# ON THE THETA OPERATOR FOR MODULAR FORMS MODULO PRIME POWERS

IMIN CHEN, IAN KIMING

Abstract. We consider the classical theta operator $\theta$ on modular forms modulo $p^m$ and level $N$ prime to $p$ where $p$ is a prime greater than 3. Our main result is that $\theta$ mod $p^m$ will map forms of weight $k$ to forms of weight $k+2+2p^{m-1}(p-1)$ and that this weight is optimal in certain cases when $m$ is at least 2. Thus, the natural expectation that $\theta$ mod $p^m$ should map to weight $k+2+p^{m-1}(p-1)$ is shown to be false.

The primary motivation for this study is that application of the $\theta$ operator on eigenforms mod $p^m$ corresponds to twisting the attached Galois representations with the cyclotomic character. Our construction of the $\theta$-operator mod $p^m$ gives an explicit weight bound on the twist of a modular mod $p^m$ Galois representation by the cyclotomic character.

## 1. Introduction

Let $p$ be a prime number. We shall assume $p \geq 5$ throughout the paper in order to avoid certain technicalities when $p$ is 2 or 3.

Further let $m \in \mathbb{N}$ and denote by $M_k(N, \mathbb{Z}_p)$ the $\mathbb{Z}_p$-module of modular forms of weight $k$ for $\Gamma_1(N)$ over $\mathbb{Z}_p$ and let $M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ be the $\mathbb{Z}/p^m\mathbb{Z}$-module of modular forms for $\Gamma_1(N)$ over $\mathbb{Z}/p^m\mathbb{Z}$ as defined classically by $M_k(N, R) = M_k(N, \mathbb{Z}) \otimes R$. Note that this definition relies on the existence of an integral structure on $M_k(N, \mathbb{C})$ (see for instance [1]).

Let $k_1, \ldots, k_t$ be a collection of weights and let $f_i \in M_{k_i}(N, \mathbb{Z}_p)$. The $q$-expansion of an element in a direct sum of the $M_{k_i}(N, \mathbb{Z}_p)$'s or $M_{k_i}(N, \mathbb{Z}/p^m\mathbb{Z})$'s is defined by extending linearly on each component. When we write $f_1 + \ldots + f_t \equiv 0 \pmod{p^m}$, we shall mean that the $q$-expansion $f_1(q) + \ldots + f_t(q)$ lies in $p^m\mathbb{Z}_p[[q]]$. Similarly for $f_i \in M_{k_i}(N, \mathbb{Z}/p^m\mathbb{Z})$, we write $f_1 + \ldots + f_t \equiv 0 \pmod{p^m}$ if the $q$-expansion of $f_1 + \ldots + f_t$ equals 0 in $(\mathbb{Z}/p^m\mathbb{Z})[[q]]$. In such a case, we say that $f_1 + \ldots + f_t$ is congruent to 0 modulo $p^m$.

Let us recall the definition and basic properties of the standard Eisenstein series on $\mathrm{SL}_2(\mathbb{Z})$, cf. §1 of [13], for instance: the series

$$G_k := -\frac{B_k}{2k} + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

with $B_k$ the $k$-th Bernoulli number and $\sigma_t(n) := \sum_{d|n} d^t$ the usual divisor sum, is (with $q := e^{2\pi i z}$) for $k$ an even integer $\geq 4$ a modular form on $\mathrm{SL}_2(\mathbb{Z})$. Defining $E_k$ as the normalization

$$E_k := -\frac{2k}{B_k} \cdot G_k$$

one has $E_k \equiv 1 \pmod{p^t}$ when (and only when) $k \equiv 0 \pmod{p^{t-1}(p-1)}$.

There are natural inclusions (preserving $q$-expansions)

$$M_k(N,\mathbb{Z}/p^m\mathbb{Z}) \hookrightarrow M_{k+p^{m-1}(p-1)}(N,\mathbb{Z}/p^m\mathbb{Z}),$$

induced by multiplication by $E_{p-1}^{p^{m-1}}$, using the fact that $E_{p-1}^{p^{m-1}} \equiv 1 \pmod{p^m}$. Note that $E_{p^{m-1}(p-1)} = E_{p-1}^{p^{m-1}}$ in $M_k(N,\mathbb{Z}/p^m\mathbb{Z})$, again by the $q$-expansion principle, so that the map can also be seen as induced by multiplication by $E_{p^{m-1}(p-1)}$.

As is well-known, when we specialize the above series for $G_k$ to $k=2$ and define

$$G_2 := -\frac{B_2}{4} + \sum_{n=1}^{\infty} \sigma_1(n)q^n = -\frac{1}{24} + \sum_{n=1}^{\infty} \sigma_1(n)q^n$$

then $G_2$ does not represent a modular form in the usual sense, but does so in the $p$-adic sense, cf. [13], §2. One defines $E_2$ as the normalization of $G_2$, i.e., $E_2 := -24G_2$. Thus, $E_2$ is also a $p$-adic modular form.

Consider the classical theta operator $\theta f = \frac{k}{12}E_2 + \frac{1}{12}\partial f$ of Ramanujan. Its effect on $q$-expansions is $\sum_n a_n q^n \mapsto \sum_n n a_n q^n$. Since $E_2$ is a $p$-adic modular form it is for any $m \in \mathbb{N}$ congruent modulo $p^m$ to a classical modular form of some weight. Thus we have $E_2 \equiv E_{p+1} \pmod{p}$, for example, and since we classically know that $\partial$ maps modular forms of weight $k$ to modular forms of weight $k+2$, one obtains the classical operator $\theta$ that maps $M_k(N,\mathbb{F}_p)$ to $M_{k+p+1}(N,\mathbb{F}_p)$. Studying this operator as well as its interaction with the 'weight filtration' (see below) is a key tool in the theory of modular forms modulo $p$; cf. for instance Jochnowitz' proof of finiteness of systems of Hecke eigenvalues mod $p$ across all weights in [10], or Edixhoven's results on the optimal weight in Serre's conjectures [6].

As we have launched a framework for the study of modular forms mod $p^m$ in [2] and [1], it is natural to ask whether a $\theta$ operator with similar properties can be defined on such forms. We have discussed in [1] how to attach Galois representations to eigenforms mod $p^m$, and it is clear from the properties of those attached representations that applying the $\theta$ operator corresponds on the Galois side to twisting by the cyclotomic character mod $p^m$. Hence, the construction of the theta operator mod $p^m$ yields an immediate application to mod $p^m$ Galois representations (see Corollary 1).

Notice that Serre shows in [13, Théorème 5] that there exists a $\theta$ operator on $p$-adic modular forms (of level 1) whose effect on $q$-expansions is $\sum_n a_n q^n \mapsto \sum_n n a_n q^n$ and that sends a form of ($p$-adic) weight $k$ to a form of weight $k+2$. One can view our results as finding a partially explicit expression with explicit weights for the mod $p^m$ reduction of this operator.

Hence, our results are then that on the one hand an extension of the $\theta$ operator from the mod $p$ to the mod $p^m$ situation is indeed possible, but that the interplay of the $\theta$ operator with the weights of the forms becomes much more complicated when $m>1$ and that, in fact, there are certain genuine qualitative differences between the case $m=1$ and the general cases $m>1$. Let us explain in detail.

We show that a $\theta$ operator on modular forms mod $p^m$ can be defined such that $\theta$ maps $M_k(N,\mathbb{Z}/p^m\mathbb{Z}) \rightarrow M_{k+k(m)}(N,\mathbb{Z}/p^m\mathbb{Z})$ with $k(m) := 2+2p^{m-1}(p-1)$, such that the effect on $q$-expansions is $\sum_n a_n q^n \mapsto \sum_n n a_n q^n$, and such that $\theta$ satisfies simple commutation rules with Hecke operators $T_\ell$ for primes $\ell \neq Np$, cf. the first part of Theorem 1 below. The proofs use a number of results from [13] plus the observation that $f \mid V \equiv f^p \pmod{p}$ where $V$ is the classical $V$ operator.

Define the weight $w_{p^m}(f)$ of an modular form $f$ mod $p^m$ with $f \not\equiv 0 \pmod{p}$ to be the smallest $k \in \mathbb{Z}$ such that $f$ is congruent modulo $p^m$ to an element of $M_k(N, \mathbb{Z}/p^m\mathbb{Z})$. A classical fact, crucial for instance in the work [10], is that when $m = 1$ we have $w_p(\theta f) \leq w_p(f) + p + 1$ with equality if (and only if) $p \nmid w_p(f)$.

One might expect the generalization of this to be that $w_{p^m}(\theta f) \leq w_{p^m}(f) + 2 + p^{m-1}(p-1)$ (perhaps with equality in some cases). However, as the second part of Theorem 1 shows, this is false:

**Theorem 1.** *Let $p \geq 5$ be a prime. Put*

$$k(m) := 2 + 2p^{m-1}(p-1).$$

*(i) The classical theta operator $\theta$ induces an operator*

$$\theta : M_k(N, \mathbb{Z}/p^m\mathbb{Z}) \to M_{k+k(m)}(N, \mathbb{Z}/p^m\mathbb{Z})$$

*whose effect on $q$-expansions is $\sum a_n q^n \mapsto \sum n a_n q^n$.*

*(ii) If $\ell \neq p$ is a prime and $T_\ell$ denotes the $\ell$-th Hecke operator, then*

$$T_\ell \theta = \ell \cdot \theta T_\ell$$

*as linear maps $M_k(N, \mathbb{Z}/p^m\mathbb{Z}) \to M_{k+k(m)}(N, \mathbb{Z}/p^m\mathbb{Z})$.*

*(iii) Assume $N \geq 5$ is prime to $p$. Let $m \geq 2$ and $f \in M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ with $f \not\equiv 0 \pmod{p}$. Suppose further that $p \nmid k$ and $w_p(f) = k$. Then*

$$w_{p^m}(\theta f) = k + 2 + 2p^{m-1}(p-1).$$

The proof of the second part of the Theorem uses some results of [11], which are applicable to general level $N$ prime to $p$. In particular, a main point is that if we consider the Eisenstein series $E_{p-1}$ and $E_{p+1}$ as modular forms modulo $p$ in the sense of Katz then $E_{p-1}$ (Hasse invariant) has only simple zeros and no zero common with $E_{p+1}$; cf. [11, Remark on p. 57]. This point allows one to compute the weight filtration of the last term of the expression below for the $\theta$ operator mod $p^m$, which is the controlling term. We would like to thank Nadim Rustom for a useful discussion about this point.

We have in fact conducted an in-depth study of the relation between and $w_{p^m}(f)$ and $w_{p^m}(\theta f)$ in the case of level $N = 1$ and for $m = 2$. However, the results are rather complicated and will not be stated in this article.

For simplicity, we have stated results in this paper for modular forms with coefficients in $\mathbb{Z}_p$ and hence reductions with coefficients in $\mathbb{Z}/p^m\mathbb{Z}$. The above theorem however is valid for coefficients in $\overline{\mathbb{Z}/p^m\mathbb{Z}}$ (see e.g. [1], section 2.4 for a definition of this ring) using the same proofs.

An immediate consequence of Theorem 1 to Galois representations is the following. We use the notation and terminology from [1].

**Corollary 1.** *Let $p \geq 5$ be a prime, $\rho : G_\mathbb{Q} \to \mathrm{GL}_2(\overline{\mathbb{Z}/p^m\mathbb{Z}})$ be a residually absolutely irreducible Galois representation, and $\chi : G_\mathbb{Q} \to \mathbb{Z}/p^m\mathbb{Z}^\times$ be the reduction modulo $p^m$ of the $p$-adic cyclotomic character. Suppose $\rho \cong \rho_f$ for some weak eigenform $f \in S_k(N)(\overline{\mathbb{Z}/p^m\mathbb{Z}})$. Then $\rho \otimes \chi \cong \rho_g$ for some weak eigenform $g \in S_{k+k(m)}(N)(\overline{\mathbb{Z}/p^m\mathbb{Z}})$.*

*Proof.* Note first that $\theta$ maps cusp forms to cusp forms. Suppose $f \in S_k(\overline{\mathbb{Z}/p^m\mathbb{Z}})$ is a weak eigenform in the sense that $T_\ell f \equiv f(T_\ell) f \pmod{p^m}$ for all $\ell \nmid D$ for some integer $D$. Then $g = \theta f \in S_{k+k(m)}(\overline{\mathbb{Z}/p^m\mathbb{Z}})$ is a weak eigenform in the sense that $T_\ell g \equiv g(T_\ell)g \pmod{p^m}$ for all $\ell \nmid NDp$, and we have that $f(T_\ell)\chi(\ell) \equiv g(T_\ell) \pmod{p^m}$ for all $\ell \nmid Dp$. The hypothesis of residual absolute irreducibility then allows us to conclude that $\rho_f \otimes \chi \cong \rho_{\theta f}$. □

Regarding Corollary 1, the main result of the paper [3] implies an analogous statement about twisting with the mod $p^m$ reduction of the Teichmüller character, with some differences: In [3], they consider the twist of $f$ by the mod $p^m$ reduction of the Teichmüller character where $f$ is a strong eigenform (again using the terminology of [1]). One then finds a strong (and not merely weak) eigenform $g$ as in the Corollary, but apparently without control over the weight $k + k(m)$. The proof uses different methods (Coleman $p$-adic families of modular forms.)

## 2. The theta operator modulo prime powers

2.1. **Eisenstein series.** We shall now develop an explicit expression for the truncation modulo $p^m$ of the $p$-adic Eisenstein series $G_2$.

**Proposition 1.** *Let $m \in \mathbb{N}$. Define the positive even integers $k_0, \ldots, k_{m-1}$ as follows: If $m \geq 2$, define:*

$$k_j := 2 + p^{m-j-1}(p^{j+1} - 1) \quad \text{for } j = 0, \ldots, m-2$$

*and*

$$k_{m-1} := p^{m-1}(p+1)$$

*and define just $k_0 := p + 1$ if $m = 1$.*

*Then $k_0 < \ldots < k_{m-1}$ and there are modular forms $f_0, \ldots, f_{m-1}$, depending only on $p$ and $m$, of level 1 and of weights $k_0, \ldots, k_{m-1}$, respectively, that have rational $q$-expansions, satisfy $v_p(f_j) = 0$ for all $j$, and are such that*

$$G_2 \equiv \sum_{j=0}^{m-1} p^j f_j \pmod{p^m}$$

*as a congruence between $q$-expansions.*

*The form $f_{m-1}$ can be chosen to be $f_{m-1} = G_{p+1}^{p^{m-1}}$.*

When $m = 2$ we can, and will, be a bit more explicit:

**Proposition 2.** *We have:*

$$G_2 \equiv f_0 + p \cdot f_1 \pmod{p^2}$$

*with modular forms $f_0$ and $f_1$ of weights $2 + p(p-1)$ and $p(p+1)$, respectively, explicitly:*

$$G_2 \equiv G_{2+p(p-1)} + p \cdot G_{p+1}^p \pmod{p^2}.$$

It is amusing to note the following consequence of the Proposition: For $p \neq 2, 3$, we have the following congruence of Bernoulli numbers,

$$\frac{B_2}{2} \equiv \frac{B_{p(p-1)+2}}{p(p-1)+2} + p\frac{B_{p+1}}{p+1} \pmod{p^2}.$$

However, this can also be seen in terms of $p$-adic continuity of Bernoulli numbers (cf. [14], Cor. 5.14 on p. 61, for instance).

Before the proofs of these propositions we need a couple of preparations.

Let $k$ be an even integer $\geq 2$. Recall from [13] that if we choose a sequence of even integers $k_i$ such that $k_i \to \infty$ in the usual, real metric, but $k_i \to k$ in the

$p$-adic metric, then the sequence $G_{k_i}$ has a $p$-adic limit denoted by $G_k^*$. This series $G_k^*$ is a $p$-adic modular form of weight $k$. It does not depend on the choice of the sequence $k_i$. In particular, we can, and will, choose $k_i := k + p^{i-1}(p-1)$, because if we chose another $k_i' = k + \lambda p^{i-1}(p-1)$, where $\lambda \geq 2$, then $G_{k_i'} \equiv G_{k_i} \pmod{p^i}$ so that $G_{k_i'} = G_{k_i} E_{p-1}^{p(\lambda-1)}$ in $M_{k_i'}(\mathbb{Z}/p^i\mathbb{Z})$ is not essentially different.

**Lemma 1.** *Let $k$ be an even integer $\geq 2$ and assume $(p-1) \nmid k$. Let $t \in \mathbb{N}$. Then $G_k^* \equiv G_{k+p^{t-1}(p-1)} \pmod{p^t}$.*

*Proof.* Let $u, v \geq t$. We claim that $G_{k_u} \equiv G_{k_v} \pmod{p^t}$. Since the series $G_{k_i}$ converges $p$-adically to $G_k^*$, the claim clearly implies the Lemma.

If $u = v$ the claim is trivial, so suppose not, say $u < v$. Then $k_v - k_u = p^{v-1}(p-1) - p^{u-1}(p-1)$ is a multiple of $p^{t-1}(p-1)$, say $k_v - k_u = s \cdot p^{t-1}(p-1)$. We also have $k_v - k_u \geq 4$. Hence, we find that $G := G_{k_u} \cdot E^s_{p^{t-1}(p-1)}$ is a modular form of weight $k_v$, and we have $G_{k_u} \equiv G \pmod{p^t}$.

Now notice that, when $i \geq t$, we have:

$$\sigma_{k_i-1}(n) = \sum_{d|n} d^{k-1+p^{i-1}(p-1)} \equiv \sum_{\substack{d|n \\ p\nmid d}} d^{k-1} \pmod{p^t}$$

as $d^{p^{i-1}(p-1)} \equiv 1 \pmod{p^t}$ when $p \nmid d$ and $i \geq t$, and as $d^{p^{i-1}(p-1)} \equiv 0 \pmod{p^t}$ when $p \mid d$ and $i \geq t$ (because $p^{t-1}(p-1) \geq t$ as long as $p \neq 2$).

We conclude that the nonconstant terms of the series $G_{k_u}$ and $G_{k_v}$ are termwise congruent modulo $p^t$. The same is then true of the forms $G$ and $G_{k_v}$ that are both forms of weight $k_v$. Hence, the nonconstant terms of the form $(G - G_{k_v})/p^t$ are all $p$-integral. As $k_v \equiv k \not\equiv 0 \pmod{(p-1)}$, it follows from Théorème 8 of [12] that the constant term of this form is in fact also $p$-integral. Hence,

$$G_{k_u} \equiv G \equiv G_{k_v} \pmod{p^t}$$

as desired. □

Recall that the $V$ operator is defined on formal $q$-expansions as

$$\left(\sum a_n q^n\right) \mid V := \sum a_n q^{np}.$$

**Corollary 2.** *We have*

$$G_2 \equiv \sum_{j=0}^{m-1} p^j \cdot \left(G_{2+p^{m-j-1}(p-1)} \mid V^j\right) \pmod{p^m}$$

*as a congruence between formal $q$-expansions.*

*Proof.* Recall from [13], §2, the identity, valid for any even integer $k \geq 2$, that

$$G_k = G_k^* + p^{k-1}\left(G_k^* \mid V\right) + \ldots + p^{t(k-1)}\left(G_k^* \mid V^t\right) + \ldots.$$

The identity is first an identity of formal $q$-expansions, but then shows that $G_k$ is a $p$-adic modular form as $V$ acts on $p$-adic modular forms, cf. [13], §2.

If we specialize this identity to the case $k = 2$, reduce modulo $p^m$, and note that the previous Lemma applies since $(p-1) \nmid 2$, the claim follows immediately. □

In the next paragraph and lemma, we use the notation $M_k(\Gamma, F)$ to mean the $F$-module of modular forms of weight $k$ over $F$, where $\Gamma$ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and $F$ is a subring of $\mathbb{C}$ or $\mathbb{C}_p$.

We can also see the $V$ operator as an operator on modular forms: Suppose that $f \in M_k(\mathrm{SL}_2(\mathbb{Z}),\mathbb{C})$. Then $(f \mid V)(z) = f(pz)$, and as is well-known $f \mid V \in M_k(\Gamma_0(p);\mathbb{C})$. The proof of the next lemma is a simple application of section 3.2 of [13]. Recall that if $f \in M_k(\mathrm{SL}_2(\mathbb{Z}),\mathbb{Q}_p)$ is nonzero with $q$-expansion $\sum_n a_n q^n$ we define $v_p(f) := \min\{v_p(a_n) \mid \ n \in \mathbb{N}\}$ where $v_p(a_n)$ is the usual (normalized) $p$-adic valuation of $a_n$.

**Lemma 2.** *Let $f \in M_k(\mathrm{SL}_2(\mathbb{Z}),\mathbb{Q})$ and suppose $v_p(f) = 0$. Let $t \in \mathbb{N}$ and suppose that $s \in \mathbb{Z}_{\geq 0}$ is such that*

$$\inf(s+1, p^s+1-k) \geq t.$$

*Then there is $h \in M_{k+p^s(p-1)}(\mathrm{SL}_2(\mathbb{Z}),\mathbb{Q})$ with $v_p(h) = 0$, and such that*

$$f \mid V \equiv h \pmod{p^t}.$$

*Proof.* As we noted above, $f \mid V$ is a modular form of weight $k$ on $\Gamma_0(p)$. Since $f \mid V = \sum a_n q^{np}$ if $f = \sum a_n q^n$ we have $v_p(f \mid V) = 0$. Recall the Fricke involution for modular forms on $\Gamma_0(p)$ given by the action of the matrix

$$W = \left(\begin{smallmatrix} 0 & -1 \\ p & 0 \end{smallmatrix}\right).$$

Since

$$f \mid V = p^{-k/2} f \mid_k \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right),$$

(recall that the weight $k$ action is normalized so that diagonal matrices act trivially), since

$$\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ p & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} p & 0 \\ 0 & p \end{smallmatrix}\right),$$

and since $f$ is on $\mathrm{SL}_2(\mathbb{Z})$ we see that $f \mid VW = p^{-k/2} f$ so that $v_p(f \mid VW) = -k/2$.

Now let $E := E_{p-1}$ and put

$$g := E - p^{p-1}(E \mid V)$$

so that $g$ is a modular form of weight $p-1$ on $\Gamma_0(p)$. Then, if we put

$$f_s := \mathrm{tr}((f \mid V) \cdot g^{p^s})$$

for $s \in \mathbb{Z}_{\geq 0}$ where tr denotes the trace from $\Gamma_0(p)$ to $\mathrm{SL}_2(\mathbb{Z})$, it follows from section 3.2 of [13] that $f_s$ is a modular form on $\mathrm{SL}_2(\mathbb{Z})$ of weight $k + p^s(p-1)$ and rational $q$-expansion. Furthermore, Lemme 9 of [13] implies that

$$\begin{aligned} v_p(f_s - (f \mid V)) &\geq \inf(s+1, p^s+1+v_p(f \mid VW) - k/2) \\ &= \inf(s+1, p^s+1-k) \geq t. \end{aligned}$$

Hence, we can choose $h := f_s$. As $f \mid V \equiv h \pmod{p^t}$ and $v_p(f \mid V) = 0$, we must have $v_p(h) = 0$. □

*Proof of Proposition 1:* That the defined weights $k_0, \ldots, k_{m-1}$ satisfy $k_0 < \ldots < k_{m-1}$ is verified immediately.

Thus, starting with Corollary 2 we see that it suffices to show that for each $j \in \{0, \ldots, m-1\}$ there is a modular form $f_j$ of weight $k_j$ with rational $q$-expansion and $v_p(f_j) = 0$, and such that

$$G_{2+p^{m-j-1}(p-1)} \mid V^j \equiv f_j \pmod{p^{m-j}}.$$

If $m = 1$, $j = 0$ we just take $f_0 := G_{p+1}$, so assume $m \geq 2$. Then, if $j = m-1$, note that

$$G_{p+1} \mid V^{m-1} \equiv G_{p+1}^{p^{m-1}} \pmod{p}$$

so that we can take $f_{m-1} := G_{p+1}^{p^{m-1}}$ that is of weight $k_{m-1} = p^{m-1}(p+1)$.

So, suppose that $j \leq m-2$. We claim that for $r = 0, \ldots, j$ there is a modular form $g_r$ of weight $2 + p^{m-j-1}(p^{r+1} - 1)$, rational $q$-expansion with $v_p(g_r) = 0$, and such that

$$G_{2+p^{m-j-1}(p-1)} \mid V^r \equiv g_r \pmod{p^{m-j}}$$

which is the desired when $r = j$.

We prove the last claim by induction on $r$ noting that the case $r = 0$ is trivial. So, suppose that $r < j$ and that we have already shown the existence of a modular form $g_r$ as above. Notice that

$$p^{m-j+r} + 1 - (2 + p^{m-j-1}(p^{r+1} - 1)) = p^{m-j-1} - 1 \geq m - j$$

holds because $m - j \geq 2$ (we used here that $p > 2$). Thus we see that Lemma 2 applies (taking $s = m-j+r$) and shows the existence of a modular form $g_{r+1}$ with rational $q$-expansion and $v_p(g_{r+1}) = 0$, such that

$$g_r \mid V \equiv g_{r+1} \pmod{p^{m-j}},$$

and such that $g_{r+1}$ has weight

$$2 + p^{m-j-1}(p^{r+1} - 1) + p^{m-j+r}(p-1) = 2 + p^{m-j-1}(p^{r+2} - 1),$$

and we are done. □

**Remark 1.** *It is interesting to note that in the induction, the inequalities do not allow us to deal with the case $j = m-1$, but then we use the congruence $f \mid V \equiv f^p \pmod{p}$ to take care of the last term.*

*Proof of Proposition 2:* Again by Corollary 2 we have:

$$G_2 \equiv G_{2+p(p-1)} + p \cdot (G_{p+1} \mid V) \pmod{p^2}.$$

Noting again that $G_{p+1} \mid V \equiv G_{p+1}^p \pmod{p}$ so that

$$p \cdot (G_{p+1} \mid V) \equiv p \cdot G_{p+1}^p \pmod{p^2},$$

we are done. □

2.2. **Definition and properties of the $\theta$ operator.** Recall the classical $\theta$ operator acting on formal $q$-expansions as $q\frac{d}{dq}$, i.e.,

$$\theta\left(\sum a_n q^n\right) := \sum n a_n q^n,$$

and the operator $\partial$ defined by

$$\frac{1}{12}\partial f := \theta f - \frac{k}{12} E_2 \cdot f = \theta f + 2k G_2 \cdot f$$

when $f = \sum a_n q^n \in M_k(N, \mathbb{C})$ is a modular form of weight $k$ (we have $B_2 = \frac{1}{6}$ so that $E_2 = -24 G_2$). Then $\partial f$ is in $M_{k+2}(N, \mathbb{C})$, and $\partial$ defines a derivation on $M(N, \mathbb{C}) := \oplus_k M_k(N, \mathbb{C})$ (as follows by writing $\theta = \frac{1}{2\pi i} \cdot \frac{d}{dz}$ (as $q = e^{2\pi i z}$) and combining with the classical transformation properties of $E_2$ under the weight 2 action of $\mathrm{SL}_2(\mathbb{Z})$ given in (4))

The definition of $\partial$ implies that $\partial$ defines a derivation on $\oplus_k M_k(N, \mathbb{Z})$ and hence also on $M(N, \mathbb{Z}_p) := \oplus_k M_k(N, \mathbb{Z}_p)$.

*Proof of Theorem 1.* *(i)* Retain the notation of Proposition 1 so that

$$k_j := 2 + p^{m-j-1}(p^{j+1} - 1) \quad \text{for } j = 0, \ldots, m-2,$$

and

$$k_{m-1} := p^{m-1}(p+1).$$

Then $k_0 < \ldots < k_{m-1}$ and by Proposition 1 we have modular forms $f_0, \ldots, f_{m-1}$ (of level 1 and) of weights $k_0, \ldots, k_{m-1}$, respectively, that have rational $q$-expansions, satisfy $v_p(f_j) = 0$ for all $j$, and are such that

$$G_2 \equiv \sum_{j=0}^{m-1} p^j f_j \pmod{p^m}.$$

With $k(m) := 2 + 2p^{m-1}(p-1)$ one checks that each number $k(m) - k_j$ is a multiple of $p^{m-j-1}(p-1)$, say $k(m) = k_j + t_j \cdot p^{m-j-1}(p-1)$ for $j = 0, \ldots, m-1$.

Since $E_{p-1}^{p^{m-j-1}} \equiv 1 \pmod{p^{m-j}}$ we find that $p^j E_{p-1}^{p^{m-j-1}t_j} \equiv p^j \pmod{p^m}$, and so the above congruence for $G_2$ can also be written as

$$G_2 \equiv \sum_{j=0}^{m-1} p^j E_{p-1}^{p^{m-j-1}t_j} f_j \pmod{p^m}$$

where now each summand is a form of weight $k(m)$.

Hence, for any an element $f \in M_k(N, \mathbb{Z}_p)$ we find that

$$\theta f = \frac{1}{12}\partial f - 2kG_2 \cdot f \tag{1}$$

$$\equiv \frac{1}{12}E_{p-1}^{2p^{m-1}}\partial f - 2kf\sum_{j=0}^{m-1} p^j E_{p-1}^{p^{m-j-1}t_j} f_j \pmod{p^m} \tag{2}$$

$$=: \theta_{p^m} f \in M_{k+k(m)}(\mathbb{Z}_p) \tag{3}$$

where now each summand on the right hand side is an element of $M_{k+k(m)}(N, \mathbb{Z}_p)$. Thus the classical theta operator induces a linear map

$$\theta = \theta_{p^m} : M_k(N, \mathbb{Z}/p^m\mathbb{Z}) \to M_{k+k(m)}(N, \mathbb{Z}/p^m\mathbb{Z})$$

the effect of which on $q$-expansions is $\sum a_n q^n \mapsto \sum n a_n q^n$. We still denote this operator (by abuse of notation) by $\theta$, but later when we need to distinguish from $\theta := \frac{1}{2\pi i} \cdot \frac{d}{dz}$, we will denote it by $\theta_{p^m}$.

*(ii)* First assume for the prime $\ell$ that we have $\ell \nmid N$ (in addition to $\ell \neq p$). Recall that the diamond operator $\langle \ell \rangle_k$ on a modular form $f$ of weight $k$ is defined by $\langle \ell \rangle_k f = f \mid_k \gamma$ for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, with $c \equiv 0 \pmod{N}$, $d \equiv \ell \pmod{N}$. (Note that we write the action of $\langle \cdot \rangle_k$ from left, though it is given by the stroke operator which is from the right. This is fine because $(\mathbb{Z}/N\mathbb{Z})^\times$ is abelian). As $\ell \neq p$ the operator $\langle \ell \rangle_k$ induces a linear action on $M_k(N, \mathbb{Z}_p)$ (as follows from the well-known formula $\ell^{k-1}\langle \ell \rangle_k = T_\ell^2 - T_{\ell^2}$ and the fact that the Hecke operators $T_n$ preserve $M_k(N, \mathbb{Z})$), and hence also on $M_k(N, \mathbb{Z}/p^m\mathbb{Z})$.

Using $\frac{1}{12}\partial f = \theta f - \frac{k}{12}E_2 f$ as well as the transformation property of $E_2$ given by

$$(E_2 \mid_2 \gamma)(z) = E_2(z) + \frac{12}{2\pi i} c j(\gamma, z)^{-1} \tag{4}$$

(see for instance [5]), a computation shows that we have $\frac{1}{12}\partial \langle \ell \rangle_k f = \langle \ell \rangle_{k+2} \frac{1}{12}\partial f$ for $f \in M_k(N, \mathbb{C})$.

Now recall the above definition of $\theta_{p^m} : M_k(N, \mathbb{Z}/p^m\mathbb{Z}) \to M_{k+k(m)}(N, \mathbb{Z}/p^m\mathbb{Z})$, as well as the fact that the forms $E_{p-1}$ and the $f_j$ occurring in the definition all have level 1 and thus are fixed under the action of the operators $\langle \ell \rangle_{p-1}$ and $\langle \ell \rangle_{k_j}$, respectively. We deduce that:

$$\theta_{p^m} \langle \ell \rangle_k f = \langle \ell \rangle_{k+k(m)} \theta_{p^m} f$$

for all $f \in M_k(N, \mathbb{Z}/p^m\mathbb{Z})$.

Let now $f = \sum a_n q^n \in M_k(N, \mathbb{Z}/p^m\mathbb{Z})$. Then we have (with the usual convention that $a_{\frac{n}{\ell}} := 0$ if $\ell \nmid n$):

$$\begin{aligned}
\langle \ell \rangle_k f &= \sum b_n q^n \\
T_\ell f &= \sum \left( a_{\ell n} + \ell^{k-1} \sum b_{\frac{n}{\ell}} \right) q^n \\
\langle \ell \rangle_{k+k(m)} \theta_{p^m} f &= \theta_{p^m} \langle \ell \rangle_k f = \sum n b_n q^n \\
T_\ell \theta_{p^m} f &= \sum \left( \ell n a_{\ell n} + \ell^{k+k(m)-1} \frac{n}{\ell} b_{\frac{n}{\ell}} \right) q^n \\
&= \ell n \sum \left( a_{\ell n} + \ell^{(k-1)+k(m)-2} b_{\frac{n}{\ell}} \right) q^n \\
\ell \theta_{p^m} T_\ell f &= \ell n \sum \left( a_{\ell n} + \ell^{k-1} b_{\frac{n}{\ell}} \right) q^n.
\end{aligned}$$

For $\ell \mid N$, $\ell \neq p$, a similar calculation holds (the second term involving $b_{\frac{n}{\ell}}$ is omitted throughout).

Thus, we have that

$$T_\ell \theta_{p^m} f = \ell \theta_{p^m} T_\ell f$$

for all $f \in M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ using the fact that $k(m) - 2 = 2p^{m-1}(p-1)$ is divisible by $p^{m-1}(p-1)$ so that

$$\ell^{k(m)-2} \equiv 1 \pmod{p^m}.$$

*(iii)* Now suppose that $m \geq 2$, that $f \in M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ with $f \not\equiv 0 \pmod p$, and suppose further that $p \nmid k$ and $w_p(f) = k$.

Assume that we had $w_{p^m}(\theta f) < k + 2 + 2p^{m-1}(p-1) = k + k(m)$, i.e., that there exist a form $g \in M_{k'}(N, \mathbb{Z}/p^m\mathbb{Z})$ with $g = f$ as forms with coefficients in $\mathbb{Z}/p^m\mathbb{Z}$, and where $k' < k + k(m)$. We then know that

$$k' \equiv k + k(m) \pmod{p^{m-1}(p-1)},$$

say $k + k(m) = k' + t \cdot p^{m-1}(p-1)$ with $t \geq 1$ (see [1, Corollary 22]; note we use the fact that modular forms in $M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ can be lifted to classical modular forms over $\mathbb{Z}_p$ which is what is used in loc. cit. and that $N$ is prime to $p$). Putting

$$h := E_{p-1}^{p^{m-1}(t-1)} g$$

we find that

$$\theta f = E_{p-1}^{p^{m-1}} h$$

as an equality of forms in $M_{k+k(m)}(N, \mathbb{Z}/p^m\mathbb{Z})$. If we combine this with (1), we obtain:

$$2kp^{m-1} E_{p-1}^{t_{m-1}} f_{m-1} f = -E_{p-1}^{p^{m-1}} h + \frac{1}{12} E_{p-1}^{2p^{m-1}} \partial f - 2kf \sum_{j=0}^{m-2} p^j E_{p-1}^{p^{m-j-1} t_j} f_j.$$

If we now use the fact that $p \nmid k$, that $p$ is odd, and that, as is easily checked, we have

$$t_{m-1} = (k(m) - k_{m-1})/(p-1) = (p^m - 3p^{m-1} + 2)/(p-1) < p^{m-1},$$

as well as $t_{m-1} < p^{m-j-1} t_j$ for $j = 0, \ldots, m-2$, we deduce that

$$p^{m-1} E_{p-1}^{t_{m-1}} f_{m-1} f = E_{p-1}^{t_{m-1}+1} h'$$

for some $h' \in M_{k+k(m)-(p-1)(t_{m-1}+1)}(N, \mathbb{Z}/p^m\mathbb{Z})$. Hence we must have $h' \in p^{m-1} M_{k+k(m)-(p-1)(t_{m-1}+1)}(N, \mathbb{Z}/p^m\mathbb{Z})$, say $h' = p^{m-1}h''$, so that

$$E_{p-1}^{t_{m-1}} f_{m-1} f \equiv E_{p-1}^{t_{m-1}+1} h'' \pmod{p},$$

and hence

$$f_{m-1} f \equiv E_{p-1} h'' \pmod{p}.$$

It follows that

$$w_p(f_{m-1} f) < k + k(m) - t_{m-1}(p-1) = k + k_{m-1} = k + p^{m-1}(p+1).$$

Now recall (from Proposition 1) that $f_{m-1} = G_{p+1}^{p^{m-1}}$. As $G_{p+1} = -\frac{B_{p+1}}{2(p+1)} E_{p+1}$ with $\frac{B_{p+1}}{2(p+1)}$ invertible modulo $p$, we deduce

$$w_p(E_{p+1}^{p^{m-1}} f) < k + p^{m-1}(p+1).$$

However, as $w_p(f) = k$ by assumption, this conclusion contradicts the following general fact:

Suppose that $N \geq 5$ and let $0 \neq \phi \in M_\kappa(N, \mathbb{Z}/p\mathbb{Z})$ where $\kappa \geq 1$, but $\kappa \neq p$. Then, for $a \in \mathbb{N}$ we have:

$$w_p(E_{p+1}^a \phi) = w_p(\phi) + a(p+1).$$

To prove this general fact, by induction on $a$ it is clearly enough to prove the case $a = 1$. Hence, let us assume $a = 1$.

Assume without loss of generality that $w_p(\phi) = \kappa$. Assume for a contradiction that we had $w_p(E_{p+1}\phi) < \kappa + p + 1$. Then

$$E_{p+1}\phi = E_{p-1}\psi$$

for some $\psi \in M_{2+\kappa}(N, \mathbb{Z}/p\mathbb{Z})$.

Let $\mathcal{M}_k(N, R)$ denote the space of modular forms of weight $k$ on $\Gamma_1(N)$ with coefficients in $R$ as defined in [4] using Katz's definition. One has an injection

$$M_k(N, R) \to \mathcal{M}_k(N, R)$$

sending classical modular forms over $R$ to Katz modular forms over $R$.

Under the hypothesis $N \geq 5$, we have from [4, Theorem 12.3.7]) that:

(K1) $\mathcal{M}_k(N, \mathbb{Z}_p) \cong M_k(N, \mathbb{Z}_p)$ (as $\mathbb{Z}_p$ is flat over $\mathbb{Z}$)
(K2) $\mathcal{M}_k(N, \mathbb{Z}/p^m\mathbb{Z}) \cong M_k(N, \mathbb{Z}/p^m\mathbb{Z})$ if $k > 1$ and $N$ is prime to $p$.

Regard the above identity as an identity of Katz modular forms on $X_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$ and let us use some results from [11]: By the remark after Lemma 1 of [11], the forms $E_{p-1}$ and $E_{p+1}$ are without a common zero (the results in [11] are for modular forms on $X(N)$ which implies the result for $X_1(N)$).

But, by a theorem of Igusa [9], $E_{p-1}$ vanishes to order 1 at every supersingular point of $X_1(N)$ (see [8, second paragraph following (4.6)]). Hence the equality $E_{p+1}\phi = E_{p-1}\psi$ means that $\phi$ vanishes at every zero of $E_{p-1}$. Thus, we must have ($\kappa > p-1$ and) $\phi = E_{p-1}\eta$ for some $\eta \in \mathcal{M}_{\kappa-(p-1)}(N, \mathbb{Z}/p\mathbb{Z})$.

The $\eta$ cannot be classical in the sense that $\eta \in M_{\kappa-(p-1)}(N, \mathbb{Z}/p\mathbb{Z})$ or else $w_p(\phi) < \kappa$, contrary to hypothesis. Hence, $\eta$ is a non-classical modular form in the space $\mathcal{M}_{\kappa-(p-1)}(N, \mathbb{Z}/p\mathbb{Z})$ of Katz modular forms. By (K2), this can only happen if $\kappa - (p-1) = 1$, which means $\kappa = p$, contrary to hypothesis. □

## References

[1] I. Chen, I. Kiming, G. Wiese: 'On modular Galois representations modulo prime powers', Int. J. Number Theory **9** (2013), 91–113.

[2] I. Chen, I. Kiming, J. B. Rasmussen: 'On congruences mod $\mathfrak{p}^m$ between eigenforms and their attached Galois representations', J. Number Theory (3) **130** (2010), 608–619.

[3] R. Colemen and W. Stein: 'Approximation of eigenforms of infinite slope by eigenforms of finite slope', Geometric aspects of Dwork theory. Vol. I, II, 437449, Walter de Gruyter GmbH & Co. KG, Berlin, 2004.

[4] F. Diamond and J. Im: 'Modular forms and modular curves', Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994), CMS Conf. Proc. **17** (1995), 39–133. Amer. Math. Soc. 1995.

[5] F. Diamond and J. Shurman: 'A first course in modular forms', Grad. Texts in Math. **228**, Springer 2005.

[6] B. Edixhoven: 'The weight in Serre's conjectures on modular forms', Invent. Math. (3) **109** (1992), 563–594.

[7] B. Edixhoven: 'Serre's Conjecture', in Modular forms and Fermat's Last Theorem (Boston, MA, 1995), 209–242, Springer, New York, 1997.

[8] B. Gross: 'A tameness criterion for Galois representations associated to modular forms mod $p$', Duke Math. J. **61**, no.2 (1990), 445–517.

[9] J.-I. Igusa: 'Class number of a definite quaternion with prime discriminant', Proc. Natl. Acad. Sci. USA **44** (1958), 312–314.

[10] N. Jochnowitz: 'Congruences between systems of eigenvalues of modular forms', Trans. Amer. Math. Soc. **270** (1982), 269–285.

[11] N. M. Katz: 'A result on modular forms in characteristic $p$', in J.-P. Serre, D. B. Zagier (eds.): Modular functions of one variable V, Lecture Notes in Math. **601** (1977), 53–61.

[12] J.-P. Serre: 'Congruences et formes modulaires [d'après H. P. F. Swinnerton-Dyer]', Séminaire Bourbaki, Lecture Notes in Math. **317**, 319–338, Springer 1973.

[13] J.-P. Serre: 'Formes modulaires et fonctions zêta $p$-adiques, in W. Kuyk, J.-P. Serre: Modular functions of one variable III, Lecture Notes in Math. **350**, Springer 1973.

[14] L. C. Washington: 'Introduction to cyclotomic fields', Grad. Texts in Math. **83**, Springer 1982.

(Imin Chen) Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, B.C., V5A 1S6, Canada

*E-mail address*: `ichen@math.sfu.ca`

(Ian Kiming) Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

*E-mail address*: `kiming@math.ku.dk`