Computing Congruences of Modular Forms and Galois Representations Modulo Prime Powers

Xavier Taixés i Ventosa[*] and Gabor Wiese[†]

21st January 2010

Abstract

This article starts a computational study of congruences of modular forms and modular Galois representations modulo prime powers. Algorithms are described that compute the maximum integer modulo which two monic coprime integral polynomials have a root in common in a sense that is defined. These techniques are applied to the study of congruences of modular forms and modular Galois representations modulo prime powers. Finally, some computational results with implications on the (non-)liftability of modular forms modulo prime powers and possible generalisations of level raising are presented.

2010 Mathematics Subject Classification: 11F33 (primary); 11F11, 11F80, 11Y40.

1 Introduction

Congruences of modular forms modulo a prime ℓ and – from a different point of view – modular forms over $\overline{\mathbb{F}}_\ell$ play an important role in modern Arithmetic Geometry. The most prominent recent example is Serre's modularity conjecture, which has just become a theorem of Khare, Wintenberger and Kisin. We particularly mention the various techniques for *Level Raising* and *Level Lowering* modulo ℓ that were already crucial for Wiles's proof of Fermat's Last Theorem.

Motivated by this, it is natural to study congruences modulo ℓ^n of modular forms and Galois representations. However, as working over non-factorial and non-reduced rings like $\mathbb{Z}/\ell^n\mathbb{Z}$ introduces many extra difficulties, one is led to first approach this subject from an algorithmic and computational point of view, which is the topic of this article.

We introduce a definition of when two algebraic integers a, b are congruent modulo ℓ^n. Our definition, which might appear non-standard at first, was forced upon us by three requirements: Firstly,

[*]Universitat Pompeu Fabra, Departament d'Economia i Empresa, Ramon Trias Fargas 25-27, 08005 Barcelona
xavier.taixes@upf.edu

[†]Universität Duisburg-Essen, Institut für Experimentelle Mathematik, Ellernstraße 29, 45326 Essen, Germany
gabor@pratum.net, http://maths.pratum.net/

1

we want it to be independent of any choice of number field containing a, b. Secondly, in the special case $n = 1$ a congruence modulo ℓ should come down to an equality in a finite field. Finally, if a, b lie in some number field K that is unramified at ℓ, then a congruence of a and b modulo ℓ^n should be a congruence modulo λ^n, where λ is a prime dividing ℓ in K.

Since algebraic integers are – up to Galois conjugacy – most conveniently represented by their minimal polynomials, we address the problem of determining for which prime powers ℓ^n two coprime monic integral polynomials have zeros which are congruent modulo ℓ^n. We prove that a certain number, called the reduced discriminant or – in our language – the congruence number of the two polynomials, in all cases gives a good upper bound and in favourable cases completely solves this problem. In the cases when the congruence number is insufficient, we use a method based on the Newton polygon of the polynomial whose roots are the differences of the roots of the polynomials we started with.

With these tools at our disposal, we target the problem of computing congruences modulo ℓ^n between two Hecke eigenforms. Since our motivation comes from arithmetic, especially from Galois representations, our main interest is in Hecke eigenforms. It quickly turns out, however, that there are several possible well justified notions of Hecke eigenforms modulo ℓ^n. We present two, which we call *strong* and *weak*. The former can be thought of as reductions modulo ℓ^n of q-expansions of holomorphic normalised Hecke eigenforms; the latter can be understood as linear combinations of holomorphic modular forms, which are in general not eigenforms, but whose reduction modulo ℓ^n becomes an eigenform (our definition is formulated in a different way, but can be interpreted to mean this). We observe that Galois representations to $\mathrm{GL}_2(R)$, where R is an extension of $\mathbb{Z}/\ell^n\mathbb{Z}$ in the sense of Section 2, can be attached to both weak and strong Hecke eigenforms (under the condition of residual absolute irreducibility).

Modular forms can be represented by their q-expansions (e.g. in $\mathbb{Z}/\ell^n\mathbb{Z}$), i.e. by power series. For computational purposes, such as uniquely identifying a modular form and comparing two modular forms, it is essential that already a finite segment of a certain length of the q-expansions suffices. We notice that a sufficient length is provided by the so-called Sturm bound, which is the same modulo ℓ^n as in characteristic 0.

The computational problem that we are mostly interested in is to determine congruences modulo ℓ^n between two newforms, i.e. equalities between strong Hecke eigenforms modulo ℓ^n. This problem is perfectly suited for applying our methods of determining congruences modulo ℓ^n of zeros of integral polynomials. The reason for this is that the Fourier coefficient a_p of a normalised Hecke eigenform is a zero of the characteristic polynomial of the Hecke operator T_p acting on a suitable integral modular symbols space (see e.g. [S] or [W2]). Thus, in order to determine the prime powers modulo which two newforms are congruent, we compute the congruences between the roots of these characteristic polynomials for a suitable number of p. One important point deserves to be mentioned here: If the two newforms that we want to compare do not have the same levels (but the same weights), one cannot expect that they are congruent at all primes; a different behaviour is to be expected at primes dividing the levels. We address this problem by applying the usual degeneracy maps 'modulo ℓ^n' in

2

order to land in the same level. All these considerations lead to an algorithm, which we sketch. We point out that this algorithm is much faster than the (naive) one which works with the coefficients of the modular forms as algebraic integers in a (necessarily big) number field.

We implemented the algorithm and performed many computations which led to observations that we consider very interesting. Some of the results are reported upon in Section 4. We are planning to investigate questions like 'Level Raising' in more detail in a subsequent work. We remark that the algorithm was already used in [DT] to determine some numerical examples satisfying the main theorem of that article.

Notation

We introduce some standard notation to be used throughout. In the article ℓ and p always refer to prime numbers. By an ℓ-adic field we shall understand a finite field extension of $\mathbb{Q}_\ell$. We fix algebraic closures $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ and $\overline{\mathbb{Q}}_\ell$ of $\mathbb{Q}_\ell$. By $\overline{\mathbb{Z}}$ and $\overline{\mathbb{Z}}_\ell$ we denote the integers of $\overline{\mathbb{Q}}$ and $\overline{\mathbb{Q}}_\ell$, respectively. If K is either a number field or a local field, then $\mathcal{O}_K$ denotes its ring of integers. In the latter case, π_K denotes a uniformiser, i.e. a generator of the maximal ideal of $\mathcal{O}_K$, and v_K is the valuation satisfying $v_K(\pi_K) = 1$. Moreover, v_ℓ denotes the valuation on K and on $\overline{\mathbb{Q}}_\ell$ normalised such that $v_\ell(\ell) = 1$.

2   Congruences modulo ℓ^n

In this section we give our definition of *congruences modulo ℓ^n* for algebraic and ℓ-adic integers and discuss how to compute them.

2.1   Definition

Since a question on congruences is a local question, we place ourselves in the set-up of ℓ-adic fields. Let $\alpha, \beta \in \overline{\mathbb{Z}}_\ell$. In our definition of congruences modulo ℓ^n we are led by three requirements: (1) If

$n = 1$, we want that $\alpha \equiv \beta \mod \ell$ if and only if the reductions of α and β are equal in $\overline{\mathbb{F}}_\ell$. (2) If α and β are elements of some finite unramified extension $K/\mathbb{Q}_\ell$, then we want $\alpha \equiv \beta \mod \ell^n$ if and only of $\alpha - \beta \in (\pi_K^n)$. (3) We want the definition to be independent of any choice of $K/\mathbb{Q}_\ell$ containing α and β.

We propose the following definition.

Definition 2.1 *Let $n \in \mathbb{N}$. Let $\alpha, \beta \in \overline{\mathbb{Z}}_\ell$. We say that α is congruent to β modulo ℓ^n, for which we write $\alpha \equiv \beta \mod \ell^n$, if and only if $v_\ell(\alpha - \beta) > n - 1$.*

Note that this definition satisfies our three requirements. Note also the trivial equivalence

$$\alpha \equiv \beta \mod \ell^n \Leftrightarrow \lceil v_\ell(\beta - \alpha) \rceil \geq n. \tag{2.1}$$

In the sequel of this article we will often speak of congruences modulo ℓ^n of (global) algebraic integers by fixing an embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_\ell$. The same notation will be used also in this situation without further comments.

2.2 Interpretation in terms of ring extensions

In this section we propose an interpretation of the above definition of congruences modulo ℓ^n in terms of ring extension of $\mathbb{Z}/\ell^n\mathbb{Z}$. This interpretation gives us a much better algebraic handle for working with such congruences because we will be able to use equality instead of congruence. We were led to Definition 2.1 by the following consideration: Let $K/\mathbb{Q}_\ell$ be a finite extension and $n \in \mathbb{N}$. What is the minimal m such that the inclusion $\mathbb{Z}_\ell \hookrightarrow \mathcal{O}_K$ induces an injection of $\mathbb{Z}/\ell^n\mathbb{Z}$ into $\mathcal{O}_K/(\pi_K^m)$? In order to formulate the answer, we introduce a function.

Definition 2.2 *Let $L/K/\mathbb{Q}_\ell$ be finite field extensions and let $e_{L/K}$ denote the ramification index of L/K. For $n \in \mathbb{N}$, let $\gamma_{L/K}(n) = (n-1)e_{L/K} + 1$.*

This function satisfies the following simple properties:

(i) For $n = 1$, we have $\gamma_{L/K}(1) = 1$.

(ii) If L/K is unramified, then $\gamma_{L/K}(n) = n$.

(iii) For extensions $M/L/K$, we have *multiplicativity*: $\gamma_{M/K}(n) = \gamma_{M/L}(\gamma_{L/K}(n))$.

(iv) For extensions L/K, the integer $\gamma_{L/K}(n)$ is the minimal one such that the embedding $\mathcal{O}_K \hookrightarrow \mathcal{O}_L$ induces an injection $\mathcal{O}_K/(\pi_K^n) \hookrightarrow \mathcal{O}_L/(\pi_L^{\gamma_{L/K}(n)})$.

(v) For $\alpha, \beta \in K/\mathbb{Q}_\ell$ we have:

$$v_K(\alpha - \beta) \geq \gamma_{K/\mathbb{Q}_\ell}(n) \Leftrightarrow v_\ell(\alpha - \beta) > n - 1 \Leftrightarrow \alpha \equiv \beta \mod \ell^n.$$

4

Note that (i)–(iii) precisely correspond to the requirements (1)–(3) from Section 2.1. By (iv) we have produced *ring extensions*

$$\mathbb{Z}/\ell^n\mathbb{Z} \hookrightarrow \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)}) \hookrightarrow \mathcal{O}_L/(\pi_L^{\gamma_{L/\mathbb{Q}_\ell}(n)}).$$

Property (v) immediately yields a reformulation of the congruence of α and β modulo ℓ^n as an equality in the residue ring $\mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$.

In order to interpret congruences as equalities without always having to choose some finite extension of $\mathbb{Q}_\ell$, we now make the following construction, which for $n = 1$ boils down to $\overline{\mathbb{F}}_\ell$. We define

$$\overline{\mathbb{Z}/\ell^n\mathbb{Z}} := \varinjlim_{K} \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)}),$$

where K runs through all subextensions of $\overline{\mathbb{Q}}_\ell$ of finite degree over $\mathbb{Q}_\ell$ and the inductive limit is taken with respect to the maps in (iv). The natural projections $\mathcal{O}_K \twoheadrightarrow \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$ give rise to a surjective ring homomorphism

$$\pi_n : \overline{\mathbb{Z}}_\ell \twoheadrightarrow \overline{\mathbb{Z}/\ell^n\mathbb{Z}}.$$

Now we can make another reformulation of our definition of congruences modulo ℓ^n: Let $\alpha, \beta \in \overline{\mathbb{Z}}_\ell$. Then we have

$$\alpha \equiv \beta \mod \ell^n \Leftrightarrow \pi_n(\alpha) = \pi_n(\beta).$$

In the sequel, we will always choose the π_n in a compatible way, i.e. if $m < n$ we want π_m to be the composition of π_n with the natural map $\overline{\mathbb{Z}/\ell^n\mathbb{Z}} \twoheadrightarrow \overline{\mathbb{Z}/\ell^m\mathbb{Z}}$.

Remark 2.3 We also point out a disadvantage of our choice of $\gamma_{K/\mathbb{Q}_\ell}(n)$, namely that it is not additive. This fact prevents us from defining a valuation on $\overline{\mathbb{Z}}_\ell$ by saying that the valuation of $a \in \overline{\mathbb{Z}}$ is equal to the maximal n such that $\pi_n(a) = 0$. Defining $\gamma_{K/\mathbb{Q}_\ell}(n)$ as n times the ramification index $e_{K/\mathbb{Q}_\ell}$ would have avoided that problem. But then $\gamma(1) = e_{K/\mathbb{Q}_\ell} \neq 1$, in general, which is not in accordance with the usual usage of modulo ℓ. This other possibility can be understood as $\overline{\mathbb{Z}}_\ell/\ell^n\overline{\mathbb{Z}}_\ell$.

2.3   Computing congruences modulo ℓ^n

If one does not require one fixed embedding into the complex numbers, algebraic integers are most easily represented by their minimal polynomials. Thus, it is natural to study congruences between algebraic integers entirely through their minimal polynomials. This is the point of view that we adapt and it leads us to consider the following problem.

Problem 2.4 *We fix, once and for all, for every n compatibly, ring homomorphisms $\pi_n : \overline{\mathbb{Z}} \hookrightarrow \overline{\mathbb{Z}}_\ell \twoheadrightarrow \overline{\mathbb{Z}/\ell^n\mathbb{Z}}$. Let $P, Q \in \mathbb{Z}[X]$ be two coprime monic polynomials and let $n \in \mathbb{N}$.*

How can we decide the validity of the following assertion?

"There exist $\alpha, \beta \in \overline{\mathbb{Z}}$ such that

(i) $P(\alpha) = Q(\beta) = 0$ *and*

(ii) $\pi_n(\alpha) = \pi_n(\beta)$ *(i.e. $\alpha \equiv \beta \mod \ell^n$)."*

In this article, we will give two algorithms for treating this problem. The first one arose from the idea that one could try to use greatest common divisors. This notion seems to be the right one for $n = 1$, but it is not well behaved for $n > 1$ since the ring $\mathbb{Z}/\ell^n\mathbb{Z}[X]$ is not a principal ideal domain. However, the algorithm for approximating greatest common divisors of two polynomials over $\mathbb{Z}_\ell$ presented in Appendix A of [FPR] led us to consider the notion of *congruence number* or *reduced resultant*. It can be used to give quite a fast algorithm, which, however, does not always give a complete answer.

The second algorithm, which we call the *Newton polygon method*, always solves Problem 2.4 but tends to be slower (experimentally). Its basic idea was suggested to us by Michael Stoll after a talk of the second author and was immediately put into practice. However, since the first version of this article had already been finished, the algorithm was not included in it, so that it was again suggested to us by one of the referees. In this section we will present both algorithms in detail.

It should be pointed out explicitly that Problem 2.4 cannot be solved completely by considering only the reductions of P and Q mod ℓ^n if $n > 1$. This is a major difference to the case $n = 1$. The difference is due to the fact that in the problem we want α and β to be zeros of P and Q: if $\overline{\alpha}$ and $\overline{\beta}$ are elements in $\overline{\mathbb{Z}/\ell^n\mathbb{Z}}$ such that inside that ring $P(\overline{\alpha}) = Q(\overline{\beta}) = 0$, then it is not clear if they are reductions of zeros of P and Q.

Congruence number

The congruence number of two integral polynomials provides an upper bound for congruences in the sense of Problem 2.4. It is defined in such a way that it can easily be calculated on a computer.

Definition 2.5 *Let R be any commutative ring. By $R[X]_{<n}$ we denote the R-module of polynomials of degree less than n. Let $P, Q \in R[X]$ be two polynomials of degrees m and n, respectively. The* Sylvester map *is the R-module homomorphism*

$$R[X]_{<n} \oplus R[X]_{<m} \to R[X]_{<(m+n)}, (r, s) \mapsto rP + sQ.$$

If R is a field, then the monic polynomial of smallest degree in the image of the Sylvester map is the greatest common divisor of P and Q. In particular, with R a factorial integral domain and P, Q primitive polynomials, the Sylvester map is injective if and only if P and Q are coprime. Consequently, if $P, Q \in \mathbb{Z}[X]$ are primitive coprime polynomials, then any non-zero polynomial of smallest degree is a constant polynomial.

Definition 2.6 *Let $P, Q \in \mathbb{Z}[X]$ be coprime polynomials. We define the* congruence number $c(P, Q)$ *of P and Q as the smallest positive integer c such that the constant polynomial c is in the image of the Sylvester map of P and Q.*

We remark that for monic coprime polynomials P and Q via polynomial division the principal ideal $(c(P,Q))$ can be seen to be equal to the intersection of the ideal of constant integral polynomials with the ideal in $\mathbb{Z}[X]$ generated by all polynomials $rP + sQ$ when r, s run through all of $\mathbb{Z}[X]$. In [Pohst] the congruence number is called the *reduced resultant*. Note that in general the reduced resultant is a proper divisor of the resultant. It makes sense to replace $\mathbb{Z}$ by $\mathbb{Z}_\ell$ everywhere and to define a congruence number as a constant polynomial in the image of the Sylvester map having the lowest ℓ-adic valuation. Although this element is not unique, its valuation is.

The congruence number gives an upper bound for the n in Problem 2.4:

Proposition 2.7 *Let $P, Q \in \mathbb{Z}[X]$ be coprime polynomials and let ℓ^n be the exact power of ℓ dividing $c(P,Q)$. Then there are no $\alpha, \beta \in \overline{\mathbb{Z}}$ such that*

(i) $P(\alpha) = Q(\beta) = 0$ *and*

(ii) $\pi_m(\alpha) = \pi_m(\beta)$ *(i.e. $\alpha \equiv \beta \mod \ell^n$) for any $m > n$.*

Proof. By assumption there exist $r, s \in \mathbb{Z}[X]$ such that $c = c(P,Q) = rP + sQ$. Let $\alpha, \beta \in \overline{\mathbb{Z}}$ be zeros of P and Q, respectively, such that $\pi_m(\alpha) = \pi_m(\beta)$. We obtain

$$\pi_m(c) = \pi_m\big(r(\alpha)P(\alpha) + s(\alpha)Q(\alpha)\big) = \pi_m\big(s(\alpha)\big)\pi_m\big(Q(\alpha)\big) = \pi_m\big(s(\beta)\big)\pi_m\big(Q(\beta)\big) = 0.$$

This means that ℓ^m divides c, whence $m \leq n$. $\qquad\square$

On the computation of the congruence number

The idea for the computation of the congruence number is very simple: we use basic linear algebra and the Sylvester matrix. The point is that the Sylvester map is described by the standard Sylvester matrix S of P and Q (or rather its transpose if one works with column vectors) for the standard bases of the polynomial rings. We describe in words the straight forward algorithm for computing the congruence number $c(P,Q)$ as well as for finding polynomials r, s such that $c(P,Q) = rP + sQ$ with $\deg(r) < \deg(Q)$ and $\deg(s) < \deg(P)$. The algorithm consists of bringing S into row echelon (or Hermite) form, i.e. one computes an invertible integral matrix B such that BS has no entries below the diagonal. The congruence number $c(P,Q)$ is (the absolute value of) the bottom right entry of BS and the coefficients of r and s are the entries in the bottom row of B. This algorithm works over the integers and over ℓ-adic rings with a certain precision, i.e. $\mathbb{Z}/\ell^n\mathbb{Z}$.

We note that by reducing BS modulo ℓ, one can read off the greatest common divisor of the reductions of P and Q modulo ℓ: its coefficients (up to normalization) are the entries in the last non-zero row of the reduction of BS modulo ℓ. This has the following trivial, but noteworthy consequence.

Corollary 2.8 *Suppose that P and Q are primitive coprime polyomials in $\mathbb{Z}[X]$. Then P and Q have a non-trivial common divisor modulo ℓ if and only if the congruence number of P and Q is divisible by ℓ.* $\qquad\square$

Applications of the congruence number

We now examine when the congruence number is enough to solve Problem 2.4 for given P, Q and for all n. In cases when it is not, we will give a lower bound for the maximum n for which the assertions of the problem are satisfied.

We start with the observation that the congruence number suffices to solve our problem for $n = 1$.

Proposition 2.9 *Let* $n = 1$. *Assume that* P *and* Q *are coprime monic polynomials in* $\mathbb{Z}[X]$. *The assertion in Problem 2.4 is satisfied if and only if the congruence number* $c(P, Q)$ *is divisible by* ℓ.

Proof. The calculations of the proof of Proposition 2.7 show that if the assertion is satisfied, then ℓ divides $c(P, Q)$. Conversely, if ℓ divides $c(P, Q)$ then by Corollary 2.8 the reductions of P and Q have a non-trivial common divisor and thus a common zero in $\overline{\mathbb{F}}_\ell$. All zeros in $\overline{\mathbb{F}}_\ell$ lift to zeros in $\overline{\mathbb{Z}}_\ell$.
$\square$

We fix an embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_\ell$. Our further treatment will be based on the following simple observation. Let $M \subset \overline{\mathbb{Q}}$ be any number field containing all the roots of the monic coprime polynomials $P, Q \in \mathbb{Z}[X]$ and let $c = c(P, Q) = rP + sQ$ with $r, s \in \mathbb{Z}[X]$, $\deg(r) < \deg(Q)$, $\deg(s) < \deg(P)$ and factor $Q(X) = \prod_i (X - \beta_i)$ in $\overline{\mathbb{Z}}[X]$. Then for $\alpha \in \overline{\mathbb{Z}}$ such that $P(\alpha) = 0$ we have

$$v_M(c) = v_M\big(s(\alpha)\big) + \sum_i v_M(\alpha - \beta_i). \tag{2.2}$$

Our aim now is to find a lower bound for the maximum of $v_M(\alpha - \beta_i)$ depending on $\pi_M(c)$. For that we discuss the two summands in the equation separately.

We first treat $v_M\big(s(\alpha)\big)$. By $\overline{F}$ we denote the reduction modulo ℓ of an integral polynomial F.

Proposition 2.10 *Suppose that* ℓ *divides* $c(P, Q)$.

(a) *If* $\overline{s}$ *and* $\overline{Q}$ *are coprime, then* $v_M\big(s(\alpha)\big) = 0$ *for all* $\alpha \in \overline{\mathbb{Z}}$ *with* $\pi_1(Q(\alpha)) = 0$.

(b) *If one of* $\overline{P}$ *or* $\overline{Q}$ *does not have any multiple factors, then there is* $\alpha \in \overline{\mathbb{Z}}$ *such that* $P(\alpha) = 0$, $\pi_1(Q(\alpha)) = 0$ *and* $v_M\big(s(\alpha)\big) = 0$, *or there is* $\beta \in \overline{\mathbb{Z}}$ *such that* $Q(\beta) = 0$, $\pi_1(P(\beta)) = 0$ *and* $v_M(r(\beta)) = 0$.

(c) *If* $\overline{P}$ *is an irreducible polynomial in* $\mathbb{F}_\ell[X]$ *and* Q *is irreducible in* $\mathbb{Z}_\ell[X]$, *then* $\overline{s}$ *and* $\overline{Q}$ *are coprime and* $v_M\big(s(\alpha)\big) = 0$ *for all* $\alpha \in \overline{\mathbb{Z}}$ *with* $\pi_1(Q(\alpha)) = 0$.

Proof. (a) Since $\overline{s}$ and $\overline{Q}$ are coprime, the reduction of α cannot be a root of both of them.

(b) We prove that there exists $y \in \overline{\mathbb{F}}_\ell$ which is a common zero of $\overline{P}$ and $\overline{Q}$, but not a common zero of $\overline{r}$ and $\overline{s}$ at the same time. Assume the contrary, i.e. that $\overline{r}(y) = \overline{s}(y) = 0$ for all $y \in \overline{\mathbb{F}}_\ell$ with $\overline{P}(y) = \overline{Q}(y) = 0$. Let $\overline{G} \in \mathbb{F}_\ell[X]$ be the monic polynomial of smallest degree annihilating all $y \in \overline{\mathbb{F}}_\ell$ with the property $\overline{P}(y) = \overline{Q}(y) = 0$. Then $\overline{G}$ divides $\overline{P}, \overline{Q}$ as well as by assumption $\overline{r}$ and $\overline{s}$. Hence, we have

$$0 = \overline{r}\,\overline{P} + \overline{s}\,\overline{Q} = \overline{G}^2\big(\overline{r_1}\,\overline{P_1} + \overline{s_1}\,\overline{Q_1}\big)$$

with certain polynomials $\overline{r_1}, \overline{P_1}, \overline{s_1}, \overline{Q_1} \in \mathbb{F}_\ell[X]$. We obtain the equation

$$0 = \overline{r_1}\overline{P_1} + \overline{s_1}\overline{Q_1} \tag{2.3}$$

and we also have $\deg(\overline{r_1}) < \deg(\overline{Q_1})$ and $\deg(\overline{s_1}) < \deg(\overline{P_1})$. As either $\overline{P}$ or $\overline{Q}$ does not have any multiple factor, it follows that $\overline{P_1}$ and $\overline{Q_1}$ are coprime. This contradicts Equation 2.3.

Hence, we have $y \in \overline{\mathbb{F}}_\ell$ with $\overline{P}(y) = \overline{Q}(y) = 0$ and $\overline{r}(y) \neq 0$ or $\overline{s}(y) \neq 0$. If $\overline{r}(y) \neq 0$ then we lift y to a zero β of Q. In the other case we lift y to a zero α of P.

(c) The assumptions imply that $\overline{Q} = \overline{P}^a$ for some a. As the degree of s is smaller than the degree of P, it follows that $\overline{s}$ and $\overline{P}$ are coprime. Thus also, $\overline{s}$ and $\overline{Q}$ are coprime and we conclude by (a).

□

We now treat the term $\sum_i v_M(\alpha - \beta_i)$.

Proposition 2.11 *Suppose that ℓ divides $c(P,Q)$ and that α is a root of P which is congruent to some root of Q modulo ℓ (which exists by Proposition 2.9). Assume without loss of generality that β_1 is a root of Q which is closest to α, i.e. such that $v_M(\alpha - \beta_1) \geq v_M(\alpha - \beta_i)$ for all i.*

(a) Suppose that $\overline{Q}$ has no multiple factors (i.e. the discriminant of Q is not divisible by ℓ, or, equivalently, the congruence number of Q and Q' is not divisible by ℓ).

Then $\sum_i v_M(\alpha - \beta_i) = v_M(\alpha - \beta_1)$.

(b) In general we have $v_M(\alpha - \beta_1) \geq \lceil \frac{1}{\deg(Q)} \left(\sum_i v_M(\alpha - \beta_i) \right) \rceil$.

Proof. (a) If $\overline{Q}$ does not have any multiple factors, then $v_M(\beta_1 - \beta_i) = 0$ for all $i \neq 1$. Consequently, $v_M(\alpha - \beta_i) = v_M(\alpha - \beta_1 + \beta_1 - \beta_i) = 0$ for $i \neq 1$.

(b) is trivial. □

We summarise of the preceding discussion in the following corollary, solving Problem 2.4 if $\overline{P}$ and $\overline{Q}$ do not have any multiple factors, and giving a partial answer in the other cases.

Corollary 2.12 *Let P, Q be coprime monic polynomials in $\mathbb{Z}[X]$ (or $\mathbb{Z}_\ell[X]$) and let ℓ^n be the highest power of ℓ dividing the congruence number $c := c(P,Q)$ and let $r, s \in \mathbb{Z}[X]$ (or $\mathbb{Z}_\ell[X]$) be polynomials such that $c = rP + sQ$ with $\deg(r) < \deg(Q)$ and $\deg(s) < \deg(P)$.*

(a) If $n = 0$, then no root of P is congruent modulo ℓ to a root of Q.

(b) If $n = 1$, then there are α, β in $\overline{\mathbb{Z}}$ (in $\overline{\mathbb{Z}}_\ell$, respectively) with $P(\alpha) = Q(\beta) = 0$ such that they are congruent modulo ℓ, and there are no α_1, β_1 in $\overline{\mathbb{Z}}$ (in $\overline{\mathbb{Z}}_\ell$, respectively) with $P(\alpha) = Q(\beta) = 0$ such that they are congruent modulo ℓ^2.

(c) Suppose now that $n \geq 1$ and that one of the following properties holds:

(i) $\overline{P}$ does not have any multiple factors and $\overline{Q}$ does not have any multiple factors (i.e. $\ell \nmid c(P, P')$ and $\ell \nmid c(Q, Q')$).

9

(ii) $\overline{Q}$ *does not have any multiple factors and* $\overline{s}$ *and* $\overline{Q}$ *are coprime.*

(iii) $\overline{P}$ *does not have any multiple factors and* $\overline{r}$ *and* $\overline{P}$ *are coprime.*

Then there are α, β *in* $\overline{\mathbb{Z}}$ *(in* $\overline{\mathbb{Z}}_\ell$, *respectively) with* $P(\alpha) = Q(\beta) = 0$ *such that they are congruent modulo* ℓ^n *and there are no* α_1, β_1 *in* $\overline{\mathbb{Z}}$ *(in* $\overline{\mathbb{Z}}_\ell$, *respectively) with* $P(\alpha_1) = Q(\beta_1) = 0$ *such that they are congruent modulo* ℓ^{n+1}.

(d) Suppose that $n \geq 1$.

(i) If $\overline{s}$ *and* $\overline{Q}$ *are coprime, let* $m = \lceil \frac{n}{\deg(Q)} \rceil$.

(ii) If $\overline{r}$ *and* $\overline{P}$ *are coprime, let* $m = \lceil \frac{n}{\deg(P)} \rceil$.

(iii) If (i) and (ii) do not hold, let $m = 1$

Then there are α, β *in* $\overline{\mathbb{Z}}$ *(in* $\overline{\mathbb{Z}}_\ell$, *respectively) with* $P(\alpha) = Q(\beta) = 0$ *such that they are congruent modulo* ℓ^m *and there are no* α_1, β_1 *in* $\overline{\mathbb{Z}}$ *(in* $\overline{\mathbb{Z}}_\ell$, *respectively) with* $P(\alpha_1) = Q(\beta_1) = 0$ *such that they are congruent modulo* ℓ^{n+1}.

Proof. In the proof we use the notation introduced above. The upper bounds in (b)-(d) were proved in Proposition 2.7.

(a) follows from Proposition 2.9.

(b) The existence of a congruence follows from Corollary 2.8.

(c) In case (i), by Proposition 2.10 (b) we can choose $\alpha, \beta \in \overline{\mathbb{Z}}$ congruent modulo ℓ with $P(\alpha) = 0$ and $\beta \in \overline{\mathbb{Z}}$ with $Q(\beta) = 0$ such that $v_M(s(\alpha)) = 0$ or $v_M(r(\beta)) = 0$. Without loss of generality (after possibly exchanging the roles of (P, r) and (Q, s)) we may assume the former case. In case (ii), by Proposition 2.10 (a) any $\alpha \in \overline{\mathbb{Z}}$ with $P(\alpha) = 0$ and $\pi_1(Q(\alpha)) = 0$ will satisfy $v_m(s(\alpha)) = 0$. In both cases, from Proposition 2.11 and Equation 2.2 we obtain the equality

$$v_M(c) = v_M(\ell^n) = v_M(\alpha - \beta_1),$$

where β_1 comes from Proposition 2.11. This gives the desired result. Case (iii) is just case (ii) with the roles of (P, r) and (Q, s) interchanged.

(d) also follows from Propositions 2.10 and 2.11 and Equation 2.2. More precisely, in case (i) we have the inequality

$$v_M(\alpha - \beta_1) \geq \lceil \frac{v_M(c)}{\deg(Q)} \rceil = \lceil \frac{en}{\deg(Q)} \rceil \geq \left(\lceil \frac{n}{\deg(Q)} \rceil - 1 \right)e + 1 = \gamma_{M/\mathbb{Q}_\ell}\left(\lceil \frac{n}{\deg(Q)} \rceil \right),$$

where e is the ramification index of $M/\mathbb{Q}_\ell$. Hence, $\pi_m(\alpha - \beta_1) = 0$ with $m = \lceil \frac{n}{\deg(Q)} \rceil$. Case (ii) is case (i) with the roles of (P, r) and (Q, s) interchanged. $\square$

Remark 2.13 It is straightforward to turn Corollary 2.12 into an algorithm. Say, $P, Q \in \mathbb{Z}[X]$ are coprime monic polynomials. First we compute the congruence numbers $c(P, P')$ and $c(Q, Q')$. If any of these is zero, then we factor P (respectively, Q) in $\mathbb{Z}[X]$ into irreducible polynomials $P = \prod_i P_i$

(respectively, $Q = \prod_j Q_j$). We then treat any pair (P_i, Q_j) separately and return the maximum upper and the maximum lower bound for congruences of zeros. For simplicity of notation, we now call the pair (P, Q).

Now we compute the congruence numbers $c = c(P, Q)$ and $c_P = c(P, P')$ as well as $c_Q = c(Q, Q')$, all of which are non-zero by assumption. Along the way we also compute polynomials $r, s \in \mathbb{Z}[X]$ such that $c = rP + sQ$ and $\deg(r) < \deg(Q)$ and $\deg(s) < \deg(P)$. For each prime power ℓ^n (with $n \geq 1$) exactly dividing c we do the following. If ℓ does not divide $c_P c_Q$, then we are in case (c)(i) and we know that there are $\alpha, \beta \in \overline{\mathbb{Z}}$ such that $P(\alpha) = 0 = Q(\beta)$ and $\pi_n(\alpha) = \pi_n(\beta)$. This is best possible and we have obtained a complete answer to Problem 2.4. If ℓ is coprime to c_P or c_Q, we check whether we are in case (c)(ii) or (c)(iii). Then we also obtain equality of the upper and lower bound and thus a complete answer to Problem 2.4. If we are in neither of these cases, then we use the much weaker lower bounds of part (d). In order to get a best possible result in this case, too, one can make use of the Newton polygon method to be described next.

Newton polygon method

We now present the second algorithm for treating Problem 2.4. The basic idea of this algorithm was suggested to us by Michael Stoll. Let still $P, Q \in \mathbb{Z}[X]$ be coprime monic polynomials. Consider factorisations in $\overline{\mathbb{Z}}[X]$:

$$P(X) = \prod_{i=1}^{u}(X - \alpha_i) \text{ and } Q(X) = \prod_{j=1}^{v}(X - \beta_j).$$

Now take $Q(X + Y) = \prod_{j=1}^{v}(X - (\beta_j - Y))$, considered as a polynomial in X with coefficients in $\mathbb{Z}[Y]$ and let $F(Y)$ be the resultant of $P(X)$ and $Q(X + Y)$ with respect to the variable X. By well known properties of the resultant one has

$$F(Y) = \pm \prod_{i=1}^{u}\prod_{j=1}^{v}(Y - (\beta_j - \alpha_i)).$$

Hence, the roots of $F(Y)$ are precisely the differences of the roots of P and Q. Thus, the slopes of the Newton Polygon of $F(Y) \in \mathbb{Z}_\ell[Y]$ are the $v_\ell(\beta_j - \alpha_i)$. We obtain the following result, solving Problem 2.4.

Proposition 2.14 *Let $P, Q \in \mathbb{Z}[X]$ be coprime monic polynomials and set $n := \lceil s \rceil$, where s is the biggest slope of the Newton polygon of the polynomial $F \in \mathbb{Z}_\ell[Y]$ defined above.*

Then there are $\alpha, \beta \in \overline{\mathbb{Z}}$ such that

(i) $P(\alpha) = Q(\beta) = 0$ and

(ii) $\pi_n(\alpha) = \pi_n(\beta)$ (i.e. $\alpha \equiv \beta \mod \ell^n$).

Moreover, n is the biggest integer satisfying this property.

Proof. Let $\alpha, \beta \in \overline{\mathbb{Z}}$ with $P(\alpha) = Q(\beta) = 0$ such that the slope of $\beta - \alpha$ is equal to s, i.e. $v_\ell(\beta - \alpha) = s$ (subject to the fixed embedding $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_\ell$). The proposition is an immediate consequence of Definition 2.1 and Equation 2.1. □

3 Modular forms and Galois representations modulo ℓ^n

In this section, we apply the methods from Section 2 to the study of congruences of modular forms and modular Galois representations modulo ℓ^n.

As in Section 2, we keep ring homomorphisms $\pi_n : \overline{\mathbb{Z}} \hookrightarrow \overline{\mathbb{Z}}_\ell \twoheadrightarrow \overline{(\mathbb{Z}/\ell^n\mathbb{Z})}$, compatibly for n, fixed. In this section, we restrict to $\Gamma_0(N)$ for simplicity. Everything can be generalised without any problems to $\Gamma_1(N)$ with the obvious modifications. Moreover, also for the simplicity of the exposition all our modular forms are cusp forms.

3.1 Modular forms modulo ℓ^n

For studying the notion of congruences modulo ℓ^n of modular forms it is useful to introduce the terminology of modular forms over $\mathbb{Z}/\ell^n\mathbb{Z}$ or, in abuse of language, modular forms modulo ℓ^n. In contrast to the case $n = 1$, one must be aware that lifting of modular forms over $\mathbb{Z}/\ell^n\mathbb{Z}$ to characteristic zero is not automatic. This will be reflected in our notions. We let $S_k(\Gamma_0(N))$ denote the $\mathbb{C}$-vector space of holomorphic cuspidal modular forms of weight k and level N.

Definition 3.1 *Let $\mathbb{T} := \mathbb{T}_k(\Gamma_0(N))$ be the $\mathbb{Z}$-subalgebra of $\mathrm{End}_\mathbb{C}(S_k(\Gamma_0(N)))$ generated by all the Hecke operators T_n, $n \in \mathbb{N}$.*

  (i) A modular form of weight k and level N over $\mathbb{Z}/\ell^n\mathbb{Z}$ (or modulo ℓ^n) is a $\mathbb{Z}$-module homomorphism $f : \mathbb{T} \to \overline{(\mathbb{Z}/\ell^n\mathbb{Z})}$.

  (ii) A modular form f over $\mathbb{Z}/\ell^n\mathbb{Z}$ is a weak Hecke eigenform *if f is a ring homomorphism.*

  (iii) A weak Hecke eigenform f over $\mathbb{Z}/\ell^n\mathbb{Z}$ is a strong Hecke eigenform *if f factors into ring homomorphisms $\mathbb{T} \to \overline{\mathbb{Z}}_\ell \xrightarrow{\pi_n} \overline{(\mathbb{Z}/\ell^n\mathbb{Z})}$.*

  (iv) Any normalised holomorphic Hecke eigenform $f = q + \sum_{m \geq 2} a_m(f)q^m$ (with $q = e^{2\pi i z}$ and $a_m \in \overline{\mathbb{Z}}$) gives rise to a strong Hecke eigenform over $\mathbb{Z}/\ell^n\mathbb{Z}$ via $\mathbb{T} \xrightarrow{T_m \mapsto a_m} \overline{\mathbb{Z}} \xrightarrow{\pi_n} \overline{(\mathbb{Z}/\ell^n\mathbb{Z})}$. This modular form will be referred to as the reduction *of f modulo ℓ^n.*

  (v) If the reductions modulo ℓ^n of two normalised holomorphic eigenforms f and g agree, then we say that f and g are congruent modulo ℓ^n. *This is the same as the congruence $a_m(f) \equiv a_m(g) \mod \ell^n$ for all $m \in \mathbb{N}$ with the notion of congruence from Section 2. If the congruence $a_p(f) \equiv a_p(g) \mod \ell^n$ holds for all primes p but possibly finitely many, we say that f and g are* congruent modulo ℓ^n at almost all primes.*

Remark 3.2 (a) It is often useful to think of a modular form f over $\mathbb{Z}/\ell^n\mathbb{Z}$ as the *q-expansion* $\sum_{n=1}^{\infty} f(T_n)q^n \in \overline{\mathbb{Z}/\ell^n\mathbb{Z}}[[q]]$.

(b) As $\mathbb{T}$ is a finitely generated (and free) $\mathbb{Z}$-module, every weak eigenform f can be factored as $\mathbb{T} \to \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)}) \to \overline{\mathbb{Z}/\ell^n\mathbb{Z}}$ for a suitable ℓ-adic field K.

(c) Let $f : \mathbb{T} \xrightarrow{\phi} \overline{\mathbb{Z}}_\ell \xrightarrow{\pi_n} \overline{\mathbb{Z}/\ell^n\mathbb{Z}}$ be a strong Hecke eigenform modulo ℓ^n. The kernel of ϕ is a minimal prime ideal $\mathfrak{p}$ of $\mathbb{T}$. As such, it corresponds to a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy class of holomorphic Hecke eigenforms, since $L := \mathrm{Frac}(\mathbb{T}/\mathfrak{p}) \subseteq \overline{\mathbb{Q}}$ is a number field (recall that $\mathbb{T}$ is a finitely generated free $\mathbb{Z}$-module) and $\mathfrak{p}$ is the kernel of the ring homomorphism

$$\mathbb{T} \twoheadrightarrow \mathbb{T}/\mathfrak{p} \subset L \hookrightarrow \overline{\mathbb{Q}} \subset \mathbb{C}, \ \ T_m \mapsto a_m,$$

which corresponds to the normalised holomorphic eigenform $\sum_{m \geq 1} a_m e^{2\pi imz}$ and depends on the choice of the embedding $L \hookrightarrow \overline{\mathbb{Q}}$. Hence, the notion of strong Hecke eigenform modulo ℓ^n implies that the form f is the reduction of a holomorphic Hecke eigenform modulo ℓ^n.

(d) For $n = 1$, the notion of weak and strong Hecke eigenform agree. The reason is that the kernel of $f : \mathbb{T} \to \overline{\mathbb{F}}_\ell$ is a maximal ideal, since the image of f is a (finite) field. Every maximal ideal of $\mathbb{T}$ contains a minimal prime ideal $\mathfrak{p}$ and, hence, f factors as $\mathbb{T} \to \mathbb{T}/\mathfrak{p} \hookrightarrow \overline{\mathbb{Z}} \hookrightarrow \overline{\mathbb{Z}}_\ell \twoheadrightarrow \overline{\mathbb{F}}_\ell$.

(e) Weak Hecke eigenforms need not be strong Hecke eigenforms in general. See, for instance, Section 4.2.

(f) Let R be any ring. Since $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}, \mathbb{Z}) \otimes_{\mathbb{Z}} R \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}, R)$ due to the freeness of $\mathbb{T}$ as a finitely generated $\mathbb{Z}$-module and since $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}, \mathbb{Z})$ can be identified with the holomorphic modular forms having integral Fourier expansions, any homomorphism $f : \mathbb{T} \to R$ (e.g. weak/strong eigenform) can be seen as an R-linear combination of holomorphic modular forms (which are not necessarily eigenforms).

(g) Another issue concerns the absence of a good Galois theory for the extensions of $\mathbb{Z}/\ell^n\mathbb{Z}$ discussed in Section 2: Let K be an ℓ-adic field. Not every ring homomorphism $\mathcal{O}_K \to \mathcal{O}_K/(\pi_K^m)$ comes from a field homomorphism $K \to K$. Suppose, for example, that $\mathcal{O}_K = \mathbb{Z}_\ell[X]/(P(X))$ is the ring of integers of a ramified extension of $\mathbb{Q}_\ell$. If α is a root of P and if m is big enough, then $\alpha + \pi^{m-1}$ is not a root of P, but nevertheless $P(\alpha + \pi^{m-1}) \in (\pi_K^m)$, whence sending α to $\alpha + \pi^{m-1}$ uniquely defines a ring homomorphism $\mathcal{O}_K \to \mathcal{O}_K/(\pi_K^m)$, which does not lift to a field automorphism $K \to K$. Hence, a strong Hecke eigenform modulo ℓ^n can give rise to many weak Hecke eigenforms modulo ℓ^n.

(h) Finally, we would like to point out a connection, as suggested by one of the referees, between the congruence number and the congruence exponent of modular abelian varieties defined in the paper [ARS] by Agashe, Ribet and Stein and our notions.

Let J be the Jacobian (over $\mathbb{Q}$) of some modular curve (say, $X_0(N)$) and A, B abelian subvarieties of J such that $J = A + B$ and $A \cap B$ is finite. For the moment, let $\mathbb{T}$ be the Hecke algebra of J, i.e. the subring of the endomorphism ring of J generated by all Hecke operators. Denote by $\mathbb{T}_A$ and $\mathbb{T}_B$ the Hecke algebras of A and B, respectively. The natural map $\phi : \mathbb{T} \to \mathbb{T}_A \oplus \mathbb{T}_B$ given by sending an operator T to its restrictions to A and B is injective due to the condition $J = A + B$. Thus, we can view $\mathbb{T}$ as an abelian subgroup of $\mathbb{T}_A \oplus \mathbb{T}_B$, which has finite index, since $A \cap B$ is finite. Agashe, Ribet and Stein define the *congruence exponent* (and the *congruence number*) of A as the exponent (the number of elements) of the abelian group $(\mathbb{T}_A \oplus \mathbb{T}_B)/\mathbb{T}$. Note that the definition also depends on B.

Now we establish the connection to our set-up. The Hecke algebra $\mathbb{T}$ is known to be isomorphic to the Hecke algebra $\mathbb{T}_2(\Gamma_0(N))$. Applying the functor $\mathrm{Hom}_{\mathbb{Z}}(\cdot, \overline{\mathbb{Z}/\ell^n\mathbb{Z}})$, we obtain the exact sequence

$$0 \to \mathrm{Hom}_{\mathbb{Z}}((\mathbb{T}_A \oplus \mathbb{T}_B)/\mathbb{T}, \overline{\mathbb{Z}/\ell^n\mathbb{Z}}) \xrightarrow{\alpha} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_A, \overline{\mathbb{Z}/\ell^n\mathbb{Z}}) \oplus \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_B, \overline{\mathbb{Z}/\ell^n\mathbb{Z}})$$
$$\xrightarrow{\beta} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}, \overline{\mathbb{Z}/\ell^n\mathbb{Z}}).$$

Note that the term on the right is precisely the group of weight 2 modular forms modulo ℓ^n on $\Gamma_0(N)$ in our definition. Let us now take two normalised newforms f and g in $S_2(\Gamma_0(N))$ in distinct Galois conjugacy classes such that f corresponds to a ring homomorphism $f : \mathbb{T}_A \to \mathbb{C}$ and g to $g : \mathbb{T}_B \to \mathbb{C}$. This is the case, for instance, if $A = (J/I_f J)^\vee$ and $B = I_f J$, where I_f is the kernel of the ring homomorphism $\mathbb{T} \to \mathbb{C}$ belonging to f. Assume that f and g are congruent modulo ℓ^n. This means by definition that $(f, -g)$ is in the kernel of β. We analyse the element $\psi \in \mathrm{Hom}_{\mathbb{Z}}((\mathbb{T}_A \oplus \mathbb{T}_B)/\mathbb{T}, \overline{\mathbb{Z}/\ell^n\mathbb{Z}})$ such that $\alpha(\psi) = (f, -g)$. It satisfies $\psi((T_1, 0) + \mathbb{T}) = f(T_1) - g(0) = 1$, since f is normalised. Consequently, $\mathbb{Z}/\ell^n\mathbb{Z}$ is in the image of ψ. Hence, $(\mathbb{T}_A \oplus \mathbb{T}_B)/\mathbb{T}$ contains an element of order ℓ^n. We conclude that ℓ^n divides the congruence exponent of A (and, of course, also the congruence number).

3.2 Galois Representations modulo ℓ^n

We are interested in congruences modulo ℓ^n (in the sense of Section 2) between 2-dimensional ℓ-adic Galois representations ($i = 1, 2$)

$$\rho_i : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathcal{O}_{K_i}),$$

i.e. $\mathcal{O}_{K_i}$ is the ring of integers of an ℓ-adic field. For that let K be an ℓ-adic field containing K_1 and K_2. We study the reductions of the representations modulo ℓ^n:

$$\overline{\rho}_i^{(n)} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathcal{O}_K) \xrightarrow{\text{nat. proj.}} \mathrm{GL}_2(\mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})).$$

Definition 3.3 *The representations ρ_1 and ρ_2 are called* congruent modulo ℓ^n *if $\overline{\rho}_1^{(n)}$ and $\overline{\rho}_2^{(n)}$ are isomorphic as $(\mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)}))[\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})]$-modules.*

14

Remark 3.4 The insistence on taking the natural projection is owed to the fact that there may be 'too many' maps from $\mathcal{O}_K \to \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$, as mentioned in Remark 3.2 (g).

Theorem 3.5 *If the ρ_i are residually absolutely irreducible, then they are congruent modulo ℓ^n if and only if the traces of Frobenius elements agree, i.e.* $\mathrm{Tr}(\overline{\rho}_1^{(n)}(\mathrm{Frob}_p)) = \mathrm{Tr}(\overline{\rho}_2^{(n)}(\mathrm{Frob}_p))$, *at a dense set of primes p.*

    Proof. Chebotarev's Theorem applied to the Proposition in [M2], p. 253.   □

    Subject to a fixed choice $\overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_\ell$, to a normalised holomorphic eigenform $f = \sum a_m q^m \in S_k(\Gamma_0(N))$ one can attach an ℓ-adic Galois representation $\rho_{f,\ell} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(K)$ with some (suitably large) ℓ-adic field K. This Galois representation has the properties that it is unramified outside ℓ and the level of f and the trace of Frob_p is equal to a_p at all unramified primes p.

Proposition 3.6 *Any weak or strong Hecke eigenform $f : \mathbb{T} \to \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$ of level N and weight k has an attached residual Galois representation $\overline{\rho}_{f,\ell}$. If $\overline{\rho}_{f,\ell}$ is absolutely irreducible, f gives rise to a Galois representation*

$$\overline{\rho}_{f,\ell}^{(n)} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)}))$$

which is unramified outside ℓN and satisfies for every $p \nmid \ell N$

$$\mathrm{Tr}(\overline{\rho}_{f,\ell}^{(n)}(\mathrm{Frob}_p)) = a_p, \text{ and } \det(\overline{\rho}_{f,\ell}^{(n)}(\mathrm{Frob}_p)) = p^{k-1},$$

where we write a_p for the p-th coefficient of f, i.e. $a_p = f(T_p)$.

    Proof. Any weak modular form modulo ℓ^n gives rise to a strong modular form modulo ℓ by reduction, and hence we dispose of $\overline{\rho}_{f,\ell}$. If the residual representation is absolutely irreducible, Theorem 3 (p. 225) from [C] implies the existence of a Galois representation

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{T} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell)$$

with the desired properties. Note that f factors as $\mathbb{T} \to \mathbb{T} \otimes_{\mathbb{Z}} \mathbb{Z}_\ell \xrightarrow{f_1} \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$. It hence suffices to compose ρ with the natural map coming from f_1.   □

3.3 Sturm bound modulo ℓ^n

If two Galois representations $\overline{\rho}_i^{(n)}$ ($i = 1, 2$) as in the previous subsection come from weak or strong modular forms modulo ℓ^n, then one can decide whether they are equivalent by comparing only finitely many coefficients, since one disposes of an effective bound for the two modular forms modulo ℓ^n to be equal. Such a bound is given by the Sturm bound ([Sturm]).

Theorem 3.7 *Let Γ be a congruence group containing $\Gamma_1(N)$, let $k \geq 1$ and let B be the* Sturm bound *defined by*

$$B := \frac{kb}{12} - \frac{b-1}{N},$$

15

where $b = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$. The Hecke algebra $\mathbb{T}$ acting on the space $S_k(\Gamma)$ is generated as a $\mathbb{Z}$-module by the Hecke operators T_n for $1 \le n \le B$. Moreover, for $\Gamma = \Gamma_0(N)$ the algebra $\mathbb{T}$ is generated as a $\mathbb{Z}$-algebra by the T_p for the primes $p \le B$.

Proof. Theorem 9.23 and Remark 9.24 from [S]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Theorem 3.8 *Let $f, g : \mathbb{T} \to \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_\ell}(n)})$ be two weak or strong Hecke eigenforms modulo ℓ^n on $\Gamma_0(N)$ for some weight k. Let $b = [SL_2(\mathbb{Z}) : \Gamma_0(N)]$. If for all primes*

$$p \le \frac{kb}{12} - \frac{b-1}{N}$$

we have

$$f(T_p) = g(T_p) \qquad (\text{i.e. ``} a_p(f) \equiv a_p(g) \mod \ell^n \text{''}),$$

then f is equal to g as a Hecke eigenform modulo ℓ^n.

Proof. As for $\Gamma = \Gamma_0(N)$ we have that $\mathbb{T}$ is generated as a $\mathbb{Z}$-algebra by the Hecke operators T_p for the primes $p \le B$ (Theorem 3.7), it follows that f and g are uniquely determined by their values at T_p for primes $p \le B$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Remark 3.9 The Sturm bound can easily be extended to modular forms with nebentype, see e.g. [S], Corollary 9.20.

We mention that in [CKR], the Sturm bound is proved by other means and is also extended to the situation when the two modular forms have different weights. It is also useful to remark that the Sturm bound for modular forms modulo ℓ^n is also a direct consequence of the Sturm bound for modular forms over $\mathbb{F}_\ell$ and Nakayama's Lemma: If $\mathbb{T} \otimes_{\mathbb{Z}} \mathbb{F}_\ell$ is generated as $\mathbb{F}_\ell$-vector space by the Hecke operators $T_1, \ldots, T_B$, then $\mathbb{T} \otimes_{\mathbb{Z}} \mathbb{Z}/\ell^n\mathbb{Z}$ is generated as a $\mathbb{Z}/\ell^n\mathbb{Z}$-modulo by $T_1, \ldots, T_B$, too.

3.4 Application of degeneracy maps

Theorem 3.8 gives a criterium for the Galois representations attached to two Hecke eigenforms $f \in S_k(\Gamma_0(N))$ and $g \in S_k(\Gamma_0(Nm))$ to be congruent modulo ℓ^n (under the assumption that the representations are residually irreducible). However, most of the time when studying congruences of Galois representations attached to modular forms f and g, the assumptions of Theorem 3.8 will not be fulfilled, as f and g will typically differ at some prime dividing one of the levels. Hence, we now propose a stronger criterion. In order to formulate it, we introduce some straightforward notation.

Definition 3.10 *Let R be a commutative ring (in the sequel, either $R = \mathbb{C}$, $R = \mathbb{Z}$ or R is an extension of $\mathbb{Z}/\ell^n\mathbb{Z}$ as in Section 2) and $d \in \mathbb{N}$. Let $N, m, n \in \mathbb{N}$. The* degeneracy map *for a positive divisor d of m is defined to be the map*

$$\phi_d : \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(N)), R) \to \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(Nm)), R)$$

which sends $f \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(N)), R)$ to the homomorphism in $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(Nm)), R)$ that maps T_n to $\phi(T_{n/d})$, if d divides n, and to 0 otherwise.

Let $f : \mathbb{T}_k(\Gamma_0(N)) \to R$ be a modular form over R. The old space *of f over R in level Nm is defined as the R-span of the image of f under the degeneracy maps for each positive $d \mid m$ inside $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(Nm)), R)$.*

On q-expansions, the degeneracy map for d corresponds to the R-module endomorphism of $R[[q]]$ given by $q \mapsto q^d$. The degeneracy map ϕ_d is well defined with $R = \mathbb{Z}$ by the classical theory of modular forms (via the identification of $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(N)), \mathbb{Z})$ with those holomorphic cusp forms in $S_k(\Gamma_0(N))$ having integral Fourier expansions) and due to the isomorphism $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(N)), \mathbb{Z}) \otimes_{\mathbb{Z}} R \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{T}_k(\Gamma_0(N)), R)$ it is well defined for all rings R.

Proposition 3.11 *Let f and g be weak Hecke eigenforms modulo ℓ^n of weight k for $\Gamma_0(N)$ and $\Gamma_0(Nm)$, respectively, and assume that their residual Galois representations are absolutely irreducible.*

Then the Galois representations modulo ℓ^n attached to f and g are isomorphic if there is a weak Hecke eigenform $\tilde{f}$ modulo ℓ^n in the oldspace of f modulo ℓ^n in level Nm such that $g(T_p) = \tilde{f}(T_p)$ (i.e. "$a_p(g) \equiv a_p(\tilde{g}) \mod \ell^n$") for the primes p up to the Sturm bound for weight k and $\Gamma_0(Nm)$.

Proof. The assumptions imply that the equality $g(T_p) = f(T_p)$ holds for all primes p except possibly those with p dividing m. Hence, we can conclude by Theorem 3.5. $\qquad\square$

Proposition 3.11 gives rise to a straightforward algorithm (see Section 3.5), since the characteristic polynomials of the Hecke operators at $p \mid m$ on the oldspace of f can be described explicitly as follows. Let $f \in S_k(\Gamma_0(N))$ and $g \in S_k(\Gamma_0(Nm))$ be Hecke eigenforms. Suppose that r is the maximum exponent such that $p^r \mid m$. Then T_p acts on the old space of f in level $p^r N$ as the $(r+1) \times (r+1)$ matrix

$$\tilde{T}_p = \begin{pmatrix} a_p(f) & 1 & 0 & 0 & \ldots & 0 \\ -\delta p^{k-1} & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots & 0 & 0 & 0 & 1 \\ 0 & \ldots & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.4}$$

where $\delta = 0$ if $p \mid N$ and $\delta = 1$ otherwise (see [W1]).

Let $[f]$ be the $\mathbb{Z}$-span of the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy class of f; say that its rank is d. The operator T_p acts on the image of $[f]$ in level mN as the $d \cdot (r+1) \times d \cdot (r+1)$ matrix resulting from (3.4), in which we substitute every 0 by the $d \times d$ dimensional 0_d matrix, 1 becomes the d-identity 1_d, the entry $a_p(f)$ is replaced by the $d \times d$ matrix of the Hecke operator T_p on $[f]$, and δ is either 0_d or 1_d. Since all the elements below the diagonal are 0 for all the blocks under the second line of blocks, we know that the characteristic polynomial of this big matrix will be the product of $X^{d(r-1)}$ and the

characteristic polynomial of the block matrix

$$
\left(
\begin{array}{c|c}
T_p & 1_d \\
\hline
-\delta p^{k-1} \cdot 1_d & 0_d
\end{array}
\right).
\tag{3.5}
$$

We now compute the characteristic polynomial of (3.5). Let $P_{f,p} = \sum_{i=0}^{d} c_i X^i = \prod_{j=1}^{d}(X - a_j)$ be the characteristic polynomial of the upper left block, where the a_j lie in some algebraic closure. With two polynomial variables $\tilde{X}, \tilde{Y}$ we hence have $\prod_j (\tilde{X} - a_j \tilde{Y}) = \sum_i c_i \tilde{X}^i \tilde{Y}^{d-i}$. We now plug in $\tilde{X} = X^2 + \delta p^{k-1}$ and $\tilde{Y} = X$ and obtain

$$
\prod_{j=1}^{d} (X^2 - a_j X + \delta p^{k-1}) = \sum_{i=0}^{d} \left(c_i X^{d-i}(X^2 + \delta p^{k-1})^i \right).
$$

By taking the Jordan normal form (over an algebraic closure) and rearranging the matrix, we see that this is the characteristic polynomial of (3.5). Hence, the characteristic polynomial $\tilde{P}_{f,p}$ of 3.4 is

$$
\tilde{P}_{f,p} = \sum_{i=0}^{d} \left(c_i X^{dr-i}(X^2 + \delta p^{k-1})^i \right),
\tag{3.6}
$$

which can be computed very quickly from $P_{f,p}$. Let us remark that, if $p \mid N$, this polynomial is simply $X^{dr} \cdot P_{f,p}$ and, if $p \nmid N$ and $d = 1$, then $\tilde{P}_{f,p}$ is X^{r-1} times the characteristic polynomial of the p-Frobenius element.

Remark 3.12 (a) It appears worthwhile to investigate the existence of a partial converse to Proposition 3.11. A true converse cannot hold if f is in the lowest possible level, since it is easy to construct a counter example if $n = 1$, $k = 2$ and $\ell = 2$ and there is a weight-1 form embedded into weight 2. Under certain conditions (e.g. $k < \ell$ and $\ell \nmid Nm$) a converse could conceivably exist.

To illustrate the problem with a particular example, let us consider the unique Hecke eigenform f modulo 2 in level $\Gamma_0(23)$ of weight one. It satisfies $a_2(f) = 1 \in \mathbb{F}_2$. It can be embedded into weight 2 for the same level in two different ways (multiplying by the Hasse invariant, which does not change the q-expansion, and applying the Frobenius, which sends q to q^2). Consequently, there are two distinct Hecke eigenforms over $\mathbb{F}_2$ in weight 2 for $\Gamma_0(23)$ whose coefficients at 2 are precisely the roots of $X^2 + X + 1 \in \mathbb{F}_2[X]$. The coefficients at the other primes are equal to the coefficients of f, whence the attached mod 2 Galois representations are equal. Consequently, a converse to Proposition 3.11 cannot exist (since in this case $m = 1$).

(b) The trick used in [CKR] will always work for deciding whether the representations attached to f and g are congruent modulo ℓ^n: By applying degeneracy maps at all primes dividing Nm one can force all coefficients $a_p(f)$ and $a_p(g)$ to be congruent to zero modulo ℓ^n for all $p \mid Nm$. This allows the application of the Sturm bound. But, usually the level and hence the bound will be bigger than the bound in Proposition 3.11.

(c) We mention a point which will be discussed in more detail in Section 4.3. We are mostly interested in congruences of Galois representations modulo ℓ^n attached to holomorphic eigenforms, hence, it seems natural to stick to *strong* Hecke eigenforms. However, since we formulated Proposition 3.11 for *weak* Hecke eigenforms, we do not need to have a congruence mod ℓ^n of ℓ-adic zeros at $p \mid m$, but a simple equality in the residue ring is enough. Currently, in the algorithm we are not using this subtle distinction, but, as we will see in the example, it can make a difference.

3.5 Algorithm

The aim is to study the following problem algorithmically.

Problem 3.13 *Let f_1, f_2 be newforms in levels N_1, N_2 and weights k_1, k_2.*

Determine a finite list of prime powers $\{\ell_1^{n_1}, \ldots, \ell_r^{n_r}\}$ such that for all $i \in \{1, \ldots, r\}$ the ℓ_i-adic Galois representations attached to the modular forms f_1 and f_2 are congruent modulo $\ell_i^{n_i}$ and are incongruent modulo $\ell_i^{n_i+1}$, and for any ℓ distinct from all the ℓ_i the ℓ-adic Galois representations of f_1 and f_2 are incongruent modulo ℓ.

Towards this problem we employ the methods developed in the Section 2. Due to its greater speed we first apply the congruence number method, which by Proposition 2.7 gives an upper bound for the possible congruences. Only if in one of the applications of Corollary 2.12 the upper bound is unequal to the lower bound we make use of the Newton polygon method.

We hence start by computing the congruence numbers $c_p = c(P_{f_1,p}, P_{f_2,p})$ for all primes $p \nmid N_1 N_2$ up to some bound (e.g. the Sturm bound), where $P_{f_i,p}$ denotes the characteristic polynomial (in $\mathbb{Z}[X]$) of the Hecke operator T_p acting on the span of the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy class $[f_i]$ of f_i. Let us number the primes $p_1, p_2, \ldots$. We compute a slightly modified greatest common divisor of all c_p, taking in account only the prime-to-p part of c_p, because we want to disregard the coefficient a_p when reducing modulo powers of p. More precisely, if we have two c_{p_1} and c_{p_2}, the first greatest common divisor that we compute will be $c = \gcd(c_{p_1} \cdot p_1^{v_{p_1}(c_{p_2})}, c_{p_2} \cdot p_2^{v_{p_2}(c_{p_1})})$. Once we have one c computed, we can improve it for the next p_i with $c' = \gcd(c_{p_i} \cdot p_i^{v_{p_i}(c)}, c)$. The significance of the number c' is that it gives an upper bound for Problem 3.13: if a prime power ℓ^n does not divide c', then there cannot exist any congruence modulo ℓ^n between the ℓ-adic Galois representations attached to f_1 and f_2.

Our approach to a solution of Problem 3.13 is based on Theorem 3.8 and Proposition 3.11 in order to obtain a lower bound, which in favourable cases equals the upper bound c'. However, whether we use the congruence number method or the Newton polygon method for computing congruences between zeros of the characteristic polynomials of the Hecke operators, we have to assume the following hypothesis, which – roughly speaking – says that it is no loss to work with $P_{f,p}$ instead of with its roots.

Hypothesis 3.14 *Let f_1 and f_2 be two newforms and $n \in \mathbb{N}$. Suppose that for all primes p there are embeddings $\sigma_{i,p} : K \hookrightarrow \overline{\mathbb{Q}}$ $(i = 1, 2)$ such that*

$$\sigma_{1,p}\big(a_p(f_1)\big) \equiv \sigma_{2,p}\big(a_p(f_2)\big) \mod \ell^n.$$

Then there are embeddings σ_1, σ_2 such that $\sigma_1(f_1) \equiv \sigma_2(f_2) \mod \ell^n$.

An equivalent formulation is the following: If $P_{f_1,p}$ and $P_{f_2,p}$ have roots congruent modulo ℓ^n (in the sense of Section 2) for all p, then there are members $\tilde{f}_i$ in the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy class of f_i for $i = 1, 2$ such that f_1 is congruent to f_2 modulo ℓ^n.

In the sequel we shall assume this hypothesis to be satisfied. Note that by using characteristic polynomials of Hecke operators we lose track of which form in the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy class really satisfies a congruence. By abuse of language we will nevertheless speak of a congruence between $\rho_{f,\ell}$ and $\rho_{g,\ell}$ modulo ℓ^n when indeed we only have a congruence of $\rho_{\tilde{f},\ell}$ and $\rho_{\tilde{g},\ell}$ for some members $\tilde{f}$ and $\tilde{g}$ of the conjugacy classes of f and g, respectively. We now sketch our algorithm for treating Problem 3.13.

Input: $f \in S_k(\Gamma_0(N_f))$ and $g \in S_k(\Gamma_0(N_g))$ be two normalised eigenforms.

Output: (L^-, L^+) (for an explanation see below).

- (Upper bound) For every prime $p \nmid N_f N_g$ up to the Sturm bound B (see Theorem 3.7), we compute the congruence number $c_p = c(P_{f,p}, P_{g,p})$ and we calculate $L^+ = \gcd_{p \leq B}(c_p)$ with the modified greatest common divisor described above. We recall that $P_{f,p}$ denotes the characteristic polynomial of the Hecke operator T_p acting on the span $[f]$ of the Galois conjugacy class of f, which can for instance be obtained as the characteristic polynomial of the action of T_p on a suitable modular symbols space.

- For every $\ell \mid L^+$, we compute $L^-_{1,\ell} = \min_{p \leq B}(\ell^{d_p})$, where ℓ^{d_p} is the maximal power of ℓ modulo which $P_{f,p}$ and $P_{g,p}$ have a root in common. This number is obtained from the congruence number method if the value returned by it is best possible, i.e. if we are in case (c) or (b) of Corollary 2.12. Otherwise, the Newton polygon method is employed. We then form the product $L^-_1 = \prod_{\ell \mid L^+} L^-_{1,\ell}$.

- Suppose for this step that $N_g = mN_f$ and that $\overline{\rho}_{f,\ell}$ and $\overline{\rho}_{g,\ell}$ are absolutely irreducible. Then, for every $\ell \mid L^+$ such that $v_\ell(L^+) \neq v_\ell(L^-_1)$, we compute $L^-_{2,\ell} = \min_{p \leq B}(\ell^{\tilde{d}_p})$ as follows: If $p \nmid m$, then we put $\tilde{d}_p = d_p$. If $p \mid m$, we let $\ell^{\tilde{d}_p}$ be the maximal power of ℓ modulo which $\tilde{P}_{f,p}$ and $P_{g,p}$ have a root in common with $\tilde{P}_{f,p}$ as in Equation (3.6). This number is again calculated by the congruence number method or the Newton polygon method as in the previous step. Again we compute $L^-_2 = \prod_{\ell \mid L^+} L^-_{2,\ell}$.

- We compute $L^- = \prod_{\ell \mid L^+} \max(L^-_{1,\ell}, L^-_{2,\ell})$.

- Return (L^-, L^+).

Proposition 2.7 ensures that L^+ is an upper bound, i.e. that $\rho_{f,\ell}$ and $\rho_{g,\ell}$ are incongruent modulo ℓ^m (more precisely, this holds for any members of the conjugacy classes of f and g) if $\ell^m \nmid L^+$. Theorem 3.8 guarantees that L^-_1 is a lower bound (under Hypothesis 3.14), meaning that under the hypothesis $\rho_{f,\ell}$ and $\rho_{g,\ell}$ are congruent modulo ℓ^n if $\ell^n \mid L^-_1$ (with the slight abuse of language pointed out above). The lower bound L^-_1 will in general be very bad (e.g. 1) due to the Hecke operators T_p for

$p \mid m$ (in the situation of the third step). This is taken care of in the third step and Proposition 3.11 tells us that L_2^- is a lower bound in the same sense as before (still under Hypothesis 3.14). Consequently, L^- is a lower bound under Hypothesis 3.14.

Remark 3.15 We point out that this algorithm might miss a congruence modulo ℓ^n due to the Hecke operator T_ℓ. Hence, one might want to exclude the operators T_ℓ in all the steps. Then, however, we do not have the congruence of g with an oldform of f (as in Proposition 3.11), hence, the congruence of the Galois representations suggested by the output of the algorithm will not be a proved result even under Hypothesis 3.14 (but the correct one in most cases).

4 Examples and numerical data

In this section we present some cases which were computed using the algorithm described above and which we consider interesting. Several more examples can be found in [T]. For our calculations we used the computer algebra system MAGMA ([Magma]).

4.1 Examples of congruences in the same level

We computed all congruences between modular forms of weight 2 and the same level up to level 2000. In Table 1, (N_j, i_j) means the i_j-th form in level N_j for $j = 1, 2$ (according to an internal ordering in MAGMA), where in these cases we have $N_1 = N_2$. In all these cases, we found $L^- = L^+$ so that under Hypothesis 3.14 we obtained all congruences.

- The biggest exponents that we found appear in 2^7 and 2^5.

- For $n = 4$, we find some congruences modulo 3^4 (also modulo 2^4).

- For $n = 3$, the primes $\ell = 5$ and $\ell = 7$ appear.

- For $n = 2$ we already have many different primes, 47^2 being the biggest square of a prime that we found.

- For $n = 1$ we just listed some of the biggest congruences that we found. $2 \cdot 8581981 = 17163962$ and $1933 \cdot 8713 = 16842229$ are just two examples of congruences, but in this case we had several primes to choose from.

4.2 Simple example for strong $\neq$ weak

We now analyse the example with the smallest level in the above table more thoroughly. On $\Gamma_0(71)$ there are two $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy classes of newforms in weight 2. The coefficient fields of both of them are isomorphic; they have degree 3, discriminant 257 and are non-Galois. The prime 3 factors in

N_1	i_1	N_2	i_2	lower bound	upper bound
1479	16	1479	8	2^7	2^7
1027	2	1027	1	2^5	2^5
602	8	602	7	2^5	2^5
1454	7	1454	1	3^4	3^4
1171	4	1171	2	3^4	3^4
1147	6	1147	5	7^3	7^3
1726	6	1726	3	5^3	5^3
1629	4	1629	3	5^3	5^3
613	2	613	1	$7 \cdot 47^2$	$7 \cdot 47^2$
1939	4	1939	2	$37^2 \cdot 4423$	$37^2 \cdot 4423$
1906	5	1906	3	19^2	19^2
1763	8	1763	5	$3 \cdot 13^2$	$3 \cdot 13^2$
1761	8	1761	7	$2 \cdot 8581981$	$2 \cdot 8581981$
1241	2	1241	1	$1933 \cdot 8713$	$1933 \cdot 8713$
71	2	71	1	$2 \cdot 3^2$	$2 \cdot 3^2$
109	3	109	1	2^2	2^2
155	4	155	2	2^4	2^4
233	3	233	1	3^3	3^3
785	2	785	1	7^3	7^3
1073	6	1073	3	$2 \cdot 17^2$	$2 \cdot 17^2$
1481	3	1481	1	$5^2 \cdot 2833$	$5^2 \cdot 2833$

Table 1: Extract from the computational results.

two prime ideals $\mathfrak{P}_1$ and $\mathfrak{P}_2$ of residue degrees 1 and 2. This means that each of the two $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy classes gives us precisely one strong Hecke eigenform f_i modulo 3^n with coefficients in $\mathbb{Z}/3^n\mathbb{Z}$ for $i = 1, 2$; the others taken modulo 3 have coefficients in $\mathbb{F}_9$.

We compute that f_1 and f_2 are congruent modulo 9, but incongruent modulo 27. Let $\mathbb{T} \subset \mathrm{End}_{\mathbb{C}}(S_2(\Gamma_0(71)))$ be the Hecke algebra, i.e. the subring generated by the Hecke operators. The above discussion shows that there is a maximal ideal $\mathfrak{m}$ of $\hat{\mathbb{T}} := \mathbb{T} \otimes_{\mathbb{Z}} \mathbb{Z}_3$ such that the localisation $\hat{\mathbb{T}}_{\mathfrak{m}}$ has two minimal prime ideals, corresponding to the two strong Hecke eigenforms f_1 and f_2. A computer calculation yields that $\hat{\mathbb{T}}_{\mathfrak{m}} \otimes_{\mathbb{Z}_3} \mathbb{Z}/9\mathbb{Z} \cong \mathbb{Z}/9\mathbb{Z}[X]/(X^2)$. Thus, we have three weak Hecke eigenforms modulo 9 coming from $\hat{\mathbb{T}}_{\mathfrak{m}}$, namely

$$\hat{\mathbb{T}}_{\mathfrak{m}} \twoheadrightarrow \hat{\mathbb{T}}_{\mathfrak{m}} \otimes_{\mathbb{Z}_3} \mathbb{Z}/9\mathbb{Z} \cong \mathbb{Z}/9\mathbb{Z}[X]/(X^2) \xrightarrow{X \mapsto 0 \text{ or } X \mapsto 3 \text{ or } X \mapsto 6} \mathbb{Z}/9\mathbb{Z}.$$

Since we know that there is only one strong Hecke eigenform modulo 9, two of them cannot be strong.

4.3 Example in levels 149 and $149 \cdot 13$

On $\Gamma_0(149)$ for weight 2 there are two $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-conjugacy classes of newforms. The degrees of the coefficient fields are 3 and 9. Let f be any of the forms whose coefficient field $\mathbb{Q}_f$ has degree 9. The prime 3 is unramified in $\mathbb{Q}_f$ and there is a prime $\mathfrak{P}$ of residue degree 1 in the ring of integers $\mathcal{O}_f$ of $\mathbb{Q}_f$.

Mazur's Eisenstein ideal ([M1]) shows that the residual representation $\overline{\rho}_{f,\mathfrak{P}}$ of f modulo $\mathfrak{P}$ is irreducible, since 149 is a prime number and 3 does not divide $149 - 1$. We first want to determine the image of the residual representation. A quick computation of a couple of coefficients of f shows that the image of $\overline{\rho}_{f,\mathfrak{P}}$ contains all possible combinations of trace and determinant. Consulting the list of subgroups of $\mathrm{GL}_2(\mathbb{F}_3)$ tells us that next to the full $\mathrm{GL}_2(\mathbb{F}_3)$ there is only one other subgroup satisfying this property. That subgroup, however, does not contain any element of order 3. Due to the semistability at 13 and 149 this group is excluded, whence the image is the full $\mathrm{GL}_2(\mathbb{F}_3)$.

There is a newform g of weight 2 on $\Gamma_0(13 \cdot 149)$ and a prime ideal Λ dividing 3 in its coefficient field such that the strong Hecke eigenform of g obtained by reducing its q-expansion modulo Λ is equal to the strong Hecke eigenform of f modulo $\mathfrak{P}$ at all prime coefficients except at 13. In fact, our algorithm gives us a congruence modulo 3^{10} (in the sense defined before) at all primes up to the Sturm bound, except 13. Moreover, 3^{10} is also an upper bound. At the prime 13 we want to apply Proposition 3.11 (i.e. the third item of the algorithm), and we hence apply the methods from Corollary 2.12 to $P_{g,13}$ and $\tilde{P}_{f,13}$. However, the upper and the lower bounds we obtain with this method are 3^9. Hence, the output of our algorithm would be a congruence modulo 3^9 of the Galois representations attached to f and g as lower bound and 3^{10} as upper bound. We analyse the situation a bit more closely by hand. The polynomial $P_{g,13}$ is equal to $(X + 1)^{80}$. The polynomial $\tilde{P}_{f,13} = Q^2$ with $Q \in \mathbb{Z}[X]$ an irreducible polynomial of degree 18. Evaluating Q at -1 (the zero of $P_{g,13}$) gives $2^6 \cdot 3^{10} \cdot 6869$. This means that there is a *weak* Hecke eigenform $\tilde{f}$ in the oldspace of f modulo 3^{10} such that $\tilde{f}(T_{13}) = -1$. Hence, Proposition 3.11 yields that $\tilde{f}$ and g are congruent modulo 3^{10} as weak Hecke eigenforms. Consequently, the attached Galois representations of f and g are congruent modulo 3^{10}.

We give a more formal argument for the existence of the weak Hecke eigenform modulo 3^{10}. Let $\mathbb{T}$ be the Hecke algebra on $S_2(\Gamma_0(149 \cdot 13))$ (as $\mathbb{Z}$-algebra) and let $\mathbb{T}_{[f]}^{\mathrm{old}}$ be the Hecke algebra (as $\mathbb{Z}$-algebra) on the image of $[f]$ under the 13-degeneracy map, where as before $[f]$ denotes the span of the Galois conjugacy classes of f. By restricting Hecke operators, we obtain a surjective ring homomorphism $\mathbb{T} \twoheadrightarrow \mathbb{T}_{[f]}^{\mathrm{old}}$. The algebra $\mathbb{T}_{[f]}^{\mathrm{old}}$ is generated by the identity matrix and $\tilde{T}_{13}$ (see Equation (3.4)). Since the minimal polynomial of $\tilde{T}_{13}$ is either Q or Q^2, the composition

$$\mathbb{T} \twoheadrightarrow \mathbb{T}_{[f]}^{\mathrm{old}} \xrightarrow{\tilde{T}_{13} \mapsto -1} \mathbb{Z}/3^{10}\mathbb{Z}$$

is a well-defined ring homomorphism, i.e. the desired weak Hecke eigenform modulo 3^{10}.

4.4 Congruences with Eisenstein series modulo ℓ^n

Let $f \in S_2(\Gamma_0(N))$ such that $\overline{\rho}_{f,\ell}$ is reducible (and semi-simple by definition). This means that f is congruent modulo ℓ to an Eisenstein series in the same level and weight at almost all primes. The converse of this statement also holds. In the context of this article, it is natural to study congruences between newforms and Eisenstein series modulo ℓ^n and to do so via the congruence number and the Newton polygon method. By computing congruences modulo ℓ^n with Eisenstein series, we study up to which ℓ^n the representation $\overline{\rho}_{f,\ell^n}$ has the same traces at the first couple of Frobenius elements at good primes as an extension of the cyclotomic character modulo ℓ^n by the trivial representation.

Let f be a newform of weight k and level N. We implemented an algorithm, which for all primes $p \nmid N$ up to the Sturm bound computes the maximal prime powers modulo which $P_{f,p}$ (as before, this is the characteristic polynomial of T_p acting on $[f]$) and the characteristic polynomial of T_p acting on the Eisenstein subspace in the given level and weight have a root in common. We then proceed as earlier, obtaining an upper bound for a congruence with an Eisenstein series as well as an unproved lower bound (note that we do not take all operators into account).

A famous theorem of Mazur's ([M1]) states that in weight 2 and prime level N there is a cusp form which is congruent to the Eisenstein series modulo ℓ at almost all primes for every ℓ dividing the numerator of $\frac{N-1}{12}$. One can ask in how far this theorem holds modulo ℓ^n. It quickly turns out that a too naive generalisation is false. We propose to study the following in a subsequent paper. Let $f_1, \ldots, f_r$ be all newforms in prime level N and weight 2 for the trivial Dirichlet character. For $i = 1, \ldots, r$ let ℓ^{n_i} be the highest power of ℓ such that f_i is congruent at almost all primes to the Eisenstein series of level N and weight 2 modulo ℓ^{n_i}. Put $n := n_1 + \ldots + n_r$.

Question 4.1 *Is n at least as big as (or even equal to) the ℓ-valuation of the numerator of $\frac{N-1}{12}$?*

4.5 Level raising modulo ℓ^n

Let $f \in S_2(\Gamma_0(N))$ be a newform. The term *level raising modulo ℓ^n* in the simplest case refers to the problem of identifying primes $p \nmid N$ such that there is a newform g in $S_2(\Gamma_0(Np))$ with the property that f and g are congruent modulo ℓ^n at almost all primes. A necessary condition for level raising of the form f modulo ℓ at the prime $p \nmid N$ when its Galois representation is residually irreducible, is that ℓ divides the congruence number $c(P_{f,p}, X - (p+1))$ or the congruence number $c(P_{f,p}, X + (p+1))$. It is a famous theorem of Ribet's ([R]) that the converse also holds (modulo ℓ).

It is natural to ask whether or in which sense level raising generalises to congruences modulo ℓ^n. We start by an observation which we consider very interesting. Let f be the only newform on $\Gamma_0(17)$ in weight 2 and let $p = 59$. The coefficient $a_{59}(f) = -12$ and we find that 9 divides $c(P_{f,59}, X - 60) = c(X + 12, X - 60) = 72$ and that 3 divides $c(P_{f,59}, X + 60) = c(X + 12, X + 60) = 48$. However, there does not seem to be a congruence modulo 9 of f with any form in level $17 \cdot 59$. Instead, there appear to be three newforms in that level which are congruent to f modulo 3 at almost all primes. Hence, we conclude that the condition that ℓ^n divides one of the above congruence numbers is not

a sufficient one for level raising of strong Hecke eigenforms. This confirms a remark by Richard Taylor.[1]

We propose to study the following question in a subsequent paper. Let $f \in S_2(\Gamma_0(N))$ be some newform and let $p \nmid N$ be a prime. Further, let $g_1, \ldots, g_r$ be all newforms in $S_2(\Gamma_0(Np))$. For $i = 1, \ldots, r$ let ℓ^{n_i} be the highest power of ℓ such that g_i is congruent to f modulo ℓ^{n_i} at almost all primes. Put $n := n_1 + \ldots + n_r$ and let c be the maximum integer such that $P_{f,p}$ and $X^2 - (p+1)^2$ have a root in common modulo ℓ^c.

Question 4.2 *Is n equal to the ℓ-valuation of c?*

An inequality (in a greater generality) is provided by Theorem 2 of [D].

References

[ARS] A. Agashe, K. Ribet, W. Stein. *The Modular Degree, Congruence Primes and Multiplicity One*, 2009, to appear in a volume in honor of Serge Lang.

[C] H. Carayol. *Formes Modulaires et Représentations Galoisiennes à valeurs dans un Anneau Local complet*. Contemporary Mathematics **165** (1994), 213–237.

[CKR] I. Chen, I. Kiming, J. B. Rasmussen. *On Congruences mod $\mathfrak{p}^m$ Between Eigenforms and Their Attached Galois Representations*. Journal of Number Theory, in press, 2010.

[D] F. Diamond. *Congruence primes for cusp forms of weight $k \geq 2$*, in *Courbes modulaires et courbes de Shimura (Orsay, 1987/1988)*, Astérisque **196–197** (1991), 205–213.

[DT] L. Dieulefait, X. Taixés i Ventosa. *Congruences between modular forms and lowering the level mod ℓ^n*. Journal de Théorie des Nombres de Bordeaux **21** (2009), no.1, 109–118.

[FPR] D. Ford, S. Pauli, X.-F. Roblot. *A Fast Algorithm for Polynomial Factorization over $\mathbb{Q}_p$*. J. Théor. Nombres Bordeaux **14** (2002), no. 1, 151–169.

[M1] B. Mazur. *Modular curves and the Eisenstein ideal.* Inst. Hautes Études Sci. Publ. Math. No. 47 (1977), 33–186 (1978).

[M2] B. Mazur. *An introduction to the deformation theory of Galois representations*, in *Modular Forms and Fermat's Last Theorem*. Tata Institute of Fundamental Research Studies in Mathematics, Springer, New York, 1997, 243–311.

[Magma] W. Bosma, J.J. Cannon, C. Playoust. *The Magma Algebra System I: The User Language.* J. Symbolic Comput. **24** (1997), 235–265.

[Pohst] M. Pohst. *A note on index divisors.* in Computational number theory (Debrecen, 1989), 173–182, de Gruyter, Berlin, 1991.

---

[1]This remark was made in the Problem Book for the MSRI Modular Forms Summer Workshop organised by William Stein in 2006.

[R] K. A. Ribet. *Raising the levels of modular representations.* Séminaire de Théorie des Nombres, Paris 1987–88, 259–271, Progr. Math., 81, Birkhäuser Boston, Boston, MA, 1990.

[S] W. Stein. *Explicitly Computing with Modular Forms*. Graduate Studies in Mathematics, American Math Society, 2007.

[Sturm] J. Sturm. *On the congruence of modular forms.* Number theory (New York, 1984–1985), 275–280, Lecture Notes in Math., 1240, Springer, Berlin, 1987.

[T] X. Taixés i Ventosa. *Theoretical and algorithmic aspects of congruences between modular Galois representations.* PhD thesis, Universität Duisburg-Essen, 2009.

[W1] G. Wiese. *Dihedral Galois Representations and Katz Modular Forms*. Documenta Math. **9** (2004), 123–133.

[W2] G. Wiese. *On modular symbols and the cohomology of Hecke triangle surfaces.* International Journal of Number Theory (2009) **5(1)**, 89–108.