\colon V_l - \{0\} \to \overline{\mathbb{Q}}, \quad x \mapsto \Sigma_d(f(Q_{x,1}), \ldots, f(Q_{x,d_l})), \quad \text{where } D''_x = \sum_{1 \le i \le d_l} Q_{x,i}, \tag{13.6}$$

and where, as before, Σ_d denotes the elementary symmetric polynomial of degree d.

The next issue that we address is the choice of f; we want the divisors $f_*D''_x$ on $\mathbb{A}^1_{\mathbb{Q}(\zeta_l)}$ to be distinct.

We start by giving an explicit description of the curve $Y_1(5)$ over $\mathbb{Z}[1/5]$. In order to do that, we determine a universal triple $(E/S, P)$ where E/S is an elliptic curve over an arbitrary scheme, and P in $E(S)$ is everywhere of order 5, i.e., for every $\mathrm{Spec}(A) \to S$ with A non-zero, the image of P in $E(A)$ has order 5. The base of this triple is the open part $Y_1(5)'$ of the model $Y_1(5)$ over $\mathbb{Z}$ (constructed in Chapter 8 of [61]) where the P has order 5 (i.e., $Y_1(5)'$ is the complement of the irreducible component of $Y_1(5)_{\mathbb{F}_5}$ where the point P generates the kernel of Frobenius). The equation of this universal triple can also be found on page 7 of Tom Fischer's thesis, see [42].

13.7 Proposition *Let E/S be an elliptic curve, and $P \in E(S)$ a point that is everywhere of order 5. Then $(E/S, P)$ arises via a unique base change from the following triple:*

$$\begin{cases} Y_1(5)' = \mathrm{Spec}(\mathbb{Z}[b, 1/\mathrm{discr}(E)]), \quad \mathrm{discr}(E) = -b^5(b^2 + 11b - 1) \\ E : y^2 + (b+1)xy + by = x^3 + bx^2 \\ P = (0,0). \end{cases}$$

The j-invariant of E is given by:

$$j(E) = -(b^4 + 12b^3 + 14b^2 - 12b + 1)^3/b^5(b^2 + 11b - 1).$$

Proof Our proof is modelled on Section 2.2 of [61]; basic properties of Weierstrass equations for elliptic curves are used without being mentioned.

Let $(E/S, P)$ be given, with P everywhere of order 5. Choose a parameter t at 0, up to order 2, i.e., a trivialisation of $\omega_{E/S}$. Note: we are working locally on S, here; in the end, as we will succeed in making things unique, our construction will be global. Note: t is unique up to $t' = ut$, with $u \in R^\times$ ($S = \mathrm{Spec}(R)$ now).

Choose x a global function on $E - 0(S)$ such that $x = t^{-2} + \cdots$. Then x is unique up to $x' = x + a$, $a \in R$. Make x unique by demanding that $x(P) = 0$ (this is alright because 0 and P are disjoint).

Choose $y = t^{-3} + \cdots$ regular on $E - 0(S)$. Then y is unique up to $y' = y + ax + b$. Make y unique by demanding that $y(P) = 0$ and that the tangent of E at P is the line given by the equation $y = 0$. (Indeed, use b (uniquely) to get $y(P) = 0$, then note that the tangent at P is nowhere the line given by $x = 0$ because P is nowhere annihilated by 2).

We have $5{\cdot}P \equiv 5{\cdot}0$, hence there is a unique f on $E - 0(S)$ of the form

$$f = xy + \alpha y + \beta x^2 + \gamma x + \delta$$

such that the divisor of f is $5{\cdot}P - 5{\cdot}0$. The conditions that x and y have order one and two, respectively, at 0 imply that $\gamma = \delta = 0$. The equation for E is of the form:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2,$$

because the usual coefficients a_4 and a_6 are zero. We also see that a_3 is a unit because E is smooth at $(0, 0)$. The coefficient a_2 is a unit, because P is nowhere annihilated by 3.

Now we try to get rid of u (the ambiguity in the choice of t). If $t' = u^{-1}t$, then $a'_i = t^i a_i$, hence we can make t unique by demanding that $a_2 = a_3$. We do that, and then we have the following equations.

The elliptic curve E and the point P are given by:

$$y^2 + axy + by = x^3 + bx^2, \quad P = (0, 0).$$

The function f with divisor $5{\cdot}P - 5{\cdot}0$ is given by:

$$f = xy + \alpha y + \beta x^2.$$

Here we know that b, α and β are in $R^\times$, because $v_P(x) = 1$ and $v_P(y) = 2$ everywhere on S. Now we have to compute what it means that $v_P(f) = 5$. This means that the intersection multiplicity of the two curves E and $V(f)$ at $(0,0)$ is 5. A systematic way to compute that is to do successive blow-ups; that works nicely, but we will not do that here. A much faster way to do the computation is to take suitable linear combinations of the equations for E and f directly. One finds the equations:

$$\beta = -\alpha, \quad a = \alpha^{-1}b + 1, \quad \alpha = 1.$$

□

The reason we give such a detailed description of $Y_1(5)$ is that it gives us functions on all the $Y_1(5l)$, at least over $\mathbb{Z}[1/5l]$, as stated in the following proposition.

13.8 Proposition *Let $l \neq 5$ be prime. Let $E/Y_1(5)'$ be the elliptic curve given in Proposition 13.7. Then $Y_1(5l)$ and $E[l] - \{0\}$ agree over $\mathbb{Z}[1/5l]$. In particular, b, x and y generate the coordinate ring of $Y_1(5l)$ over $\mathbb{Z}[1/5l]$.*

Proof This is standard. The first part follows immediately from the universality of $Y_1(5)'$ and the fact that $E[l]$ is finite étale over $Y_1(5)'$ away from characteristic l. The second statement follows from the fact that $E[l] - \{0\}$ is a closed subscheme of the affine scheme $E - \{0\}$. □

Using these functions b, x and y, we can now say how we will choose the function f.

We let D'' be the sum of all the D''_x. Then D'' is an effective divisor on $X_{l,\overline{\mathbb{Q}}}$, disjoint from the cusps, of degree $d_l(l^2-1)$. We then take a suitable "small" linear combination

$$f := \alpha b + \beta x + \gamma y \tag{13.9}$$

with α, β and γ in $\mathbb{Z}[\zeta_l]$, such that under f all geometric points of D'' have distinct images. Finding such α, β and γ will be done via reduction modulo a suitable small finite place of $\mathbb{Q}(\zeta_l)$. The choice of f implies that the divisors $f_*D''_x$ on $\mathbb{A}^1_{\mathbb{Q}(\zeta_l)}$ are distinct, for x running through $V_l - \{0\}$. Equivalently, this means that the l^2-1 points $k_{D,f}(x)$ in $\overline{\mathbb{Q}}^{d_l}$ with coordinates $k_{D,f,d}(x)$ are distinct. This means that the $k_{D,f,d}$ together generate $K'_{l,\mathbb{Q}(\zeta_l)}$. Then a suitable $\mathbb{Z}[\zeta_l]$-linear combination:

$$k := \sum_{1 \leq d \leq d_l} a_d k_{D,f,d} \tag{13.10}$$

will be a generator of $K'_{l,\mathbb{Q}(\zeta_l)}$ over $\mathbb{Q}(\zeta_l)$. Finding such a_d can be done using reduction modulo the same finite place as the one that was used to find f. Indeed, the condition on such an

$a = (a_1, \ldots, a_{d_l})$ is that it must avoid the hyperplanes of $\overline{\mathbb{Q}}^{d_l}$ orthogonal to the differences between the $k_{D,f}(x)$. Over $\mathbb{Q}(\zeta_l)$, this means that a must avoid the subspaces (defined over $\mathbb{Q}(\zeta_l)$) obtained by intersecting each of these hyperplanes with all their $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$-conjugates. We let:

$$P_l := \prod_{x \in V_l - \{0\}} (T - k(x)) \quad \text{in } \mathbb{Q}(\zeta_l)[T] \tag{13.11}$$

be the minimal polynomial over $\mathbb{Q}(\zeta_l)$ of k.

14 Very short introduction to heights and Arakelov theory

In the previous section it has been explained how the computation of the residual Galois representations V_l associated to Δ should proceed. The essential step is to approximate the minimal polynomial P_l of (13.11) with sufficient precision so that P_l itself can be obtained. The topic to be addressed now is to bound from above the precision that is needed for this. This means that we must bound the heights of the coefficients of P_l. As was hinted to in Section 11, we get such bounds using Arakelov theory, a tool that we discuss in this section. It is not at all excluded that a direct approach to bound the coefficients of P_l, thus avoiding the complicated theory that we use, could work. On the other hand, it is clear that the use of Arakelov theory provides a way to split the work to be done in smaller steps, and that the quantities occurring in each step are intrinsic in the sense that they do not depend on coordinate systems or other choices that one could make. We also want to point out that our method does not depend on cancellations of terms in the estimates that we will do; all contributions encountered can be bounded appropriately.

A good reference for a more detailed introduction to heights is Chapter 6 of [12]. Good references for the Arakelov theory that we will use are [39] and [84].

14.1 Heights on $\mathbb{Q}$ and $\overline{\mathbb{Q}}$

The definition of the height of an element of $\mathbb{Q}$ has already been given in Section 11; for $x = a/b$ with a and $b \neq 0$ relatively prime integers, we have $h(x) = \log \max\{|a|, |b|\}$. We will now give an equivalent definition in terms of absolute values $|\cdot|_v$ on $\mathbb{Q}$ associated to all places v of $\mathbb{Q}$, the finite places, indexed by the prime numbers, and the infinite place denoted ∞.

The absolute value $|\cdot|_\infty$ is just the usual absolute value on $\mathbb{R}$, restricted to $\mathbb{Q}$. We note that $\mathbb{R}$ is the completion of $\mathbb{Q}$ for $|\cdot|_\infty$. For p prime, we let v_p be the p-adic valuation:

$$v_p \colon \mathbb{Z} \longrightarrow \mathbb{Z} \cup \{\infty\}, \tag{14.1.1}$$

sending an integer to the maximal number of times that it can be divided by p. This valuation v_p extends uniquely to $\mathbb{Q}$ subject to the condition that $v_p(xy) = v_p(x) + v_p(y)$; we have $v_p(a/b) = v_p(a) - v_p(b)$ for integers a and $b \neq 0$. We let $|\cdot|_p$ denote the absolute value on $\mathbb{Q}$ defined by:

$$|x|_p = p^{-v_p(x)}, \quad |0|_p = 0. \tag{14.1.2}$$

The completion of $\mathbb{Q}$ with respect to $|\cdot|_p$ is the locally compact topological field $\mathbb{Q}_p$. An important property of these absolute values is that all together they satisfy the product formula:

$$\prod_v |x|_v = 1, \quad \text{for all } x \text{ in } \mathbb{Q}^\times. \tag{14.1.3}$$

With these definitions, we have:

$$h(x) = \sum_v \log \max\{1, |x|_v\}, \quad \text{for all } x \in \mathbb{Q}, \tag{14.1.4}$$

where v ranges over the set of all places of $\mathbb{Q}$ (note that almost all terms in the sum are equal to 0).

The height function on $\mathbb{Q}$ generalises as follows to number fields. First of all, for a local field F we define the natural absolute value $|\cdot|$ on it by letting, for x in $F^\times$, $|x|_F$ be the factor by which all Haar measures on F are scaled by the homothecy $y \mapsto xy$ on F. For example, for $F = \mathbb{C}$ we have $|z|_{\mathbb{C}} = z\overline{z} = |z|^2$, the square of the usual absolute value. Let now K be a number field. By a finite place of K we mean a maximal ideal of O_K. An infinite place of K is an embedding of K into $\mathbb{C}$, up to complex conjugation. For each place v of K, let K_v be its completion at v; as K_v is a local field, we have the natural absolute value $|\cdot|_v := |\cdot|_{K_v}$ on K_v and on K. In this case, the product formula is true (this can be shown easily by considering the adèles of K, see Chapter IV, Section 4, Thm 5 of [111]). The height function on $\mathbb{Q}$ also generalises to K:

$$h_K(x) := \sum_v \log \max\{1, |x|_v\} = \sum_{v \text{ finite}} \log \max\{1, |x|_v\} + \sum_{\sigma\colon K \to \mathbb{C}} \log \max\{1, |\sigma(x)|\}, \tag{14.1.5}$$

for all $x \in K$. This function h_K is called the *height function of* K. For $K \to K'$ an extension of number fields, and for x in K, we have $h_{K'}(x) = (\dim_K K')h_K(x)$. Therefore one has the *absolute* height function h on $\overline{\mathbb{Q}}$ defined by:

$$h\colon \overline{\mathbb{Q}} \to \mathbb{R}, \quad h(x) = \frac{h_K(x)}{\dim_{\mathbb{Q}} K}, \tag{14.1.6}$$

where $K \subset \overline{\mathbb{Q}}$ is any number field that contains x.

14.2 Heights on projective spaces and on varieties

For $n \geq 0$ and for K a number field, we can define height functions on $\mathbb{P}^n(K)$ by:

$$h_K((x_0 : \cdots : x_n)) := \sum_v \log \max\{|x_0|, \ldots, |x_n|\}, \quad h(x) = \frac{h_K(x)}{\dim_{\mathbb{Q}} K}, \tag{14.2.1}$$

where v ranges through the set of all places of K. We note that it is because of the product formula that $h_K(x)$ is well-defined, and that this definition is compatible with our earlier definition of the height and absolute height on K if we view K as the complement of ∞ in $\mathbb{P}^1(K)$. The functions h on $\mathbb{P}^n(K)$ for varying K naturally induce the absolute height function on $\mathbb{P}^n(\overline{\mathbb{Q}})$.

A fundamental result, not difficult to prove, but too important to omit here (even though we will not use it), is Northcott's finiteness theorem.

14.2.2 Theorem (Northcott) *Let n, d and C be integers. Then*

$$\{x \in \mathbb{P}^n(\overline{\mathbb{Q}}) \mid h(x) \leq C \text{ and } \dim_{\mathbb{Q}}(\mathbb{Q}(x)) \leq d\}$$

is a finite set.

For a proof the reader is referred to Chapter 6 of [12], or to Section 2.4 of [98].

For any algebraic variety X embedded in a projective space $\mathbb{P}^n_K$ over some number field K, we get height functions h_K on $X(K)$ and h on $X(\overline{\mathbb{Q}})$ by restricting those from $\mathbb{P}^n$ to X.

14.3 The Arakelov perspective on height functions

We have just defined height functions h_K and h on a variety X over a number field K, embedded into some projective space $\mathbb{P}^n_K$. Such an embedding determines a line bundle $\mathcal{L}$ on X: the restriction of the line bundle $\mathcal{O}(1)$ of $\mathbb{P}^n_K$ that corresponds to homogeneous forms of degree 1, in the variables $x_0, \ldots, x_n$, say. The embedding of X into $\mathbb{P}^n_K$ is given by the global sections $s_0, \ldots, s_n$ of $\mathcal{L}$ obtained by restricting the global sections $x_0, \ldots, x_n$ to X. Now any finite set of generating global sections $t_0, \ldots, t_m$ of $\mathcal{L}$ determines a morphism $f\colon X \to \mathbb{P}^m_K$, inducing height functions $h_{K,f}$ and h_f via pullback along f. For f and f' two such morphisms, the difference $|h_f - h_{f'}|$ is bounded on $X(\overline{\mathbb{Q}})$ (see Theorem 3.1 of Chapter 6 of [12]). For this reason, one usually associates to a line bundle $\mathcal{L}$ on a variety X a class of height functions $f_{\mathcal{L}}$, i.e., an element in the set of functions $X(\overline{\mathbb{Q}}) \to \mathbb{R}$ modulo bounded functions; this map is then a morphism of groups on $\mathrm{Pic}(X)$: $f_{\mathcal{L}_1 \otimes \mathcal{L}_2} \equiv f_{\mathcal{L}_1} + f_{\mathcal{L}_2}$. However, in our situation, we cannot permit ourselves to work just modulo bounded functions on each variety, as we have infinitely many curves $X_1(l)$ to deal with.

There is a geometric way to associate to a line bundle a specific height function, not just a class of functions modulo bounded functions. For this, the contributions from the finite as well as the infinite places must be provided. Those from the finite places come from a model of X over the ring of integers O_K of K, i.e., an O_K-scheme X_{O_K} whose fibre over K is X, together with a line bundle $\mathcal{L}$ on X_{O_K} whose restriction to X is the line bundle that we had. The O_K-scheme X_{O_K} is required to be proper (e.g., projective). The contributions from the infinite places are provided by a hermitian metric (or inner product) on $\mathcal{L}$, a notion that we will briefly explain.

A hermitian metric on a locally free $\mathcal{O}_X$-module of finite rank $\mathcal{E}$ consists of a hermitian metric $\langle \cdot, \cdot \rangle_x$ on all $\mathbb{C}$-vector spaces $x^*\mathcal{E}$, where x runs through $X(\mathbb{C}) = \{x\colon \operatorname{Spec}(\mathbb{C}) \to X\}$. Each x in $X(\mathbb{C})$ induces a morphism $\operatorname{Spec}(\mathbb{C}) \to \operatorname{Spec}(K)$, i.e., an embedding of K into $\mathbb{C}$. Therefore, $X(\mathbb{C})$ is the disjoint union of the complex analytic varieties X_σ, indexed by the $\sigma\colon K \to \mathbb{C}$. A hermitian metric on $\mathcal{E}$ consists of hermitian metrics on all the holomorphic vector bundles $\mathcal{E}_\sigma$ that $\mathcal{E}$ induces on the X_σ. The metrics to be used are required to be continuous, i.e., for U open in X and s and t in $\mathcal{E}(U)$, the function $x \mapsto \langle s(x), t(x) \rangle_x$ on $U(\mathbb{C})$ must be continuous. Actually, the metrics that we will use will live on non-singular X, and will be required to be smooth (infinitely differentiable). Another condition that is usually imposed is a certain compatibility between the metrics at a point x in $X(\mathbb{C})$ and its complex conjugate $\overline{x}$. We do not give this condition in detail, but note that it will be fulfilled by the metrics that we will use. It is also customary to denote a hermitian metric $\langle \cdot, \cdot \rangle$ by its norm $\|\cdot\|$, given by $\|s\|^2 = \langle s, s \rangle$. Indeed, a suitable polarisation identity expresses the hermitian metric in terms of its norm. A pair $(\mathcal{E}, \|\cdot\|)$ of a locally free $\mathcal{O}_{X_{O_K}}$-module with a hermitian metric $\|\cdot\|$ is called a metrised vector bundle on X_{O_K}. Metrised vector bundles can be pulled back via morphisms $f\colon W_{O_K} \to X_{O_K}$ between O_K-schemes of the type considered.

An important example of the above is the case where $X = \operatorname{Spec}(K)$, just a point, and $X_{O_K} = \operatorname{Spec}(O_K)$. A metrised line bundle $(\mathcal{L}, \|\cdot\|)$ then corresponds to an invertible O_K-module, L, say, with hermitian metrics on the $L_\sigma := \mathbb{C} \otimes_{\sigma, O_K} L$. The *Arakelov degree* of $(\mathcal{L}, \|\cdot\|)$ is the real number defined by:

$$\deg(\mathcal{L}, \|\cdot\|) = \log \#(L/O_K s) - \sum_{\sigma\colon K \to \mathbb{C}} \log \|s\|_\sigma, \tag{14.3.1}$$

where s is any nonzero element of L (independence of the choice of s follows from the product formula). This definition should be compared to that of the degree of a line bundle on a smooth projective curve over a field: there one takes a rational section, and counts zeros and poles. The first term in (14.3.1) counts the zeros of s at the finite places. Interpreting this term in terms of valuations, and then norms, at the finite places, then leads to the second term which "counts" the "zeros" (or minus the "poles", for that matter) at the infinite places. For a finite extension

$K \to K'$, and $(\mathcal{L}, \|\cdot\|)$ on $\mathrm{Spec}(O_K)$ as above, the pullback $(\mathcal{L}', \|\cdot\|)$ to $\mathrm{Spec}(O_{K'})$ has degree $\dim_K K'$ times that on $\mathrm{Spec}(O_K)$.

We can now give the definition of the height given by a proper O_K-scheme X together with a hermitian line bundle $(\mathcal{L}, \|\cdot\|)$. Let x be in $X(K)$. Then, by the properness of X over O_K, x extends uniquely to an O_K-valued point, also denoted x, and one defines:

$$h_K(x) := \deg x^*(\mathcal{L}, \|\cdot\|). \tag{14.3.2}$$

The same method can be applied to get an absolute height $h\colon X(\overline{\mathbb{Q}}) \to \mathbb{R}$. For $K \to K'$ a finite extension, each x in $X(K')$ extends uniquely to an x in $X(O_{K'})$, and one defines:

$$h(x) := \frac{\deg x^*(\mathcal{L}, \|\cdot\|)}{\dim_{\mathbb{Q}} K'}. \tag{14.3.3}$$

It is not hard to verify that this height function h is in the class (modulo bounded functions) that is associated to X_K and $\mathcal{L}_K$ (without metric); see Proposition 7.2 of Chapter 6 of [12], or Theorem. 4.5 of Chapter V in [37]. In fact, for $X = \mathbb{P}^n_{O_K}$, and $\mathcal{L} = \mathcal{O}(1)$ with a suitable metric, the height h just defined is *equal* to the one defined in (14.2.1).

14.4 Arithmetic Riemann-Roch and intersection theory on arithmetic surfaces

The context in which we are going to apply Arakelov theory is that of smooth projective curves X over number fields K. In [3] Arakelov defined an intersection theory on the *arithmetic surfaces* associated to such curves, with the aim of proving certain results, known in the case of functions fields, in the case of number fields. The idea is to take a regular projective model $\mathcal{X}$ over $B := \mathrm{Spec}(O_K)$ of X, and try to develop an intersection theory on the surface $\mathcal{X}$, analogous to the theory that one has when K is a function field. If K is a function field over a finite field k, say, one gets a projective surface $\mathcal{X}$ over k, fibred over the nonsingular projective curve B over k that corresponds to K. On such a projective surface, intersecting with principal divisors gives zero, hence the intersection pairing factors through the Picard group of $\mathcal{X}$, the group of isomorphism classes of invertible $\mathcal{O}_{\mathcal{X}}$-modules. In the number field case one "compactifies" B by formally adding the infinite places of K; the product formula then means that principal divisors have degree zero. Instead of the Picard group of $\mathcal{X}$, one considers the group of isomorphism classes of certain metrised line bundles on $\mathcal{X}$, as defined above. In [39], see also Chapters II, III and I of [107], Faltings extended Arakelov's work by establishing results such as a Grothendieck-Riemann-Roch theorem in this context. Since then, Arakelov theory has been generalised by

Gillet and Soulé (see [103] and [40]). Below, we will use the theory as given in [39] and Chapter II of [107]. We start with some preparations concerning Riemann surfaces. The aim of this subsection is to give the arithmetic Riemann-Roch theorem as stated and proved by Faltings.

Let X be a compact Riemann surface of genus $g > 0$. The space of holomorphic differentials $\mathrm{H}^0(X, \Omega^1_X)$ carries a natural hermitian inner product:

$$(\omega, \eta) \mapsto \frac{i}{2} \int_X \omega \wedge \overline{\eta}. \tag{14.4.1}$$

Let $(\omega_1, \ldots, \omega_g)$ be an orthonormal basis with respect to this inner product. This leads to a positive $(1,1)$-form μ on X given by:

$$\mu = \frac{i}{2g} \sum_{k=1}^{g} \omega_k \wedge \overline{\omega_k}, \tag{14.4.2}$$

independent of the choice of orthonormal basis. Note that $\int_X \mu = 1$. We refer to [3] for a proof of the following proposition. Denote by $\mathcal{C}^\infty$ the sheaf of complex valued C^∞-functions on X, and by $\mathcal{A}^1$ the sheaf of complex C^∞ 1-forms on X. Recall that we have a tautological differential operator $d\colon \mathcal{C}^\infty \to \mathcal{A}^1$. It decomposes as $d = \partial + \overline{\partial}$ where, for any local C^∞ function f and any holomorphic local coordinate z, with real and imaginary parts x and y, one has $\partial f = \frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}) \cdot dz$ and $\overline{\partial} f = \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}) \cdot d\overline{z}$.

14.4.3 Proposition *For each a in X, there exists a unique real-valued $g_{a,\mu}$ in $\mathcal{C}^\infty(X - \{a\})$ such that the following properties hold:*

1. *we can write $g_{a,\mu} = \log|z - z(a)| + h$ in an open neighbourhood of a, where z is a local holomorphic coordinate and where h is a C^∞-function;*
2. *$\partial\overline{\partial} g_{a,\mu} = \pi i \mu$ on $X - \{a\}$;*
3. *$\int_X g_{a,\mu}\mu = 0$.*

We refer to μ and the $g_{a,\mu}$ as the Arakelov $(1,1)$-form and the Arakelov-Green function, respectively. We note that Stokes' theorem implies $g_{a,\mu}(b) = g_{b,\mu}(a)$ for all a and b in X. The Arakelov-Green functions determine certain metrics, called admissible metrics, on all line bundles $\mathcal{O}_X(D)$, where D is a divisor on X, as well as on the holomorphic cotangent bundle Ω^1_X. To start, consider line bundles of the form $\mathcal{O}_X(a)$ with a a point in X (the general case with D follows by taking tensor products). Let s be the tautological section of $\mathcal{O}_X(a)$, i.e. the constant function 1. We define a smooth hermitian metric $\|\cdot\|_{\mathcal{O}_X(a)}$ on $\mathcal{O}_X(a)$ by putting $\log\|s\|_{\mathcal{O}_X(a)}(b) = g_{a,\mu}(b)$ for any b in X. By property 2 of the Arakelov-Green function, the curvature form of $\mathcal{O}_X(a)$ is

equal to μ. Second, it is clear that the functions $g_{a,\mu}$ can be used to put a hermitian metric on the line bundle $\mathcal{O}_{X\times X}(\Delta_X)$, where Δ_X is the diagonal on $X \times X$, by putting $\log \|s\|(a,b) = g_{a,\mu}(b)$ for the tautological section s of $\mathcal{O}_{X\times X}(\Delta_X)$. Restricting to the diagonal we have a canonical adjunction isomorphism $\mathcal{O}_{X\times X}(-\Delta_X)|_{\Delta_X} \xrightarrow{\sim} \Omega^1_X$. We define a hermitian metric $\|\cdot\|_{\mathrm{Ar}}$ on Ω^1_X by insisting that this adjunction isomorphism be an isometry. It is proved in [3] that this gives a smooth hermitian metric on Ω^1_X, and that its curvature form is a multiple of μ. From now on we will work with these metrics on $\mathcal{O}_X(P)$ and Ω^1_X (as well as on tensor product combinations of them) and refer to them as Arakelov metrics. A metrised line bundle $\mathcal{L}$ in general is called admissible if, up to a constant scaling factor, it is isomorphic to one of the admissible bundles $\mathcal{O}_X(D)$, or, equivalently, if its curvature form $\mathrm{curv}(\mathcal{L})$ is a multiple of μ. Note that then necessarily we have $\mathrm{curv}(\mathcal{L}) = (\deg \mathcal{L}) \cdot \mu$ by Stokes' theorem.

For any admissible line bundle $\mathcal{L}$, Faltings defines a certain metric on the determinant of cohomology $\lambda(\mathcal{L}) = \det \mathrm{H}^0(X,\mathcal{L}) \otimes \det \mathrm{H}^1(X,\mathcal{L})^\vee$ of the underlying line bundle. This metric is the unique metric satisfying a set of axioms. We recall these axioms (cf. [39], Theorem 1): (i) any isometric isomorphism $\mathcal{L}_1 \xrightarrow{\sim} \mathcal{L}_2$ of admissible line bundles induces an isometric isomorphism $\lambda(\mathcal{L}_1) \xrightarrow{\sim} \lambda(\mathcal{L}_2)$; (ii) if we scale the metric on $\mathcal{L}$ by a factor α, the metric on $\lambda(\mathcal{L})$ is scaled by a factor $\alpha^{\chi(\mathcal{L})}$, where $\chi(\mathcal{L}) = \deg \mathcal{L} - g + 1$ is the Euler-Poincaré characteristic of $\mathcal{L}$; (iii) for any divisor D and any point P on X, the exact sequence $0 \to \mathcal{O}_X(D-P) \to \mathcal{O}_X(D) \to P_*P^*\mathcal{O}_X(D) \to 0$ induces an isometry $\lambda(\mathcal{O}_X(D)) \xrightarrow{\sim} \lambda(\mathcal{O}_X(D-P)) \otimes P^*\mathcal{O}_X(D)$; (iv) for $\mathcal{L} = \Omega^1_X$, the metric on $\lambda(\mathcal{L}) \cong \det \mathrm{H}^0(X,\Omega^1_X)$ is defined by the hermitian inner product $(\omega,\eta) \mapsto (i/2)\int_X \omega \wedge \overline{\eta}$ on $\mathrm{H}^0(X,\Omega^1_X)$. In particular, for an admissible line bundle $\mathcal{L}$ of degree $g-1$, the metric on the determinant of cohomology $\lambda(\mathcal{L})$ is independent of scaling.

It was proved by Faltings that we can relate the metric on the determinant of cohomology to theta functions on the Jacobian of X. Let $\mathbb{H}_g$ be the Siegel upper half space of complex symmetric g-by-g-matrices with positive definite imaginary part. Let τ in $\mathbb{H}_g$ be the period matrix associated to a symplectic basis of $\mathrm{H}_1(X,\mathbb{Z})$ and consider the analytic Jacobian $J_\tau(X) = \mathbb{C}^g/(\mathbb{Z}^g + \tau\mathbb{Z}^g)$ associated to τ. On $\mathbb{C}^g$ one has a theta function $\vartheta(z;\tau) = \sum_{n\in\mathbb{Z}^g} \exp(\pi i^t n\tau n + 2\pi i^t nz)$, giving rise to a reduced effective divisor Θ_0 and a line bundle $\mathcal{O}(\Theta_0)$ on $J_\tau(X)$. Now consider on the other hand the set $\mathrm{Pic}_{g-1}(X)$ of divisor classes of degree $g-1$ on X. It comes with a canonical subset Θ given by the classes of effective divisors. A fundamental theorem of Abel-Jacobi-Riemann says that there is a canonical bijection $\mathrm{Pic}_{g-1}(X) \xrightarrow{\sim} J_\tau(X)$ mapping Θ onto Θ_0. As a result, we can equip $\mathrm{Pic}_{g-1}(X)$ with the structure of a compact complex manifold, together with a divisor Θ and a line bundle $\mathcal{O}(\Theta)$.

The function ϑ is not well-defined on $\mathrm{Pic}_{g-1}(X)$ or $J_\tau(X)$. We can remedy this by putting $\|\vartheta\|(z;\tau) = (\det \Im(\tau))^{1/4} \exp(-\pi^t y(\Im(\tau))^{-1}y)|\vartheta(z;\tau)|$, with $y = \Im(z)$. One can check that

$\|\vartheta\|$ descends to a function on $J_\tau(X)$. By our identification $\mathrm{Pic}_{g-1}(X) \xrightarrow{\sim} J_\tau(X)$ we obtain $\|\vartheta\|$ as a function on $\mathrm{Pic}_{g-1}(X)$. It can be checked that this function is independent of the choice of τ. Note that $\|\vartheta\|$ gives a canonical way to put a metric on the line bundle $\mathcal{O}(\Theta)$ on $\mathrm{Pic}_{g-1}(X)$. For any line bundle $\mathcal{L}$ of degree $g-1$ there is a canonical isomorphism $\lambda(\mathcal{L}) \xrightarrow{\sim} \mathcal{O}(-\Theta)[\mathcal{L}]$, the fibre of $\mathcal{O}(-\Theta)$ at the class in $\mathrm{Pic}_{g-1}(X)$ determined by $\mathcal{L}$. Faltings proves that when we give both sides the metrics discussed above, the norm of this isomorphism is a constant independent of $\mathcal{L}$; he writes it as $\exp(\delta(X)/8)$. This invariant appears in the Noether formula, see below.

We will now turn to intersections on an arithmetic surface. For us, an arithmetic surface is a proper, flat morphism $p\colon \mathcal{X} \to B$ with $\mathcal{X}$ a regular scheme, with B the spectrum of the ring of integers O_K in a number field K, and with generic fibre a geometrically connected and smooth curve X/K. We say that $\mathcal{X}$ is of genus g if the generic fibre is of genus g. We will always assume that p is a semi-stable curve, unless explicitly stated otherwise. After extending the base field if necessary, any geometrically connected, smooth proper curve X/K of positive genus with K a number field is the generic fibre of a unique semi-stable arithmetic surface.

An Arakelov divisor on $\mathcal{X}$ is a finite formal integral linear combination of integral closed subschemes of codimension 1 of $\mathcal{X}$ plus a contribution $\sum_\sigma \alpha_\sigma \cdot F_\sigma$ running over the complex embeddings of K. Here α_σ is a real number, and the symbols F_σ correspond to the compact Riemann surfaces X_σ obtained by base changing X/K to $\mathbb{C}$ via σ. We have an $\mathbb{R}$-valued intersection product $(\cdot,\cdot)$ for such divisors, respecting linear equivalence. The notion of principal divisor is given as follows: let f be a non-zero rational function in $K(X)$, then $(f) = (f)_{\mathrm{fin}} + (f)_{\mathrm{inf}}$ with $(f)_{\mathrm{fin}}$ the usual Weil divisor of f on $\mathcal{X}$, and with $(f)_{\mathrm{inf}} = \sum_\sigma v_\sigma(f)\cdot F_\sigma$ with $v_\sigma(f) = -\int_{X_\sigma} \log|f|_\sigma \mu_\sigma$. For a list of properties of this intersection product we refer to [3], [39] or Chapter II of [107].

It is proved in [3] that the group of linear equivalence classes of Arakelov divisors is canonically isomorphic to the group $\widehat{\mathrm{Pic}}(\mathcal{X})$ of isometry classes of admissible line bundles on $\mathcal{X}$. By an admissible line bundle on $\mathcal{X}$ we mean the datum of a line bundle $\mathcal{L}$ on $\mathcal{X}$, together with admissible metrics on the restrictions $\mathcal{L}_\sigma$ of $\mathcal{L}$ to the X_σ. In particular we have a canonical admissible line bundle $\omega_{\mathcal{X}/B}$ whose underlying line bundle is the relative dualising sheaf of p. In many situations it is convenient to treat intersection numbers from the point of view of admissible line bundles.

For example, if $P\colon B \to \mathcal{X}$ is a section of p, and D is a divisor on $\mathcal{X}$, the pull-back $P^*\mathcal{O}_\mathcal{X}(D)$ is a metrised line bundle on B, and we have $(D,P) = \deg P^*\mathcal{O}_\mathcal{X}(D)$, where the degree $\deg$ of a metrised line bundle is as defined in (14.3.1). As a second example, we mention that by definition of the metric on $\omega_{\mathcal{X}/B}$, we have for each section $P\colon B \to \mathcal{X}$ of p an adjunction formula $(P, P+\omega_{\mathcal{X}/B}) = 0$.

For an admissible line bundle $\mathcal{L}$ on $\mathcal{X}$, we have the notion of determinant of cohomology on B, in this context denoted by $\det \mathrm{R}p_*\mathcal{L}$ (see Chapter II of [107]). By using the description

above for its metrisation over the complex numbers, we obtain the determinant of cohomology on B as a metrised line bundle. One of its most important features is a metrised Riemann-Roch formula (cf. [39], Theorem 3), also called *arithmetic Riemann-Roch formula*:

$$\deg \det \mathrm{R}p_*\mathcal{L} = \frac{1}{2}(\mathcal{L}, \mathcal{L} \otimes \omega_{\mathcal{X}/B}^{-1}) + \deg \det p_*\omega_{\mathcal{X}/B} \tag{14.4.4}$$

for any admissible line bundle $\mathcal{L}$ on $\mathcal{X}$.

The term $\deg \det p_*\omega_{\mathcal{X}/B}$ is also known as the *Faltings height* of X, the definition of which we will now recall. We let J_K be the Jacobian variety of X, and J its Néron model over B. Then we have the locally free O_K-module $\mathrm{Cot}_0(J) := 0^*\Omega^1_{J/O_K}$ of rank g, and hence the invertible O_K-module of rank one

$$\omega_J := \bigwedge^g 0^* \mathrm{Cot}_0(J).$$

For each $\sigma\colon K \to \mathbb{C}$ we have the scalar product on $\mathbb{C} \otimes_{O_K} \omega_J$ given by

$$\langle \omega | \eta \rangle_\sigma = (i/2)^g (-1)^{g(g-1)/2} \int_{J_\sigma(\mathbb{C})} \omega \wedge \overline{\eta}.$$

The Faltings height $h_K(X)$ is then defined to be the Arakelov degree of this metrised line bundle:

$$h_K(X) = \deg(\omega_J), \tag{14.4.5}$$

and the absolute Faltings height (also called stable Faltings height) $h_{\mathrm{abs}}(X)$ of X is defined as:

$$h_{\mathrm{abs}}(X) = [K : \mathbb{Q}]^{-1} \deg(\omega_J). \tag{14.4.6}$$

We remark that the stable Faltings height of X does not change after base change to larger number fields; that is why it is called stable. Therefore, $h_{\mathrm{abs}}(X)$ can be computed from any model of X over a number field as long as that model has stable reduction over the ring of integers of that number field.

As $\mathcal{X} \to B$ is semi-stable, a result of Raynaud gives that the connected component of 0 of J is the Picard scheme $\mathrm{Pic}^0_{\mathcal{X}/B}$, whose tangent space at 0 is $\mathrm{R}^1 p_* \mathcal{O}_{\mathcal{X}}$. Therefore, $\mathrm{Cot}_0(J)$ is the same as $p_*\omega$, as locally free O_K-modules. A simple calculation (see lemme 3.2.1 in Chapter I of [107]) shows that, with these scalar products, ω_J and $\det p_*\omega$ are the same as metrised O_K-modules. Therefore we have:

$$h_K(X) = \deg(\omega_J) = \deg \det p_*\omega. \tag{14.4.7}$$

One may derive from (14.4.4) the following projection formula: let E be a metrised line bundle on B, and $\mathcal{L}$ an admissible line bundle on $\mathcal{X}$. Then the formula:

$$\deg \det \mathrm{R}p_*(\mathcal{L} \otimes p^*E) = \deg \det \mathrm{R}p_*\mathcal{L} + \chi(\mathcal{L}) \cdot \deg E \tag{14.4.8}$$

holds. Here again $\chi(\mathcal{L})$ is the Euler-Poincaré characteristic of $\mathcal{L}$ on the fibres of p.

15 Applying Arakelov theory

In this section we start applying Arakelov theory in order to derive a bound for the height of the coefficients of the polynomials P_l as in (13.11). We proceed in a few steps. The first step, taken in Section 15.1, is to relate the height of the $f(Q_{x,i})$ as in Section 13 to intersection numbers on X_l. The second step, taken in Section 15.2, is to get some control on the difference of the divisors D and D'_x as in (11.5). Certain intersection numbers concerning this difference are bounded in Theorem 15.2.5, in terms of a number of invariants in the Arakelov theory on modular curves X_l. These invariants will then be bounded in terms of l in Sections 16, 17, and 18. Finally, in Section 19, the height of the coefficients of the P_l will be bounded.

15.1 Relating heights to intersection numbers

We pick up the notation as at the end of Section 13, so we have a modular curve $X_l = X_1(5l)_{\mathbb{Q}(\zeta_l)}$, non-constant morphisms b, x, $y\colon X_l \to \mathbb{P}^1_{\mathbb{Q}(\zeta_l)}$, and certain divisors $D''_x = \sum_{i=1}^{d_l} Q_{x,i}$ on X_l that have support outside the cusps. It is our objective in this subsection to link the absolute height $h(f(Q_{x,i}))$ of an $f(Q_{x,i})$, where f is either b, x or y, to certain quantities coming from Arakelov intersection theory.

15.1.1 Theorem *There are integers c_1 and c_2 such that the following holds. Let $l > 5$ be a prime. Let K be a number field containing $\mathbb{Q}(\zeta_{5l})$ and such that $Q_{x,i}$ is defined over K. Let $\mathcal{X}$ be the minimal regular model of X_l over K. Take $f \in \{b, x, y\}$. Then we have:*

$$h(f(Q_{x,i})) \leq \frac{1}{[K:\mathbb{Q}]}\left((\overline{Q_{x,i}}, \overline{f^*\infty})_{\mathcal{X}} + c_1 l^2 \sum_{\sigma} \sup_{X_\sigma} g_\sigma\right) + c_2 l\,.$$

Here σ runs through the embeddings of K into $\mathbb{C}$ and g_σ is the Arakelov-Green function on $X_{l,\sigma}$.

In the next sections we shall derive polynomial bounds for the terms $(\overline{Q_{x,i}}, \overline{f^*\infty})_{\mathcal{X}}$ and $\sum_\sigma \sup_{X_\sigma} g_\sigma$ in the above estimate.

Theorem 15.1.1 will be derived from Theorem 15.1.2 below, which states a fairly general result. We start with a definition. Let K be a number field and consider $\mathbb{P}^1_{O_K}$. For any section $\overline{P}$ in $\mathbb{P}^1_{O_K}(O_K)$ we define by $(\overline{P}, \overline{\infty})_{\mathbb{P}^1}$ the degree (see (14.3.1)) of $\overline{P}^* O_{\mathbb{P}^1}(\overline{\infty})$, where $O_{\mathbb{P}^1}(\overline{\infty})$ has the Fubini-Study metric, i.e. the metric $\|\cdot\|_{\mathbb{P}^1}$ given by:

$$\|1\|_{\mathbb{P}^1}(x_0 : x_1) := \frac{|x_0|}{(|x_0|^2 + |x_1|^2)^{1/2}}$$

over $\mathbb{P}^1_{\mathbb{C}}$. Here 1 is the tautological section of $O_{\mathbb{P}^1}(\infty)$.

15.1.2 Theorem *Let X be a geometrically irreducible, smooth and complete curve over a number field K and let $\mathcal{X}$ be a semi-stable model of X over the ring of integers O_K of K. Suppose that we have a non-constant morphism $f\colon X \to \mathbb{P}^1_K$ and a K-rational point Q of X with $f(Q) \neq \infty$. Assume the following: both the zero divisor $\mathrm{div}(f)_+$ and the polar divisor $\mathrm{div}(f)_-$ of f on X have as their support only K-rational closed points, and the Zariski closure of $\mathrm{Supp}(\mathrm{div}(f)_+) \cup \mathrm{Supp}(\mathrm{div}(f)_-)$ in $\mathcal{X}$ is étale over O_K. For any closed point s of $\mathrm{Spec}(O_K)$, denote by $m_s(f)$ the supremum of the multiplicities of $\mathrm{div}(f)_-$ on $\mathcal{X}$ along the irreducible components of the fibre at s of $\mathcal{X}$. Then we have the inequality:*

$$(\overline{f(Q)}, \overline{\infty})_{\mathbb{P}^1} \leq (\overline{Q}, \overline{f^*\infty})_{\mathcal{X}} + \deg f \sum_{\sigma} \sup_{X_\sigma} g_\sigma + \sum_s m_s(f) \log \#k(s)\,.$$

Here the first sum runs over the embeddings of K into $\mathbb{C}$, and the second sum runs over the closed points of $\mathrm{Spec}(O_K)$.

Proof Note that the locus of indeterminacy of f on $\mathcal{X}$ consists of finitely many closed points. This implies that there exists a blow-up $\tilde{\mathcal{X}} \to \mathcal{X}$ of $\mathcal{X}$ such that f extends to a regular map $f\colon \tilde{\mathcal{X}} \to \mathbb{P}^1_{O_K}$. For any such $\tilde{\mathcal{X}}$ we have by construction:

$$\begin{aligned}(\overline{f(Q)}, \overline{\infty})_{\mathbb{P}^1} &= \deg \overline{f(Q)}^*(O_{\mathbb{P}^1}(\overline{\infty}), \|\cdot\|_{\mathbb{P}^1}) \\ &= \deg \overline{Q}^* f^*(O_{\mathbb{P}^1}(\overline{\infty}), \|\cdot\|_{\mathbb{P}^1}) \\ &= \deg \overline{Q}^*(O_{\tilde{\mathcal{X}}}(f^*\overline{\infty}), \|\cdot\|_{\mathbb{P}^1})\,.\end{aligned}$$

If we let $\|\cdot\|_X$ denote the canonical Arakelov metric on $O_X(f^*\infty)$ then we can write:

$$\begin{aligned}\deg \overline{Q}^*(O_{\tilde{\mathcal{X}}}(f^*\overline{\infty}), \|\cdot\|_{\mathbb{P}^1}) &= \deg \overline{Q}^*(O_{\tilde{\mathcal{X}}}(f^*\overline{\infty}), \|\cdot\|_X \cdot \frac{\|\cdot\|_{\mathbb{P}^1}}{\|\cdot\|_X}) \\ &= \deg \overline{Q}^*(O_{\tilde{\mathcal{X}}}(f^*\overline{\infty}), \|\cdot\|_X) - \sum_{\sigma} \log((\frac{\|\cdot\|_{\mathbb{P}^1}}{\|\cdot\|_X})(Q_\sigma)) \\ &= (\overline{Q}, f^*\overline{\infty})_{\tilde{\mathcal{X}}} + \sum_{\sigma} \log((\frac{\|\cdot\|_X}{\|\cdot\|_{\mathbb{P}^1}})(Q_\sigma))\,.\end{aligned}$$

A bound for $\log((\frac{\|\cdot\|_X}{\|\cdot\|_{\mathbb{P}^1}})(Q_\sigma))$ follows by testing on the tautological section 1, giving:

$$\log\|1\|_X(Q_\sigma) - \log\|1\|_{\mathbb{P}^1}(Q_\sigma) = g_\sigma(f^*\infty, Q_\sigma) - \frac{1}{2}\log(|f(Q)|^2 + 1) \leq (\deg f) \sup_{X_\sigma} g_\sigma\,.$$

This accounts for the second term in the bound of the theorem. We are finished once we prove that $(\overline{Q}, f^*\overline{\infty})_{\tilde{\mathcal{X}}} - (\overline{Q}, \overline{f^*\infty})_{\mathcal{X}}$ is bounded by $\sum_s m_s(f) \log \#k(s)$ for a particular choice of $\tilde{\mathcal{X}}$. (The usual projection formula shows that in fact $(\overline{Q}, f^*\overline{\infty})_{\tilde{\mathcal{X}}}$ is independent of the choice

of $\tilde{\mathcal{X}}$.) On any $\tilde{\mathcal{X}}$ we write $f^*\overline{\infty}$ as a sum $f^*\overline{\infty} = (f^*\overline{\infty})_{\mathrm{hor}} + (f^*\overline{\infty})_{\mathrm{vert}}$ of a horizontal and a vertical part. Note that $(f^*\overline{\infty})_{\mathrm{hor}} = \overline{f^*\infty}$, Zariski closure now taken in $\tilde{\mathcal{X}}$. Since the local intersection multiplicities of $\overline{Q}$ and $\overline{f^*\infty}$ do not go up when passing from $\mathcal{X}$ to $\tilde{\mathcal{X}}$, we have $(\overline{Q}, (f^*\overline{\infty})_{\mathrm{hor}})_{\tilde{\mathcal{X}}} = (\overline{Q}, \overline{f^*\infty})_{\tilde{\mathcal{X}}} \leq (\overline{Q}, \overline{f^*\infty})_{\mathcal{X}}$ and thus we are reduced to proving that $(\overline{Q}, (f^*\overline{\infty})_{\mathrm{vert}})_{\tilde{\mathcal{X}}}$ is bounded from above by $\sum_s m_s(f) \log \#k(s)$ for a particular choice of $\tilde{\mathcal{X}}$.

We exhibit a specific blow-up, and we calculate which multiplicities f acquires along the irreducible components of the vertical fibres of this blow-up. Note that the locus of indeterminacy of f on $\mathcal{X}$ consists precisely of the closed points of $\mathcal{X}$ where an irreducible component from the zero divisor $\mathrm{div}(f)_+$ of f on $\mathcal{X}$ and an irreducible component from its polar divisor $\mathrm{div}(f)_-$ meet. Now since by assumption the Zariski closure of $\mathrm{Supp}(\mathrm{div}(f)_+) \cup \mathrm{Supp}(\mathrm{div}(f)_-)$ in $\mathcal{X}$ is étale over O_K, an intersection of two irreducible components C and C' of $\mathcal{X}$ that have different sign in $\mathrm{div}(f)$ on $\mathcal{X}$ can only occur when at least one of C, C' is vertical. In such points where this happens we have to perform a sequence of successive blowings-up until a component arises with multiplicity 0 for f, so that the components with positive multiplicities and the components with negative multiplicities are separated from each other.

We begin by observing that it will cause no harm if we pass to a finite extension $K \to K'$. Indeed, both the left hand side and the right hand side of the inequality that we wish to prove get multiplied by $[K' : K]$ if we do this. Here is why: for the terms $(\overline{f(Q)}, \overline{\infty})_{\mathbb{P}^1}$ and $(\overline{Q}, \overline{f^*\infty})$ the scaling by a factor $[K' : K]$ follows from general properties of the Arakelov intersection product, cf. [39], p. 404 for example. (Note that it is understood that over K', intersection products are taken on the minimal resolution of the pullback of the model $\mathcal{X}$.) That the term $\sum_\sigma \sup_{X_\sigma} g_\sigma$ scales by a factor $[K' : K]$ is obvious. Finally, fix a closed point s of $\mathrm{Spec}(O_K)$ and let s' be any closed point of $\mathrm{Spec}(O_{K'})$ above it. Denoting by $e_{s'}$ the ramification index of s' over s and by $f_{s'}$ the degree of the residue field extension of s' over s, we see that for any s' above s, the integer $m_s(f)$ gets multiplied by $e_{s'}$, and the number $\log \#k(s)$ gets multiplied by $f_{s'}$. Using that $\sum_{s'} e_{s'} f_{s'} = [K' : K]$, the sum running over the closed points s' above s, we see finally that also the term $\sum_s m_s(f) \log \#k(s)$ gets multiplied by $[K' : K]$.

Starting with $\mathcal{X}$ over O_K, we first do the following. Let x be a closed point on $\mathcal{X}$ that is the intersection of a vertical component C and a horizontal component C' having non-zero multiplicities m and m' for f that have different sign. After blowing up in x, we obtain an exceptional divisor E whose multiplicity for f is $m + m'$. We have two distinguished points on E, one lying on the strict transform of C, and one lying on the strict transform of C'. At exactly one of them there is a sign change for the multiplicities, or $m + m' = 0$. If a sign change happens at the double point lying on the strict transform of C', then we repeat the process. If a sign change happens at the double point lying on the strict transform of C or if $m + m' = 0$, we stop, and continue with a new point x', if available.

We end up with a blow-up $\mathcal{X}' \to \mathcal{X}$ such that an intersection of two irreducible components C, C' that have different sign in $\mathrm{div}(f)$ on $\mathcal{X}'$ only occurs for C, C' both vertical. For this, we did not yet need to extend the ground field K. In order to continue, we note the following. Suppose that we have a closed point x on the model $\mathcal{X}'$ of X over O_K which is a double point of a vertical fibre, and two irreducible components C, C' of that vertical fibre pass through x, having non-zero multiplicities m and m' for f that differ in sign. Assume that $K \to K'$ is a Galois extension that ramifies over the image of x in $\mathrm{Spec}(O_K)$, with a ramification index e that is a multiple of $m-m'$. In passing to the minimal resolution $\tilde{\mathcal{X}}$ of $\mathcal{X}'_{O_{K'}}$, the point x gets replaced by a chain of $e-1$ projective lines of self-intersection -2. The multiplicities of f along these components change in e steps from em to em', so that the steps are $m-m'$ and a multiplicity 0 will appear somewhere, because $m-m'$ is a divisor of em.

Thus we see how we can reach our goal: take a Galois extension $K \to K'$ that ramifies as specified above over the images in $\mathrm{Spec}(O_K)$ of the double points where components meet with a different sign for f. (This is always possible.) By our remarks above, it suffices to prove the inequality over K'. By construction, the morphism f extends over the model $\tilde{\mathcal{X}}$ that arises in this way. Moreover, it follows from the construction that for s' a closed point of $\mathrm{Spec}(O_{K'})$ and s its image in $\mathrm{Spec}(O_K)$ we have $m_{s'}(f) \leq e_{s'} m_s(f)$. Hence the sum of the local intersection numbers $(\overline{Q}, (f^*\overline{\infty})_{\mathrm{vert}})_{s'}$ for all s' over s is bounded from above by $[K' : K] m_s(f) \log \#k(s)$. This is what we needed to prove. □

Proof [Proof of Theorem 15.1.1] If for P in $\mathbb{P}^1_{O_K}$ we put $h'(P) = (P, \infty)/[K : \mathbb{Q}]$ then we have $h(P) \leq h'(P)$. In order to bound $h'(P)$ from above we want to apply Theorem 15.1.2. It follows from the definitions of the morphisms b, x and y that for each $f \in \{b, x, y\}$ both the zero divisor $\mathrm{div}(f)_+$ and the polar divisor $\mathrm{div}(f)_-$ of f on X_l have as their support only K-rational closed points, namely, cusps. In particular, we never have $f(Q_{x,i}) = \infty$, by construction of D''_x. We have also seen that the Zariski closure in $\mathcal{X}$ of $\mathrm{Supp}(\mathrm{Cusps})$ is étale over O_K, and hence the same holds for the Zariski closure in $\mathcal{X}$ of $\mathrm{Supp}(\mathrm{div}(f)_+) \cup \mathrm{Supp}(\mathrm{div}(f)_-)$. Theorem 15.1.2 now gives us that:

$$h(f(Q_{x,i})) \leq \frac{1}{[K : \mathbb{Q}]} \left((\overline{Q_{x,i}}, \overline{f^*\infty})_{\mathcal{X}} + \deg f \sum_{\sigma} \sup_{X_\sigma} g_\sigma + \sum_s m_s(f) \log \#k(s) \right)$$

with $m_s(f)$ the supremum of the multiplicities of $\mathrm{div}(f)_-$ on $\mathcal{X}$ along the irreducible components of the fibres of $\mathcal{X}$ at s. We are done if we can prove that $\deg f$ is bounded by a constant times l^2, and that $\frac{1}{[K:\mathbb{Q}]} \sum_s m_s(f) \log \#k(s)$ is bounded by a constant times l. For this we need information on the divisors $\mathrm{div}(b)$, $\mathrm{div}(x)$ and $\mathrm{div}(y)$ on $\mathcal{X}$.

We start with working over $\mathbb{Q}(\zeta_{5l})$. From the discussion in Section 12 we recall that there is a fine moduli scheme $Y_1(5l)_{\mathbb{Z}[\zeta_{5l}]}$ over $\mathbb{Z}[\zeta_{5l}]$ of elliptic curves with balanced level structure

(terminology from [61]). Let $E(5l) \to Y_1(5l)_{\mathbb{Z}[\zeta_{5l}]}$ be the universal elliptic curve and let P_5, P_l be the tautological points of order 5 and l. From Section 13 we recall the elliptic curve $E \to Y_1(5)'$ with $Y_1(5)' = \mathrm{Spec}(\mathbb{Z}[b, 1/\mathrm{discr}(E)])$. The elliptic curve $E(5l) \to Y_1(5l)_{\mathbb{Z}[\zeta_{5l},1/5]}$ arises from $E \to Y_1(5)'$ by a unique base change. Via P_l the scheme $Y_1(5l)_{\mathbb{Z}[\zeta_{5l},1/5]}$ is mapped into E, and thus we get the functions b, x and y on it. We claim that they are invertible regular functions on $Y_1(5l)_{\mathbb{Z}[\zeta_{5l},1/5]}$. Indeed, first of all b is invertible on $Y_1(5)'$, hence it is invertible on $Y_1(5l)_{\mathbb{Z}[\zeta_{5l},1/5]}$. Next, x and y have polar divisor on E supported on the image of the zero-section (with multiplicity 2 and 3, respectively), and their zero divisors are certain linear combinations of points of order 5 on E (to be precise, $\mathrm{div}(x)_+ = (0,0) + (0,-b)$ and $\mathrm{div}(y)_+ = 2 \cdot (0,0) + (-b,0)$). Looking at the image under P_l of $Y_1(5l)_{\mathbb{Z}[\zeta_{5l},1/5]}$ in E, we see that it is disjoint from both the zero-section and the points of order 5. We conclude that $\mathrm{div}(b)$, $\mathrm{div}(x)$ and $\mathrm{div}(y)$ on $X_1(5l)_{\mathbb{Z}[\zeta_{5l}]}$ are certain linear combinations of the irreducible components of the closed subschemes Cusps, $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$ and $X_1(5l)_{\mathbb{F}_l[\zeta_5]}$. In order to find these linear combinations, we just consider our functions one by one and examine their multiplicities along the irreducible components that we have isolated.

We start with b, and calculate its multiplicities along the irreducible components of Cusps first. It is sufficient to study the situation over $\mathbb{C}$, and here we can make a beginning by looking at $\mathrm{div}(b)$ on $X_1(5)_{\mathbb{C}}$. From the equations in Proposition 13.7 we obtain that over the cusp 0 of $X_0(5)_{\mathbb{C}}$ lie two cusps, say c_1 and c_2, with c_1, say, corresponding to the 5-gon with the tautological point of order 5 being on a component adjacent to the connected component of 0, and the other, c_2, corresponding to the 5-gon with the tautological point of order 5 being on a component that is not adjacent to the connected component of 0. We have $\mathrm{div}(b) = \pm(c_1 - c_2)$ on $X_1(5)_{\mathbb{C}}$; we could compute the exact sign but that is not important for us. Now note that the universal elliptic curve $E \to Y_1(5)_{\mathbb{C}}$ extends naturally to a generalised elliptic curve $\pi\colon \overline{E} \to X_1(5)_{\mathbb{C}}$. The inclusion $Y_1(5l)_{\mathbb{C}} \to \overline{E}$ given by the map P_l extends to a map $\varphi\colon X_1(5l)_{\mathbb{C}} \to \overline{E}$, and the composite $\pi \cdot \varphi$ is just the forgetful map $X_1(5l)_{\mathbb{C}} \to X_1(5)_{\mathbb{C}}$. Pulling back the divisor $c_1 - c_2$ we get plus or minus the divisor of b on $X_1(5l)_{\mathbb{C}}$; the multiplicities are just the ramification indices above the cusps c_1 and c_2. Since these are in $\{1, l\}$, we obtain that the multiplicities of b along the irreducible components of Cusps are just 1 or l in absolute value.

Next we calculate the multiplicities of b along the irreducible components of $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$ and $X_1(5l)_{\mathbb{F}_l[\zeta_5]}$. The structure of a connected component of the scheme $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$ is as follows: it consists of two components, one having an open part where P_5 has order 1, and one component having an open part where P_5 has order 5. These two components intersect (transversally) in the supersingular points. A similar description, with the roles of 5 and l interchanged, holds for the connected components of the scheme $X_1(5l)_{\mathbb{F}_l[\zeta_5]}$.

We denote by Γ the union of the components over $\mathbb{F}_5$ where P_5 has order 1. By construction

of the scheme $Y_1(5)'$ there is a forgetful map $X_1(5l)_{\mathbb{Z}[\zeta_{5l}]} - \Gamma - \mathrm{Supp}(\mathrm{Cusps}) \to Y_1(5)'$. Since b is invertible on $Y_1(5)'$, the same holds for b along the irreducible components of $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$ and $X_1(5l)_{\mathbb{F}_l[\zeta_5]}$, except possibly for the components in Γ. But the multiplicity of b along such a component is then also zero, as can be seen by the following argument. Let $C \cup C'$ be a connected component of $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$, with the irreducible component C corresponding to P_5 having order 1. All the horizontal components of $\mathrm{div}(b)$ on $X_1(5l)_{\mathbb{Z}[\zeta_{5l}]}$ specialise to C'. We know that b has multiplicity 0 along C' and hence it restricts to a non-trivial rational function, also denoted b, on C'. The degree of b on C' is zero, or equivalently $m(C', C) + (C', \mathrm{div}(b)_{\mathrm{hor}}) = 0$, where m is the multiplicity of b along C. Now, since $(C', \mathrm{div}(b)_{\mathrm{hor}})$ is zero and (C', C) isn't, we get $m = 0$.

We continue with the functions x and y. Again we start with the cusps, and again it will be sufficient to work over $\mathbb{C}$. The functions x and y have some multiplicities along the irreducible components of the fibres of $\overline{E} \to X_1(5)_{\mathbb{C}}$ above the cusps, but note that these do not depend on l. Let a be an upper bound for the absolute values of these multiplicities. By an argument analogous to the argument above with b, we find that the cusps of $X_1(5l)_{\mathbb{C}}$ arise in the divisors of x and y on $X_1(5l)_{\mathbb{C}}$ with multiplicities that are bounded in absolute value by $a \cdot l$.

Next we study x and y on a connected component of $X_1(5l)_{\mathbb{F}_l[\zeta_5]}$. Write the connected component as $C \cup C'$, where P_l has order 1 on C and order l on C'. Along C' the multiplicity of both x and y is zero. Indeed, the image of the generic point of C' in E is a point of order l, which is disjoint from the 5-torsion, the locus where x and y have their zeroes and poles.

The difficult part lies in dealing with the other component C. We claim that the multiplicity of x along C is -2, and that the multiplicity of y along C is -3. We reason as follows. Let η be the generic point of C and write A for the local ring $O_{X_1(5l)_{\mathbb{Z}[\zeta_{5l}]},\eta}$. Then A is a discrete valuation ring with uniformiser $\zeta_l - 1$. The multiplicities of x and y along C are the valuations of x and y. We denote by $\widehat{A^{\mathrm{sh}}}$ the completion of the strict Henselisation of A. We have our elliptic curve E over $\widehat{A^{\mathrm{sh}}}$, and it has ordinary reduction at the residue field. The l-divisible groups $E[l^\infty]$ and μ_{l^∞} over $\widehat{A^{\mathrm{sh}}}$ are isomorphic, and thus, by completing along the zero-section, the formal groups of E and $\mathbb{G}_{\mathrm{m}}$ are isomorphic. Under this isomorphism, the point P_l of order l is identified with ζ_l on $\mathbb{G}_{\mathrm{m}}$. Moreover, a formal uniformiser u of E at the zero-section is identified with a uniformiser $t - 1$ at 1 on $\mathbb{G}_{\mathrm{m}}$, where t is the standard coordinate on $\mathbb{G}_{\mathrm{m}}$. Combining, we find that the valuations of x and y with respect to $\zeta_l - 1$ are just the valuations of x and y with respect to $t - 1$, that is, with respect to u. But we know that these valuations are -2 and -3, respectively. This proves our claim.

Finally we consider x and y on a connected component of $X_1(5l)_{\mathbb{F}_5[\zeta_l]}$. Again write the connected component as $C \cup C'$, where P_5 has order 1 on C and order 5 on C'. Along C' the multiplicity of both x and y is just zero: indeed, the section P_l has image in $E/Y_1(5)'$ and there it is not equal to a 5-torsion point, so that x and y are regular along the image.

Again most of the work is to be done for the other component C. What we will do is just give an estimate for the multiplicities of x and y along C, using classical intersection theory in the fibre $C \cup C'$. Say we start with the function x; the case of y will be just analogous. We know that x has multiplicity 0 along C', and hence it restricts to a non-trivial rational function, also denoted x, on C'. The degree of x on C' is zero, or equivalently $m(C', C) + (C', \mathrm{div}(x)_{\mathrm{hor}}) = 0$, where m is the multiplicity of C for x. Earlier we have defined a to be the maximum of the absolute values of the multiplicities of x and y along the irreducible components of the fibres of $\overline{E} \to X_1(5)_{\mathbb{C}}$ above the cusps of $X_1(5)_{\mathbb{C}}$; we concluded then that the multiplicity of a cusp of $X_1(5l)$ in $\mathrm{div}(x)$ or $\mathrm{div}(y)$ is bounded in absolute value by $a{\cdot}l$. The number of cusps on $X_1(5l)_{\mathbb{C}}$ is also linear in l, say bounded by $c{\cdot}l$. We conclude that (using classical local intersection numbers) the term $(C', \mathrm{div}(x)_{\mathrm{hor}})$ is bounded from above, in absolute value, by $ac{\cdot}l^2$. On the other hand, the term (C', C) is bounded from below by the number of supersingular j-invariants in $\overline{\mathbb{F}_l}$, and this, in turn, is bounded from below by $d{\cdot}l$ for some positive number d. Combining we find that $|m| \leq (ac/d){\cdot}l$.

All in all we conclude that the absolute values of the multiplicities of the irreducible components in $\mathrm{div}(b)$, $\mathrm{div}(x)$ and $\mathrm{div}(y)$ on $X_1(5l)_{\mathbb{Z}[\zeta_{5l}]}$ are bounded by a constant times l, and those in characteristic l are at most 3. The same holds then for $\frac{1}{[K:\mathbb{Q}]} \sum_s m_s(f) \log \#k(s)$ in the situation of the theorem. Since the number of cusps is bounded by a constant times l, we get that $\deg f$ is bounded by a constant times times l^2. This completes the proof of Theorem 15.1.1. □

15.2 Controlling $D'_x - D$

In this subsection, the hypotheses are as follows (unless stated otherwise). We let K be a number field, O_K its ring of integers, $B := \mathrm{Spec}(O_K)$, $p\colon \mathcal{X} \to B$ a regular, split semi-stable curve over B whose generic fibre $X \to \mathrm{Spec}\, K$ is geometrically irreducible and of genus $g \geq 1$. We let D be the closure in $\mathcal{X}$ of an effective divisor of degree g (also denoted D) on X. We let x be a K-rational torsion point of the Jacobian of X, i.e., a torsion element of $\mathrm{Pic}(X)$, which has the property that there is a unique effective divisor D'_x on X such that $x = [D'_x - D]$. Finally, we let $P\colon B \to \mathcal{X}$ be a section of p, i.e., an element of $\mathcal{X}(B)$.

We denote by $\Phi_{x,P}$ the unique finite vertical *fractional* divisor Φ (i.e., with rational coefficients that are not necessarily integral) on $\mathcal{X}$ such that $P(B)$ is disjoint from the support of Φ and $(D'_x - D - \Phi, C) = 0$ for all irreducible components C of finite fibres of p. It is not difficult to see that such a Φ exists and that it is unique. We denote by δ_s the number of singular points in the geometric fibre at a closed point s of B.

15.2.1 Theorem *The O_B-module $\mathrm{R}^1p_*O_{\mathcal{X}}(D'_x)$ is a torsion module on B, and we have:*

$$\begin{aligned}(D'_x, P) + \log \#\mathrm{R}^1p_*O_{\mathcal{X}}(D'_x) + \frac{1}{8}(\omega_{\mathcal{X}/B}, \omega_{\mathcal{X}/B}) + \frac{1}{8}\sum_s \delta_s \log \#k(s)\\ = (D, P) - \frac{1}{2}(D + \Phi_{x,P}, D + \Phi_{x,P} - \omega_{\mathcal{X}/B}) + \frac{1}{2}\deg\det p_*\omega_{\mathcal{X}/B}\\ + \sum_\sigma \int_{X_\sigma} \log\|\vartheta\|(D'^{\sigma}_x - Q)\cdot\mu_\sigma(Q) + \frac{g}{2}[K:\mathbb{Q}]\log(2\pi).\end{aligned}$$

Here s runs over the closed points of B, and σ runs through the complex embeddings of K.

We derive Theorem 15.2.1 from three lemmas. For the moment we work in $\mathbb{Q}\otimes_{\mathbb{Z}}\widehat{\mathrm{Pic}}(\mathcal{X})$.

15.2.2 Lemma *The admissible line bundles $O_{\mathcal{X}}(D'_x - D)\otimes p^*P^*O_{\mathcal{X}}(D'_x - D)^\vee$ and $O_{\mathcal{X}}(\Phi_{x,P})$ are numerically equivalent. That is, for any admissible line bundle F on $\mathcal{X}$ we have:*

$$(O_{\mathcal{X}}(D'_x - D)\otimes p^*P^*O_{\mathcal{X}}(D'_x - D)^\vee, F) = (O_{\mathcal{X}}(\Phi_{x,P}), F).$$

Proof In this proof we just write Φ for $\Phi_{x,P}$. We denote the first line bundle in the lemma by Ψ. Since $D'_x - D$ is torsion, there is a positive integer N such that $\Psi^{\otimes N}$ is trivial on the generic fibre as a classical line bundle (that is, without taking the metrics into account). We have a canonical isomorphism $P^*\Psi \xrightarrow{\sim} O_B$ on B. Combining, we find that $\Psi^{\otimes N}$ has a rational section s with $\mathrm{div}_{\mathcal{X}}(s)$ vertical and with $P^*s \mapsto 1$. The latter condition implies that P intersects to zero with $\mathrm{div}_{\mathcal{X}}(s)$ for the Arakelov intersection product. On the other hand, as $p^*P^*O_{\mathcal{X}}(D'_x - D)^\vee$ is trivial on the fibres of p over finite places of B, we have $(N(D'_x - D) - \mathrm{div}_{\mathcal{X}}(s), C) = 0$ for all irreducible components C of fibres of p. Hence in fact $\Phi = \frac{1}{N}\mathrm{div}_{\mathcal{X}}(s)$. To prove the lemma, it suffices now to prove that $\Psi^{\otimes N} \xrightarrow{\sim} O_{\mathcal{X}}(\mathrm{div}_{\mathcal{X}}(s))$ given by $s \mapsto 1$, with 1 the tautological section, is an isometry. Because of admissibility, it suffices to check that this is so when restricted to P; but here we get the canonical isomorphism $P^*\Psi \xrightarrow{\sim} O_B$. This is indeed an isometry by the definition of Ψ. □

15.2.3 Lemma *Let X be a compact Riemann surface of genus $g \geq 1$. Let D' be an effective divisor on X of degree g satisfying $h^0(D') = 1$. Then the determinant of cohomology $\lambda(O_X(D'))$ of D' is identified with $H^0(X, O_X(D'))$. Further, the formula:*

$$\log\|1\| + \frac{\delta(X)}{8} + \int_X \log\|\vartheta\|(D' - Q)\cdot\mu_X(Q) = 0$$

holds for the length (with respect to Faltings' metrisation of the determinant of cohomology) of the tautological section 1 of $H^0(X, O_X(D'))$.

Proof Since $h^0(D') = 1$, the set of points Q on X such that $h^0(D' - Q) > 0$ is finite. Let Q be a point with $h^0(D' - Q) = 0$. According to the axioms for the metrisation of the determinant of cohomology, the exact sequence:

$$0 \to O_X(D' - Q) \to O_X(D') \to Q_* Q^* O_X(D') \to 0$$

gives rise to an isometry:

$$\lambda(O_X(D')) \xrightarrow{\sim} \lambda(O_X(D' - Q)) \otimes Q^* O_X(D') \cong O(-\Theta)[O_X(D' - Q)] \otimes Q^* O_X(D').$$

Taking the norm on left and right of a tautological section we obtain:

$$\|1\| = \exp(-\delta(X)/8) \cdot \|\vartheta\|(D' - Q)^{-1} \cdot G(D', Q).$$

Taking logarithms and then integrating against $\mu_X(Q)$ gives the result. □

15.2.4 Lemma *(Noether formula) We have:*

$$12 \deg \det p_* \omega_{\mathcal{X}/B} = (\omega_{\mathcal{X}/B}, \omega_{\mathcal{X}/B}) + \sum_s \delta_s \log \#k(s) + \sum_\sigma \delta(X_\sigma) - 4g[K : \mathbb{Q}] \log(2\pi)\,,$$

the first sum running over the closed points of B, the second sum running over the complex embeddings of K.

Proof See [39] and [85]. □

Proof [Proof of Theorem 15.2.1] The fact that $\mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x)$ is a torsion module follows since generically $h^1(D'_x) = 0$. So let us turn to the formula. We start by noting that

$$(D'_x - D, P) = \deg P^* O_{\mathcal{X}}(D'_x - D).$$

According to Lemma 15.2.2, the admissible line bundles $O_{\mathcal{X}}(D'_x) \otimes p^* P^* O_{\mathcal{X}}(D'_x - D)^\vee$ and $O_{\mathcal{X}}(D + \Phi_{x,P})$ are numerically equivalent. By the Riemann-Roch theorem we then have:

$$\deg \det \mathrm{R}p_*(O_{\mathcal{X}}(D'_x) \otimes p^* P^* O_{\mathcal{X}}(D'_x - D)^\vee) = \frac{1}{2}(D + \Phi_{x,P}, D + \Phi_{x,P} - \omega_{\mathcal{X}/B}) + \deg \det p_* \omega_{\mathcal{X}/B}\,.$$

By the projection formula for the determinant of cohomology we can write the left-hand side as:

$$\deg \det \mathrm{R}p_*(O_{\mathcal{X}}(D'_x) \otimes p^* P^* O_{\mathcal{X}}(D'_x - D)^\vee) = \deg \det \mathrm{R}p_* O_{\mathcal{X}}(D'_x) - \deg P^* O_{\mathcal{X}}(D'_x - D)\,.$$

Since $p_* O_{\mathcal{X}}(D'_x)$ is canonically trivialised by the function 1, the term $\deg \det \mathrm{R}p_* O_{\mathcal{X}}(D'_x)$ can be computed as:

$$\deg \det \mathrm{R}p_* O_{\mathcal{X}}(D'_x) = -\sum_\sigma \log \|1\|_\sigma - \log \#\mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x),$$

where for each complex embedding σ, the norm $\|1\|_\sigma$ is the length of the tautological section 1 of $\lambda(O_{X_\sigma}(D'_x)) = H^0(O_{X_\sigma}(D'_x))$. By Lemma 15.2.3 we can then write:

$$\deg\det \mathrm{R}p_*O_{\mathcal{X}}(D'_x) = \sum_\sigma \int_{X_\sigma} \log\|\vartheta\|_\sigma(D'^\sigma_x - Q)\cdot\mu_\sigma(Q) + \sum_\sigma \delta(X_\sigma)/8 - \log\#\mathrm{R}^1p_*O_{\mathcal{X}}(D'_x)\,.$$

Combining everything gives:

$$\begin{aligned}(D'_x - D, P) = &-\frac{1}{2}(D+\Phi_{x,P}, D+\Phi_{x,P}-\omega_{\mathcal{X}/B}) - \deg\det p_*\omega_{\mathcal{X}/B}\\ &+\sum_\sigma\int_{X_\sigma}\log\|\vartheta\|(D'^\sigma_x - Q)\cdot\mu_\sigma(Q) + \sum_\sigma\delta(X_\sigma)/8 - \log\#\mathrm{R}^1p_*O_{\mathcal{X}}(D'_x)\,.\end{aligned}$$

We obtain the required formula upon eliminating $\sum_\sigma \delta(X_\sigma)/8$ with the Noether formula, Lemma 15.2.4. □

15.2.5 Theorem *We have an upper bound:*

$$\begin{aligned}(D'_x, P) + \log\#\mathrm{R}^1p_*O_{\mathcal{X}}(D'_x) \le &-\frac{1}{2}(D, D-\omega_{\mathcal{X}/B}) + 2g^2\sum_{s\in B}\delta_s\log\#k(s)\\ &+\sum_\sigma\log\|\vartheta\|_{\sigma,\sup} + \frac{g}{2}[K:\mathbb{Q}]\log(2\pi)\\ &+\frac{1}{2}\deg\det p_*\omega_{\mathcal{X}/B} + (D,P)\,,\end{aligned}$$

where s runs through the closed points of B, and where the supnorm $\|\vartheta\|_{\sigma,\sup}$ is taken over $\mathrm{Pic}^{g-1}(X_\sigma)$.

The upper bound follows directly from Theorem 15.2.1 by using Lemma 15.2.6 below and the fact that $(\omega_{\mathcal{X}/B}, \omega_{\mathcal{X}/B}) \ge 0$ (cf. Thm. 5 of [39]).

15.2.6 Lemma *We have an upper bound:*

$$-\frac{1}{2}(D+\Phi_{x,P}, D+\Phi_{x,P}-\omega_{\mathcal{X}/B}) \le -\frac{1}{2}(D, D-\omega_{\mathcal{X}/B}) + 2g^2\sum_{s\in B}\delta_s\log\#k(s)\,,$$

with s running through the closed points of B.

Proof In this proof we just write Φ for $\Phi_{x,P}$. By definition we have $(D'_x - D - \Phi, \Phi) = 0$, or in other words, $(\Phi,\Phi) = (D'_x - D, \Phi)$. Using this we can write

$$-\frac{1}{2}(D+\Phi, D+\Phi-\omega_{\mathcal{X}/B}) = -\frac{1}{2}(D, D-\omega_{\mathcal{X}/B}) + \frac{1}{2}(\Phi, \omega_{\mathcal{X}/B} - D - D'_x)\,.$$

Write $\Phi = \sum_C \Phi(C) \cdot C$ and for any finite fibre F_s of p put $A_s := \sup_C |\Phi(C)|$ with C running through the irreducible components of F_s. Since $\omega_{\mathcal{X}/B}, D$ and D'_x intersect any irreducible component C with non-negative multiplicity, we find

$$\frac{1}{2}(\Phi, \omega_{\mathcal{X}/B} - D - D'_x) \leq \frac{1}{2}\left(\sum_s A_s F_s, \omega_{\mathcal{X}/B} + D + D'_x\right) \leq 2g \sum_s A_s \log \#k(s)\,.$$

We are going to prove that $A_s \leq g\delta_s$, and then we are done. So let s be a finite place of B. Let S_0 be the set of irreducible components of F_s, and let S_1 be the set of double points on F_s. Let Γ_s be the dual graph of F_s (thus, the set of vertices of Γ_s corresponds to S_0, the set of edges corresponds to S_1, and the graph is defined by the incidence relations). Choose an orientation on Γ_s. This gives rise to the usual source and target maps s and $t\colon S_1 \to S_0$. Consider the boundary and coboundary maps $d_* = t_* - s_*\colon \mathbb{Q}^{S_1} \to \mathbb{Q}^{S_0}$ and $d^* = t^* - s^* : \mathbb{Q}^{S_0} \to \mathbb{Q}^{S_1}$. Then $d_*d^*\colon \mathbb{Q}^{S_0} \to \mathbb{Q}^{S_0}$ is given by minus the intersection matrix of F_s. In particular, the map d_*d^* sends Φ to the map $u\colon C \mapsto -(\Phi, C) = (D - D'_x, C)$. The kernel of d_*d^* consists exactly of the constant functions, and the image consists of the orthogonal complement of the constant functions. Now consider the graph Γ_s as an electric circuit, where each edge has a resistance of 1 Ohm. By Ohm's law and by spelling out the maps d_* and d^* we see that if we let at each vertex C a current of $u(C)$ Ampère enter the circuit, subject to the condition that $\sum_C u(C) = 0$, the potentials $\varphi(C)$ at each vertex C will be given, up to addition of a constant function, by a solution of the equation $d_*d^*\varphi = u$. Hence Φ is the potential corresponding to the current $C \mapsto (D - D'_x, C)$, normalised by the condition that $\Phi(C_P) = 0$ with C_P the component that P specialises to. We must bound the $|\Phi(C)|$ for C varying over S_0. The worst case that may happen is that Γ_s is a chain, with D' and D specialising entirely to the beginning and end point, respectively. In this case, the biggest potential difference is $g \cdot (\#S_0 - 1)$ in absolute value, so that we arrive at $|\Phi(C)| \leq g \cdot (\#S_0 - 1)$. Now note that Γ_s is connected and that X/K has split semi-stable reduction. This gives $\#S_0 - 1 \leq \delta_s$ and hence $|\Phi(C)| \leq g\delta_s$, as required. □

16 Bounding the height of $X_1(pl)$

As before, for $l > 5$ prime, we let X_l be the modular curve $X_1(5l)$, over a suitable base that will be clear from the notation. We let g_l denote the genus of X_l; we have $g_l > 1$. A model $X_{l,\mathbb{Z}}$ is given by [61], as well as a model $X_{l,\mathbb{Z}[\zeta_{5l}]}$ that is semi-stable. The aim of this section is to prove a suitable bound for the stable Faltings height of X_l (see 14.4.5). We will in fact give such a bound for the modular curves $X_1(pl)$ with p and l distinct primes. Before we get to that, we prove some intermediate results, that will also be important in the next section.

16.1 Lemma *Let $N \geq 1$ be an integer, and let*

$$\mathcal{B}_2(N) := \coprod_{m|N} \coprod_{d|(N/M)} B^*_{N,M,d} S_2(\Gamma_1(M))^{\mathrm{new}}$$

be the basis of $S_2(\Gamma_1(N))$ as explained in (8.17). Let $f = \sum_{n\geq 1} a_n(f)q^n$ be an element of $\mathcal{B}_2(N)$. Then we have for all $n \geq 1$:

$$|a_n(f)| \leq 2n.$$

Proof As $a_n(B^*_{N,M,d}f) = a_{n/d}(f)$ (see 8.16), it suffices to treat the case that f is a newform of some level M dividing N. We use the Weil bounds on the $a_p(f)$ for all primes p. We recall from Section 1.8 of [22] that we have an equality of formal Dirichlet series:

$$\sum_{n\geq 1} a_n(f)n^{-s} = \prod_{p|M}(1 - a_p(f)p^{-s})^{-1} \prod_{p\nmid M}(1 - \alpha_p p^{-s})^{-1}(1 - \beta_p p^{-s})^{-1}$$

with the following properties. For $p \nmid M$ we have $|\alpha_p| = |\beta_p| = \sqrt{p}$. For $p|M$ we have:

$$\begin{cases} a_p(f) = 0 & \text{if } p^2|M, \\ a_p(f) = 0 & \text{if } \varepsilon_f \text{ factors through } (\mathbb{Z}/(M/p)\mathbb{Z})^\times, \\ |a_p(f)| = p^{1/2} & \text{if } \varepsilon_f \text{ does not factor through } (\mathbb{Z}/(M/p)\mathbb{Z})^\times, \\ |a_p(f)| = 1 & \text{if } p^2|M \text{ and } \varepsilon_f \text{ factors through } (\mathbb{Z}/(M/p)\mathbb{Z})^\times. \end{cases}$$

Using that:

$$(1 - a_p(f)p^{-s})^{-1} = \sum_{k\geq 0} a_p(f)^k p^{-sk},$$

and that:

$$(1 - \alpha_p p^{-s})^{-1}(1 - \beta_p p^{-s})^{-1} = (\sum_{k\geq 0} \alpha_p^k p^{-sk})(\sum_{k\geq 0} \beta_p^k p^{-sk})$$

we find that for arbitrary n we have $|a_n(f)| \leq \sigma_{0,p}(n)\sqrt{n}$, where $\sigma_{0,p}(n)$ is the number of positive divisors of n that are prime to p, and a simple estimate leads to $|a_n(f)| \leq 2n$. □

The following lemma states a very well known lower bound for the Petersson norm of a normalised cuspform.

16.2 Lemma *Let $N \geq 1$ and let $\omega = f dq/q$ be the holomorphic 1-form on $X_1(N)(\mathbb{C})$ associated to a cusp form $f = \sum_n a_n(f)q^n$ in $S_2(\Gamma_1(N))$ with $a_1(f) = 1$. Then we have:*

$$\|\omega\|^2 = \frac{i}{2}\int_{X_1(N)} \omega \wedge \overline{\omega} \geq \pi e^{-4\pi}.$$

Proof We have $\omega = \sum_{n\geq 1} a_n(f)q^n dq/q$ in the coordinate $q = e^{2\pi iz}$, where z is the standard coordinate on the upper half plane $\mathbb{H}$. If we let x and y be the real and imaginary parts of z we have:

$$\frac{i}{2}\omega \wedge \overline{\omega} = 4\pi^2 |f|^2 \, dx \, dy$$

Let F be the region in $\mathbb{H}$ given by the conditions $|x| < 1/2$ and $y > 1$. Then:

$$\begin{aligned}
\|\omega\|^2 &\geq \int_F 4\pi^2 |f(z)|^2 dx dy \\
&= 4\pi^2 \sum_{m,n\geq 1} a_m(f)\overline{a_n(f)} \int_{-1/2}^{1/2} e^{2\pi i(m-n)x} \int_1^{\infty} e^{-2\pi(m+n)y} dy \\
&= 4\pi^2 \sum_{n\geq 1} |a_n(f)|^2 e^{-4\pi n}/4\pi n \, .
\end{aligned}$$

Taking only the first term (note that $a_1(f) = 1$) we obtain $\|\omega\|^2 \geq \pi e^{-4\pi}$. □

We now specialise to a slightly less special case than our curves X_l: the curves $X_1(pl)$ with p and l two distinct prime numbers. We call an Atkin-Lehner basis for $\Omega^1(X_1(pl))$ any basis of $\Omega^1(X_1(pl))$ given by an ordering of the set $\mathcal{B}_2(pl)$. We start by describing, in a notation that is slightly different from the one used in (8.14), the degeneracy maps that are used for the definition of $\mathcal{B}_2(pl)$. This time, we call them source and target maps:

$$\begin{cases}
s_l \colon X_1(pl) \to X_1(p), & (E, P, L) \mapsto (E, P) \\
t_l \colon X_1(pl) \to X_1(p), & (E, P, L) \mapsto (E/\langle L \rangle, P) \\
s_p \colon X_1(pl) \to X_1(l), & (E, P, L) \mapsto (E, L) \\
t_p \colon X_1(pl) \to X_1(l), & (E, P, L) \mapsto (E/\langle P \rangle, L)
\end{cases}$$

where (E, P, L) denotes an elliptic curve E with a point P of order p and a point L of order l. Note that s_l and t_l have degree $l^2 - 1$, and that s_p and t_p have degree $p^2 - 1$. For any integer $M \geq 1$ we denote by $\Omega^1(X_1(M))^{\mathrm{new}}$ the set of holomorphic 1-forms in $\Omega^1(X_1(M))$ of the form $f dq/q$ with f in $S_2(X_1(M))^{\mathrm{new}}$. Our next goal is to get information on the Gram matrix of an Atkin-Lehner basis of $\Omega^1(X_1(pl))$. As described above, the contribution to $\Omega^1(X_1(pl))$ of each f in $S_2(\Gamma_1(pl))^{\mathrm{new}}$ is the subspace $\mathbb{C} f dq/q$. The contribution of an f in $S_2(\Gamma_1(p))^{\mathrm{new}}$ is the 2-dimensional space generated by $s_l^* f dq/q$ and $t_l^* f dq/q$, and, of course, each f in $S_2(\Gamma_1(l))^{\mathrm{new}}$ contributes the 2-dimensional space generated by $s_p^* f dq/q$ and $t_p^* f dq/q$.

16.3 Lemma *For f in $S_2(\Gamma_1(l))^{\mathrm{new}}$ and $\omega = f dq/q$ we have:*

$$\begin{aligned}
\langle s_p^*\omega, s_p^*\omega \rangle &= (p^2 - 1)\|\omega\|^2 \, , \\
\langle t_p^*\omega, t_p^*\omega \rangle &= (p^2 - 1)\|\omega\|^2 \, , \\
\langle s_p^*\omega, t_p^*\omega \rangle &= (p - 1)\overline{a_p(f)}\|\omega\|^2 .
\end{aligned}$$

We have similar equalities with p and l switched.

Proof The first two equalities are clear. As to the latter, note first that

$$\langle s_p^*\omega, t_p^*\omega\rangle = \frac{i}{2}\int_{X_1(pl)} s_p^*\omega \wedge \overline{t_p^*\omega} = \frac{i}{2}\int_{X_1(l)} s_{p,*}(s_p^*\omega \wedge \overline{t_p^*\omega}) = \frac{i}{2}\int_{X_1(l)} \omega \wedge \overline{s_{p,*}t_p^*\omega}\,.$$

Next note that $s_p\colon X_1(pl) \to X_1(l)$ and $t_p\colon X_1(pl) \to X_1(l)$ factor through the forget map $X_1(pl) \to X_1(l;p)$ where the latter curve corresponds to the moduli problem (E, P, G) with P of order l and G a subgroup of order p. This forget map has degree $p-1$, and the correspondence on $X_1(l)$ induced by $X_1(l;p)$ is the standard Hecke correspondence T_p. We find that $s_{p,*}t_p^*\omega = (p-1)T_p^*\omega$. By the standard relation between eigenvalues and q-coefficients we have $T_p^*\omega = a_p(f)\omega$, so finally:

$$\langle s_p^*\omega, t_p^*\omega\rangle = \frac{i}{2}\int_{X_1(l)} \omega \wedge \overline{(p-1)T_p^*\omega} = (p-1)\overline{a_p(f)}\|\omega\|^2$$

as required. □

16.4 Corollary *Let p and l be two distinct primes. The structure of the Gram matrix $(\langle\omega_i, \omega_j\rangle)_{i,j}$ of holomorphic 1-forms associated to an Atkin-Lehner basis for $\Omega^1(X_1(pl))$ is as follows. Two subspaces associated to distinct elements of the union of $S_2(\Gamma_1(pl))^{\mathrm{new}}$, $S_2(\Gamma_1(l))^{\mathrm{new}}$ and $S_2(\Gamma_1(p))^{\mathrm{new}}$ are orthogonal to each other, hence the Gram matrix decomposes into blocks corresponding to these subspaces. The contribution of an element f in $S_2(\Gamma_1(pl))^{\mathrm{new}}$ is the 1-by-1 block $\|fdq/q\|^2$. The contribution of an element f in $S_2(\Gamma_1(l))^{\mathrm{new}}$ is the 2-by-2 block:*

$$(p-1)\|fdq/q\|^2 \begin{pmatrix} p+1 & \overline{a_p(f)} \\ a_p(f) & p+1 \end{pmatrix},$$

where the norm $\|fdq/q\|^2$ is taken on $X_1(l)$). The contribution of an element f in $S_2(\Gamma_1(p))^{\mathrm{new}}$ is the 2-by-2 block:

$$(l-1)\|fdq/q\|^2 \begin{pmatrix} l+1 & \overline{a_l(f)} \\ a_l(f) & l+1 \end{pmatrix},$$

where the norm $\|fdq/q\|^2$ taken on $X_1(l)$.

16.5 Corollary *The determinant of the Gram matrix of the holomorphic 1-forms associated to an Atkin-Lehner basis for $\Omega^1(X_1(pl))$ is bounded below by $(\pi e^{-4\pi})^{g/2}$.*

Proof By the Weil bounds, the determinant of a 2-by-2 block as in Corollary 16.4 is bounded below by $\|fdq/q\|^2$. We obtain our corollary by invoking Lemma 16.2. □

16.6 Corollary *The Arakelov (1,1)-form μ on $X_1(pl)$ is given by:*

$$\mu = \sum_{\omega} \frac{i}{2g\|\omega\|^2}\omega \wedge \overline{\omega}$$
$$+ \sum_{\omega} \frac{(i/2g)\left((p+1)s_p^*\omega \wedge \overline{s_p^*\omega} + (p+1)t_p^*\omega \wedge \overline{t_p^*\omega} - a_p(f)s_p^*\omega \wedge \overline{t_p^*\omega} - \overline{a_p(f)}t_p^*\omega \wedge \overline{s_p^*\omega}\right)}{(p-1)\|\omega\|^2((p+1)^2 - |a_p(f)|^2)},$$
$$+ \sum_{\omega} \frac{(i/2g)\left((l+1)s_l^*\omega \wedge \overline{s_l^*\omega} + (l+1)t_l^*\omega \wedge \overline{t_l^*\omega} - a_l(f)s_l^*\omega \wedge \overline{t_l^*\omega} - \overline{a_l(f)}t_l^*\omega \wedge \overline{s_l^*\omega}\right)}{(l-1)\|\omega\|^2((l+1)^2 - |a_l(f)|^2)}$$

the first sum running over $\Omega^1(X_1(pl))^{\mathrm{new}}$, the second sum running over $\Omega^1(X_1(l))^{\mathrm{new}}$, the third sum running over $\Omega^1(X_1(p))^{\mathrm{new}}$.

Proof Consider first an arbitrary compact Riemann surface X and let $\omega = (\omega_1, \ldots, \omega_g)$ be an arbitrary basis of $\Omega^1(X)$. Let a be the g-by-g matrix given by $a_{i,j} = \langle \omega_i, \omega_j \rangle$. Note that $\overline{a} = a^t$. Let $b = a^{t,-1}$, the inverse of the transpose of a. Then we claim that the Arakelov $(1,1)$-form on X can be written as:

$$\mu = \frac{i}{2g}\sum_{i,j} b_{i,j}\omega_i \wedge \overline{\omega_j}.$$

To see this, note that for ω an orthonormal basis this is the correct expression, and that changing to $\omega' = \omega \cdot g$ with any invertible g does not change μ, as one may directly calculate.

In our case, the basis $(\omega_1, \ldots, \omega_g)$ that we take is an Atkin-Lehner basis as above. Using Corollary 16.4 one obtains the expression that we gave. □

We remark that Abbes and Ullmo have determined the Arakelov $(1,1)$-form on $X_0(n)$ for all square free $n \geq 1$ such that $X_0(n)$ has genus at least one in [1]. It should not be hard to generalise their result to $X_1(n)$ for square free n.

Now we arrive at the main result of this section. We recall that the Faltings height of a curve, and its stable or absolute version, have been briefly described in (14.4.6).

16.7 Theorem *For the stable Faltings height of $X = X_1(pl)$, for distinct prime numbers p and l, one has:*

$$h_{\mathrm{abs}}(X) = O((pl)^2 \log(pl)).$$

Proof This proof is an adaptation of an argument in Section 5 of [10] where the case $X_0(p)$ with p prime was treated. We may and do assume that $X_1(pl)$ has genus at least one.

We start with a general observation. For X_K a curve over a number field, and $K \to L$ a finite extension, we claim that:

$$[L:\mathbb{Q}]^{-1}h_L(X_L) \leq [K:\mathbb{Q}]^{-1}h_K(X_K).$$

This inequality simply results from the fact that for the Néron models of the Jacobians the identity morphism on the generic fibres extends to a morphism:

$$(J_{O_K})_{O_L} \longrightarrow J_{O_L}.$$

Because of this observation it suffices to establish the bound of the theorem for the $h_{\mathbb{Q}}(X_\mu(pl)_{\mathbb{Q}})$, where $X_\mu(pl)_{\mathbb{Q}}$ denotes the modular curve corresponding to elliptic curves with an embedding of μ_{pl}. The reason for considering this variant of $X_1(pl)$ is that the cusp ∞ of $X_\mu(n)$ is $\mathbb{Q}$-rational for all n. Of course, $X_1(n)_{\mathbb{Q}}$ and $X_\mu(n)_{\mathbb{Q}}$ become isomorphic over $\mathbb{Q}(\zeta_n)$. For more details about these $X_\mu(n)$ we refer to sections 9.3 and 12.3 of [28]. We let X be the model over $\mathbb{Z}$ of $X_{\mathbb{Q}}$ obtained by normalisation of the j-line $\mathbb{P}^1_{\mathbb{Z}}$ in the function field of $X_{\mathbb{Q}}$. As X is proper over $\mathbb{Z}$, the $\mathbb{Q}$-rational point ∞ extends to an element ∞ in $X(\mathbb{Z})$, which is known to lie in the open part X^{sm} of X where the structure morphism to $\mathrm{Spec}(\mathbb{Z})$ is smooth.

We let J be the Néron model over $\mathbb{Z}$ of the Jacobian of the curve $X_{\mathbb{Q}}$. Then, by the defining property, the embedding of $X_{\mathbb{Q}}$ into $J_{\mathbb{Q}}$ that sends ∞ to 0 extends to a morphism from X^{sm} to J. This morphism induces via pullback of differential forms a morphism from $\mathrm{Cot}_0(J)$ to $S(\mathbb{Z})$, the sub-$\mathbb{Z}$-module of $\Omega^1(X_{\mathbb{Q}})$ of forms whose q-expansion at ∞ has coefficients in $\mathbb{Q}$ (see around (9.9)). As $\mathrm{Cot}_0(J)$ and $S(\mathbb{Z})$ are both $\mathbb{Z}$-structures on $\Omega^1(X(\mathbb{C}))$, we have (see around (14.4.5)):

$$\begin{aligned} h_{\mathbb{Q}}(X_{\mathbb{Q}}) &= \deg(\bigwedge^g 0^* \mathrm{Cot}_0(J)) = -\log \mathrm{Vol}((\mathbb{R}\otimes \mathrm{Cot}_0(J))/\mathrm{Cot}_0(J)) \\ &\leq -\log \mathrm{Vol}(\mathbb{R}\otimes S(\mathbb{Z})/S(\mathbb{Z})), \end{aligned}$$

where the volume form on $\mathbb{R}\otimes \mathrm{Cot}_0(J)$ comes from integration over $J(\mathbb{C})$, and that on $\mathbb{R}\otimes S(\mathbb{Z})$ from integration over $X(\mathbb{C})$.

Let $\mathbb{T}\subset \mathrm{End}(J)$ be the Hecke algebra, generated by all T_i, $i\geq 1$, and the $\langle a\rangle$, a in $(\mathbb{Z}/pl\mathbb{Z})^\times$. We have a perfect pairing (see (9.9)):

$$\mathbb{T}\times S(\mathbb{Z})\to \mathbb{Z}, \quad (t,\omega)\mapsto a_1(t\omega).$$

Using the duality we can write:

$$-\log \mathrm{Vol}(\mathbb{R}\otimes S(\mathbb{Z})/S(\mathbb{Z})) = \log \mathrm{Vol}(\mathbb{R}\otimes \mathbb{T}/\mathbb{T})$$

where in the latter the volume form is dual to the earlier one. Now consider an Atkin-Lehner basis $(\omega_1,\ldots,\omega_g)$ of $\Omega^1(X)_{\mathbb{C}}$. Let Vol' denote the volume with respect to the volume form on $\mathbb{R}\otimes\mathbb{T}$ induced by the one on $\mathbb{C}\otimes S(\mathbb{Z})$ for which the basis $(\omega_1,\ldots,\omega_g)$ is an orthonormal basis. Then we have:

$$\log \mathrm{Vol}(\mathbb{R}\otimes\mathbb{T}/\mathbb{T}) = \log \mathrm{Vol}'(\mathbb{R}\otimes\mathbb{T}/\mathbb{T}) - \frac{1}{2}\log\det(\langle\omega_i,\omega_j\rangle)\,.$$

By Corollary 16.5 we have:

$$-\log\det(\langle\omega_i,\omega_j\rangle) \leq g(4\pi - \log\pi)/2 = O((pl)^2).$$

It remains to bound $\log \mathrm{Vol}'(\mathbb{R}\otimes\mathbb{T}/\mathbb{T})$. Let Γ be the set of integers $i \geq 1$ such that there exists an ω in $\Omega^1(X(\mathbb{C}))$ with a zero of exact order $i-1$ at ∞. We can write $\Gamma = \{i_1,\ldots,i_g\}$ with:

$$1 = i_1 < i_2 < \ldots < i_g < 2g-2.$$

It follows that the $\mathbb{Z}$-submodule $\mathbb{T}'$ of $\mathbb{T}$ generated by the T_{i_j} has finite index. We thus find:

$$\log \mathrm{Vol}(\mathbb{R}\otimes\mathbb{T}/\mathbb{T}) \leq \log \mathrm{Vol}'(\mathbb{R}\otimes\mathbb{T}'/\mathbb{T}') - \frac{1}{2}\log\det(\langle\omega_i,\omega_j\rangle)\,.$$

Now we have $g = r_1 + 2r_2$, where r_1 is the number of elements of our basis $(\omega_1,\ldots,\omega_g)$ of $\mathbb{C}\otimes\mathbb{T}^\vee$ that are fixed by the complex conjugation. We let:

$$\phi\colon \mathbb{R}\otimes\mathbb{T} \longrightarrow \mathbb{R}^{r_1}\times\mathbb{C}^{r_2}\times\mathbb{C}^{r_2} \longrightarrow \mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$$

be the map obtained from our basis (each ω_i gives $t\mapsto a_1(t\omega_i)$), composed with the projection. We view $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$ as $\mathbb{R}^g$ by decomposing each factor $\mathbb{C}$ as $\mathbb{R}\oplus\mathbb{R}i$. Then we have:

$$\mathrm{Vol}'(\mathbb{R}\otimes\mathbb{T}/\mathbb{T}') = 2^{r_2}|\det(\phi(T_{i_1}),\ldots,\phi(T_{i_g}))|.$$

By construction, each $\phi(T_{i_j})_k$ is the real or imaginary part of some $a_{i_j}(\omega_l)$. Hence by Lemma 16.1 we have:

$$|\phi(T_{i_j})_k| \leq 2i_j.$$

We obtain:

$$|\det(\phi(T_{i_1}),\ldots,\phi(T_{i_g}))| \leq \prod_{j=1}^{g}(2i_j\sqrt{g}) \leq (4g^2)^g.$$

Hence, finally:

$$\log \mathrm{Vol}'(\mathbb{R}\otimes\mathbb{T}/\mathbb{T}) \leq r_2\log 2 + g(\log 4 + 2\log g).$$

Noting that $r_2 \leq g/2$ and that $g = O((pl)^2)$ completes our proof. □

17 Bounding the theta function on $\mathrm{Pic}^{g-1}(X_1(pl))$

The aim of this section is to give a bound for the supnorm of the theta function that occurs in Theorem 15.2.5.

17.1 Theorem *For $X = X_1(pl)$, with p and l distinct primes for which the genus of $X_1(pl)$ is at least one, we have* $\log \|\vartheta\|_{\sup} = O((pl)^6)$.

We need two lemmas, which are possibly of independent interest.

17.2 Lemma *Let $X = V/\Lambda$ be a principally polarised complex Abelian variety and let $H\colon V \times V \to \mathbb{C}$ be its Riemann form. Let $\lambda_1, \ldots, \lambda_{2g}$ be the successive minima of the lattice Λ, with norm defined by $\|x\|^2 = H(x,x)$. Then for any symplectic basis $(e_1, \ldots, e_{2g})$ of Λ one has*

$$\sqrt{\det \Im(\tau)} \leq \frac{(2g)!}{2^g} \frac{V_{2g}}{V_g} \lambda_{g+1} \cdots \lambda_{2g} \,,$$

where τ is the period matrix in $\mathbb{H}_g$ corresponding to $(e_1, \ldots, e_{2g})$. Here V_n denotes the volume of the unit ball in $\mathbb{R}^n$ with its standard euclidean inner product.

Proof We start by considering the lattice $M = \mathbb{Z} \cdot e_1 \oplus \ldots \oplus \mathbb{Z} \cdot e_g$ in the real subvector space $W = \mathbb{R} \cdot e_1 \oplus \ldots \oplus \mathbb{R} \cdot e_g$ of V. Denote by $\mu_1, \ldots, \mu_g$ the successive minima of M, where the norm is given by restricting H to W (note that M is isotropic for the symplectic form, so that H takes real values on W). We have $(H(e_i, e_j))_{i,j} = (\Im(\tau))^{-1}$, so that the volume (with respect to the inner product on W given by H) of W/M is equal to $(\det \Im(\tau))^{-1/2}$, and hence by Minkowski's second fundamental inequality:

$$\mu_1 \cdots \mu_g \leq 2^g \frac{(\det \Im(\tau))^{-1/2}}{V_g} \,.$$

On the other hand we have:

$$\mu_1 \cdots \mu_g \geq \lambda_1 \cdots \lambda_g = (\lambda_1 \cdots \lambda_{2g}) \cdot (\lambda_{g+1} \cdots \lambda_{2g})^{-1}$$

and since the volume of V/Λ is 1 we obtain by Minkowski's first fundamental inequality:

$$\lambda_1 \cdots \lambda_{2g} \geq \frac{2^{2g}}{(2g)!} \frac{1}{V_{2g}} \,.$$

Combining we find a lower bound:

$$\mu_1 \cdots \mu_g \geq \frac{2^{2g}}{(2g)!} \frac{1}{V_{2g}} \cdot (\lambda_{g+1} \cdots \lambda_{2g})^{-1} \,.$$

Combining this with the upper bound for $\mu_1 \cdots \mu_g$ we obtain the required formula. □

17.3 Lemma *Let $N \geq 3$ be an integer. The group $\Gamma_1(N)$ is generated by its elements g whose entries are bounded from above in absolute value by $N^6/4$.*

Proof We first note the following: let G be a group, and let $S \subset G$ be a set of generators. Let X be a transitive G-set and let x be in X. For each y in X, let g_y be an element of G such that $g_y x = y$; we demand that $g_x = 1$. Then the $g_{sy}^{-1} s g_y$, for s in S and y in X, form a system of generators for the stabiliser G_x of x. To see this, first replace S by $S \cup S^{-1}$. Let g be in G_x. Write $g = s_n \cdots s_1$ with s_i in S. Then we can write:

$$g = s_n \cdots s_1 = g_{s_n y_n}^{-1} s_n g_{y_n} \cdots g_{s_1 y_1}^{-1} s_1 g_{y_1} \quad \text{with } y_i = (s_{i-1} \cdots s_1)x.$$

The equality holds because $s_i y_i = y_{i+1}$, and $g_{y_1} = 1$ and $g_{s_n y_n} = 1$. Now we apply this to our case. We take $G = \mathrm{SL}_2(\mathbb{Z})$, and we take X to be the subset of $(\mathbb{Z}/N\mathbb{Z})^2$ consisting of the elements of order N. This is a transitive G-set. We let $x = (1, 0)$; then G_x is identified with $\Gamma_1(N)$. Let S be the set consisting of $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and their inverses. Then S generates G. We now apply the previous argument to find generators of $\Gamma_1(N)$. So let $y = (\overline{a}, \overline{b})$ be in X. Then, thinking of $\mathbb{Z}/N\mathbb{Z}$ as the product of its local rings, we see that there is a u in $\mathbb{Z}$ with $|u| \leq N/2$ and $\overline{a + bu}$ in $(\mathbb{Z}/N\mathbb{Z})^\times$. Put $a_1 := a + bu$ and $b_1 := b$. Then:

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

Next, there is a v in $\mathbb{Z}$ with $|v| \leq N/2$ and $b_1 + a_1 v = 1 \bmod N$, i.e.:

$$\begin{pmatrix} a_1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \bmod N.$$

Finally, let w be in $\mathbb{Z}$ with $|w| \leq N/2$ and with image $\overline{a_1}$ in $\mathbb{Z}/N\mathbb{Z}$. Then one has:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} 1 & w \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \bmod N.$$

Writing out, we have:

$$\begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} 1 & w \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} w - u(1 - vw) & -1 - uv \\ 1 - vw & v \end{pmatrix}$$

so that we can put:

$$g_y = g_{(a,b)} = \begin{pmatrix} w - u(1 - vw) & -1 - uv \\ 1 - vw & v \end{pmatrix}.$$

The absolute values of the coefficients of $g_{(a,b)}$ are smaller than $N^3/4$, if $N \geq 3$, and the lemma follows. $\square$

Proof [Proof of Theorem 17.1] Recall that $\|\vartheta\|(z;\tau)$ is given by:

$$\|\vartheta\|(z;\tau) = (\det \Im(\tau))^{1/4} \exp(-\pi {}^t y(\Im(\tau))^{-1} y)|\vartheta(z;\tau)|,$$

where $y = \Im(z)$ and where τ is a period matrix in the Siegel upper half plane $\mathbb{H}_g$ corresponding to X. We first deal with the factor $\det \Im(\tau)$ and for this we invoke Lemma 17.2. Choose once more an Atkin-Lehner basis $(\omega_1, \ldots, \omega_g)$ for $\Omega^1(X)$. Using the dual basis in $\Omega^1(X)^\vee$ we write:

$$J(X) = \mathbb{C}^g/\Lambda \quad \text{with} \quad \Lambda = \mathrm{Image}(H_1(X,\mathbb{Z}) \to \mathbb{C}^g \colon \gamma \mapsto \int_\gamma (\omega_1, \ldots, \omega_g)).$$

The polarisation form for $J(X)$ is given by:

$$(z,w) \mapsto {}^t z \cdot (\langle \omega_i, \omega_j \rangle)_{i,j}^{-1} \cdot \overline{w}.$$

Denote by $\|\cdot\|_P$ the corresponding norm on $\mathbb{C}^g$. We also consider the standard hermitian inner product on $\mathbb{C}^g$, which is just $(z,w) \mapsto {}^t z \cdot \overline{w}$. Here we denote the corresponding norm by $\|\cdot\|_E$. From the next two lemmas we obtain

$$(\lambda_{g+1} \cdots \lambda_{2g})^2 \leq \left(g e^{4\pi} (pl)^{46}/\pi\right)^g$$

and hence, by Lemma 17.2, the estimate:

$$\log(\det \Im(\tau)) = O((pl)^2 \log(pl)).$$

17.4 Lemma *The lattice Λ is generated by its elements x with $\|x\|_E^2 \leq g \cdot (pl)^{46}$.*

Proof For the moment put $N = pl$. Following the natural surjections:

$$\Gamma_1(N) \twoheadrightarrow \Gamma_1(N)^{\mathrm{ab}} = H_1(Y_1(N),\mathbb{Z}) \twoheadrightarrow H_1(X_1(N),\mathbb{Z})$$

we see that any set of generators for $\Gamma_1(N)$ gives generators for $H_1(X_1(N),\mathbb{Z})$. We will take generators for $\Gamma_1(N)$ as given by Lemma 17.3. In particular, the absolute values of their coefficients are bounded by $N^6/4$. We have to see now what this implies for $\|x\|_E$ for corresponding elements x of Λ. Concretely, choose a $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\Gamma_1(N)$. The image in Λ can be given as follows: choose any z in $\mathbb{H}$ and any path in $\mathbb{H}$ from z to gz. This gives us a loop in $X_1(N)$, and the class of that loop is the image of g in $H_1(X_1(N),\mathbb{Z})$. In order to get to Λ we compute the periods of $(\omega_1, \ldots, \omega_g)$ around this loop. We want to get bounds for these periods. In order to do this, note first that we can assume that $c \neq 0$. Indeed, the g with $c = 0$ are unipotent, hence have

trivial image in $H_1(X_1(N),\mathbb{Z})$. Now we make the following choices: first we take a z in $\mathbb{H}$ with $\Im(z)=\Im(gz)$. Since $\Im(gz)=|cz+d|^{-2}\Im(z)$ this implies $|cz+d|=1$, i.e. $z=-d/c+i/|c|$. Second, the path that we take is the straight line from z to gz. We have $|c|\geq N$ because g is in $\Gamma_1(N)$. Using furthermore that the absolute values of the coefficients of g are bounded by $N^6/4$ we get $|gz-z|\leq |az+b|+|z|\leq N^{11}/10$. For the period of an element $\omega=fdq/q$ of the Atkin-Lehner basis we obtain from this that:

$$\left|\int_z^{gz}\omega\right|=\left|\int_z^{gz}f(dq)/q\right|\leq\frac{2\pi N^{11}}{10}\|f\|,$$

where $\|f\|$ denotes the supnorm of f on the straight line from z to gz. When writing $f=\sum_{n\geq 1}a_n(f)q^n$ we have $|a_n(f)|\leq 2n$ by Lemma 16.1. Noting furthermore that that we have $\Im z=|c|^{-1}\geq 4/N^6$ it follows that:

$$\|f\|\leq 2\sum_{n\geq 1}ne^{-8\pi N^{-6}n}=\frac{2r}{(1-r)^2},\quad\text{where}\quad r=e^{-8\pi/N^6}.$$

Hence:

$$\|f\|\leq(N^6/4\pi)^2,\quad\text{and}\quad\left|\int_z^{gz}f(dq)/q\right|\leq\frac{2\pi}{10}N^{11}\left(\frac{N^6}{4\pi}\right)^2\leq N^{23}$$

and the lemma follows. □

17.5 Lemma *For any x in $\mathbb{C}^g$ we have the estimate $\|x\|_P^2\leq(e^{4\pi}/\pi)\|x\|_E^2$.*

Proof By Lemma 16.4 the matrix $(\langle\omega_i,\omega_j\rangle)^{-1}$ is almost diagonal, having in fact diagonal elements $1/\|\omega\|^2$ corresponding to newforms ω on $X_1(pl)$, and 2-by-2 blocks corresponding to newforms ω on $X_1(l)$ and $X_1(p)$, these blocks being:

$$\frac{1}{(p-1)\|\omega\|^2((p+1)^2-|a_p(\omega)|^2)}\begin{pmatrix}p+1 & -\overline{a_p(\omega)}\\ -a_p(\omega) & p+1\end{pmatrix},$$

and similarly for $X_1(p)$. A short calculation shows that for any (z_1,z_2) in $\mathbb{C}^2$ one has:

$$\frac{1}{(p-1)((p+1)^2-|a_p(\omega)|^2)}(z_1,z_2)\begin{pmatrix}p+1 & -\overline{a_p(\omega)}\\ -a_p(\omega) & p+1\end{pmatrix}\begin{pmatrix}\overline{z_1}\\ \overline{z_2}\end{pmatrix}\leq\left(|z_1|^2+|z_2|^2\right),$$

and similarly for $X_1(p)$, so that all in all one gets:

$$\|(z_1,\ldots,z_g)\|_P^2\leq\frac{|z_1|^2}{\|\omega_1\|^2}+\cdots+\frac{|z_g|^2}{\|\omega_g\|^2}.$$

The lemma follows by the lower bound from Lemma 16.2. □

Next we consider the factor $\exp(-\pi^t y(\Im(\tau))^{-1}y)|\vartheta(z;\tau)|$. Since in our previous estimates the choice of τ was irrelevant, it will cause no loss of generality here if we restrict to τ lying in the so-called Siegel fundamental domain F_g, which is the set of $\tau = x+iy$ satisfying the conditions:

1. for each entry x_{ij} of x one has $|x_{ij}| \leq \frac{1}{2}$,
2. for all τ in F_g and all γ in $\mathrm{Sp}(2g,\mathbb{Z})$ one has $\det \Im(\gamma\cdot\tau) \leq \det \Im(\tau)$,
3. y is Minkowski reduced, i.e. for each $\xi = (\xi_1,\ldots,\xi_g) \in \mathbb{Z}^g$ and each i such that $\xi_i,\ldots,\xi_g$ are non-zero, one has $\xi\, y^t\xi \geq y_{ii}$ and moreover, for each $1 \leq i \leq g-1$ one has $y_{i,i+1} \geq 0$.

It is well known that F_g contains at least one representative from each $\mathrm{Sp}(2g,\mathbb{Z})$-orbit on $\mathbb{H}_g$. We claim that for τ in F_g the estimate:

$$\exp(-\pi^t y(\Im(\tau))^{-1}y)|\vartheta(z;\tau)| \leq 2^{3g^3+5g}$$

holds, for all z in $\mathbb{C}^g$. Thus, this factor gives us a contribution $O((pl)^6)$. In order to prove the estimate, write $y = \Im(z) = (\Im(\tau))\cdot b$ with b in $\mathbb{R}^g$. Then it is easy to see that:

$$\exp(-\pi^t y(\Im(\tau))^{-1}y)|\vartheta(z;\tau)| \leq \sum_{n\in\mathbb{Z}^g} \exp(-\pi^t(n+b)(\Im(\tau))(n+b))\,.$$

Since the $\Im(\tau)$ are Minkowski reduced we have, for any m in $\mathbb{R}^g$ (cf. [57], V §4):

$$^t m\Im(\tau)m \geq c(g)\sum_{i=1}^g m_i^2(\Im(\tau))_{ii} \quad \text{with} \quad c(g) = \left(\frac{4}{g^3}\right)^{g-1}\left(\frac{3}{4}\right)^{g(g-1)/2}.$$

Moreover, we have $(\Im(\tau))_{i,i} \geq \sqrt{3}/2$ for $i = 1,\ldots,g$. From this we derive:

$$\begin{aligned}
\sum_{n\in\mathbb{Z}^g} \exp(-\pi^t(n+b)(\Im(\tau))(n+b)) &\leq \sum_{n\in\mathbb{Z}^g} \exp\left(-\sum_{i=1}^g \pi c(g)(n_i+b_i)^2(\Im(\tau))_{i,i}\right)\\
&\leq \prod_{i=1}^g \sum_{n_i\in\mathbb{Z}} \exp(-\pi c(g)(n_i+b_i)^2(\Im(\tau))_{i,i})\\
&\leq \prod_{i=1}^g \frac{2}{1-\exp(-\pi c(g)(\Im(\tau))_{ii})}\\
&\leq 2^g\left(1+\frac{2}{\sqrt{3}\pi c(g)}\right)^g.
\end{aligned}$$

From this and the formula for $c(g)$ the required estimate follows and the proof of Theorem 17.1 is finished. □

18 Upper bounds for Arakelov Green functions on the $X_1(pl)$

The aim of this section is to give an upper bound for the Arakelov Green functions on the curves $X_1(pl)$ that will enable us to bound from above the contributions of the intersection numbers in the right hand side of the inequality in Theorem 15.2.5. As the $X_l(\mathbb{C})$ are compact, it is clear that for each l such an upper bound exists, but we need such upper bounds that grow as most as a power of l.

In order to establish such upper bounds we will use a result of Franz Merkl on Green functions on arbitrary Riemann surfaces that is given in the next subsection. The bound for the Arakelov Green functions on the modular curves $X_1(pl)$ will then be given in the second subsection of this section.

Instead of using the result of Merkl for our work we could certainly also have used recent work by Jorgenson and Kramer in [59]. The results of Jorgenson and Kramer date back to the same time as those of Merkl (early Spring 2004). We chose to use Merkl's results because his approach is more elementary, and we had the details earlier than those of Jorgenson and Kramer.

18.1 An upper bound for Green functions on Riemann surfaces, by Franz Merkl

The contents of this section have been provided by Franz Merkl, already in February 2004.

We begin with explaining the setup and the results of this subsection. Let X be a compact Riemann surface, endowed with a 2-form $\mu \geq 0$ that fulfills $\int_X \mu = 1$. Let $*$ denote rotation by 90° in the cotangential spaces (with respect to the holomorphic structure); in a coordinate z this means $*dz = -idz$, $*d\bar{z} = id\bar{z}$. In particular, the Laplace operator on real C^∞ functions on X can be written as $d * d = 2i\partial\bar{\partial}$. For $a, b \in X$, let $g_{a,b}$ denote the (unique) solution on X (in the sense of distributions) of:

$$d * dg_{a,b} = \delta_a - \delta_b$$

with the normalising condition:

$$\int_X g_{a,b}\mu = 0.$$

Furthermore, let:

$$g_{a,\mu}(x) := \int_{b\in X} g_{a,b}(x)\,\mu(b).$$

Then we have:

$$d * dg_{a,\mu}(x) = \delta_a - \mu$$

in the sense of distributions, and:

$$\int_X g_{a,\mu}\mu = 0.$$

We consider an atlas of X consisting of n local coordinates:

$$z^{(j)}\colon U^{(j)} \to \mathbb{C}, \quad j = 1, \dots, n,$$

such that each range $z^{(j)}[U^{(j)}]$ contains the closed unit disk. For any radius $0 < r \leq 1$ and $j \in \{1, \dots, n\}$, we define the disk $U_r^{(j)} = \{P \in U^{(j)} \mid |z^{(j)}(P)| < r\}$. We fix a radius $0 < r_1 < 1$ once and for all. Our aim is to prove the following result.

18.1.1 Theorem *Assume that the open sets $U_{r_1}^{(j)}$ with j in $\{1, \dots, n\}$ cover X. Next, assume that c_1 is a positive real number such that for all j in $\{1, \dots, n\}$ we have:*

$$\mu \leq c_1 |dz^{(j)} \wedge d\bar{z}^{(j)}| \quad \text{on } U_1^{(j)}.$$

Finally, assume that for all j and k in $\{1, \dots, n\}$:

$$\sup_{U_1^{(j)} \cap U_1^{(k)}} \left| \frac{dz^{(j)}}{dz^{(k)}} \right| \leq M$$

holds with some constant $M \geq 1$. Then for some positive constants c_7, c_9, c_{10} and c_{11}, depending only on r_1, we have, for all a in X:

$$g_{a,\mu} \leq n(c_{10} + c_1 c_{11} + c_7 \log M) + \frac{\log 2}{2\pi}$$

and, for all j such that $a \in U_{r_1}^{(j)}$:

$$\lim_{x \to a} \left| g_{a,\mu}(x) - \frac{1}{2\pi} \log |z^{(j)}(x) - z^{(j)}(a)| \right| \leq n(c_{10} + c_1 c_{11} + c_7 \log M) + \frac{\log M}{2\pi} + c_9.$$

We start by considering just one coordinate $z = z^{(j)}$ for a fixed j. To simplify the notation in this section, we drop the superscript (j) in $U = U^{(j)}$, $z = z^{(j)}$, and so on. We fix three radii $0 < r_1 < r_2 < r_3 < 1$ once and for all. The radii r_2 and r_3 should depend only on r_1; e.g. $r_2 = (2r_1 + 1)/3$, $r_3 = (r_1 + 2)/3$ is an admissible choice. Furthermore, we fix a partition of unity: let $\chi : X \to [0, 1]$ be a C^∞ function with $\chi = 0$ on $X - U_1$, $\chi = 1$ on $\overline{U_{r_2}}$, and set $\chi^c = 1 - \chi$. More specifically, we take $\chi = \tilde{\chi}(|z|)$ on U_1 with a smooth function $\tilde{\chi} : \mathbb{R} \to [0, 1]$ such that $\tilde{\chi}(r) = 0$ for $r \geq 1$, and $\tilde{\chi}(r) = 1$ for $r \leq r_2$. The shape function $\tilde{\chi}$ may be taken independently of X and the choice of the coordinate z, only depending on r_2.

We shall use the 2-norm of a (real valued) 1-form ω over a measurable set $Y \subseteq X$ defined by:

$$\|\omega\|_Y := \left(\int_Y \omega \wedge *\omega\right)^{1/2} = \left(2i \int_Y \omega_{1,0} \wedge \omega_{0,1}\right)^{1/2},$$

where $\omega = \omega_{1,0} + \omega_{0,1}$ is the decomposition of ω in its components in $T_{(1,0)}X$ and $T_{(0,1)}X$. In the case $Y = X$, we just write $\|\omega\|$ for $\|\omega\|_X$.

Given a and b in U_{r_1}, we define the following function, having logarithmic singularities in a and b:

$$f_{a,b} := \frac{1}{2\pi} \log \left| \frac{(z - z(a))(\overline{z(a)}z - 1)}{(z - z(b))(\overline{z(b)}z - 1)} \right| \quad \text{on } U_1.$$

Note that the singularities at $1/\overline{z(a)}$ and $1/\overline{z(b)}$ do not lie within the unit disk. We note that:

$$d * df_{a,b} = \delta_a - \delta_b$$

holds on U_1 in the sense of distributions, and that $f_{a,b}$ fulfills Neumann boundary conditions on ∂U_1. To see this, note that the meromorphic function on U_1:

$$\frac{(z - z(a))(\overline{z(a)}z - 1)}{(z - z(b))(\overline{z(b)}z - 1)} = \frac{(z - z(a))(\frac{1}{z} - \overline{z(a)})}{(z - z(b))(\frac{1}{z} - \overline{z(b)})}$$

takes positive real values on ∂U_1.

Finally, for $a \in U_{r_1}$, we set:

$$l_a := \frac{1}{2\pi} \chi \log |z - z(a)|$$

on U_1, extended by 0 to X.

Our first step in the proof of Theorem 18.1.1 is the following key lemma.

18.1.2 Lemma *For a and b in U_{r_1}, the supremum $\sup_X |g_{a,b} - l_a + l_b|$ is bounded by a constant $c_2 = c_4 + c_1 c_5$, with c_4, c_5 depending only on r_1.*

18.1.3 Remark Note that $g_{a,b} - l_a + l_b$ has removable singularities at a and b, since the logarithmic singularities cancel. The constant c_2 is uniform in the choice of the Riemann surface X, and uniform in the choice of $a, b \in U_{r_1}$. The choice of the coordinate z influences c_2 only via the dependence of c_1 on the choice of z. The radii r_2, r_3 and the shape function $\tilde{\chi}$ are viewed as r_1-dependent parameters; this is why we need not emphasise in the lemma that c_2 also depends on these quantities.

Proof (of Lemma 18.1.2) We define the 2-form:

$$u_{a,b} := d * d(\chi^c f_{a,b}) \quad \text{on } U_1 - U_{r_1}$$

and extend it by 0 to the whole surface X. Note that $u_{a,b}$ is supported in $\overline{U_1} - U_{r_2}$, since χ^c varies only there, and since $f_{a,b}$ is harmonic. Consider the variational principle on square integrable 1-forms:

$$\|\omega\|^2 \stackrel{!}{=} \text{minimal}$$

with the constraint:

$$d * \omega = u_{a,b}$$

in the sense of distributions. Writing the constraint with test functions, we see that the minimisation problem is taken over the following closed affine linear subspace of $L^2(X, T^*X)$:

$$V = \{\omega \in L^2(X, T^*X) : \ -\int_X dg \wedge *\omega = \int_X g u_{a,b} \text{ for all } g \in C^\infty(X)\}.$$

The space V is nonempty, since $\tilde{\omega}_{a,b} \in V$ holds for the following 1-form:

$$\tilde{\omega}_{a,b} = \begin{cases} d(\chi^c f_{a,b}) & \text{on } U_1 - U_{r_1}, \\ 0 & \text{otherwise.} \end{cases}$$

Indeed, using Stokes' theorem, we have:

$$\begin{aligned} -\int_X dg \wedge *\tilde{\omega}_{a,b} &= -\int_{U_1} dg \wedge *\tilde{\omega}_{a,b} \\ &= -\int_{\partial U_1} g * \tilde{\omega}_{a,b} + \int_{U_1} g\, d * \tilde{\omega}_{a,b}. \end{aligned}$$

The first summand in the last expression vanishes by the Neumann boundary conditions of $\chi^c f_{a,b} = f_{a,b}$ on ∂U_1, and the second summand equals:

$$\int_{U_1} g\, d * \tilde{\omega}_{a,b} = \int_X g u_{a,b}$$

by the definition of $u_{a,b}$.

Our minimisation problem has a unique solution $\omega_{a,b} \in V$. It fulfills:

$$\int_X \omega_{a,b} \wedge \sigma = 0 \quad \text{for all closed } C^\infty \text{ 1-forms } \sigma. \tag{18.1.4}$$

Indeed: if $d\sigma = 0$, then $\omega_{a,b} + t * \sigma \in V$ holds for all $t \in \mathbb{R}$, since $\omega_{a,b} \in V$ and $d * (*\sigma) = -d\sigma = 0$. Thus:

$$0 = \frac{d}{dt}\|\omega_{a,b} + t * \sigma\|^2\Big|_{t=0} = -2\int_X \omega_{a,b} \wedge \sigma.$$

In particular,

$$\int_X \omega_{a,b} \wedge dg = 0$$

for all $g \in C^\infty(X)$, i.e. $d\omega_{a,b} = 0$ in the sense of distributions. Since $d * \omega_{a,b} = u_{a,b}$ and $d\omega_{a,b} = 0$, we get that $\omega_{a,b}$ is smooth by elliptic regularity, and then (18.1.4) implies that $\omega_{a,b}$ is exact:

$$\omega_{a,b} = d\tilde{g}_{a,b}$$

for some $\tilde{g}_{a,b} \in C^\infty(X)$; see for example [Fo], Corollary 19.13. We normalise $\tilde{g}_{a,b}$ such that:

$$\int_X \tilde{g}_{a,b}\mu = 0, \tag{18.1.5}$$

to make it uniquely determined.

Now we remark that:

$$g_{a,b} = \tilde{g}_{a,b} + \chi f_{a,b} - \int_X \chi f_{a,b}\mu.$$

Indeed, first of all:

$$d * dg_{a,b} = \delta_a - \delta_b.$$

This is clear outside the support of χ, since there $g_{a,b} = \tilde{g}_{a,b} - \int_X \chi f_{a,b}\mu$ is harmonic. On U_1, we have:

$$d * dg_{a,b} = u_{a,b} - u_{a,b} + \delta_a - \delta_b = \delta_a - \delta_b;$$

note that $d * d(\chi^c f_{a,b}) = u_{a,b}$ holds on U_1, since $f_{a,b}$ is harmonic on $U_1 \cap \mathrm{Supp}(\chi^c)$. Furthermore:

$$\int_X g_{a,b}\mu = 0$$

holds by construction.

The function:

$$g^{(1)}_{a,b} = \tilde{g}_{a,b} + \chi f_{a,b}$$

is harmonic on $X - \{a, b\}$, and:

$$g^{(2)}_{a,b} = \tilde{g}_{a,b} - \chi^c f_{a,b}$$

is harmonic on U_1; in particular both functions are harmonic on the annulus $A := U_1 - \overline{U_{r_2}}$. Now for every harmonic function g on A, we have a bound:

$$\max_{\partial U_{r_3}} g - \min_{\partial U_{r_3}} g \leq c_3 \|dg\|_A$$

with some positive constant c_3 depending only on r_2 and r_3; note that the circle ∂U_{r_3} is relatively compact in the annulus A. We bound $\|dg^{(2)}_{a,b}\|_A$ from above:

$$\|dg^{(2)}_{a,b}\|_A \leq \|d\tilde{g}_{a,b}\|_A + \|d(\chi^c f_{a,b})\|_A.$$

We estimate the first summand as follows, using that $\omega_{a,b} = d\tilde{g}_{a,b}$ solves the above variational problem:

$$\|d\tilde{g}_{a,b}\|_A \leq \|d\tilde{g}_{a,b}\| = \|\omega_{a,b}\| \leq \|\tilde{\omega}_{a,b}\| = \|\tilde{\omega}_{a,b}\|_A = \|d(\chi^c f_{a,b})\|_A;$$

we used that $\tilde{\omega}_{a,b}$ is supported in A. Thus we have:

$$\|dg^{(2)}_{a,b}\| \leq 2\|d(\chi^c f_{a,b})\|_A,$$

which is bounded by a constant, uniformly in a and b in U_{r_1}.

This also allows us to estimate $g^{(1)}_{a,b}$: on A, we know $g^{(1)}_{a,b} = g^{(2)}_{a,b} + f_{a,b}$, hence:

$$\|dg^{(1)}_{a,b}\|_A \leq \|dg^{(2)}_{a,b}\|_A + \|df_{a,b}\|_A \leq 2\|d(\chi^c f_{a,b})\|_A + \|df_{a,b}\|_A.$$

Both summands on the right hand side are bounded by constants, only depending on r_1 and r_2, but uniformly in a and b in U_{r_1}. To summarise, we have shown that:

$$\max_{\partial U_{r_3}} g^{(j)}_{a,b} - \min_{\partial U_{r_3}} g^{(j)}_{a,b}$$

($j = 1, 2$) are uniformly bounded by a constant depending only on r_1, r_2, and r_3. However, $g^{(1)}_{a,b}$ is harmonic on $X - U_{r_3}$, and $g^{(2)}_{a,b}$ is harmonic on $\overline{U_{r_3}}$, which both have the same boundary ∂U_{r_3}. Thus, by the maximum principle,

$$\max_{X-U_{r_3}} g^{(1)}_{a,b} - \min_{X-U_{r_3}} g^{(1)}_{a,b} = \max_{\partial U_{r_3}} g^{(1)}_{a,b} - \min_{\partial U_{r_3}} g^{(1)}_{a,b}$$

and:

$$\max_{\overline{U_{r_3}}} g^{(2)}_{a,b} - \min_{\overline{U_{r_3}}} g^{(2)}_{a,b} = \max_{\partial U_{r_3}} g^{(2)}_{a,b} - \min_{\partial U_{r_3}} g^{(2)}_{a,b}.$$

Furthermore, $\max_{X-U_{r_3}} |\chi f_{a,b}|$ and $\max_{\overline{U_{r_3}}} |\chi^c f_{a,b}|$ are bounded, uniformly in a and b in U_{r_1}, by a constant only depending on r_1 and r_3. Using $\tilde{g}_{a,b} = g^{(1)}_{a,b} - \chi f_{a,b}$ on $X - U_{r_3}$ and $\tilde{g}_{a,b} = g^{(2)}_{a,b} + \chi^c f_{a,b}$ on $\overline{U_{r_3}}$, we conclude that $\max_X \tilde{g}_{a,b} - \min_X \tilde{g}_{a,b}$ is bounded on $X = (X - U_{r_3}) \cup \overline{U_{r_3}}$ by a constant c_6 only depending on the radii r_1, r_2 and r_3. Using the normalisation condition (18.1.5), we know that:

$$\max_X \tilde{g}_{a,b} \geq 0 \geq \min_X \tilde{g}_{a,b}$$

holds; thus:

$$\max_X |\tilde{g}_{a,b}| \leq \max_X \tilde{g}_{a,b} - \min_X \tilde{g}_{a,b}$$

is also bounded by the same constant.

From this we get a bound for:

$$g_{a,b} - \chi f_{a,b} = \tilde{g}_{a,b} - \int_X \chi f_{a,b} \mu.$$

Indeed, we estimate:

$$\left|\int_X \chi f_{a,b}\mu\right| \leq \int_{U_1} |f_{a,b}|\mu \leq c_1 \int_{U_1} |f_{a,b}\, dz \wedge d\bar{z}|,$$

which is uniformly bounded for $a, b \in U_{r_1}$ by a constant $c_1 c_5$ with c_5 depending only on r_1; note that the logarithmic singularities are integrable. Combining the bounds for $\max_X |\tilde{g}_{a,b}|$ and $\left|\int_X \chi f_{a,b}\mu\right|$, we conclude that $\sup_X |g_{a,b} - \chi f_{a,b}|$ is bounded by a constant $c_6 + c_1 c_5$ with c_6, c_5 depending on r_1. Since:

$$\sup_X |\chi f_{a,b} - l_a + l_b| = \frac{1}{2\pi} \sup_{U_1} \left| \chi \log \left| \frac{\overline{z(a)}z - 1}{\overline{z(b)}z - 1} \right| \right|$$

is bounded, uniformly in $a, b \in U_{r_1}$ and X, the key lemma follows (with c_4 being the sum of c_6 and the uniform upper bound last mentioned). □

Proof (of Theorem 18.1.1) Since we now work with varying coordinates, we include again the superscript coordinate index (j) in the coordinate $z^{(j)}$, its domain $U^{(j)}$, but also in $U_r^{(j)}$, $\chi^{(j)}$, and $l_a^{(j)}$.

18.1.6 Lemma *Consider two coordinates $z^{(j)}$ and $z^{(k)}$, with $1 \leq j, k \leq n$. Assume that $x \in U_{r_1}^{(j)} \cap U_{r_1}^{(k)}$ and $y \in U_{r_2}^{(j)}$ with $|z^{(j)}(y) - z^{(j)}(x)| < (r_2 - r_1)/M$. Then $y \in U_{r_2}^{(k)}$ holds.*

Proof The intersection $U_{r_1}^{(j)} \cap U_{r_1}^{(k)}$ is an open neighbourhood of x. Assume that there exists $y \in \overline{U}_{r_2}^{(j)}$ with $|z^{(j)}(y) - z^{(j)}(x)| < (r_2 - r_1)/M$ and $y \notin U_{r_2}^{(k)}$. Then there is also such a point y with minimal distance $|z^{(j)}(y) - z^{(j)}(x)|$ from x, since $\overline{U}_{r_2}^{(j)} - U_{r_2}^{(k)}$ is compact. For this point y, we conclude $y \in \partial U_{r_2}^{(k)} \subseteq \overline{U}_{r_2}^{(k)}$, and the straight line from x to y in the $z^{(j)}$-coordinate is contained in $\overline{U}_{r_2}^{(j)} \cap \overline{U}_{r_2}^{(k)}$. By the mean value theorem, we conclude $|z^{(k)}(y) - z^{(k)}(x)| \leq M|z^{(j)}(y) - z^{(j)}(x)| < r_2 - r_1$, hence $|z^{(k)}(y)| < r_2$, since $|z^{(k)}(x)| \leq r_1$. This contradicts $y \in \partial U_{r_2}^{(k)}$. □

We choose a smooth partition of unity $\phi^{(j)} : X \to [0, 1]$, $j = 1, \ldots, n$, such that $\phi^{(j)}$ is supported in $U_{r_1}^{(j)}$. For $a \in X$, we set:

$$h_a := \sum_j \phi^{(j)}(a) l_a^{(j)}.$$

18.1.7 Lemma *Let $a \in U_{r_1}^{(k)}$, $y \in X$, $y \neq a$. Then we have:*

$$l_a^{(k)}(y) \leq \frac{\log 2}{2\pi}.$$

Proof This follows immediately from the definition of $l_a^{(k)}$, since $|z^{(k)}(y) - z^{(k)}(a)| \leq 2$ whenever $y \in \mathrm{Supp}(\chi^{(k)})$. □

18.1.8 Lemma *For all $a, b \in X$ we have the inequality:*

$$\sup_X |g_{a,b} - h_a + h_b| \leq n(c_{10} + c_1 c_5 + c_7 \log M)$$

with constants c_{10}, c_5, and c_7 depending only on r_1.

Proof We first show for $a \in U_{r_1}^{(k)} \cap U_{r_1}^{(j)}$ that:

$$\sup_X |l_a^{(k)} - l_a^{(j)}| \leq \frac{1}{2\pi}[\log M + |\log(r_2 - r_1)| + \log 2].$$

To prove this, let $y \in X$. We distinguish 3 cases to prove that $l_a^{(k)}(y) - l_a^{(j)}(y)$ is bounded from above by the right hand side.

case 1: $y \in U_1^{(j)}$ with $|z^{(j)}(y) - z^{(j)}(a)| < (r_2 - r_1)/M$. In particular, we have $|z^{(j)}(y)| < |z^{(j)}(a)| + (r_2 - r_1)/M \leq r_2$ (recall that $M \geq 1$), hence $a, y \in U_{r_2}^{(j)}$. Consequently, the straight line $[a, y]^{(j)}$ from a to y in the $z^{(j)}$-coordinate is contained in $U_{r_2}^{(j)}$. Then Lemma 18.1.6 implies that $[a, y]^{(j)} \subseteq U_{r_2}^{(k)}$. Using $\chi^{(j)}(y) = \chi^{(k)}(y) = 1$, we conclude by the mean value theorem that:

$$l_a^{(k)}(y) - l_a^{(j)}(y) = \frac{1}{2\pi} \log \left| \frac{z^{(k)}(y) - z^{(k)}(a)}{z^{(j)}(y) - z^{(j)}(a)} \right| \leq \frac{\log M}{2\pi},$$

which is bounded by the right hand side.

case 2: $y \notin U_1^{(j)}$. Then $l_a^{(j)}(y) = 0$, and we conclude, using Lemma 18.1.7, that:

$$l_a^{(k)}(y) - l_a^{(j)}(y) = l_a^{(k)}(y) \leq \frac{\log 2}{2\pi}.$$

case 3: $y \in U_1^{(j)}$ and $|z^{(j)}(y) - z^{(j)}(a)| \geq (r_2 - r_1)/M$; thus:

$$l_a^{(k)}(y) - l_a^{(j)}(y) \leq \frac{\log 2}{2\pi} - l_a^{(j)}(y) \leq \frac{1}{2\pi}(\log 2 - \chi^{(j)}(y) \log[(r_2 - r_1)/M]),$$

which is also bounded by the right hand side.

The upper bound for $l_a^{(j)}(y) - l_a^{(k)}(y)$ in our claim is obtained by exchanging j and k. Thus the claim is proven.

We conclude:

$$|h_a - l_a^{(j)}| \leq \sum_k \phi^{(k)}(a)|l_a^{(k)} - l_a^{(j)}| \leq \frac{1}{2\pi}\left(\log M + |\log(r_2 - r_1)| + \log 2\right). \tag{18.1.9}$$

Combining this with Lemma 18.1.2, we conclude for $a, b \in U_{r_1}^{(j)}$:

$$\begin{aligned} |g_{a,b} - h_a + h_b| &\leq |g_{a,b} - l_a^{(j)} + l_b^{(j)}| + |h_a - l_a^{(j)}| + |h_b - l_b^{(j)}| \\ &\leq c_{10} + c_1 c_5 + c_7 \log M \end{aligned}$$

with some constants c_{10}, c_5, c_7 depending only on r_1 (one can take $c_7 = (\log M)/\pi$ and $c_{10} = (|\log(r_2 - r_1)| + \log 2)/\pi + c_4$).

Finally, for general $a, b \in X$, we choose a finite sequence of points $a = a_0, a_1, \ldots, a_m = b$ in X and indices $j_1, \ldots, j_m$ with $m \leq n$ and $a_{i-1}, a_i \in U_{r_1}^{(j_i)}$ for all $i = 1, \ldots, m$. Using:

$$g_{a,b} = \sum_{i=1}^{m} g_{a_{i-1},a_i},$$

we get by estimating:

$$|g_{a,b} - h_a + h_b| \leq \sum_{i=1}^{m} |g_{a_{i-1},a_i} - h_{a_{i-1}} + h_{a_i}| \leq n(c_{10} + c_1 c_5 + c_7 \log M)$$

the claim of the lemma. □

We define:

$$h_\mu(x) := \int_{b \in X} h_b(x)\, \mu(b), \qquad (x \in X).$$

18.1.10 Lemma *We have:*

$$\sup_X |h_\mu| \leq n c_1 c_8,$$

with some universal constant c_8. *Furthermore, we have:*

$$\sup_{\substack{b,x \in X \\ b \neq x}} h_b(x) \leq \frac{\log 2}{2\pi}.$$

Proof We observe first that for all $w \in \mathbb{C}$ with $|w| \leq 1$ the integral:

$$\frac{1}{2\pi} \int_{|z| \leq 1} |\log|z - w|| \, |dz \wedge d\bar{z}|$$

is bounded from above by a universal constant c_8. We conclude that for all $x \in X$ we have:

$$\int_{b\in U_{r_1}^{(j)}} |l_b^{(j)}(x)|\phi^{(j)}(b)\,\mu(b) \le c_1 \int_{b\in U_{r_1}^{(j)}} |l_b^{(j)}(x)|\,|dz^{(j)} \wedge d\overline{z^{(j)}}| \le c_1 c_8.$$

Let $x \in X$. We get the first estimate:

$$|h_\mu(x)| \le \sum_{j=1}^{n} \int_{U_1^{(j)}} |l_b^{(j)}(x)|\phi^{(j)}(b)\,\mu(b) \le n c_1 c_8.$$

Finally, the second estimate follows from Lemma 18.1.7:

$$h_b = \sum_{j=1}^{n} \phi^{(j)} l_b^{(j)} \le \frac{\log 2}{2\pi},$$

as required. □

18.1.11 Proposition *For some positive constants c_{10}, c_7, and c_{11} that depend only on r_1 we have, uniformly in a and $x \neq a$ on X:*

$$|g_{a,\mu}(x) - h_a(x)| \le n(c_{10} + c_1 c_{11} + c_7 \log M).$$

Proof Indeed, averaging Lemma 18.1.8 over b with respect to μ, we obtain:

$$\sup_X |g_{a,\mu} - h_a + h_\mu| \le n(c_{10} + c_1 c_5 + c_7 \log M).$$

By Lemma 18.1.10, one has $|h_\mu| \le n c_1 c_8$. Combining gives what we want (we can take $c_{11} = c_5 + c_8$). □

18.1.12 Proposition *Let c_{10}, c_7, and c_{11} be as in Proposition 18.1.11, and let a be in X. Then $\lim_{x\to a} |g_{a,\mu}(x) - h_a(x)|$ exists, and we have:*

$$\lim_{x\to a} |g_{a,\mu}(x) - h_a(x)| \le n(c_{10} + c_1 c_{11} + c_7 \log M).$$

Proof The functions $g_{a,\mu}$ and h_a have the same logarithmic singularity at a; hence the limit exists. The estimate then follows from Proposition 18.1.11. □

We can now finish the proof of Theorem 18.1.1. We have seen in (18.1.9) that:

$$|h_a - l_a^{(j)}| \le \frac{1}{2\pi}\left(\log M + |\log(r_2 - r_1)| + \log 2\right).$$

Combining this with Proposition 18.1.12 and using the definition of $l_a^{(j)}$ gives the second estimate of the theorem. As to the first estimate, using:

$$g_{a,\mu} \le h_a + |g_{a,\mu} - h_a|$$

we obtain it by applying the upper bound for h_a in Lemma 18.1.10 and the upper bound for $|g_{a,\mu} - h_a|$ in Proposition 18.1.11. This ends the proof of Theorem 18.1.1. □

18.2 Application of Merkl's result to the modular curves $X_1(pl)$

In this subsection we will give a suitable upper bound for the Arakelov-Green functions $g_{a,\mu}$ (see Proposition 14.4.3) on the modular curves $X_1(pl)$ with p and l distinct primes.

18.2.1 Theorem *There is a real number c such that for all pairs of distinct prime numbers p and l for which the genus of $X_1(pl)$ is at least one and for all distinct a and b on $X_1(pl)(\mathbb{C})$ we have:*

$$g_{a,\mu}(b) \leq c \cdot (pl)^6.$$

Let ∞ denote the cusp ∞ on $X_1(pl)$, and let q be the standard local coordinate around ∞ given by the map $\tau \mapsto \exp(2\pi i \tau)$ from the region $\Im\tau > 1$ in $\mathbb{H}$ to $\mathbb{C}$. Then we have:

$$|\log \|dq\|_{\mathrm{Ar}}(\infty)| = O((pl)^6),$$

where $\|\cdot\|_{\mathrm{Ar}}$ denotes the Arakelov metric on Ω^1 (see Section 14.4).

Proof We write for the moment N for pl. We will apply Theorem 18.1.1, but we will carry out the estimates on the more symmetrical modular curve $X(N)$ which for us is $\Gamma(N)\setminus(\mathbb{H}\cup\mathbb{P}^1(\mathbb{Q}))$. Let $h\colon X(N) \to X_1(N)$ be the canonical map; it has degree N. We let μ denote the Arakelov $(1,1)$-form on $X_1(N)$, and we define $\mu' = h^*\mu/N$. The characterising properties of Green functions directly imply that:

$$h^* g_{a,\mu} = \sum_{h(b)=a} g_{b,\mu'}, \tag{18.2.2}$$

where the b are counted with multiplicity.

As in Section 18.1 we fix a constant r_1 with $0 < r_1 < 1$; we take $r_1 := 3/4$. We need to construct an atlas with charts $z^{(j)}\colon U^{(j)} \to \mathbb{C}$ for $X(N)$ with all $z^{(j)}(U^{(j)})$ containing the closed unit disk and with the $U^{(j)}_{r_1}$ covering $X(N)$.

We start with a construction of a local coordinate $z\colon U \to \mathbb{C}$ in a neighbourhood of the standard cusp ∞. Since $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ acts transitively on the set of cusps of $X(N)$, this construction will suffice to give the full atlas. Our initial coordinate is induced by the map z from $\mathbb{H}$ to $\mathbb{C}$ that sends τ to $e^{2\pi i\tau/N}$.

18.2.3 Lemma *Let $n \geq 1$. Then the subset in $\mathbb{H}$ given by the conditions $-1/2 \leq \Re\tau < n - 1/2$ and $\Im\tau > 1/n$ is mapped injectively to $X(n)(\mathbb{C})$.*

Proof First we note that for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z})$ and for τ in $\mathbb{H}$ we have:

$$\Im\frac{a\tau+b}{c\tau+d} = \frac{\Im\tau}{|c\tau+d|^2}.$$

Suppose then that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is in $\Gamma(n)$, that $c \neq 0$ and that $\Im\tau > y$ for some real $y > 0$. Then we have $|c| \geq n$ because $n|c$, and:

$$\Im\frac{a\tau+b}{c\tau+d} = \frac{\Im\tau}{|c\tau+d|^2} \leq \frac{\Im\tau}{(\Im c\tau)^2} \leq \frac{\Im\tau}{n^2(\Im\tau)^2} < \frac{1}{n^2}\frac{1}{y}.$$

It follows that the optimal choice for y is $1/n$. □

In particular, the region of τ with $-1/2 \leq \Re\tau < N - 1/2$ and $\Im\tau > 1/2$ is mapped injectively into $X(N)$ to give an open neighbourhood U of ∞. We could replace the condition "$\Im\tau > 1/2$" replaced by "$\Im\tau > 1/N$" but that would not make the work to be done significantly easier. The map $\tau \mapsto e^{2\pi i\tau/N}$ gives an isomorphism:

$$z\colon U \longrightarrow D(0, e^{-\pi/N}) \subset \mathbb{C}. \tag{18.2.4}$$

The region of τ with $-1/2 \leq \Re\tau < N - 1/2$ and $\Im\tau > 3/4$ gives an open neighbourhood V of ∞, contained in U, such that the translates of V under $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ cover $X(N)$ (note that $3/4 < \sqrt{3}/2$). The image of V under z is the disk $D(0, e^{-3\pi/2N})$. However, the quotient of the radii $e^{-\pi/N}$ and $e^{-3\pi/2N}$ tends to 1 as N tends to infinity, hence we cannot work with these disks directly.

What we do is the following. We define $z' := e^{3\pi/2N}z$ to get $z'V = D(0,1)$. Then we have $z'U = D(0, e^{\pi/2N})$. Let $\varepsilon := \varepsilon(N) := e^{\pi/2N} - 1$ be the difference between the two new radii. Then $\varepsilon > \pi/2N$. We can choose $O(\varepsilon^{-2})$ open disks $D(a,\varepsilon)$ with centre a in $D(0,1)$, such that the union of the $D(a,\varepsilon/2)$ contains $D(0,1)$; we let A denote the set of these a. The $D(a,\varepsilon)$ are contained in $z'U = D(0, 1+\varepsilon)$.

The group $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ acts transitively on the set of cusps of $X(N)$. For each cusp c, we choose a g_c in $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ such that $c = g_c\infty$. The open sets of our atlas for $X(N)$ are then the $U^{(a,c)}$ with a in A and c a cusp, defined by:

$$U^{(a,c)} := g_c \cdot (z')^{-1} D(a,\varepsilon).$$

The required coordinates $z^{(a,c)}$ on the $U^{(a,c)}$ are defined by the composition of isomorphisms:

$$z^{(a,c)}\colon U^{(a,c)} \xrightarrow{g_c^{-1}} U^{(a,\infty)} \xrightarrow{z'} D(a,\varepsilon) \xrightarrow{-a} D(0,\varepsilon) \xrightarrow{\cdot 3/2\varepsilon} D(0,3/2).$$

Indeed, the images $z^{(a,c)}U^{(a,c)}$ contain the unit disk, and $U_{r_1}^{(a,c)}$ corresponds via $z' \circ g_c^{-1}$ to the subdisk $D(a,\varepsilon/2)$ of $D(a,\varepsilon)$, hence the $U_{r_1}^{(a,c)}$ cover $X(N)$. The exact number $n = n(N)$ of $U^{(a,c)}$ is the cardinality of A times the number of cusps, hence $n = O(N^4)$. We choose a numbering of $A \times \{\text{cusps}\}$ with the integers $\{1,\ldots,n\}$, and we will denote our charts as:

$$z^{(j)}\colon U^{(j)} \to \mathbb{C}. \tag{18.2.5}$$

18.2.6 Lemma *For the local coordinates $z^{(j)}\colon U^{(j)} \to \mathbb{C}$ on $X(N)$ that we have just defined, the following holds. For all j and k in $\{1,\dots,n\}$ we have:*

$$\sup_{U_1^{(j)} \cap U_1^{(k)}} \left| \frac{dz^{(j)}}{dz^{(k)}} \right| \leq M,$$

with $M = 6$.

Proof Let j and k be in $\{1,\dots,n\}$. If j and k arise from the same cusp, then $z^{(j)}$ and $z^{(k)}$ differ by a translation, hence $dz^{(j)}/dz^{(k)} = 1$. Now suppose that j and k arise from two distinct cusps. We may suppose then, by acting with an element of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, that k arises from the standard cusp ∞. Let x denote the cusp that j arises from. The coordinate $z^{(j)}$ is then obtained as above from an element g_x of $\mathrm{SL}_2(\mathbb{Z})$ that sends ∞ to x. Let us write $g_x^{-1} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Note that $c \neq 0$, hence $|c| \geq 1$. Let z be a point in $\mathbb{H}$ with $-1/2 \leq \Re z < N - 1/2$ that maps to an element in $U_1^{(j)} \cap U_1^{(k)}$. Then we know that $1/2 < \Im z < 1$ because disks given by $\Im z > 1$ around different cusps do not meet at all. Likewise, we then know that:

$$1/2 < \Im \frac{az+b}{cz+d} = \frac{\Im z}{|cz+d|^2} < 1.$$

Hence, as $(\Im z)/|cz+d|^2 \leq (\Im z)/c^2(\Im z)^2$, we have $\Im z < 2/c^2$ which gives in fact $|c| = 1$. Under these conditions, we estimate:

$$\left| \log \left| \frac{d \exp(2\pi i \frac{az+b}{cz+d}/N)}{d \exp(2\pi i z/N)} \right| \right| = \left| \log \left| \frac{\exp(2\pi i \frac{az+b}{cz+d}/N) d\frac{az+b}{cz+d}}{\exp(2\pi i z/N) dz} \right| \right| =$$

$$= \left| \log \left| \exp(2\pi i \frac{az+b}{cz+d}/N) \right| - \log |\exp(2\pi i z/N)| - \log |cz+d|^2 \right| \leq$$

$$\leq 4\pi/N + 4\pi/N + \log 4.$$

So indeed, for as $N \geq 6$, we can take $M = 6$. Some explanations are perhaps in order here: as $\Im z$ and $\Im(az+b)/(cz+d)$ are between $1/2$ and 2, $|\exp(2\pi i z/N)|$ and $|\exp(2\pi i \frac{az+b}{cz+d}/N)|$ are between $\exp(-4\pi/N)$ and $\exp(-\pi/N)$. As $\Im(az+b)/(cz+d) = (\Im z)/|cz+d|^2$, we see that $|cz+d|^2$ is between $1/4$ and 4. □

Our next task is to produce a suitable bound of the type $\mu \leq c_1 |dz^{(j)} \wedge d\bar{z}^{(j)}|$ as in Theorem 18.1.1. We start with a bound for μ on disks around ∞ on $X_1(pl)$.

18.2.7 Lemma *Let r be a real number such that $0 < r < 1$. We map $D(0,r)$ to $X_1(pl)$ by sending $q \neq 0$ to $(\mathbb{C}^\times/q^{\mathbb{Z}}, \zeta_{pl})$. The image of this map is the image in $X_1(pl)$ of the region*

in $\mathbb{H}$ defined by the condition "$\Im\tau > -(\log r)/2\pi$", plus the cusp ∞. We still denote by μ the $(1,1)$-form on $D(0,r)$ induced by μ. Then we have, on $D(0,r)$:

$$\mu \leq \frac{28e^{4\pi}}{\pi}\frac{1}{(1-r)^4}\cdot\frac{i}{2}dqd\overline{q}.$$

Proof We first bound the functions $\sum_{n\geq 1}a_n(f)q^{n-1}$ and $\sum_{n\geq 1}a_n(f)q^{nl-1}$ for a newform f on the disk $D(0,r)$. We have, for $|q|<r$ by Lemma 16.1:

$$|\sum_{n\geq 1}a_n(\omega)q^{n-1}| \leq \sum_{n\geq 1}|a_n(\omega)|r^{n-1} \leq 2\sum_{n\geq 1}nr^{n-1} = \frac{2}{(1-r)^2}\,,$$

and next:

$$|\sum_{n\geq 1}a_n(\omega)q^{nl-1}| \leq 2\sum_{n\geq 1}nr^{nl-1} = \frac{2r^{l-1}}{(1-r^l)^2} \quad \text{for } |q|<r.$$

Now recall from Corollary 16.6 that for μ we have the expression:

$$\begin{aligned}\mu = &\sum_{\omega}\frac{i}{2g\|\omega\|^2}\omega\wedge\overline{\omega}\\ &+\sum_{\omega}\frac{(i/2g)\left((p+1)s_p^*\omega\wedge\overline{s_p^*\omega}+(p+1)t_p^*\omega\wedge\overline{t_p^*\omega}-a_p(f)s_p^*\omega\wedge\overline{t_p^*\omega}-\overline{a_p(f)}t_p^*\omega\wedge\overline{s_p^*\omega}\right)}{(p-1)\|\omega\|^2((p+1)^2-|a_p(f)|^2)}\\ &+\sum_{\omega}\frac{(i/2g)\left((l+1)s_l^*\omega\wedge\overline{s_l^*\omega}+(l+1)t_l^*\omega\wedge\overline{t_l^*\omega}-a_l(f)s_l^*\omega\wedge\overline{t_l^*\omega}-\overline{a_l(f)}t_l^*\omega\wedge\overline{s_l^*\omega}\right)}{(l-1)\|\omega\|^2((l+1)^2-|a_l(f)|^2)},\end{aligned}$$

the first sum running over $\Omega^1(X_1(pl))^{\rm new}$, the second sum running over $\Omega^1(X_1(l))^{\rm new}$, and the third sum running over $\Omega^1(X_1(p))^{\rm new}$. We bound the different terms from the above expression for μ. The first term yields for $|q|<r$:

$$\begin{aligned}\frac{i}{2g\|\omega\|^2}\omega\wedge\overline{\omega} &= \frac{1}{g\|\omega\|^2}\left|\sum_{n\geq 1}a_n(f)q^{n-1}\right|^2\cdot\frac{i}{2}dqd\overline{q}\\ &\leq \frac{1}{g}\frac{e^{4\pi}}{\pi}\left(\frac{2}{(1-r)^2}\right)^2\cdot\frac{i}{2}dqd\overline{q}\\ &= \frac{1}{g}\frac{e^{4\pi}}{\pi}\frac{4}{(1-r)^4}\cdot\frac{i}{2}dqd\overline{q}\end{aligned}$$

The contribution of an element ω of $\Omega^1(p)^{\rm new}$ is:

$$\begin{aligned}&\frac{(i/2g)\left((l+1)s_l^*\omega\wedge\overline{s_l^*\omega}+(l+1)t_l^*\omega\wedge\overline{t_l^*\omega}-a_l(f)s_l^*\omega\wedge\overline{t_l^*\omega}-\overline{a_l(f)}t_l^*\omega\wedge\overline{s_l^*\omega}\right)}{(l-1)\|\omega\|^2((l+1)^2-|a_l(f)|^2)} \leq\\ &\leq \frac{1}{g}\frac{e^{4\pi}}{\pi(l-1)^3}\left(\frac{4(l+1)}{(1-r)^4}+\frac{4(l+1)r^{2(l-1)}}{(1-r^l)^4}+\frac{16\sqrt{l}r^{l-1}}{(1-r)^2(1-r^l)^2}\right)\cdot\frac{i}{2}dqd\overline{q}.\end{aligned}$$

Here one uses the Weil bounds on $a_l(f)$. Symmetrically (in p and l), the contribution to μ of an element ω of $\Omega^1(l)^{\mathrm{new}}$ is:

$$\frac{(i/2g)\left((p+1)s_p^*\omega\wedge\overline{s_p^*\omega}+(p+1)t_p^*\omega\wedge\overline{t_p^*\omega}-a_p(f)s_p^*\omega\wedge\overline{t_p^*\omega}-\overline{a_p(f)}t_p^*\omega\wedge\overline{s_p^*\omega}\right)}{(p-1)\|\omega\|^2((p+1)^2-|a_p(f)|^2)}\leq$$
$$\leq\frac{1}{g}\frac{e^{4\pi}}{\pi(p-1)^3}\left(\frac{4(p+1)}{(1-r)^4}+\frac{4(p+1)r^{2(p-1)}}{(1-r^p)^4}+\frac{16\sqrt{p}r^{p-1}}{(1-r)^2(1-r^p)^2}\right)\cdot\frac{i}{2}dqd\overline{q}.$$

Now we sum this all up, over the elements of $\Omega^1(X_1(pl))^{\mathrm{new}}$, $\Omega^1(X_1(p))^{\mathrm{new}}$, and $\Omega^1(X_1(l))^{\mathrm{new}}$. We get for $|q|<r$:

$$\mu\leq\frac{4e^{4\pi}}{\pi}\left(\frac{1}{(1-r)^4}+\frac{1}{(1-r)^4}+\frac{r^{2(l-1)}}{(1-r^l)^4}+\frac{r^{l-1}}{(1-r)^2(1-r^l)^2}+\right.$$
$$\left.\frac{1}{(1-r)^4}+\frac{r^{2(p-1)}}{(1-r^p)^4}+\frac{r^{p-1}}{(1-r)^2(1-r^p)^2}\right)\cdot\frac{i}{2}dqd\overline{q},$$

and finally:

$$\mu\leq\frac{28e^{4\pi}}{\pi}\frac{1}{(1-r)^4}\cdot\frac{i}{2}dqd\overline{q},\quad\text{for }|q|<r.$$

□

Our next step is to consider the disks gU in $X(pl)$, where g is in $\mathrm{SL}_2(\mathbb{Z}/pl\mathbb{Z})$ and where U is as in (18.2.4).

18.2.8 Lemma *Let g be in $\mathrm{SL}_2(\mathbb{Z}/pl\mathbb{Z})$ and let $z\colon U\to D(0,e^{-\pi/pl})$ be as in (18.2.4). Let $z_g:=z\circ g^{-1}\colon gU\to D(0,e^{-\pi/pl})$. Then we have, on gU:*

$$h^*\mu\leq c(pl)^4|dz_gd\overline{z_g}|,$$

with c independent of p and l.

Proof To prove this, we consider the map $h\circ g\circ z^{-1}$ from $D(0,e^{-\pi/pl})$ to $X_1(pl)$ and the pullback of μ to $D(0,e^{-\pi/pl})$. We observe that μ is invariant under all automorphisms of $X_1(pl)$. This applies in particular to the diamond operators and the Atkin-Lehner pseudo-involutions (defined in (10.10)). As the group generated by these automorphisms permutes the cusps of $X_1(pl)$ transitively, we can take such an automorphism α such that $\alpha\circ h\circ g\circ z^{-1}$ sends $D(0,e^{-\pi/pl})$ to a disk around the cusp ∞, where we can then apply Lemma 18.2.7. The pullbacks of μ via $h\circ g\circ z^{-1}$ and $\alpha\circ h\circ g\circ z^{-1}$ are the same.

The map $h\circ g\circ z^{-1}$ sends a point $0\neq q\in D(0,e^{-\pi/pl})$ to the point of $X_1(pl)$ corresponding to $(\mathbb{C}^\times/q^{pl\mathbb{Z}},\zeta_p^a(q^l)^b,\zeta_l^c(q^p)^d)$ for certain a and b in $\mathbb{F}_p$ and c and d in $\mathbb{F}_l$ depending on g. After

replacing h with h composed with a suitable diamond operator, we are in one of four cases, that we will treat one by one.

In the first case, q is mapped to $(\mathbb{C}^\times/q^{pl\mathbb{Z}}, \zeta_p, \zeta_l)$. Then the map $h\circ g\circ z^{-1}$ factors as the cover $D(0,e^{-\pi/pl}) \to D(0,e^{-\pi})$ of degree pl sending z to z^{pl}, followed by the map of Lemma 18.2.7 that sends q to $(\mathbb{C}^\times/q^{\mathbb{Z}}, \zeta_{pl})$. Then we have, on $D(0,e^{-\pi/pl})$:

$$h^*\mu \le \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi})^4}\cdot h^*\left(\frac{i}{2}dqd\overline{q}\right) \le \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi})^4}(pl)^2\cdot\frac{i}{2}dzd\overline{z}.$$

In the second case, q is mapped to $(\mathbb{C}^\times/q^{pl\mathbb{Z}}, q^l, \zeta_l)$. In this case, we compose it with the pseudo-involution w_{ζ_p}, which brings us to $(\mathbb{C}^\times/q^{l\mathbb{Z}}, \zeta_p, \zeta_l)$. The map then factors as the lth power map from $D(0,e^{-\pi/pl})$ to $D(0,e^{-\pi/p})$, followed by the map of Lemma 18.2.7. We find:

$$h^*\mu \le \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi/p})^4}\cdot h^*\left(\frac{i}{2}dqd\overline{q}\right) \le \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi/p})^4}l^2\cdot\frac{i}{2}dzd\overline{z}.$$

The third case is obtained by interchanging the roles of p and l, so we will not make it explicit.

In the fourth case, q is mapped to $(\mathbb{C}^\times/q^{pl\mathbb{Z}}, q)$ (q is now a point of order pl). We compose with the pseudo-involution $w_{\zeta_{pl}}$, which brings us to $(\mathbb{C}^\times/q^{\mathbb{Z}}, \zeta_{pl})$. This is the map of Lemma 18.2.7. We find:

$$h^*\mu \le \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi/pl})^4}\cdot h^*\left(\frac{i}{2}dqd\overline{q}\right) = \frac{28e^{4\pi}}{\pi}\frac{1}{(1-e^{-\pi/pl})^4}\cdot\frac{i}{2}dzd\overline{z}.$$

In these four cases, we see that the factor in front of $(i/2)dzd\overline{z}$ in the upper bound for $h^*\mu$ is $O((pl)^4)$. □

18.2.9 Lemma *For the local coordinates $z^{(j)}$ and the real $(1,1)$-form μ' on $X(pl)$ as defined in (18.2.5) we have:*

$$\mu' \le c_1|dz^{(j)}\wedge d\overline{z}^{(j)}|$$

on $U_1^{(j)}$ with $c_1 = c_1(pl) = O(pl)$.

Proof First of all, we have, by definition: $\mu' = (1/pl)h^*\mu$. The definition of the $z^{(j)}$ (see 18.2.5) plus the definitions $z' = e^{3\pi/2pl}$ and $\varepsilon = e^{\pi/2pl} - 1$ give:

$$dz = \frac{2\varepsilon}{3}e^{3\pi/2pl}\cdot dz^{(j)}.$$

The factor $e^{3\pi/2pl}$ tends to 1 if pl gets large, and $\varepsilon = (\pi/2pl)(1+O(1/pl))$. Combining all this with Lemma 18.2.8 finishes the proof. □

We can now finish the proof of Theorem 18.2.1. We apply Theorem 18.1.1 on $X(pl)$ with the $(1,1)$-form μ'. Then we have $n = O((pl)^4)$, $M = 6$ (Lemma 18.2.6) and $c_1 = O(pl)$ (Lemma 18.2.9). We obtain that there exists a constant c such that $g_{b,\mu'}(b') \leq c{\cdot}(pl)^5$ for all distinct primes p and l, and all distinct b and b' on $X(pl)$. For distinct a and a' on $X_1(pl)$ we then have (see (18.2.2)):

$$g_{a,\mu}(a') = \sum_{h(b)=a} g_{b,\mu'}(h(a')) \leq c{\cdot}(pl)^6.$$

As to the statement that $\log \|dz^{(j)}\|_{\mathrm{Ar}}(P) = O((pl)^6)$, this also follows from Theorem 18.1.1. Indeed, if locally we write $g_{a,\mu} = \log|z - z(a)| + f$ then $f(a) = -\log \|dz\|_{\mathrm{Ar}}(a)$. □

18.3 Bounds for intersection numbers on $X_1(pl)$

In this subsection, we will bound the intersection numbers occurring in the right hand side of the inequality in Theorem 15.2.5, in the situation described in Section 13.

18.3.1 Theorem *Let p and l be two distinct prime numbers such that $X_1(pl)$ is of genus at least two, and let $\mathcal{X}$ be the semistable model over $B := \operatorname{Spec} \mathbb{Z}[\zeta_{pl}]$ provided by [61]. For two cusps P and Q (possibly equal) in $\mathcal{X}(B)$ we have:*

$$(P,P) \leq 0, \quad \textit{and} \quad |(P,Q)| = O((pl)^7).$$

For a cuspidal effective divisor D of degree g on $\mathcal{X}$ we have:

$$|(D, D - \omega_{\mathcal{X}/B})| = O((pl)^{11}).$$

Proof By the adjunction formula we have $-(P,P) = (P, \omega_{\mathcal{X}/B})$, and by [39], Theorem 5, we have $(P, \omega_{\mathcal{X}/B}) \geq 0$, hence $(P,P) \leq 0$.

Let us now derive an upper bound for $(P, \omega_{\mathcal{X}/B})$. As the automorphism group of $\mathcal{X}$ over B acts transitively on the cusps, it suffices to do this for the standard cusp ∞. There, $1/j$ is a parameter over all of B, hence it gives us a generator $d(1/j)$ of $\infty^*\omega_{\mathcal{X}/B}$. As $1/j$ is in $\mathbb{Z}[[q]]$ we have $d(1/j) = dq$ in $\infty^*\omega_{\mathcal{X}/B}$. By definition of the Arakelov intersection product and Theorem 18.2.1 we then have:

$$(\infty, \omega_{\mathcal{X}/B}) = -[\mathbb{Q}(\zeta_{pl})) : \mathbb{Q}] \log \|dq\|_{\mathrm{Ar}}(\infty) = O((pl)^7).$$

We now know $|(P,P)| = O((pl)^7)$ for all cuspidal P in $\mathcal{X}(B)$. We will now show that $|(P,Q)| = O((pl)^7)$. By the Theorem of Manin-Drinfeld, see [31], the image of the divisor $P - Q$ in $J_1(pl)(\mathbb{Q}(\zeta_{pl}))$ is of finite order. Let Φ be a vertical fractional divisor such that for

any irreducible component C of a fibre of $\mathcal{X}$ over B we have $(P - Q - \Phi, C) = 0$. By [54] or Theorem 4 of [39] we have $(P - Q - \Phi, P - Q - \Phi) = 0$. Equivalently, we have:

$$2(P, Q) = (P, P) + (Q, Q) - (P - Q, \Phi).$$

The term $(P - Q, \Phi)$ can be dealt with by a method analogous to the one used in Lemma 15.2.6, but now one can use that the geometric fibres of $\mathcal{X}$ at p and at l consist of two irreducible components, smooth and meeting transversally in at least one point. A short computation gives that $|(P - Q, \Phi)| = O(pl)$. The estimate $|(P, Q)| = O((pl)^7)$ now follows.

To get to the second statement of the theorem, note that

$$(D, D - \omega_{\mathcal{X}/B}) = (D, D + \omega_{\mathcal{X}/B}) - 2(D, \omega_{\mathcal{X}/B}) = \sum_{k \neq l} (P_k, P_l) - 2(D, \omega_{\mathcal{X}/B}),$$

where we have written $D = P_1 + \cdots + P_g$, with repetitions allowed. By our previous estimates, we get $|(D, D - \omega_{\mathcal{X}/B})| = O(g^2(pl)^7) = O((pl)^{11})$. □

We will also need a lower bound for the intersection number of two distinct points on $X_1(pl)$.

18.3.2 Theorem *There is an integer c such that for all pairs of distinct primes p and l such that $X_1(pl)$ has genus at least one, for any extension K of $\mathbb{Q}(\zeta_{5l})$ and for P and Q distinct points in $X_1(pl)(K)$ we have:*

$$\frac{1}{[K : \mathbb{Q}]}(P, Q) \geq c(pl)^6,$$

where (P, Q) is the intersection number of P and Q on the minimal regular model of $X_1(pl)$ over O_K.

Proof We have:

$$(P, Q) = (P, Q)_{\text{fin}} + (P, Q)_\infty,$$

with $(P, Q)_{\text{fin}}$ the contribution from the finite places of K, and $(P, Q)_\infty$ the contribution from the infinite places. As $P \neq Q$, we have $(P, Q)_{\text{fin}} \geq 0$. On the other hand, we have:

$$(P, Q)_\infty = \sum_\sigma -g_\sigma(P_\sigma, Q_\sigma).$$

By Theorem 18.2.1 we have:

$$g_\sigma(P_\sigma, Q_\sigma) \leq c \cdot (pl)^6$$

for some absolute constant c. This finishes the proof. □

19 Final estimates of the Arakelov contribution

We will now put the estimates of the preceding sections together, in the situation of Section 13. We briefly recall this situation. We have a prime number $l > 5$, and X_l denotes the modular curve $X_1(5l)$, over $\mathbb{Q}$, and g_l its genus. The Jacobian variety of X_l is denoted by J_l. In $J_l(\overline{\mathbb{Q}})[l]$ we have the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-module V_l that gives the representation associated to Δ. We have an effective divisor D on $X_{l,\mathbb{Q}(\zeta_l)}$, of degree g_l, supported on the cusps. For every non-zero x in V_l there is a unique effective divisor $D'_x = Q_{x,1} + \cdots + Q_{x,g_l}$ of degree g_l such that x is the class of $D'_x - D$. We have written $D'_x = D''_x + (D'_x)^{\mathrm{rest}}$, where $(D'_x)^{\mathrm{rest}}$ is the part of D'_x supported on the cusps. The numbering of the $Q_{x,i}$ is such that $D''_x = Q_{x,1} + \cdots + Q_{x,d_l}$. We have morphisms b_l, x_l and y_l from $X_{l,\mathbb{Q}}$ to $\mathbb{P}^1_{\mathbb{Q}}$ that, seen as rational functions, have their poles contained in the set of cusps of X_l.

19.1 Proposition *There is an integer c, such that for all $l > 5$ and for all x in $V_l - \{0\}$, the absolute heights of all $b_l(Q_{x,i})$, $x_l(Q_{x,i})$ and $y_l(Q_{x,i})$ are bounded from above by $c \cdot l^{12}$.*

Proof If $Q_{x,i}$ is a cusp, then this follows from Theorems 15.1.1 and 18.3.1. For the general case, we apply Theorem 15.2.5, where (in the notation of that Theorem) we assume that all $Q_{x,i}$ are K-rational and that K contains $\mathbb{Q}(\zeta_{5l})$, where $\mathcal{X}$ is the minimal regular model of X_l over $B := \mathrm{Spec}\, O_K$ and where P is a cusp. That theorem gives:

$$\begin{aligned}\frac{1}{[K:\mathbb{Q}]}(D'_x, P) \leq \frac{1}{[K:\mathbb{Q}]}\Bigg(&-\frac{1}{2}(D, D - \omega_{\mathcal{X}/B}) + 2g_l^2 \sum_{s\in B} \delta_s \log \#k(s) \\ &+ \sum_{\sigma} \log \|\vartheta\|_{\sigma,\sup} + \frac{g_l}{2}[K:\mathbb{Q}]\log(2\pi) \\ &+ \frac{1}{2}\deg\det p_*\omega_{\mathcal{X}/B} + (D, P)\Bigg)\ .\end{aligned}$$

Theorem 18.3.1, applied with $B = \mathrm{Spec}(\mathbb{Z}[\zeta_{5l}])$, gives that:

$$\frac{1}{[K:\mathbb{Q}]}|(D, D - \omega_{\mathcal{X}/B})| = O(l^{10}),$$

and that:

$$\frac{1}{[K:\mathbb{Q}]}|(D, P)| = O(l^8),$$

as D is an effective cuspidal divisor of degree g_l and $g_l = O(l^2)$. By Theorem 17.1 we have:

$$\frac{1}{[K:\mathbb{Q}]}\sum_{\sigma} \log \|\vartheta\|_{\sigma,\sup} = O(l^6).$$

By Theorem 16.7 we have:

$$\frac{1}{[K:\mathbb{Q}]} \deg \det p_* \omega_{\mathcal{X}/B} = O(l^2 \log l).$$

Finally, we have:

$$\frac{g_l^2}{[K:\mathbb{Q}]} \sum_{s \in B} \delta_s \log \#k(s) = O(l^6),$$

by the following argument. The only non-trivial contributions come from s over 5 and over l. The total contribution at 5 is independent of which extension K of $\mathbb{Q}(\zeta_5)$ we use, and for $\mathbb{Q}(\zeta_5)$ there is only one s over 5, $k(s) = \mathbb{F}_5$, and δ_s equals the number of supersingular points in $X_1(l)(\overline{\mathbb{F}}_5)$, which is $O(l^2)$. The contribution from l can be computed over $\mathbb{Q}(\zeta_l)$. Then there is one s, and $k(s) = \mathbb{F}_l$, and δ_s is the number of supersingular points in $X_1(5)(\overline{\mathbb{F}}_l)$, which is $O(l)$.

Putting these last estimates together, we get that there is an integer c such that for all l and x we have:

$$\frac{1}{[K:\mathbb{Q}]}(D'_x, P) \leq c \cdot l^{10}. \tag{19.1.1}$$

Now D'_x is the sum of g_l points $Q_{x,i}$. In order to get upper bounds for the individual $(Q_{x,i}, P)$ we need a lower bound for these. Theorem 18.3.2 gives us a lower bound if $Q_{x,i} \neq P$ and Theorem 18.3.1 gives us one if $P = Q_{x,i}$. Putting these together, we get a universal constant c such that:

$$\frac{1}{[K:\mathbb{Q}]}(Q_{x,i}, P) \geq c l^6.$$

The last two estimates together imply that there is an integer c such that for all l, x and i we have:

$$\frac{1}{[K:\mathbb{Q}]}(Q_{x,i}, P) \leq c \cdot l^{10}.$$

We can now apply Theorem 15.1.1. We note that $f^*\infty$ is an effective cuspidal divisor on X_l, of degree $O(l^2)$. This finishes the proof. □

Before we continue our serious work, we state and prove a simple lemma.

19.2 Lemma *Let $d \geq 1$ and $n \geq d$ be integers. Let Σ_d denote the elementary symmetric polynomial of degree d in n variables. Let $y_1, \ldots, y_n$ be in $\overline{\mathbb{Q}}$. Then we have:*

$$h\left(\Sigma_d(y_1, \ldots, y_n)\right) \leq n \log 2 + n \sum_{1 \leq i \leq n} h(y_i).$$

Proof Let K be the compositum of the fields $\mathbb{Q}(y_i)$ for $i = 1, \ldots, n$. For each place v of K, we let $|\cdot|_v$ be the natural absolute value on K_v and on K as at the end of Section 14.1. By the triangle inequality we obtain:

$$|\Sigma_d(y_1, \ldots, y_n)|_v \leq c(v,d,n) \max_{1 \leq i_1 < \cdots < i_d \leq n} |y_{i_1} \cdots y_{i_d}|_v \leq c(v,d,n) \max_{1 \leq i \leq n} |y_i|_v^d$$

for each place v of K, where $c(v,d,n) = \binom{n}{d} \leq 2^n$ if v is Archimedean, and $c(v,d,n) = 1$ if v is non-Archimedean. It follows that:

$$\max\{1, |\Sigma_d(y_1, \ldots, y_n)|_v\} \leq c(v,n) \prod_{i=1}^{n} \max\{1, |y_i|_v\}^n$$

with $c(v,n) = 2^n$ if v is Archimedean, and $c(v,n) = 1$ if v is non-Archimedean. The lemma follows by taking logarithms, summing over the places v, and dividing by $[K : \mathbb{Q}]$. □

19.3 Theorem *There are integers c_1, d_1, c_2, d_2 and c_3 and d_3 with the following property. Let $l > 5$ be prime, let $f = \alpha b_l + \beta x_l + \gamma y_l$ be a function as in (13.9). Let the a_d in $\mathbb{Z}[\zeta_l]$ be as in (13.10) and $P_l = \sum_m P_{l,m} T^m$ in $\mathbb{Q}(\zeta_l)[T]$ be as in (13.11). Then for all m we have:*

$$h(P_{l,m}) \leq c_1 l^{24} + c_2 l^6 \sum_{1 \leq d \leq d_l} h(a_d) + c_3 l^{12} (h(\alpha) + h(\beta) + h(\gamma)).$$

Proof Let $x_1, \ldots, x_{l^2-1}$ run through $V_l - \{0\}$. We have $P_{l,m} = \pm \Sigma_m(k(x_1), \ldots, k(x_{l^2-1}))$ and thus, by Lemma 19.2:

$$h(P_{l,m}) \leq (l^2 - 1) \log 2 + (l^2 - 1) \sum_{1 \leq j \leq l^2 - 1} h(k(x_j)) \,.$$

Note that the $k(x_j)$ are all conjugates of each other; this means that they all have the same height. It suffices therefore to continue with a single x in $V_l - \{0\}$. We have $k(x) = \sum_{d=1}^{d_l} a_d k_{D,f,d}(x)$ and this gives, using Lemma 19.2 and the rule $h(yz) \leq h(y) + h(z)$, that:

$$h(k(x)) \leq d_l \log 2 + d_l \sum_{1 \leq d \leq d_l} \left(h(a_d) + h(k_{D,f,d}(x)) \right) \,.$$

Next we have $k_{D,f,d}(x) = \Sigma_d(f(Q_{x,1}), \ldots, f(Q_{x,d_l}))$ and hence, once again by Lemma 19.2,

$$h(k_{D,f,d}(x)) \leq d_l \log 2 + d_l \sum_{1 \leq i \leq d_l} h(f(Q_{x,i})) \,.$$

By definition, $f = \alpha b_l + \beta x_l + \gamma y_l$, and hence

$$h(f(Q_{x,i})) \leq 3 \log 2 + h(\alpha) + h(b_l(Q_{x,i})) + h(\beta) + h(x_l(Q_{x,i})) + h(\gamma) + h(y_l(Q_{x,i})) \,,$$

for all i. Combining everything we get:

$$\begin{aligned} h(P_{l,m}) \leq & \left((l^2-1)+(l^2-1)^2 d_l + (l^2-1)^2 d_l^3 + 3(l^2-1)^2 d_l^4\right)\log 2 \\ & + d_l(l^2-1)^2 \sum_{d=1}^{d_l} h(a_d) + (l^2-1)^2 d_l^4 \left(h(\alpha)+h(\beta)+h(\gamma)\right) \\ & + (l^2-1)^2 d_l^4 \left(h(b_l(Q_{x,i})) + h(x_l(Q_{x,i})) + h(y_l(Q_{x,i}))\right) , \end{aligned}$$

for all m. Noting that $d_l \leq g_l \leq c_4 \cdot l^2$ for some universal constant c_4 and that $h(b_l(Q_{x,i}))+h(x_l(Q_{x,i}))+h(y_l(Q_{x,i})) \leq c_5 l^{12}$ for some universal constant c_5 by Proposition 19.1, we finally get:

$$h(P_{l,m}) \leq c_1 l^{24} + c_2 l^6 \sum_{d=1}^{d_l} h(a_d) + c_3 l^{12}(h(\alpha)+h(\beta)+h(\gamma))$$

for all m, as we wanted. □

A last consequence of all Arakelovian estimates is the following upper bound for the term $\log \#\mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x)$ in Theorem 15.2.5.

19.4 Theorem *There is an integer c such that for all $l > 5$ prime and all x in $V_l - \{0\}$ we have:*

$$\frac{1}{[K:\mathbb{Q}]} \log \#\mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x) \leq c \cdot l^{10}.$$

Proof We have already seen, in the proof of Proposition 19.1, that the right hand side of the inequality in Theorem 15.2.5, divided by $[K:\mathbb{Q}]$, is bounded from above by a constant times l^{10}. We have also seen that the term $(D'_x, P)/[K:\mathbb{Q}]$ on the left hand side is bounded from below by a constant times l^8 (recall that D'_x is of degree $O(l^2)$). This proves the inequality. □

This upper bound will be very useful for us, as the next interpretation shows.

19.5 Theorem *There is an integer c with the following property. Let $l > 5$ be a prime number, and let D and the D'_x be as before. We recall that X_l has good reduction outside $5l$. A prime number $p \not| 5l$ is said to be l-*good *if for all x in $V_l - \{0\}$ the following two conditions are satisfied:*

1. *at all places v of $\overline{\mathbb{Q}}$ over p the specialisation $(D'_x)_{\overline{\mathbb{F}}_p}$ at v is the unique effective divisor on the reduction $X_{l,\overline{\mathbb{F}}_p}$ such that the difference with $D_{\overline{\mathbb{F}}_p}$ represents the specialisation of x;*
2. *the specialisations of the non-cuspidal part D''_x of D'_x at all v above p are disjoint from the cusps.*

Then we have:

$$\sum_{p \text{ not } l\text{-good}} \log p \leq cl^{14}.$$

Proof First of all, a prime number p satisfies conditions (1) and (2) for all x if and only if it satisfies them for one x, as $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$ acts transitively on the set of x.

We take K to be the extension of $\mathbb{Q}(\zeta_l)$ that corresponds to the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$-set $V_l - \{0\}$; this is the field of definition of one x in $V_l - \{0\}$. Then $[K : \mathbb{Q}] = O(l^3)$. We define $S(l)$ to be the image in $\mathrm{Spec}(\mathbb{Z}[1/5l])$ of the support in $\mathrm{Spec}(O_K)$ of the finite O_K-module $\mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x)$. By Theorem 19.4 we have:

$$\log \#(\mathbb{Z}[1/5l] \otimes \mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x)) = O(l^{13}),$$

and hence:

$$\sum_{p \in S(l)} \log p = O(l^{13}).$$

We claim that the p in $S(l)$ are precisely the primes $p \notin \{5, l\}$ such that condition (1) is not satisfied for x. To see this, we first note that for a residue field $O_K \to \mathbb{F}$ at p we have $\mathrm{H}^1(X_{l,\mathbb{F}}, O_{\mathcal{X}_{\mathbb{F}}}((D'_x)_{\mathbb{F}})) = \mathbb{F} \otimes_{O_K} \mathrm{R}^1 p_* O_{\mathcal{X}}(D'_x)$ (see Theorem III.12.1 of [53]; cohomology and base change in top dimension). The divisor $(D'_x)_{\mathbb{F}}$ is the unique effective divisor in its linear equivalence class if and only if $h^0(X_{l,\mathbb{F}}, O_{\mathcal{X}_{\mathbb{F}}}((D'_x)_{\mathbb{F}})) = 1$, which, by Riemann-Roch, is equivalent to $h^1(X_{l,\mathbb{F}}, O_{\mathcal{X}_{\mathbb{F}}}((D'_x)_{\mathbb{F}})) = 0$.

Now we let $T(l)$ denote the set of primes $p \notin \{5, l\}$ such that at least one specialisation of D''_x at a place of K above p is not disjoint from the cusps. Taking into account that $[K : \mathbb{Q}] = O(l^3)$, equation (19.1.1) gives us an upper bound:

$$(D'_x, P) = O(l^{13})$$

As the degree of $(D'_x)^{\mathrm{rest}}$ is at most $O(l^2)$, Theorem 18.3.1 gives us:

$$|((D'_x)^{\mathrm{rest}}, P)| = O(l^9), \quad \text{hence} \quad (D''_x, P) = O(l^{13}).$$

As the divisor Cusps has degree $O(l)$, we have:

$$(D''_x, \mathrm{Cusps}) = O(l^{14}).$$

The intersection number (D''_x, Cusps) is the sum of the contribution $(D''_x, \mathrm{Cusps})_{\mathrm{fin}}$ of the finite places and the contribution $(D''_x, \mathrm{Cusps})_\infty$ of the infinite places. We have:

$$(D''_x, \mathrm{Cusps})_\infty = \sum_{i,P,\sigma} -g_\sigma(Q_{x,i}, P),$$

where the sum is taken over the i with $1 \leq i \leq d_l$, over the cusps P and the $\sigma\colon K \to \mathbb{C}$. Then Theorem 18.2.1 gives the upper bound:

$$(D''_x, \mathrm{Cusps})_{\mathrm{fin}} = (D''_x, \mathrm{Cusps}) - (D''_x, \mathrm{Cusps})_{\infty} = O(l^{14}).$$

By the definition of our set $T(l)$ and the definition of $(D''_x, \mathrm{Cusps})_{\mathrm{fin}}$ we get:

$$\sum_{p \in T(l)} \log p \leq (D''_x, \mathrm{Cusps})_{\mathrm{fin}} = O(l^{14}).$$

The proof of the theorem is then finished by noticing that the set of primes $p \notin \{5, l\}$ that are not l-good is precisely the union of $S(l)$ and $T(l)$. □

20 Ultra short description of Couveignes's finite field approach

The results of the previous section provide the necessary input for the computational part of our whole program. More precisely, Theorem 19.3 gives an upper bound for the height of the coefficients of our polynomials P_l in $\mathbb{Q}(\zeta_l)[T]$. This upper bound is polynomial in l if we can show that we can take the coefficients α, β and γ small enough; in Section 22 it will be explained that we can do that. Theorem 19.5 shows the existence of sufficiently many small primes p that are l-good, which implies that the coefficients of P_l are integral at all places of $\mathbb{Q}(\zeta_l)$ over p and moreover that the reduction of P_l at these places can be computed over the corresponding finite fields.

In the next section (Section 21), Couveignes shows that there is a probabilistic algorithm that computes for p a prime that is l-good, the reductions $(D'_x)_{\mathbb{F}}$ of the divisors D'_x on $X_{l,\mathbb{F}}$, where $\mathbb{F}$ is a suitable extension of any residue field of $\mathbb{Z}[\zeta_l]$ at p, with an expected running time that is polynomial in l and p. A few remarks should be made, at this point.

20.1 Remark Couveignes starts by showing that the standard operations on divisors on $X_{l,\mathbb{F}}$, such as computing $\mathrm{H}^0(X_{l,\mathbb{F}}, O(D_1 - D_2))$, as well as the standard operations on $J_{l,\mathbb{F}}(\mathbb{F})$ such as addition and applying Hecke operators T_n, can be computed in probabilistic polynomial time in l, $\log \#\mathbb{F}$, $\deg D_i$, and n.

20.2 Remark The reason for which the expected running time of Couveignes algorithm is not polynomial in l and $\log p$ is simply that he needs to compute the numerator of the zeta function of $X_{l,\mathbb{F}}$. Modular symbols methods are used to compute the characteristic polynomial of the

Hecke operator T_p on the Jacobian J_l of X_l; the Eichler-Shimura relation then gives the characteristic polynomial of the Frobenius endomorphism of $J_{l,\mathbb{F}_p}$. It is only the modular symbols part of Couveignes algorithm of which the running time is not polynomial in $\log p$.

We note that general p-adic algorithms as described in Section 6 could have been used for computing the zeta-function of $X_{l,\mathbb{F}}$. These p-adic methods are not restricted to modular curves.

20.3 Remark Readers who are surprised by the probabilistic aspect of Couveignes algorithm should realise that already no deterministic polynomial time algorithm is presently known for factoring polynomials in one variable over finite fields, or, for that matter, for finding an element of $\mathbb{F}_p$ that is not a square. We stress that the notion of probabilistic algorithm that we use here means that the output is correct, but that the algorithm must make some random choices during its execution. The two examples just given are typical cases where a probabilistic polynomial time algorithm is known, but no deterministic polynomial time algorithm. Technically speaking, our notion of probabilistic algorithm is that of *Las Vegas algorithm*, as opposed to the notion of *Monte Carlo algorithm*. Monte Carlo algorithms are allowed to give incorrect output, which is certainly not what we want. More details can be found by searching "randomised algorithm" in the Wikipedia. A good reference for foundational material concerning algorithms in number theory is Lenstra's article [74].

21 Linearising torsion classes in the Picard group of algebraic curves over finite fields, by Jean-Marc Couveignes

This section has been written by Jean-Marc Couveignes. It addresses the problem of computing in the group of ℓ^k-torsion rational points of the Jacobian variety of algebraic curves over finite fields, with a view toward computing modular representations.

21.1 Introduction

Let $\mathbb{F}_q$ be a finite field of characteristic p and $\mathbb{A}^2 \subset \mathbb{P}^2$ the affine and projective planes over $\mathbb{F}_q$ and $C \subset \mathbb{P}^2$ a plane projective absolutely irreducible reduced curve and $\mathcal{X}$ its smooth projective model and $\mathcal{J}$ the Jacobian variety of $\mathcal{X}$. Let g be the genus of $\mathcal{X}$ and d the degree of C.

We assume we are given the numerator of the zeta function of the function field $\mathbb{F}_q(\mathcal{X})$. So we know the characteristic polynomial of the Frobenius endomorphism F_q of $\mathcal{J}$. This is a unitary degree $2g$ polynomial $\chi(X)$ with integer coefficients.

Let $\ell \neq p$ be a prime integer and let $n = \ell^k$ be a power of ℓ. We look for a *nice generating set* for the group $\mathcal{J}[\ell^k](\mathbb{F}_q)$ of ℓ^k-torsion points in $\mathcal{J}(\mathbb{F}_q)$. By *nice* we mean that the generating set

$(g_i)_{1\leq i\leq I}$ should induce a decomposition of $\mathcal{J}[\ell^k](\mathbb{F}_q)$ as a direct product $\prod_{1\leq i\leq I}\langle g_i\rangle$ of cyclic subgroups with non-decreasing orders.

Given such a generating set and an $\mathbb{F}_q$-endomorphism of $\mathcal{J}$, we also want to describe the action of this endomorphism on $\mathcal{J}[\ell^k](\mathbb{F}_q)$ by an $I \times I$ integer matrix.

In Section 21.2 we recall how to compute in the Picard group $\mathcal{J}(\mathbb{F}_q)$. Section 21.3 gives a naive algorithm for picking random elements in this group. Pairings are useful when looking for relations between divisor classes. So we recall how to compute pairings in Section 21.4. Section 21.5 is concerned with characteristic subspaces for the action of Frobenius inside the ℓ^∞-torsion of $\mathcal{J}(\overline{\mathbb{F}}_q)$. In Section 21.6 we look for a convenient surjection from $\mathcal{J}(\mathbb{F}_q)$ onto its ℓ^k-torsion subgroup. We use the Kummer exact sequence and the structure of the ring generated by the Frobenius endomorphism. In Section 21.7 we give an algorithm that, on input a degree d plane projective curve over $\mathbb{F}_q$, plus some information on its singularities, and the zeta function of its function field, returns a nice generating set for the group of ℓ^k-torsion points inside $\mathcal{J}(\mathbb{F}_q)$ in probabilistic polynomial time in $\log q$, d and ℓ^k. Sections 21.8 and 21.9 are devoted to two families of modular curves. We give a nice plane model for such curves. The general algorithms presented in Section 21.7 are then applied to these modular curves in Section 21.10 in order to compute explicitly the modular representations associated with the discriminant modular form (level 1 and weight 12). This makes a connection with Edixhoven's program for computing coefficients of modular forms.

21.1.1 Remark The symbol $\mathcal{O}$ in this section (Section 21) stands, when used for "big O" notation, for a positive effective absolute constant. So any statement containing this symbol becomes true if the symbol is replaced in every occurrence by some well chosen positive real number, that may be every time different.

21.1.2 Remark By an algorithm in this paper we usually mean a probabilistic (Las Vegas) algorithm. This is an algorithm that succeeds with probability $\geq 1/2$. When it fails, it gives no answer. In some places we shall give deterministic algorithms or probabilistic (Monte-Carlo) algorithms, but this will be stated explicitly. A Monte-Carlo algorithm gives a correct answer with probability $\geq 1/2$. But it may give an incorrect answer with probability $\leq 1/2$. A Monte-Carlo algorithm can be turned into a Las Vegas one, provided we can efficiently check the correctness of the result. The reason for using probabilistic Turing machines is that in many places it will be necessary (or at least wiser) to decompose a divisor as a sum of places. This is the case in particular for the conductor of some plane curve. Another more intrinsically probabilistic algorithm in this paper is the one that searches for generators of the Picard group.

21.2 Basic algorithms for plane curves

We recall elementary results about computing in the Picard group of an algebraic curve over a finite field. See [52, 109].

21.2.1 Finite fields

We should first explain how finite fields are represented. The base field $\mathbb{F}_q$ is given by an irreducible polynomial $f(X)$ with degree a and coefficients in $\mathbb{F}_p$ where p is the characteristic and $q = p^a$. So $\mathbb{F}_q$ is $\mathbb{F}_p[X]/f(X)$. An extension of $\mathbb{F}_q$ is given similarly by an irreducible polynomial in $\mathbb{F}_q[X]$. Polynomial factoring in $\mathbb{F}_q[X]$ is probabilistic polynomial time in $\log q$ and the degree of the polynomial to be factored.

21.2.2 Plane projective curves and their smooth model

We now explain how curves are supposed to be represented in this paper.

To start with, a projective plane curve $C/\mathbb{F}_q$ is given by a degree d homogeneous polynomial $E(X, Y, Z)$ in the three variables X, Y and Z, with coefficients in $\mathbb{F}_q$.

The smooth model $\mathcal{X}$ of C is not given as a projective variety. Indeed, we shall only need a nice local description of $\mathcal{X}$ at every singularity of C. This means we need a list (a labelling) of all places above every singularity of C and a uniformising parameter at every place. We also need the Laurent series expansions of affine plane coordinates in terms of all these uniformising parameters.

More precisely, let P be a place above a singular point S and v the corresponding valuation. We say that P is a singular branch. The field of definition of P is an extension field $\mathbb{F}_P$ of $\mathbb{F}_q$ with degree $\leq (d-1)(d-2)/2$. Let x and y be affine coordinates that vanish at the singular point S on C. We need a local parameter t at P and expansions $x = \sum_{k \geq v(x)} a_k t^k$ and $y = \sum_{k \geq v(y)} b_k t^k$ with coefficients in $\mathbb{F}_P$.

Because these expansions are not finite, we just assume we are given an oracle that on input a positive integer n returns the first n terms in all these expansions.

This is what we mean when we say the smooth model $\mathcal{X}$ is given.

We may also assume that we are given the conductor $\mathfrak{C}$ of C as a combination of singular branches with integer coefficients. The degree $\deg(\mathfrak{C})$ is even and we have $\deg(\mathfrak{C}) = 2\delta$ where δ is the difference between the arithmetic genus $(d-1)(d-2)/2$ and the geometric genus g. In fact it suffices to have a divisor $\mathfrak{D}$ that is greater than the conductor and has polynomial degree in d. Such a divisor can be computed from the equation of C in deterministic polynomial time in $\log q$ and d.

There are many families of curves for which such a smooth model can be given as a Turing machine that answers in probabilistic polynomial time in the size $\log q$ of the field and the degree d of C and the number n of requested significant terms in the parameterisations of singular branches. This is the case for curves with ordinary multiple points for example. We shall show in Sections 21.8 and 21.9 that this is also the case for two nice families of modular curves.

21.2.3 Points, forms, and functions

Smooth points on C can be represented by their affine or projective coordinates. Labelling for the branches above singular points is given in the description of $\mathcal{X}$. So we know how to represent divisors on $\mathcal{X}$.

For any integer $h \geq 0$ and extension field K of $\mathbb{F}_q$, we set:

$$\mathcal{S}_h/K = \mathrm{H}^0(\mathbb{P}^2/K, \mathcal{O}_{\mathbb{P}^2/K}(h))$$

the K-linear space of degree h homogeneous polynomials in X, Y, and Z. It is a vector space of dimension $(h+1)(h+2)/2$ over K. A basis for it is made of all monomials of the form $X^aY^bZ^c$ with $a, b, c \in \mathbb{N}$ and $a + b + c = h$.

We denote by:

$$\mathcal{H}^h/K = \mathrm{H}^0(\mathcal{X}/K, \mathcal{O}_{\mathcal{X}/K}(h))$$

the space of forms of degree h on $\mathcal{X}/K$. Here $\mathcal{O}_{\mathcal{X}/K}(h)$ is the pullback of $\mathcal{O}_{\mathbb{P}^2/K}(h)$ to $\mathcal{X}$. Let W be a form on $\mathbb{P}^2$ having non-zero pull back $W_{\mathcal{X}}$ on $\mathcal{X}$. Let $H = (W_{\mathcal{X}})$ be the divisor of this restriction. The map $f \mapsto f/W_{\mathcal{X}}$ is a bijection from $\mathrm{H}^0(\mathcal{X}/K, \mathcal{O}_{\mathcal{X}/K}(h))$ to the linear space $\mathcal{L}(H)$. If Δ is a divisor on $\mathcal{X}/K$ we note $\mathcal{H}^h(\Delta)/K$ the subspace of forms in $\mathcal{H}^h/K$ with divisor $\geq \Delta$. The dimension of $\mathcal{H}^h(\mathfrak{C})$ is at least $dh + 1 - g - \deg(\mathfrak{C})$ and is equal to this number when it exceeds $g - 1$. This is the case if $h \geq d$. The dimension of $\mathcal{H}^h(\mathfrak{C})$ is greater than $2g$ if $h \geq 2d$.

The image of the restriction map $\rho\colon \mathcal{S}_h \to \mathcal{H}^h$ contains $\mathcal{H}^h(\mathfrak{C})$ according to Noether's residue theorem [49, Theorem 7].

We set $h_C = 2d$ and $S_C = \mathcal{S}_{h_C}$ and $\mathcal{H}_C = \mathcal{H}^{h_C}(\mathfrak{C})$, and $H_C = \rho^{-1}(\mathcal{H}_C) \subset S_C$ and $K_C = \ker(\rho) \subset H_C$.

So we have $0 \to K_C \to H_C \to \mathcal{H}_C \to 0$.

In order to find linear equations for $H_C \subset S_C$ we consider a generic homogeneous form $F(X, Y, Z) = \sum_{a+b+c=h_C} \epsilon_{a,b,c} X^aY^bZ^c$ of degree h_C in X, Y and Z.

For every branch P above a singular point S with homogeneous coordinates $[X, Y, Z]$ (assuming for example that S has non-zero Z-coordinate) we replace in $F(X/Z, Y/Z, 1)$ the affine

coordinates $x = X/Z$ and $y = Y/Z$ by their expansions as series in the local parameter t at this branch.

We ask the resulting series in t to have valuation at least the multiplicity of P in the conductor $\mathfrak{C}$.

Every singular branch thus produces linear equations in the $\epsilon_{a,b,c}$. The collection of all such equations defines the subspace H_C.

A basis for the subspace $K_C \subset H_C \subset S_C$ consists of all $X^a Y^b Z^c E(X,Y,Z)$ with $a+b+c = h_C - d$. We fix a supplementary space M_C to K_C in H_C and assimilate $\mathcal{H}_C$ to it.

Given a homogeneous form in three variables one can compute its divisor using resultants and the given expansions of affine coordinates in terms of the local parameters at every singular branch.

A function is given as a quotient of two forms.

21.2.4 The Brill-Noether algorithm

Linear spaces of forms computed in the previous paragraph allow us to compute in the group $\mathcal{J}(\mathbb{F}_q)$ of $\mathbb{F}_q$-points in the Jacobian of $\mathcal{X}$. We fix an effective $\mathbb{F}_q$-divisor ω with degree g on $\mathcal{X}$. A point $\alpha \in \mathcal{J}(\mathbb{F}_q)$ is represented by a divisor $A - \omega$ in the corresponding linear equivalence class, where A is an effective $\mathbb{F}_q$-divisor with degree g. Given another point $\beta \in \mathcal{J}(\mathbb{F}_q)$ by a similar divisor $B - \omega$, we can compute the space $\mathcal{H}^{h_C}(\mathfrak{C} + A + B)$ which is non-trivial and pick a non-zero form f_1 in it. The divisor of f_1 is $(f_1) = A + B + \mathfrak{C} + R$ where R is an effective divisor with degree $(2g-2)h_C - 2g - 2\delta$. The linear space $\mathcal{H}^{h_C}(\mathfrak{C} + R + \omega)$ has dimension at least 1. We pick a non-zero form f_2 in it. It has divisor $(f_2) = \mathfrak{C} + R + \omega + D$ where D is effective with degree g. And $D - \omega$ is linearly equivalent to $A - \omega + B - \omega$.

In order to invert the class α of $A - \omega$ we pick a non-zero form f_1 in $\mathcal{H}^{h_C}(\mathfrak{C} + 2\omega)$. The divisor of f_1 is $(f_1) = 2\omega + \mathfrak{C} + R$ where R is an effective divisor with degree $(2g-2)h_C - 2g - 2\delta$. The linear space $\mathcal{H}^{h_C}(\mathfrak{C} + R + A)$ has dimension at least 1. We pick a non-zero form f_2 in it. It has divisor $(f_2) = \mathfrak{C} + R + A + B$ where B is effective with degree g. And $B - \omega$ is linearly equivalent to $-(A - \omega)$.

This algorithm works just as well if we replace $\mathfrak{C}$ by some $\mathfrak{D} \geq \mathfrak{C}$ having polynomial degree in d.

21.2.5 Lemma (Arithmetic operations in the jacobian) *Let $C/\mathbb{F}_q$ be a degree d plane projective absolutely irreducible reduced curve. Let g be the geometric genus of C. Assume we are given the smooth model $\mathcal{X}$ of C and a degree g origin $\mathbb{F}_q$-divisor ω on $\mathcal{X}$. Arithmetic opera-*

tions in the Picard group $\mathrm{Pic}^0(\mathcal{X}/\mathbb{F}_q)$ *can be performed in polynomial time in* $\log q$ *and* d*. This includes addition, subtraction and comparison of divisor classes.*

We now recall the principle of the Brill-Noether algorithm for computing complete linear series. Functions in $\mathbb{F}_q(\mathcal{X})$ are represented as quotients of forms.

21.2.6 Lemma (Brill-Noether) *There exists an algorithm that on input a degree* d *plane projective absolutely irreducible reduced curve* $C/\mathbb{F}_q$ *and the smooth model* $\mathcal{X}$ *of* C *and two effective* $\mathbb{F}_q$*-divisors* A *and* B *on* $\mathcal{X}$*, computes a basis for* $\mathcal{L}(A-B)$ *in time polynomial in* d *and* $\log q$ *and the degrees of* A *and* B*.*

Proof We assume $\deg(A) \geq \deg(B)$, otherwise $\mathcal{L}(A-B) = 0$. We let h be the smallest integer such that $h \geq h_C$ and $\dim(\mathcal{H}^h(\mathfrak{C}+A)) > 0$.

The space $\mathcal{H}^h(\mathfrak{C}+A)$ is non-zero and is contained in the image of the restriction map $\rho\colon \mathcal{S}_h \to \mathcal{H}^h$ so that we can represent it as a subspace of $\mathcal{S}_h$. We pick a non-zero form f in $\mathcal{H}^h(\mathfrak{C}+A)$ and compute its divisor $(f) = \mathfrak{C}+A+D$.

The space $\mathcal{H}^h(\mathfrak{C}+B+D)$ is contained in the image of the restriction map $\rho\colon \mathcal{S}_h \to \mathcal{H}^h$ so that we can represent it as a subspace of $\mathcal{S}_h$. We compute forms γ_1, γ_2, ..., γ_k in $\mathcal{S}_h$ such that their images by ρ provide a basis for $\mathcal{H}^h(\mathfrak{C}+B+D)$. A basis for $\mathcal{L}(A-B)$ is made of the functions γ_1/f, γ_2/f, ..., γ_k/f. Again this algorithm works just as well if we replace $\mathfrak{C}$ by some $\mathfrak{D} \geq \mathfrak{C}$ having polynomial degree in d. □

We deduce an explicit moving lemma for divisors.

21.2.7 Lemma (Moving divisor lemma I) *There exists an algorithm that on input a degree* d *plane projective absolutely irreducible reduced curve* $C/\mathbb{F}_q$ *and the smooth model* $\mathcal{X}$ *of* C *and a degree zero* $\mathbb{F}_q$*-divisor* $D = D^+ - D^-$ *and an effective divisor* A *with degree* $< q$ *on* $\mathcal{X}$ *computes a divisor* $E = E^+ - E^-$ *linearly equivalent to* D *and disjoint from* A *in time polynomial in* d *and* $\log q$ *and the degrees of* D^+*, and* A*. Further the degree of* E^+ *and* E^- *can be taken to be* $\leq 2gd$*.*

Proof Let O be an $\mathbb{F}_q$-rational divisor on $\mathcal{X}$ having degree $\leq d$ and disjoint from A. We may take O to be a well chosen fibre of some plane coordinate function on $\mathcal{X}$. We compute the linear space $\mathcal{L} = \mathcal{L}(D^+ - D^- + 2gO)$. The subset of functions f in $\mathcal{L}$ such that $(f) + D^+ - D^- + 2gO$ is not disjoint from A is contained in a union of at most $\deg(A) < q$ hyperplanes. We conclude invoking lemma 21.2.8 below. □

There remains to state and prove the following result.

21.2.8 Lemma (Solving inequalities) *Let q be a prime power, $d \geq 2$ and $n \geq 1$ two integers and let H_1, ..., H_n be hyperplanes inside $V = \mathbb{F}_q^d$, each given by a linear equation. Assume $n < q$. There exists a deterministic algorithm that finds a vector in $U = V - \bigcup_{1 \leq k \leq n} H_k$ in time polynomial in $\log q$, d and n.*

Proof This is proved by lowering the dimension d. For $d = 2$ we pick any affine line L in V not containing the origin. We observe that there are at least $q-n$ points in $U \cap L = L - \bigcup_{1 \leq k \leq n} L \cap H_k$. We enumerate points in L until we find one which is not in any H_k. This requires at most $n + 1$ trials.

Assume now d is greater than 2. Hyperplanes in V are parametrised by the projective space $\mathbb{P}(\hat{V})$ where $\hat{V}$ is the dual of V. We enumerate points in $\mathbb{P}(\hat{V})$ until we find a hyperplane K distinct from every H_k. We compute a basis for K and an equation for every $H_k \cap K$ in this basis. This way, we have lowered the dimension by 1. □

We can strengthen a bit the moving divisor algorithm by removing the condition that A has degree $< q$. Indeed, in case this condition is not met, we pick two distinct primes α and β such that $q^\alpha > \deg(A)$ and $q^\beta > \deg(A)$. We apply Lemma 21.2.7 after base change to the field with q^α element and find a divisor E_α. We call e_α the norm of E_α from $\mathbb{F}_{q^\alpha}$ to $\mathbb{F}_q$. It is equivalent to αD. We similarly construct a divisor e_β that is equivalent to βD. Let u and v be positive integers such that $\alpha u - \beta v = 1$. We return the divisor $E = ue_\alpha - ve_\beta = E^+ - E^-$. We observe that we can take $\alpha \leq 2 + 2\log_q \deg(A)$ and $\beta \leq 2\alpha$ and $u < \beta$ and $v < \alpha$ so the degree of E^+ is $\leq 12gd(\log_q(\deg(A)) + 1)$.

21.2.9 Lemma (Moving divisor lemma II) *There exists an algorithm that on input a degree d plane projective absolutely irreducible curve $C/\mathbb{F}_q$ and the smooth model $\mathcal{X}$ of C and a degree zero $\mathbb{F}_q$-divisor $D = D^+ - D^-$ and an effective divisor A on $\mathcal{X}$ computes a divisor $E = E^+ - E^-$ linearly equivalent to D and disjoint from A in time polynomial in d and $\log q$ and the degrees of D^+, and A. Further the degree of E^+ and E^- can be taken to be $\leq 12gd(\log_q(\deg(A)) + 1)$.*

21.3 A first approach to picking random divisors

Given a finite field $\mathbb{F}_q$ and a plane projective absolutely irreducible reduced curve C over $\mathbb{F}_q$ with projective smooth model $\mathcal{X}$, we call $\mathcal{J}$ the Jacobian of $\mathcal{X}$ and we consider the two related problems: picking a random element in $\mathcal{J}(\mathbb{F}_q)$ with (close to) uniform distribution and finding a generating set for (a large subgroup of) $\mathcal{J}(\mathbb{F}_q)$.

We will assume the size q of the field is greater than or equal to $4g^2$. This condition ensures the existence of a $\mathbb{F}_q$-rational point[3].

Picking efficiently and provably random elements in $\mathcal{J}(\mathbb{F}_q)$ with uniform distribution seems difficult to us. We first give here an algorithm for efficiently constructing random divisors with a distribution that is far from uniform but still sufficient to construct a generating set for a large subgroup of $\mathcal{J}(\mathbb{F}_q)$. Once given generators, picking random elements becomes much easier.

Let r be the smallest prime integer bigger than 30, $2g-2$ and d. We observe r is less than $\max(4g-4, 2d, 60)$.

The set $\mathcal{P}(r,q)$ of $\mathbb{F}_q$-places with degree r on $\mathcal{X}$ has cardinality:

$$\#\mathcal{P}(r,q) = \frac{\#\mathcal{X}(\mathbb{F}_{q^r}) - \#\mathcal{X}(\mathbb{F}_q)}{r}.$$

So:

$$(1-10^{-2})\frac{q^r}{r} \leq \#\mathcal{P}(r,q) \leq (1+10^{-2})\frac{q^r}{r}.$$

Indeed, $|\#\mathcal{X}(\mathbb{F}_{q^r}) - q^r - 1| \leq 2gq^{r/2}$ and $|\#\mathcal{X}(\mathbb{F}_q) - q - 1| \leq 2gq^{1/2}$. Therefore we have $|\#\mathcal{P}(r,q) - q^r/r| \leq ((4g+3)/r)q^{r/2} \leq 8q^{r/2}$ and $8rq^{-r/2} \leq r2^{3-r/2} \leq 10^{-2}$ since $r \geq 31$.

Since we are given a degree d plane model C for the curve $\mathcal{X}$, we have a degree d map $x\colon \mathcal{X} \to \mathbb{P}^1$. Since $d < r$, the function x maps $\mathcal{P}(r,q)$ to the set $\mathcal{U}(r,q)$ of unitary prime polynomials of degree r over $\mathbb{F}_q$. The cardinality of $\mathcal{U}(r,q)$ is $(q^r-q)/r$ so:

$$(1-10^{-9})\frac{q^r}{r} \leq \#\mathcal{U}(r,q) \leq \frac{q^r}{r}.$$

The fibres of the map $x\colon \mathcal{P}(r,q) \to \mathcal{U}(r,q)$ have cardinality between 0 and d.

We can pick a random element in $\mathcal{U}(r,q)$ with uniform distribution in the following way: we pick a random unitary polynomial of degree r with coefficients in $\mathbb{F}_q$, with uniform distribution. We check whether it is irreducible. If it is, we output it. Otherwise we start again. This is polynomial time in r and $\log q$.

Given a random element in $\mathcal{U}(r,q)$ with uniform distribution, we can compute the fibre of x above it and, provided it is non-empty, pick a random element in it with uniform distribution. If the fibre is empty, we pick another element in $\mathcal{U}(r,q)$ until we find a non-empty fibre. At least one over $d \times (0.99)^{-1}$ fibre is non-empty. We thus define a distribution μ on $\mathcal{P}(r,q)$ and prove the following result.

[3]If we are given a curve over $\mathbb{F}_q$ with genus g such that $q < 4g^2$ then we find ourselves in a much better situation. By the Riemann hypothesis the Picard group is generated by the classes of prime divisors of degree $\leq d$ where d is the first integer bigger than or equal to $2\log_q(4g-2)$. See [47]. The number of such divisors is then bounded by a polynomial in the genus. So the problems treated in that section become trivial in this special case.

21.3.1 Lemma (A very rough measure) *There exists a probabilistic algorithm that picks a random element in $\mathcal{P}(r,q)$ with distribution μ in time polynomial in d and $\log q$. For every subset Z of $\mathcal{P}(r,q)$ the measure $\mu(Z)$ is related to the uniform measure $(\#Z)/(\#\mathcal{P}(r,q))$ by:*

$$\frac{\#Z}{d\#\mathcal{P}(r,q)} \leq \mu(Z) \leq \frac{d\#Z}{\#\mathcal{P}(r,q)}.$$

Now let $\mathcal{D}(r,q)$ be the set of effective $\mathbb{F}_q$-divisors with degree r on $\mathcal{X}$.

Since we have assumed $q \geq 4g^2$ we know that $\mathcal{X}$ has at least one $\mathbb{F}_q$-rational point.

Let Ω be a degree r effective divisor on $\mathcal{X}/\mathbb{F}_q$. Now we associate to every α in $\mathcal{D}(r,q)$ the class of $\alpha - \Omega$ in $\mathcal{J}(\mathbb{F}_q)$. This defines a surjection $\zeta\colon \mathcal{D}(r,q) \to \mathcal{J}(\mathbb{F}_q)$ with all its fibres having cardinality $\#\mathbb{P}^{r-g}(\mathbb{F}_q)$.

So the set $\mathcal{D}(r,q)$ has cardinality $((q^{r-g+1}-1)/(q-1))\#\mathcal{J}(\mathbb{F}_q)$.

So:

$$\#\mathcal{P}(r,q) \leq \#\mathcal{D}(r,q) \leq q^{r-g}\frac{1-\frac{1}{q^{r-g+1}}}{1-\frac{1}{q}}q^g(1+\frac{1}{\sqrt{q}})^{2g}.$$

Since $q \geq 4g^2$ we have $\#\mathcal{D}(r,q) \leq 2eq^r$.

Assume G is a finite group and ψ an epimorphism of groups $\psi\colon \mathcal{J}(\mathbb{F}_q) \to G$. We look for some divisor $\Delta \in \mathcal{D}(r,q)$ such that $\psi(\zeta(\Delta)) \neq 0 \in G$. Since all the fibres of $\psi \circ \zeta$ have the same cardinality, the fibre above 0 has at most $2eq^r/\#G$ elements. So the number of prime divisors $\Delta \in \mathcal{P}(r,q)$ such that $\psi(\zeta(\Delta))$ is not 0 is at least $q^r(0.99/r - 2e/\#G)$. We assume $\#G$ is at least $12r$. Then at least half of the divisors in $\mathcal{P}(r,q)$ are not mapped onto 0 by $\psi \circ \zeta$. The μ-measure of the subset consisting of these elements is at least $1/2d$.

So if we pick a random Δ in $\mathcal{P}(r,q)$ with μ-measure as in Lemma 21.3.1, the probability of success is at least $1/2d$.

21.3.2 Lemma (Finding non-zero classes) *There exists a probabilistic (Monte-Carlo) algorithm that takes as input:*

1. *a degree d and geometric genus g plane projective absolutely irreducible reduced curve C over $\mathbb{F}_q$, such that $q \geq 4g^2$,*
2. *the smooth model $\mathcal{X}$ of C,*
3. *a degree g effective divisor ω, as origin,*
4. *an epimorphism $\psi\colon \mathrm{Pic}^0(\mathcal{X}/\mathbb{F}_q) \to G$ (that need not be computable) such that the cardinality of G is at least $\max(48g, 24d, 720)$,*

and outputs a sequence of $2d$ elements in $\mathrm{Pic}^0(\mathcal{X}/\mathbb{F}_q)$ such that at least one of them is not in the kernel of ψ with probability $\geq 1/2$. The algorithm is polynomial time in d and $\log q$.

As a special case we take $G = G_0 = \mathcal{J}(\mathbb{F}_q)$ and $\psi = \psi_0$ the identity. Applying Lemma 21.3.2 we find a sequence of elements in $\mathcal{J}(\mathbb{F}_q)$ out of which one at least is non-zero. We take G_1 to be quotient of G by the subgroup generated by these elements and ψ_1 the quotient map. Applying the lemma again we construct another sequence of elements in $\mathcal{J}(\mathbb{F}_q)$ out of which one at least is not in G_0. We go on like that and produce a sequence of subgroups in $\mathcal{J}(\mathbb{F}_q)$ that increase with constant probability until the index in $\mathcal{J}(\mathbb{F}_q)$ becomes smaller than $\max(48g, 24d, 720)$.

21.3.3 Lemma (Finding an almost generating set) *There exists a probabilistic (Monte-Carlo) algorithm that takes as input:*

1. *a degree d and geometric genus g plane projective absolutely irreducible reduced curve C over $\mathbb{F}_q$, such that $q \geq 4g^2$,*
2. *the smooth model $\mathcal{X}$ of C,*
3. *a degree g effective divisor ω, as origin,*

and outputs a sequence of elements in $\mathrm{Pic}^0(\mathcal{X}/\mathbb{F}_q)$ that generate a subgroup of index at most:

$$\max(48g, 24d, 720)$$

with probability $\geq 1/2$. The algorithm is polynomial time in d and $\log q$.

21.4 Pairings

Let m be a prime to p integer and $\mathcal{J}$ a Jacobian variety over $\mathbb{F}_q$. The Weil pairing relates the full m-torsion subgroup $\mathcal{J}(\overline{\mathbb{F}}_q)[m]$ with itself. It can be defined using Kummer theory and is geometric in nature. The Tate-Lichtenbaum-Frey-Rück pairing is more cohomological and relates the m-torsion $\mathcal{J}(\mathbb{F}_q)[m]$ in the group of $\mathbb{F}_q$-rational points and the quotient $\mathcal{J}(\mathbb{F}_q)/m\mathcal{J}(\mathbb{F}_q)$. In this section, we quickly review the definitions and algorithmic properties of these pairings, following work by Weil, Lang, Menezes, Okamoto, Vanstone, Frey and Rück.

For every Abelian variety A, we denote by $Z_0(A)_0$ the group of 0-cycles with degree 0 and by $S\colon Z_0(A)_0 \to A$ the summation map, that associates to every 0-cycle of degree 0 the corresponding sum in A.

Let V and W be two projective non-singular absolutely irreducible varieties over an algebraically closed field k with characteristic p, and let $\alpha\colon V \to A$ and $\beta\colon W \to B$ be canonical maps into their Albanese varieties. Let D be a correspondence on $V \times W$.

Let $n \geq 2$ be a prime to p integer. Let $\mathfrak{a}$ (resp. $\mathfrak{b}$) be a 0-cycle of degree 0 on V (resp. W) and let $a = S(\alpha(\mathfrak{a}))$ (resp. $b = S(\beta(\mathfrak{b}))$) be the associated point in A (resp. B). Assume $na = nb = 0$.

Following [66, VI, §4, Theorem 10] one can define the Weil pairing $e_{n,D}(a, b)$. It is an n-th root of unity in k. It depends linearly in a, b and D.

Assume $V = W = \mathcal{X}$ is a smooth projective absolutely irreducible reduced curve and $A = B = \mathcal{J}$ is its Jacobian and $\alpha = \beta = f\colon \mathcal{X} \to \mathcal{J}$ is the Jacobi map (once chosen an origin on $\mathcal{X}$). If we take D to be the diagonal on $\mathcal{X} \times \mathcal{X}$ we define a pairing $e_{n,D}(a, b)$ that will be denoted $e_n(a, b)$ or $e_{n,\mathcal{X}}(a, b)$. It does not depend on the origin for the Jacobi map. It is non-degenerate.

If $\mathcal{Y}$ is another smooth projective absolutely irreducible reduced curve and $\mathcal{K}$ its Jacobian and $\phi\colon \mathcal{X} \to \mathcal{Y}$ a non-constant map with degree d, and $\phi^*\colon \mathcal{K} \to \mathcal{J}$ the associated map between Jacobians, then for a and b of order dividing n in $\mathcal{K}$ one has $e_{n,\mathcal{X}}(\phi^*(a), \phi^*(b)) = e_{n,\mathcal{Y}}(a, b)^d$.

The Frey-Rück pairing can be constructed from the Lichtenbaum version of the Tate pairing [76] as was shown in [46]. Let $\mathbb{F}_q$ be a finite field with characteristic p. Let again $n \geq 2$ be a prime to p integer and $\mathcal{X}$ a smooth projective absolutely irreducible reduced curve over $\mathbb{F}_q$. Let g be the genus of $\mathcal{X}$. We assume n divides $q - 1$. Let $\mathcal{J}$ be the Jacobian of $\mathcal{X}$. The Frey-Rück pairing $\{,\}_n\colon \mathcal{J}(\mathbb{F}_q)[n] \times \mathcal{J}(\mathbb{F}_q)/n\mathcal{J}(\mathbb{F}_q) \to \mathbb{F}_q^*/(\mathbb{F}_q^*)^n$ is defined as follows. We take a class of order dividing n in $\mathcal{J}(\mathbb{F}_q)$. Such a class can be represented by an $\mathbb{F}_q$-divisor D with degree 0. We take a class in $\mathcal{J}(\mathbb{F}_q)$ and pick a degree zero $\mathbb{F}_q$-divisor E in this class, that we assume to be disjoint from D. The pairing evaluated at the classes $[D]$ and $[E] \bmod n$ is $\{[D], [E] \bmod n\}_n = f(E) \bmod (\mathbb{F}_q^*)^n$ where f is any function with divisor nD.

This is a non-degenerate pairing.

We now explain how one can compute the Weil pairing, following work by Menezes, Okamoto, Van Stone, Frey and Rück. The Tate-Lichtenbaum-Frey-Rück pairing can be computed similarly.

The Weil pairing is computed as follows. As usual, we assume we are given a degree d plane model C for $\mathcal{X}$. Assume $\mathfrak{a}$ and $\mathfrak{b}$ have disjoint support (otherwise we may replace $\mathfrak{a}$ by some linearly equivalent divisor using the explicit moving Lemma 21.2.7.)

We compute a function ϕ with divisor $n\mathfrak{a}$. We similarly compute a function ψ with divisor $n\mathfrak{b}$. Then $e_n(a, b) = \psi(\mathfrak{a})/\phi(\mathfrak{b})$. This algorithm is polynomial in the degree d of C and the order n of the divisors, provided the initial divisors $\mathfrak{a}$ and $\mathfrak{b}$ are given as differences between effective divisors with polynomial degree in d.

Using an idea that appears in a paper by Menezes, Okamoto and Vanstone [79] in the context of elliptic curves, and in [46] for general curves, one can make this algorithm polynomial in $\log n$ in the following way.

We write $\mathfrak{a} = \mathfrak{a}_0 = \mathfrak{a}_0^+ - \mathfrak{a}_0^-$ where $\mathfrak{a}_0^+$ and $\mathfrak{a}_0^-$ are effective divisors. Let ϕ be the function computed in the above simple minded algorithm. One has $(\phi) = n\mathfrak{a}_0^+ - n\mathfrak{a}_0^-$. We want to express ϕ as a product of small degree functions. We use a variant of fast exponentiation. Using Lemma 21.2.7 we compute a divisor $\mathfrak{a}_1 = \mathfrak{a}_1^+ - \mathfrak{a}_1^-$ and a function ϕ_1 such that $\mathfrak{a}_1$ is disjoint from $\mathfrak{b}$ and $(\phi_1) = \mathfrak{a}_1 - 2\mathfrak{a}_0$ and such that the degrees of $\mathfrak{a}_1^+$ and $\mathfrak{a}_1^-$ are $\leq 12gd(\log_q(\deg(\mathfrak{b})) + 1)$. We go on and compute, for $k \geq 1$ an integer, a divisor $\mathfrak{a}_k = \mathfrak{a}_k^+ - \mathfrak{a}_k^-$ and a function ϕ_k such that $\mathfrak{a}_k$ is disjoint from $\mathfrak{b}$ and $(\phi_k) = \mathfrak{a}_k - 2\mathfrak{a}_{k-1}$ and such that the degrees of $\mathfrak{a}_k^+$ and $\mathfrak{a}_k^-$ are $\leq 12gd(\log_q(\deg(\mathfrak{b})) + 1)$.

We write the base 2 expansion of $n = \sum_i \epsilon_k 2^k$ with $\epsilon_k \in \{0, 1\}$. We compute the function Ψ with divisor $\sum_k \epsilon_k \mathfrak{a}_k$. We claim that the function ϕ can be written as a product of the ϕ_k, for $k \leq \log_2 n$, and Ψ with suitable integer exponents bounded by n in absolute value. Indeed we write $\mu_1 = \phi_1$, $\mu_2 = \phi_2\phi_1^2$, $\mu_3 = \phi_3\phi_2^2\phi_1^4$ and so on. We have $(\mu_k) = \mathfrak{a}_k - 2^k\mathfrak{a}$ and $\Psi \prod_k \mu_k^{-\epsilon_k}$ has divisor $n\mathfrak{a}$ so is the ϕ we were looking for.

21.4.1 Lemma (Computing the Weil pairing) *There exists an algorithm that on input a prime to q integer $n \geq 2$ and a degree d absolutely irreducible reduced plane projective curve C over $\mathbb{F}_q$ and its smooth model $\mathcal{X}$ and two $\mathbb{F}_q$-divisors on $\mathcal{X}$ given as differences of disjoint effective divisors, denoted $\mathfrak{a} = \mathfrak{a}^+ - \mathfrak{a}^-$ and $\mathfrak{b} = \mathfrak{b}^+ - \mathfrak{b}^-$, with degree 0, and order dividing n in the Jacobian, computes the Weil pairing $e_n(\mathfrak{a}, \mathfrak{b})$ in time polynomial in d, $\log q$, $\log n$ and the degrees of $\mathfrak{a}^+$, $\mathfrak{a}^-$, $\mathfrak{b}^+$, $\mathfrak{b}^-$.*

21.4.2 Lemma (Computation of Tate-Lichtenbaum-Frey-Rück pairings) *There exists an algorithm that on input a prime to q integer $n \geq 2$ and a degree d absolutely irreducible reduced plane projective curve C over $\mathbb{F}_q$ and its smooth model $\mathcal{X}$ and two $\mathbb{F}_q$-divisors on $\mathcal{X}$ given as differences of disjoint effective divisors, denoted $\mathfrak{a} = \mathfrak{a}^+ - \mathfrak{a}^-$ and $\mathfrak{b} = \mathfrak{b}^+ - \mathfrak{b}^-$, with degree 0, and such that the class of $\mathfrak{a}$ has order dividing $n \geq 2$ in the Jacobian, computes the Tate-Lichtenbaum-Frey-Rück pairing $\{\mathfrak{a}, \mathfrak{b}\}_n$ in time polynomial in d, $\log q$, $\log n$ and the degrees of $\mathfrak{a}^+$, $\mathfrak{a}^-$, $\mathfrak{b}^+$, $\mathfrak{b}^-$.*

21.5 Divisible groups

For ℓ a prime integer, it is convenient to introduce ℓ-divisible subgroups inside the ℓ^∞-torsion of a Jacobian $\mathcal{J}$, that may or may not correspond to subvarieties. We see how to define such subgroups and control their rationality properties.

21.5.1 Definition (Divisible group) Let $\mathbb{F}_q$ be a finite field with characteristic p and let $\mathcal{X}$ be a projective smooth absolutely irreducible reduced algebraic curve over $\mathbb{F}_q$. Let g be the genus of

$\mathcal{X}$ and let ℓ be a prime integer and $n = \ell^k$ a power of ℓ. We assume $g \geq 1$. Let $\mathcal{J}$ be the Jacobian of $\mathcal{X}$ and let $\mathrm{End}(\mathcal{J}/\mathbb{F}_q)$ be the ring of endomorphisms of $\mathcal{J}$ over $\mathbb{F}_q$. Let $\Pi\colon J[\ell^\infty] \to J[\ell^\infty]$ be a group homomorphism whose restriction to its image $\mathbb{G}$ is a bijection. Multiplication by ℓ is then a surjection from $\mathbb{G}$ to itself. We denote by $\mathbb{G}[\ell^k]$ the ℓ^k-torsion in $\mathbb{G}$. There is an integer w such that $\mathbb{G}[\ell^k]$ is a free $\mathbb{Z}/\ell^k\mathbb{Z}$ module of rank w for every k. We assume that Π commutes with the Frobenius endomorphism F_q. We then say $\mathbb{G}$ is the divisible group associated with Π. From Tate's theorem [108] Π is induced by some endomorphism in $\mathrm{End}(\mathcal{J}/\mathbb{F}_q) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ and we can define Π^* the Rosati dual of Π and denote by $\mathbb{G}^* = \mathrm{im}(\Pi^*)$ the associated divisible group, that we call the adjoint of $\mathbb{G}$.

Let $\chi(X)$ be the characteristic polynomial of the Frobenius endomorphism $F_q \in \mathrm{End}(\mathcal{J}/\mathbb{F}_q)$. Let $F(X) = F_1(X)$ and $G(X) = G_1(X)$ be two unitary coprime polynomials in $\mathbb{F}_\ell[X]$ such that $\chi(X) = F_1(X)G_1(X) \bmod \ell$. From Bezout's theorem we have two polynomials $H_1(X)$ and $K_1(X)$ in $\mathbb{F}_\ell[X]$ such that $F_1H_1 + G_1K_1 = 1$ and $\deg(H_1) < \deg(G_1)$ and $\deg(K_1) < \deg(F_1)$. From Hensel's lemma, for every positive integer k there exist four polynomials $F_k(X)$, $G_k(X)$, $H_k(X)$ and $K_k(X)$ in $(\mathbb{Z}/\ell^k\mathbb{Z})[X]$ such that F_k and G_k are unitary and $\chi(X) = F_k(X)G_k(X) \bmod \ell^k$ and $F_kH_k + G_kK_k = 1 \bmod \ell^k$ and $\deg(H_k) < \deg(G_1)$ and $\deg(K_1) < \deg(F_1)$ and $F_1 = F_k \bmod \ell$, $G_1 = G_k \bmod \ell$, $H_1 = H_k \bmod \ell$, $K_1 = K_k \bmod \ell$. The sequences $(F_k)_k$, $(G_k)_k$, $(H_k)_k$, and $(K_k)_k$ converge in $\mathbb{Z}_\ell[X]$ to F_0, G_0, H_0, and K_0. If we substitute X by F_q in F_0H_0 we define a map $\Pi_G\colon \mathcal{J}[\ell^\infty] \to \mathcal{J}[\ell^\infty]$ and similarly, if we substitute X by F_q in G_0K_0 we define a map Π_F. It is clear that $\Pi_F^2 = \Pi_F$ and $\Pi_G^2 = \Pi_G$ and $\Pi_F + \Pi_G = 1$ and $\Pi_F\Pi_G = 0$. We call $\mathbb{G}_F = \mathrm{im}(\Pi_F)$ and $\mathbb{G}_G = \mathrm{im}(\Pi_G)$ the associated supplementary ℓ-divisible groups. The Rosati dual to F_q is q/F_q. Let $\mathcal{O} = \mathbb{Z}[X]/\chi(X)$ and $\mathcal{O}_\ell = \mathbb{Z}_\ell[X]/\chi(X)$. We set $\varphi_q = X \bmod \chi(X) \in \mathcal{O}$. Mapping φ_q onto F_q defines an epimorphism from the ring $\mathcal{O}$ onto $\mathbb{Z}[F_q]$.

21.5.2 Definition (Characteristic subspaces) For every non-trivial unitary factor F of $\chi(X)$ mod ℓ such that the cofactor $G = \chi/F \bmod \ell$ is prime to F, we write $\chi = F_0G_0$ the corresponding factorisation in $\mathbb{Z}_\ell[X]$. The ℓ-divisible group $\mathbb{G}_F$ is called the F_0-torsion in $\mathcal{J}[\ell^\infty]$ and is denoted $\mathcal{J}[\ell^\infty, F_0]$. It is the characteristic subspace of F_q associated with the factor F. If $F = (X-1)^e$ is the largest power of $X-1$ dividing $\chi(X) \bmod \ell$ we abbreviate $\mathbb{G}_{(X-1)^e} = \mathbb{G}_1$. If $\ell \neq p$ and $F = (X-q)^e$ then we write similarly $\mathbb{G}_{(X-q)^e} = \mathbb{G}_q = \mathbb{G}_1^*$.

We now compute fields of definitions for torsion points.

We assume ℓ is prime to q. The element φ_q belongs to the unit group $\mathcal{U}_1 = (\mathcal{O}/\ell\mathcal{O})^*$ of the quotient algebra $\mathcal{O}/\ell\mathcal{O} = \mathbb{F}_\ell[X]/\chi(X)$. Let the prime factorisation of $\chi(X) \bmod \ell$ be $\prod_i \chi_i(X)^{e_i}$ with $\deg(\chi_i) = f_i$. The order of $\mathcal{U}_1$ is $\prod_i (\ell^{f_i} - 1)\ell^{(e_i-1)f_i}$. Let γ be the small-

est integer such that ℓ^γ is bigger than or equal to $2g$. Then the exponent of the group $\mathcal{U}_1$ divides $A_1 = \ell^\gamma \prod_i (\ell^{f_i} - 1)$. We set $B_1 = \prod_i (\ell^{f_i} - 1)$ and $C_1 = \ell^\gamma$. There is a unique polynomial $M_1(X)$ with degree $< 2g$ such that $(\varphi_q^{A_1} - 1)/\ell = M_1(\varphi_q) \in \mathcal{O}$.

Now for every positive integer k, the element φ_q belongs to the unit group $\mathcal{U}_k = (\mathcal{O}/\ell^k\mathcal{O})^*$ of the quotient algebra $\mathcal{O}/\ell^k\mathcal{O} = \mathbb{Z}[X]/(\ell^k, \chi(X))$. The prime factorisation of $\chi(X) \bmod \ell$ is lifted modulo ℓ^k as $\prod_i \Xi_i(X)$ with $\deg(\Xi_i) = e_i f_i$ and the order of $\mathcal{U}_k$ is $\prod_i (\ell^{f_i} - 1)\ell^{f_i(ke_i - 1)}$. The exponent of the latter group divides $A_k = A_1\ell^{k-1}$. So we set $B_k = B_1 = \prod_i (\ell^{f_i} - 1)$ and $C_k = C_1\ell^{k-1} = \ell^{k-1+\gamma}$. There is a unique polynomial $M_k(X)$ with degree $< \deg(\chi)$ such that $\varphi_q^{A_k} - 1)/\ell^k = M_k(\varphi_q) \in \mathcal{O}$. For every integer $N \geq 2$ we can compute $M_k(X) \bmod N$ from $\chi(X)$ in probabilistic polynomial time in $\log q$, $\log \ell$, $\log N$, k, and g: we first factor $\chi(X) \bmod \ell$ then compute χ and the e_i and f_i. We compute X^{A_k} modulo $(\chi(X), \ell^k N)$ using fast exponentiation. We remove 1 and divide by ℓ^k.

21.5.3 Lemma (Frobenius and ℓ-torsion) *Let k be a positive integer and $\ell \neq p$ a prime. Let $\chi(X)$ be the characteristic polynomial of the Frobenius F_q of $\mathcal{J}$. Let e_i and f_i be the multiplicities and inertiae in the prime decomposition of $\chi(X) \bmod \ell$. Let γ be the smallest integer such that ℓ^γ is bigger than or equal to $2g$. Let $B = \prod_i (\ell^{f_i} - 1)$. Let $C_k = \ell^{k-1+\gamma}$ and $A_k = BC_k$. The ℓ^k-torsion in $\mathcal{J}$ decomposes over the degree A_k extension of $\mathbb{F}_q$. There is a degree $< 2g$ polynomial $M_k(X) \in \mathbb{Z}[X]$ such that $F_q^{A_k} = 1 + \ell^k M_k(F_q)$. For every integer N one can compute such a $M_k(X) \bmod N$ from $\chi(X)$ in probabilistic polynomial time in $\log q$, $\log \ell$, $\log N$, k, and g.*

We obtain sharper rationality results if we restrict to ℓ-divisible groups. So let $\chi = FG \bmod \ell$ with F and G unitary coprime and let $\chi = F_0G_0$ be the corresponding factorisation in $\mathbb{Z}_\ell[X]$. The action of F_q on the ℓ^k-torsion $\mathbb{G}_F[\ell^k] = \mathcal{J}[\ell^k, F_0]$ inside $\mathbb{G}_F$ factors through the ring $\mathcal{O}_\ell/(\ell^k, F_0) = \mathbb{Z}_\ell[X]/(\ell^k, F_0)$. We deduce the following result.

21.5.4 Lemma (Frobenius and F_0-torsion) *Let k be a positive integer and $\ell \neq p$ a prime. Let $\chi(X)$ be the characteristic polynomial of the Frobenius F_q of $\mathcal{J}$. Let $\chi = FG \bmod \ell$ with F and G unitary coprime. Let e_i and f_i be the multiplicities and inertiae in the prime decomposition of $F(X) \bmod \ell$. Let γ be the smallest integer such that ℓ^γ is bigger than or equal to $2g$. Let $B(F) = \prod_i (\ell^{f_i} - 1)$. Let $C_k(F) = \ell^{k-1+\gamma}$ and $A_k(F) = B(F)C_k(F)$. The ℓ^k-torsion in $\mathbb{G}_F$ decomposes over the degree $A_k(F)$ extension of $\mathbb{F}_q$. There is a degree $< \deg(F)$ polynomial $M_k(X) \in \mathbb{Z}_\ell[X]$ such that $\Pi_F F_q^{A_k(F)} = \Pi_F + \ell^k \Pi_F M_k(F_q)$. For every power N of ℓ, one can compute such an $M_k(X) \bmod N$ from $\chi(X)$ and $F(X)$ in probabilistic polynomial time in $\log q$, $\log \ell$, $\log N$, k, and g.*

If we take for F the largest power of $X - 1$ dividing $\chi(X) \bmod \ell$ in the above lemma, we have $B(F) = 1$ so $A_k(F)$ is an ℓ power $\leq 2g\ell^k$.

If we take for F the largest power of $X-q$ dividing $\chi(X) \bmod \ell$ in the above lemma, we have $B(F) = \ell - 1$ so $A_k(F)$ is $\leq 2g(\ell-1)\ell^k$.

So the characteristic spaces associated with the eigenvalues 1 and q decompose over small degree extensions of $\mathbb{F}_q$.

21.6 The Kummer map

Let $\mathcal{X}$ be a smooth projective algebraically irreducible reduced curve over $\mathbb{F}_q$ of genus g and $\mathcal{J}$ the Jacobian of $\mathcal{X}$. Let $n \geq 2$ be an integer dividing $q-1$. We assume $g \geq 1$. In this section, we construct a convenient surjection from $\mathcal{J}(\mathbb{F}_q)$ to $\mathcal{J}(\mathbb{F}_q)[n]$.

If P is in $\mathcal{J}(\mathbb{F}_q)$ we take some R such that $nR = P$ and form the 1-cocycle $({}^{\sigma}R - R)_{\sigma}$ in $\mathrm{H}^1(\mathbb{F}_q, \mathcal{J}[n])$. Using the Weil pairing we deduce an element:

$$\square \mapsto (e_n({}^{\sigma}R - R, \square))_{\sigma}$$

of

$$\begin{aligned}\mathrm{Hom}(\mathcal{J}[n](\mathbb{F}_q), \mathrm{H}^1(\mu_n)) &= \mathrm{Hom}(\mathcal{J}[n](\mathbb{F}_q), \mathrm{Hom}(\mathrm{Gal}(\mathbb{F}_q), \mu_n)) \\ &= \mathrm{Hom}(\mathcal{J}[n](\mathbb{F}_q), \mathbb{F}_q^*/(\mathbb{F}_q^*)^n).\end{aligned}$$

The map that sends $P \bmod n\mathcal{J}(\mathbb{F}_q)$ to $\square \mapsto (e_n({}^{\sigma}R - R, \square))_{\sigma}$ is injective because the Frey-Rück pairing is non-degenerate. We observe that $\mathrm{Hom}(\mathrm{Gal}(\mathbb{F}_q), \mu_n)$ is isomorphic to μ_n since an homomorphism from $\mathrm{Gal}(\mathbb{F}_q)$ to μ_n is characterised by the image of the Frobenius generator F_q. We obtain a bijection $T_{n,q}$ from $\mathcal{J}(\mathbb{F}_q)/n\mathcal{J}(\mathbb{F}_q)$ to the dual $\mathrm{Hom}(\mathcal{J}[n](\mathbb{F}_q), \mu_n)$ of $\mathcal{J}[n](\mathbb{F}_q)$ that we call the *Tate map*. It maps P onto $\square \mapsto e_n({}^{F_q}R - R, \square)$. If $\mathcal{J}[n]$ decomposes over $\mathbb{F}_q$ we set $K_{n,q}(P) = {}^{F_q}R - R$ and define a bijection $K_{n,q} \colon \mathcal{J}(\mathbb{F}_q)/n\mathcal{J}(\mathbb{F}_q) \to \mathcal{J}[n](\mathbb{F}_q) = \mathcal{J}[n]$ that we call the *Kummer map*.

We now assume that $n = \ell^k$ is a power of some prime integer $\ell \neq p$. We also make the (strong!) assumption that $\mathcal{J}[n]$ decomposes over $\mathbb{F}_q$. We want to compute the Kummer map $K_{n,q}$ explicitly. Let P be an $\mathbb{F}_q$-rational point in $\mathcal{J}$. Let R be such that $nR = P$. Since $F_q - 1$ kills $\mathcal{J}[n]$, there is an $\mathbb{F}_q$-endomorphism κ of $\mathcal{J}$ such that $F_q - 1 = n\kappa$. We note that κ belongs to $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}[F_q]$ and therefore commutes with F_q. We have $\kappa(P) = (F_q - 1)(R) = K_{n,q}(P)$ and $\kappa(P)$ is $\mathbb{F}_q$-rational.

Computing the Kummer map will show very useful but it requires that $\mathcal{J}[n]$ decomposes over $\mathbb{F}_q$. In general, we shall have to base change to some extension of $\mathbb{F}_q$.

Let $\chi(X)$ be the characteristic polynomial of F_q and let $B = \prod_i (\ell^{f_i} - 1)$ where the f_i are the degrees of prime divisors of $\chi(X) \pmod{\ell}$. Let ℓ^{γ} be the smallest power of ℓ that is bigger than

or equal to $2g$. Let $C_k = \ell^{\gamma+k-1}$ and $A_k = BC_k$. Set $Q = q^{A_k}$. From Lemma 21.5.3 there is a polynomial $M_k(X)$ such that $F_Q = 1 + \ell^k M_k(F_q)$. So, for P an $\mathbb{F}_Q$-rational point in $\mathcal{J}$ and R such that $nR = P$ the Kummer map $K_{n,Q}$ applied to P is $M_k(F_q)(P) = (F_Q - 1)(R) = K_{n,Q}(P)$ and this is an $\mathbb{F}_Q$-rational point.

21.6.1 Lemma (The Kummer map) *Let $\mathbb{F}_q$ be a finite field and let $\mathcal{X}$ be a projective smooth absolutely irreducible reduced algebraic curve over $\mathbb{F}_q$. Let g be the genus of $\mathcal{X}$ and let $\ell \neq p$ be a prime integer and $n = \ell^k$ a power of ℓ. We assume that $g \geq 1$. Let $\chi(X)$ be the characteristic polynomial of F_q and let $B = (\ell - 1)\prod_i (\ell^{f_i} - 1)$ where the f_i are the degrees of prime divisors of $\chi(X) \pmod{\ell}$. Let ℓ^γ be the smallest power of ℓ that is bigger than or equal to $2g$. Let $C_k = \ell^{\gamma+k-1}$ and $A_k = BC_k$. Set $Q = q^{A_k}$ and observe that n divides $Q - 1$. There exists an endomorphism $\kappa \in \mathbb{Z}[F_q]$ of $\mathcal{J}$ such that $n\kappa = F_Q - 1$ and for every $\mathbb{F}_Q$-rational point P and any R with $nR = P$ one has $\kappa(P) = (F_Q - 1)(R) = K_{n,Q}(P)$. This endomorphism κ induces a bijection between $\mathcal{J}(\mathbb{F}_Q)/n\mathcal{J}(\mathbb{F}_Q)$ and $\mathcal{J}[n](\mathbb{F}_Q) = \mathcal{J}[n]$. Given $\chi(X)$ and a positive integer N one can compute $\kappa \bmod N$ as a polynomial in F_q with coefficients in $\mathbb{Z}/N\mathbb{Z}$ in probabilistic polynomial time in g, $\log \ell$, $\log q$, k, and $\log N$.*

This lemma is not of much use in practice because the the field $\mathbb{F}_Q$ is too big. On the other hand, we may not be interested in the whole n-torsion in $\mathcal{J}$ but just a small piece in it, that may or may not correspond to a subvariety.

Let $\ell \neq p$ be a prime integer and $\mathbb{G}$ an ℓ-divisible group in $\mathcal{J}[\ell^\infty]$ and $\Pi = \Pi^2 \colon \mathcal{J}[\ell^\infty] \to \mathbb{G}$ a projection onto it. Let $n = \ell^k$ and let Q be a power of q such that $\mathbb{G}[n]$ decomposes over $\mathbb{F}_Q$. Let P be an $\mathbb{F}_Q$-rational point in $\mathbb{G}$. Let $R \in \mathbb{G}$ be such that $nR = P$. We set $K_{\mathbb{G},n,Q}(P) = {}^{F_Q}R - R$ and define an isomorphism $K_{\mathbb{G},n,Q} \colon \mathbb{G}(\mathbb{F}_Q)/n\mathbb{G}(\mathbb{F}_Q) \to \mathbb{G}(\mathbb{F}_Q)[n] = \mathbb{G}[n]$.

In order to make this construction explicit, we now assume that there exists some $\kappa \in \mathbb{Z}_\ell[F_q]$ such that $\Pi(F_Q - 1 - n\kappa) = 0$.

Lemma 21.5.4 provides us with such a Q and such a κ when $\mathbb{G} = \mathcal{J}[\ell^\infty, F_0]$ is some characteristic subspace.

We now can compute this new Kummer map $K_{\mathbb{G},n,Q}$. Let P be an $\mathbb{F}_Q$-rational point in $\mathbb{G}$. Let $R \in \mathbb{G}$ be such that $nR = P$. From $(F_Q - 1 - n\kappa)\Pi(R) = 0 = (F_Q - 1 - n\kappa)(R)$ we deduce that $K_{\mathbb{G},n,Q}(P) = \kappa(P)$. Hence the following result.

21.6.2 Lemma (The Kummer map) *Let $\mathbb{F}_q$ be a finite field and let $\mathcal{X}$ be a projective smooth absolutely irreducible reduced algebraic curve over $\mathbb{F}_q$. Let g be the genus of $\mathcal{X}$ and let $\ell \neq p$ be a prime integer and $n = \ell^k$ a power of ℓ. We assume that $g \geq 1$. Let $\chi(X)$ be the characteristic polynomial of F_q. Assume $\chi(X) = F(X)G(X) \bmod \ell$ with F and G unitary coprime polynomials in $\mathbb{F}_\ell[X]$ and let $\mathbb{G}_F$ be the associated ℓ-divisible group. Let*

$B = (\ell-1)\prod_i(\ell^{f_i}-1)$ *where the* f_i *are the degrees of prime divisors of* $F(X) \pmod{\ell}$. *Let* ℓ^γ *be the smallest power of* ℓ *that is bigger than or equal to* $2g$. *Let* $C_k = \ell^{k-1+\gamma}$ *and* $A_k = BC_k$. *Set* $Q = q^{A_k}$. *From Lemma 21.5.4 there exists an endomorphism* $\kappa \in \mathbb{Z}_\ell[F_q]$ *such that* $\Pi_F(n\kappa - F_Q + 1) = 0$ *and for every* $\mathbb{F}_Q$*-rational point* $P \in \mathbb{G}_F$ *and any* $R \in \mathbb{G}_F$ *with* $nR = P$ *one has* $\kappa(P) = (F_Q - 1)(R) = K_{\mathbb{G},n,Q}(P)$. *This endomorphism* κ *induces a bijection between* $\mathbb{G}_F(\mathbb{F}_Q)/n\mathbb{G}_F(\mathbb{F}_Q)$ *and* $\mathbb{G}_F[n](\mathbb{F}_Q) = \mathbb{G}_F[n]$. *Given* $\chi(X)$ *and* F *and a power* N *of* ℓ, *one can compute* $\kappa \bmod N$ *as a polynomial in* F_q *with coefficients in* $\mathbb{Z}/N\mathbb{Z}$ *in probabilistic polynomial time in* g, $\log \ell$, $\log q$, k, *and* $\log N$.

21.7 Linearisation of torsion classes

Let C be a degree d plane projective absolutely irreducible reduced curve C over $\mathbb{F}_q$ with geometric genus $g \geq 1$, and assume we are given the smooth model $\mathcal{X}$ of C. Let $\mathcal{J}$ be the Jacobian of $\mathcal{X}$. We assume $\ell \neq p$ is a prime integer that divides $\#\mathcal{J}(\mathbb{F}_q)$. Let $n = \ell^k$ be a power of ℓ. We want to describe $\mathcal{J}(\mathbb{F}_q)[\ell^k]$ by generators and relations.

If $x_1, x_2, \ldots, x_I$ are elements in a finite commutative group G we let $\mathcal{R}$ be the kernel of the map $\xi\colon \mathbb{Z}^I \to G$ defined by $\xi(a_1, \cdots, a_I) = \sum_i a_i x_i$. We call $\mathcal{R}$ the lattice of relations between the x_i.

We first give a very general and rough algorithm for computing relations in any finite commutative group.

21.7.1 Lemma (Finding relations in blackbox groups) *Let* G *be a finite and commutative group and let* $x_1, x_2, \ldots, x_I$ *be elements in* G. *A basis for the lattice of relations between the* x_i *can be computed at the expense of* $3I(\#G)^2$ *operations (or comparisons) in* G.

Proof We first compute and store all the multiples of x_1. So we list 0, x_1, $2x_1, \ldots$, until we find the first multiple e_1x_1 that is equal to zero. This gives us the relation $r_1 = (e_1, 0, \ldots, 0) \in \mathcal{R}$. This first step requires at most $o = \#G$ operations in G and o comparisons.

We then compute successive multiples of x_2 until we find the first one e_2x_2 that is in $L_1 = \{0, x_1, \ldots, (e_1-1)x_1\}$. This gives us a second relation r_2. The couple (r_1, r_2) is a basis for the lattice of relations between x_1 and x_2. Using this lattice, we compute the list L_2 of elements in the group generated by x_1 and x_2. This second step requires at most $2o$ operations and o^2 comparisons.

We then compute successive multiples of x_3 until we find the first one e_3x_3 that is in L_2. This gives us a third relation r_3. The triple (r_1, r_2, r_3) is a basis for the lattice of relations between x_1, x_2 and x_3. Using this lattice, we compute the list L_3 of elements in the group generated by x_1,

x_2 and x_3. This third step requires at most $2o$ operations and o^2 comparisons. And we go on like this. □

This is far from efficient unless the group is very small.

We come back to the computation of generators and relations for $\mathcal{J}(\mathbb{F}_q)[\ell^k]$.

Let $B = \ell - 1$. Let ℓ^γ be the smallest power of ℓ that is bigger than or equal to $2g$ and let $A_k = B\ell^{\gamma+k-1}$. We set $Q_k = q^{A_k}$.

If we take for F a power of $X - 1$ in Definition 21.5.2 and Lemma 21.6.2 we obtain two surjective maps $\Pi_1 \colon \mathcal{J}(\mathbb{F}_{Q_k})[\ell^\infty] \to \mathbb{G}_1(\mathbb{F}_{Q_k})$ and $K_{\mathbb{G}_1,\ell^k,Q_k} \colon \mathbb{G}_1(\mathbb{F}_{Q_k}) \to \mathbb{G}_1[\ell^k]$.

If we now take for F a power of $X - q$ in Definition 21.5.2 and Lemma 21.6.2 we obtain two surjective maps $\Pi_q \colon \mathcal{J}(\mathbb{F}_{Q_k})[\ell^\infty] \to \mathbb{G}_q(\mathbb{F}_{Q_k})$ and $K_{\mathbb{G}_q,\ell^k,Q_k} \colon \mathbb{G}_q(\mathbb{F}_{Q_k}) \to \mathbb{G}_q[\ell^k]$.

We observe that Π_1 and Π_q are Rosati dual to each other (as elements in $\mathrm{End}(\mathcal{J}/\mathbb{F}_q) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$), and adjoint for the Weil pairing.

Using Lemma 21.3.3 we produce a sequence $\gamma_1, \ldots, \gamma_I$ of elements in $\mathcal{J}(\mathbb{F}_{Q_k})$ that generate a subgroup of index at most $\iota = \max(48g, 24d, 720)$. Let N be the largest prime to ℓ divisor of $\#\mathcal{J}(\mathbb{F}_{Q_k})$. We set $\alpha_i = K_{\mathbb{G}_1,\ell^k,Q_k}(\Pi_1(N\gamma_i))$ and $\beta_i = K_{\mathbb{G}_q,\ell^k,Q_k}(\Pi_q(N\gamma_i))$.

The group $\mathcal{A}_k$ generated by the α_i has index at most ι in $\mathbb{G}_1[\ell^k]$. The group $\mathcal{B}_k$ generated by the β_i has index at most ι in $\mathbb{G}_q[\ell^k]$.

Let ℓ^δ be smallest power of ℓ that is bigger than ι and assume $k > \delta$. Then $\mathcal{A}_k$ contains $\mathbb{G}_1[\ell^{k-\delta}]$.

We now explain how to compute the lattice of relations between given elements $\rho_1, \ldots, \rho_J$ in $\mathbb{G}_1[\ell^k]$. We denote by $\mathcal{R}$ this lattice. We notice that the restriction of the Weil pairing to $\mathbb{G}_1[\ell^k] \times \mathbb{G}_q[\ell^k]$ is non-degenerate. We fix an isomorphism between the group $\mu_{\ell^k}(\overline{\mathbb{F}}_q)$ of ℓ^k-th roots and $\mathbb{Z}/\ell^k\mathbb{Z}$. We regard the matrix $(e_{\ell^k}(\beta_i, \rho_j))$ as a matrix with I lines, J columns and coefficients in $\mathbb{Z}/\ell^k\mathbb{Z}$. This matrix defines a morphism from $\mathbb{Z}^J$ to $(\mathbb{Z}/\ell^k\mathbb{Z})^I$ whose kernel is a lattice $\mathcal{R}'$ that contains $\mathcal{R}$. The index of $\mathcal{R}$ in $\mathcal{R}'$ is at most ι. Indeed $\mathcal{R}'/\mathcal{R}$ is isomorphic to the orthogonal of $\mathcal{B}_k$ in $\langle \rho_1, \ldots, \rho_J \rangle \subset \mathbb{G}_1[\ell^k]$ and the latter group has order $\leq \iota$. Once given a basis of $\mathcal{R}'$, the sublattice $\mathcal{R}$ can be computed using Lemma 21.7.1 at the expense of $\leq 3J\iota^2$ operations.

Assume k is bigger than δ. We apply this method to the generators $(\alpha_i)_i$ of $\mathcal{A}_k$. Once given the lattice $\mathcal{R}$ of relations between the α_i it is a matter of linear algebra to find a basis $(b_1, \ldots, b_w)$ for $\mathcal{A}_k[\ell^{k-\delta}] = \mathbb{G}_1[\ell^{k-\delta}]$. The latter group is a rank w free module over $\mathbb{Z}/\ell^{k-\delta}\mathbb{Z}$ and is acted on by the q-Frobenius F_q. For every b_j we can compute the lattice of relations between $F_q(b_j)$, b_1, $b_2, \ldots, b_w$ and deduce the matrix of F_q in the basis $(b_1, \ldots, b_w)$. From this matrix we deduce a nice generating set for the kernel of $F_q - 1$ in $\mathbb{G}_1[\ell^{k-\delta}]$. This kernel is $\mathcal{J}[\ell^{k-\delta}](\mathbb{F}_q)$. We deduce the following result.

21.7.2 Theorem *There is a probabilistic Monte-Carlo algorithm that on input:*

1. *a degree d and geometric genus g plane projective absolutely irreducible reduced curve C over $\mathbb{F}_q$,*
2. *the smooth model $\mathcal{X}$ of C,*
3. *a degree 1 divisor $O = O^+ - O^-$ where O^+ and O^- are effective, $\mathbb{F}_q$-rational and have degree bounded by a constant times g,*
4. *a prime ℓ different from the characteristic p of $\mathbb{F}_q$ and a power $n = \ell^k$ of ℓ,*
5. *the zeta function of $\mathcal{X}$;*

outputs a set $g_1, \ldots, g_W$ of divisor classes in the Picard group of $\mathcal{X}/\mathbb{F}_q$, such that the ℓ^k torsion $\mathrm{Pic}(\mathcal{X}/\mathbb{F}_q)[\ell^k]$ is the direct product of the $\langle g_i \rangle$, and the orders of the g_i form a non-decreasing sequence. Every class g_i is given by a divisor $G_i - gO$ in it, where G_i is a degree g effective $\mathbb{F}_q$-divisor on $\mathcal{X}$.

The algorithm runs in probabilistic polynomial time in d, g, $\log q$ and ℓ^k. It outputs the correct answer with probability $\geq 1/2$. Otherwise, it may return either nothing or a strict subgroup of $\mathrm{Pic}(\mathcal{X}/\mathbb{F}_q)[\ell^k]$.

If one is given a degree zero $\mathbb{F}_q$-divisor $D = D^+ - D^-$ of order dividing ℓ^k, one can compute the coordinates of the class of D in the basis $(g_i)_{1 \leq i \leq W}$ in polynomial time in d, $\log q$, ℓ^k and the degree of D^+. These coordinates are integers x_i such that $\sum_{1 \leq i \leq W} x_i g_i = [D]$.

21.8 An example: modular curves

In this subsection we consider a family of modular curves for which we can easily provide and study a plane model. Let $\ell \geq 5$ be a prime. We set $d_\ell = (\ell^2 - 1)/4$ and $m_\ell = (\ell - 1)/2$. We denote by $\mathcal{X}_\ell = X(2)_1(\ell)$ the moduli of elliptic curves with full 2-torsion plus one non-trivial ℓ-torsion point. We first describe a homogeneous singular plane model C_ℓ for this curve. We enumerate the places of $\mathcal{X}_\ell$ above every singularity of C_ℓ and compute the adjoint divisor $\mathfrak{C}_\ell$ using the Tate elliptic curve.

We let λ be an indeterminate and we consider the Legendre elliptic curve with equation $y^2 = x(x-1)(x-\lambda)$. Call $\mathcal{T}_\ell(\lambda, x)$ the ℓ-division polynomial of this curve. It is a polynomial in $\mathbb{Q}[\lambda][x]$ with degree $2d_\ell = (\ell^2 - 1)/2$ in x.

As a polynomial in x we have:

$$\mathcal{T}_\ell(\lambda, x) = \sum_{0 \leq k \leq 2d_\ell} a_{2d_\ell - k}(\lambda) x^k$$

where $a_0(\lambda)$ has degree 0 in λ so that we normalise by setting $a_0(\lambda) = \ell$.

We can compute the $2d_\ell$ roots of $\mathcal{T}_\ell(\lambda, x)$ in the field $\overline{\mathbb{Q}}\{\{\lambda^{-1}\}\}$ of Puiseux series in λ^{-1}. We set:

$$j = j(\lambda) = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2} = 2^8\lambda^2(1 - \lambda^{-1} + 3\lambda^{-2} + 3\lambda^{-4} + \cdots)$$

so:

$$j^{-1} = 2^{-8}(\lambda^{-2} + \lambda^{-3} - 2\lambda^{-4} - 5\lambda^{-5} + \cdots).$$

We introduce Tate's q-parameter, defined implicitly by $j = q^{-1} + 744 + 196884q + \cdots$ so:

$$\begin{aligned} q &= j^{-1} + 744j^{-2} + 750420j^{-3} + \cdots \\ &= \frac{1}{256}\lambda^{-2} + \frac{1}{256}\lambda^{-3} + \frac{29}{8192}\lambda^{-4} + \frac{13}{4096}\lambda^{-5} + \cdots . \end{aligned}$$

We set $x = x' + (1+\lambda)/3$ and $y' = y$ and find the reduced Weierstrass equation for the Legendre curve:

$$y'^2 = x'^3 - \frac{\lambda^2 - \lambda + 1}{3}x' - \frac{(\lambda - 2)(\lambda + 1)(2\lambda - 1)}{27}$$

We want to compare the latter curve and the Tate curve with equation:

$$y''^2 = x''^3 - \frac{E_4(q)}{48}x'' + \frac{E_6(q)}{864}$$

where $E_4(q) = 1 + 240q + \cdots$ and $E_6(q) = 1 - 504q + \cdots$.

The quotient $(E_4(q)(dq)^2)/((\lambda^2 - \lambda + 1)q^2)$ is a quadratic differential on the curve $X(2)$ with divisor $-2(0) - 2(1)$ in the λ coordinate. Examination of the leading terms of its expansion shows that:

$$E_4\left(\frac{dq}{q}\right)^2 = \frac{4(\lambda^2 - \lambda + 1)(d\lambda)^2}{\lambda^2(1 - \lambda)^2}$$

and similarly:

$$E_6\left(\frac{dq}{q}\right)^3 = \frac{4(\lambda - 2)(\lambda + 1)(2\lambda - 1)(d\lambda)^3}{\lambda^3(1 - \lambda)^3}.$$

We deduce the isomorphism $x' = \gamma^2 x''$ and $y' = \gamma^3 y''$ with:

$$\gamma^2 = 2\lambda(\lambda - 1)\left(\frac{dq}{qd\lambda}\right) = -4\lambda + 2 + \frac{3}{8}\lambda^{-1} + \frac{3}{16}\lambda^{-2} + \cdots .$$

Set $\zeta_\ell = \exp(2i\pi/\ell)$. For a and b integers such that either $b = 0$ and $1 \leq a \leq (\ell - 1)/2$ or $1 \leq b \leq (\ell - 1)/2$ and $0 \leq a \leq \ell - 1$ we set $w = \zeta_\ell^a q^{b/\ell}$ in the expansion:

$$x''(w, q) = \frac{1}{12} + \sum_{n \in \mathbb{Z}} \frac{wq^n}{(1 - wq^n)^2} - 2\sum_{n \geq 1} \frac{nq^n}{1 - q^n}$$

and find:

$$x''_{a,b} = \frac{1}{12} + \zeta_\ell^a q^{b/\ell} + O(q^{(b+1)/\ell})$$

if $b \neq 0$ and:

$$x''_{a,0} = \frac{1}{12} + \frac{\zeta_\ell^a}{(1-\zeta_\ell^a)^2} + O(q).$$

So:

$$x_{a,b} = \gamma^2 x'' + \frac{1+\lambda}{3} = -4\zeta_\ell^a 2^{-8b/\ell}\lambda^{1-2b/\ell} + O(\lambda^{1-(2b+1)/\ell})$$

if $b \neq 0$ and:

$$x_{a,0} = \frac{-4\zeta_\ell^a}{(1-\zeta_\ell^a)^2}\lambda + O(1).$$

The $x_{a,b}$ are the roots of $\mathcal{T}_\ell(\lambda, x)$ in the field $\overline{\mathbb{Q}}\{\{\lambda^{-1}\}\}$ of Puiseux series.

We deduce that for $1 \leq k \leq (\ell-1)/2$ the polynomial $a_k(\lambda)$ has degree at most k. Further:

$$a_{(\ell-1)/2}(\lambda) = 2^{\ell-1}(-\lambda)^{(\ell-1)/2} + O(\lambda^{(\ell-3)/2}).$$

For $k > (\ell-1)/2$ the polynomial $a_k(\lambda)$ has degree $< k$ and $\leq d_\ell$.

The coefficients in all the series expansions above are in $\overline{\mathbb{Z}}[1/6\ell]$. The coefficients in $\mathcal{T}_\ell(\lambda, x)$ are in $\mathbb{Z}[1/6\ell]$. In fact $\mathcal{T}_\ell(\lambda, x)$ is in $\mathbb{Z}[\lambda, x]$ but this is not needed here.

Since $\mathcal{T}_\ell \in \mathbb{Q}[\lambda, x]$ is absolutely irreducible, the equation $\mathcal{T}_\ell(\lambda, x) = 0$ defines a plane absolutely irreducible affine curve $\mathcal{C}_\ell$.

We denote by $T_\ell(\Lambda, X, Y) = \mathcal{T}_\ell(\Lambda/Y, X/Y)Y^{2d_\ell}$ the associated homogeneous polynomial and call $C_\ell \subset \mathbb{P}^2$ the corresponding projective curve.

For every place P on $\mathcal{X}_\ell$ such that $\lambda(P) \notin \{0, 1, \infty\}$, the function $\lambda - \lambda(P)$ is a uniformising parameter at P. Further $x(P)$ is finite and P is the only place of $\mathcal{X}_\ell$ above the point $(\lambda(P), x(P))$ of $\mathcal{C}_\ell$. So the only possible singularities of C_ℓ lie on one of the three lines with equations $\Lambda = 0$, $Y = 0$ and $\Lambda - Y = 0$.

The points at infinity are given by the degree $2d_\ell$ form:

$$2^{\ell-1}(-1)^{\frac{\ell-1}{2}}\Lambda^{\frac{\ell-1}{2}}X^{\frac{\ell^2-\ell}{2}} + \cdots + \ell X^{\frac{\ell^2-1}{2}} = X^{\frac{\ell^2-\ell}{2}} \prod_{0 \leq a \leq \frac{\ell-1}{2}} (-4\Lambda - (\zeta_\ell^a + \zeta_\ell^{-a} - 2)X).$$

We call $\Sigma_\infty = [1, 0, 0]$ the unique singular point at infinity and for every $1 \leq b \leq (\ell-1)/2$ we call $\sigma_{\infty,b}$ the point above Σ_∞ on $\mathcal{X}_\ell$ associated with the orbit $\{x_{0,b}, x_{1,b}, \ldots, x_{\ell-1,b}\}$ for the inertia group. We call $\mu_{\infty,a}$ the point on $\mathcal{X}_\ell$ corresponding to the expansion $x_{a,0}$. The ramification index of the covering map $\lambda\colon \mathcal{X}_\ell \to X(2)$ is ℓ at $\sigma_{\infty,b}$ and 1 at $\mu_{\infty,a}$. Since $\ell - 2b$ and ℓ are coprime, there exist two integers α_b and β_b such that $\alpha_b(\ell - 2b) - \beta_b\ell = 1$ and $1 \leq \alpha_b \leq \ell - 1$ and $1 \leq \beta_b \leq \ell - 1$. The monomial $x^{\alpha_b}\lambda^{-\beta_b} \in \overline{\mathbb{Q}}(\mathcal{X}_\ell)$ is a local parameter at $\sigma_{\infty,b}$. Of course, $\lambda^{-1/\ell}$ is also a local parameter at this point, and it is much more convenient, although it is not in $\overline{\mathbb{Q}}(\mathcal{X}_\ell)$.

The morphism $\phi\colon \mathcal{X}_\ell \to X_1(\ell)$ corresponding to forgetting the 2-torsion structure is Galois with group $\mathcal{S}_3$ generated by the two transpositions $\tau_{(0,\infty)}$ and $\tau_{(0,1)}$ defined in homogeneous coordinates by $\tau_{(0,\infty)}\colon [\Lambda, X, Y] \to [Y, X, \Lambda]$ and $\tau_{(0,1)}\colon [\Lambda, X, Y] \to [Y - \Lambda, Y - X, Y]$. We observe that these act on $\mathcal{X}_\ell$, $\mathbb{P}^2$ and C_ℓ in a way compatible with the maps $\mathcal{X}_\ell \to C_\ell$ and $C_\ell \subset \mathbb{P}^2$.

We set $\Sigma_0 = \tau_{(0,\infty)}(\Sigma_\infty) = [0, 0, 1]$ and $\Sigma_1 = \tau_{(0,1)}(\Sigma_0) = [1, 1, 1]$. We set $\sigma_{0,b} = \tau_{(0,\infty)}(\sigma_{\infty,b})$ and $\sigma_{1,b} = \tau_{(0,1)}(\sigma_{0,b})$, $\mu_{0,a} = \tau_{(0,\infty)}(\mu_{\infty,a})$ and $\mu_{1,a} = \tau_{(0,1)}(\mu_{0,a})$.

The genus of $\mathcal{X}_\ell$ is $g_\ell = (\ell - 3)^2/4 = (m_\ell - 1)^2$. The arithmetic genus of C_ℓ is $g_a = (m_\ell^2 + m_\ell - 1)(2m_\ell^2 + 2m_\ell - 1)$. We now compute the conductor of C_ℓ. Locally at Σ_∞ the curve C_ℓ consists of m_ℓ branches (one for each place $\sigma_{\infty,b}$) that are cusps with equations $(X/\Lambda)^\ell = -2^{2\ell-8b}(Y/\Lambda)^{2b} + \cdots$. The conductor of this latter cusp is $\sigma_{\infty,b}$ times $(\ell-1)(2b-1)$ which is the next integer to the last gap of the additive semigroup generated by ℓ and $2b$. The conductor of the full singularity Σ_∞ is now given by Gorenstein's formula [49, Theorem 2] and is:

$$\sum_{1\le b\le m_\ell} \{b(4m_\ell^2 + 4m_\ell - 1) - 2m_\ell - (2m_\ell + 1)b^2\} \cdot \sigma_{\infty,b}.$$

The full conductor $\mathfrak{C}_\ell$ is the sum of this plus the two corresponding terms to the isomorphic singularities Σ_0 and Σ_1. Some authors call it the adjunction divisor.

The degree $\deg(\mathfrak{C}_\ell)$ of $\mathfrak{C}_\ell$ is $2m_\ell(2m_\ell^3 + 4m_\ell^2 - 2m_\ell - 1)$. So we have:

$$\delta(\mathfrak{C}_\ell) = m_\ell(2m_\ell^3 + 4m_\ell^2 - 2m_\ell - 1)$$

and we check that $g_a = g_\ell + \delta(\mathfrak{C}_\ell)$.

Now let $p \notin \{2, 3, \ell\}$ be a prime. Let $\mathbb{C}_p$ be the field of p-adics and $\overline{\mathbb{F}}_p$ its residue field. We embed $\overline{\mathbb{Q}}$ in $\mathbb{C}_p$ and also in $\mathbb{C}$. In particular $\zeta_\ell = \exp(2i\pi/\ell)$ and $2^{1/\ell}$ are well defined as p-adic numbers.

We observe that in the calculations above, all coefficients belong to $\overline{\mathbb{Z}}[1/6\ell]$.

More precisely, the curves C_ℓ and $\mathcal{X}_\ell$ are defined over $\mathbb{Z}[1/6\ell]$. We note $C_\ell \bmod p = C_\ell/\mathbb{F}_p = C_\ell \otimes_{\mathbb{Z}[1/6\ell]} \mathbb{F}_p$ the reduction of C_ℓ modulo p and define similarly $\mathcal{X}_\ell \bmod p$. The points $\sigma_{\infty,b} \in \mathcal{X}_\ell$ are defined over $\mathbb{Z}[1/6\ell]$ and their reductions $\sigma_{\infty,b} \bmod p = \sigma_{\infty,b} \otimes_{\mathbb{Z}[1/6\ell]} \mathbb{F}_p$ are defined over $\mathbb{F}_p$. The points $\mu_{\infty,a} \in \mathcal{X}_\ell$ are defined over $\mathbb{Z}[\zeta_\ell, 1/6\ell]$ and their reductions $\mu_{\infty,a} \bmod p = \mu_{\infty,a} \otimes_{\mathbb{Z}[\zeta_\ell,1/6\ell]} \mathbb{F}_p(\zeta_\ell)$ are defined over $\mathbb{F}_p(\zeta_\ell)$.

We deduce the following result.

21.8.1 Lemma (Computing C_ℓ and resolving its singularities) *There exists a deterministic algorithm that given a prime $\ell \ge 5$ and a prime $p \notin \{2, 3, \ell\}$ and a finite field $\mathbb{F}_q$ with characteristic p such that $\zeta_\ell \bmod p$ and $2^{1/\ell} \bmod p$ belong to $\mathbb{F}_q$, computes the equation $\mathcal{T}_\ell(\lambda, x)$ modulo p and the expansions of all $x_{a,b}$ as series in $\lambda^{-1/\ell}$ with coefficients in $\mathbb{F}_q$, in time polynomial in ℓ, $\log q$ and the required $\lambda^{-1/\ell}$-adic accuracy.*

21.9 Another example of modular curves

In this subsection we consider another family of modular curves for which we can easily provide and study a plane model. This family will be useful in the calculation of modular representations as sketched in the next section. Let $\ell > 5$ be a prime. This time we set $\mathcal{X}_\ell = X_1(5\ell)$ the moduli of elliptic curves with one point of order 5ℓ. We first describe a homogeneous singular plane model C_ℓ for this curve. We enumerate the places of $\mathcal{X}_\ell$ above every singularity of C_ℓ and provide series expansions for affine coordinates at every such place.

Let b be an indeterminate and form the elliptic curve E_b in Tate normal form with equation:

$$y^2 + (1-b)xy - by = x^3 - bx^2.$$

The point $P = (0,0)$ has order 5 and its multiples are $2P = (b, b^2)$, $3P = (b, 0)$, and $4P = (0, b)$. The multiplication by ℓ isogeny induces a degree ℓ^2 rational fraction on x-coordinates: $x \mapsto \mathcal{N}(x)/\mathcal{M}(x)$ where $\mathcal{N}(x)$ is a unitary degree ℓ^2 polynomial in $\mathbb{Q}(b)[x]$. Recursion formulae for division polynomial (see [38], Section 3.6) provide a quick algorithm for computing this polynomial, and also show that the coefficients actually lie in $\mathbb{Z}[b]$. If ℓ is congruent to ± 1 modulo 5 then $\ell P = \pm P$ and x divides $\mathcal{N}(x)$. Otherwise $\mathcal{N}(x)$ is divisible by $x - b$.

Call $\mathcal{T}_\ell(b, x)$ the quotient of $\mathcal{N}(x)$ by x or $x - b$. This is a unitary polynomial in $\mathbb{Z}[b][x]$ with degree $\ell^2 - 1$ in x.

As a polynomial in x we have:

$$\mathcal{T}_\ell(b, x) = \sum_{0 \le k \le \ell^2 - 1} a_{\ell^2 - 1 - k}(b) x^k$$

where $a_0(\lambda) = 1$. Let d be the total degree of $\mathcal{T}_\ell$.

We denote by $T_\ell(B, X, Y) = \mathcal{T}_\ell(B/Y, X/Y)Y^d$ the associated homogeneous polynomial and call $C_\ell \subset \mathbb{P}^2$ the corresponding projective curve.

We set:

$$j = j(b) = \frac{(b^4 - 12b^3 + 14b^2 + 12b + 1)^3}{b^5(b^2 - 11b - 1)}.$$

Let $\sqrt{5} \in \mathbb{C}$ be the positive square root of 5 and let $\zeta_5 = \exp(2i\pi/5)$. Let $s = (11 + 5\sqrt{5})/2$ and $\bar{s}$ be the two roots of $b^2 - 11b - 1$.

The forgetting map $X_1(5\ell) \to X_1(5)$ is unramified except at $b \in \{0, \infty, s, \bar{s}\}$. For every place P on $\mathcal{X}_\ell$ such that $b(P) \notin \{0, s, \bar{s}, \infty\}$, the function $b - b(P)$ is a uniformising parameter at P.

Let $\mathcal{U}$ be the affine open set with equation $YB(B^2 - 11BY + Y^2) \neq 0$. Every point on $C_\ell \cap \mathcal{U}$ is smooth and all places of $\mathcal{X}_\ell$ above points in $C_\ell - \mathcal{U}$ are cusps in the modular sense (i.e. the modular invariant at these places is infinite).

In order to desingularise C_ℓ at a given cusp, we shall construct an isomorphism between the Tate q-curve and the completion of E_b at this cusp.

We call A_∞, A_0, A_s, $A_{\bar{s}}$ the points on $X_1(5)$ corresponding to the values ∞, 0, s and $\bar{s}$ of b.

We first study the situation locally at A_∞. A local parameter is b^{-1} and we have $j^{-1} = b^{-5} + 25b^{-6} + \cdots$.

We introduce Tate's q-parameter, defined implicitly by $j = q^{-1} + 744 + 196884q + \cdots$ so:

$$\begin{aligned} q &= j^{-1} + 744j^{-2} + 750420j^{-3} + \cdots \\ &= b^{-5} + 25b^{-6} + \cdots \end{aligned}$$

and we fix an embedding of the local field at A_∞ inside $\mathbb{C}\{\{q\}\}$ by setting:

$$b^{-1} = q^{1/5} - 5q^{2/5} + \cdots .$$

We set $x' = 36x + 3(b^2 - 6b + 1)$ and $y' = 108(2y + (1-b)x - b)$ and find the reduced Weierstrass equation:

$$y'^2 = x'^3 - 27(b^4 - 12b^3 + 14b^2 + 12b + 1)x' + 54(b^2+1)(b^4 - 18b^3 + 74b^2 + 18b + 1).$$

We want to compare the latter curve and the Tate curve with equation:

$$y''^2 = x''^3 - \frac{E_4(q)}{48}x'' + \frac{E_6(q)}{864}$$

where $E_4(q) = 1 + 240q + \cdots$ and $E_6(q) = 1 - 504q + \cdots$. See [56, Theorem 10.1.6].

From the classical (see [97, Proposition 7.1]) identities:

$$\left(\frac{qdj}{dq}\right)^2 = j(j - 1728)E_4$$

and:

$$\left(\frac{qdj}{dq}\right)^3 = -j^2(j - 1728)E_6$$

we deduce:

$$\left(\frac{qdb}{dq}\right)^2 = \frac{b^2(b^2 - 11b - 1)^2 E_4}{25(b^4 - 12b^3 + 14b^2 + 12b + 1)}$$

and:

$$\left(\frac{qdb}{dq}\right)^3 = -\frac{b^3(b^2 - 11b - 1)^3 E_6}{125(b^2+1)(b^4 - 18b^3 + 74b^2 + 18b + 1)}.$$

We deduce the isomorphism $x' = \gamma^2 x''$ and $y' = \gamma^3 y''$ with:

$$\gamma^2 = -\frac{36b(b^2 - 11b - 1)dq}{5qdb}.$$

The point P has (x, y) coordinates equal to $(0, 0)$. So:

$$x''(P) = 3(b^2 - 6b + 1)/\gamma^2 = \frac{1}{12} + b^{-2} + 11b^{-3} + \cdots = \frac{1}{12} + q^{\frac{2}{5}} + O(q^{\frac{3}{5}}).$$

Since on the Tate curve we have:

$$x''(w, q) = \frac{1}{12} + \sum_{n \in \mathbb{Z}} \frac{wq^n}{(1 - wq^n)^2} - 2\sum_{n \geq 1} \frac{nq^n}{1 - q^n} \tag{21.9.1}$$

we deduce that $w(P) = q^{\pm 2/5} \bmod \langle q \rangle$. We may decide for the sign in the exponent because we may choose any of the two isomorphisms corresponding to either possible values for γ. We decide that $w(P) = q^{2/5} \bmod \langle q \rangle$.

Set $\zeta_\ell = \exp(2i\pi/\ell)$. For α and β integers such that $0 \leq \alpha, \beta \leq \ell - 1$ we set $w = \zeta_\ell^\alpha q^{\beta/\ell} q^{2/5\ell}$ in the expansion 21.9.1 and find:

$$x''_{\alpha,\beta} = \frac{1}{12} + \zeta_\ell^\alpha q^{\beta/\ell} q^{2/5\ell}(1 + O(q^{1/5\ell}))$$

if $0 \leq \beta \leq (\ell - 1)/2$ and:

$$x''_{\alpha,\beta} = \frac{1}{12} + \zeta_\ell^{-\alpha} q^{1-\beta/\ell-2/5\ell}(1 + O(q^{1/5\ell}))$$

if $(\ell + 1)/2 \leq \beta \leq \ell - 1$.

Since:

$$x_{\alpha,\beta} = (\gamma^2 x''_{\alpha,\beta} - 3(b^2 - 6b + 1))/36$$

and:

$$\gamma^2 = 36b^2 - 216b - 396 + O(b^{-1}) = 36q^{-2/5} + 144q^{-1/5} + 144 + \cdots$$

we deduce that:

$$x_{\alpha,\beta} + 1 = \zeta_\ell^\alpha q^{\beta/\ell+2/5\ell-2/5}(1 + O(q^{1/5\ell}))$$

if $0 \leq \beta \leq (\ell - 1)/2$ and:

$$x_{\alpha,\beta} + 1 = \zeta_\ell^{-\alpha} q^{1-\beta/\ell-2/5\ell-2/5}(1 + O(q^{1/5\ell}))$$

if $(\ell + 1)/2 \leq \beta \leq \ell - 1$.

In particular, the degree of $\mathcal{T}_\ell(b, x)$ in b is $\leq 2(\ell^2 - 1)$.

For $0 \leq \alpha < \ell$ and $0 \leq \beta < \ell$ we set $\tilde{\alpha} = 5\alpha \bmod \ell$ and $\tilde{\beta} = 5\beta + 2 \bmod \ell$.

If $\tilde{\beta}$ is non-zero, the inertia group permutes cyclically the ℓ roots $x_{\alpha,\beta}$ for $0 \leq \alpha < \ell$. We call $\sigma_{\infty,\tilde{\beta}}$ the corresponding branch on $\mathcal{X}_\ell$. On the other hand, if $\beta = -2/5 \bmod \ell$ then $\tilde{\beta} = 0 \bmod \ell$ and every $x_{\alpha,-2/5 \bmod \ell}$ is fixed by inertia. We observe that $x_{0,-2/5 \bmod \ell}$ is either b or 0 and is

not a root of $\mathcal{T}_\ell(b,x)$. For $\tilde{\alpha}$ a non-zero residue modulo ℓ, we denote by $\mu_{\infty,\tilde{\alpha}}$ the branch on $\mathcal{X}_\ell$ corresponding to $x_{\alpha,-2/5 \bmod \ell}$.

So we have $\ell-1$ unramified places on $\mathcal{X}_\ell$ above A_∞ and $\ell-1$ ramified places with ramification index ℓ.

The coefficients in all the series expansions above are in $\mathbb{Z}[1/30,\zeta_\ell]$. The coefficients in $\mathcal{T}_\ell(b,x)$ are in $\mathbb{Z}$.

The curves C_ℓ and $\mathcal{X}_\ell$ are defined over $\mathbb{Z}[1/30\ell]$. Let $p \notin \{2,3,5,\ell\}$ be a prime. We note $C_\ell \bmod p = C_\ell/\mathbb{F}_p = C_\ell \otimes_{\mathbb{Z}[1/30\ell]} \mathbb{F}_p$ the reduction of C_ℓ modulo p and define similarly $\mathcal{X}_\ell \bmod p$.

From the discussion above we deduce the following result.

21.9.2 Lemma (Computing C_ℓ and resolving its singularities, I) *There exists a deterministic algorithm that given a prime $\ell \geq 7$ and a prime $p \notin \{2,3,5,\ell\}$ and a finite field $\mathbb{F}_q$ with characteristic p such that $\zeta_\ell \bmod p$ belongs to $\mathbb{F}_q$, computes the equation $\mathcal{T}_\ell(b,x)$ modulo p and the expansions of all $x_{\alpha,\beta}$ as series in $b^{-1/\ell}$ with coefficients in $\mathbb{F}_q$, in time polynomial in $\ell, \log q$ and the required $b^{-1/\ell}$-adic accuracy.*

Proof The equation is computed using recursion formulae for division polynomials. The q-series for the modular function j is given by:

$$j(q) = 1728 E_4^3(q)(E_4^3(q) - E_6^2(q))^{-1}$$

where:

$$E_4(q) = 1 + 240 \sum_{n \geq 1} \frac{n^3 q^n}{1-q^n}$$

and:

$$E_6(q) = 1 - 504 \sum_{n \geq 1} \frac{n^5 q^n}{1-q^n}.$$

The expansions for the $x_{\alpha,\beta}$ are then obtained through standard operations on series like product, sum, reversion, composition. And this is done in polynomial time in the absolute accuracy. □

Here are a few lines of GP-PARI code:

```
{ser(aa,bb,prec,ell,p,z,b,jc,E4,E6,D,jq,qc,gc,w,x)=if(1,
ell=7;
p=953;
z=Mod(431,p);
b=1/c;
```

```
jc=(b^4-12*b^3+14*b^2+12*b+1)^3/b^5/(b^2-11*b-1);
E4=sum(n=1,prec, n^3*q^n/(1-q^n))*240+1+O(q^prec);
E6=sum(n=1,prec, -n^5*q^n/(1-q^n))*504+1+O(q^prec);
D=(E4^3-E6^2)/1728;
jq=E4^3/D;
qc=subst(serreverse(1/jq),q,1/jc+O(c^prec));
gc= -36*b*(b^2-11*b-1)*deriv(qc)*(-c^2)/5/qc;
w=z^aa*Q^(2+5*bb);
xabs=Mod(1,p)*(1/12
+sum(n=1,prec,w*Q^(5*ell*n)/(1-w*Q^(5*ell*n))^2
+O(Q^(5*ell*prec))) +w/(1-w)^2
+sum(n=1,prec,Q^(5*ell*n)/w/(1-(w)^(-1)*Q^(5*ell*n))^2
+O(Q^(5*ell*prec)))
-2*sum(n=1,prec,n*Q^(5*ell*n)/(1-Q^(5*ell*n))
+O(Q^(5*ell*prec )) ));
cQ=subst(serreverse((qc/c^5)^(1/5)*c),c,Q^ell);
bQ=1/cQ;
gQ=subst(gc,c,cQ);
XabQ=(gQ*xabs-3*(bQ^2-6*bQ+1) )/36;
QC=subst(serreverse(1/((bQ*Q^ell)^(1/ell)/Q)),Q,C);
XabC=subst(XabQ,Q,QC);
,)}
```

The same holds for singular places above A_0. A local parameter at A_0 is b and:

$$j^{-1} = -b^5 + 25b^6 + \dots \quad \text{so} \quad q = -b^5 + 25b^6 + \dots$$

and we fix an embedding of the local field at A_0 inside $\mathbb{C}\{\{q\}\}$ by setting $b = -q^{1/5} + 5q^{2/5} + \dots$. From $\gamma^2 = 36 - 216q^{1/5} + \dots$ we deduce that the coordinate $x''(P)$ of the 5-torsion point P is $x''(P) = 1/12 + q^{1/5} + O(q^{2/5})$ so the parameter w at P can be taken to be $w(P) = q^{1/5} \bmod \langle q \rangle$ this time. For α and β integers such that $0 \leq \alpha, \beta \leq \ell - 1$ we set $w = \zeta_\ell^\alpha q^{\beta/\ell} q^{1/5\ell}$ in the expansion 21.9.1 and we finish as above.

Now, a local parameter at A_s is $b - s$ and $j^{-1} = (1/2 - 11\sqrt{5}/50)(b-s) + O((b-s)^2)$ so $q = (1/2 - 11\sqrt{5}/50)(b-s) + O((b-s)^2)$ and we fix an embedding of the local field at A_s inside $\mathbb{C}\{\{q\}\}$ by setting $b - s = ((125 + 55\sqrt{5})/2)q + O(q^2)$. We deduce that the coordinate $x''(P)$ of the 5-torsion point P is $x''(P) = 1/12 + w/(1-w)^2 + O(q)$ where $w = \exp(4i\pi/5) = \zeta_5^2$ so the parameter w at P can be taken to be $w(P) = \zeta_5^2 \bmod \langle q \rangle$ this time.

Altogether we have proved the following.

21.9.3 Lemma (Computing C_ℓ and resolving its singularities, II) *There exists a deterministic algorithm that given a prime $\ell \geq 7$ and a prime $p \notin \{2, 3, 5, \ell\}$ and a finite field $\mathbb{F}_q$ with characteristic p such that $\zeta_\ell \bmod p$ and $\zeta_5 \bmod p$ belong to $\mathbb{F}_q$, computes the equation $\mathcal{T}_\ell(b, x)$ modulo p and expansions (with coefficients in $\mathbb{F}_q$) at every singular branch of C_ℓ in time polynomial in ℓ, $\log q$ and the required number of significant terms in the expansions.*

We also need the following result due to Manin, Shokurov, Merel and Cremona [77, 80, 15, 45].

21.9.4 Lemma (Manin, Shokurov, Merel, Cremona) *For ℓ a prime and $p \notin \{5, \ell\}$ another prime, the zeta function of $\mathcal{X}_\ell \pmod p$ can be computed in time polynomial in ℓ and p.*

Proof We first compute the action of the Hecke operator T_p on the space of Manin symbols for the congruence group $\Gamma_1(5\ell)$ associated with $\mathcal{X}_\ell$. Then, from the Eichler-Shimura identity $T_p = F_p + p\langle p\rangle/F_p$ we deduce the characteristic polynomial of the Frobenius F_p. □

The lines below are written in the Magma language.

```
ZZ:=IntegerRing();
l:=11;
N:=5*11;
QN:=CyclotomicField(EulerPhi(N));
R1<T>:=PolynomialRing(QN,1);
R2<T,U>:=PolynomialRing(QN,2);
G := DirichletGroup(N,QN);
chars := Elements(G);
gen4:=chars[2];
gen10:=chars[5];
Genus(Gamma1(N));
charsmc:=[gen4,gen4^2,gen4^4, gen4*gen10,gen4^2*gen10,gen10,
gen4*gen10^2,gen4^2*gen10^2,gen10^2,gen4*gen10^5,
gen4^2*gen10^5,gen10^5];
p:=101;
PT:= R2 ! 1;
W:=1;
g:=1;
```

```
for eps in charsmc do

M := ModularForms([eps],2);
P:= R2 ! Evaluate(HeckePolynomial(CuspidalSubspace(M),p), T);
g:=Degree(P,T);
W :=   Evaluate(P,[ T+Evaluate(eps,p)*p/T,  1])*T^g;
PT:=PT*W;

end for;

PT := R2 ! PT;

k:=2;
PTk:= Resultant(PT, T^k-U,T);
```

21.10 Computing the Ramanujan subspace

This section explains the connection between the methods given here and Edixhoven's program for computing coefficients of modular forms. Recall the definition of the Ramanujan arithmetic τ function, related to the sum expansion of the discriminant form:

$$\Delta(q) = q\prod_{k\geq 1}(1-q^k)^{24} = \sum_{k\geq 1}\tau(k)q^k.$$

We call $\mathbb{E} \subset \mathrm{End}(J_1(\ell)/\mathbb{Q})$ the algebra of endomorphisms of $J_1(\ell)$ generated by the Hecke operators T_n for all integers $n \geq 2$. In view of Theorem 10.9 we make the following definition.

21.10.1 Definition (The Ramanujan ideal) We denote by $\mathfrak{m}$ the maximal ideal in $\mathbb{E}$ generated by ℓ and the $T_n - \tau(n)$. The subspace $J_1(\ell)[\mathfrak{m}]$ of the ℓ-torsion of $J_1(\ell)$ cut out by all $T_n - \tau(n)$ is called the Ramanujan subspace and denoted V_ℓ.

This V_ℓ is a 2-dimensional vector space over $\mathbb{F}_\ell$ and the characteristic polynomial of the Frobenius endomorphism F_p on it is $X^2 - \tau(p)X + p^{11} \bmod \ell$.

In this section, we address the problem of computing $\mathfrak{m}$-torsion divisors on modular curves over some extension field $\mathbb{F}_q$ of $\mathbb{F}_p$. The definition field $\mathbb{F}_q$ for such divisors can be predicted from the characteristic polynomial of F_p on V_ℓ. So the strategy is to pick random $\mathbb{F}_q$-points in the ℓ-torsion of the Jacobian $J_1(\ell)$ and to project them onto V_ℓ using Hecke operators.

In Section 21.9 we have defined the modular curve $\mathcal{X}_\ell = X_1(5\ell)$ and the degree 24 covering $\phi\colon \mathcal{X}_\ell \to X_1(\ell)$ of $X_1(\ell)$. We prefer $\mathcal{X}_\ell$ to $X_1(\ell)$ because we are able to construct a natural and convenient plane model for it, and anyway, we need to work with this curve in order to connect to the setup described in Section 13. The covering map $\phi\colon \mathcal{X}_\ell \to X_1(\ell)$ corresponds to forgetting the 5-torsion structure. It induces two morphisms $\phi^*\colon J_1(\ell) \to \mathcal{J}_\ell$ and $\phi_*\colon \mathcal{J}_\ell \to J_1(\ell)$ such that $\phi_* \circ \phi^* = [24]$ on $J_1(\ell)$.

The curve $\mathcal{X}_\ell$ provides a convenient computational model for the group of $\mathbb{F}_q$-points of the Jacobian of $X_1(\ell)$. Indeed, the Jacobian $J_1(\ell)$ of $X_1(\ell)$ and the Jacobian $\mathcal{J}_\ell$ of $\mathcal{X}_\ell$ are related by the natural map $\phi^*\colon J_1(\ell) \to \mathcal{J}_\ell$ induced by ϕ.

We denote by $\mathcal{A}_\ell \subset J_\ell$ the image of $\nu = \phi^* \circ \phi_*$. This is a subvariety of J_ℓ isogenous to $J_1(\ell)$. The restriction of ν to $\mathcal{A}_\ell$ is multiplication by 24.

The maps ϕ^* and ϕ_* induce Galois equivariant bijections between the N-torsion subgroups $J_1(\ell)[N]$ and $\mathcal{A}_\ell[N]$ for every prime to 6 integer N.

We call $W_\ell \subset \mathcal{A}_\ell \subset \mathcal{J}_\ell$ the image of the Ramanujan subspace by ϕ^*. We choose an integer $\widehat{24}$ such that $24 \times \widehat{24}$ is congruent to 1 modulo ℓ. We set $\hat{T}_n = [\widehat{24}] \circ \phi^* \circ T_n \circ \phi_*$ and notice that $\hat{T}_n \circ \phi^* = \phi^* \circ T_n$ on $J_1(\ell)[\ell]$. This way, the map $\phi^*\colon J_1(\ell) \to \mathcal{J}_\ell$ induces a Galois equivariant bijection of Hecke modules between the $J_1(\ell)[\ell]$ and $\mathcal{A}_\ell[\ell]$ and $W_\ell = \phi^*(V_\ell)$ is the subspace in $\mathcal{A}_\ell[\ell]$ cut out by all $\hat{T}_n - \tau(n)$. So W_ℓ will also be called the Ramanujan subspace whenever there is no risk of confusion.

We notice that ϕ^*, ϕ_*, T_n, and $\hat{T}_n$ can be seen as correspondences as well as morphisms between Jacobians, and we state the following lemma.

21.10.2 Lemma (Computing the Hecke action) *Let ℓ, p, n be primes such that $p \not| 5\ell$. Let q be a power of p and let D be an effective $\mathbb{F}_q$-divisor of degree d on $\mathcal{X}_\ell \pmod p$. The divisors $\phi^* \circ \phi_*(D)$ and $\phi^* \circ T_n \circ \phi_*(D)$ can be computed in deterministic polynomial time in ℓ, d, n and $\log q$.*

Proof Let $x = (E, u)$ be a point on $X_1(\ell)$ representing and elliptic curve E with a point u of order ℓ. Let n be an integer. The Hecke operator T_n maps x onto the sum of all $(E_\phi, \phi(u))$, where $\phi\colon E \to E_\phi$ runs over the set of all isogenies (actually, over the set of their kernels) of degree n from E such that $\phi(u)$ still has order ℓ. So we can compute the action of Hecke correspondences on points $x = (E, u)$ using Vélu's formulae.

There remains to treat the case of cusps.

We call $\sigma_{\tilde\beta}$ for $1 \le \tilde\beta \le (\ell-1)/2$ and $\mu_{\tilde\alpha}$ for $1 \le \tilde\alpha \le (\ell-1)/2$ the cusps on $X_1(\ell)$ images by ϕ of the $\sigma_{\infty,\tilde\beta}$ and $\mu_{\infty,\tilde\alpha}$.

For n prime to ℓ we have $T_n(\sigma_{\tilde\beta}) = \sigma_{\tilde\beta} + n\sigma_{n\tilde\beta}$ and $T_n(\mu_{\tilde\alpha}) = n\mu_{\tilde\alpha} + \mu_{n\tilde\alpha}$, where $n\tilde\alpha$ in $\mu_{n\tilde\alpha}$ (resp. $n\tilde\beta$ in $\sigma_{n\tilde\beta}$) should be understood as the class of this integer in $(\mathbb{Z}/\ell\mathbb{Z})^*/\{1,-1\}$.

Similarly $T_\ell(\sigma_{\tilde{\beta}}) = \sigma_{\tilde{\beta}} + 2\sum_{1\le\tilde{\alpha}\le(\ell-1)/2}\mu_{\tilde{\alpha}}$ and $T_\ell(\mu_{\tilde{\alpha}}) = \ell\mu_{\tilde{\alpha}}$. The identity 8.11 implies that $T_{l^i} = T_l^i$ for all $i \ge 0$. Using this one can compute the effect of T_n on cusps for all n. □

We can now state the following theorem.

21.10.3 Theorem *There is a probabilistic (Las Vegas) algorithm that on input a prime ℓ and a prime $p \ge 7$ such that ℓ is prime to p, computes the Ramanujan subspace $W_\ell = \phi^*(V_\ell)$ inside the ℓ-torsion of the Jacobian of $\mathcal{X}_\ell/\mathbb{F}_p$. The answer is given as a list of ℓ^2 degree g_ℓ effective divisors on $\mathcal{X}_\ell$, the first one being the origin ω. The algorithm runs in probabilistic polynomial time in p and ℓ.*

Proof Lemma 21.9.3 gives us a plane model for $\mathcal{X}_\ell \pmod p$ and a resolution of its singularities. From Lemma 21.9.4 we obtain the zeta function of $\mathcal{X}_\ell \pmod p$. The characteristic polynomial of F_p on the Ramanujan space V_ℓ is $X^2 - \tau(p)X + p^{11} \bmod \ell$. So we compute $\tau(p) \pmod \ell$ using the expansion of the discriminant form. We deduce some small enough field of decomposition $\mathbb{F}_q$ for $V_\ell \pmod p$. We then apply Theorem 21.7.2 and obtain a basis for the ℓ-torsion in the Picard group of $\mathcal{X}_\ell/\mathbb{F}_q$. The same theorem allows us to compute the matrix of the endomorphism $\nu = \phi^* \circ \phi_*$ in this basis. We deduce a basis for the image $\mathcal{A}[\ell](\mathbb{F}_q)$ of ν. Using Theorem 21.7.2 again, we now write down the matrices of the Hecke operators $\hat{T}_n$ in this basis for all $n < \ell^2$. It is then a matter of linear algebra to compute a basis for the intersection of the kernels of all $\hat{T}_n - \tau(n)$ in $\mathcal{A}[\ell](\mathbb{F}_q)$. The algorithm is Las Vegas rather than Monte-Carlo because we can check the result, the group W_ℓ having known cardinality ℓ^2. □

21.10.4 Remark In the above theorem, one may impose an origin ω rather than letting the algorithm choose it. For example, one may choose as origin a well designed linear combination of the cusps, as in Section 12. Such an adapted choice of the origin may insure that the $\ell^2 - 1$ divisors representing the non-zero classes in W_ℓ are unique in characteristic zero and thus remain unique modulo p for all but finitely many primes p.

22 Computing the mod l Galois representations associated to τ in time polynomial in l

22.1 Theorem *There exists a probabilistic algorithm that computes the mod l Galois representation associated to Δ in time polynomial in l. More precisely, on input a prime number l it gives:*

1. *a Galois extension K_l of $\mathbb{Q}$, given as a $\mathbb{Q}$-basis e and the products e_ie_j (i.e., the $a_{i,j,k}$ in $\mathbb{Q}$ such that $e_ie_j = \sum_k a_{i,j,k}e_k$ are given);*

2. *a list of the elements σ of $\mathrm{Gal}(K_l/\mathbb{Q})$, where each σ is given as its matrix with respect to e;*

3. *an injective morphism ρ_l from $\mathrm{Gal}(K_l/\mathbb{Q})$ into $\mathrm{GL}_2(\mathbb{F}_l)$,*

such that K_l is unramified outside $\{l\}$, and for all prime numbers p different from l we have $\mathrm{trace}(\rho_l\mathrm{Frob}_p) = \tau(p)$ mod l and $\det(\rho_l\mathrm{Frob}_p) = p^{11}$mod l. The expected running time of the algorithm is polynomial in l.

Proof We may and do assume that $l > 11$ and that the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{Aut}(V_l)$ contains all elements with determinant 1 (just take $l \notin \{23, 691\}$). As this proof is rather long, we divide it into subsections.

22.2 Computing the $\mathbb{Q}(\zeta_l)$-algebra corresponding to $V_l - \{0\}$

We start with choosing a divisor D as in Theorem 12.8. Then the D'_x on $X_{l,\overline{\mathbb{Q}}}$ are uniquely defined. Using Couveignes' Theorem 21.10.3, we try to compute the $(D'_x)_{\mathbb{F}}$ over suitable extensions of the residue fields of $\mathbb{Z}[\zeta_l]$ at successive prime numbers p up to some bound B_l that will be polynomial in l, skipping 5 and l. If some $(D'_x)_{\mathbb{F}}$ is not unique, this will be detected by our computations, and we throw the corresponding prime p away. By Theorem 19.5, at most $O(l^{14})$ primes p are thrown away. For the l-good primes $p \leq B_l$ we then have computed all $(D'_x)_{\mathbb{F}}$.

We split each such $(D'_x)_{\mathbb{F}}$ in a cuspidal part and a non-cuspidal part. By Theorem 19.5, there are at most $O(l^{14})$ primes p where at some $\mathbb{F}$ the degree of the cuspidal part of $(D'_x)_{\mathbb{F}}$ is larger than the degree $g_l - d_l$ of the cuspidal part $(D'_x)^{\mathrm{rest}}$ of D'_x. This means that we have computed d_l. We discard the primes p where some degree of the cuspidal part of a $(D'_x)_{\mathbb{F}}$ is larger than d_l. For the remaining primes, we have computed the $(D''_x)_{\mathbb{F}}$ for all x in $V_l - \{0\}$.

For one residue field $\mathbb{Z}[\zeta_l] \twoheadrightarrow \mathbb{F}_q$ that is left, we first take a suitable $\mathbb{F}_q$-linear combination f of the functions b, x and y, with coefficients $\overline{\alpha}$, $\overline{\beta}$ and $\overline{\gamma}$, such that $f\colon X_l(\overline{\mathbb{F}}_q) \to \mathbb{P}^1(\overline{\mathbb{F}}_q)$ is injective on $D''(\overline{\mathbb{F}}_q)$, where $D'' = \sum_x D''_x$; we take q large enough with respect to $\deg D'' = O(l^4)$ so that such a linear combination is guaranteed to exist and easy to find. Then we lift these coefficients to small elements α, β and γ of $\mathbb{Z}[\zeta_l]$ (e.g., we can use the $\mathbb{Z}$-basis $(1, \zeta_l, \ldots, \zeta_l^{l-2})$ and lift elements from $\mathbb{F}_p$ to $\mathbb{Z}$ with minimum absolute value). The divisors $f_*D''_x$ are pairwise distinct when x varies.

As we have now chosen our function f, we have the elements $k_{D,f,d}$, $1 \leq d \leq d_l$, of the $\mathbb{Q}(\zeta_l)$-algebra $K'_{l,\mathbb{Q}(\zeta_l)}$ corresponding to the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$-set $V_l - \{0\}$, as defined in (13.6). We

recall that:

$$k_{D,f,d}\colon V_l - \{0\} \to \overline{\mathbb{Q}}, \quad x \mapsto \Sigma_d(f(Q_{x,1}), \ldots, f(Q_{x,d_l})), \quad \text{where } D''_x = \sum_{1 \le i \le d_l} Q_{x,i},$$

and where Σ_d denotes the elementary symmetric polynomial of degree d. We note that Σ_d, even though it has many terms, can be evaluated efficiently by expressing it in power sums. By construction, these $k_{D,f,d}$ generate $K'_{l,\mathbb{Q}(\zeta_l)}$. Even better, these $k_{D,f,d}$ are integral at the place $\mathbb{Z}[\zeta_l] \twoheadrightarrow \mathbb{F}_q$ that we used for constructing f, and their reductions generate the $\mathbb{F}_q$-algebra $\mathbb{F}_q \otimes_{\mathbb{Z}[\zeta_l]} O_{K'_{l,\mathbb{Q}(\zeta_l)}}$. As q is sufficiently large with respect to the dimension of this $\mathbb{F}_q$-algebra (which is $O(l^2)$), this algebra is generated by a linear combination of the $k_{D,f,d}$. We compute such a linear combination and lift it with small coefficients to a linear combination:

$$k_{d,f} := \sum_{1 \le d \le d_l} a_d k_{D,f,d}, \quad a_d \in \mathbb{Z}[\zeta_l].$$

The minimal polynomial $P_{l,f,D}$ in $\mathbb{Q}(\zeta_l)[T]$ of $k_{D,f}$ over $\mathbb{Q}(\zeta_l)$ is given as:

$$P_{l,D,f} = \prod_{x \neq 0} (T - k_{D,f}(x)).$$

Theorem 19.3 tells us that the heights of the coefficients of $P_{l,D,f}$ are bounded by a fixed power of l.

Our construction shows that $P_{l,D,f}$ is integral at all primes p that are l-good, and Couveignes has shown that for those with $p \le B_l$ we can compute all reductions $(P_{l,D,f})_{\mathbb{F}}$ in probabilistic time polynomial in l. If the sum of the $\log \#\mathbb{F}$ where we have computed the $(P_{l,D,f})_{\mathbb{F}}$ is large enough with respect to the height of the coefficients of $P_{l,D,f}$, then $P_{l,D,f}$ is determined by these reductions, and there is an efficient algorithm to compute it. So we have now computed $P_{l,D,f}$ in time polynomial in l. Of course, we then have $K'_{l,\mathbb{Q}(\zeta_l)} = \mathbb{Q}(\zeta_l)[T]/(P_{l,D,f})$.

22.3 Computing the $\mathbb{Q}$-algebra corresponding to $V_l - \{0\}$

At this moment, we must finally pay the price for working with a divisor D on $X_{l,\mathbb{Q}(\zeta_l)}$ and not on X_l itself. We have $K'_{l,\mathbb{Q}(\zeta_l)} = \mathbb{Q}(\zeta_l) \otimes K'_l$, hence we have a semi-linear action of $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q})$ on the $\mathbb{Q}(\zeta_l)$-algebra $K'_{l,\mathbb{Q}(\zeta_l)}$, and K'_l is precisely the subset of elements that are fixed by this action.

We take a generator σ of $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q}) = \mathbb{F}_l^\times$. We want to compute $\sigma(k_{D,f})$, as a $\mathbb{Q}(\zeta_l)$-linear combination of the $k_{D,f}^i$:

$$\sigma(k_{D,f}) = \sum_{0 \le i < l^2 - 1} c_i k_{D,f}^i. \tag{22.3.1}$$

We can compute the c_i by computing their reductions at sufficiently many residue fields $\mathbb{Z}[\zeta_l] \twoheadrightarrow \mathbb{F}$. In order to do that, we will first give a formula for the c_i in order to get a bound for their heights. We start with a lemma.

22.3.2 Lemma *Let K be a field and let L be a K-algebra of the form $K[T]/(f)$ with $f = T^d + f_{d-1}T^{d-1} + \cdots + f_0$ a separable polynomial with coefficients in K. We identify L with the K-vector space of polynomials of degree at most $d-1$ in T. Let the c_i be the coordinate functions of L with respect to the K-basis $(1, T, \ldots, T^{d-1})$: for g in L we have $g = \sum_i c_i(g)T^i$. Let $K \to \overline{K}$ be an algebraic closure, and let $r_1, \ldots, r_d$ be the roots of f in $\overline{K}$. Then we have, for each g in L:*

$$g = \sum_{j<d} g(r_j)\frac{\prod_{k\neq j}(T-r_k)}{\prod_{k\neq j}(r_j-r_k)}, \quad \textit{hence} \quad c_i(g) = \sum_{0\leq i<d} g(r_j)(-1)^{d-1-i}\frac{\Sigma_{d-1-i}(r^{(j)})}{\prod_{k\neq j}(r_j-r_k)},$$

where $r^{(j)}$ denotes the $d-1$-tuple obtained from $(r_1, \ldots, r_d)$ by removing the jth-coordinate.

Proof One simply uses the isomorphism:

$$L_{\overline{K}} = \overline{K}[T]/(f) \longrightarrow \overline{K}^d, \quad g \mapsto (g(r_1), \ldots, g(r_d)),$$

also called Lagrange interpolation. □

With the formulas of Lemma 22.3.2 it is an easy matter to derive a bound for the heights of the c_i that is a fixed power of l. Indeed, for x in $\overline{\mathbb{Q}}^\times$ we have $h(1/x) = h(x)$, hence the denominators $\prod_{k\neq j}(r_j - r_k)$ need not scare us.

Now that we have a bound on the height of the c_i, we explain how we can compute reductions of them at finite places. We simply let σ act on all of our construction of $k_{D,f}$. We let σ denote a lift of σ to $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

As x is represented by $D'_x - D$, σx is represented by $\sigma D'_x - \sigma D$. As $D'_x = D''_x + (D'_x)^{\mathrm{rest}}$, we have $\sigma D'_x = \sigma D''_x + (\sigma D'_x)^{\mathrm{rest}}$ where the first term is the cuspidal part. We have $\sigma D''_x = \sum_{i\leq d_l} \sigma Q_{x,i}$. We have $\sigma f = \sigma(\alpha)b + \sigma(\beta)x + \sigma(\gamma)y$, and we have:

$$\begin{aligned}(\sigma k_{D,f,d})(\sigma x) = \sigma(k_{D,f,d}(x)) = \Sigma_d(\sigma(f(Q_{x,1})), \ldots, \sigma(f(Q_{x,d_l}))) = \\ = \Sigma_d((\sigma f)(\sigma Q_{x,1}), \ldots, (\sigma f)(\sigma Q_{x,d_l})).\end{aligned}$$

Hence:

$$(\sigma k_{D,f,d})(x) = \Sigma_d((\sigma f)(\sigma Q_{\sigma^{-1}x,1}), \ldots, (\sigma f)(\sigma Q_{\sigma^{-1}x,d_l})).$$

We can interpret the last equality as:

$$(\sigma k_{D,f,d})(x) = k_{\sigma D,\sigma f,d}(\sigma^{-1}x).$$

Likewise:

$$(\sigma k_{D,f})(x) = \sum_{1 \leq d \leq d_l} \sigma(a_d) k_{\sigma D, \sigma f, d}(\sigma^{-1} x).$$

From this last identity we see that we can compute the reductions at the places of $\mathbb{Z}[\zeta_l]$ at p of left hand side of (22.3.1) for all p that are l-good, in time polynomial in l and p. This gives us the reductions of the c_i in the right hand side of (22.3.1), unless $k_{D,f}$ fails to be a generator at all places above p, in which case we just continue with the next prime. As we have produced $k_{D,f}$ in polynomial time, it can only fail to be a generator at a set of primes that we can allow ourselves to discard.

It follows that we have computed $\sigma(k_{D,f})$ in time polynomial in l. From this, we can compute the $\sigma(k_{D,f}^i)$ for $1 \leq i < l^2 - 1$. We can now produce a small generator of the $\mathbb{Q}$-algebra K_l' by taking the trace under the action of $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q})$ of a suitable element $\sum_i \lambda_i k_{D,f}^i$. Such a suitable element can be found by forcing it to be a generator modulo a suitable small prime p. We let $P_{l,\mathbb{Q}}$ denote the minimal polynomial of our generator of K_l'. The $\mathbb{Q}$-algebra corresponding to $V_l - \{0\}$, with its given generator, is then $\mathbb{Q}[T]/(P_{l,\mathbb{Q}})$.

22.3.3 Remark An alternative way to compute the action of σ on $K'_{l,\mathbb{Q}(\zeta_l)}$ would be to factor the minimal polynomial $\sigma(P_{l,D,f})$ over $K'_{l,\mathbb{Q}(\zeta_l)}$. However, it is easy to see that this factorisation gives precisely $l - 1$ linear factors, and to choose the correct one, one needs to do a check at a finite place, or, possibly, at an infinite place.

22.4 Computing the $\mathbb{F}_l^\times$-action on K_l'

The group $\mathbb{F}_l^\times$ acts on the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-set V_l, hence on K_l'. We want to compute this action, i.e., for every a in $\mathbb{F}_l^\times$ we want to give the image of our generator under that automorphism. We note that it suffices to compute the corresponding automorphisms of $K'_{l,\mathbb{Q}(\zeta_l)}$, and then deduce from that the image of our generator of K_l'. For a in $\mathbb{F}_l^\times$ we let $\langle a \rangle$ denote the corresponding automorphism of K_l' as well as that of $K'_{l,\mathbb{Q}(\zeta_l)}$.

We recall that $K'_{l,\mathbb{Q}(\zeta_l)}$ is the set of functions g from $V_l - \{0\}$ to $\overline{\mathbb{Q}}$ such that for all σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_l))$ and all x in $V_l - \{0\}$ we have $g(\sigma x) = \sigma(g(x))$. The action of a in $\mathbb{F}_l^\times$ is then defined by: $(\langle a \rangle g)x = g(ax)$. For example, we have:

$$(\langle a \rangle k_{D,f,d})x = k_{D,f,d}(ax) = \Sigma_d(f(Q_{ax,1}, \ldots, f(Q_{ax,d_l}))).$$

We can compute $\langle a \rangle k_{D,f}$ as a linear combination of the $k_{D,f}^i$ with $0 \leq i < l^2 - 1$ just as in the previous subsection we computed the action of $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q})$.

22.5 Computing K_l when $l \not\equiv 1 \bmod 11$

For simplicity, we will now first assume that l is not 1 modulo 11, so that the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{Aut}(V_l)$ is all of $\mathrm{Aut}(V_l)$ (at the end of the proof we will indicate how the l that are 1 mod 11 can be treated).

We recall that K_l denotes the Galois extension of $\mathbb{Q}$ given by the mod l representation associated to Δ. But now, under the hypothesis that that representation is surjective to $\mathrm{Aut}(V_l)$, K_l is the $\mathbb{Q}$-algebra corresponding to the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-set $\mathrm{Isom}(\mathbb{F}_l^2, V_l)$. As explained before, we have two maps from $\mathrm{Isom}(\mathbb{F}_l^2, V_l)$ to $V_l - \{0\}$, sending ϕ to $\phi(1,0)$ and to $\phi(0,1)$, respectively. Clearly, the map from $\mathrm{Isom}(\mathbb{F}_l^2, V_l)$ to $(V_l - \{0\})^2$ that they give together is injective. This means that the corresponding algebra morphism from $K_l' \otimes K_l'$ to K_l is surjective. We let t denote our generator of K_l'. Then the elements t_1 and t_2 of K_l defined by:

$$t_1, t_2 \colon \mathrm{Isom}(\mathbb{F}_l^2, V_l) \longrightarrow \overline{\mathbb{Q}}, \quad t_1 \colon \phi \mapsto t(\phi(1,0)), \quad t_2 \colon \phi \mapsto t(\phi(0,1))$$

generate K_l. We can now compute the minimal polynomial of t_2 over the subfield $\mathbb{Q}(t_1)$ of K_l, by the method already used twice above, and then the $t_1^i t_2^j$ with $0 \leq i < l^2 - 1$ and $0 \leq j < l^2 - l$ form a $\mathbb{Q}$-basis of K_l.

Finally, we can compute the action of $\mathrm{GL}_2(\mathbb{F}_l)$. This action is given as follows. For an element g in $\mathrm{GL}_2(\mathbb{F}_l)$ and an element h of K_l we have, for all ϕ in $\mathrm{Isom}(\mathbb{F}_l^2, V_l)$:

$$(gh)\phi = h(\phi \circ g).$$

For such a g, the coefficients of gt_1 and gt_2 with respect to our basis can be found by computing them at many residue fields, as before.

22.6 Computing K_l when $l \equiv 1 \bmod 11$

In this case, the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{Aut}(V_l)$ is the subgroup of index 11 of elements whose determinant is an 11th power in $\mathbb{F}_l^\times$. Johan Bosman explained that by twisting we can reduce again to the case where the representation is surjective.

We let $\chi_l \colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{F}_l^\times$ denote the character giving the action on $\mu_l(\overline{\mathbb{Q}})$. For i in $\mathbb{Z}/(l-1)\mathbb{Z}$, we let $\mathbb{F}_l(i)$ denote $\mathbb{F}_l$ with $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acting on it by χ_l^i, and we let $V_l(i)$ denote the twist $V_l \otimes \mathbb{F}_l(i)$ of V_l. The determinant of $V_l(i)$ is χ_l^{11+2i}. In particular, $V_l(-5)$ has determinant χ_l, hence the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\mathrm{Aut}(V_l(-5))$ is all of $\mathrm{Aut}(V_l(-5))$.

The $\mathbb{Q}$-algebra $K_l(-5)$ associated to $\mathrm{Isom}(\mathbb{F}_l^2, V_l(-5))$ can be computed by the same methods as used above for $V_l - \{0\}$. One just changes the action of $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q})$ on the $\mathbb{Q}(\zeta)$-algebra $K'_{l,\mathbb{Q}(\zeta_l)}$ as follows. We write $K'_{l,\mathbb{Q}(\zeta_l)} = \mathbb{Q}(\zeta_l) \otimes K_l'$. Then σ in $\mathrm{Gal}(\mathbb{Q}(\zeta_l)/\mathbb{Q})$ acts as

$(\sigma \otimes \mathrm{id}) \circ \langle \chi_l(\sigma)^{-5} \rangle$. In other words, one just descends $K'_{l,\mathbb{Q}(\zeta_l)}$ from $\mathbb{Q}(\zeta_l)$ to $\mathbb{Q}$ in another way than before. We have now finished the proof of Theorem 22.1. □

22.7 Remark The algorithm given above is probabilistic. It is almost certain that replacing Couveignes finite field approach by his complex approach as in [13] will give a deterministic algorithm; the problem for the moment is that [13] only deals with the modular curves $X_0(l)$ for l prime.

22.8 Remark The operation of twisting the representation V_l by a power of the mod l cyclotomic character can probably be profitably used to make the image of the representation smaller, namely, of order $O(l^3)$. It needs to be seen if this is useful.

22.9 Remark The algorithm in Theorem 22.1 does not really give what was promised in Section 11: the minimal polynomial of a generator and the Galois action in terms of matrices with respect to the $\mathbb{Q}$-basis given by that generator. However, standard algorithms can produce such a generator, its minimal polynomial and the matrices giving the Galois action in polynomial time; see Section 2.9 of [74].

23 Computing $\tau(p)$ in time polynomial in $\log p$

We recall that Ramanujan's τ-function is defined by the following identity of formal power series with integer coefficients:

$$x \prod_{n \geq 1} (1 - x^n)^{24} = \sum_{n \geq 1} \tau(n) x^n.$$

23.1 Theorem *There exists a probabilistic algorithm that on input a prime number p gives $\tau(p)$, in expected running time polynomial in $\log p$.*

Proof Deligne has proved that for prime numbers p we have $|\tau(p)| < 2p^{11/2}$. Therefore, it suffices to compute $\tau(p)$ mod l for all primes $l < x$, if the product of these l is at least $4p^{11/2}$. Analytic number theory tells us that we can take $x = O(\log p)$, hence the proof is reduced to showing that there is a probabilistic algorithm that computes $\tau(p)$ mod l for prime numbers p and l in expected time polynomial in $\log p$ and l. (Of course, the slightly weaker but much more elementary bound in [82, Cor. 2.1.6] also suffices here for our purposes.)

We fix an algorithm as in Theorem 22.1. We must now show that from the output of that algorithm we can compute $\tau(p)$ mod l in time polynomial in $\log p$ and l. Theorem 1.3 of [7] gives the existence of a polynomial time algorithm that produces an order A in O_{K_l} that is maximal

at p (orders are given by specifying a $\mathbb{Z}$-basis of it). Then we replace A by the order generated by all σA, for σ in $\mathrm{Gal}(K_l/\mathbb{Q})$. This new order A is preserved by the action of $\mathrm{Gal}(K_l/\mathbb{Q})$ and we can compute it in polynomial time. Then $\mathrm{Gal}(K_l/\mathbb{Q})$ acts on the étale $\mathbb{F}_p$-algebra $\overline{A} := A/pA$, and $\mathrm{Hom}(\overline{A}, \overline{\mathbb{F}}_p)$ is a $\mathrm{Gal}(K_l/\mathbb{Q})$-torsor. Moreover, $\overline{A}$ is the product of its finitely many residue fields $\overline{A}/m$, where m ranges through the maximal ideals of $\overline{A}$.

We let Frob denote the absolute Frobenius endomorphism of $\overline{A}$; it sends a to a^p, it is an automorphism and it induces the absolute Frobenius automorphism on each of the residue fields. The matrix of Frob can be computed in polynomial time.

The Frobenius element σ_m associated to a maximal ideal m of $\overline{A}$ is the unique element σ of $\mathrm{Gal}(K_l/\mathbb{Q})$ that fixes m and induces the absolute Frobenius on $\overline{A}/m$. For varying m, the σ_m form the Frobenius conjugacy class (at p) in $\mathrm{Gal}(K_l/\mathbb{Q})$.

For each σ in $\mathrm{Gal}(K_l/\mathbb{Q})$ we let $\overline{A}_\sigma$ be the quotient of $\overline{A}$ by the ideal generated by the image of $\mathrm{Frob} - \sigma$. Such an $\overline{A}_\sigma$ can be computed in polynomial time. The σ in the Frobenius conjugacy class are precisely those σ for which $\overline{A}_\sigma$ is non-zero. We can try the σ one by one until we have found a σ in the Frobenius conjugacy class. Then we apply the map ρ_l given by Theorem 22.1 to it, and get an element $\rho_l(\sigma)$ of $\mathrm{GL}_2(\mathbb{F}_l)$. The trace of $\rho_l(\sigma)$ is then $\tau(p) \bmod l$. □

23.2 Remark Theorem 23.1 above proves the existence of a probabilistic algorithm with polynomial expected running time. The proof shows in fact that if Theorem 22.1 can be strengthened to give the existence of a deterministic algorithm, then in Theorem 23.1 we can replace the word "probabilistic" by "deterministic". I thank Hendrik Lenstra for the observation that one does not need to factor a polynomial over $\mathbb{F}_p$ in order to find the Frobenius conjugacy class, which would have made the final step probabilistic given the present state of affairs.

24 Some explicit examples, by Johan Bosman

This section has been written by Johan Bosman, in March 2006.

In this section we will show some explicit calculations that were performed. It must be emphasised that the calculations we did are not at all rigorous, i.e. we do not know a rigorous proof that the results are correct. The point is that rigorous methods so far only work in theory and are still too complex to be performed on actual computers. And conversely, of methods that turn out to work pretty well on computers by experimentation, we do not know how to prove some important statements concerning the complexity and the correctness. All the calculations here concern the mod l representations for the Ramanujan τ function for several prime numbers $l \geq 13$ (the case $l = 11$ is easy, as $J_1(11)$ is an elliptic curve).

Let g be the genus of the modular curve $X_1(l)$. As described in Section 11 we will approximate l-torsion points in $J_1(l)$. Because l is prime there are no nonzero oldforms in $S_2(\Gamma_1(l))$, so the space $S_2(\Gamma_1(l))$ has a basis $f_1, \ldots, f_g$ consisting of newforms. One can calculate the coefficients of the q-expansion of a newform efficiently using modular symbols, see [104]. One can also find methods there how to calculate the period lattice $\Lambda \subset \mathbb{C}^g$, defined as:

$$\Lambda := \left\{ \int_\gamma (f_1, \ldots, f_g) \frac{dq}{q} : [\gamma] \in \mathrm{H}_1(X_1(l)(\mathbb{C}), \mathbb{Z}) \right\}.$$

Once we have calculated $\int_\gamma (f_1, \ldots, f_g) \frac{dq}{q}$ along a basis for $\mathrm{H}_1(X_1(l), \mathbb{Z})$ we can immediately express the any integral of a newform over a path between two cusps in terms of the integrals over this basis. This is because the Manin-Drinfeld theorem tells us that linear equivalence classes cuspidal divisors of degree 0 are torsion elements of $J_1(l)$, hence we can identify them with elements of $\mathrm{H}_1(X_1(l), \mathbb{Q})$. The corresponding $\mathbb{Q}$-coefficients of such an element of $\mathrm{H}_1(X_1(l), \mathbb{Q})$ with respect to a given basis can be calculated using modular symbols.

We want to be able to calculate the Abel-Jacobi integration map $J_1(l)(\mathbb{C}) \to \mathbb{C}^g/\Lambda$. For this we need to calculate $\int_P^Q (f_1, \ldots, f_g) dq/q \in \mathbb{C}^g/\Lambda$ for any two points $P, Q \in X_1(l)$. So in fact we need to calculate $\int_\infty^z f dq/q$ for $z \in \mathbb{H}$ and $f \in S_2(\Gamma_1(l))$ a newform. Choose a fundamental domain for $\Gamma_0(l)$ in $\mathbb{H}$ which has only the two cusps ∞ and 0 and write $z = \gamma z'$ with $\gamma \in \Gamma_0(l)$ and z' in this fundamental domain. Then:

$$\int_\infty^{\gamma z'} f \frac{dq}{q} = \int_\infty^{\gamma\infty} f \frac{dq}{q} + \int_{\gamma\infty}^{\gamma z'} f \frac{dq}{q} = \int_\infty^{\gamma\infty} f \frac{dq}{q} + \varepsilon(\gamma) \int_\infty^{z'} f \frac{dq}{q},$$

where ε is the Nebentypus of f. The first term here is an integral over a path between cusps, so this can be performed. We can use an Atkin-Lehner operator on $S_2(\Gamma_1(l))$ now, if necessary. It is defined as $W_l(f) := f|\left(\begin{smallmatrix} 0 & -1 \\ l & 0 \end{smallmatrix}\right)$. The transformation $z \mapsto -1/lz$ on $X_1(l)$ is also denoted by W_l. By our choice of fundamental domain, $\max\{\Im z', \Im(-1/lz')\} \geq \sqrt{3}/2l$. So either we can calculate the right term integral directly by a formal integration of q-series or we do it by:

$$\int_\infty^{z'} f \frac{dq}{q} = \int_\infty^0 f \frac{dq}{q} + \int_0^{z'} f \frac{dq}{q} = \int_\infty^0 f \frac{dq}{q} + \int_\infty^{-1/lz'} W_l(f) \frac{dq}{q}$$

and then formally integrating the q-series of $W_l(f) dq/q$. In conclusion, once we calculated the period integrals, we can calculate $\int_\infty^z f dq/q$ to a precision of p decimals using about $\frac{pl \log 10}{\sqrt{3}\pi} \approx 0.42pl$ terms of the q-expansion of f. With similar methods, one can evaluate $f(z)$ to a precision of p decimals using about $\frac{pl \log 10}{\sqrt{3}\pi} \approx 0.42pl$ terms of the q-expansion as well.

Now, following the strategy of section 11, we pick a divisor D of degree g on $X_1(l)$ which is defined over $\mathbb{Q}$ and consider the map:

$$\phi' \colon X_1(l)^g \to J_1(l) : (Q_1, \ldots, Q_g) \mapsto [Q_1 + \cdots + Q_g - D].$$

Since the cusp 0 on $X_1(l)$ is defined over $\mathbb{Q}$, we simply take $D = g \cdot (0)$. The map ϕ factors as $X_1(l)^g \to X_1(l)^{(g)} \to J_1(l)$. The map $X_1(l)^{(g)} \to J_1(l)$ is an isomorphism outside a closed subset of codimension 1. It is highly unlikely that preimages of some l-torsion points of $J_1(l)$ lie in this set. So we just assume that this is not the case, without knowing how to prove it. Identifying $J_1(l)$ with $\mathbb{C}^g/\Lambda$, the map ϕ' becomes

$$\phi : X_1(l)^g \to \mathbb{C}^g/\Lambda : (Q_1, \ldots, Q_g) \mapsto \sum_{i=1}^{g} \int_0^{Q_i} (f_1, \ldots, f_g) \frac{dq}{q} \mod \Lambda.$$

The idea is now, given a point in $P \in \mathbb{C}^g/\Lambda$, of which we assume that it is not a branch point of the map ϕ, and which we have calculated to a high precision, we want to calculate a point $Q = (Q_1, \ldots, Q_g) \in X_1(l)^g$ mapping to P, also to a high precision.

We write all our points in $X_1(l)(\mathbb{C})$ as γz with $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $z \in F$, where:

$$F = \left\{ z \in \mathbb{H} : |z| \geq 1 \text{ and } |\Re z| \leq \frac{1}{2} \right\}.$$

This works well because of the transformation behaviour of a modular form. The approximation that we use here consists of two steps. First of all, we want to find a point Q for which $\phi(Q)$ is very close to P. Then, we use a Newton-Raphson process that roughly doubles the accuracy at each step. This Newton-Raphson process is very easy to perform since the partial derivatives of the Abel-Jacobi integrals can be calculated in terms of modular forms directly.

The first approximation is done with a low calculation precision, because of speed. We pick a number of random points in $X_1(l)^g$ and map them to $\mathbb{C}/\Lambda$ and we proceed with the point which maps closest to P. So we start with a point $Q = (Q_1, \ldots, Q_g) = (\gamma_1 z_1, \ldots, \gamma_g z_g)$ and then make small adjustments to the z_i so that $\phi(Q)$ gets closer to P. We have that:

$$\phi(\gamma_1(z_1 + h_1), \ldots, \gamma_g(z_g + h_g)) = \phi(\gamma_1 z_1, \ldots, \gamma_g z_g) + \left(\frac{\partial \phi_i(\gamma_1 z_1, \ldots, \gamma_g z_g)}{\partial z_j} \right)_{i,j} \cdot h + O(\|h\|^2).$$

Now choose a point P' on the segment between P and $\phi(Q)$, very close to $\phi(Q)$. Let v be a smallest representative of $P' - \phi(Q)$ in $\mathbb{C}^g$ and put:

$$h = \left(\frac{\partial \phi_i(\gamma_1 z_1, \ldots, \gamma_g z_g)}{\partial z_j} \right)_{i,j}^{-1} \cdot v.$$

then for $Q' = (\gamma_1(z_1 + h_1), \ldots, \gamma_g(z_g + h_g))$, the point $\phi(Q')$ is very close to P'. This means that it is highly likely that $\phi(Q')$ is closer to P than $\phi(Q)$ is. If we repeat this process we can often get very close to P after several steps. It could be that this process fails because we are in $X_1(l)^g$ close to points where the Jacobian matrix of the map ϕ is singular. In that case we try another random point at the beginning.

Inside $J_1(l)[l]$ there is a 2-dimensional $\mathbb{F}_l$-subspace V_l defined over $\mathbb{Q}$ such that the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on it gives the Galois representation for the Ramanujan τ function, see Section 10. We assume here that the map ϕ is étale over $V_l - \{0\}$ hence in particular that the preimage of each point of $V_l - \{0\}$ in $X_1(l)^{(g)}$ consists of exactly one point. Let us denote, by abuse of notation, the preimage of V_l in $X_1(l)^{(g)}$ also by V_l.

Finding the torsion points in $\mathbb{C}^g/\Lambda$ that we have to approximate is not so difficult. We can identify Λ with $\mathrm{H}_1(X_1(l), \mathbb{Z})$ and we can use modular symbols algorithms to calculate Hecke operators on $\mathrm{H}_1(X_1(l), \mathbb{Z})$, hence also on $l^{-1}\mathrm{H}_1(X_1(l), \mathbb{Z})/\mathrm{H}_1(X_1(l), \mathbb{Z}) \cong J_1(l)(\mathbb{C})[l]$. By Theorem 10.9, we can thus calculate which points of $J_1(l)(\mathbb{C})[l]$ are relevant for the mod l Galois representation of the Ramanujan τ function.

Now that we have calculated the points of $V_l - \{0\} \subset X_1(l)^{(g)}$ to a high precision, we want to choose a function $f \in \mathbb{Q}(X_1(l))$ such that the height of $k(Q) = f(Q_1) + \cdots + f(Q_g)$ is low for $Q \in \phi^{-1}(V_l - \{0\})$. In theory it could happen that $k(Q)$ does not generate the field of definition of the corresponding point in $J_1(l)[l]$, but in practise the chance that this happens is extremely small.

So how to choose f? The strategy that we use here is to let f be a quotient $W_l(f_1)/W_l(f_2)$ of two elements of $S_2(\Gamma_1(l))$. We apply the Atkin-Lehner transformation W_l here because we focus on the cusp 0 instead of the cusp ∞. This transformation has the nice property that it is an isomorphism between $X_1(l)_{\mathbb{Q}}$ and $X_\mu(l)_{\mathbb{Q}}$, defined over $\mathbb{Q}$. The space $X_\mu(l)$ here denotes the compactification of the moduli space of elliptic curves with a given embedding of the group scheme μ_l. In particular the cusp ∞ is rational as a point of $X_\mu(l)_{\mathbb{Q}}$. To make the height of $k(Q)$ small it is wise to let f have a pole in the cusp 0, and furthermore to let the degree of $f\colon X_1(l) \to \mathbb{P}^1_{\mathbb{Q}}$ be small. We also want the q-series of f around the cusp 0 to have integer coefficients and the coefficient of the lowest term of this q-expansion to be equal to 1, otherwise this gives us some unnecessary prime contributions in the height.

This can be achieved by writing down a basis for $S_2(\Gamma_1(l))$ which consists of forms having integer q-coefficients around ∞. The forms f_1 and f_2 will then be constructed by choosing suitable linear combinations having a zero of high degree at the cusp ∞. Since $S_2(\Gamma_1(l))$ is isomorphic to $\mathrm{H}^0(X_1(l)_{\mathbb{C}}, \Omega^1_{X_1(l)})$ the divisors on $X_1(l)$ of the elements of $S_2(\Gamma_1(l))$ are effective of degree $2g-2$ so this way we find a function of degree at most g.

We have chosen l so that the image of $\overline{\rho}_{\Delta,l}\colon \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_l)$ contains $\mathrm{SL}_2(\mathbb{F}_l)$, hence $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts transitively on $V_l - \{0\}$. Furthermore, since we are assuming that $k(Q)$ generates the field of definition of the corresponding point of $J_1(l)$, it follows that:

$$P_l(x) = \prod_{Q \in V_l - \{0\}} (x - k(Q)) \in \mathbb{Q}[x]$$

is a polynomial for the field of definition of a point of $V_l - \{0\}$. The group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ also acts on the set of lines in V_l and this gives us a homomorphism $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}(V_l) \cong \mathrm{PGL}_2(\mathbb{F}_l)$, which is the same as the composition of ρ_l with the projection $\mathrm{GL}(V_l) \to \mathrm{PGL}(V_l)$. The field of definition of a line in V_l is equal to the field defined by the polynomial:

$$P_l'(x) = \prod_{L \in \mathbb{P}(V_l)} (x - \sum_{Q \in L - \{0\}} k(Q)).$$

We can compute the polynomials P_l' and P_l numerically, and from this we want to find the actual rational coefficients that they have. It is too hard to calculate to the precision that the theory proves to be sufficient. Luckily, in practise the height is much smaller than the theoretical upper bounds given, the only problem is that we cannot rigorously prove this. Using continued fractions, we can find approximations of the coefficients by rational numbers. If a is a rational number of bounded height, then it is highly unlikely that there exist coprime integers p and q such that $p/q \neq a$ but $|p/q - a|$ is much smaller than $1/q^2$. So if a is some coefficient of P_l or P_l' and we find integers p and q with $|p/q - a| \ll 1/q^2$ then we have quite some evidence that the guess $a = p/q$ is correct. The polynomial P_l is the minimal polynomial of an algebraic number of bounded height. It is thus quite natural to expect that the coefficients of P_l have a relatively small common denominator. So if the denominators of our guesses p/q for the coefficients are equal or differ only by a small factor this is even more evidence for the correctness of these guesses, since this is not very likely to happen for a random polynomial with rational coefficients of bounded height.

Let us now show some explicit results. We begin with $l = 13$. The calculations shown here were performed using MAGMA on a Xeon CPU with 2.4 Ghz of speed and 2 GB of memory. The genus of $X_1(13)$ is equal to 2 so there are 2 newforms f_1 and f_2 in $S_2(\Gamma_1(13))$. They are the complex conjugates of the form whose q-expansion starts with:

$$q - (\zeta_3 + 2)q^2 + 2\zeta_3 q^3 + (\zeta_3 + 1)q^4 + \cdots,$$

so if we put $\zeta_3 = \exp(2\pi i/3)$ then we set $f_1 = q - (\zeta_3 + 2)q^2 + 2\zeta_3 q^3 + (\zeta_3 + 1)q^4 + \cdots$ and $f_2 = q - (\overline{\zeta_3} + 2)q^2 + 2\overline{\zeta_3} q^3 + (\overline{\zeta_3} + 1)q^4 + \cdots$. The period lattice $\Lambda \subset \mathbb{C}^2$ can be calculated to a precision of 200 decimals within a few seconds. Shown to a smaller precision here, it is approximately equal to:

$$\begin{aligned}\Lambda = &(1.718687 + 2.225030i, -1.718687 + 2.225030i)\mathbb{Z} \\ &\oplus (2.786276 - 0.375912i, -2.786276 - 0.375912i)\mathbb{Z} \\ &\oplus (1.438957 - 0.194137i, 4.225233 + 0.570049i)\mathbb{Z} \\ &\oplus (0.551351 - 1.343242i, 1.618940 + 3.944184i)\mathbb{Z}.\end{aligned}$$

Approximation to 200 decimals of the 168 points in $X_1(13)^{(2)}$ mapping to $V_{13} - \{0\}$ was performed in less than half an hour.

As function $f : X_1(13)_{\mathbb{Q}} \to \mathbb{P}^1_{\mathbb{Q}}$ we choose:

$$f = W_{13}\left(3\frac{(\zeta_3+1)f_1 + (\overline{\zeta_3}+1)f_2}{(2\zeta_3+1)f_1 + (2\overline{\zeta_3}+1)f_2}\right) = W_{13}\left(q^{-1} + 2 + q + \cdots\right).$$

If we calculate P_{13} using this function f, using our heuristics about finding the rational coefficients, then we find a remarkably small common denominator for the coefficients, namely $2535853 = 251 \cdot 10103$. Although, we have no proof that this number is correct, let us assume that it is. If we multiply P_{13} with this number, then we get a polynomial in $\mathbb{Z}[x]$ whose largest coefficient has 89 digits and whose constant coefficient has 80 digits. Compared to the theoretical upper bounds, this is very small, but keeping in mind that $l = 13$ should be a very simple case, these numbers are actually very large.

The polynomial P'_{13} that belongs to the $\mathrm{PGL}_2(\mathbb{F}_{13})$-extension is a bit smaller and looks as follows.

$$\begin{aligned}
2535853P'_{13} = {} & 2535853x^{14} - 127713190x^{13} - 9947603692x^{12} + 795085450224x^{11} \\
& - 29425303073920x^{10} + 667684302673440x^9 - 9974188441308416x^8 \\
& + 106364914419352576x^7 - 1012336515218109952x^6 \\
& + 9094902359324720640x^5 - 60847891441699468288x^4 \\
& + 324814691085008943104x^3 - 1761495929112889016320x^2 \\
& + 6235371687080448827392x - 10767442738728520761344.
\end{aligned}$$

However, there exists a polynomial with smaller coefficients defining the same number field, which we can obtain by first calculating the ring of integers and then applying a lattice reduction technique:

$$\begin{aligned}
& x^{14} + 7x^{13} + 26x^{12} + 78x^{11} + 169x^{10} + 52x^9 - 702x^8 - 1248x^7 \\
& + 494x^6 + 2561x^5 + 312x^4 - 2223x^3 + 169x^2 + 506x - 215,
\end{aligned}$$

which looks already a lot nicer.

For $l = 17$, the calculations took already a few days and the results are a lot uglier. The polynomial P_l that came out has a common denominator of 182 decimals and the largest coefficient has 393 digits. So the required precision to obtain this conjectural result is about 400 digits.

For $l = 19$ we distributed the calculation over several machines and it still took a couple of months. The common denominator of P_{19} has 375 digits and the largest coefficient has 822 digits, so a precision of about 830 digits is required to get this.

Several sanity checks can be performed. First of all we can of course check whether the discriminants of P_l and P_l' are powers of l times a square. We can also check whether the cycle structure of $\overline{\rho}(\mathrm{Frob}_p)$ coincides with the decomposition structure of P_l modulo p. This all turns out to be correct.

References

[1] A. Abbes and E. Ullmo. *Comparaison des métriques d'Arakelov et de Poincaré sur $X_0(N)$.* Duke Math. J. 80 (1995), no. 2, 295–307.

[2] L.M. Adleman and M-D. Huang. *Counting points on curves and abelian varieties over finite fields.* J. Symbolic Comput. 32 (2001), no. 3, 171–189.

[3] S.Y. Arakelov, *An intersection theory for divisors on an arithmetic surface*, Math. USSR Izvestija **8** (1974), 1167–1180.

[4] P. Bayer and J. Neukirch. *On automorphic forms and Hodge theory.* Math. Ann. 257 (1981), no. 2, 137–155.

[5] J-F. Boutot and H. Carayol. *Uniformisation p-adique des courbes de Shimura: les théorèmes de Cerednik et de Drinfeld.* Courbes modulaires et courbes de Shimura (Orsay, 1987/1988). Astérisque No. 196-197, (1991), 7, 45–158 (1992).

[6] S. Bosch, W. Lütkebohmert and M. Raynaud. *Néron models.* Springer Verlag, Ergebnisse 3, 21 (1990).

[7] J.A. Buchmann and H.W. Lenstra. *Approximating rings of integers in number fields.* J. Théor. Nombres Bordeaux 6 (1994), no. 2, 221–260.

[8] H. Carayol. *Sur les représentations l-adiques associées aux formes modulaires de Hilbert.* Annales Scientifiques de l'École Normale Supérieure Sér. 4, 19 no. 3 (1986), p. 409-468.

[9] *Algebraic number theory.* Proceedings of the instructional conference held at the University of Sussex, Brighton, September 1–17, 1965. Edited by J. W. S. Cassels and A. Fröhlich. Reprint of the 1967 original. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1986. xviii+366 pp. ISBN: 0-12-163251-2.

[10] R.F. Coleman and S.J. Edixhoven. *On the semi-simplicity of the U_p-operator on modular forms.* Math. Ann. 310 (1998), no. 1, 119–127.

[11] B. Conrad. *Modular forms and the Ramanujan conjecture.* Book in preparation. See: http://www.math.lsa.umich.edu/~bdconrad/

[12] G. Cornell, J.H. Silverman, *Arithmetic geometry*, Springer 1986.

[13] J-M. Couveignes. *Jacobiens, jacobiennes et stabilité numérique.* Preprint, available on the author's home page: http://www.univ-tlse2.fr/grimm/couveignes/

[14] J-M. Couveignes, S.J. Edixhoven and R. de Jong. Article in preparation.

[15] J.E. Cremona. *Algorithms for modular elliptic curves.* Cambridge University Press, 1997.

[16] C.W. Curtis and I. Reiner. *Representation theory of finite groups and associative algebras.* Pure and Applied Mathematics, Vol. XI Interscience Publishers, a division of John Wiley & Sons, New York-London 1962.

[17] S. David. *Fonctions thêta et points de torsion des variétés abéliennes.* Compositio Math. 78 (1991), no. 2, 121–160.

[18] S. David and P. Philippon. *Minorations des hauteurs normalisées des sous-variétés de variétés abeliennes. II.* Comment. Math. Helv. 77 (2002), no. 4, 639–700.

[19] P. Deligne. *Formes modulaires et représentations l-adiques.* Séminaire Bourbaki, 355, Février 1969.

[20] P. Deligne *La conjecture de Weil. I.* Inst. Hautes Études Sci. Publ. Math. No. 43 (1974), 273–307.

[21] P. Deligne and M. Rapoport. *Les schémas de modules des courbes elliptiques.* In Modular Functions of One Variable II. Springer Lecture Notes in Mathematics 349 (1973).

[22] P. Deligne and J-P. Serre. *Formes modulaires de poids* 1. Ann. Sci. École Norm. Sup. (4) 7 (1974), 507–530.

[23] M. Dickinson. *On the modularity of certain 2-adic Galois representations.* Duke Math. J. 109, no. 2 (2001), 319–382.

[24] L.E. Dickson. *Linear groups: With an exposition of the Galois field theory.* With an introduction by W. Magnus Dover Publications, Inc., New York 1958.

[25] P. Deligne. *Cohomologie étale.* Séminaire de Géométrie Algébrique du Bois-Marie SGA $4\frac{1}{2}$. Avec la collaboration de J. F. Boutot, A. Grothendieck, L. Illusie et J. L. Verdier. Lecture Notes in Mathematics, Vol. 569. Springer-Verlag, Berlin-New York, 1977.

[26] P. Deligne. *Groupes de monodromie en géométrie algébrique. II.* Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II). Dirigé par P. Deligne et N. Katz. Lecture Notes in Mathematics, Vol. 340. Springer-Verlag, Berlin-New York, 1973.

[27] J. Denef and F. Vercauteren. *Computing zeta functions of hyper-elliptic curves over finite fields of characteristic 2.* Advances in cryptology—CRYPTO 2002, 369–384, Lecture Notes in Comput. Sci., 2442, Springer, Berlin, 2002.

[28] F. Diamond and J. Im. *Modular forms and modular curves.* Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994), 39–133, CMS Conf. Proc., 17, Amer. Math. Soc., Providence, RI, 1995.

[29] F. Diamond and J. Shurman. *A first course in modular forms.* GTM 228, Springer-Verlag, 2005.

[30] L.E. Dickson. *Linear groups: With an exposition of the Galois field theory.* With an introduction by W. Magnus. Dover Publications, Inc., New York 1958.

[31] V.G. Drinfeld. *Two theorems on modular curves.*Funkcional. Anal. i Priložen. 7 (1973), no. 2, 83–84.

[32] S.J. Edixhoven. *The weight in Serre's conjectures on modular forms.* Inventiones Mathematicae **109**, 563–594 (1992).

[33] S.J. Edixhoven. *The modular curves $X_0(N)$.* Series of lectures at the ICTP Summer school on rational torsion of elliptic curves over number fields, held from 11–29 August, 1997. Available at:
`www.math.leidenuniv.nl/~edix/public_html_rennes/cours/trieste.html`

[34] S.J. Edixhoven. *Rational elliptic curves are modular (after Breuil, Conrad, Diamond and Taylor).* Séminaire Bourbaki, Vol. 1999/2000. Astérisque No. 276 (2002), 161–188.

[35] *On the computation of coefficients of modular forms.* Workshop Computational Arithmetic Geometry, MSRI, Berkeley, December 2000. Notes and streaming video available at the website of the MSRI.

[36] S.J. Edixhoven. *Point counting after Kedlaya.* Syllabus for the EIDMA-Stieltjes Graduate course "Mathematics of Cryptology", at the Lorentz Center in Leiden, September 2003. Available at:
`www.math.leidenuniv.nl/~edix/oww/mathofcrypt/carls_edixhoven/kedlaya.pdf`

[37] S.J. Edixhoven and J-H. Evertse *Diophantine approximation and abelian varieties*, Lecture Notes in Mathematics 1566 (Edixhoven and Evertse, eds.), Springer-Verlag (1993, 2nd printing 1997).

[38] A. Enge. *Elliptic curves and their applications to cryptography, an introduction.* Kluwer Academic Publishers, 1999. **— N° 844**.

[39] G. Faltings, *Calculus on arithmetic surfaces*, Ann. of Math. **119** (1984), 387–424.

[40] G. Faltings. *Lectures on the arithmetic Riemann-Roch theorem.* Notes taken by Shouwu Zhang. Annals of Mathematics Studies, 127. Princeton University Press, 1992.

[41] G. Faltings and B.W. Jordan. *Crystalline cohomology and* $\mathrm{GL}(2, Q)$. Israel J. Math. 90 (1995), no. 1-3, 1–66.

[42] T. Fisher. *On 5 and 7 descents for elliptic curves.* Available on the author's home page: `http://www.dpmms.cam.ac.uk/~taf1000/thesis.html`

[Fo] O. Forster, *Riemannsche Flächen*. Springer-Verlag, Berlin (1977).

[43] M. Fouquet, P. Gaudry and R. Harley. *An extension of Satoh's algorithm and its implementation.* J. Ramanujan Math. Soc. **15** (2000), no. 4, 281–318.

[44] E. Freitag and R. Kiehl. *Étale cohomology and the Weil conjecture.* Translated from the German by Betty S. Waterhouse and William C. Waterhouse. With an historical introduction by J. A. Dieudonné. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 13. Springer-Verlag, Berlin, 1988.

[45] G. Frey and M. Müller. *Arithmetic of modular curves and applications.* In "On Artin's conjecture for odd 2-dimensional representations", number 1585 in Lecture Notes in Math. Springer, 1994.

[46] G. Frey and H-G. Rück. *A remark concerning* m*-divisibility and the discrete logarithm in the divisor class group of curves.* Mathematics of computation, 62(206):865–874, 1994.

[47] S.D. Galbraith, S. Paulus, and N.P. Smart. *Arithmetic of superelliptic curves.* Mathematics of computation, 71(237):393–405, 2002.

[48] P. Gaudry and N. Gürel. *An extension of Kedlaya's point-counting algorithm to superelliptic curves.* In C. Boyd (ed.), Advances in Cryptology — ASIACRYPT 2001, Lecture Notes in Computer Science 1807, Springer-Verlag (2000), 19–34.

[49] D. Gorenstein. *An arithmetic theory of adjoint plane curves.* Trans. Amer. Math. Soc., 72:414–436, 1952.

[50] B.H. Gross. *A tameness criterion for Galois representations associated to modular forms (mod p).* Duke Mathematical Journal 61, No. 2, (1990).

[51] A. Grothendieck. *Théorie des topos et cohomologie étale des schémas.* Séminaire de Géométrie Algébrique du Bois-Marie 1963–1964 (SGA 4). Dirigé par M. Artin, A. Grothendieck, et J. L. Verdier. Avec la collaboration de N. Bourbaki, P. Deligne et B. Saint-Donat. Lecture Notes in Mathematics, Volumes 269, 270 and 305. Springer-Verlag, Berlin-New York, 1972 and 1973.

[52] Gaétan Haché. *Computation in algebraic function fields for effective construction of algebraic-geometric codes.* In "Proceedings of the 11th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes", pages 262–278. 1995.

[53] R. Hartshorne. *Algebraic geometry.* Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977.

[54] P. Hriljac, *Heights and Arakelov's intersection theory.* Amer. J. Math. **107**, 1 (1985), 193–218.

[55] H. Hubrechts. *Point counting in families of hyperelliptic curves.* Available on arxiv: math.NT/0601438.

[56] D. Husemoller. *Elliptic curves.* Springer, 1987.

[57] J. I. Igusa, *Theta functions*. Grundlehren der Math. Wissenschaften **194**, Springer-Verlag, Berlin (1972).

[58] L. Illusie. Cohomologie l-adique et fonctions L. (French) Séminaire de Géometrie Algébrique du Bois-Marie 1965–1966 (SGA 5). Edité par Luc Illusie. Lecture Notes in Mathematics, Vol. 589. Springer-Verlag, Berlin-New York, 1977.

[59] J. Jorgenson and J. Kramer. *Bounds on canonical Green's functions.* To appear in Compositio Mathematica.

[60] G. Kato and S. Lubkin. *Zeta matrices of elliptic curves.* Journal of Number Theory **15**, 318–330 (1982).

[61] N.M. Katz, B. Mazur. *Arithmetic moduli of elliptic curves.* Annals of Mathematics Studies 108, Princeton University Press (1985).

[62] K. Kedlaya. *Counting points on hyper-elliptic curves using Monsky-Washnitzer cohomology.* J. Ramanujan Math. Soc. **16** (2001), no. 4, 323–338.

[63] I. Kiming and H. Verrill. *On modular mod l Galois representations with exceptional images.* J. Number Theory 110 (2005), no. 2, 236–266.

[64] S. Landau. *Factoring polynomials over algebraic number fields.* SIAM J. Comput. 14 (1985), no. 1, 184–195.

[65] S. Lang. *Algebraic number theory.* Second edition. Graduate Texts in Mathematics, 110. Springer-Verlag, New York, 1994. xiv+357 pp. ISBN: 0-387-94225-4.

[66] Serge Lang. *Abelian varieties*, volume 7 of "Interscience Tracts in Pure and Applied Mathematics". Interscience Publishers, 1959. **— N° 751**.

[67] J-C. Lario and R. Schoof. *Some computations with Hecke rings and deformation rings.* With an appendix by Amod Agashe and William Stein. Experiment. Math. 11 (2002), no. 2, 303–311.

[68] A.G.B. Lauder. *Computing zeta functions of Kummer curves via multiplicative characters.* Preprint, 2002,
Available at: `http://web.comlab.ox.ac.uk/oucl/work/alan.lauder`

[69] A.G.B. Lauder. *Deformation theory and the computation of zeta functions.* Proceedings of the London Mathematical Society, Vol. 88 Part 3, (2004), 565-602

[70] A.G.B. Lauder and D. Wan. *Counting points on varieties over finite fields of small characteristic.* Preprint, 2001,
Available at: `http://web.comlab.ox.ac.uk/oucl/work/alan.lauder`

[71] A.G.B. Lauder and D. Wan. *Computing zeta functions of Artin-Schreier curves over finite fields.* To appear in LMS J. Comp. Math.,
Available at: `http://web.comlab.ox.ac.uk/oucl/work/alan.lauder`

[72] A.K. Lenstra, H.W. Lenstra, and L. Lovász. *Factoring polynomials with rational coefficients.* Math. Ann. 261 (1982), no. 4, 515–534.

[73] A.K. Lenstra. *Factoring polynomials over algebraic number fields.* Computer algebra (London, 1983), 245–254, Lecture Notes in Comput. Sci., 162, Springer, Berlin, 1983.

[74] *Algorithms in algebraic number theory.* Bull. Amer. Math. Soc. (N.S.) 26 (1992), no. 2, 211–244. Available on-line at:
`http://www.ams.org/bull/bull-pre1996.html`

[75] R. Lercier and D. Lubicz. *A Quasi Quadatric Time Algorithm for Hyperelliptic Curve Point Counting.*
Available at: `www.math.u-bordeaux.fr/~lubicz/`

[76] S. Lichtenbaum. *Duality theorems for curves over p-adic fields.* Invent. Math., 7:120–136, 1969.

[77] Yuri Manin. *Parabolic points and zeta function of modular curves.* Math. USSR Izvestija, 6(1):19–64, 1972.

[78] B. Mazur and K.A. Ribet. *Two-dimensional representations in the arithmetic of modular curves.* Courbes modulaires et courbes de Shimura (Orsay, 1987/1988). Astérisque No. 196-197, (1991), 6, 215–255 (1992).

[79] A. Menezes, S. Vanstone, and T. Okamoto. *Reducing elliptic curve logarithms to logarithms in a finite field.* IEEE Trans. Inf. Theory, IT-39(5):1639–1646, 1993.

[80] Loïc Merel. *Universal Fourier expansions of modular forms.* In "On Artin's conjecture for odd 2-dimensional representations", number 1585 in Lecture Notes in Math. Springer, 1994.

[81] J.S. Milne. *Étale cohomology.* Princeton Mathematical Series, 33. Princeton University Press, Princeton, N.J., 1980.

[82] T. Miyake. *Modular forms.* Springer-Verlag, Berlin, 1989.

[83] B.M. Moret. *The theory of computation.* Addison-Wesley, 1998.

[84] L. Moret-Bailly. *Métriques permises.* Chapter II of [107].

[85] L. Moret-Bailly. *La formule de Noether pour les surfaces arithmétiques.* Inventiones Mathematicae **98** (1989), 491–498.

[86] J. Nekovar. *DEA 2003/04: Elliptic functions and elliptic curves.* Available on: www.math.jussieu.fr/~nekovar/co/ln/el/

[87] P. Parent. *Bornes effectives pour la torsion des courbes elliptiques sur les corps de nombres.* J. Reine Angew. Math. 506 (1999), 85–116.

[88] J. Pila. *Frobenius maps of abelian varieties and finding roots of unity in finite fields.* Math. Comp. **55** (1990), no. 192, 745–763.

[89] K. Ribet. *Galois representations attached to eigenforms with Nebentypus.* Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976), pp. 17–51. Lecture Notes in Math., Vol. 601, Springer, Berlin, 1977.

[90] K. Ribet. *On l-adic representations attached to modular forms. II.* Glasgow Math. J. 27 (1985), 185–194.

[91] K. Ribet. *Images of semistable Galois representations.* Olga Taussky-Todd: in memoriam. Pacific J. Math. 1997, Special Issue, 277–297.

[92] T. Saito. *Modular forms and p-adic Hodge theory.* Invent. Math. 129 (1997), no. 3, 607–620.

[93] T. Saito. *Hilbert modular forms and p-adic Hodge theory (preliminary version).* Preprint, available from the author's home page: http://www.ms.u-tokyo.ac.jp/~t-saito/pp.html

[94] T. Satoh. *The canonical lift of an ordinary elliptic curve over a finite field and its point counting.* J. Ramanujan Math. Soc. **15** (2000), no. 4, 247–270.

[95] A.J. Scholl. *Motives for modular forms.* Invent. Math. 100 (1990), no. 2, 419–430.

[96] R.J. Schoof. *Elliptic curves over finite fields and the computation of square roots mod p.* Math. Comp. **44** (1985), no. 170, 483–494.

[97] R.J. Schoof. *Counting points on elliptic curves over finite fields.* Les Dix-huitièmes Journées Arithmétiques (Bordeaux, 1993). J. Théor. Nombres Bordeaux 7 (1995), no. 1, 219–254.

[98] J-P. Serre. *Une interprétation des congruences relatives à la fonction τ de Ramanujan.* 1969 Séminaire Delange-Pisot-Poitou: 1967/68, Théorie des Nombres, Fasc. 1, Exp. 14, 17 pp.

[99] J-P. Serre. *Quelques applications du théorème de densité de Chebotarev.* Publications Mathématiques de l'IHES **54** (1981), 123–202.

[100] J-P. Serre. *Représentations linéaires des groupes finis.* Third revised edition. Hermann, Paris, 1978.

[101] J-P. Serre, *Lectures on the Mordell-Weil theorem*, Asp. Math. E15, Vieweg, 1989.

[102] G. Shimura. *Introduction to the arithmetic theory of automorphic functions.* Iwanami Shoten and Princeton University Press, Princeton, 1971.

[103] C. Soulé. *Lectures on Arakelov geometry.* With the collaboration of D. Abramovich, J.-F. Burnol and J. Kramer. Cambridge Studies in Advanced Mathematics, 33. Cambridge University Press, 1992.

[104] W.A. Stein *Algorithms for Computing with Modular Forms*, course notes, downloadable at `http://modular.ucsd.edu/257/notes/257.pdf`

[105] J. Sturm. *On the congruence of modular forms.* Number Theory (New York, 1984–1985), 275–280, Lecture Notes in Mathematics 1240, Springer, 1987.

[106] H. P. F. Swinnerton-Dyer. *On l-adic representations and congruences for coefficients of modular forms.* Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, 1972), pp. 1–55. Lecture Notes in Math., Vol. 350, Springer, Berlin, 1973.

[107] L. Szpiro. *Séminaire sur les pinceaux arithmétiques: la conjecture de Mordell.* Astérisque No. 127 (1985), Société Mathématique de France, 1990.

[108] John Tate. *Endomorphisms of abelian varieties over finite fields.* Invent. Math., 2:134–144, 1966.

[109] E.J. Volcheck. *Computing in the jacobian of a plane algebraic curve.* In "Algorithmic number theory, ANTS I", number 877 in lecture notes in computer science, pages 221–233. Springer, 1994.

[110] A. Weil. *Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen.* Math. Ann. 168 1967 149–156.

[111] A. Weil, *Basic number theory*, Springer 1967.

[112] G. Wiese. *On the faithfulness of parabolic cohomology as a Hecke module over a finite field.* Available on arxiv: math.NT/0511115.