Dimension Variation of Classical and p-adic Modular Forms

Daqing Wan

Abstract: A quadratic bound is obtained for a conjecture of Gouvêa-Mazur on arithmetic variation of dimensions of classical and p-adic modular forms.

1. Introduction

Let $p \geq 5$ be a fixed prime and let N be a fixed positive integer not divisible by p. Let K be a finite extension field of the p-adic rational numbers $\mathbf{Q}_p$ and let B be the ring of integers in K. For a non-negative integer k, denote by $\mathbf{M}_k(Np, B)$ (resp. $\mathbf{M}_k(Np, K)$) the space of classical modular forms of weight k on $\Gamma_1(N) \cap \Gamma_0(p)$ whose Fourier coefficients at infinity belong to B (resp. K). This is a finitely generated free B-module, whose rank is denoted by $\mathbf{d}(k)$. The space $\mathbf{M}_k(Np, K)$ is a K-vector space of dimension $\mathbf{d}(k)$.

The Atkin U operator acts on $\mathbf{M}_k(Np, B)$. Its effect on q-expansion is given by

$$U(\sum_n a_n q^n) = \sum_n a_{np} q^n.$$

The characteristic polynomial $\mathbf{P}_k(t)$ of U acting on $\mathbf{M}_k(Np, B)$ has coefficients in B. One can factor $\mathbf{P}_k(t)$ over B in terms of the p-adic valuations of the characteristic roots (reciprocal roots) of $\mathbf{P}_k(t)$. Namely, one has the factorization

$$\mathbf{P}_k(t) = \prod_\alpha \mathbf{P}_k(t)^{(\alpha)},$$

where α runs over the set $\mathbf{Q}_{\geq 0}$ of non-negative rational numbers and each characteristic root ρ of $\mathbf{P}_k(t)^{(\alpha)}$ has p-adic valuation α: $\mathrm{ord}_p(\rho) = \alpha$. Denote the degree of $\mathbf{P}_k(t)^{(\alpha)}$ by $\mathbf{d}(k, \alpha)$. Then, we have the direct sum decomposition

$$\mathbf{M}_k(Np, K) = \bigoplus_{\alpha \in \mathbf{Q}_{\geq 0}} \mathbf{M}_k(Np, K)^{(\alpha)},$$

where the characteristic polynomial of U acting on the K-vector space $\mathbf{M}_k(Np, K)^{(\alpha)}$ is given by $\mathbf{P}_k(t)^{(\alpha)}$. The dimension of the K-vector space $\mathbf{M}_k(Np, K)^{(\alpha)}$ is $\mathbf{d}(k, \alpha)$.

Partially supported by NSF

A natural interesting question is how the dimension $\mathbf{d}(k,\alpha)$ varies as a function of the weight k, where the slope α is fixed. It is a conjecture of Gouvêa [Go1-2] and a recent theorem of Coleman [Co2] that the function $\mathbf{d}(k,\alpha)$ is p-adically locally constant for k in each residue class modulo $(p-1)$. Let $\mathbf{m}(\alpha)$ be the smallest non-negative integer (or $+\infty$ if it does not exist) such that whenever k_1 and k_2 are two positive integers $\geq 2\alpha+2$ satisfying

$$k_1 \equiv k_2 \pmod{p^{\mathbf{m}(\alpha)}(p-1)},$$

then we have the equality

$$\mathbf{d}(k_1,\alpha) = \mathbf{d}(k_2,\alpha).$$

Coleman's result then says that $\mathbf{m}(\alpha)$ is always finite. Based on numerical computations and motivated by Hida's work in the ordinary case $\alpha = 0$, Gouvêa-Mazur [GM1-2] conjectured an amazingly precise linear bound on $\mathbf{m}(\alpha)$. Recall that $\lceil x \rceil$ denotes the smallest integer which is at least as large as x.

Gouvêa-Mazur Conjecture: $\mathbf{m}(\alpha) \leq \lceil \alpha \rceil$.

The full strength of the conjecture, if correct, seems to be difficult to prove. It would involve a miraculous cancellation yet to be discovered. Our goal here is much more modest. We shall give a quadratic upper bound for $\mathbf{m}(\alpha)$. Namely,

Theorem 1.1. *There are three explicit constants A, B and C depending only on Np such that*

$$\mathbf{m}(\alpha) \leq A\alpha^2 + B\alpha + C.$$

The proof is based on Katz's theory [Ka] of overconvergent p-adic modular forms, Gouvêa-Mazur's continuity result [GM2] and Coleman's recent work [Co1-2] on classical and p-adic modular forms. The idea consists of two simple steps. Roughly speaking, the first step is to get a uniform quadratic lower bound for the Newton polygon of the characteristic polynomial $\mathbf{P}_k(t)$ as k varies. The second step is to transform the uniform quadratic lower bound for Newton polygons to a quadratic upper bound for the quantity $\mathbf{m}(\alpha)$.

Our method and result clearly extend to modular forms (or cusp forms) on $\Gamma_1(Np)$ with a given character at p. The present paper corresponds to the case when the character at p is trivial.

The quantity $\mathbf{m}(\alpha)$ contains important information about congruences of modular forms. It is also related to constructing p-adic Galois representations of the absolute Galois group of $\mathbf{Q}$. See [GM1], [GM3] and [Co2] for further conjectures, motivations and related results.

2. p-adic Modular forms

In this section, we recall various basic results on p-adic modular forms that will be needed for our proof. We shall not work in maximal generality. The reader is referred to [Ka], [Go1] and [Co1-2] for more systematic information.

Let $p \geq 5$ be a fixed prime number and let N be a fixed positive integer not divisible by p. Let B be the ring of integers in a finite extension K of the field $\mathbf{Q}_p$ of p-adic rational numbers. For a non-negative integer $k \geq 0$ and an element $r \in B$, let $M_k(N, B, r)$ * be the space of r-overconvergent p-adic modular forms of weight k on $\Gamma_1(N)$ whose q-expansions have coefficients in B. The space $M_k(N, B, r)$ can be concretely described in terms of the Katz expansion of a p-adic modular form.

Let E_{p-1} be the classical Eisenstein series of weight $p-1$. Let $\mathbf{M}_k(N, B)$ denote the space of classical modular forms on $\Gamma_1(N)$ whose q-expansions have coefficients in B. This is a free B-module of finite rank. For each integer $j > 0$, Katz showed that there is a free B-submodule $\mathbf{W}_j(N, B)$ of $\mathbf{M}_{k+j(p-1)}(N, B)$ such that

$$\mathbf{M}_{k+j(p-1)}(N, B) = E_{p-1} \cdot \mathbf{M}_{k+(j-1)(p-1)}(N, B) \oplus \mathbf{W}_j(N, B). \tag{2.1}$$

We simply set $\mathbf{W}_0(N, B) = \mathbf{M}_k(N, B)$. Fix such a submodule for each $j \geq 0$. Any r-overconvergent modular form $f \in M_k(N, B, r)$ can then be written in a unique way in the form (the Katz expansion of f):

$$f = \sum_{j \geq 0} r^j \frac{b_j}{E_{p-1}^j}, \quad b_j \in \mathbf{W}_j(N, B), \quad \lim_{j \to \infty} b_j = 0, \tag{2.2}$$

where $b_j \longrightarrow 0$ means that the q-expansions of b_j are more and more divisible by p as j goes to infinity. Conversely, every element f of the form (2.2) is in $M_k(N, B, r)$. It is clear that if $r = r_1 r_2$ in B, then we have an inclusion

$$M_k(N, B, r) \longrightarrow M_k(N, B, r_1).$$

In the case that r is a unit in B, the space $M_k(N, B, r)$ reduces to the space of p-adic modular forms studied by Serre [Se2]. This space is very large. From spectral point of view for the U operator, one needs to restrict to a smaller subspace such as $M_k(N, B, r)$ with $\mathrm{ord}_p(r) > 0$. This is the truly

* Convention: Classical objects are denoted by boldface symbols and their p-adic analogues are denoted by normal type.

overconvergent case. The space $M_0(N,B,r)$ is a commutative noetherian ring. Each $M_k(N,B,r)$ is a finite $M_0(N,B,r)$-module.

There are two important operators (the Frobenius operator F and its one sided inverse called the U operator) acting on the full space $M_k(N,B,1)$ of p-adic modular forms. On q-expansion, the Frobenius operator F takes the form

$$F(\sum_n a_n q^n) = \sum_n a_n q^{pn}.$$

The Frobenius operator reduces the rate of overconvergence and thus in general is not stable on $M_k(N,B,r)$ if $\mathrm{ord}_p(r) > 0$. For our purpose, we are interested in the quotient

$$G(q) = \frac{E_{p-1}(q)}{E_{p-1}(q^p)}.$$

This is easily seen to be a 1-unit in the full ring $M_0(N,B,1)$. The following more precise result and its proof are due to Coleman.

Lemma 2.1. *The function $G(q)$ is a 1-unit in the ring $M_0(N,B,r)$ for every r satisfying the inequality $0 \leq \mathrm{ord}_p(r) < 1/(p+1)$.*

Proof. Following Coleman, we shall use the notations from [Ka] for the proof of this lemma. We shall assume that $\mathrm{ord}_p(r) < p/(p+1)$. By the modular definition of the Frobenius operator ([Ka, line 10 on p114]) and the congruence equation (3.9.1) ([Ka, p121]), we have

$$E_{p-1}(q)^{p-1} E_{p-1}(q^p) \equiv (\frac{E_{p-1}(q)}{\lambda})^{p-1} E_{p-1}(q)^p \pmod{p/r},$$

where we claim that

$$(\frac{E_{p-1}(q)}{\lambda})^{p-1} \equiv 1 \pmod{p/r^2}. \tag{2.3}$$

Cancelling the factor $E_{p-1}(q)^{p-1}$ from the first equation, we deduce

$$E_{p-1}(q^p) \equiv (\frac{E_{p-1}(q)}{\lambda})^{p-1} E_{p-1}(q) \pmod{p/r^p}.$$

This and (2.3) together show that

$$E_{p-1}(q^p) \equiv E_{p-1}(q) \pmod{p/r^p}.$$

It follows that

$$G(q) = \frac{E_{p-1}(q)}{E_{p-1}(q^p)} \equiv 1 \pmod{p/r^{p+1}}.$$

This proves that $G(q)$ is a 1-unit if $\mathrm{ord}_p(r^{p+1}) < 1$, which is the lemma.

As for the claim in (2.3), Katz ([Ka, p114]) pointed out that a simple tangent calculation shows that λ and $E_{p-1}(q)$ are essentially equal. Using the notations there, the quantity λ is defined by

$$\check{\pi}^*(\omega) = \lambda\omega' \tag{2.4}$$

for a nowhere vanishing differential ω'. Since $\mathrm{ord}_p(p/r) > 0$, we can reduce modulo p/r and work in characteristic p. Assume then $p/r = 0$, by the definition of the Hasse invariant $A(E,\omega)$ on [Ka, p97],

$$F^*_{abs}(\eta) = A(E,\omega)\eta,$$

where η is the dual of ω in $H^1(E, O_E)$ and F^*_{abs} is the p-linear map on $H^1(E, O_E)$ induced by the p-th power map on the characteristic p structure sheave O_E of E. The dual of the last equation is given (see [Ka, p115]) by

$$\check{\pi}^*(\omega) = A(E,\omega)\omega^{(p)}. \tag{2.5}$$

We first choose ω' such that its reduction modulo p/r is $\omega^{(p)}$. Then, (2.4) and (2.5) show

$$\lambda \equiv A(E,\omega) \pmod{p/r}.$$

Since $A(E,\omega) \equiv E_{p-1}(q) \pmod{p}$, we deduce

$$\frac{E_{p-1}(q)}{\lambda} \equiv 1 \pmod{p/r^2}$$

which is stronger than (2.3). For a different choice of ω', one always has the weaker result in (2.3) because of an extra $(p-1)$-th root of unity modulo p/r. . The proof of Lemma 2.1 is complete.

The U operator

$$U(\sum_n a_n q^n) = \sum_n a_{np} q^n$$

is stable on the full space $M_k(N, B, 1)$ but in general not stable on the smaller space $M_k(N, B, r)$ for non-unit r. Define a p-adic Banach space over K by

$$M_k(N, K, r) = M_k(N, B, r) \otimes K,$$

where $M_k(N, B, r)$ is considered to be the unit ball in $M_k(N, K, r)$. Then, for $0 < \mathrm{ord}_p(r) < 1/(p+1)$, one has

$$pU\big(M_k(N, B, r)\big) \subset M_k(N, B, r^p). \tag{2.6}$$

Composing with the completely continuous inclusion

$$\iota : \ M_k(N, B, r^p) \longrightarrow M_k(N, B, r),$$

one deduces the following result for $U = U \circ \iota$.

Lemma 2.2. *Assume that* $0 < \mathrm{ord}_p(r) < p/(p+1)$. *Then the* U *operator induces a completely continuous endomorphism of the* p*-adic Banach space* $M_k(N, K, r)$.

In the rest of the paper, we shall assume that $0 < \mathrm{ord}_p(r) < p/(p+1)$ unless otherwise specified. Since U is completely continuous, the usual p-adic spectral theory [Se1] applies. In particular, the Fredholm determinant $\det(1 - tU | M_k(N, K, r))$ is well defined and is a p-adic entire function with coefficients in B. It is called the characteristic series of U acting on $M_k(N, K, r)$, denoted by $P_k(t)$. This entire series can be factorized over B in terms of the slopes of its Newton polygon. Write

$$P_k(t) = \prod_{\alpha \in \mathbf{Q}_{\geq 0}} P_k(t)^{(\alpha)},$$

where $P_k(t)^{(\alpha)} \in B[t]$ is the slope α part of $P_k(t)$. Denote the degree of $P_k(t)^{(\alpha)}$ by $d(k, \alpha)$. Then, for each non-negative rational number α, we have the spectral U-equivariant decomposition

$$M_k(N, K, r) = F_\alpha \oplus M_k(N, K, r)^{(\alpha)},$$

where $M_k(N, K, r)^{(\alpha)}$ denotes the generalized eigenspace of U whose eigenvalues have p-adic valuation α. The characteristic polynomial of U acting on the slope α subspace $M_k(N, K, r)^{(\alpha)}$ is given by $P_k(t)^{(\alpha)}$. The dimension of the K-vector subspace $M_k(N, K, r)^{(\alpha)}$ is $d(k, \alpha)$. The connection between classical modular forms and p-adic modular forms is made clear in Coleman [Co1], resolving a stronger version of a conjecture of Gouvêa [Go2]. In particular, it implies

Lemma 2.3. *If* $k \geq \alpha + 2$, *then* $\mathbf{d}(k, \alpha) = d(k, \alpha)$.

We are thus led to investigating how the dimension $d(k, \alpha)$ varies as a function of the weight k, where the slope α is fixed. Again, it is a conjecture of Gouvêa and a recent theorem of Coleman [Co2] that the function $d(k, \alpha)$ is p-adically locally constant for k in each residue class modulo $(p - 1)$. Let $m(\alpha)$ be the smallest non-negative integer (or $+\infty$ if it does not exist) such that whenever k_1 and k_2 are two non-negative integers satisfying

$$k_1 \equiv k_2 \pmod{p^{m(\alpha)}(p-1)},$$

then we have the equality

$$d(k_1,\alpha) = d(k_2,\alpha).$$

Coleman's result then says that $m(\alpha)$ is always finite. A precise linear bound about $m(\alpha)$ is also conjectured in [GM2].

p-adic Gouvêa-Mazur Conjecture: $m(\alpha) \leq \lceil \alpha \rceil$.

Write

$$P_k(t) = \sum_{n=0}^{\infty} a_n(k) t^n.$$

Generalizing earlier result of Dwork-Katz, Coleman [Co2] showed that $P_k(t)$ can be expressed as an infinite product of certain unit root zeta functions arising from the modular family of elliptic curves. An immediate consequence of this formula is the following continuity result, first proved by Gouvêa-Mazur [GM2] using more elementary methods.

Lemma 2.4. *If k_1 and k_2 are two non-negative integers satisfying*

$$k_1 \equiv k_2 \pmod{p^m(p-1)},$$

then for each $n \geq 0$, we have

$$a_n(k_1) \equiv a_n(k_2) \pmod{p^{m+1}}.$$

By Coleman's result (Lemma 2.3), Theorem 1.1 is reduced to the following p-adic version.

Theorem 2.5. *There are three explicit constants A, B and C depending only on Np such that*

$$m(\alpha) \leq A\alpha^2 + B\alpha + C.$$

As explained in the introduction, the proof consists of two steps, which will be given in the next two sections. Section 3 gives a uniform quadratic lower bound for the Newton polygon of $P_k(t)$. Section 4 gives a general reciprocity lemma which shows that the two Newton polygons of $P_{k_1}(t)$ and $P_{k_2}(t)$ coincide for all sides with small slopes if $P_{k_1}(t)$ and $P_{k_2}(t)$ are also congruent modulo a suitable power of p. The needed congruence is provided by Lemma 2.4.

3. A uniform lower bound for Newton polygons

In this section, we show that for a fixed k, there is a uniform quadratic lower bound for the Newton polygon of the characteristic series $P_{k+j(p-1)}(t)$ as j varies. We shall now fix a non-negative integer k. For instance, we may assume that $0 \leq k < p-1$.

For $j \geq 0$, let d_j be the rank of the free B-module $\mathbf{M}_{k+j(p-1)}(N, B)$ and let m_j be the rank of the free B-submodule $\mathbf{W}_j(N, B)$ in (2.1). It is clear that $m_0 = d_0$ and $m_j = d_j - d_{j-1}$ for $j > 0$. Let $a_n(k + j(p-1))$ be the coefficient of t^n in $P_{k+j(p-1)}(t)$.

Lemma 3.1. *For all $j \geq 0$, we have the following uniform lower bound. If $d_\ell \leq n < d_{\ell+1}$ for some $\ell \geq 0$, then*

$$\mathrm{ord}_p\big(a_n(k + j(p-1))\big) \geq \frac{p-1}{p+1}\Big(\sum_{u=0}^{\ell} u m_u + (\ell+1)(n - d_\ell)\Big) - n. \tag{3.1}$$

Proof. Using the Katz expansion, one checks easily that the multiplication map

$$E_{p-1}^j : \ M_k(N, K, r) \longrightarrow M_{k+j(p-1)}(N, K, r)$$

is an isomorphism of p-adic Banach spaces. Since conjugate operators have the same characteristic series, it follows that

$$P_{k+j(p-1)}(t) = \det\big(I - tE_{p-1}^{-j} \circ U \circ E_{p-1}^j | M_k(N, K, r)\big).$$

As the operator U is Frobenius linear, we have

$$E_{p-1}^{-j}(q) \circ U = U \circ E_{p-1}^{-j}(q^p).$$

This shows that

$$\begin{aligned} P_{k+j(p-1)}(t) &= \det\Big(I - tU \circ \big(\frac{E_{p-1}(q)}{E_{p-1}(q^p)}\big)^j | M_k(N, K, r)\Big) \\ &= \det\big(I - tU \circ G^j | M_k(N, K, r)\big), \end{aligned}$$

as first observed by Coleman [Co2]. This formula reduces the calculation of $P_{k+j(p-1)}(t)$ to the calculation of the characteristic series of the completely continuous operator $U \circ G^j$ acting on the constant space $M_k(N, K, r)$ independent of j.

Now, the quadratic uniform lower bound in Lemma 3.1 can be obtained in a standard manner by working with an orthonormal basis of $M_k(N, K, r)$. Fix r in B such that $0 < \mathrm{ord}(r) < 1/(p+1)$ (this can be done by going to a finite extension of B if necessary). For each i, let $\{b_{i,1}, \cdots, b_{i,m_i}\}$ be a basis of the free B-module $\mathbf{W}_i(N, B)$. The elements

$$e_{i,s} = \frac{r^i}{E_{p-1}^i} b_{i,s}$$

form an orthonormal basis of $M_k(N, K, r)$. Write

$$U \circ G^j(e_{i,s}) = \sum_{u,v} A_{i,s}^{u,v}(j) e_{u,v},$$

where the $A_{i,s}^{u,v}(j)$ are elements in K. Since G^j is in $M_0(N,B,r)$, multiplication by G^j is stable on the unit ball $M_k(N,B,r)$ of $M_k(N,K,r)$. By (2.6), the operator pU maps $M_k(N,B,r)$ to the smaller space $M_k(N,B,r^p)$, where $0<\mathrm{ord}_p(r)<1/(p+1)$. We deduce that

$$\begin{aligned} U\circ G^j(e_{i,s}) &= \frac{1}{p}\sum_u \frac{(r^p)^u}{E_{p-1}^u} b_u(i,s,j) \\ &= \frac{1}{p}\sum_u r^{(p-1)u}(\frac{r}{E_{p-1}})^u b_u(i,s,j), \end{aligned}$$

where $b_u(i,s,j)\in \mathbf{W}_u(N,B)$. This immediately gives the bound

$$\mathrm{ord}_p\big(A_{i,s}^{u,v}(j)\big)\geq u(p-1)\mathrm{ord}_p(r)-1.$$

A standard determinant argument shows that for $d_\ell\leq n<d_{\ell+1}$, we have

$$\mathrm{ord}_p\big(a_n(k+j(p-1))\big)\geq (p-1)\mathrm{ord}_p(r)\big(\sum_{u=0}^{\ell} um_u+(\ell+1)(n-d_\ell)\big)-n.$$

As this estimate is true for any rational number r (extending B if necessary) satisfying $0<\mathrm{ord}_p(r)<1/(p+1)$. The proof is complete.

Remarks. In all non-trivial cases (namely $d_j\to\infty$), the lower bound in (3.1) is easily seen to be quadratic. As we have only finitely many initial k ($0\leq k<p-1$), we obtain a uniform quadratic lower bound for the Newton polygons.

A non-uniform quadratic lower bound for the Newton polygon is given in Gouvêa [Go1]. One gets a quadratic bound because the ring $M_0(N,K,r)$ is of dimension 1 (a curve). If one works with a completely continuous operator on an overconvergent ring of dimension d (>0), one would get a lower bound for the Newton polygon which grows like $x^{1+1/d}$. In the curve case $d=1$, the exponent reduces to 2.

4. A reciprocity lemma

In this section, we show how to transform a uniform lower bound for the Newton polygon of $P_k(t)$ (now $0\leq k<\infty$) into an upper bound for $m(\alpha)$. For possible future applications, we state and prove the result for more general growth functions. The special quadratic case needed for this paper is given explicitly in Corollary 4.2. In particular, Theorem 2.5 follows immediately from Corollary 4.2, Lemma 2.4 and Lemma 3.1.

Lemma 4.1. *Let $Q_1(t)$ and $Q_2(t)$ be two elements in $\mathbf{Z}_p[[t]]$ with $Q_1(0) = Q_2(0) = 1$. Let $N_i(x)$ be the function on $\mathbf{R}_{\geq 0}$ whose graph is the Newton polygon of $Q_i(t)$. Assume that $\nu(x)$ is a strictly increasing continuous function on $\mathbf{R}_{\geq 0}$ such that*

$$\nu(0) \leq 0, \ N_i(x) \geq x\nu(x) \ (1 \leq i \leq 2, x \geq 1), \ \lim_{x \to \infty} \nu(x) = \infty.$$

Assume further that the function $x\nu^{-1}(x)$ is increasing on $\mathbf{R}_{\geq 0}$, where $\nu^{-1}(x)$ denotes the inverse function of $\nu(x)$ defined at least on $\mathbf{R}_{\geq 0}$ as $\nu(0) \leq 0$. For $x \geq 0$, we define the integer valued increasing function

$$m_\nu(x) = [x\nu^{-1}(x)].$$

If the congruence

$$Q_1(t) \equiv Q_2(t) \ (\mathrm{mod}\ p^{m_\nu(\alpha)+1})$$

holds for some $\alpha \geq 0$, then the two Newton polygons $N_i(x)$ coincide for all the sides with slopes at most α.

Proof. If γ is the slope of a side of the Newton polygon $N_i(x)$ with end-vertex at $(n, N_i(n))$ for some integer $n > 0$, it is clear that

$$\gamma \geq \frac{N_i(n)}{n} \geq \frac{n\nu(n)}{n} \geq \nu(n). \tag{4.1}$$

Since the inverse function $\nu(x)$ is also increasing, we have

$$n = \nu^{-1}(\nu(n)) \leq \nu^{-1}(\gamma).$$

This and (4.1) together imply that

$$N_i(n) \leq \gamma n \leq \gamma\nu^{-1}(\gamma). \tag{4.2}$$

By our assumption, the function $x\nu(x)$ grows faster than a linear function. Thus, both $Q_1(t)$ and $Q_2(t)$ are p-adic entire functions. Let $\alpha_1, \cdots, \alpha_s$ be the finitely many non-negative rational numbers which are at most α and which occur as a slope of one of the two Newton polygons $N_i(x)$. Since $m_\nu(x)$ is increasing for $x \geq 0$, using induction we may assume that the two polygons $N_i(x)$ coincide for all the sides with slopes at most α_{s-1}. We need to prove that the two Newton polygons $N_i(x)$ also coincide for the side with slope α_s. Without loss of generality, we may assume that α_s

occurs as a slope of $N_1(x)$ with endpoint at $(n_s, N_1(n_s))$. Let β be the next slope of $N_2(x)$ after α_{s-1}. If such β does not exist, we define $\beta = \infty$. It is clear that $\beta \geq \alpha_s$.

By (4.2) and the fact that $m_\nu(x)$ is increasing, we have

$$N_1(n_s) \leq \alpha_s \nu^{-1}(\alpha_s) < m_\nu(\alpha_s) + 1 \leq m_\nu(\alpha) + 1. \tag{4.3}$$

Write

$$Q_i(t) = 1 + \sum_{n=1}^{\infty} a_n(i) t^n.$$

Our congruence assumption shows that

$$a_{n_s}(1) \equiv a_{n_s}(2) \pmod{p^{m_\nu(\alpha)+1}}. \tag{4.4}$$

Since $(n_s, N_1(n_s))$ is a vertex of $N_1(x)$, we have

$$\mathrm{ord}_p(a_{n_s}(1)) = N_1(n_s). \tag{4.5}$$

It follows from (4.3)-(4.5) that

$$\mathrm{ord}_p(a_{n_s}(2)) = N_1(n_s).$$

This shows that $\beta \leq \alpha_s$ and thus $\beta = \alpha_s$. Furthermore, the side of slope α_s of $N_2(x)$ is at least as long as the side of slope α_s of $N_1(x)$. By a symmetric argument, we conclude that $N_1(x)$ and $N_2(x)$ have the same side of slope α_s. The proof is complete.

Corollary 4.2. *Let $Q_1(t)$ and $Q_2(t)$ be two elements in $\mathbf{Z}_p[[t]]$ with $Q_1(0) = Q_2(0) = 1$. Let $N_i(x)$ be the function on $\mathbf{R}_{\geq 0}$ whose graph is the Newton polygon of $Q_i(t)$. Suppose that there are three real numbers $a(>0)$, b and c such that for $1 \leq i \leq 2$,*

$$N_i(x) \geq ax^2 + bx + c \ (x \geq 1).$$

For $x \geq 0$, we define

$$m'(x) = [\frac{x(x + |b| + |c|)}{a}].$$

If the congruence

$$Q_1(t) \equiv Q_2(t) \pmod{p^{m'(\alpha)+1}}$$

holds for some $\alpha \geq 0$, then the two Newton polygons $N_i(x)$ coincide for all the sides with slopes at most α.

Proof. Taking $v(x) = ax - |b| - |c|$, one checks that $\nu(x)$ satisfies the conditions in Lemma 4.1 and that $xv^{-1}(x) = x(x + |b| + |c|)/a$. We can thus take $m'(\alpha) = m_\nu(\alpha)$ as in Lemma 4.1.

Remarks. Lemma 4.1 may be viewed as some sort of reciprocity law. It transforms a uniform lower bound for the Newton polygons to an upper bound for $m(\alpha)$. For instance, if the uniform lower bound in Lemma 3.1 were given by a polynomial of degree d ($d > 1$), then Lemma 4.1 shows that $m(\alpha)$ would be bounded by $O(\alpha^{1+1/(d-1)})$. In the case $d = 2$, the exponent $1 + \frac{1}{d-1}$ reduces to 2. However, I would not hope for a general uniform lower bound for Newton polygons which is better than quadratic in our current modular form situation.

On the other hand, for a global L-function of Drinfeld modules [TW], there is a trick (that works only in characteristic p situation) that can be used to improve the quadratic convergence of the characteristic series to exponential convergence. See [W1] for an explicit exponential lower bound in a special case. The same characteristic p trick also implies a drastic improvement for the continuity result in characteristic p situation. This improved continuity result and Lemma 4.1 then give an upper bound of the type $O(\log \alpha)$ for the characteristic p analogue of the quantity $m(\alpha)$. Lemma 4.1 is also useful for explicit study of the distribution of zeroes and poles of the L-functions attached to a certain family of p-adic representations of the arithmetic fundamental group of an algebraic variety over a finite field as studied in [W2].

We return to the original p-adic Gouvêa-Mazur conjecture. As indicated above, there is no reason to expect a qualitative improvement for the uniform quadratic lower bound of the Newton polygons. If the full form of the conjecture is true and if our arguments could be made to work, one would need to have a quadratic improvement of the continuity result. Namely, one would need to have something like: $P_{k_1}(t) \equiv P_{k_2}(t) \pmod{p^{Q(m)}}$ whenever $k_1 \equiv k_2 \pmod{p^m(p-1)}$, where $Q(m)$ is a quadratic polynomial instead of being linear. Such a result would involve a great cancellation in computing the coefficients of the characteristic series $P_k(t)$.

Another possible obstruction to the full form of the Gouvêa-Mazur conjecture is the problem about independence in p. Even in the special case $N = 1$ as studied in next section, the author does not know how to remove the p factor from our quadratic bound of $m(\alpha)$ if it is removable.

5. An example

We now work out the various constants involved in the special case $N = 1$ and $p \equiv 1 \pmod{12}$. No attempt is made to optimize the constants.

We shall restrict to each congruence class modulo $(p-1)$. For this purpose, let k be a fixed integer with $0 \le k < p - 1$. We may assume that k is even, as $M_k(N, K, r) = 0$ for odd k. Let

$m(\alpha, k)$ be the smallest non-negative integer such that for all non-negative integers k_1 and k_2 in the residue class k modulo $(p-1)$ satisfying $k_1 \equiv k_2 (\mathrm{mod}\ p^{m(\alpha,k)}(p-1))$, we have the equality $d(k_1, \alpha) = d(k_2, \alpha)$. It is clear that

$$m(\alpha) = \max_{0 \le k < p-1} m(\alpha, k).$$

Assume $N = 1$, $p \equiv 1 \pmod{12}$, k even and $0 \le k < p-1$. It is well known that

$$m_0 = d_0 = \begin{cases} [k/12] + 1, & \text{if } k \not\equiv 2 (\mathrm{mod}\ 12), \\ [k/12], & \text{if } k \equiv 2 (\mathrm{mod}\ 12). \end{cases} \tag{5.1}$$

For $j > 0$, one has $d_j = m_0 + j(p-1)/12$ and $m_j = (p-1)/12$. If n is a positive integer such that

$$d_\ell \le n < d_{\ell+1}$$

for some $\ell \ge 0$, then we solve ℓ in terms of n and get

$$\ell + 1 > \frac{12(n - m_0)}{p-1}.$$

One checks that

$$\begin{aligned} \ell\frac{p-1}{24} + n - d_\ell &= n - m_0 - \ell\frac{p-1}{24} \\ &= \frac{n-m_0}{2} + \frac{n-m_0}{2} - \ell\frac{p-1}{24} \\ &\ge \frac{n-m_0}{2}. \end{aligned}$$

These inequalities and Lemma 3.1 together show that

$$\begin{aligned} \mathrm{ord}_p\big(a_n(k+j(p-1))\big) &\ge \frac{p-1}{p+1}\Big(\sum_{u=1}^{\ell} u\frac{p-1}{12} + (\ell+1)(n-d_\ell)\Big) - n \\ &= \frac{p-1}{p+1}(\ell+1)\big(\ell\frac{p-1}{24} + n - d_\ell\big) - n \\ &\ge \frac{6(n-m_0)^2}{p+1} - n. \end{aligned}$$

This is true for all $n \ge m_0$. Removing the non-negative constant term, one sees that the slightly weaker inequality

$$\mathrm{ord}_p\big(a_n(k+j(p-1))\big) \ge \frac{6}{p+1}n^2 - \big(\frac{12m_0}{p+1} + 1\big)n$$

holds for all $n \ge 0$, noting that the quadratic polynomial of the right hand side is actually negative for $1 \le n \le m_0$. By Corollary 4.2, we deduce the bound

$$m(\alpha, k) \le \frac{p+1}{6}\alpha^2 + \big(2m_0 + \frac{p+1}{6}\big)\alpha, \tag{5.2}$$

where m_0 is given in (5.1). This immediately implies a crude but simple bound for $m(\alpha)$:

$$m(\alpha) \leq \frac{p+1}{6}\alpha^2 + \frac{p}{3}\alpha. \tag{5.3}$$

6. Bounding the denominators of slopes

Another result from Coleman [Co2] is that the set of slopes of all classical modular forms on $\Gamma_1(N)$ (and arbitrary weight) is a discrete subset of the real numbers. This can be made more precise. For a real number $x \geq 0$, let $c(N, p; x)$ be the largest denominator among the rational numbers $\alpha \leq x$ which occur as slopes of classical modular forms (of any weight) on $\Gamma_1(N)$. Then, $c(N, p; x)$ is bounded by a linear polynomial of x. In fact, the proof of Corollar 4.2 and the second inequality in the proof of Lemma 4.1 show that

$$\begin{aligned} c(N, p, x) &\leq \nu^{-1}(x) \\ &= (x + |b| + |c|)/a. \end{aligned} \tag{6.1}$$

We believe that something much stronger is true. Namely,

Conjecture 6.1. *The number $c(N, p; x)$ is bounded by a constant depending only on N and p.*

We have no real evidence for it. But it fits into the general p-adic Riemann Hypothesis philosophy which we believe and which has been checked to be true in a non-trivial function field situation [W1].

Acknowledgements. We thank R. Coleman for valuable communications, G. Stevens for helpful conversations and the referee for useful comments on the proof of Lemma 4.1.

References

[Co1] R. Coleman, Classical and overconvergent modular forms, Invent. Math., 124(1996), 215-241.

[Co2] R. Coleman, p-Adic Banach spaces and families of modular forms, Invent. Math., 127(1997), 417-479.

[Go1] F. Gouvêa, Arithmetic of p-adic modular forms, Lecture Notes in Math., Vol. 1304, 1988.

[Go2] F. Gouvêa, Continuity properties of modular forms, in Elliptic Curves and Related Topics, CRM Proceedings and Lecture Notes, AMS, 4(1994), 85-99.

[GM1] F. Gouvêa and B. Mazur, Families of modular eigenforms, Math. Comp., 58(1992), 793-806.

[GM2] F. Gouvêa and B. Mazur, On the characteristic power series of the U operator, Ann. Inst. Fourier, Grenoble, 43, 2(1993), 301-302.

[GM3] F. Gouvêa and B. Mazur, Searching for p-adic eigenforms, Math. Research Letter, 2(1995), 515-536.

[Ka] N. Katz, p-Adic properties of modular schemes and modular forms, in Modular Forms in One Variable III (SLN 350), Springer-Verlag, 1973, 69-190.

[Se1] J-P. Serre, Endomorphismes complètement continus des espaces de Banach p-adiques, Publ. Math., IHES, 12(1962), 69-85.

[Se2] J-P. Serre, Formes modulaires et fonctiones zeta p-adiques, in Modular Forms in One Variable III (SLN 350), Springer-Verlag, 1973, 191-268.

[TW] Y. Taguchi and D. Wan, L-functions of φ-sheaves and Drinfeld modules, J. Amer. Math. Soc., 9(1996), 755-781.

[W1] D. Wan, On the Riemann hypothesis for the characteristic p zeta function, J. Number Theory, Vol. 58, 1(1996), 196-212.

[W2] D. Wan, Meromorphic continuation of L-functions of p-adic representations, Ann. Math., 143(1996), 469-498.

Department of Mathematics, University of California, Irvine, CA 92697-3875

E-mail: dwan@math.uci.edu