Invent. math. 127, 417–479 (1997)

Inventiones mathematicae

P-adic Banach spaces and families of modular forms

Robert F. Coleman

UC Berkeley, Department of Mathematics, Berkeley, CA 94720, USA

Oblatum 21-VII-1994 & 15-V-1996

This paper is dedicated to Bernard Dwork who has been a friend and an inspiration for many years.

0. Introduction

Let p be a prime, $\mathbf{C}_p$ the completion of an algebraic closure of the p-adic numbers $\mathbf{Q}_p$ and K a finite extension of $\mathbf{Q}_p$ contained in $\mathbf{C}_p$. Let v be the valuation on $\mathbf{C}_p$ such that $v(p) = 1$ and let $|\,|$ be the absolute value on $\mathbf{C}_p$ such that $|x| = p^{-v(x)}$ for $x \in \mathbf{C}_p$.

Suppose N is a positive integer prime to p. Let $X_1(Np)$ denote the modular curve over K which represents elliptic curves with $\Gamma_1(Np)$-structure and let U_p be the Hecke operator on modular forms on $X_1(Np)$ which takes a form with q-expansion $\sum_n a_n q^n$ to the modular form with q-expansion $\sum_n a_{np} q^n$.

A modular form F is said to have slope $\alpha \in \mathbf{Q}$ if there is a polynomial $R(T)$ over $\mathbf{C}_p$ such that $R(U_p)F = 0$ and such that the Newton polygon of $R(T)$ has only one side and its slope is $-\alpha$.

For simplicity of notation, now suppose p is odd.

Now $(\mathbf{Z}/Np\mathbf{Z})^*$ acts on modular forms on $\Gamma_1(Np)$ via the diamond operators and we identify $(\mathbf{Z}/p\mathbf{Z})^*$ with a subgroup of $(\mathbf{Z}/Np\mathbf{Z})^*$ in the natural way. If χ is a $\mathbf{C}_p^*$-valued character on $(\mathbf{Z}/p\mathbf{Z})^*$, we will say a modular form F is of $(\mathbf{Z}/p\mathbf{Z})^*$-character χ if $F \,|\, \langle d \rangle = \chi(d)F$ for $d \in (\mathbf{Z}/p\mathbf{Z})^*$. Let $\tau : (\mathbf{Z}/p\mathbf{Z})^* \to \mu(\mathbf{Q}_p)$ denote the Teichmüller character.

We prove in Sect. B3:

Theorem A. *Suppose $\alpha \in \mathbf{Q}$ and $\varepsilon : (\mathbf{Z}/p\mathbf{Z})^* \to \mathbf{C}_p^*$ is a character. Then there exists an $M \in \mathbf{Z}$ which depends only on p, N, ε and α with the following property: If $k \in \mathbf{Z}$, $k > \alpha + 1$ and there is a unique normalized cusp form F on $X_1(Np)$ of weight k, $(\mathbf{Z}/p\mathbf{Z})^*$-character $\varepsilon\tau^{-k}$ and slope α and if $k' > \alpha + 1$ is an integer congruent to k modulo p^{M+n}, for any non-negative integer n, then there exists a unique normalized cusp form F' on $X_1(Np)$ of weight*

k', $(\mathbf{Z}/p\mathbf{Z})^*$-character $\varepsilon\tau^{-k'}$ and slope α. Moreover, this form satisfies the congruence

$$F'(q) \equiv F(q) \bmod p^{n+1} .$$

Both F and F' must be eigenforms for the full Hecke algebra of the respective weight because these algebras are commutative and therefore preserve the space of forms of a given slope. It is (a slight generalization of) a conjecture of Gouvêa–Mazur [GM-F] that M may be taken to be zero as long as $n \geqq \alpha$. When $\alpha = 0$ this is a theorem of Hida [H-GR]. In this paper, we obtain no information about M (except in one example, discussed in Appendix II). However, using recent results of Daqing Wan, we have been able to give an upper bound, quadratic in α, on the minimal allowable M for fixed N and p. We are also able to obtain results in the case when there exists more than one normalized form of a given slope, character and level in Sect. B5. That is, we prove, Corollary B5.7.1, the existence of what Gouvêa and Mazur call "**R**-families" in [GM-F].

For example, let Δ be the unique normalized weight 12, level 1, cusp form. Write

$$\Delta(q) = \sum_{n \geqq 1} \tau(n) q^n .$$

Then $\tau(7) = -7.2392$. The above theorem implies, for any positive integer k divisible by 6 and close enough 7-adically to 12, that there exists a unique normalized weight k, level 1, cuspidal eigenform F_k over $\mathbf{Q}_p$ such that $F_k \mid T_7 = a(k)F_k$ for some $a(k) \in \mathbf{Z}_7$ with valuation 1. Moreover, for any positive integer n, if k is sufficiently large and close 7-adically to 12,

$$F_k(q) \equiv \Delta(q) - \rho\Delta(q^7) \bmod 7^n$$

where ρ is the root of $X^2 - \tau(7)X + 7^{11}$ in $\mathbf{Z}_7$ with valuation 10. (We prove a similar, more precise, statement for $p = 2$ in Appendix II.)

The following is one important ingredient in the proof of Theorem A:

Theorem B. *For integers $0 \leqq i < p-1$ there exist series $P_{N,i}(s,T) \in \mathbf{Z}_p[[s,T]]$ which converge for all T and s such that $|s| < p^{(p-2)/(p-1)}$ such that for integers k, $P_{N,i}(k,T)$ is the characteristic series of Atkin's U-operator acting on overconvergent forms of weight k and $(\mathbf{Z}/p\mathbf{Z})^*$-character τ^{i-k}.*

This considerably strengthens the main result of [GM-CS].

Now let $M_{k,cl}$ denote the space of classical modular forms of weight k on $X_1(Np)$ defined over K. Then the eigenvalues of U_p on $M_{k,cl}$ have valuation at most $k-1$. For a character ε on $(\mathbf{Z}/p\mathbf{Z})^*$ we also let $M_{k,cl}(\varepsilon)$ denote the subspace of forms of weight k and $(\mathbf{Z}/p\mathbf{Z}^*)$-character ε and set $\mathbf{d}(k,\varepsilon,\alpha)$ equal to the dimension of the subspace of $M_{k,cl}(\varepsilon\tau^{-k})$ consisting of forms of slope α.

As a corollary of Theorem 8.1 of [C-CO] we obtain:

Theorem C. *With notation as above, the set of zeroes of $P_{N,i}(k,T^{-1})$ in $\mathbf{C}_p^*$ with valuation strictly less than $k-1$ is the same as the set of eigenvalues* *with valuation strictly less than* $k-1$ *of* U_p *acting on* $M_{k,cl}(\tau^{i-k})$ *(counting multiplicities in both cases).*

We are able to deduce from this, in Sect. B3, another result conjectured in more precise form by Gouvêa and Mazur:

Theorem D. *If* ε *is a* $\mathbf{C}_p^*$*-valued character on* $(\mathbf{Z}/p\mathbf{Z})^*$*, and* k *and* k' *are integers strictly bigger than* $\alpha+1$ *and sufficiently close p-adically*

$$\mathbf{d}(k,\varepsilon,\alpha)=\mathbf{d}(k',\varepsilon,\alpha)\,.$$

Moreover, the closeness sufficient for this equality only depends on α.

Wan's result implies a lower bound, quadratic in α, of how valuation of $k-k'$ must be for the equality in the above theorem to be true. Since $\mathbf{Z}_p$ is compact this implies what is called a "control theorem", that is, for a fixed $\alpha \in \mathbf{Q}$, the dimension of the space of forms of a given weight and slope α is bounded independently of the weight. We are also able to deal with the prime 2.

We show, in Sect. B3, that the set of slopes of modular forms on $\Gamma_1(N)$ (and arbitrary weight) is a discrete subset of the real numbers which is a consequence of the Gouvêa–Mazur conjectures. In the Appendix I, we show how to use the trace formulas of Eichler–Selberg and Reich–Monsky to prove that there exist overconvergent forms of any given integral weight of arbitrarily large slope. In a future article with Barry Mazur, we will begin to develop the connections between the results in this paper and p-adic representations of the Galois group of $\bar{\mathbf{Q}}/\mathbf{Q}$.

What foreshadows our proofs is the study of the k-th Hecke polynomial (see Eichler [E], Sato [Sa], Kuga [Ku] and Ihara [I]),

$$\det((1-T_pu+p^{k-1}u^2)\,|\,S_k)\,,$$

where S_k is the space of weight k cusp forms of level one defined over $\mathbf{C}$ and T_p is the p-th Hecke operator. It was used to relate the Ramanujan–Petersson conjecture to the Weil conjectures (see the Introduction to [I] for more history). Ihara applied the Eichler–Selberg trace formula to this effort. This line of research was continued by Morita, Hijikata and Koike ([M], [Hj], [Ko1] and [Ko2]). Dwork began another approach to the study of these polynomials using what is now known as the theory of overconvergent p-adic modular forms and also the Reich–Monsky trace formula ([D1] and [D2]). This work was continued by Katz [K] and Adolphson [A]. Both of these lines of research seem to have stopped in the mid-seventies. Hida [H1], [H2] developed to great utility the theory of "ordinary" modular forms which in this optic are modular forms of slope zero. In particular, he proved what may now be interpreted as the slope zero part of the aforementioned conjectures.

Our approach is a continuation of that introduced by Dwork. In fact, the inspiration for this paper arose in an attempt to interpret Dwork's paper

"On Hecke polynomials" [D1] in terms of the point of view developed in [C-CO]. The key idea in Gouvêa–Mazur's paper "On the characteristic series of the U operator" [GM-CS] provided the bridge between [D1] and [C-CO].

We will now give an outline of this paper. It naturally breaks into two parts.

In Part A, which we entitle "Families of Banach Spaces", we show how Serre's p-adic Banach–Fredholm–Riesz theory [S] works in a family, i.e., may be extended over complete, normed rings, which we call Banach algebras. We define and prove some basic results about these algebras, Banach modules over them, orthonormal bases for these modules and completely continuous maps between Banach modules in Sect. A1. In Sect. A2, we show that a completely continuous operator on a Banach module has a "Fredholm determinant" which behaves well under a contractive base change (such as the restriction to a residue field). We define the resultant of a monic polynomial and an entire series and prove some basic results about it in Sect. A3. This will be necessary for us to extend Serre's Riesz theory to this more general situation, in Sect. A4. I.e., given a factorization of the Fredholm determinant into relatively prime factors, one of which is polynomial with unit leading coefficient, we will be able to find in Theorem A4.3, a corresponding direct sum decomposition of the Banach module. In Sect. A5, we specialize our theory and consider Banach modules over reduced affinoid algebras (which are Banach algebras). In Subsect. A5.i, we show, Proposition A5.2, that a homomorphism of affinoid algebras over an affinoid algebra A can be interpreted as a completely continuous map of Banach modules over A, when the associated map between affinoid spaces over A is what we call "inner". We also indicate how our Riesz theory can be strengthened over an affinoid algebra. In Subsect. A5.ii, we prove our main technical result, Proposition 5.3, about quasi-finite morphisms from an affinoid to the closed unit disk. As a corollary of this result we may conclude that if Z is the zero locus of the characteristic power series of a completely continuous operator over the ring of rigid analytic functions on an affinoid disk B, then for each $z \in Z$ there exists an affinoid open neighborhood X of z in Z whose image Y in B is an affinoid disk and is such that the morphism from Z to Y is finite.

In Part B, which we entitle, "Families of Modular Forms", we apply the results of Part A to elliptic modular forms. In Sect. B1, we recall or derive some basic results on Eisenstein series and define the weight space $\mathcal{W}$ (which is the union of a finite number of open disks). The rings of rigid analytic functions on affinoid open subspaces of $\mathcal{W}$ will be our Banach algebras. In Sect. B2, we introduce the basic set up notation and explain how to extend the results of [C-CO] and [C-HCO] to all primes and levels. Sect. B3 is the heart of the paper. Multiplication by an appropriate Eisenstein series E_k of weight k gives an isomorphism from the space of overconvergent forms of weight 0 to the space of overconvergent forms of weight k. Thus one can study the U-operator on weight k forms, $U_{(k)}$ by studying a twist of the U-operator on weight 0 forms, $U_{(0)}$. The key observation is that this twist can be viewed as an "internal multiplication". I.e., there is an overconvergent rigid analytic function e_k so that, if F is a weight 0 form (i.e. a function),

$$(E_k)^{-1}U_{(k)}(E_kF) = U_{(0)}(e_kF)\,. \tag{1}$$

Moreover, these functions e_k vary analytically in k, for k in a subspace $\mathscr{W}^*$ of W (as we point out in Sect. B4, with a more judicious choice of E_k (and more work which we will carry out in another article [C-CPS]) one can replace the e_k with functions which vary analytically over all of $\mathscr{W}$). This will allow us to consider the family of operators (1) as one completely continuous operator on a Banach module over the rigid analytic functions on any affinoid disk in $\mathscr{W}^*$ (the ring of rigid analytic functions of $\mathscr{W}^*$ is not itself a Banach algebra.) This allows us to prove Theorem B3.2 and its refinement Theorem B3.3 (which is Theorem B above extended to the prime 2), and this implies that the Fredholm determinants of the U-operator acting on weight k overconvergent modular forms, for integers k, are specializations of a Fredholm determinant of a completely continuous operator over the Banach algebra of rigid analytic functions on any sufficiently large closed disk in $\mathscr{W}^*$. This, combined with Theorem C and the corollary to Proposition 5.3 discussed above, yields Theorem D. We are also able to prove Theorem A, as well as its extension to $p = 2$, Theorem B3.5, in this section. In Sect. B4, we give a definition of the q-expansion of an overconvergent modular form of non-integral weight as well as of a family of such objects which is forced on us by the considerations of Sect. B3 although we do not have a geometric interpretation of either. We also show that the q-expansions of Eisenstein series, introduced in Section B1, live in an overconvergent family. In Sect. B5, we define a Hecke algebra which acts on families of q-expansions of overconvergent modular forms and use it together with our Riesz theory and a basic duality result, Proposition 10.3, to prove a qualitative version of Gouvêa–Mazur's R-family conjecture. We discuss further results, including generalizations of some of our results to higher level and the connections of our families of modular forms with Galois representations, which will be proved elsewhere (eg. in [C-CPS] and [C-HCO]), in Sect. B6. Appendix I contains explicit formulas for the Fredholm determinants of our operators as well as a proof of the existence of infinitely many non-classical overconvergent eigenforms of any integral weight. We point out that we have not been able to prove any of Theorems A–D using these formulas. Finally, in Appendix II, we show, by considering the special case $p = 2$ and $N = 1$, how our general results combined with the explicit formulas of Appendix I can be used, in specific cases, to make the estimates in our theorems explicit.

We are grateful to the following people who provided comments helpful in writing and revising this paper: de Shalit, Iovita, Liu, Mazur, Serre, Stevens and Teitelbaum. We thank Brady Kahn for providing the illustration displayed in Sect. A5 and we also thank the referees whose careful reading of the manuscript and insightful remarks helped make the paper substantially more readable.

Contents

A. Families of Banach spaces

A1. Banach algebras and Banach modules

Suppose A is a commutative ring with a unit element, complete and separated with respect to a non-trivial ultrametric norm $|\,|$ (see [BGR, Sect. 1.2]).

I.e., $|1| = 1$,

$$|a+b| \leqq \operatorname{Max}\{|a|,|b|\}, \quad |ab| \leqq |a|\,|b|\,,$$

for a and $b \in A$, and moreover, $|a| = 0$ if and only if $a = 0$. We will call such a ring, a *Banach algebra*. We will call an ultrametrically normed complete module E over A, such that $|ae| \leqq |a|\,|e|$ if $a \in A$ and $e \in E$ a *Banach module* over A.

An element a in A is called *multiplicative* if $|ab| = |a|\,|b|$ for all $b \in A$. We say $|\,|$ is a *multiplicative norm* if every element in A is multiplicative (in [BGR] such a norm is called a valuation). Let A^m be the group of multiplicative units in A, A^0 denote the subring of A consisting of elements a such that $|a| \leqq 1$ and E^0 the A^0 submodule in E consisting of all e such that $|e| \leqq 1$. Let $\mathscr{I}(A)$ denote the set of finitely generated ideals I of A^0 such that $\{I^n : n \in \mathbf{Z}, n \geqq 0\}$ is a basis of open neighborhoods of the origin in A^0. We will suppose throughout this article that $\mathscr{I}(A) \neq \emptyset$. Clearly, if there exists an $a \in A^m$ such that $|a| < 1$, then $aA^0 \in \mathscr{I}(A)$. (We also point out that $(A^0)^* = \{a \in A^m : |a| = 1\}$.) Suppose N is closed submodule of E. Then by the induced norm $|\,|_N$ on E/N, we mean

$$|a \bmod N|_N = \operatorname{Inf}\{|b| : b - a \in N\}\,.$$

It is clear that E/N is complete with respect to this norm. We will also frequently make the following hypothesis:

Hypothesis M (for multiplicative).

$$|A^m| \cup \{0\} = |A| .$$

Since $|\ |$ is non-trivial, it follows from this hypothesis that $|A^m| \neq 1$.

Examples. (i) *The ring $A = \mathbf{Q}_p$ with its standard norm is a Banach algebra which satisfies hypothesis M while the subring $A^0 = \mathbf{Z}_p$ is a Banach algebra which does not.* (ii) *If A is a Banach algebra, we let $A\langle T_1, \ldots, T_n \rangle$ denote the ring of restricted power series over A, i.e., power series over A whose coefficients tend to zero in A with their degree. Then, if $F \in A\langle T_1, \ldots, T_n \rangle$, we set $|F|$ equal to the supremum of the absolute vales of the coefficients of F. (This is called the Gauss norm.) Then $A\langle T_1, \ldots, T_n \rangle$ is a Banach algebra with respect to this norm and satisfies hypothesis M if and only if A does.* (iii) *While we will later see many more examples of Banach algebras satisfying hypothesis M, one which does not and which will be very important for us in the future is the Iwasawa algebra, $\mathbf{\Lambda} := \mathbf{Z}_p[[\mathbf{Z}_p^*]]$. Some complete norms on $\mathbf{\Lambda}$ may be described as follows: Suppose $\kappa : \mathbf{Z}_p^* \to \mathbf{C}_p^*$ is a continuous character. Then κ extends by linearity and continuity to a continuous ring homomorphism of $\mathbf{\Lambda}$ into $\mathbf{C}_p$. For $0 < r < 1 \in \mathbf{R}$ and $\lambda \in \mathbf{\Lambda}$, let*

$$|\lambda|_r = \sup_{\kappa} |\kappa(\lambda)|$$

where κ runs over the set of continuous characters on $\mathbf{Z}_p^*$ with values in the closed ball of radius r around 1. For example, if $r \in |\mathbf{C}_p|$, $|1 - [1 + \mathbf{q}]|_r = r$. Then $|\ |_r$ is a complete multiplicative norm on $\mathbf{\Lambda}$, $\mathbf{\Lambda}^0 = \mathbf{\Lambda}$, $|\mathbf{\Lambda}^m|_r = 1$ and where $\mathbf{q} = 4$ if $p = 2$ and $\mathbf{q} = p$ otherwise $(p, 1 - [1 + \mathbf{q}]) \in \mathscr{I}(\mathbf{\Lambda})$. All these norms can be shown to be equivalent and induce the "adic" topology corresponding to the maximal ideal of $\mathbf{\Lambda}$, as we will verify in [C-CPS].

If K is a multiplicatively normed field and A is a K-algebra such that the structural morphism from K to A is an isometry onto its image, we will call A a *K-Banach algebra.*

An *orthonormal basis* for a Banach module E over A is a set $\{e_i : i \in I\}$ of elements of E, for some index set I, such that every element m in E can be written uniquely in the form $\sum_{i \in I} a_i e_i$ with $a_i \in A$ such that $\lim_{i \to \infty} |a_i| = 0$ (this means that for any $\varepsilon \in \mathbf{R}_{>0}$ there exist a finite subset S of I such that $|a_i| < \varepsilon$ for $i \in I - S$) and

$$|m| = \operatorname{Sup} \{|a_i| : i \in I\} .$$

We say E is *orthonormizable* if it has an orthonormal basis. Clearly, if E is orthonormizable $|E| = |A|$ and $|ae| = |a||e|$ if $a \in A$ is multiplicative.

Lemma A1.1. *If E is orthonormizable, and $I \in \mathscr{I}(A)$, $I^n E^0$ is a basis of open neighborhoods of the origin in E^0.*

Proof. For $\varepsilon \in \mathbf{R}$, $\varepsilon > 0$, let $E(\varepsilon) = \{e \in E : |e| < \varepsilon\}$. What we must prove is that the collections $\{E(\varepsilon)\}$ and $\{I^n E^0\}$ are cofinal. Since the collection$\{I^n\}_{n \geq 0}$ is a basis of neighborhoods of 0 in A^0, it is clear that if $\varepsilon > 0$, there exist an $n \in \mathbf{Z}$, $n \geqq 0$ such that $I^n E^0 \subseteq E(\varepsilon)$. Now suppose $n \in \mathbf{Z}$, $n \geqq 0$. Then there exists an $\varepsilon > 0$ such that $A(\varepsilon) \subseteq I^n$. Claim: $E(\varepsilon) \subset I^n E^0$. Suppose $a_1, \ldots, a_m$ generate I^n over A^0 and $\{e_i : i \in J\}$ is an orthonormal basis for E. Let $e \in E(\varepsilon)$. Write,

$$e = \sum_J b_i e_i \,.$$

Then $b_i \in I^n$ for all i. In fact, since $|b_i| \to 0$, $b_i \in I^{n+m(i)}$ where $m(i) \geqq 0$ and $m(i) \to \infty$ as $i \to \infty$, so we may write

$$b_i = \sum_{j=1}^{m} c_{ij} a_j$$

where $c_{ij} \in I^{m(i)}$. It follows that $|c_{ij}| \to 0$ as $i \to \infty$. Hence

$$e = \sum_{j=1}^{m} a_j \sum_{i \in J} c_{ij} e_i$$

and the inner sums converge in E. Thus $e \in I^n E^0$. □

Using this, we see that if $B = \{e_i\}$ is an orthonormal basis for E and $I \in \mathscr{I}(A)$ then the reduction of B modulo $I^n E^0$ is an algebraic basis for $E^0/I^n E^0$ over $A^0/I^n A^0$.

One has, using the same line of reasoning as in the proof of [S, Lemma 1].

Lemma A1.2. *Suppose $(A, |\,|)$ satisfies hypothesis M and $|A| = |E|$. Then a subset B of E is an orthonormal basis for E if and only if $B \subset E^0$ and the image of B in E^0/aE^0 is a basis of this module over A^0/aA^0 for some $a \in A^m$, $|a| < 1$.*

Suppose M and N are Banach modules over A. Then we put a semi-norm on $N \otimes_A M$ by letting $|c|$, for $c \in N \otimes_A M$, equal the greatest lower bound over all representations $c = \sum_i n_i \otimes m_i$ of

$$\mathrm{Max}_i\{|n_i||m_i|\} \,.$$

We then let $N \hat{\otimes}_A M$ denote the completion of $N \otimes_A M$ with respect to this semi-norm.

If B is a complete normed A algebra such that the structural morphism from A to B is contractive, then B is a Banach module over A and it is easy to see that $B \hat{\otimes}_A M$ is, naturally, a Banach module over B.

Proposition A1.3. *If $\{e_i : i \in I\}$ is an orthonormal basis for M over A, for some index set I, then $\{1 \otimes e_i : i \in I\}$ is an orthonormal basis for $B \hat{\otimes}_A M$ over B.*

Proof. First, every element n in the image of $B \otimes_A M$ in $B \hat{\otimes}_A M$ can be written in the form $\sum_{i \in I} b_i \otimes_A e_i$ where $b_i \in B$ and $b_i \to 0$. We claim $|n| = \mathrm{Sup}|b_i| =: P$. We have, for each $\varepsilon \in \mathbf{R}_{>0}$, there exists a finite subset T of I such that

$$||n| - |n_S|| < \varepsilon$$

for all finite subsets S of I containing T, where

$$n_S = \sum_S b_i \otimes e_i \, .$$

It follows that $|n| \leqq P$. Now, fix $j \in I$ and let $h : M \to A$ be the A homomorphism which takes $\sum_I a_i e_i$ to a_j. Then h is continuous, in fact $|h(m)| \leqq |m|$. Let h_B denote the extension by scalars of h to a morphism from $B \otimes_A M$ to B. Suppose S is a finite subset of I and $j \in S$. Then $h(n_S) = b_j$. If $n_S = \sum_i c_i \otimes m_i$ where $c_i \in B$ and $m_i \in M$. Then $b_j = h_B(n_S) = \sum_i c_i h_B(m_i)$. Hence, using the contractivity of $A \to B$,

$$|b_j| \leqq \mathrm{Max}_i\{|c_i|_B |h_B(m_i)|_B\} \leqq \mathrm{Max}_i\{|c_i|_B |h(m_i)|_A\} \leqq \mathrm{Max}_i\{|c_i|_B |m_i|\} \, .$$

Since this is true for all representations of n_S, all $j \in S$ and all S, it follows that $P \leqq |n|$. This establishes the claim. The proposition follows easily. □

If J is a closed ideal of A then we call the induced norm on A/J the *residual norm*. Then $(A/J, |\ |_J)$ is a Banach algebra. We set $E_J = (A/J)\hat{\otimes} E$. We note that, since A is complete, maximal ideals are automatically closed [BGR, 1.2.4/5]. Since the map $A \to A/J$ is contractive, as a corollary of the previous proposition, we obtain:

Corollary A1.3.1. *If J is a closed ideal of A and E is orthonormizable, $E_J = E/J$. Moreover, if $\{e_i : i \in I\}$ is an orthonormal basis for E, $\{e_i \bmod JE\}$ is an orthonormal basis for E_J over A/J.*

Proof. By the proposition, we know $\{1 \otimes e_i\}$ is an orthonormal basis for E_J over A/J. Clearly, $E/JE = (A/J) \otimes E \subseteq A/J \hat{\otimes} E$. On the other hand, if $x = \sum_{i \in I} a_i \hat{\otimes} e_i \in (A/J)\hat{\otimes} E$ where $|a_i|_J \to 0$, we can choose $\alpha_i \in A$ such that $\alpha_i \equiv a_i \bmod J$ and $|\alpha_i|_A \leqq 2|a_i|_k$. Hence $e := \sum_{i \in I} \alpha_i e_i \in E$. Let $y = 1 \otimes e \in E/JE$. Then $|x - y| < \varepsilon$ for every positive ε. Hence $x = y$ and so $E/JE = E_J$. The last part of the corollary follows immediately. □

Continuous homomorphisms. If M and N are Banach modules over A, and $L : M \to N$ is a continuous A-homomorphism we set

$$|L| = \sup_{m \neq 0} \frac{|L(m)|}{|m|} \, .$$

This determines a topology on the set of continuous A-homomorphisms. The homomorphism L is said to be *completely continuous* if

$$L = \lim_{j \to \infty} L_j$$

where L_j is a continuous A-homomorphism from M to N whose image is contained in a finitely generated submodule of N. If $f : M' \to M$ and $g : N \to N'$ are continuous A-homomorphisms of A-Banach modules then it is easy to see that $g \circ L \circ f$ is also completely continuous. Let $\mathscr{C}_A(M,N)$ denote the Banach module of completely continuous A-homomorphisms from M to N.

It is also easy to see:

Lemma A1.4. *If $A \to B$ is a contractive map of Banach algebras, M and N are Banach algebras over A and $L \in \mathscr{C}_A(M,N)$, then $1 \hat{\otimes} L \in \mathscr{C}_B(B \hat{\otimes} M, B \hat{\otimes} N)$.*

Remarks A1.5. (i) *When A is field, $|\ |$ is multiplicative and $|A^*| \neq \{1\}$ the above is the theory discussed by Serre [S]. However, Serre's theory works without change even when $|A^*| = \{1\}$. It's only easier. Indeed, in this case, an orthonormal basis is a basis and a completely continuous linear map is a linear map of finite rank.* (ii) *When A is a field, $|\ |$ is multiplicative, $|A^*| \neq 1$, and V is a finite dimensional subspace of M, then Serre proves that there exists a continuous projector from M onto V with norm less than 1 whose kernel is orthonormizable. We cannot prove this in our more general context.*

Suppose $\{e_i\}_I$ is an orthonormal basis for M and $\{d_j\}_J$ is an orthonormal basis for N. Suppose

$$L(e_i) = \sum_j n_{i,j} d_j \,.$$

Then, as in [S], we have the following useful lemma:

Lemma A1.6. *The linear map L is completely continuous if and only if*

$$\lim_{j \to \infty} \operatorname{Sup}_{i \in I} |n_{i,j}| = 0 \,.$$

or equivalently, for $S \subset I$, let $\pi_S : E \to E$ be the projector

$$\sum_{i \in I} a_i e_i \mapsto \sum_{i \in S} a_i e_i \,.$$

Then, L is completely continuous if and only if the net $\{\pi_S \circ L\}$, where S ranges over the directed set of finite subsets of I, converges to L.

Proof. First suppose the matrix for L is as above. Then for each finite set S of J let

$$L_S(e_i) = \sum_{j \in S} n_{i,j} d_j \,.$$

It is clear that the L_S converge to L.

Now suppose L is completely continuous. Then for each $\varepsilon > 0$ there exists an A-linear map $L' : M \to N$ whose image is contained in a finitely generated submodule N' and is such that $|L - L'| < \varepsilon$. Since N' is finitely generated there exists a finite subset T of J such that if π_T is the projection from N onto the span of $\{d_j\}_{j \in T} \mid \pi_T|_{N'} - \mathrm{id}_{N'}| < \varepsilon$. It follows that

$$|L - \pi_T \circ L'| < \varepsilon \,.$$

This implies $|n_{i,j}| < \varepsilon$ for $j \notin T$ which concludes the proof. □

For an orthonormizable Banach module E, let $E^\vee$ denote the continuous dual of E with the norm $|\ |^\vee$ defined by

$$|h|^\vee = \sup\{|h(x)| : x \in E^0\}$$

for $h \in E^\vee$. This is well defined and if B is an orthonormal basis for E, $|h|^\vee = \sup\{|h(e)| : e \in B\}$.

Lemma A1.7. *If M and N are orthonormizable Banach modules over A, $M^\vee \hat{\otimes} N$ is naturally isomorphic to $\mathscr{C}_A(M,N)$.*

Proof. Suppose $\{e_i\}_I$ is an orthonormal basis for M and $\{d_j\}_J$ is an orthonormal basis for N.

We can write any $y \in M^\vee \hat{\otimes} N$, uniquely, as

$$\sum_J h_j \otimes d_j$$

where $h_j \in M^\vee$, $|h_j|^\vee \to 0$. Now if $m \in M$, we set

$$y(m) = \sum_J h_j(m) d_j\,.$$

This clearly well defines a linear map from M to N and, since $|h_j|^\vee \to 0$, is completely continuous by the previous lemma.

Now let $e_i^\vee$ be the element of $E^\vee$ such that $e_i^\vee(e_j) = \delta_{i,j}$. We can represent any $h \in M^\vee$ as $\sum_I a_i e_i^\vee$ where $a_i \in A$ and the set $\{|a_i| : i \in I\}$ is bounded. If, on the other hand, $L \in \mathscr{C}_A(M,N)$ has the matrix $(n_{i,j})_{I,J}$ let $y = \sum_J(\sum_I n_{i,j} e_i^\vee) \otimes d_j$ which, using Lemma A1.6, we see is an element of $M^\vee \hat{\otimes} N$. Clearly, y maps to L.

The map $M^\vee \hat{\otimes} N$ to $\mathscr{C}_A(M,N)$ is independent of the choice of the bases because it is the natural map on $M^\vee \otimes N$ and is continuous. □

We say a normed ring A is semi-simple if:

The intersection of the maximal ideals of A is 0 and if m is a maximal ideal, the residual norm on A/m is multiplicative.

Examples. (i) *If A is a reduced affinoid algebra over a complete multiplicatively normed field and the norm on A is the supremum norm* [*BGR, Definition* 3.8/2], *then A is semi-simple* (*see* [*BGR, Proposition* 6.1.1/3 *and Corollary* 6.1.2/3]). *It also satisfies hypothesis M.* (ii) *The ring Λ with any of the norms described above is not semi-simple.*

Probably, the hypothesis on residual norms in a definition of semi-simple can be weakened, for our applications, to the assumption that that the residual norms are equivalent to a multiplicative norm (two norms on a ring are said to be equivalent, if they induce the same topology), as George Bergman has shown, based on results in [B], if a norm $|\ |_1$ on a field is equivalent to a multiplicative norm $|\ |_2$, then there exists a positive constant c such that $|\ |_2 \leqq c|\ |_1$. We do not know an example of a complete normed field whose norm is not equivalent to a multiplicative norm.

A2. The Fredholm determinant

Suppose A is a Banach algebra and E is a Banach module over A with an orthonormal basis B. If L is a completely continuous operator on E, and

$$\text{there exists a } c \in A^M \quad \text{such that } |cL| \leqq 1\,, \tag{$*$}$$

one can translate the discussion in Serre to produce a characteristic series $P_L(T)$ of L, with respect to B, which we will also denote by $\det(1 - TL)$ (which it morally is). The key point is: By means of $(*)$ we may suppose $|L| \leqq 1$ and observe, if $I \in \mathscr{I}(A)$, Lemma A1.6 implies $L(E^0) \bmod I^n E^0$ is contained in a free direct factor of $E^0/I^n E^0$ of finite rank over $A^0/I^n A^0$. We will suppose all completely continuous operators mentioned in this section satisfy property $(*)$ (which is automatic if $|A^m| \neq 1$). We can also prove:

Theorem A2.1. *If L has norm at most $|a|$ where $a \in A^m$ then $P_L(T)$ is an element of $A^0[[aT]]$ and is entire in T (i.e., if $P_L(T) = \sum_{m \geqq 0} c_m T^m$, $|c_m| M^m \to 0$ for any real number M). Also, $P_L(T)$ is characterized by:*

(i) *If $\{L_n\}_{n \geqq 0}$ is a sequence of completely continuous operators on E, and $L_n \to L$ then $P_{L_n} \to P_L$ coefficientwise.*
(ii) *If the image of L in E is contained in an orthonormizable direct factor F of finite rank over A of E such that the projection from E onto F has norm at most* 1 *then*

$$P_L(T) = \det(1 - TL \,|\, F)\,.$$

Proof. This follows by translating the arguments in [S]. E.g., suppose the hypotheses of (ii). Let $\pi : E \to F$ be the projection. After changing L by a homothety in A^m, if necessary, we may assume $|L| \leqq 1$. Let $F^0 = \{x \in F : |x| \leqq 1\}$. Let I be an element of $\mathscr{I}(A)$. Then, since $|\pi| \leqq 1$, $F_I =: F^0/IF^0$ injects onto a free direct factor of finite rank of E^0/IE^0 over A^0/I. It follows that

$$P_L(T) \equiv \det(1 - TL \,|\, F_I) \bmod I\,.$$

Assertion (ii) follows upon taking a limit. □

Remark A2.2. *If follows from* (i) *and* (ii) *of the theorem that $P_L(T)$ does not depend on the choice of the orthonormal basis but, as far as we know, it may depend, in general, on the norm on E and not just the topology. However below, Corollary A2.6.2, we show that, when A is semi-simple, it does only depend on the topology.*

Just as in [S, Sect. 5] (see the remark after *Corollaire* 1), one may deduce from the theorem,

Corollary A2.1.1. *If u and v are completely continuous operators on E,*

$$\det(1 - Tu)\det(1 - Tv) = \det((1 - Tu)(1 - Tv))\,.$$

Also, one may deduce similarly to the proof of *Corollaire* 2 of [S, Sect. 5].

Proposition A2.3. *Suppose E_1 and E_2 are orthonormizable Banach modules over A. Suppose u is a completely continuous homomorphism from E_1 to E_2 and $v : E_2 \to E_1$ is a continuous homomorphism. Then*

$$P_{u\circ v}(T) = P_{v\circ u}(T)\,.$$

Lemma A2.4. *Suppose N is a closed orthonormizable Banach submodule of M over A such that the quotient module $F := M/N$, with the induced norm, is also orthonormizable and moreover that there is an isometric section $\psi : F \to M$ of $M \to F$. Then M is orthonormizable and if L is a completely continuous operator on M stabilizing N, its restriction to N and the induced operator, L_F, on F are also completely continuous and*

$$P_L(T) = P_{L|_N}(T)P_{L_F}(T)\,.$$

Proof. Let $E := \{e_i : i \in I\}$ be an orthonormal basis for N and $D := \{d_j : j \in J\}$ be an orthonormal basis for F. Then, we claim,

$$B := \{e_i : i \in I\} \cup \{\psi(d_j) : j \in J\}\,.$$

is an orthonormal basis for M. First, it is clear that if $m \in M$, there exist unique $a_i, b_j \in A$, for $i \in I,\ j \in J$, such that

$$\sum_i a_i e_i + \sum_j b_j\psi(d_j) = m\,.$$

Since ψ is an isometry,

$$|m| \leqq \operatorname*{Max}_{i\in I,\, j\in J}\{|a_i|, |b_j|\} =: K\,.$$

Suppose $|m| < K$. It follows, from the fact that the norm on F is the induced norm, that $|\sum_j b_j d_j| < K$. From the fact that D is an orthonormal basis for F, we see that $|b_j| < K$ for all $j \in J$ and hence that $K = \mathrm{Max}_{i\in I}\{|a_i|\}$. Since E is an orthonormal basis for N, this latter equals $|\sum_i a_i e_i|$. Now, since $|\sum_i a_i e_i| > |\sum_j b_j d_j|$, we deduce that $|m| = K$, a contradiction. Thus, B is an orthonormal basis.

Now we know we can compute $P_L(T)$ with respect to B. For a subset S of an orthonormal basis for a Banach algebra W over A, let π_S be the projection of W onto the span of S, as described in the last section, and for an operator U on W, let $U^S = \pi_S \circ U$. Now, for a subset S of B, let $E_S = S \cap E$ and $D_S = \psi(S) \subseteq D$. Now, since $L|_N = \lim_S (L|_N)^{E_S}$ and $L_F = \lim_S L_F^{D_S}$ as S ranges over finite sets, these operators are completely continuous. It is elementary algebra to check, for finite subsets S of B, that,

$$P_{L^S}(T) = P_{(L|_N)^{E_S}}(T)P_{L_F^{D_S}}(T)\,.$$

The lemma follows from the fact, which is a consequence of Theorem A2.1, that

$$P_L(T) = \lim_S P_{L^S}(T), \qquad P_{L|_N}(T) = \lim_S P_{(L|_N)^{E_S}}(T) \quad \text{and}$$

$$P_{L_F}(T) = \lim_S P_{L_F^{D_S}}(T),$$

as S ranges over finite subsets of B. □

We remark that the hypothesis of this lemma about an isometric splitting is automatic when the absolute value on A is *discrete* (by this we mean that the subset of the real numbers $\{\log|a| : a \in A, a \neq 0\}$ is discrete). Indeed, in this case, with notation as in the proof of the lemma, for each $j \in J$, there exists an $e'_j \in M$ such that $e'_j = e_j \bmod N$ and $|e'_j| = 1$. Then we can define ψ as follows,

$$\psi\left(\sum_j b_j e_j\right) = \sum_j b_j e'_j .$$

We will see below that we can also eliminate this hypothesis when A is semi-simple.

It follows easily using Proposition A1.3 and Lemma A1.4 that

Lemma A2.5. *Suppose $\phi : A \to B$ is a contractive map of Banach algebras, then*

$$\phi(\det(1 - TL \mid E)) = \det(1 - T(1 \hat{\otimes} L) \mid B \hat{\otimes}_A E).$$

Proposition A2.6. *Suppose A is semi-simple, E is an orthonormizable Banach module over A and L is a completely continuous linear operator on E whose image is contained in a free submodule of finite rank F such that there is a continuous projector from E onto F. Then $P_L(T) = \det(1 - TL \mid F)$.*

Proof. Let $\{e_i\}$ be an orthonormal basis for E. Let m be a maximal ideal of $A, k = A/m$ and $|\ |_m$ the residual norm. Then the natural map from A to k is contractive so $1 \otimes e_i$ is an orthonormal basis for $k \hat{\otimes} E$ by Proposition A1.3. Also, if $\pi : E \to F$ is a continuous projector, $\mathrm{id} \otimes \pi : k \otimes E \to k \otimes F$ is a continuous projector and the elements $1 \otimes \pi(e_i)$ are bounded in $k \otimes F$. We also know $k \otimes E = k \hat{\otimes} E$ by Corollary A1.3.1. The result now follows from Lemma A2.5 applied to $B = A/m$ and Remark (1) of [S, Sect. 5] and [S, Propostion 7d)]. □

This proposition together with part (i) of the theorem implies that

Corollary A2.6.1. *When A is semi-simple, $P_L(T)$ only depends on the topology of E.*

We do not know whether or not this is true more generally.

Corollary A2.6.2. *When A is semi-simple and M is orthonormizable, the conclusion of Lemma* A2.4 *remains true even without the assumption that there is an isometric section from F to M.*

Proof. Let $E = \{e_i : i \in I\}$ and $D = \{d_j : j \in J\}$ be as in the proof of Lemma A2.4. Let $\varepsilon \in \mathbf{R}$ such that $0 < \varepsilon < 1$. For each $j \in J$ let d'_j be an element of M which maps to d_j such that $|d'_j| < (1+\varepsilon)$. Let ϕ be the unique continuous section of $M \to F$ which takes d_j to d'_j. Let C be the set $E \cup \{d'_j : j \in J\}$. Clearly, if $m \in M, m$ can be uniquely written in the form,

$$\sum_i a_i e_i + \sum_j b_j \phi(d_j) .$$

We now let $|\,|'$ be the unique absolute value on M such that C is an orthonormal basis. Then since ϕ is an isometric section with respect to $|\,|'$, we may apply Lemma A2.4 to the characteristic series of L with respect to this absolute value. But it is clear that for $m \in M$,

$$|m|' \leqq |m| \leqq (1+\varepsilon)|m|' .$$

Thus $|\,|$ and $|\,|'$ induce the same topology on M and so by the previous corollary, the characteristic series of L defined with respect to $|\,|'$ is the same as that defined with respect to $|\,|, P_L(T)$. Thus the conclusion of Lemma A2.4 applies to $P_L(T)$. □

Remark A2.7. *The Fredholm determinant may be defined and many of its properties proven when the condition "orthonormizable" is replaced by "locally orthonormizable."*

Example. *Suppose A is a Banach algebra, M is an orthonormizable Banach module over A and u and v are two completely continuous operators on M over A. Then if $A\langle X, Y\rangle$ is the ring of restricted power series over A, the operator $Xu + Yv$ is a completely continuous on $\tilde{M} =: M \hat{\otimes} A\langle X, Y\rangle$ over $A\langle X, Y\rangle$ (which is given the Gauss Norm). Hence we have a characteristic series $P_{u,v}(X, Y, T) = \det((1 - T(Xu + Yv))\,|\,\tilde{M})$ such that*

$$P_{u,v}(x, y, T) = \det((1 - T(xu + yv))\,|\,M)$$

whenever $x, y \in A$ and both $|x|$ and $|y|$ are at most 1. Clearly when $|A^m| \neq 1$, $P(X, Y, T)$ continues to a series entire in X and Y.

Now suppose A is an algebraically closed field with a multiplicative norm. Then if u and v commute

$$P_{u,v}(X, Y, T) = \prod_i (1 - (a_i X + b_i Y)T)$$

where a_i and b_i are elements of A which tend to zero. This is a consequence of the fact that the generalized eigenspaces of u are stabilized by v and vice versa.

A3.. Resultants

In this section we extend many of the classical results about resultants (see [L-A, Ch. IV, Sect. 8]) to our analytic situation. This is necessary for us to be able to prove analogues for completely continuous operators over a Banach algebra of Serre's Riesz theory results [S, Sect. 7] for completely continuous operators over a complete normed field.

Suppose $(A, |\,|)$ is a Banach algebra and $|A^m| \neq 1$.

Lemma A3.1. *If $G(T)$ is a polynomial whose leading coefficient is multiplicative and $H(T) \in A\langle T\rangle$ such that $G(T)H(T) \in A$ then $G(T) \in A$ or $H(T) = 0$.*

Proof. Let $a \in A^m$, $|a| > 1$. Replacing $G(T)$ by $G(a^M T)$ for some positive integer M we may assume that the absolute value of the leading coefficient c of G is greater than all its other coefficients. Suppose $n = \deg(G) > 0$ and $H \neq 0$. Suppose $H(T) = \sum_k b_k T^k$ and $m \geqq 0$ is such that $|b_m| \geqq |b_k|$ for all k with strict inequality for $k > m$. It follows that the coefficient of T^{n+m} has absolute value equal to $|cb_m| = |c||b_m| \neq 0$. □

For $I = (i_1, \ldots, i_n) \in \mathbf{N}^n, s(I) = i_1 + \cdots + i_n$ and if $(T_1, \ldots, T_n)$ is an n-tuple of elements in a ring, we set $T^I = T_1^{i_1} \cdots T_n^{i_n}$. Let $A\{\{T_1, \ldots, T_n\}\}$ be the ring of power series

$$\sum_I B_I T^I$$

over A in $(T_1, \ldots, T_n)$ where I ranges over $\mathbf{N}^n$, such that

$$|B_I| M^{s(I)} \to 0$$

as $s(I) \to \infty$ for all $M \in \mathbf{R}$. This is the ring of power series over A which converge on affine N-space over A. In particular, if $P(T)$ is the characteristic series of a completely continuous operator on a Banach module over $A, P(T) \in A\{\{T\}\}$.

Remark A3.2. *The above lemma is also true if we suppose instead of the hypothesis $G(T) \in A[T]$ that $G(T) \in A\{\{T\}\}$ and either all the coefficients of G are multiplicative or A is semi-simple.*

Suppose $e_1, \ldots, e_n$ are the elementary symmetric polynomials in $T_1, \ldots, T_n$.

Lemma A3.3. *The subring of $A[[T_1, \ldots, T_n]]$, $A\{\{e_1, \ldots, e_n\}\}$, is equal to the subring of $A\{\{T_1, \ldots, T_n\}\}$ consisting of elements which are left invariant under permutation of the variables T_i.*

Proof. For an element $I = (i_1, \ldots, i_n)$, let $t(I) = i_1 + 2i_2 + \cdots + ni_n$. Now if $I \in \mathbf{N}^n, e^I$ is a linear combination of T^J where $s(J) = t(I)$. Since

$$s(I) \leqq t(I) \leqq ns(I)\,,$$

it follows that if

$$\sum_I A_I T^I = \sum_J B_J e^J\,,$$

where the sums run over $\mathbf{N}^n$ and the A_I and B_J are elements of A, then

$$\max_{s(J)=m} \{|B_J|\} \leqq \max_{m \leqq s(I) \leqq nm} \{|A_I|\} \tag{1}$$

and

$$\operatorname*{Max}_{s(I)=m} \{|A_I|\} \leqq \operatorname*{Max}_{m/n \leqq s(J) \leqq m} \{|B_J|\} . \tag{2}$$

The containment of rings $A\{\{e_1,\ldots,e_n\}\} \subseteq A\{\{T_1,\ldots,T_n\}\}$ follows from estimate (2). It is clear that elements of $A\{\{e_1,\ldots,e_n\}\}$ are invariant under permutation of the T_i. If $f \in A\{\{T_1,\ldots,T_n\}\}$ is invariant under permutation of the T_i it follows that f equals $g(e_1,\ldots,e_n)$ for some $g \in A[[X_1,\ldots,X_n]]$. It now follows from estimate (1) that g is in fact in $A\{\{X_1,\ldots,X_n\}\}$ which completes the proof. □

Let

$$Q(T) = T^n - a_1 T^{n-1} + \cdots + (-1)^n a_n$$

be an element of $A[T]$.

Lemma A3.4. *If $S(e_1,\ldots,e_n)$ is in*

$$\left(\sum_{i=1}^{n} Q(T_i) A\{\{T_1,\ldots,T_n\}\}\right) \cap A\{\{e_1,\ldots,e_n\}\}$$

then $S(a_1,\ldots,a_n) = 0$.

Proof. First, suppose C is a ring and $K(T) = \sum_{i=1}^{n} (-1)^i c_i T^{n-i}$ and $R(e_1,\ldots,e_n)$ is in $(\sum_{i=1}^{n} K(T_i) C[T_1,\ldots,T_n]) \cap C[e_1,\ldots,e_n]$. Consider the ring

$$B = C[b_1,\ldots,b_n] / \left(K(T) - \prod_i (T - b_i)\right) .$$

We can write

$$R(e_1,\ldots,e_n) = \sum_i K(T_i) f_i(T_1,\ldots,T_n) ,$$

where $f_i(T_1,\ldots,T_n) \in C[T_1,\ldots,T_n]$. Then we may conclude

$$R(c_1,\ldots,c_n) = \sum_i K(b_i) f_i(b_1,\ldots,b_n) = 0 .$$

Now we assume the hypotheses of the lemma. Replace $Q(T)$ with $b^n Q(T/b)$ for some appropriate $b \in A^m$ so that all the a_i are in A^0. We can also scale S so that $S(e_1,\ldots,e_n)$ is in

$$\left(\sum_{i=1}^{n} Q(T_i) A^0\{\{T_1,\ldots,T_n\}\}\right) \cap A^0\{\{e_1,\ldots,e_n\}\}$$

Write S as

$$\sum_i Q(T_i) f_i(T_1,\ldots,T_n)$$

with $f_i \in A^0\{\{T_1,\ldots,T_n\}\}$. Let f_{iN} be the sum of the terms of f of degree at most N and $g_{iN} = f_i - f_{iN}$. Then if $\varepsilon > 0 \in \mathbf{R}$ for large N the coefficients of g_{iN} have absolute value at most ε. Let I_ε be the ideal in $A^0, \{a \in A : |a| \leqq \varepsilon\}$. Then we may apply the above argument with the ring C equal to A^0/I_ε and R equal to S mod I_ε, to conclude that $|S(a_1,\ldots,a_n)| \leqq \varepsilon$ for all $\varepsilon > 0$. Hence, $S(a_1,\ldots,a_n) = 0$. □

Suppose $P(T) \in A\{\{T\}\}$. We know $P(T_1)\cdots P(T_n) = H(e_1,\ldots,e_n)$ for some $H \in A\{\{X_1,\ldots,X_n\}\}$ by Lemma 3.3. Then, for Q as above, we define the *resultant* of Q and P to be

$$\mathrm{Res}(Q,P) = H(a_1,\ldots,a_n)\,.$$

(See also [L-A, Ch. IV, Sect. 8].) Then

$$\mathrm{Res}(Q,1) = 1 \tag{3}$$

$$\mathrm{Res}(Q,aP) = a^n \mathrm{Res}(Q,P) \tag{4}$$

$$\mathrm{Res}(Q,PR) = \mathrm{Res}(Q,P)\mathrm{Res}(Q,R) \tag{5}$$

$$\mathrm{Res}(Q,P+BQ) = \mathrm{Res}(Q,P) \tag{6}$$

if $a \in A$ and $R,B \in A\{\{T\}\}$. If $P(T) = \sum_{n\geqq 0} b_n T^n$, one can show $\mathrm{Res}(Q,P)$ is the limit as m goes to infinity of the determinants of the $(n+m)\times(n+m)$ matrices,

$$\begin{matrix} m\left\{\begin{matrix} \\ \\ \\ \\ \end{matrix}\right. \\ n\left\{\begin{matrix} \\ \\ \\ \\ \end{matrix}\right. \end{matrix}
\begin{pmatrix}
1 & -a_1 & \cdots & \cdots & (-1)^n a_n & & & \\
 & 1 & -a_1 & \cdots & \cdots & (-1)^n a_n & & \\
 & & \cdots & \cdots & \cdots & \cdots & \cdots & \\
 & & & 1 & -a_1 & \cdots & \cdots & (-1)^n a_n \\
b_m & b_{m-1} & \cdots & \cdots & b_0 & & & \\
 & b_m & b_{m-1} & \cdots & \cdots & b_0 & & \\
 & & \cdots & \cdots & \cdots & \cdots & \cdots & \\
 & & & b_m & b_{m-1} & \cdots & \cdots & b_0
\end{pmatrix}$$

in which there are m rows of a's and n rows of b's. If S is a monic polynomial of degree m,

$$\mathrm{Res}(SQ,P) = \mathrm{Res}(S,P)\mathrm{Res}(Q,P) \tag{7}$$

$$\mathrm{Res}(Q,S) = (-1)^{mn}\mathrm{Res}(S,Q) \tag{8}$$

$$\mathrm{Res}(Q,S^*) = \mathrm{Res}(S,Q^*) \tag{9}$$

where if $F(T)$ is a polynomial of degree $d, F^*(T) = T^d F(T^{-1})$. We can also interpret the resultant as a norm. Indeed, consider the extension $B := A\{\{T\}\}/(Q(T))$ of A. This extension is isomorphic to $A[T]/(Q(T))$ which is finite and free and the resultant of Q and P is the norm of the image of P in B to A.

Lemma A3.5. *The resultant of Q and P is a linear combination of Q and P. If Q and P have a non-constant polynomial common factor G whose leading term is multiplicative, then the resultant of Q and P is zero.*

Proof. When P is a polynomial, the first statement follows from [L-A Ch. IV, Sect. 8]. In general, we can write P as $BQ+R$ where R is a polynomial and $B \in A\{\{T\}\}$ and then apply formula (6) above.

Now it follows that $G(T)$ divides the resultant. However, the resultant lies in A, and this together with Lemma A3.1 implies the resultant is zero. □

Remark A3.6. *By Remark* A3.2, *the conclusion of this lemma is still true if we only assume* $G(T) \in A\{\{T\}\}$ *as long as* A *is semi-simple.*

Lemma A3.7. $\mathrm{Res}(Q,P)$ *is a unit if and only if* Q *and* P *are relatively prime in* $A\{\{T\}\}$.

Proof. One direction follows immediately from the previous lemma. Therefore, suppose $fQ + gP = 1$ where $f, g \in A\{\{T\}\}$. Then using (3), (5) and (6)

$$1 = \mathrm{Res}(Q, fQ + gP) = \mathrm{Res}(Q, gP) = \mathrm{Res}(Q, g)\mathrm{Res}(Q, P) . \qquad \square$$

We now want to explain the relationship between the characteristic series of a completely continuous operator and that of an entire series in that operator with zero constant term (which we know is also completely continuous).

Suppose B and P are polynomials over A and

$$P(T) = 1 - a_1 T + \cdots + (-1)^n a_n T^n .$$

Then we set

$$D(B,P)(T) = \prod_{i=1}^{n} (1 - TB(T_i)) ,$$

where on the right hand side we set $e_i(T_1, \ldots, T_n) = a_i$. If B and P are in $A\{\{T\}\}$, $B(0) = 0$ and $P(0) = 1$ then we set

$$D(B,P)(T) = \lim_{n \to \infty} D(B_n, P_n)(T)$$

where, for an element $F(T) = \sum_{k=0}^{\infty} c_k T^k \in A[[T]]$, $F_n(T) = \sum_{k=0}^{n} c_k T^k$. It is easy to see that $D(B,P)(T) \in A\{\{T\}\}$. Moreover,

Lemma A3.8. *If* $P(T) = R(T)S(T)$, R, $S \in A\{\{T\}\}$ *and* $R(0) = S(0) = 1$, *then we have,*

$$D(B,P) = D(B,R)D(B,S) . \tag{10}$$

and if Q *is a monic polynomial,*

$$D(1 - Q^*, P)(1) = \mathrm{Res}(Q,P) . \tag{11}$$

Proof. The first formula is obvious. For the second, observe that it follows from the definitions and (9) that

$$\begin{aligned} D(1 - Q^*, P_n)(1) &= \mathrm{Res}(T^n P_n(T^{-1}), Q^*(T)) \\ &= \mathrm{Res}(Q, P_n) . \end{aligned}$$

Hence the lemma follows by taking a limit. $\square$

Theorem A3.9. *If* u *is a completely continuous operator on an orthonormizable Banach module* E *over* A *and* $B \in TA\{\{T\}\}$ *then*

$$P_{B(u)}(T) = D(B, P_u)(T) . \tag{12}$$

Proof. As we remarked above $B(u)$ is completely continuous, so $P_{B(u)}$ makes sense. We may apply a homothety and assume that the norms of u and B are at most one. Suppose $I \in \mathscr{I}(A)$. Consider the operator u_I induced by u on $E_I := E^0/IE^0$. The corresponding formula is true for $\det(1 - u_I T \mid E_I)$ which is congruent to $P_u(T)$ modulo I. Hence (12) follows by a limiting argument. □

A4. Riesz theory

Suppose $(A, |\,|)$ satisfies hypothesis M. Let u be a completely continuous operator on an orthonormizable Banach module E over A. As in Serre, we can define the Fredholm resolvant $FR(T,u) := \det(1 - Tu)/(1 - Tu)$ of u, which is an element of $A[u]\{\{T\}\}$, and use it and the theory of resultants to prove:

Lemma A4.1. *Suppose $Q(T) \in A[T]$ is a monic polynomial. Then Q and P_u are relatively prime if and only if $Q^*(u)$ is an invertible operator on E.*

Proof. Let $v = 1 - Q^*(u)$. Then v is completely continuous and we have,

$$(1 - vT)FR(T,v) = P_v(T) = D(1 - Q^*, P_u)(T)\,,$$

by Theorem A3.9, and so using Lemma A3.8,

$$Q^*(u)FR(1,v) = (1 - v)FR(1,v) = \mathrm{Res}(Q, P_u)\,.$$

Thus it follows from Lemma A3.7 that if Q and P_u are relatively prime, $Q^*(u)$ is invertible. If, on the other hand, there exists an operator w on E such that $Q^*(u)(1 - w) = 1$, then we find that w is completely continuous and we deduce using Corollary A2.2.1,

$$\det(1 - v)\det(1 - w) = 1$$

but by Theorem A3.9 and Lemma A3.8,

$$\det(1 - v) = D(1 - Q^*, P_u)(1) = \mathrm{Res}(Q, P_u)\,.$$

Hence Q and P_u are relatively prime by Lemma A3.7. □

Let Δ^s denote the operator on power series in T which takes $\sum_n a_n T^n$ to $\sum_n \binom{n}{s} a_n T^{n-s}$. We also let $\Delta = \Delta^1$. Suppose $a \in A$. Then we say a is a zero of $H(T) \in A\{\{T\}\}$ of order h if $\Delta^s H(a) = 0$ for $s < h$ and $\Delta^h H(a)$ is invertible. (With this definition, some zeroes do not have an order.)

Using the previous lemma and following the same line of reasoning as in [S, Sect. 7] one obtains:

Proposition A4.2. *Suppose $a \in A$ is a zero of $P_u(T)$ of order h. Then we have a unique decomposition*

$$E = N(a) \oplus F(a)$$

into closed submodules such that $1 - au$ is invertible on $F(a)$ and $(1 - au)^h N(a) = 0$.

Theorem A4.3. *Suppose $P_u(T) = Q(T)S(T)$ where $S \in A\{\{T\}\}$ and Q is a polynomial such that $Q(0) = 1$ whose leading coefficient is a unit and which is relatively prime to S. Then there is a unique direct sum decomposition*

$$E = N_u(Q) \oplus F_u(Q)$$

of E into closed submodules such that $Q^*(u)N_u(Q) = 0$ and $Q^*(u)$ is invertible on $F_u(Q)$.

Proof. We note that $S(0) = 1$. Let $B(T) = 1 - Q^*(T)/Q^*(0)$ and $v = B(u)$. Then, by (A3.10)

$$P_v = D(B, P_u) = D(B, Q)D(B, S)\,.$$

We have $D(B, Q)(T) = (1 - T)^n$, where $n = \deg Q$ and

$$D(B, S)(1) = \operatorname{Res}(Q/Q^*(0), S)$$

by (A3.11) which is a unit using Lemma A3.7. Now apply Proposition A4.2 to the operator v and the zero 1 of $P_v(T)$. □

Remarks A4.4. (i) *Let $R_Q = A[X]/Q^*(X) \cong A[Y]/Q(Y)$. Then $N_u(Q)$ is a R_Q module, via*

$$Xm = um$$

for $m \in N_u(Q)$. (ii) *Following Serre we have explicit formulas for the projectors from E onto the subspaces $N_u(Q)$ and $F_u(Q)$. For example, let v be as above, then*

$$\left(\frac{(1-v)\Delta^n FR(1,v)}{\Delta^n P_v(1)}\right)^n$$

is a formula for the projector onto $F_u(Q)$ with kernel $N_u(Q)$.

Since projective modules over a ring are locally free, one can define the determinant of an operator on such a module if it has locally finite rank.

Theorem A4.5. *Suppose A is semi-simple and Q has degree r. Then under the hypotheses of Theorem* A4.3 *the A module $N_u(Q)$ is projective of rank r. Moreover,* $\det(1 - Tu|N_u(Q)) = Q(T)$.

Proof. First suppose A is a field, then $|\,|$ is multiplicative. The result [S, Proposition 12] of Serre applies and establishes our result in this case.

Let $N = N_u(Q)$ and $F = F_u(Q)$. Let m be a maximal ideal of A. Then because $E = N + F$, $E_m = N_m + F_m$ and $Q^*(u)$ is zero on N_m and invertible on F_m so that this decomposition is the one established by Theorem A4.3. It follows from the above and the hypotheses on A that N_m is a vector space of dimension r over k_m, the residue field at m. Now, let

$$f_i = \sum_{j \in I} a_{i,j} e_j \quad \text{for } 1 \leqq i \leqq r$$

be elements of N which form a basis of N_m modulo m. Then, there exist $j_1, \ldots, j_r$ in I such that

$$g = \det((a_{i,j_k})_{i,k})$$

is not zero at m. Let U be the affine open subscheme of Spec(A) where g is invertible. It follows that the f_i are a basis for N_P for every closed point P of U. We claim $\{f_i\}$ is a basis for N_U.

Indeed let $h \in N_U$. Then because g is invertible on U, there exist $a_i \in A_U$ such that the coefficient of e_{j_k} in the expansion of

$$a_1 f_1 + \cdots + a_r f_r - h$$

is zero for $1 \leqq k \leqq r$. If follows that this element vanishes at every closed point P of U. Thus by the hypotheses this element vanishes on U. If $h = 0$, it follows that the a_i vanish at every closed point P in U and hence $a_i = 0$ for all i. Thus N is locally free, so projective.

Finally, by Corollary A2.6.2,

$$\det(1 - Tu|E) = \det(1 - Tu|N)\det(1 - Tu|F)\,.$$

Now since $Q(T)$ divides $P_u(T)$ and $Q^*(u)$ is invertible on F, it follows, using Lemma 4.1, that $Q(T)$ differs from $\det(1 - Tu|N)$ by an element of A^*. Equality follows from the fact that $Q(0) = 1$. □

Corollary A4.5.1. *Suppose A is semi-simple. If R_Q is étale over A (i.e., if $(Q(T), \Delta Q(T)) = 1$) than $N_u(Q)$ is a locally free R_Q module of rank* 1.

Proof. This is true when A is a field. It follows more generally when A is semi-simple, by the same kind of reasoning which established the theorem. □

Remark A4.6. *One can show, when A is semi-simple, that $F_u(Q)$ is locally orthonormizable.*

A5. Rigid Theory

In this section, we will show how the results of the previous sections apply in the rigid category. We will be able to obtain much more precise results, which will be essential to us when we begin to discuss modular forms. A good emcyclopedic reference for the foundations of rigid analysis is the book *Non- Archimedean Analysis* by Bosch, Guntzer and Remmert. A more low key introduction to the subject can be found in the book *Géométrie Analytique*

Rigide et Applications by Fresnel and Van der Put and the original paper "Rigid analytic spaces" [T] by Tate is quite accessible.

Let K be either $\mathbf{C}_p$ or a complete discretely valued subfield of $\mathbf{C}_p$ and $|\,|$ be the absolute value on K such that $|p| = p^{-1}$ (or more generally we may suppose that K is a complete stable valued field (see [BGR, Sect. 3.6.1, Definition 1])). Let $K^0 = \{a \in K : |a| \leqq 1\}$ be the ring of integers in K and $\wp = \{a \in R : |a| < 1\}$ the maximal ideal of K^0.

If Y is a rigid space over K, $A(Y)$ will denote the ring of rigid analytic functions on Y, we let $|\,|$ also denote the supremum semi-norm on $A(Y)$ [BGR, Sect. 3.8] and $A^0(Y)$ will denote the subring in $A(Y)$ of power bounded functions, $\{f \in A(Y) : |f| \leqq 1\}$, on Y. The supremum semi-norm is a non-trivial ultrametric norm on $A(Y)$ if $A(Y)$ is reduced [BGR, Proposition 6.2.1/4]. As we have pointed out, $A(Y)$ is semi-simple in this case. We set $t(Y) = \{f \in A(Y) : |f| < 1\}$, the topologically nilpotent elements of $A(Y)$, and $\bar{Y} = \mathrm{Spec}(A^0(Y)/t(Y))$. In general, if $X \to Y$ is a morphism of rigid spaces and Z is a subspace of Y, then X_Z will denote the pullback of X to Z (the "fiber" of $X \to Y$ over Z).

In particular, $\mathbf{B}_K^n$ will denote the n-dimensional affinoid polydisk over K. Then $A(\mathbf{B}_K^n) \cong K\langle T_1, \ldots, T_n\rangle$ and $A^0(\mathbf{B}_K^n) \cong K^0\langle T_1, \ldots, T_n\rangle$. Finally, if $a \in K$ and $r \in |\mathbf{C}_p|$ we let $B_K[a,r]$ and $B_K(a,r)$ denote the affinoid and wide open disks of radius r about a in $\mathbf{A}_K^1$. When $K = \mathbf{C}_p$ we will drop the subscript K, and we will sometimes abuse notation and let these latter symbols denote the $\mathbf{C}_p$-valued points of the corresponding rigid space.

(i) *Fredholm and Riesz theory over affinoid algebras.* Suppose $X \to Y$ is a morphism of reduced affinoids over K. Then $(A(Y), |\,|)$ is a Banach algebra and $(A(X), |\,|)$ is a Banach module over $(A(Y), |\,|)$.

If $A^0(Y)/\wp A^0(Y)$ is reduced then $|A(Y)| = |K|$ so $(A(Y), |\,|)$ satisfies hypothesis M. In this case, $\wp A^0(Y) = t(Y)$ so the reduction of Y, $\bar{Y}$, equals $\mathrm{Spec}(A^0(Y)/\wp A^0(Y)) =: \tilde{Y}$. If Y is reduced, this occurs after a finite base extension. We will suppose for the rest of this section that Y is a reduced irreducible affinoid such that $\tilde{Y}$ is also reduced and we will regard $A(Y)$ as a Banach algebra with respect to the supremum norm.

One can show, using Lemma A1.2,

Lemma A5.1. *Suppose K is discretely valued, $X \to Y$ is a morphism of reduced affinoids over K and $A^0(X)/\wp A^0(X)$ is free over $A^0(Y)/\wp A^0(Y)$. Then the Banach module $A(X)$ over $A(Y)$ is orthonormizable.*

The simplest case of this phenomenon is: $X = Z \times_K Y$ where Z is a reduced affinoid over K. This will, in fact, be the case of interest to us.

Definition. *If $f : Z \to X$ is a morphism of affinoids over Y then we say, f is* **inner** *over Y if the image of $\overline{Z}$ in $\overline{X}$ is finite over $\overline{Y}$.*

This is a slight generalization of Kiehl's notion of inner which is called relatively compact in [BGR, Sect. 9.6.2].

Proposition A5.2. *Suppose $f : Z \to X$ is an inner map of reduced affinoids over Y, $\tilde{X}$ is reduced and $A(X)$ is orthonormizable over $A(Y)$. Then the map f^* from $A(X)$ to $A(Z)$ is a completely continuous homomorphism of Banach modules over $A(Y)$.*

Proof. Let $B = A^0(Y)$, $C = A^0(Z)$ and $D = A^0(X)$. Let $x_1, \ldots, x_n$ be elements of D such that the map from $B\langle T_1, \ldots, T_n\rangle$, $T_i \mapsto x_i$ is surjective onto D (these exist by [BGR, Theorem 6.4.3/1] using the fact that under our hypotheses $|D| = |K|$). The hypotheses that f is inner implies that the image of $\overline{X}$ is finite over $\overline{Y}$ which is equivalent to the existence of monic polynomials $g_i(S) \in B[S]$, $1 \leqq i \leqq n$ such that $f^* g_i(x_i) \in \pi C$ for some $\pi \in K^0$ such that $|\pi| < 1$. We can write any element of D as

$$\sum_{I,N} a_{I,N} x^I g(x)^N ,$$

where $x = (x_1, \ldots, x_n)$, $g = (g_1, \ldots, g_n)$, I and N are multi-indices in $\mathbf{N}^n$ ordered lexographically, $I < \deg(g)$ and $a_{I,N} \in B$. It follows that the image of D in $C/\pi^n C$ is spanned by the images of $f^*(x^I g(x)^N)$ where $I < \deg(g)$ and $S(N) < n$. Now let $\{e_i\}_{i \in I}$ be an orthonormal basis for $A(X)$ over $A(Y)$. Then $e_i \in D$. Let $F_{i,n}$ be an element in the B-span of $\{f^*(x^I g(x)^N) : I < \deg g$ and $S(N) < n\}$ such that $F_{i,n} \equiv f^* e_i \bmod \pi^n C$. There exists a unique continuous B-linear map $L_n : A(X) \to A(Z)$ such that $L_n(e_i) = F_{i,n}$. Then L_n converges to f^* and the image of L_n is contained in a submodule of C finitely generated over $A(Y)$. □

We will also need in Sects. B4 and B5, the following notion of relative over-convergence:

Definition. *If $X \to Y$ is a morphism of rigid spaces over K, we say that X is* **affinoid** *over Y if for each affinoid subdomain Z in Y, X_Z is an affinoid. Suppose $W \to Y$ is a map of rigid spaces and $X \subseteq W$ is affinoid over Y, then we say that a rigid space $V \subseteq W$ is a* **strict** *neighborhood of X over Y in W if for each affinoid subdomain Z of Y there exists a neighborhood U of X_Z in V affinoid over Y such that $X_Z \to U_Z$ is inner over Y. Finally, if X, W and Y are as above, we say that a rigid function f on X is* **overconvergent** *in W over Y if f extends to some strict neighborhood of X in W over Y. When Y is* Spec(K), *we just say f is overconvergent on X in W.*

Now suppose E is a Banach module over $A(Y)$. Suppose $P(T)$ is the characteristic series of a completely continuous operator u on E and $P(T) = Q(T)S(T)$ where $S \in A(Y)\{\{T\}\}$ and Q is a polynomial, whose leading coefficient is a unit and whose constant term is one, such that $(Q, S) = 1$.

Proposition A5.3. *Suppose Q has degree r. Then the $A(Y)$ module $N(Q)$ is projective of rank r and* $\det(1 - Tu|N_u(Q)) = Q(T)$.

Proof. Indeed, this follows from Theorem A4.5 since $A(Y)$ is semi-simple. □

In fact, in the rigid context, we can strengthen Corollary A4.5.1. Suppose $R_Q = A(Y)[Z]/Q(Z)$ is étale over $A(Y)$. Then R_Q is also a reduced affinoid algebra and the supremum norm on R_Q extends the supremum norm on A. The operator $1 \otimes u$ on $R_Q \otimes E$ over R_Q is completely continuous. Then Z is a zero of $P_{1\otimes u}(T) = P_u(T)$ of order 1 as

$$\Delta P_u(Z) = \Delta Q(Z)S(Z)$$

which is a unit since R_Q is étale over $A(Y)$ and $(Q, S) = 1$ so the subspace $N_{1\otimes u}(Z^{-1}T - 1)$ of $R_Q \otimes E$ is locally free of rank one over R_Q. Summarizing,

Proposition A5.4. *Suppose $R_Q = A(Y)[Z]/(Q(Z))$ is étale over $A(Y)$. Then, if $1 \otimes u$ is the extension of scalars of u to $R_Q \otimes E$, Z is a zero of $P_{1\otimes u}(T)$ of order one and, locally on $R_Q, N_{1\otimes u}(Z^{-1}T - 1)$ is freely generated by an element m such that*

$$(1 \otimes u)m = Z^{-1}m\,.$$

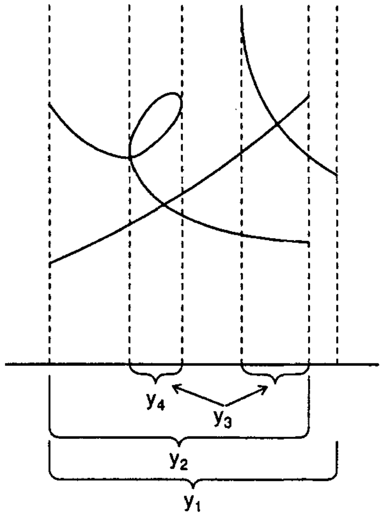
This is the genesis of our work on R-families (see Sect. B3 and Sect. B5).

More generally, suppose $Q = F^m$ where $m \deg F = \deg Q$, $F^*(u)N(Q) = 0$ and R_F is étale over A. Let $C = (R_F)^m$. Then C is a reduced affinoid algebra and the supremum norm on C extends the supremum norm on $A(Y)$.

(ii) *The zero locus of an entire series.* Suppose $P(T)$ is a non-zero entire power series over Y (like the characteristic series of a completely continuous operator on a Banach space over Y). Suppose $r \leqq s$ are real numbers in $|K|$. Then the subset of $Y \times \mathbf{A}^1_K$ determined by the inequalities $r \leqq |T| \leqq s$ is the affinoid $Y \times A[r, s]$, the fiber product of Y and the annulus of radii r and s, which is irreducible. The subspace of this affinoid determined by $P(T) = 0$ is an affinoid Z of dimension equal to that of Y. Moreover, the projection $Z \to Y$ is finite to one if $P(0) = 1$. We will investigate this situation in the abstract. I.e., suppose $f : Z \to Y$ is a quasi-finite morphism of affinoids over K. Then for a closed point x of Y, the fiber over $x, f^{-1}(x)$, is scheme of dimension 0 over the residue field of x. By $\deg(f^{-1}(x))$, we mean the dimension of its ring of functions over this field (its degree as a divisor). We will prove,

Proposition A5.5. *Let notation be as above. Suppose $Y = \mathbf{B}^1_K$. For each integer $i \geqq 0$ the set of closed points x of $\mathbf{B}^1_K$ such that $\deg(f^{-1}(x)) \geqq i$ is the set of closed points of an affinoid subdomain Y_i of Y. Moreover, $Y_i = \emptyset$ for large i.*

The following is a pictorial explanation of Proposition A5.5. Regard closed intervals in the interval representing Y as affinoid disks.



Before we begin the proof we point out the following corollaries:

Corollary A5.5.1. *For each $x \in \mathbf{B}^1_K(K)$, there exists an affinoid ball $B \subseteq \mathbf{B}^1_K$ over K containing x such that $g : Z_B \to B$ is finite.*

Corollary A5.5.2. *Suppose K is discretely valued. Let T be an invertible rigid function on Z defined over K. Then the set of valuations,*

$$\{v(T(z)) : z \in Z(\mathbf{C}_p), f(z) \in Y(K)\},$$

is finite.

Proof. Since the degree of $f^{-1}(y)$ for $y \in Y(\mathbf{C}_p)$ is bounded and for $y \in Y(K)$ the set of points of $f^{-1}(y)$ is closed under $\mathrm{Gal}(\bar{K}/K)$, the points in $f^{-1}(y)$ for $y \in Y(K)$ are all defined over a finite extension of K. The result follows from this and the fact that T is bounded above and below on Z. □

To prove Proposition A5.5, we will need,

Lemma A5.6. *Suppose $g : W \to \mathbf{B}^1_K$ is a non-constant morphism of affinoids over K and W is irreducible. Then the image of g is an affinoid subdomain of $\mathbf{B}^1_K$.*

Proof. We may suppose W is reduced and absolutely irreducible. We may also extend scalars to $\mathbf{C}_p$ so that $K = \mathbf{C}_p$ and $\widetilde{W} = \overline{W}$. After a translation and a homothety we may suppose $\bar{g}$ is non-constant. Since W is irreducible, $\bar{W}$ is connected and so the image of $\bar{g}$ is connected and thus an affine open. If every point whose reduction is in the image of $\bar{g}$ is in the image of g we have nothing to prove since the image of $\bar{g}$ is an affine open and its inverse image under reduction is an affinoid subdomain. Therefore suppose 0 is not in the image of g but is in the image of $\bar{g}$. Then there exists a $b \in \mathbf{C}_p^*$ such that $|b| < 1$ and $|b/g| = 1$. Let $h = b/g$. Then as $h, g \in A^0(W)$, $|g| = |h| = 1$ and $|gh| = |b| < 1$, it follows that $\overline{W}$ is not irreducible. Thus the lemma is true in the case when $\overline{W}$ is irreducible and in this case $g(W) = B[0,1] - \bigcup_{a \in T} B(a,1)$ where T is some finite subset of $B[0,1]$.

Now let Z be an irreducible component of $\overline{W}$. Let Z^0 be the complement in Z of the other irreducible components of $\overline{W}$ and $\tilde{Z}^0 = \mathrm{red}^{-1} Z^0$. Then the rigid space $\tilde{Z}^0$ is an irreducible open in W and since Z^0 is an affine open in $\overline{W}$, $\tilde{Z}$ is an affinoid subdomain with irreducible reduction. It follows from the argument in the previous paragraph (after undoing the translation and homothety) that $g(\tilde{Z}^0) = B[a_Z, r_Z] - \cup B(b_{Z,j}, r_Z)$ for some $r_Z \in |\mathbf{C}_p|$, $a_Z \in B[0,1]$ and some finite set $\{b_{Z,j}\}$ of $B[a_Z, r_Z]$.

Let $S = \{B(x,r) : x \notin g(W), r = |(g-x)^{-1}|^{-1}\}$. Thus S is the collection of maximal wide open disks in $B[0,1]$ contained in the complement of the image of g. We also note that the radii of the disks in S are elements of $|\mathbf{C}_p^*|$. Clearly, $g(W) = B[0,1] - \bigcup S$. We claim:

$$S \subseteq \{B(b_{Z,j}, r_Z) : Z \text{ is an irreducible component of } \overline{W}\} .$$

This will complete the proof of the lemma as the latter set is finite. Let $B(x,r) \in S$. In particular, $r \leqq 1$. After a translation we may suppose $x = 0$. Let $|b| = r$ and $h = b/g$. Then $\bar{h}$ is non-constant by the reasoning in the first paragraph of this proof if $r < 1$ and as an immediate consequence of the conclusions of this paragraph in the case $r = 1$. Therefore, there exists an irreducible component Z of $\overline{W}$ such that $\bar{h}|_Z$ is non-constant. It follows, that $\overline{g/b}|_{Z^0}$ is defined and non-constant. This implies, $|g|_{\tilde{Z}^0} = r$ and thus $B[a_Z, r_Z] = B[0,r]$ and since 0 is not in the image of g, $B(0,r) = B(b_{Z,i}, r_Z)$ for some i. This establishes the claim and completes the proof. □

Now we define a descending tower Z_i, $i \geqq 1$ of affinoid subdomains of Z such that, if $Y_i = f(Z_i)$, $x \in Y_i$ if and only if $\deg f^{-1}(x) \geqq i$. The Y_i are affinoid subdomains of $\mathbf{B}_K^1$ by the lemma as quasi-finiteness implies f is not constant on any irreducible component of Z_i. We take $Y_0 = Y$.

Let X denote the affinoid subspace of Z^k, $k \geqq 1$, determined by the equations $f \circ \pi_i(x) = f \circ \pi_j(x)$, $1 \leqq i \leqq j \leqq k$, where the $\pi_1 \cdots \pi_k$ are the k projections from Z^k to Z. Since f is quasi-finite, X is one dimensional. Let X_k denote the one dimensional affinoid consisting of the union of the irreducible components of X not contained in any hyperdiagonal, $\pi_i(x) = \pi_j(x)$ for some $i \neq j$, of Z^k and set $Z_k = \pi_1(X_k)$. It follows that Z_k satisfies the required conditions.

Finally, we sketch two proofs of the fact that $Y_i = \emptyset$ for large i. First, extend scalars to a maximally complete algebraically closed field Γ containing K. Maximal completeness implies there exists an $x \in \bigcap Y_i(\Gamma)$ if $Y_i \neq \emptyset$ for all i. But then $\deg f^{-1}(x) = \infty$ which contradicts the quasi-finiteness of f.

The other proof uses the stable reduction theory of curves. There exists a semi-stable model of f over a finite extension of K. I.e. there exist semi-stable formal scheme models $\mathscr{Y}$ and $\mathscr{Z}$ of Y and Z over K^0 and an extension of f to a morphism $\mathscr{F}$ from $\mathscr{Z}$ to $\mathscr{Y}$ such that $\bar{\mathscr{F}}$ is quasi-finite. It follows that for each irreducible component X of $\overline{\mathscr{Z}}$, the map $\overline{\mathscr{F}}_X$ has finite generic degree $d(X)$ for some non-negative integer $d(X)$. Suppose $x \in X$. Let $\bar{x}$ denote its image in $\overline{\mathscr{X}}$. Then one can show

$$\deg f^{-1}(x) \leqq \sum_X d(X)$$

where X runs over the irreducible components of $\overline{\mathscr{Z}}$ which meet $\overline{\mathscr{F}}^{-1}(\bar{x})$.

Questions and Remarks A5.7. (i) *Using the stable reduction theory of curves, one can check this proposition remains true whenever* $\dim Y = 1$. (ii) *Is the proposition true when Y has dimension greater than one if the phrase "an affinoid subdomain" in this proposition is replaced with "a finite union of affinoid subdomains?"* (iii) *It is clear that the results of this section can be globalized to arbitrary rigid spaces over K. One only has to replace the notion of orthonormizability with local orthonormizability.* (iv) *Suppose X is an irreducible component of the zero locus of $P(T)$. Liu has observed that the image of X in $\mathbf{B}^1(\mathbf{C}_p)$ is the complement of a finite set of points.* (v) *The projection from X to Y is not necessarily quasi-finite. In general, X corresponds to an irreducible factor of $P(T)$. Suppose*

$$P(x,T) = 1 + xT\prod_{i=1}^{\infty}(1 - p^iT)\,.$$

Then, $P(x,T)$ is an irreducible element of $A(B[0,1])\{\{T\}\}$, whereas $P(x_0,T)$ has infinitely many zeroes, for $x_0 \neq 0 \in B[0,1]$. (Note, however, that $x + T\prod_{i=1}^{\infty}(1 - p^iT)$ has infinitely many distinct irreducible factors.)

Although, we will not use the following result in this paper it will be crucial in constructing an important geometric object which encodes much of the theory of "families of modular forms" and related objects which we call the eigencurve.

Proposition A5.8. *Suppose $P(X,T)$ is a rigid analytic function on $\mathbf{B}^1_K \times \mathbf{A}^1_K$ such that $P(X,0) = 1$. Let Z be the zero locus of $P(X,T)$ and $f : Z \to \mathbf{B}^1_K$ the natural map. Let $\mathscr{C}$ be the collection of affinoid subdomains Y of Z such that Y is finite over $f(Y)$ and the collection $\{Y, Z_{f(Y)} - Y\}$ makes up an admissible open cover of $Z_{f(Y)}$ (i.e., Y is disconnected from its complement $Z_{f(Y)}$). Then $\mathscr{C}$ is an admissible open cover of Z.*

Proof. Let $r \in |K|$, $Y_r = Z \cap (\mathbf{B}^1_K \times B_K(0,r))$ and let f_r be the restriction of f to Y_r. Now, let notation be as in Proposition A5.5. Suppose V is an affinoid in $\mathbf{B}^1_K$ such that $\deg(f_r^{-1}(x)) = d > 0$ for all $x \in V$.

Let $\mathbf{B} = \mathbf{B}^1_K$. If $X \subseteq Y$ are affinoids in $\mathbf{B}$ we say Y is a strict affinoid neighborhood of X in $\mathbf{B}$ if there exists a strict affinoid neighborhood U of X

in $\mathbf{A}_K^1$ such that $Y = U \cap \mathbf{B}$. We will only complete the proof of the above proposition when $K = \mathbf{C}_p$. □

Lemma A5.9. *There exists an $s \in |\mathbf{C}_p|$ such that $s > r$ and there exists a strict affinoid neighborhood W of V in $\mathbf{B}$ such that the affinoid $Y := \{f_s^{-1}(x) : x \in W\}$ lies in $\mathscr{C}$ and has degree d over W.*

Proof. Write $P(X,T) = \sum_{i=0}^{\infty} a_i(X)T^i$, where $a_i(X) \in A(\mathbf{B})$ and $a_0(X) = 1$. Let $\alpha = \log_p(r)$. It follows, from the fact that $\deg(f_r^{-1}(x)) = d$ for all $x \in V$, that, for all $x \in V$ and all i,

$$v(a_d(x)) - v(a_i(x)) \leqq (d-i)\alpha\,,$$

with strict inequality for $i > d$. (Otherwise, there would exist a side of the Newton polygon of $P(x,T)$ of slope less than or equal to α, extending to the right of the point $(d, v(a_d(x)))$.) Now from the entirety of $P(X,T)$ it follows that there exists a real number $\beta > \alpha$ such that $v(a_d(x)) - v(a_i(x)) \leqq (d-i)\beta$, for $i > d$. It also follows from the above inequalities that $a_d(x)$ is invertible on V and so there exist real numbers $\delta_2 > \delta_1$ in $v(\mathbf{C}_p)$ such that $\delta_2 > v(a_d(x)) > \delta_1$ for all $x \in V$. Suppose, $\alpha < \gamma_1 < \gamma_2 < \beta$ for some γ_1 and γ_2 in $v(\mathbf{C}_p)$. Let W be the subspace of $\mathbf{B}$ determined by the inequalities,

$$\delta_1 \leqq v(a_d(x)) \leqq \delta_2 \tag{1}$$

$$v(a_d(x)) - v(a_i(x)) \leqq (d-i)\gamma_1 \quad \text{for } i < d \tag{2}$$

$$v(a_i(x)) - v(a_d(x)) \geqq (i-d)\gamma_2 \quad \text{for } i > d\,. \tag{3}$$

The entirety of $P(X,T)$ in T forces all but finitely many of the inequalities in (3) to be true for all $x \in \mathbf{B}$. Hence, W is a rational affinoid subdomain of Y_s, where $s = p^{\gamma_1}$, in the sense of [BGR, Sect. 7.2]. Since the affinoids defined by each of the inequalities in (1)–(3) are strict affinoid neighborhoods of V, W is as well. It is easily checked that Y has degree d over W. The fact that Y lies in $\mathscr{C}$ follows from the fact that Y and $\{y \in Z_W : |T| \geqq s\}$ make up an admissible open cover of Z_W by two disjoint admissible open subsets with respect to the strong topology. (See [BGR, Proposition 9.1.4/6].) □

Now to prove Proposition A5.8, first observe that the collection $\{Y_r : r \in |\mathbf{C}_p|, r > 0\}$ is an admissible open cover of Z. Thus all we have to do is find a finite cover of Y_r by elements of $\mathscr{C}$. We know, by Proposition A5.5, that the set $\{a \in \mathbf{B} : \deg(f_r^{-1}(a)) \geqq i\}$ is the set of points of an affinoid subdomain U_i in $\mathbf{B}$ and $U_i = \emptyset$ for i large. Let $Z_i = f_r^{-1}(U_i)$ which is an affinoid subdomain of Y_r. Let d be the largest integer such that $U_d \neq \emptyset$. Then Z_d is finite over U_d of degree d. By the lemma, there is a strict affinoid open neighborhood W_d of U_d in $\mathbf{B}$ and an $s > r$ such that $T_d =: f_s^{-1}(W_d)$ is finite over W_d of degree d and T_d is a finite union of connected components of Z_{W_d}. Suppose we have affinoid subdomains of Z, $T_i, T_{i+1}, \ldots, T_d$ satisfying

H(i) $\qquad T_i \in \mathscr{C}$ and if $S_i =: \bigcup_{j \geqq i} T_i$, $\;S_i \supseteq Z_i$ and

$f(S_i)$ is a strict affinoid neighborhood of U_i.

Since $f(S_i)$ is a strict affinoid neighborhood of U_i, there is an affinoid subdomain V of $U_{i-1} - U_i$ such that $V \cup f(S_i) \supseteq U_{i-1}$. Then, by the lemma, there exists a strict affinoid open neighborhood W of V in $\mathbf{B}$ such that there is an affinoid subdomain T_{i-1} of Z containing $f_r^{-1}(V)$ which is finite of degree $i-1$ over W and is a union of connected components of Z_W. It follows that $T_{i-1}, \ldots, T_d$ satisfy (H(i − 1)). Hence we may construct a cover $T_1, \ldots, T_d$ of Y_r satisfying (H(1)) and this completes the proof. □

B. Families of modular forms

B1. Eisenstein series

For the statements about Eisenstein series discussed in this section see [H-LE, Ch. 5 Sect. 1 and Ch. 9 Sect. 4] as well as [S-MZp, Sect. 3]. For the statements on p-adic L-functions see [L-CF, Ch. 4] and [W, Chs. 5 and 7].

For a character $\chi : \mathbf{Z}_p^* \to \mathbf{C}_p^*$, let f_χ denote the smallest positive integer such that χ is trivial on $1 + f_\chi \mathbf{Z}_p$ if one exists, if not, let $f_\chi = \infty$. We call f_χ the conductor of χ. For a ring R, let $\mu(R)$ denote the group of roots of unity in R. Let $w_p = |\mu(\mathbf{Q}_p)|$, $\tau : \mathbf{Z}_p^* \to \mu(\mathbf{Q}_p)$ be the character of smallest conductor which restricts to the identity of $\mu(\mathbf{Q}_p)$ and $\mathbf{q} = f_\tau$. Then $w_2 = 2$, τ is the character $d \mapsto (-1)^{(d-1)/2}$ and $\mathbf{q} = 4$, if $p = 2$. Otherwise, $w_p = p - 1$, τ is the composition of reduction and the Teichmüller character and $\mathbf{q} = p$. For $d \in \mathbf{Z}_p^*$, let $\langle\langle d \rangle\rangle = d/\tau(d)$ which is congruent to 1 modulo $\mathbf{q}$. Also fix a $(p-1)$-st root π of $-p$ in $\mathbf{C}_p$. We summarize this notation in the following table:

p	w_p	$\mathbf{q}$	$\tau(d)$	$\langle\langle d \rangle\rangle$	π
2	2	4	$(-1)^{(d-1)/2}$	$d/\tau(d)$	-2
> 2	$p-1$	p	$\lim_{n\to\infty} d^{p^n}$	$d/\tau(d)$	$(-p)^{1/(p-1)}$

We let $\mathscr{W}$ equal the rigid analytic space over $\mathbf{Q}_p$ whose points over $\mathbf{C}_p$ are the continuous $\mathbf{C}_p^*$-valued characters on $\mathbf{Z}_p^*$. We note that $\mathbf{Z}$ injects naturally into $\mathscr{W}(\mathbf{Q}_p)$; indeed, send $k \in \mathbf{Z}$ to the character which maps $a \in \mathbf{Z}_p^*$ to a^k. Let $\mathbf{1}$ denote the trivial character $a \mapsto 1$. We think of $\mathscr{W}$ as our weight space (it has been known for some time that, p-adically, a weight should be thought of as a continuous $\mathbf{C}_p^*$-valued character on $\mathbf{Z}_p^*$ (see [K-pMF, Sect. 4.5] or [G-ApM, Sect. I.3.4]).) For $\kappa \in \mathscr{W}(\mathbf{C}_p)$, $\kappa \neq \mathbf{1}$, and $n \geqq 1 \in \mathbf{Z}$, let

$$\sigma_\kappa^*(n) = \sum_{\substack{d|n \\ (d,p)=1}} \kappa(d) d^{-1} \quad \text{and} \quad \zeta^*(\kappa) = \frac{1}{\kappa(c) - 1} \int_{\mathbf{Z}_p^*} \kappa(a) a^{-1} dE_{1,c}(a)$$

in the notation of [L-CF, Ch. 4 Sect. 3] for any $c \in \mathbf{Z}_p^*$ such that $\kappa(c)$ is not 1. So that, when $\kappa(a) = \langle\langle a \rangle\rangle^s \chi(a)$ where $s \in \mathbf{C}_p$, $|s| < |\pi/\mathbf{q}|$, and χ is a character of finite order

$$\zeta^*(\kappa) = L_p(1-s,\chi)$$

where L_p is the p-adic L-function. (This number is zero when $\chi(-1) = -1$.) If $\kappa \neq 1$ let

$$G^*_\kappa(q) = \frac{\zeta^*(\kappa)}{2} + \sum_{n \geqq 1} \sigma^*_\kappa(n) q^n .$$

Then when $\kappa(a) = \langle\langle a \rangle\rangle^k \chi(a)$, where k is an integer and χ is a character of finite order on $\mathbf{Z}_p^*$ such that $\chi(-1) = 1$, $G^*_\kappa(q)$ is the q-expansion of a weight k overconvergent modular form G^*_κ on $\Gamma_1(\mathrm{LCM}(\mathbf{q}, f_\chi))$ and character $\chi\tau^{-k}$. We call such characters $|K|$, *arithmetic characters.* It is classical if k is at least 1 (see [Mi]). (To prove that G^*_κ is the q-expansion of an overconvergent modular form, in general one first invokes Theorem 4.5.1 of [K-pMF] to conclude that it is the q-expansion of a p-adic modular form. Next one observes that this modular form is an eigenvector for the U-operator with eigenvalue 1. Finally, one invokes a generalization of Proposition II.3.22 of [G-ApM] to conclude that this p-adic modular form is overconvergent.)

Whenever $\zeta^*(\kappa) \neq 0$ and $\kappa \neq 1$, let $E^*_\kappa(q) = 2G^*_\kappa(q)/\zeta^*(\kappa)$. We also set $E^*_1(q) = 1$. Suppose $\kappa \in \mathscr{W}(\mathbf{C}_p)$ and κ is trivial on $\mu(\mathbf{Q}_p)$, then $|\zeta^*(\kappa)/2| > 1$ and

$$|E^*_\kappa(q) - 1| < 1 .$$

Remark B1.1. *We may regard $\mathscr{W}$ as a rigid analytic covering space of $\mathbf{A}^1_{\mathbf{Q}_p}$ whose fibers are principal homogeneous spaces for the group* $\mathrm{Hom}(\mathbf{Z}_p^*, \mu(\mathbf{C}_p))$. *Indeed the covering map is given by*

$$\kappa \in \mathscr{W}(\mathbf{C}_p) \mapsto \log(\kappa(a))/\log(a)$$

for any $a \in 1 + \mathbf{q}\mathbf{Z}_p$, $a \neq 1$ ($|\kappa(a) - 1| < 1$ since κ is a continuous homomorphism). The space $\mathscr{W}$ has w_p connected components (one for each element of $\hat{D} :=$ $\mathrm{Hom}(D, \mu(\mathbf{Q}_p))$, *where $D = (\mathbf{Z}/\mathbf{q}\mathbf{Z})^*$) each conformal to the open unit disk over $\mathbf{Q}_p$. In view of this, ζ^* may be thought of as a rigid analytic function on a covering space of $\mathbf{C}_p$. (We may think of $\mathscr{W}$ as $\mathscr{B} \times \hat{D}$ where* $\mathscr{B}(\mathbf{C}_p) = \mathrm{Hom}_{\mathrm{cont}}(1 + \mathbf{q}Z_p, \mathbf{C}_p^*)$.).

Let $\mathscr{B}^* = B_{\mathbf{Q}_p}(0, |\pi/\mathbf{q}|)$ and $\mathscr{W}^* = \mathscr{B}^* \times Z/w_p\mathbf{Z}$. We identify a point $s = (t, i)$ in $\mathscr{W}^*(\mathbf{C}_p)$ with the character $\kappa_s : a \mapsto \langle\langle a \rangle\rangle^t \tau^i(a)$ (and will denote this latter expression a^s) and will also write, in this case, $G_s = G^*_{\kappa_s}$ and $E_s = E^*_\kappa$. Thus both $\mathbf{Z}$ and $\mathscr{W}^*$ sit inside $\mathscr{W}$ and in fact $\mathbf{Z} \subset \mathscr{W}^*(\mathbf{Q}_p)$. More directly, an element $n \in Z$ corresponds to the element $(n, n \bmod w_p)$ of $\mathscr{W}^*$. Let E denote the weight one modular form $E_{(1,0)}$ which naturally lives on $X_1(\mathbf{q})$. We signal,

$$E(q) = 1 + \frac{2}{L_p(0,\mathbf{1})} \sum_{n \geqq 1} \left(\sum_{\substack{d|n \\ (d,p)=1}} \tau^{-1}(d) \right) q^n . \qquad (1)$$

Note that $E(q) \equiv 1 \bmod \mathbf{q}$ because $L_p(0, \mathbf{1}) \equiv 1/p \bmod \mathbf{Z}_p$.

For an integer $m \geqq 0$ and a positive integer N prime to p let $Z_1(Np^m)$ denote the rigid connected component of the ordinary locus in $X_1(Np^m)$ containing the cusp ∞. In particular, $Z_1(Np^m)$ is an affinoid.

Lemma B1.2. *Suppose $\kappa(a) = \langle\langle a\rangle\rangle^k \chi(a)$ where k is an integer and χ is a character in $\mathcal{W}$ of finite order which is trivial on $\mu(\mathbf{Q}_p)$. Then E_κ^* (which converges on) does not vanish on $Z_1(p^m)$ where $p^m = \mathrm{LCM}(\mathbf{q}, f_\chi)$.*

Proof. First E_κ^* converges on $Z_1(p^m)$ because it is overconvergent. Next, the lemma is true for E; i.e., E does not vanish on $Z_1(\mathbf{q})$, because E^{p-1} reduces to the Hasse invariant on the component of the reduction of the Deligne–Rapoport/Katz–Mazur model of $X_1(\mathbf{q})$ containing ∞. Now observe that $F = E_\kappa^*/E^k$ is a function on $Z_1(p^m)$ whose q-expansion is congruent to 1. It follows that F is congruent to 1 on all of $Z_1(p^m)$ and so doesn't vanish there. Hence, $E_\kappa^* = FE^k$ does not vanish on this affinoid. □

B2. General setup

In this section, we will set the groundwork needed to be able to study overconvergent forms in all levels for all primes.

Suppose $N > 4$ and $n \geqq 1$ are integers such that $(Nn, p) = 1$ and there is a lifting A of the Hasse invariant to $X_1(Nn)$. Such a lifting always exists if $p > 3$ (indeed, in such a case, one can take $A = E_{p-1}$) and exists for suitable n for $p = 2$ or 3. For $v \geqq 0 \in \mathbf{Q}$ let $X_1(Nn)(v)$ denote the affinoid subdomain of $X_1(Nn)$, $v(A(y)) \leqq v$. (In particular, $X_1(Nn)(0) = Z_1(Nn)$.) Let $E_1(Nn)$ be the universal elliptic curve over $X_1(Nn)$ and $E_1(Nn)(v)$ denote its pullback to $X_1(Nn)(v)$. Then by Katz, [K-pMF], if $v < 1/(p+1)$ there is a commutative diagram of rigid morphisms;

$$\begin{array}{ccc} E_1(Nn)(v) & \xrightarrow{\ \mathbf{\Phi}\ } & E_1(Nn)(pv) \\ \downarrow & & \downarrow \\ X_1(Nn)(v) & \xrightarrow{\ \phi\ } & X_1(Nn)(pv) \end{array}$$

We will think of this diagram as a morphism, labeled $\mathbf{\Phi}/\phi$, from $E_1(Nn)(v)/X_1(Nn)(v)$ to $E_1(Nn)(pv)/X_1(Nn)(pv)$, which it is, in the category of morphisms of rigid spaces. For $w \geqq 0 \in \mathbf{Q}$ let $X_1(N)(w)$ be the affinoid subdomain of $X_1(N)$ which is the image of $X_1(Nn)(w)$ in $X_1(N)$ and $E_1(N)(v)$ the pullback of $E_1(N)$ to $X_1(N)(v)$.

Proposition B2.1. *If $0 \leqq v \leqq 1/(p+1)$, there is a unique morphism*

$$\mathbf{\Phi}'/\phi' : E_1(N)(v)/X_1(N)(v) \to E_1(N)(pv)/X_1(N)(pv)$$

such that

$$\begin{array}{ccc} E_1(Nn)(v)/X_1(Nn)(v) & \xrightarrow{\Phi/\phi} & E_1(Nn)(pv)/X_1(Nn)(pv) \\ \downarrow & & \downarrow \\ E_1(N)(v)/X_1(N)(v) & \xrightarrow{\Phi'/\phi'} & E_1(N)(pv)/X_1(N)(pv) \end{array}$$

commutes, where the vertical arrows are the natural forgetful projections.

Proof. It is enough to check this on $\mathbf{C}_p$-valued points. Let x be a point of $X_1(N)(v)$ and $(E, \alpha : \mu_N \hookrightarrow E)$ the corresponding elliptic curve with level structure. Let y be a point of $X_1(Nn)(v)$ above x. Then y corresponds to (E, β) where $\beta : \mu_{Nn} \hookrightarrow E$ is an injective homomorphism such that $\beta|_{\mu N} = \alpha$. It follows that $\phi(y)$ corresponds to (E', β') where $E' = E/\mathscr{K}_y$ and β' is the composition of β and the natural map $\rho : E \to E/\mathscr{K}_y$ where $\mathscr{K}_y$ is the canonical subgroup of E. Moreover, $\rho = \Phi_E$. The proposition follows from the fact that $\mathscr{K}_y$ is independent of the choice of y. Indeed, $\mathscr{K}_y(\mathbf{C}_p)$ is the set of points of $E[p]$ closest to the origin. (If $v(A) = w < p/(p+1)$ and X is a local uniformizer on E at zero the points of $\mathscr{K}_y$ are the points P of E such that $v(X(P)) \geqq (1-w)/(p-1)$ and at the other points P of order p, $X(P)$ has valuation $w/(p^2-p)$ [K-pMF, Theorem 3.10.7].) □

Henceforth we will denote Φ'/ϕ' by Φ/ϕ.

This proposition is already enough to allow us to establish the main results of [C-CO] for the primes 2 and 3. In particular, for any prime p we can define an operator, $U_{(k)}$, on overconvergent forms of level N and weight k and assert that any such form of weight k and slope strictly less than $k-1$ is classical.

As in [C-CO, Sect. 8], we may and will regard $X_1(N)(v)$ as an affinoid subdomain of the modular curve $X(N; p) = X(\Gamma_1(N) \cap \Gamma_0(p))$.

For m a positive integer, denote the set $\{v \in \mathbf{Q} : 0 \leqq p^{m-1}v < p/(p+1)\}$ by the expression I_m and $I_m - \{0\}$ by I_m^*. Fix a subfield K of $\mathbf{C}_p$ equal to $\mathbf{C}_p$ or to a complete discretely valued subfield. All our constructions will be over K. We will employ the notation and definitions of [C-CO]. For $v \in I_1$, let $X_1(Np)(v)$ be the affinoid subdomain of $X_1(Np)$ which is the inverse image of $X_1(N)(v)$ under the natural forgetful map to $X(N; p)$. For $k \in \mathbf{Z}$ let $M_{Np,k}(v)$ be the space of modular forms of weight k on $\Gamma_1(N_p)$ which converge on $X_1(Np)(v)$. In other words, $M_{Np,k}(v) = \omega^k(X_1(Np)(v))$. Now, $M_{Np,k}(v)$ has a natural structure as a Banach space and when $v > 0$ there is a completely continuous operator on this space denoted by $U_{(k)}$ in [C-CO].

Now we will recall and modify some constructions carried out in [C-HCO]. Let $\pi_{m,1} : X_1(Np^m) \to X_1(Np)$ denote the map which sends the point corresponding to triples (E, α, β), where E is an elliptic curve and where $\alpha : \mu_{p^m} \hookrightarrow E$ and $\beta : \mu_N \hookrightarrow E$ are embeddings, to the point corresponding to $(E, \alpha|_{\mu_p}, \beta)$. Suppose, $v \in I_m$. Let $X_1(Np^m)(v)$ denote the affinoid subdomain of $X_1(Np^m)$ consisting of points x corresponding to triples (E, α, β) such that $\pi_{m,1}(x) \in X_1(Np)(v)$, $\Phi^{m-1}(\alpha(\mu_{p^{m-1}})) = 0$ and $(\Phi^{m-1}E, \Phi^{m-1} \circ \alpha, \Phi^{m-1} \circ \beta)$

corresponds to $\phi^{m-1}(\pi_{m,1}(x))$. Let $E_1(Np^m)(v)$ denote the pullback of $E_1(Np^m)$ to $X_1(Np^m)(v)$.

If $v \in I_m$, we have a lifting of Φ/ϕ to a morphism from $E_1(Np^m)(v/p)/X_1(Np^m)(v/p)$ to $E_1(Np^m)(v)/X_1(Np^m)(v)$, which takes (E,α,β) to $(\Phi E, \alpha', \Phi \circ \beta)$ where $\alpha' : \mu_{p^m} \hookrightarrow E$ is determined by the requirements that $(\Phi E, \alpha', \Phi \circ \beta)$ corresponds to a point in $X_1(Np^m)(v)$ and $\alpha'(\zeta) = \Phi(Q)$, if Q is a point of E such that $pQ = \alpha(\zeta)$ for $\zeta \in \mu_{p^m}$, $\zeta \neq 1$. We will denote these liftings by the same symbols. The context will make it clear which spaces we are dealing with.

Let $\omega := \omega_{Np^m}$ equal the direct image on $X_1(Np^m)$ of the sheaf $\Omega^1_{E_1(Np^m)/X_1(Np^m)}$. For $k \in \mathbf{Z}$, $v \in I_m$, we set

$$M_{Np^m,k}(v) := \omega^k(X_1(Np^m)(v)) .$$

These spaces may be considered as Banach spaces over K and when $v > 0$, we have a completely continuous operator, which we will still denote by $U_{(k)}$, acting on $M_{Np^m,k}(v)$ defined as in [C-CO] (see also [C-HCO]).

We can deal with $N \leqq 4$, $(N,p) = 1$ along the same lines as those discussed in the remark at the end of Sect. 6 of [C-CO]. In particular, if $A, B \in \mathbf{Z}$, $A, B > 4$, $(AB, p) = 1$ and $(A, B) = N$, we identify $M_{Np^m,k}(v)$ with the intersection of the images (via the maps which preserve q-expansions) of $M_{Ap^m,k}(v)$ and $M_{Bp^m,k}(v)$ in $M_{ABp^m,k}(v)$.

B3. Twists of U

In this section we prove Theorems A, B and D of the introduction as well as their extensions to the prime 2.

Fix a positive integer N, $(N,p) = 1$. For $v \in I_1$, let $X(v) = X_1(N\mathbf{q})(v)$ and $M_k(v) := M_{N\mathbf{q},k}(v)$. Recall, $I_m = \{v \in \mathbf{Q} : 0 \leqq p^{m-1}v < p/(p+1)\}$ and $I_m^* = I_m - \{0\}$.

Suppose $v \in I_1$ and $F \in M_{k-r}(v)$ is an overconvergent form of weight $k - r$ which has an inverse in $M_{r-k}(v)$ (we will see an example of such a form below). Then the map from $M_r(v)$ to $M_k(v)$, $h \mapsto hF$, is an isomorphism of Banach spaces. Moreover, the pullback of $U_{(k)}$ on $M_k(v/p)$ to $M_r(v/p)$ is the map

$$h \mapsto F^{-1}U_{(k)}(hF)$$

which equals $U_{(r)}(hF/\sigma(F))$, by [C-C0, 3.3].[1] Thus, in this case, since the restriction map $M_k(v) \to M_k(v/p)$ is completely continuous, this formula, together with Proposition A2.3, implies the Fredholm theory of the operator $U_{(k)}$ on $M_k(v/p)$ is equivalent to that of the completely continuous operator $U_{(r)} \circ m_f$ on $M_r(v/p)$ where $f = F/\sigma(F)$ and m_f is the operator "multiplication by f". (Note that $I_1/p = I_2$.)

[1] If $F(q)$ is the q-expansion of F, $\sigma(F)$ is the overconvergent form in $M_{k-r}(v/p)$ whose q-expansion is $F(q^p)$.

Recall, E is the weight one modular form $E_{(1,0)}$ on $\Gamma_1(\mathbf{q})$ with character τ^{-1} described in Sect. B1. It follows that there is an analytic function e on $\bigcup_{v\in I_2} X(v)$ with q-expansion $E(q)/E(q^p)$. Since $E(q) \equiv 1 \bmod \mathbf{q}$, we see that $|e-1|_{X(0)} \leqq |\mathbf{q}|$. As $|e-1|_{X(0)} = \lim_{v\to 0+} |e-1|_{X(v)}$ (the $X(v)$, $v \in I_2^*$, form a basis of neighborhoods of $X(0)$), we have,

Lemma B3.1. *For any $\varepsilon \in \mathbf{R}$, $|\mathbf{q}| < \varepsilon$, there exists a $v \in I_2^*$ such that e is defined on $X(v)$ and $|e-1|_{X(v)} \leqq \varepsilon$.*

Recall, π is a $(p-1)$-st root of $-p$. For $s \in \mathbf{C}_p$, $|s| < |\pi/\mathbf{q}|$,[2] let u_s be the operator on $M_0(v)$, for any $v \in I_2^*$ such that $|e-1|_{X(v)} < |\pi/s|$, defined by

$$u_s(h) = U_{(0)}(he^s)\,.$$

Then from the discussion in the previous two paragraphs, we see that if $k \in \mathbf{Z}$

$$\det(1 - Tu_k|M_0(v)) = \det(1 - TU_{(k)}|M_k(v))\,. \tag{1}$$

Recall, $\mathscr{B}^* = B_{\mathbf{Q}_p}(0, |\pi/\mathbf{q}|)$. Now we think of s as a parameter on $\mathscr{B}^*$ so that we may view e^s as a rigid analytic function on the rigid analytic subspace $\mathscr{V}^*$ of $\mathbf{A}^1 \times X_1(N\mathbf{q})$ which we define to be that subspace admissibly covered by the affinoids $Z_t(v) := B_K[0,t] \times_K X(v)$ where $v \in I_2^*$ and $t \in |\mathbf{C}_p| \cap [1, |\pi/\mathbf{q}|)$ such that e is defined on $X(v)$ and $|e-1|_{X(v)} \leqq |\pi|/t$. Let $\mathscr{T}^*$ be the set of ordered pairs (t,v) satisfying these conditions. (The set $\mathscr{T}^*$ is not empty, in fact, by the previous lemma, we see that the first projection to $|\mathbf{C}_p| \cap [1, |\pi/\mathbf{q}|)$ is a surjection.) Since $U_{(0)}$ extends uniquely to a continuous $A(B[0,t])$-linear map from $A(Z_t(v))$ to $A(Z_t(v))$ for $(t,v) \in \mathscr{T}^*$, we may now view u_s as a family of operators, i.e. there is a compatible collection of operators $\{U_{(t,v)} : (t,v) \in \mathscr{T}^*\}$, where $U_{(t,v)}$ is the operator on $A(Z_t(v))$, whose restriction to the fiber above s is u_s. This operator is nothing more than the composition of $id \otimes U_0$ and the operator, m_{e^s} multiplication by the function e^s, restricted to $Z_t(v)$. By Proposition A5.2, if $M(t,v) := A(Z_t(v))$, for $(t,v) \in \mathscr{T}^*$, $U_{(t,v)}$ is a completely continuous operator on $M(t,v)$ over $A(B[0,t])$. We will abuse notation and write U^* for $U_{(t,v)}$ when the context makes it clear we are talking about an operator action on $M(t,v)$. Also, as remarked after Lemma A5.1, $M(t,v)$ is orthonormizable over $A(B[0,t]) = \mathbf{C}_p\langle X/m\rangle$ where $m \in \mathbf{C}_p$ such that $|m| = t$. Thus we have characteristic series $P_{(t,v)} := \det(1 - TU^*|M(t,v))$ for any $(t,v) \in \mathscr{T}^*$.

We claim this series is independent of (t,v), in the sense that if (t,v) and (t',v') lie in $\mathscr{T}^*$ and $t \leqq t' < |\pi/\mathbf{q}|$ then the restriction of $P_{(t',v')}$, which is analytic on $B[0,t'] \times \mathbf{C}_p$, to $B[0,t] \times \mathbf{C}_p$ is $P_{(t,v)}$. Indeed, we first observe that if $(t,v) \in \mathscr{T}^*$, $0 < w \leqq v$, $s \leqq t$, $w \in \mathbf{Q}$ and $s \in |\mathbf{C}_p|$, $(s,w) \in \mathscr{T}^*$. From this, it follows that we only need to establish the claim when $t = t'$ or $v = v'$. When $v = v'$, it follows from Lemma A2.5. Now suppose $t = t'$. We may also suppose $v \geqq v' \geqq v/p$. For $u \leqq w$ such that $(t,u), (t,w) \in \mathscr{T}^*$, let R_u^w denote $1/p$ times the restriction map from $M(t,w)$ to $M(t,u)$ (which

[2] The series $\sum_{n=0}^{\infty} \binom{s}{n} T^n$ converges for $|T| \leqq |\mathbf{q}|$ if and only if $|s| \leqq |\pi/\mathbf{q}|$.

is completely continuous over $A(B[0,t])$ by Proposition A5.2 if $w > u$) and $T_u^{u/p} : M(t, u/p) \to M(t,u)$ the trace with respect to the restriction of $1 \otimes \phi$ to $M(t, u/p)$ (ϕ is the Frobenius morphism described in the last section which restricts to a finite morphism from $X(u/p)$ to $X(u)$). Then on $M(t,u)$, U^* is the operator $T_u^{u/p} \circ R_{u/p}^u \circ m_{e^s}$ (see [CO, Sect. 2]). Now we observe that it follows from the aforementioned finiteness of ϕ that,

$$T_{v'}^{v'/p} \circ R_{v'/p}^{v'} = R_{v'}^{v} \circ T_v^{v/p} \circ R_{v/p}^{v'} .$$

As

$$(T_v^{v/p} \circ R_{v/p}^{v'} \circ m_{e^s}) \circ R_{v'}^{v} = T_v^{v/p} \circ R_{v/p}^{v} \circ m_{e^s} = U^* ,$$

the claim follows from Proposition A2.3.

Theorem B3.2. *There is a unique rigid analytic function $P(s,T) = P_N(s,T)$ on $\mathscr{B}^* \times \mathbf{C}_p$ defined over $\mathbf{Q}_p$, i.e. $P(s,T)$ is a power series over $\mathbf{Q}_p$ in s and T, which converges for $|s| < |\pi/\mathbf{q}|$, such that for $k \in \mathbf{Z}$ and $v \in \mathbf{Q}$ such that $0 < v < p/(p+1)$,*

$$P(k,T) = \det(1 - TU_{(k)}|M_k(v)) .$$

Proof. The existence of the function $P(s,T)$ defined over $\mathbf{C}_p$ follows from the discussion in the previous two paragraphs combined with formula (1) and Lemma A2.5. That it is defined over $\mathbf{Q}_p$ follows from the fact that it equals $\det(1 - TU^*|M(t,v))$ for any $(t,v) \in \mathscr{T}^*$. Indeed, $M(t,v)$ is the extension of scalars of an orthonormizable Banach module $M_L(t,v)$ over $A(B_L[0,t])$ such that U^* restricts to a completely continuous operator on $M_L(t,v)$, for any finite extension L of $\mathbf{Q}_p$ contained in $\mathbf{C}_p$ such that t and p^v lie in $|L|$. Since we may choose $t = 1$, it follows that $P(s,t)$ is defined over any complete extension of $\mathbf{Q}_p$ containing an element with valuation less than $1/(p+1)$ and since the intersection of these is $\mathbf{Q}_p$, we see that $P(s,T)$ is, in fact, defined over $\mathbf{Q}_p$.

Suppose now $Q(s,T)$ is an analytic function on $\mathscr{B}^* \times \mathbf{C}_p$ such that

$$Q(k,T) = \det(1 - TU_{(k)}|M_k(v)) ,$$

for $k \in \mathbf{Z}$ and $v \in \mathbf{Q}$ such that $0 < v < p/(p+1)$, then $R(s,T) := P(s,T) - Q(s,T)$ vanishes on the set $S = \{(k,T) : k \in \mathbf{Z}\}$. Now consider the two dimensional affinoid balls in $\mathscr{B}^* \times \mathbf{C}_p$, $Y(a,b)$ where $a,b \in |\mathbf{C}_p^*|$ such that $a \leqq |\pi/\mathbf{q}|$ defined by the inequalities $|s| \leqq a$ and $|T| \leqq b$. Then the restriction of $R(s,T)$ to $Y(a,b)$ vanishes on $S \cap Y(a,b)$ which is a union of infinitely many one-dimensional affinoid balls defined by the equations $s - k = 0$, where $k \in \mathbf{Z}$. It follows that $R(s,T) \in \bigcap_{k \in \mathbf{Z}}(s-k)A(Y(a,b)) = 0$. Thus, $R(s,T)$ vanishes on $Y(a,b)$ and since $\bigcup_{a,b} Y(a,b) = \mathscr{B}^* \times \mathbf{C}_p$, $R(s,T) = 0$. Thus Q must equal P, which establishes the uniqueness. □

Let

$$P(s,T) = \sum_{n \geqq 0} f_n(s) T^n .$$

At this point we know that the coefficients of the series $f_n(s)$ lie in $\mathbf{Q}_p$ and the numbers $|f_n|_{\mathscr{B}^*}$ are bounded independently of n. In fact, if $p \geqq 7$, using known properties of $U_{(0)}$ (e.g. Lemma 3.11.7 of [K-pMF]), we could show they are bounded by 1, but later, in the Appendix I, we will give explicit formulas, derived using the Monsky–Riech trace formula, for the $f_n(s)$ which imply that they are Iwasawa functions. We also give a conceptual proof of this in [C-CPS] as well as proof that $P(s, T)$ extends to a rigid analytic function of $\mathscr{B} \times \mathbf{C}_p$.

We now explain how to factor $P(s, T)$ into series depending on nebentype. As $Z_t(v) = B[0, t] \times X(v)$ for $(t, v) \in \mathscr{T}^*$, the diamond operators $\langle b \rangle$, $b \in (\mathbf{Z}/N\mathbf{q}\mathbf{Z})^*$, act on $M(t, v)$. Recall, $D = (\mathbf{Z}/\mathbf{q}\mathbf{Z})^*$. We will regard D as a subgroup of $(\mathbf{Z}/N\mathbf{q}\mathbf{Z})^*$ in the natural way and also as a quotient of $\mathbf{Z}_p^*$. Then D acts via the diamond operators on all the spaces $M_k(v)$ and $M(t, v)$. For $a \in \mathbf{Z}_p^*$ we set $\langle a \rangle = \langle a \bmod \mathbf{q} \rangle$. For each integer k, character $\varepsilon \in \hat{D}$ and $v \in I_2^*$, let $M_k(v, \varepsilon)$ denote the subspace of $M_k(v)$ of forms with eigencharacter ε for this action. Similarly, let $M(t, v, \varepsilon)$ denote the subspace of $M(t, v)$ with eigencharacter ε. Then,

$$M_k(v) = \bigoplus_{\varepsilon} M_k(v, \varepsilon)$$

and

$$M(t, v) = \bigoplus_{\varepsilon} M(t, v, \varepsilon)$$

where the direct sums range over $\varepsilon \in \hat{D}$. Moreover these direct sums are stabilized by the respective operators $U_{(k)}$ and U^*. We thus have, by Lemma A2.4, the formulas

$$\det(1 - TU_{(k)}|M_k(v)) = \prod_{\varepsilon} \det(1 - TU_{(k)}|M_k(v, \varepsilon))$$

and

$$\det(1 - TU^*|M(t, v)) = \prod_{\varepsilon} \det(1 - Tu_s|M(t, v, \varepsilon)) .$$

Let $P_\varepsilon(s, T)$ be the function on $\mathscr{B}^* \times \mathbf{C}_p$ characterized by the identities:

$$P_\varepsilon(s, T)|_{B[0,t] \times \mathbf{C}_p} = \det(1 - TU^*|M(t, v, \varepsilon))$$

for all $(t, v) \in \mathscr{T}^*$. Then, arguing as in the proof of Theorem B3.2, we see that $P_\varepsilon(s, T)$ is defined over $\mathbf{Q}_p$,

$$P_\varepsilon(k, T) = \det(1 - TU_{(k)}|M_k(v, \varepsilon\tau^{-k})) \tag{2}$$

and

$$P(s, T) = \prod_{\varepsilon} P_\varepsilon(s, T) \tag{3}$$

for $k \in \mathbf{Z}$. This implies Theorem B of the introduction (except for the assertion that $P_{N,i}(s, T) \in \mathbf{Z}_p[[s, T]]$ which will follow from Corollary I.2.1), as well as its extension to $p = 2$. That is, we have proven

Theorem B3.3. *For each character* $0 \leqq i < w_p$ *there exists a series* $P_{N,i}(s,T) \in \mathbf{Q}_p[[s,T]]$ *which converges on the region* $|s| < |\pi/\mathbf{q}$ *in* $\mathbf{C}_p^2$ *such that for integers* k, $P_{N,\varepsilon}(k,T)$ *is the characteristic series of Atkin's* U*-operator acting on overconvergent forms of weight* k *and character* τ^{i-k}.

Indeed, we may take $P_{N,i} = P_{\tau^i}$. (Note that, when p is odd, $|\pi/\mathbf{q}| = p^{(p-2)/(p-1)}$.) While N is fixed we set $P_i(s,T) = P_{\tau^i}(s,T)$.

Recall, K equals $\mathbf{C}_p$ or is complete and discretely valued subfield. Let $M_{k,cl} = M_{k,cl}(N)$ denote the space of classical modular forms of weight k on $X_1(N\mathbf{q})$ defined over K. For a character $\varepsilon \in \hat{D}$, $M_{k,cl}(\varepsilon)$ denotes the subspace of forms of weight k and D-character (i.e. character for the action of D) ε. Also, $\mathbf{d}(k,\varepsilon,\alpha)$ equals the dimension of the subspace of $M_{k,cl}(\varepsilon\tau^{-k})$ consisting of forms of slope α. Theorem D, extended to $p=2$, is,

Theorem B3.4. *If* $\varepsilon \in \hat{D}$, $\alpha \in \mathbf{Q}$ *and* k *and* k' *are integers strictly bigger than* $\alpha + 1$ *and sufficiently close p-adically,*

$$\mathbf{d}(k,\varepsilon,\alpha) = \mathbf{d}(k',\varepsilon,\alpha)\,.$$

Moreover, the closeness sufficient for this equality only depends on α.

Proof. The first assertion follows from (2), Theorem C (which is the assertion that the set of zeroes of $P_\varepsilon(k,T^{-1})$ with valuation strictly less than $k-1$ is the same as the set of eigenvalues of classical weight k eigenforms with D-character $\varepsilon\tau^{-k}$ of slope strictly less than $k-1$) and Proposition A5.5. The second assertion follows from this and the fact that $\mathbf{Z}_p$ is compact. □

The fact that the set of slopes of classical modular forms on $\Gamma_1(N\mathbf{q})$ is discrete in $\mathbf{R}$ follows from Corollary A5.2.2.

Let $S(t,v,\varepsilon)$ denote the subspace of cusp forms in $M(t,v,\varepsilon)$ (i.e. the subspace of functions vanishing at the cusps in $X(0)$). Then $S(t,v,\varepsilon)$ is stable under U^* and we can proceed as above and let $P_\varepsilon^0(s,T)$ be the function characterized by the identities:

$$P_\varepsilon^0(s,T) = \det(1 - TU^*|S(t,v,\varepsilon))$$

for all $(t,v) \in \mathscr{T}^*$. Moreover,

$$P_\varepsilon^0(k,T) = \det(1 - TU^*|S_k(v,\varepsilon)) \tag{4}$$

for $k \in \mathbf{Z}$. We also $P_i^0(s,T) = P_{\tau^i}^0(s,T)$.

We will now prove Theorem A of the introduction, its extension to the prime 2, as well as a qualitative version of the Gouvêa–Mazur conjecture about "R-families" [GM-F, Conj. 3] in the case in which $U_{(k)}$ acts semi-simply on the slope α subspace of $M_{k,cl}(\varepsilon\tau^{-k})$. To treat the general case, we will use the ring of Hecke operators to be defined in Sect. B5.

Theorem B3.5. *Suppose $\alpha \in \mathbf{Q}$ and $\varepsilon : (\mathbf{Z}/\mathbf{q}\mathbf{Z})^* \to \mathbf{C}_p^*$ is a character. Then there exists an $M \in \mathbf{Z}$ which depends only on p, N, ε and α with the following property: If $k \in \mathbf{Z}$, $k > \alpha + 1$ and there is a unique normalized cusp form F on $X_1(N\mathbf{q})$ of weight k, $(\mathbf{Z}/p\mathbf{Z})^*$-character $\varepsilon\tau^{-k}$ and slope α and if $k' > \alpha + 1$ is an integer congruent to k modulo p^{M+n}, for any non-negative integer n, then there exists a unique normalized cusp form F' on $X_1(N\mathbf{q})$ of weight k', $(\mathbf{Z}/p\mathbf{Z})^*$-character $\varepsilon\tau^{-k'}$ and slope α. Moreover, this form satisfies the congruence*

$$F'(q) \equiv F(q) \bmod \mathbf{q}p^n .$$

Let $d(k, \varepsilon, \alpha)$ denote the dimension of the slope α subspace of $M_k(\varepsilon\tau^{-k})$. Also let $\mathbf{d}^0(k, \varepsilon, \alpha)$ and $d^0(k, \varepsilon, \alpha)$ denote the dimensions of the spaces of cusp forms of slope α in $M_{k,cl}(\varepsilon\tau^{-k})$ and $M_k(\varepsilon\tau^{-k})$. Then, by [C-CO, Theorem 8.1 and Lemma 8.7], we know,

$$\begin{aligned} \mathbf{d}(k, \varepsilon, \alpha) &= d(k, \varepsilon, \alpha) \\ \mathbf{d}^0(k, \varepsilon, \alpha) &= d^0(k, \varepsilon, \alpha) . \end{aligned} \qquad \text{if } k > \alpha + 1 \qquad (5)$$

Fix $\varepsilon \in \hat{D}$. Let Z_ε^0 be the zero locus of $P_\varepsilon^0(s, T)$ in $\mathscr{B}^* \times \mathbf{A}^1$ and for $\alpha \geqq 0 \in \mathbf{Q}$ let $Z_\varepsilon^0(\alpha)$ be the affinoid subdomain of Z_ε^0 whose closed points are $\{z \in Z_\varepsilon^0 : v(T(z)) = -\alpha\}$. This affinoid is, in fact, defined over $\mathbf{Q}_p$. Let $r \geqq 0 \in \mathbf{Z}$ and $T_r(\varepsilon, \alpha)$ be the subset of $j \in \mathbf{Z}_p$ such that $d(j, \varepsilon, \alpha) = r$. It follows from Corollary A5.5.1 that $T_r(\varepsilon, \alpha)$ is compact and if $k \in T_r(\varepsilon, \alpha)$, there exists an affinoid ball $B := B[k, p^{-m}] \subset \mathscr{B}^*$ containing k such that the map $Z^0(\alpha)_B \to B$ is finite of degree r. Thus there is a monic polynomial $Q(T)$ of degree r with coefficients in $K\langle(s-k)/p^m\rangle$ such that $P_\varepsilon(s, T)_B = Q(T)S(T)$ where $S(T) \in K\langle(s-k)/p^m\rangle\{\{T\}\}$ prime to $Q(T)$. By Theorem A4.3,

$$M_B = N_{U_B}(Q) \oplus F_{U_B}(Q) ,$$

where U_B is the restriction of U^* to M_B. Let

$$R_Q = \mathbf{Q}_p\langle(s-k)/p^m\rangle[X]/Q^*(X) .$$

We know $N_{U_B}(Q)$ is a R_Q-module.

Suppose that $(\Delta Q(k, T), Q(k, T)) = 1$. This will automatically be true when $r = 1$ and more generally when the eigenvalues of $U_{(k)}$ as an operator on the slope α subspace of $M_k(\varepsilon)$ over $\mathbf{C}_p$ are distinct. Then after shrinking B, if necessary, we may suppose that $(\Delta Q(T), Q(T)) = 1$ (now regarding $Q(T)$ as a polynomial over $A(B)$) and using Corollary A4.5.1, we may suppose that $N_{U_B}(Q)$ is a locally principal R_Q module.

Suppose, for the moment, that $r = 1$ and suppose k is an integer and F is a non-zero overconvergent cusp form on $\Gamma_1(N\mathbf{q})$ of weight k, character $\varepsilon\tau^{-k}$ and slope α. Because all the Hecke operators preserve the space of slope α, character $\varepsilon\tau^{-k}$ modular forms, F must be an eigenform. It is non-constant because it vanishes on the cusps in $X(0)$, so we may suppose it is normalized. If $F|U_{(k)} = aF$, then $x = (k, a)$ is a point of $Z^0(\alpha)$. Thus the map $Z^0(\alpha)_B \to B$

has degree 1. It follows that there exists a function f on B such that $f(k) = a$, $v(f(s)) = \alpha$ and $1 - f(s)T$ divides the restriction of $P_\varepsilon(s,T)$ to $B \times \mathbf{A}^1$. This latter is the characteristic series of the restriction of U^* to the space $M\hat{\otimes}B$ by Lemma A2.5 and $\Delta(P_\varepsilon(s, 1/f(s))$ is invertible on B. Hence, since $A(B)$ is a PID, our Riesz theory implies that there exists an analytic function G on $B \times X(0)^\dagger$ which vanishes at the cusps and spans the kernal of $U^* - f(s)$ in $S(\varepsilon)_B$. Thus if $k' \in \mathbf{Z}$, $k' \equiv k \bmod p^m, F_{k'} := E^{k'}G(k')$ is a non-zero overconvergent modular form of weight k', slope α and character $\varepsilon\tau^{-k'}$. Moreover, if $k' > \alpha + 1$, $F_{k'}$ is classical by [C-CO, Theorem 8.1]. Now let

$$G(s) = \sum_{n \geqq 1} a_n(s)q^n$$

be the q-expansion of $G(s)$. The $a_n(s)$ are rigid analytic functions on B. We must have $F_k = a_1(k)F$ and so $a_1(k) \neq 0$. Hence after shrinking B, we may suppose $a_1(s)$ is invertible and therefore we may suppose it equals 1. In particular, now $F_k = F$. Since G is bounded on the affinoid $B \times X(0)$, being a rigid function, the a_n's are uniformly bounded on B. Hence, there exists a constant $M \geqq 0$ such that for all $t \geqq 0$ and all $n \geqq 0$ and all $a \in \mathbf{Z}$

$$|a_n(k + p^{t+M}a) - a_n(k)| \leqq |\mathbf{q}p^t| \, .$$

As $E^{p^t}(q) \equiv 1 \bmod \mathbf{q}p^t$, this implies

$$F_{k'}(q) \equiv F_k(q) \bmod \mathbf{q}p^t$$

if $k' \equiv k \bmod p^{t+M}$. Since $T_r(\varepsilon, \alpha)$ is compact we see that we may pick an M that only depends on α. This yields Theorem B3.5.

Remarks B3.6. (i) *Using the Hecke operators to be introduced in Sect.* B5, *we will show in Lemma* B5.3 *that it is unnecessary to shrink B before assuming that $a_1(s) = 1$ and also that, then, the functions $|a_k(s)|$ are bounded by* 1. *This means that if $m \geqq 0$ is an integer such that $B(k, |p^m|) \subseteq B$, then we can take $M = m + v(\mathbf{q})$.*
(ii) *All of the above will go through with $M(\varepsilon)$ and $P_\varepsilon(s,T)$ in place of $S(\varepsilon)$ and $P^0_\varepsilon(s,T)$ if we suppose ε is not trivial. When $r = 1$, all we needed to know was that our form F is not constant.*

Now allow r to be arbitrary. By Proposition A5.4 (note that here $X = Z^{-1}$), shrinking B if necessary, there exists a generator $H \in R_Q \otimes_{A(B)} N_{U_B}(Q)$ such that,

$$(1 \otimes U)H = XH \, .$$

Suppose

$$E^s(q)H(q) = \sum_{n \geqq 0} b_n q^n$$

where the $b_n \in R_Q$. Let Y_Q be the rigid space sitting over B whose ring of functions is R_Q. If $k \in \mathbf{Z}$, and y is a point of Y_Q above k', $\sum_{n \geqq 0} b_n(y)q^n$ is the q-expansion of an overconvergent modular form $F_{k'}$ of weight k', character

$\varepsilon\tau^{-k}$ and slope α. If $k > \alpha + 1$ then $F_{k'}$ is classical. In fact, because R_Q is étale over B, $F_{k'}$ is an eigenform. In this way, we get a weak version of an "R-family" in the sense [GM-F] where $R = R_Q$.

Recall, $\mathscr{W}^* = \mathscr{B}^* \times \mathbf{Z}/w_p\mathbf{Z}$. It will sometimes be convenient for us to replace our base $\mathscr{B}^*$ with $\mathscr{W}^*$. Indeed, as we will see in Sect. B6, the ring Λ embeds naturally into the ring of rigid analytic functions on $\mathscr{W}^*$ (in fact, Λ is naturally isomorphic to the ring of rigid analytic functions defined over $\mathbf{Q}_p$ and bounded by 1 on $\mathscr{W}$). First, we identify $\mathbf{Z}/w_p\mathbf{Z}$ with $\hat{D}$ via $i \in \mathbf{Z}/w_p\mathbf{Z} \mapsto \tau^i$. Then we may view $A(\mathscr{W}^*)$ as $A(\mathscr{B}^*)[D]$ and, if $t < |\pi/\mathbf{q}|$ write $A(\mathscr{W}^*(t))$ for $A(B[0,t])[D] = A(B[0,t] \times \hat{D})$. For $(t,v) \in \mathscr{T}^*$, we make $M(t,v)$ into a Banach $A(\mathscr{W}^*(t))$-module as follows: If $f = \sum_{d\in D} f_d d$ is an element of $A(\mathscr{W}^*(t))$ where $f_d \in A(B[0,t])$ and G is an element of $M(t,v)$ we set

$$fG = \sum_{d\in D} f_d G|\langle d\rangle \,.$$

Henceforth, we will write f as $\sum_{d\in D} f_d\langle d\rangle$. Now for $\varepsilon \in \hat{D}$, let $\iota_\varepsilon \in (1/w_p)\mathbf{Z}_p[D]$ be the element $(1/w_p)\sum_{d\in D}\varepsilon^{-1}(d)\langle d\rangle$. Then any element m in an $A(\mathscr{W}^*(t))$ module equals $\sum_{\varepsilon\in\hat{D}} m_\varepsilon$ where $m_\varepsilon = \iota_\varepsilon m$. We put "new" norms $|\ |^*$ on $M(t,v)$, for $(t,v) \in \mathscr{T}^*$, as follows: Suppose H is in $M(t,v)$. Then we set

$$|H|^* = \operatorname*{Max}_{\varepsilon\in\hat{D}}\{|H_\varepsilon|_{Z_t(v)}\} \,.$$

When p is odd, $|\ |^*$ equals the supremum norm and is equivalent to it, in general (because ι_ε is defined over $\mathbf{Z}_p$ when p is odd and over $(1/2)\mathbf{Z}_2$ when $p = 2$). Moreover, $M(t,v)$ is a Banach module over $A(\mathscr{W}^*(t))$ with respect to $|\ |^*$.

If B is a $\mathbf{Z}_p$-algebra and $\varepsilon \in \hat{D}$, we also let ε denote the unique B-module homomorphism from $B[D]$ to B which takes $\langle d\rangle$ to $\varepsilon(d) \in \mathbf{Z}_p$.

Lemma B3.7. *With respect to the norm $|\ |^*$, $M(t,v)$ is orthonormizable over $A(\mathscr{W}^*(t))$. Moreover, on $M(t,v), U^*$ is a completely continuous $A(\mathscr{W}^*(t))$-operator. There is a series $Q_N(T) \in A(\mathscr{W}^*)[[T]]$ whose restriction to $A(\mathscr{W}^*(t))$ is the characteristic series for this operator. It is characterized by the identities:*

$$\varepsilon(Q_N(T)) = P_\varepsilon(s,T) \,,$$

for all $\varepsilon \in \hat{D}$.

Proof. Since this result will not be crucial in what follows, we only sketch the proof. For each $1 \leqq i \leqq w_p$, let $v_{i,1},\ldots,v_{i,n},\ldots$ be an orthonormal basis for $M(t,v,\tau^i)$ over $A(B[0,t])$ (with respect to the supremum norms). Then the set $w_1,\ldots,w_n,\ldots$, where

$$w_n = v_{1,n} + \cdots + v_{w_p,n}$$

is an orthonormal basis for $M(t,v)$ over $A(\mathscr{W}^*(t))$.

The fact that U^* is an operator over $A(\mathscr{W}^*(t))$ follows from the fact that U commutes with $\langle d\rangle$, for all $d \in D$. Complete continuity follows immediately from the facts that the operator is completely continuous over $A(\mathcal{B}^*(t))$ and that $|\ |^*$ is equivalent to the supremum norm. The proof of the existence of $Q_N(T)$ follows the same lines as the proof of the existence of $P(s,T)$. Finally, the last assertion follows from elementary linear algebra. □

By the q-expansion over $\mathcal{W}^*$ of an element $F \in A(\mathcal{V}^*)$, we mean the series $\sum_n a_n q^n \in A(\mathcal{W}^*)[[q]]$ where $a_n = \sum_{d\in D} a_{n,d}\langle d\rangle$, $a_{n,d} \in A(\mathcal{B}^*)$ and $\sum a_{n,d}q^n$ is the q-expansion of $F|\langle d^{-1}\rangle$.

As in Sect. B1, using (1), we may think of $\mathcal{W}^*$ as a subspace of $\mathcal{W}$ containing the image of $\mathbf{Z}$. When $k = (s,i) \in 2(\mathbf{Z}_p \times \mathbf{Z}/w_p\mathbf{Z}) \subset \mathcal{W}$, $k \neq 0$, $G_k(q)$ is the q-expansion of a Serre modular form of weight k [S-MZp, Sect. 1.6].

B4. Non-integral weight

Recall, K equals $\mathbf{C}_p$ or is a complete discretely valued subfield, $\mathcal{B}^* = B(0, |\pi/\mathbf{q}|)$, $w_p = \mathrm{LCM}(p-1, 2)$, $D = (\mathbf{Z}/\mathbf{q}\mathbf{Z})^*$, $\mathcal{W}^* = \mathcal{B}^* \times \mathbf{Z}/w_p\mathbf{Z}$ and $\langle\langle a\rangle\rangle = a/\tau(a)$, for $a \in \mathbf{Z}_p^*$.

In this section, we will give definitions of a q-expansion of an overconvergent form of non-integral weight and of overconvergent families of modular forms.

As in Sect. B1, $Z_1(N\mathbf{q})$ denotes the affinoid subdomain of $X_1(N\mathbf{q})$ which is the connected component of the ordinary locus containing the cusp ∞. (In the notation of Sect. B2, this is also $X_1(N\mathbf{q})(0)$.)

Definition. *We say $F(q) = \sum_{n=0}^{\infty} a_n q^n$, $a_n \in K$, is the q-expansion of an overconvergent form on $\Gamma_1(N\mathbf{q})$ with weight $k = (s,i) \in \mathcal{W}^*$ over K if $F(q)/E(q)^s$ is the q-expansion of an overconvergent function on $Z_1(N\mathbf{q})$ in $X_1(N\mathbf{q})$ of character τ^i for the action of D.*

If $\mathcal{U}$ is an admissible open subspace of $\mathcal{B}^*$ we also say that

$$F_s(q) = \sum_{n=0}^{\infty} a_n(s) q^n ,$$

*$a_n(s) \in A(\mathcal{U})$, is the q-***expansion of a family of overconvergent forms** *over $\mathcal{U}$ on $\Gamma_1(N\mathbf{q})$ if $F_s(q)/E(q)^s$ is the q-expansion of an overconvergent function on $\mathcal{U} \times Z_1(N\mathbf{q})$ over $\mathcal{U}$. We say this family has type $i \in \mathbf{Z}/w_p\mathbf{Z}$, if this function has character τ^i for the action of D and is an eigenvector for U^* with eigenvalue $f(s) \in A(\mathcal{U})$ if $U^*(H) = f(s)H$.*

For $k \in \mathcal{W}^*$, let $M_k^\dagger(N)$ denote the vector space over K of weight k overconvergent modular forms on $\Gamma_1(N)$, let $M^\dagger(N)$ denote the $A(\mathcal{B}^*)$ module of families of overconvergent forms over $\mathcal{B}^*$ on $\Gamma_1(N)$ and, for $i \in \mathbf{Z}/w_p\mathbf{Z}$, $M^\dagger(N,i)$ the subspace of those of type i. Also let $S^\dagger(N,i)$ denote the subspace of cusp forms in $M^\dagger(N,i)$. In the notation of Sect. B3,

$$M^\dagger(N,i) = \varprojlim_{t \leq |\pi/\mathbf{q}|} \varinjlim_{(t,v)\in\mathcal{T}^*} M(t,v,\tau^i)$$

and

$$S^{\dagger}(N,i) = \varprojlim_{t \leq |\pi/\mathbf{q}|} \; \varinjlim_{(t,v)\in\mathcal{T}^*} S(t,v,\tau^i)\,.$$

Clearly, if $F(q)$ is an overconvergent form of weight k and $G(q)$ is an overconvergent form of weight j, $F(q)G(q)$ is an overconvergent form of weight $k+j$. We will show, in a future article, that if $k=(s,i)$, $F(q)$ is the q-expansion of a generalized Katz modular function with weight character $z \mapsto \langle\langle a\rangle\rangle^s \tau(z)^i$ and the q-expansion of a family of modular forms with integral q-expansions over a rigid space $X \subset \mathcal{B}^*$ is the q-expansion of a Katz modular function over $A^0(X)$ [K-pIE] (see also [G-ApM, Sect. I.3]). Also,

$$M^{\dagger}(N) = \bigoplus_{i\in\mathbf{Z}/w_p\mathbf{Z}} M^{\dagger}(N,i)$$

and if $A^{\dagger}(N)$ denotes the $A(\mathcal{B}^*)$ algebra of overconvergent functions on $\mathcal{B}^* \times Z_1(N\mathbf{q})$ over $\mathcal{B}^*$, then $A^{\dagger}(N)$ is isomorphic to $M^{\dagger}(N)$ as an $A(\mathcal{B}^*)$ module. For $k=(s,i)\in\mathcal{W}^*$, we have natural homomorphisms

$$M^{\dagger}(N) \to M^{\dagger}(N,i) \to M_k^{\dagger}(N)\,,$$

where the first arrow is the projection and the second is restriction.

Theorem B4.1. *Suppose $i\in 2\mathbf{Z}/w_p\mathbf{Z}$. Then $G_{s,i}$ is an overconvergent family of eigenforms over $\mathcal{B}^*$, if $i \neq 0$, and over $\mathcal{B}^* - \{0\}$, if $i=0$, on $\Gamma_1(\mathbf{q})$ of type i with eigenvalue 1 for U^*.*

Proof. First we observe that the set of cusps C in $Z_1(\mathbf{q})$ has order $w_p/2$. For $c\in C$, let $]c[$ denote the residue disk in $Z_1(\mathbf{q})$ containing c. We may regard q as a parameter on the residue disk $]\infty[$ of the cusp ∞. Fix $(t,v)\in\mathcal{T}^*$, $t>1$, and let $A=A(Z_t(v))$. Let $I_C \subset A$ be the ideal of $B[0,t]\times C$. The homomorphism $h: A \to A/I_C A := B$ is respected by U^* and by the diamond operators.

Fix $i\in 2\mathbf{Z}/w_p\mathbf{Z}$. We will work on the τ^i eigensubspace of A for the action of D, A_i, which maps onto the τ^i eigensubspace of B, B_i, and this latter is free of rank one over $A(B[0,t])$. Since U^* commutes with the diamond operators and the constant term of the q-expansion of a form F is the same as that of $U^*(F)$, the following diagram commutes:

$$\begin{array}{ccc} A_i & \xrightarrow{\;\;h\;\;} & B_i \\ \Big\downarrow {\scriptstyle U^*} & & \Big\downarrow {\scriptstyle \mathrm{id}} \\ A_i & \xrightarrow{\;\;h\;\;} & B_i \end{array}$$

It follows from Lemma A2.4, since the absolute values are discrete, that $1-T$ divides $P_i(s,T)$. Restricting s to an integer $k>2$ and using (8.2) and [C-CO, Theorem 8.1] we see that $(1-T)^2$ does not divide $P_i(s,T)$ since every $U_{(k)}$-eigenvector in $M_k(\tau^{i-k})$ with eigenvalue 1 is a classical modular form of weight k and character τ^{i-k} and the dimension of these is one. Let

$X \subseteq B[0,t]$ be an affinoid such that $X \times \{1\}$ lies in the complement of the zero locus S of $\Delta P_i(s,T)$. Then our Riesz theory, Theorem A4.5, tells us that the eigenspace of U^* over X with eigenvalue 1 is locally free of rank one. In fact, using the map h and the above commutative diagram, we see this module is free spanned by $\Pi(H) =: F$ where Π is the Riesz theory projector onto the eigenvalue 1 subspace and H is any function in A_i which maps to the element of B_i which is 1 along $B[0,t] \times \infty$. We may suppose that X contains infinitely many integers greater than or equal to 2. Then for any such integer k, we know $F_k(q) = E_{k,i}(q)/E(q)^k$ since in this case we know $E(q)^k F_k(q)$ must be the q-expansion of a classical modular form and the q-expansions of F_k and $E_{k,i}/E^k$ have the same constant term, 1. Since the q-expansion coefficients of F must be analytic on X, We see that $F_s(q) = E_{s,i}(q)/E(q)^s$ for all $s \in X$ such that $L_p(1-s,\tau^i) \neq 0$. Since this is true for any affinoid X in $B[0,t] - S$ we conclude that $L_p(1-s,\tau^i)F_s =: H_s$ is an overconvergent analytic function on $(B[0,t]-S) \times Z_1(\mathbf{q})$. But the q-expansion of H_s clearly extends to $B[0,t] \times]C(\infty)[$ when $i \neq 0$ and to $(B[0,t] - \{0\}) \times]\infty[$ when $i = 0$. Hence as $G_{s,i}(q) = H_s(q)E(q)^s$ for $|s| \leqq t$, $G_{s,i}(q)$ is the q-expansion of a family of forms over $B[0,t]$ when $i \neq 0$ and over $B[0,t] - \{0\}$ when $i = 0$. Since this is true for any t such that $t < |\pi/\mathbf{q}|$ the theorem follows. □

Corollary B4.1.1. *For $(s,i) \in \mathcal{W}^*$, $(s,2i) \neq (0,0)$, there exists an overconvergent form of type $(s,2i)$ with q-expansion $G_{s,2i}(q)$.*

Corollary B4.1.2. *For each $i \in 2\mathbf{Z}/w_p\mathbf{Z}$ there exists an overconvergent function F_i on $\mathcal{B}^* \times Z_1(\mathbf{q})$ such that*

$$F_i(s,q) = G_{s,i}(q)/G_{s,i}(q^p)\,.$$

We also see that $E_{s,i}(q)$ is a family of overconvergent forms over the complement in $\mathcal{B}^*$ of the zeroes of $L_p(1-s,\tau^i)$. So when $i=0$, it is a family of overconvergent forms over all of $\mathcal{B}^*$. In particular, we can replace $E^s(q)$ with $E_{s,0}$ in our definition of overconvergent forms of non-integral weight and of families of overconvergent forms.

Remark B4.2. *We could now upgrade our Fredholm theory by using the function F_0 in place of e^s. Let $\mathcal{T}$ denote the subset of $\mathcal{T}^*$ consisting of pairs (t,v) such that F_0 converges on $Z_t(v)$. (The set $\mathcal{T}$ also projects onto $\mathbf{Q} \cap [1,|\pi/\mathbf{q}|)$.) Let $\mathcal{V}$ denote the rigid subspace of $\mathcal{V}^*$ admissibly covered by the affinoids $Z_t(v)$ where $(t,v) \in \mathcal{T}$. Let U be the operator $h \mapsto U_{(0)}(hF_0)$ on $A(\mathcal{V})$. It is a completely continuous operator on $A(Z_t(v))$ for each $(t,v) \in \mathcal{T}$. It also sits in a commutative diagram*

$$\begin{array}{ccc} A(\mathcal{V}) & \xrightarrow{m_f} & A(\mathcal{V}) \\ \downarrow{\scriptstyle U^*} & & \downarrow{\scriptstyle U} \\ A(\mathcal{V}) & \xrightarrow{m_f} & A(\mathcal{V}) \end{array}$$

where f is the function with q-expansion $E(q)^s/E_{s,0}(q)$ which is a unit using Theorem B4.1 *and the fact that $E_{s,0}(q)/E(q)^s$ is congruent to* 1. *It follows* *that the characteristic power series of* U *is the same as that of* U^*. *The reason why this is an improvement, is that the q-expansion coefficients of* F_0 *are Iwasawa functions, so that* U *preserves the submodule of* $A(\mathscr{V})$ *consisting of elements whose q-expansions are Iwasawa functions. This will be used, in a subsequent article* [C-CPS], *to give a conceptual proof of the fact, proven in the appendix, that the coefficients of* $Q_N(T)$ *are Iwasawa functions and to remove our restriction to the subspace* $\mathscr{W}^*$ *of* $\mathscr{W}$.

Suppose that the tame level N equals 1 for the rest of this section.

For $i \in 2\mathbf{Z}/w_p\mathbf{Z}$, we have an overconvergent function $E_{(i)}$ defined on $\mathscr{V}^*$ away from the fibers above the zeroes of $L_p(1-s,\tau^i)$ such that

$$E_{(i)}(s,q) = E_{s,i}(q)/E_{s,0}(q)\,.$$

It follows that $E_{(i)}|\langle d\rangle = \tau(d)^i E_{(i)}$ for $d \in \mathbf{Z}_p^*$.

Theorem B4.3. *Suppose* $i \in 2\mathbf{Z}/w_p\mathbf{Z}$. *Suppose* F *is an overconvergent function on* $\mathscr{V}^*$ *which satisfies*

$$F(\infty) = 1 \quad \textit{and} \quad F|\langle d\rangle = \tau^i(d)F$$

for $d \in \mathbf{Z}_p^*$, *then away from the zeroes of* $\Delta P_i(s,1)$ *and* $L_p(1-s,\tau^i)$,

$$\frac{-U^*R_{\tau^i}((s,1),U)}{\Delta P_i(s,1)}F(s) = E_{(i)}(s)\,. \tag{3}$$

Proof. We know, for k an integer at least 2 and $i \in 2\mathbf{Z}/w_p\mathbf{Z}$, the $U_{(k)}$-eigensubspace of $M_k(\tau^{i-k})$ with eigenvalue 1 is one dimensional and spanned by $G_{k,i}$. It follows, in particular, that

$$U^*E_{(i)} = E_{(i)}\,.$$

Thus $1-T$ divides $P_i(s,T)$ and since the aforementioned eigenspaces have dimension one, $1-T$ divides $P_i(s,T)$ simply. Our Riesz theory, the uniqueness of analytic continuation and the fact that the two sides of (3) agree at the cusps now implies that it holds whenever both sides are defined. □

Remark B4.4. *Overconvergent functions like* F *certainly exist, for example, we can take* F *to be the function (which is "constant in the* s *direction")* $E_{m,i}/E_{m,0}$ *where* m *is an integer at least* 1.

It is clear that $P_i(s,T)/(1-T) = P_i^0(s,T)$. Then since the polar divisor of $E_{(i)}$ is the divisor of zeroes of $L_p(1-s,\tau^i)$, equation (3) implies:

Corollary B4.3.1. *For* $i \in 2\mathbf{Z}/w_p\mathbf{Z}$, $L_p(1-s,\tau^i)$ *divides* $P_{\tau^i}^0(s,1)$ *in* $A(\mathscr{B}^*)$.

Remarks B4.5. (i) *We will show, in a future article, that when* $i \neq 0$ *or* 2, *that* $P_{\tau^i}^0(s,1)$ *is the product of a unit in* $\Lambda \subset A(\mathscr{B}^*)$ *and the function* $D(\tau^{i-2},s-2)$ *of Mazur and Wiles* [*MW*]. (ii) *Suppose* $p \equiv -1 \bmod 4$. *If* ψ *is a non-trivial* *character on the class group of* $\mathbf{Q}(\sqrt{-p})$ *then*

$$\sum_{\mathscr{A}} \psi(\mathscr{A}) q^{\mathrm{N}\mathscr{A}},$$

where $\mathscr{A}$ *runs over the ideals of* $\mathbf{Z}[\sqrt{-p}]$ *and* $\mathrm{N}\mathscr{A}$ *is the norm of* $\mathscr{A}$, *is the* q*-expansion of a weight one cusp form on* $\Gamma_1(p)$ *with character* $\chi = \tau^{(p-1)/2}$ *fixed by* U^* *and so* $\Delta^i P_\chi^0(1,T)|_{T=1} = 0$ *for* $0 \leqq i < h-1$ *where* h *is the class number of* $\mathbf{Q}(\sqrt{-p})$. (iii) *In particular, when* $p \equiv -1 \bmod 4$, $p \geqq 23$ *and* p *doesn't equal* 43, 67 *or* 161, $D_p(\tau^{(p-3)/2}, 1) = 0$.

We also deduce from the proof of the theorem,

Corollary B4.5.2. *For each* $i \in 2\mathbf{Z}/w_p\mathbf{Z}$, $E_{s,i}(q)$ *is the* q*-expansion of an over-convergent family of eigenforms of type* i *on the complement of the zero locus of* $L_p(1-s, \tau^i)$ *in* $\mathscr{B}^*$.

B5. Hecke operators and R-families

In this section we eschew the notion of "radius of overconvergence" (i.e. we ignore how far into the supersingular region an overconvergent object converges). We will prove a qualitative version of the Mazur–Gouvêa conjecture on the existence of "R-families" (Conjecture 3 of [GM]) in this section. This conjecture asserts that for any classical eigenform f of weight k, tame level N and slope α there is a finite flat $\mathbf{Z}_p[[T]]$ algebra R, a power series $F = \sum_{n=1} r_n q^n$ with coefficients in R and homomorphisms $\eta_j : R \to \mathbf{C}_p$ for j an integer such that $|j-k| \leqq p^{-\alpha}$ and $j > \alpha + 1$ such that $f_j(q) := \sum_{n=1}^{\infty} \eta_k(r_n) q^n$ is the q-expansion of a classical weight j modular form of tame level N and slope α and $f_k(q) = f(q)$.

For $d \in (\mathbf{Z}/N\mathbf{Z})^* \times \mathbf{Z}_p^*$, we will let $\langle\langle d \rangle\rangle$ denote $\langle\langle d_p \rangle\rangle$ where d_p is the projection of d into $\mathbf{Z}_p$. Recall, K is a either $\mathbf{C}_p$ or a complete discretely valued subfield and $M^\dagger(N)$ is the $A(\mathscr{B}^*)$ module of families of overconvergent forms.

We define an action of Hecke on $M^\dagger(N)$. First, if $l \in (\mathbf{Z}/N\mathbf{Z})^* \times \mathbf{Z}_p^*$ we define

$$(F|\langle l \rangle^*)_s(q) = \langle\langle l \rangle\rangle^s E^s(q) \left(\frac{F_s}{E^s} \bigg| \langle l \rangle \right)(q),$$

for $s \in \mathscr{B}^*$. When $k \in \mathbf{Z}$,

$$(F|\langle l \rangle^*)_k = l^k F_k |\langle l \rangle . \tag{0}$$

Next, generalizing the notation of Sect. B2, if n and M are relatively prime positive integers, we let $\Gamma(M;n)$ denote the congruence subgroup $\Gamma_1(M) \cap \Gamma_0(n)$ of $SL_2(\mathbf{Z})$ and $X(M;n)$ the corresponding modular curve over K. We can repeat all of our previous constructions and definitions in this situation and we will use obvious extensions of our previous notations. For example, if $(Nn, p) = 1$, $Z(Np^m; n)$ denotes the rigid connected component of the ordinary locus in $X(Np^m, n)$ containing the cusp ∞ and $A^\dagger(N;n)$ denotes the $A(\mathscr{B}^*)$-algebra of overconvergent functions on $\mathscr{B}^* \times Z(N\mathbf{q};n)$ over $\mathscr{B}^*$.

If F is modular form on $X_1(M)$ and l is a prime not dividing M, we let $F|V_l$ denote the modular form on $X(M;l)$ such that

$$F|V_l(Y,\alpha,C,\omega) = F(Y/C, l^{-1}v \circ \alpha, \check{v}^*\omega)$$

where Y is an elliptic curve, $\alpha : \mu_M \hookrightarrow Y$ is an injective homomorphism, C is a cyclic subgroup of Y of order l, ω is a non-vanishing differential on $Y, v : E \to E/C$ is the natural isogeny and $\check{v}$ is its dual. Then

$$F|V_l(q) = F(q^l)\,.$$

Since $E(q)/E(q^l)$ is congruent to 1 modulo $\mathbf{q}$, and both E and $E|V_l$ have weight $(1,0)$, there is an element e_l^s in $A^\dagger(1,l) \subset A^\dagger(l)$ whose q-expansion is $(E(q)/E(q^l))^s$ and which is invariant under the action of D.

For prime l, let ψ_l be the operator on $A(\mathscr{B}^*)[[q]]$

$$\psi_l\left(\sum_n a_n q^n\right) = \sum_n a_{nl} q^n\,.$$

Lemma B5.1. *For each prime number l there is a unique continuous operator $T(l)$ on $M^\dagger(N)$ such that, for $F \in M^\dagger(N)$, when $l = p$,*

$$(F|T(p))_s = E^s U^*\left(\frac{F_s}{E^s}\right),$$

when $l|N$

$$F|T(l)(q) = \psi_l(F(q))$$

and when $l \not| Np$

$$(F|T(l))(q) = \psi_l(F(q)) + l^{-1}(F|\langle l\rangle^*)(q^l)\,. \qquad (1)$$

Proof. If $l = p$, there is nothing to prove. When $l|N$ one may verify this lemma by first showing, using a correspondence, in the usual way, that for $g \in A^\dagger(N)$, there is an element in $A^\dagger(N)$ with q-expansion $\psi_l(g(q))$ (see [Sh, Sect. 7.3] or [C-PSI, Sect. 8]), and then observing that

$$(\psi_l(F(q)))_s = E^s(q)\psi_l\left(\left(\frac{F_s}{E^s}e_l^s\right)(q)\right).$$

Now suppose $l \not| Np$. If $G_s(q)$ is the right hand side of (1) (at s),

$$\frac{G_s(q)}{E^s(q)} = \psi_l\left(\frac{F_s}{E^s}e_l^s\right)(q) + l^{-1}\langle\langle l\rangle\rangle^s e_l^{-s}\left(\frac{F_s}{E^s}\bigg|\langle l\rangle\right)(q^l)$$

which, by the previous discussion, is the q-expansion of a function in $A^\dagger(N;l)$. Moreover, when k is an integer, G_k is clearly on $\Gamma_1(N)$, since the specialization of (1), in this case, is the classical formula for the l-th Hecke operator acting on the overconvergent modular form F_k of weight k. Now consider, the function in $A^\dagger(N;l), \mathrm{Tr}(G_s/E^s) - (l+1)G_s/E^s$, where Tr is the trace map from level $\Gamma(N;l)$ to level $\Gamma_1(N)$. By what we have said, it is zero when s is an integer. It follows that it is zero for all s, since it is an analytic function. This implies the lemma as $\mathrm{Tr}(G_s/E^s)$ is on $\Gamma_1(N)$. □

Remark B5.2. *Our proof implies that $T(l)$ acts on families of forms F_s such that $F_s(q)/E(q)^s$ converges on some strict neighborhood of $\mathscr{B}^* \otimes Z_1(N\mathbf{q})$ which depends only on l. With a little more care we can show that one can use the same neighborhood for all l (at least when $p \neq 2$). The key fact needed to prove this is: If R is a ring of characteristic p and A is the Hasse invariant form over R, then if E/R is an elliptic curve, ω generates $H^0(E, \Omega^1_{E/R})$ and $\gamma\colon E \to \gamma E$ is an isogeny of degree prime to p,*

$$A(E,\omega) = A(\gamma E, \check{\gamma}^*\omega)$$

where $\check{\gamma}\colon \gamma E \to E$ is the isogeny dual to γ.

Let $\mathbf{T} := \mathbf{T}_K$ denote the $A(\mathscr{B}^*)$-algebra generated over $A(\mathscr{B}^*)$ by the operators $\langle d\rangle^*$ for $d \in (\mathbf{Z}/N\mathbf{q}\mathbf{Z})^*$ and $T(l)$. Similarly, if L is an extension of K in $\mathbf{C}_p$ and $k \in \mathscr{W}(L)$ we may define operators $\langle d\rangle_k^*$ for $d \in (\mathbf{Z}/N\mathbf{q}\mathbf{Z})^*$ and $T_k(l)$ for primes l. We let $\mathbf{T}_{L,k}$ denote the L-algebra generated by these over L. We define additional operators $T(n)$, for positive integers n in $\mathbf{T}_K$ by the formal identity:

$$\sum_{n\geqq 1} \frac{T(n)}{n^t} = \prod_{l|Np} (1 - T(l)l^{-t})^{-1} \prod_{(l,Np)=1} (1 - T(l)l^{-t} + \langle l\rangle^* l^{-1-2t})^{-1},$$

where the products are over primes l and when $k \in \mathscr{W}(L)$, we define $T_k(n)$ in $\mathbf{T}_{L,k}$, similarly. When, $k \in \mathbf{Z}$, it follows from equation (0) that $\langle d\rangle_k^* = d^k\langle d\rangle$ and hence $\mathbf{T}_k$ is the usual Hecke algebra acting on overconvergent weight k modular forms on $\Gamma_1(N\mathbf{q})$ (see [G, Ch. II]).

We now prove the assertions in Remark B3.6(i). Let notation be as in the proof of Theorem B3.5 in Sect. B3. In particular, $\varepsilon \in \hat{D}$, α is a rational number and B is a disk in $\mathscr{B}^*$ about an integer k such that the affinoid $\{z \in \mathscr{B}^* \times \mathbf{A}^1 : P_\varepsilon^0(z) = 0, v(T(z)) = -\alpha\}$ has degree one over B and $s \mapsto (s, f(s))$ is the corresponding section. Also, G is a function on $B \times W_0^\dagger$ which vanishes on the cusps and spans the kernel of $U^* - f$ in $S(\varepsilon)_B$.

Lemma B5.3. *If the q-expansion of G is*

$$\sum_{n\geqq 1} a_n(s)q^n,$$

then function $a_1(s)$ is invertible and $|a_n(s)/a_1(s)| \leqq 1$ for $n \geqq 1$ and $s \in B$.

Proof. First, note that $F_s(q) := E^s(q)G(s)(q)$ is (the q-expansion of) an eigenform for $\mathbf{T}$. Suppose $T(n)F_s = c_n(s)F_s$. If

$$F_s(q) = \sum_{n\geqq 1} b_n(s)q^n,$$

$b_1(s) = a_1(s)$ and we see that

$$c_n(s)a_1(s) = b_n(s)\,.$$

So if $a_1(s_0) = 0$, $F_{s_0}(q) = 0$ and this implies $G(s_0) = 0$. We see this is impossible using Lemma A2.5 and our Riesz theory.

Now, it is easy to see that the operator $T(l)$ is bounded by one on the relevant Banach spaces if $l \neq p$. This and the fact that the coefficients of the characteristic power series of U^* lie in Λ imply that $|c_n(s)| \leqq 1$ for all n. This completes the proof. □

It is clear that we have a natural homomorphism, $h \mapsto h_k$, from $\mathbf{T}$ onto $\mathbf{T}_k$ for $k \in \mathscr{W}^*(L)$ which takes $\langle d\rangle^*$ to $\langle d\rangle_k^*$ and $T(n)$ to $T_k(n)$. Also,

Lemma B5.4. *If $k \in \mathscr{W}^*(L)$, $h \in \mathbf{T}$ and $F \in M^\dagger(N)$ then*

$$(hF)_k = h_k F_k\,.$$

R-families. Before we proceed, we point out that if $\alpha \in \mathbf{Q}$, $\alpha \neq 0$, the slope α subspace of $M_k(N)$ is canonically isomorphic to the slope α subspace of $M_k^\dagger(N)$ and we identify the two. For a rigid space $\mathscr{U} \subseteq \mathscr{B}^*$ and an element $F \in M^\dagger(N)_{\mathscr{U}}$, we let $a_n(F) \in A(\mathscr{U})$ denote the coefficient of q^n in its q-expansion.

Suppose α is a rational number, i is an integer such that $0 \leqq i < w_p$ and $k_0 \in \mathscr{B}^*(K)$. Suppose $0 < r < |\pi/\mathbf{q}|$ and $r \in |K|$ such that the slope α affinoid in the zero locus Z^0 of $P_i^0(s,T)$ (i.e., the affinoid whose closed points are the closed points P in Z^0 such that $v(T(P)) = -\alpha$) is finite of degree d over the affinoid disk $B = B_K[k_0, r]$. (We know $d = \mathbf{d}^0(k_0, \alpha, \tau^{i-k})$ if k_0 is an integer and $k_0 > \alpha + 1$.) This disk exists by Corollary A5.5.1. Let $A = A(B)$. Suppose Q is the corresponding factor of $P_i^0(s,T)$ over B. (Recall, $P_i^0(s,T)$ equals $P_{\tau^i}^0(s,T)$, which is morally the characteristic series of the U^* operator on $S(N,i)$.) Then, Q satisfies the hypotheses of Theorem A5.3, so the A-module $H := N_{U_B}(Q)$, where U_B is the restriction of U^* to $S(N,i)_B$, is projective of rank d over A. Since A is a PID, this module is, in fact, free. Let R denote the image of $\mathbf{T} \otimes A$ in $\mathrm{End}_A(H)$. Since $\mathrm{End}_A(H)$ is free of rank d^2 it follows that R is also free of finite rank. In particular, R is the ring of rigid analytic functions on an affinoid $X(R)$ with a finite morphism to B.

We have an A-bilinear pairing

$$\langle\ ,\ \rangle : R \times H \to A$$

$$\langle h, m\rangle = a_1(hm)\,.$$

Similarly, if $k \in \mathscr{B}^*(L)$, we have an L-bilinear pairing $\langle\ ,\ \rangle_k$ from $R_k \times H_k$ to L. (In our previous terminology, we are actually working over the point (k,i) of $\mathscr{W}(L)$.)

Proposition B5.6. *The pairing $\langle\ ,\ \rangle$ is perfect.*

Proof. First, arguing exactly as in the proof of [H-LE, Theorem 5.3.1], we see that if $h \in R$, $\langle h,m \rangle = 0$ for all $m \in H$ implies $h = 0$ and $\langle h,m \rangle = 0$ for all $h \in R$ implies $m = 0$. The key point is that if $F \in H$, $\langle T(n),F \rangle$ equals the n-th q-expansion coefficient of F.

Now if $k \in B(\mathbf{C}_p)$, the same argument yields the same conclusion for the pairing $\langle\ ,\ \rangle_k : \mathbf{T}_k \times H_k \to \mathbf{C}_p$, but since this is a pairing over a field, it follows that $\langle\ ,\ \rangle_k$ is perfect.

Since A is a PID, it suffices to check that the homomorphism,

$$\gamma : R \to \mathrm{Hom}_A(H,A)$$

$$\gamma(h)(n) = \langle h,n \rangle ,$$

is an isomorphism. By Lemma B5.4, if $h \in R$ and $m \in H$, the restriction of $\langle h,m \rangle$ to k is $\langle h_k, m_k \rangle_k$. Since H is free, $\mathrm{Hom}_A(H,A)_k = \mathrm{Hom}_{\mathbf{C}_p}(H_k, \mathbf{C}_p)$. Thus γ_k is an isomorphism for all $k \in B(\mathbf{C}_p)$. This implies γ is an isomorphism and the proposition follows. □

Corollary B5.6.1. *If k_0 is an integer and $k_0 > \alpha + 1$, the degree of $X(R) \to B$ is $\mathbf{d}^0(k_0, \alpha, \tau^{i-k})$.*

Theorem B5.7. *Suppose $L \subset \mathbf{C}_p$ is a finite extension of K. For $x \in X(R)(L)$, let $\eta_x : R \to L$ be the corresponding homomorphism and set*

$$F_x(q) = \sum_{n \geqq 1} \eta_x(T(n))q^n .$$

Now suppose k is an integer such that $k \in B(K)$ and $k > \alpha + 1$. Then the mapping from $X(R)_k(L)$ to $L[[q]]$, $x \in X(R)_k(L) \mapsto F_x(q)$, is a bijection onto the set of q-expansions of classical cuspidal eigenforms on $X_1(N\mathbf{q})$ over L of weight (k,i) and slope α.

Proof. After extending scalars we may suppose $L = K$. First, suppose $x \in X(R)_k(K)$. Then it follows from the proposition and the freeness of R that there is an $m \in H$ such that $(\langle h,m \rangle)_k = \eta_x(h)$. This equals $a_1(h_k m_k)$ by Lemma B5.4. Since η_x is homomorphism, m_k is an eigenform. It also follows that $F_x(q)$ is the q-expansion of m_k and since $k > \alpha + 1$ that m_k is classical.

Now suppose $F(q) = \sum_{n \geq 1} a_n q^n$ is the q-expansion of a weight k cuspidal eigenform on $X_1(N\mathbf{q})$ of weight (k,i) and slope α. It follows that $F(q) \in H_k$. Hence gives rise to a K-linear map $\eta : \mathbf{T}_k \to K$, $\eta(h) = \langle h, F(q) \rangle_k$. Since F is an eigenform, η is a ring homomorphism, so corresponds to a point $x \in X(R)_k(K)$. Finally, since $\langle T(n), F(q) \rangle = a_n$, $F_x(q) = F(q)$. □

We can show that the subring of R generated over $A^0(B)$ by the $T(n)$ is finite of degree d over this ring. When $r \in |K|$, $A^0(B)$ is isomorphic to $K^0[[T]]$ where K^0 is the ring of integers in K. From this, it is not hard to see that the R-family conjecture of Gouvêa–Mazur would follow from the assertion that the radius r of the disk B about k_0 can be chosen to be at least $p^{-\alpha}$.

As Glenn Stevens pointed out, we also have

Corollary B5.7.1. *Suppose k_0 is an integer, $k_0 > \alpha + 1$ and F is an eigenform, new away from p, on $X_1(N\mathbf{q})$ of weight (k_0, i) and slope α. If $i = 0$, F has character $\varepsilon = \varepsilon_N \varepsilon_p$ and $U^*F = aF$, suppose in addition that $a^2 \neq \varepsilon_N(p)p^{k_0-1}$. Then there exists an affinoid disk B' containing k_0 and rigid analytic functions $a_n(s)$ on B' such that if k is an integer strictly greater than $\alpha + 1$ in B'*

$$F_k(q) := \sum_n a_n(k)q^n$$

is the q-expansion of a classical cuspidal eigenform on $X_1(N\mathbf{q})$ of weight (k, i) and slope α which is equal to F if $k = k_0$.

Before beginning the proof we need to discuss families of new forms.

Definition. *We say an overconvergent modular form of weight k (or a family of overconvergent modular forms) on $\Gamma_1(N\mathbf{q})$ is a* **p'-new form** *(or a* **family of p'-new forms***) if its image in $M_k(d)$ (or $M^\dagger(d)$) is zero under any of the degeneracy "trace" maps for any proper divisor d of N.*

We note that the image of a classical modular form is new in this sense if and only if it is new "away from p."

We denote the Banach module of p'-new forms of weight k by $M_k^{p'-nw}(N)$ and of families of p'-new forms by $M^\dagger(N)^{p'-nw}$. Now, U^* acts completely continuously on this module. We now restrict U^* to $M^\dagger(N, i)^{p'-nw}$. Let

$$P_i^{p'-nw}(s, T) = \det(1 - TU^*|M^\dagger(N, i)^{p'-nw}) .$$

Everything we said above about $P_i(s, T)$ carries over to $P_i^{p'-nw}(s, T)$ and we will use the same notations. In particular, now B is an affinoid disk such that the slope α affinoid in the zero locus of $P_i^{p'-nw}(s, T)$ is finite over B and R now denotes the image of $\mathbf{T} \otimes A(B)$ in the endomorphism ring of the $A(B)$-module of families of p'-new forms of slope α over B. The form F corresponds to a point x of $X(R)$ by the theorem. It suffices to prove that the morphism $X(R) \to B$ is unramified at x for then we will have a section s in a neighborhood of k_0 such that $s(k_0) = x$ and we may take $F_k(q) = F_{s(k)}$ for $k \in \mathbf{Z} \cap B$. This assertion follows from the fact that the classical Hecke algebra acts semi-simply on the space of classical p'-new forms on $X_1(Np)$ satisfying the hypotheses of the corollary. This in turn follows from the well known fact that the Hecke algebra on $\Gamma_1(M)$ acts semi-simply on the space of new forms on $\Gamma_1(M)$ for each positive integer M (see [Li, Lemma 6 iii)]), the fact that the classical p'-new forms on $\Gamma_1(pN)$ is the sum of the new forms on this group and the images of the new forms on $\Gamma_1(N)$ and the Lemma 6.4 of [C-CO] which explains how the U_p operator acts on this space.[3]

We can also define a form or a family of forms to be **p'-old** if it is a sum of elements in the images of $M_k(d)$ (or $M^\dagger(d)$) (under the various natural maps)

[3] See also [CE] which proves that the exceptional case never occurs in weight 2 and discusses its likelihood in higher weights.

where d runs over the proper divisors of N. Although the corresponding statements about classical forms are true, we do not know if every overconvergent form or family of such is a sum of a p'-new form and a p'-old form or whether, if it is, this decomposition is unique.

Theorem B5.7 implies that an eigenform of slope α lives in a family of eigenforms of slope α, but in fact any form of slope α lives in a family of forms of slope α. For each $k \in B(K) \cap \mathbf{Z}$, specialization gives us a map from C into the space $M_k(N,i)_\alpha$ of slope α forms on $\Gamma_1(N\mathbf{q})$ of weight (k,i) (which are classical if $k > \alpha + 1$). We have,

Proposition B5.8. *The map from H to $M_k(N,i)_\alpha$ is a surjection.*

Proof. Suppose $F \in M_k(N,i)_\alpha$. Then we can certainly produce an element $G \in M^\dagger(N,i)$ which specializes at k to F (if we regard $M^\dagger(N,i)$ as functions on $\mathscr{V}^*$, we just take G to be the function F/E^k on $Z_1(\mathbf{q}) \times \mathscr{B}^*$ which is constant in the $\mathscr{B}^*$ direction). Let $\tilde{F}$ be the projection into H of the restriction of G to the fiber above B. Since projection commutes with specialization $\tilde{F}_k = F$. □

B6. Further results

In this section we will explain how our series $P_N(T)$ also "controls" forms on $X_1(Np^m)$ when $(N,p) = 1$ for $n \geqq 1$ (the proofs will appear in [C-CPS]) and indicate the connection between the results of this paper and the theory of representations of the absolute Galois group of $\mathbf{Q}$.

If $\gamma \in \mathbf{Z}_p^*$, let $[\gamma]$ denote the corresponding element in the completed group ring Λ of $\mathbf{Z}_p^*$ over $\mathbf{Z}_p$. Then there exists a unique injective homomorphism α from Λ into $A^0(\mathscr{W}^*)$ such that, for $\gamma \in \mathbf{Z}_p^*$,

$$\alpha([\gamma])(s,i) = \langle\langle\gamma\rangle\rangle^s \tau(\gamma)^i .$$

(In other words, $[\gamma]$ goes to the element $\langle\langle\gamma\rangle\rangle^s\langle\gamma\rangle$ in $A(\mathscr{B}^*)[D] = A(\mathscr{W}^*)$.) It follows that, for $\lambda \in \Lambda$, $\alpha(\lambda)$ is bounded on $\mathscr{W}^*$ and

$$|\alpha(\lambda)|_{\mathscr{W}^*} = |\lambda|_{|\pi|} ,$$

where $|\ |_{|\pi|}$ is one of the absolute values on Λ described in Sect. A1.

We will show,

Theorem B6.1. *The series $Q_N(T)$ lies in $\Lambda[[T]]$ and converges on $\mathscr{W} \times \mathbf{C}_p$.*

We, actually, give one proof of this in the appendix using explicit formulas, but it is also possible to give a more conceptual proof which we do in [C-CPS].

The space of overconvergent forms of level Np^m of integral weight k together with an operator U_k is defined in [C-HCO] (see also Sect. B2). For $\kappa \in \mathscr{C}$ and $F(T) = \sum_{n\geqq 0} B_n T^n \in \Lambda[[T]]$ we set $\kappa(F)(T) = \sum_{n\geqq 0} \kappa(B_n)T^n$. We can map $\mathbf{Z}_p^*$ onto $(\mathbf{Z}/p^n\mathbf{Z})^*$ which is naturally a direct factor of $(\mathbf{Z}/Np^m\mathbf{Z})^*$.

Hence, we may regard characters on $\mathbf{Z}_p^*$ of conductor p^m as characters on $(\mathbf{Z}/Np^m\mathbf{Z})^*$. We also prove in [C-CPS]:

Theorem B6.2. *If $\kappa(x) = \chi(x)\langle x\rangle^k$ where k is an integer, $\chi\colon \mathbf{Z}_p^* \to \mathbf{C}_p^*$ is a character of finite order and $p^n = LCM(\mathbf{q}, f_\chi)$, then $\kappa(P_N)(T)$ is the characteristic series of the operator $U_{(k)}$ on overconvergent modular forms of level Np^m, weight k and character χ.*

The analogue of Theorem C is true in these higher levels. In particular, we prove in [C-HOC] that any form of weight k and level Np^m of slope strictly less than $k-1$ is classical.

The next theorem describes one of the main implications of the combined results of this paper and those of [C-CPS]. For an integer j and χ and a character of finite order on $1+\mathbf{q}\mathbf{Z}$, let $s(\chi,j) = \chi(1+\mathbf{q})(1+\mathbf{q})^j - 1$ and $s(j) = s(\mathbf{1},j)$.

Theorem B6.3. *Suppose ε is a character on $(\mathbf{Z}/\mathbf{q}\mathbf{Z})^*$, $k \in \mathbf{Z}$, $\alpha \in \mathbf{Q}$ and $d(k,\varepsilon,\alpha) = 1$. Then there exists a real number R, a subset S of $B(k,R)$, a function $r\colon S \to \mathbf{R}$ such that, if $X(k,\varepsilon,\alpha) = B(k,R) - \bigcup_{a\in S} B[a,r(a)]$, $s(k) \in X(k,\varepsilon,\alpha)$, there exist rigid analytic functions $a_n(T)$, for $n \geqq 2$, on $X(k,\varepsilon,\alpha)$ bounded by 1 such that if*

$$F(T,q) = q + a_2(T)q^2 + \cdots + a_n(T)q^n + \cdots ,$$

χ is a character of finite order on $1+\mathbf{q}\mathbf{Z}_p$ and j is an integer such that $s(\chi,j) \in X(k,\varepsilon,\alpha)$, $F(s(\chi,j),q)$ is the q-expansion of an overconvergent eigenform $F_{\chi,j}$ of tame level N, weight j, finite slope and character $\tau^{-j}\varepsilon\chi$. Finally, $F_{1,k}$ has slope α.

In fact, we can show $X(k,\varepsilon,\alpha)$ and $a_n(T)$ are defined over $\mathbf{Q}_p$.

We note that one can show that if f is an analytic function on $X(k,\varepsilon,\alpha)$ bounded by 1, and if d and e are in $X(k,\varepsilon,\alpha)$ such that $|d-b| = |e-b| = |s(k)-b|$ for all $b \in S$, then

$$|f(d) - f(e)| < |d-e| \mathrm{Max}\{1/R, r(b)/|s(k)-b|^2 : b \in S\} . \qquad (2)$$

We note that these hypotheses hold when $B[s(k),t] \subset X(k,\varepsilon,\alpha)$ and $d,e \in X(k,\varepsilon,\alpha)$. This implies that Conjecture 2 of [GM] follows from the assertions (which we don't know how to prove):

(i) $B[s(k), p^{-(\alpha+1)}] \subset X(k,\varepsilon,\alpha)$,
(ii) $v(a_p(e)) = \alpha$ if $e \in B[s(k), p^{-(\alpha+1)}]$,
(iii) $R = 1$ and
(iv) $r(b) \leqq |s(k)-b|^2$ for $b \in S$.

Now, let $G(nP)$ be the Galois group of a maximal extension of $\mathbf{Q}$ unramified outside Np. With Mazur, we prove,

Theorem B6.4. *There exists a 2-dimensional pseudo-representation $\pi\colon G(Np) \to T_{\mathbf{Q}}$ such that, for primes $l \nmid Np$,*

$$\mathrm{Trace}(\pi(\mathrm{Frob}_l)) = T(l) \quad \text{and} \quad \det(\pi(\mathrm{Frob}_l)) = \langle l\rangle^*/l .$$

The proof of this is based on the Gouvêa–Hida Theorem (see [G-ApM Theorem III.5.6] and [H-NO Sect. 1]).

Appendix I: Formulas

Fix a positive integer N prime to p. Let $Q_N(T)$ be the characteristic power series of the operator U^* acting on overconvergent forms on $\Gamma_1(N\mathbf{q})$ whose coefficients are in $A(\mathcal{W}^*)$, as in Sect. B3.

For an order $\mathcal{O}$ in a number field, let $h(\mathcal{O})$ denote the class number of $\mathcal{O}$. If γ is an algebraic integer, let O_γ be the set of orders in $\mathbf{Q}(\gamma)$ containing γ. Finally, for m an integer, let $W_{p,m}$ denote the finite set of $\gamma \in \mathbf{Q}_p$ such that $\mathbf{Q}(\gamma)$ is an imaginary quadratic field, γ is an algebraic integer,

$$\mathrm{Norm}_{\mathbf{Q}}^{\mathbf{Q}(\gamma)}(\gamma) = p^m \quad \text{and} \quad v(\gamma) = 0\,. \tag{3}$$

Theorem I1. *Suppose $N \geqq 4$. Then*

$$T\frac{d}{dT}Q_N(T)/Q_N(T) = \sum_{m \geqq 1} A_m T^m$$

where A_m is the element of $\Lambda \subset A(\mathcal{W}^*)$, expressed by the finite sum,

$$A_m = \sum_{\gamma \in W_{p,m}} \sum_{\mathcal{O} \in O_\gamma} h(\mathcal{O}) B_N(\mathcal{O}, \gamma) \cdot \frac{[\gamma]}{\gamma^2 - p^m}$$

where $B_N(\mathcal{O}, \gamma)$ is the number of elements of $\mathcal{O}/N\mathcal{O}$ of order N fixed under multiplication by $\bar{\gamma}$.

(Recall, for $a \in \mathbf{Z}_p^*$, $[a]$ denotes the element of the group of which Λ is the completed group ring.)

Proof. If $\kappa \in \mathcal{W}^*$ is an arithmetic character the specialization of this formula for $\kappa(Q_N)$ may be proven using the Monsky–Reich trace formula, as in Dwork [D1], Katz [K] and Adolphson [A]. The general case follows from the fact that the coefficients of the powers of T in the series $Q_N(T)$ are an analytic functions on $\mathcal{W}^*$. □

Another version of the above theorem is:

Theorem I2. *Suppose $N \geqq 4$. Let Y be the component of the ordinary non-cuspidal locus in the reduction of $X_1(N\mathbf{q})$ containing ∞ and, for x a closed point of Y, $a(x) \in \mathbf{Z}_p^*$ the unit root of Frobenius on the fiber of $E_1(N\mathbf{q})/X_1(N\mathbf{q})$ above x. Then,*

$$Q_N(T) = \prod_{r \geqq 0} \prod_{x \in Y} (1 - \bar{a}(x)^r [a(x)] T^{\deg(x)}/a(x)^{r+2})^{-1}$$

where the second product is over closed points of Y and $\bar{a}(x)$ is the complex conjugate of $a(x)$ in $\mathbf{Z}_p$.

Corollary I2.1. *The coefficients of $Q_N(T)$, as a series in T, lie in the Iwasawa algebra $\mathbf{Z}_p[[\mathbf{Z}_p^*]]$.*

This answers a question of [GM-CS].

Also, using Hijikata's application of the Eichler–Selberg trace formula [Hj], Koike [Ko] proved the specializations of following result to arithmetic characters and the general case follows by analyticity as above.

Theorem I3. *We have the formula,*

$$T\frac{d}{dT}Q_1(T)/Q_1(T) = \sum_{m\geqq 1} B_m T^m$$

where

$$B_m = \sum_{\gamma\in W_{p,m}} \sum_{\mathcal{O}\in O_\gamma} \frac{h(\mathcal{O})}{w(\mathcal{O})} \cdot \frac{[\gamma]}{\gamma^2 - p^m}\,.$$

We note that the specializations of the B_m to $\mathcal{B}^* \times \{i\} \subset \mathcal{W}^*$ are all zero, if i is odd, as they should be, since there are no overconvergent forms of the corresponding weights.

One can generalize these formulas to the moduli problems associated to subgroups of $GL_2(\mathbf{Z}/N\mathbf{q}\mathbf{Z})$ in the sense of Katz–Mazur [KM, Ch. 7] of the form $G \times G_1(\mathbf{q})$ where G is a subgroup of $GL_2(\mathbf{Z}/N\mathbf{Z})$ and $G_1(\mathbf{q})$ is the semi-Borel in $GL_2(\mathbf{Z}/\mathbf{q}\mathbf{Z})$. We will now use the above formulas to prove the existence of nonclassical overconvergent eigenforms.

Proposition I4. *Let k be an integer. Then there exist weight k overconvergent new forms on $\Gamma_1(N\mathbf{q})$ of arbitrarily large slope.*

Proof. We must show the characteristic power series $G_N(k,T)$ of U^* acting on the space of weight k overconvergent new forms on $\Gamma_1(N\mathbf{q})$ is not a polynomial. Let

$$T\frac{d}{dT}G_N(k,T)/G_N(k,T) = \sum_{m\geqq 1} D_m(k)T^m\,.$$

It suffices to show that the numbers $D_m(k)$ are algebraic and are not all defined over a finite extension of $\mathbf{Q}$.

For a positive integer n let $f(n)$ be the number of distinct prime divisors of n and

$$t(n) = \begin{cases} (-2)^{f(n)} & \text{if } n \text{ is square free} \\ 0 & \text{otherwise}\,. \end{cases}$$

Using (10.2) and the two linearly disjoint degeneracy maps from forms on $\Gamma_1(M\mathbf{q})$ to forms on $\Gamma_1(Ml\mathbf{q})$ for primes l and positive integers M prime to p, one can show

$$G_N(k,T) := \prod_{d|N} P^0_{N/d}(k,T)^{t(N/d)}\,.$$

To simplify the argument, we will complete the proof only in the case in which $N = l^t$ where l is an odd prime and $l^{t-1} \geqq 5$.

Suppose M is any integer at least 5. Let $K \subset \mathbf{Q}_p$ be a quadratic field of discriminant D less than $-M$. Also, suppose for simplicity of exposition that $D \not\equiv 1 \bmod 4$. Then there exists an $m \in \mathbf{N}$ and an element $\gamma \in K \cap W_{p,m}$ such that $\bar{\gamma} \equiv 1 \bmod N\mathcal{O}_K$ where $\mathcal{O}_K$ is the maximal order of K. In fact, since $N \geqq 5$, $-\gamma$ is the only other element of $K \cap W_{p,m}$. Then Theorem I2 implies $D_m(k) = \alpha + \beta$ where

$$\alpha = \sum_{\mathcal{O} \in O_\gamma} h(\mathcal{O})(B_N(\mathcal{O},\gamma) - 2B_{N/l}(\mathcal{O},\gamma))\frac{\gamma^k}{\gamma^2 - p^m}$$

and β is a sum of elements contained in quadratic fields different from $K(B_N(\mathcal{O},-\gamma) = B_{N/l}(\mathcal{O},-\gamma) = 0$ for all $\mathcal{O} \in O_\gamma)$. The corollary will follow from the claim: $K = \mathbf{Q}(\alpha)$. It is easy to see that $K = \mathbf{Q}(\gamma^k/(\gamma^2 - p^m))$ and $B_N(\mathcal{O}_K,\gamma) = N^2 - 3(N/l)^2 + 2(N/l^2)^2 > 0$. Thus all we need verify is that

$$\mathbf{C}(\mathcal{O}) =: B_N(\mathcal{O},\gamma) - 2B_{N/l}(\mathcal{O},\gamma) \geqq 0$$

for all $\mathcal{O} \in O_\gamma$. We first observe that the numbers $B_H(\mathcal{O},\gamma)$ only depend on the power of l dividing $[\mathcal{O}_K : \mathcal{O}]$. Therefore suppose K has discriminant D and $\mathcal{O} = \mathbf{Z}[l^s\sqrt{D}]$ is in O_γ. Also suppose $\bar{\gamma} = 1 + \alpha$ where $\alpha = N(a + bl^r\sqrt{D})$ where $a, b \in \mathbf{Z}$ and $(l,b) = 1$. It follows that $r \geqq 0$ and $t + r \geqq s$. Suppose $x \in \mathcal{O}$ and x has order N modulo $N\mathcal{O}$. Let $x = c + dl^s\sqrt{D}$. Then

$$\alpha x \equiv Nbcl^r\sqrt{D} \bmod N\mathcal{O} .$$

Hence, $\bar{\gamma}x \equiv x \bmod (N/l)\mathcal{O}$ if and only if $cl^r \equiv 0 \bmod l^{s-1}$ (here when $s = 0$ we require no condition on c) and $\bar{\gamma}x \equiv x \bmod N\mathcal{O}$ if and only if $cl^r \equiv 0 \bmod l^s$. Suppose first that $r < s - 1$, then in either case $l|c$ so $(d,l) = 1$. Thus,

$$\begin{aligned} \mathbf{C}(\mathcal{O}) &= l^{t-(s-r)}(N - N/l) - 2l^{t-1-((s-1)-r)}(N/l - N/l^2) \\ &= l^{r-s}N(N - 3N/l + 2N/l^2) > 0 . \end{aligned}$$

Suppose now $r = s - 1$, then c may be arbitrary in the first case and $l|c$ in the second. Thus,

$$\begin{aligned} \mathbf{C}(\mathcal{O}) &= (N/l)(N - N/l) - 2((N/l)^2 - (N/l^2)^2) \\ &= N^2/l - 3(N/l)^2 + 2(N/l^2)^2 > 0 , \end{aligned}$$

because $l > 2$. Suppose finally that $r > s - 1$. Then,

$$\mathbf{C}(\mathcal{O}) = N^2 - 3(N/l)^2 + 2(N/l^2)^2 > 0 .$$

This establishes the claim. □

Remarks I5. (1) *One may deduce that the field generated by the coefficients of $Q_N(k,T)$, for any $k \in \mathbf{Z}$, equals the compositum of all the imaginary quadratic fields in $\mathbf{Q}_p$ in which p splits.* (2) *We expect that the same methods can be used to prove that there exist overconvergent forms on $\Gamma_1(Np^m)$, of weight k and character χ of arbitrarily large slope.* (3) *This proof gives no information on the distribution of the weight k slopes.*

Combining this proposition with Theorem B5.7 we deduce:

Corollary I4.1. *Given an integer j and a positive integer n there exist arbitrarily large rational numbers α such that there are infinitely many integers $k \equiv j \bmod p^n$ and classical weight k eigenforms on $\Gamma_1(N\mathbf{q})$, which are new away from p, and have slope α.*

Remark I6. *To prove the existence of arbitrarily large rational numbers α for which there exist infinitely many weights k such that there are classical forms of weight k and slope α, one could also use Theorem D combined with Gouvêa–Mazur's method of "proliferation by evil twinning." Indeed, if one has a classical eigenform F on $\Gamma_1(N\mathbf{q})$ of weight k and slope β which is either old or of non-trivial character at p, there exists another eigenform F', the "evil twin" of F, of weight k and slope $k-1-\beta$. Using Theorem D, there exists infinitely many weights j for which there is a classical eigenform F_j of weight j and slope β which is either old or has non-trivial character at p. Hence, the evil twin, F_j', of F_j has slope $j-1-\beta$ and applying Theorem D again we deduce the existence of infinitely many weights of classical eigenforms of slope $j-1-\beta$ for each j.*

Appendix II: A 2-adic example

Although, apart from the results of Appendix I, our theorems have been inexplicit, the methods used are strong enough to given explicit results in any given case. Throughout this section, we will be working over $\mathbf{C}_2$.

Theorem II1. *Suppose k is an even integer. Then there does not exist an over-convergent eigenform form on $\Gamma_0(2)$, weight k and slope in the interval $(0,3)$ and if $k \equiv 2 \bmod 4$ there does not exist one of slope 3. However, if k is an integer divisible by 4, then there exists a unique normalized overconvergent eigenform F_k on $\Gamma_0(2)$, weight k and slope 3. Moreover,*

$$F_k(q) \equiv F_{k'}(q) \bmod \frac{(k-k')}{2}\mathbf{Z}_2 .$$

Remarks II2. (i) *We know F_k is classical if $k \geqq 8$, by Theorem C. Mathew Emerton pointed out that F_4 is also. In fact, we must have $F_4(q) = G_4(q) - G_4(q^2) = G_{(4,0)}$.* (ii) *We must also have $F_{12}(q) = \Delta(q) - \beta\Delta(q^2)$ where β is the root of $X^2 + 24X + 2^{11}$ of valuation 8 $(-24 = \tau(2))$ and F_8 is the unique normalized cusp form on $X_0(2)$ of weight 8.* (iii) *As Mazur pointed out, using the facts that $\Delta(z) = \eta(z)^{24}$ and $F_8(z) = (\eta(2z)\eta(z))^8$, one can show that,*

$$F_{12}(q) \equiv F_8(q) \bmod 16$$

and using the congruences discussed in [SwD, Sect. 1], one can show

$$F_{12}(q) \equiv F_4(q) \bmod 32 .$$

This and other computations of Emerton suggest that the above congruence can be improved to be modulo $4(k-k')$ rather than $(k-k')/2$.

To prove this theorem we must establish estimates for the 2-adic sizes of the coefficients of the characteristic series of the U^* operator. The proof of its entirety, by its nature, can be used to give upper bounds for these which ultimately allow us to ignore most of them when we search for information about forms of a small slope. We then can use Koike's formula to determine the exact sizes of the remaining finite number.

We identify $\mathscr{B}^* = B(0,2)$ with $\mathscr{B}^* \times \{0\} \subset \mathscr{W}^*$ and will restrict the function $Q_1(T)$ to the region $\mathscr{B}^* \times \mathbf{A}^1$ (as we remarked after Theorem I1 it is identically 1 on $\mathscr{B}^* \times 1$), where we may regard it is the series $P(s,T)$ of Theorem B3.2, by Eq. (3) of Sect. B3.

Lemma II3. *Let* $P(s,T) = 1 + C_1(s)T + C_2(s)T^2 + \cdots$ *Then, on* $\mathscr{B}^*$, $v(C_1(s)) = 0$ *and* $v(C_2(s))$ *equals* $2 + v(s-2)$ *if* $v(s-2) < 2$ *and is at least* 4 *otherwise.*

Proof. Let

$$\gamma = \frac{-1-\sqrt{-7}}{2} \quad \text{and} \quad \rho = \frac{1+\sqrt{-15}}{2}$$

where the square roots are taken so as to be elements of $1 + 4\mathbf{Z}_2$. Note that $\gamma \equiv \rho \equiv 1 \bmod 4$. We know, from Theorem I3, that

$$C_1(s) = d_1(s) \quad \text{and} \quad C_2(s) = \frac{d_1(s)^2 + d_2(s)}{2}\,.$$

where

$$d_1(s) = \frac{\gamma^{s-2}}{1-2\gamma^{-2}} \quad \text{and} \quad d_2(s) = \frac{\gamma^{2s-4}}{1-4\gamma^{-4}} + \frac{2\rho^{s-2}}{1-4\rho^{-2}}\,.$$

Clearly, $d_1(s)$ has valuation 0 for $v(s) > -1$. We now investigate the next coefficient of $P(s,T)$. It is easy to see that

$$C_2(s) \equiv \tfrac{1}{3}(\gamma^{2s} - 9\rho^s) \bmod 16\mathcal{O}\,.$$

This element of $\mathcal{O}$ has valuation equal to $2 + v(s-2)$ if $v(s-2) < 2$. □

Proof of Theorem. Silverberg suggested considering the family of curves with a point of order 2:

$$(E_c, P_c) := \left(y^2 = x^3 + x^2 + \frac{16c}{1+64c}x,\ (0,0)\right)$$

$c \neq 0$ or $-1/64$ (c may be thought of as a parameter on $X_0(2)$). The curve isogenous to E_c after dividing out by P_c is $E_{w(c)}$, where w is the Atkin–Lehner involution; $w(c) = 1/2^{12}c$. The j-invariants of these curves are

$$j(E_c) = \frac{(1+16c)^3}{c^2} \quad \text{and} \quad j(E_{w(c)}) = \frac{(1+256c)^3}{c}\,.$$

It follows that E_c has potential supersingular reduction if and only if $-12 < v(c) < 0$ so the connected component of the ordinary locus containing 0, of the above model of $X_0(2)$, is the disk $\{c : v(c) \geqq 0\}$.

Let ϕ be the Tate–Deligne morphism near 0 (which is $w\varphi w$ where φ is the Tate–Deligne morphism near ∞) which is defined on a wide open containing $B[0,1]$. Since the point P_c of E_c is not in the kernel of reduction if $c \in B[0,1]$, we have:

$$\frac{\phi(c)}{(1+256\phi(c))^3} = \frac{c^2}{(1+16c)^3}\,. \tag{1}$$

This implies

$$\phi(c) = c^2 G(16c) \tag{2}$$

for some $G(T) \in \mathbf{Z}[[T]]$ such that $G(0) = 1$. (This means ϕ converges on the disk $v(c) > -4$, which implies that the Hasse invariant of the reduction Modulo 2 of a smooth model of E_c has valuation strictly less than 2/3 (and more importantly, the Hasse invariant of the reduction of $E_{w(c)}$ has valuation strictly less than 1/3).)

For $a \in \mathbf{C}_2$, $v(a) < 0$, let V_a be the affinoid disk $\{x \in X_0(2) : v(c(x)) \geqq v(a)\}$. Then, an orthonormal basis for $N_a := A(V_a)$ is $\{(c/a)^n : n \geqq 0\}$. For $v(a) > -4$, ϕ is a finite morphism from V_a to V_{a^2}, so we have a map $T' := \frac{1}{2}\mathrm{Tr}_\phi : N_a \to N_{a^2}$. Now let r denote the restriction map from N_{a^2} to N_a and U' be the operator on $N_{a^2}, T' \circ r$.

Let $I(Y) = Y^2/(1+Y)^3$ and $H(T) = T/(1+256T)^3$. Then we may write

$$Y^2 = A(I(Y))Y + B(I(Y))\,,$$

where $A = TA_0(T)$, $B = TB_0(T)$ with $A_0, B_0 \in \mathbf{Z}[[T]]$ and $B_0(0) = 1$. Let $e = c/a$ and $d = c/a^2$. Using the fact that $H(\phi(c)) = I(16c)/16^2$, we conclude,

$$e^2 = 16a\phi^*(d)K((16a)^2\phi^*(d))e + \phi^*(d)J((16a)^2\phi^*(d))$$

where $K, J \in \mathbf{Z}[[T]]$, $J(0) = 1$. Thus, $T'(1) = 1$, $T'(e) = 8adK((16a)^2 d)$ and, for $i \geqq 2$,

$$T'(e^i) = (16a)dK((16a)^2d)T'(e^{i-1}) + dJ((16a)^2d)T'(e^{i-2})\,.$$

Thus, if

$$U'(d^i) = \sum_{j \geqq 0} c_{ij}(a)d^j\,,$$

$c_{ij} = 0$, if $i > 2j$ or $i = 0$, and $j > 0$ and

$$v(c_{ij}(a)) \geqq \begin{cases} -2jv(a) & \text{if } i = 2j \\ 2j(4+v(a)) - i(4+2v(a)) - 1 & \text{if } i < 2j\,. \end{cases}$$

Let $r_j(a) = \min_i v(c_{ij}(a))$. Then if, $v(a) \geqq -3/2$,

$$r_j(a) \geqq -2jv(a)\,, \tag{3}$$

if $j > 0$ and $r_0(a) = 0$.

The form E_2^* (the weight 2 Eisenstein series on $X_0(2)$ whose q-expansion is $2E_2(q^2) - E_2(q)$) corresponds to a constant multiple of $\omega = \mathrm{d}c/c$. We need to compute $E^{-1}(c) := \frac{1}{2}\phi^*\omega/\omega$. From (2) we deduce, $E(c) = 1 + V(8c)$, for some $V(T) \in T\mathbf{Z}[[T]]$. In any case, $E(c) \equiv 1 \bmod 8c$ for $v(c) > -3$.

Now we investigate the operator $U'' : f(s,c) \mapsto r \circ U'(E^{s/2}(c)f(s,c))$ on the functions $f(s,c)$ on the region determined by the inequalities $v(c) > -2$ and $v(s) > -1 - v(c)$, where r is the appropriate restriction map ($E^{s/2}(c)$ makes sense on this region). Now $P(s,T)$ is also the characteristic series of this operator. Suppose $v(a) < 0$. Then d^n is an "orthonormal basis" for functions on the region (which is now an open subdomain) determined by the inequalities $v(c) \geqq v(a)$ and $v(s) > -1 - v(a)$. Suppose

$$U''(d^i) = \sum_j c_{ij}(a,s)d^j .$$

Writing $E(c)^{s/2} = \sum_{n \geqq 0} h_n(s)c^n$, we see that

$$U''(d^i) = \sum_{n \geqq 0} a^{2n}h_n(s)U'(d^{n+i}) .$$

So

$$c_{ij}(a,s) = \sum_{n \geqq 0} a^{2n}h_n(s)c_{n+i,j}(a) .$$

Now, $|a^{2n}h_n(s)| \leqq 1$ if

$$v(a) > -1 + (1 - v(s))/4 \tag{4}$$

and $v(a) > -5/4$. We see that, under these conditions, if $R_j(a) = \min_i v(c_{ij}(a,s))$, $R_j(a) = r_j(a)$ and so, using the analogue of the estimates in [S, Sect. 5] and (3),

$$v(C_m(s)) \geqq \sum_{j=0}^{m-1} R_j(a) \geqq -v(a)m(m-1) . \tag{5}$$

This implies that on the disk $v(s) > 1$, $v(C_m(s)) > 3(m-1)$ if $m > 2$ (given any s in this disk we may choose an a such that $-1 > v(a) > -5/4$ so that the inequality (4) holds). Since $E(c)^{s/2} = E(c)^{(s-2)/2}(1 + V(8c))$, we may also verify this inequality on the disk $v(s-2) > 1$. This together with Lemma II3, tells us that all, if $v(s-2) > 1$, and all but one, otherwise, of the sides of the Newton polygon $P(s,T)$ with positive slope have slope strictly greater than 3 and moreover, if $v(s) > 1$, the Newton polygon has a side of slope 3 above the interval $[1,2]$. This implies all the assertions of the theorem save the congruence. The congruence follows from Lemma B5.3. □

Remarks II4. (i) *What we have ultimately proven is that there exists a q-expansion*

$$F(s,T) := q + a_2(s)q + a_3(s)q^2 + \cdots$$

where the $a_i(s)$ *are power series which converge and are bounded by one on the disk* $v(s) > 1$ *such that* $F(k,q) = F_k(q)$. *We can show that the* $a_i(s)$ *analytically continue to rigid analytic functions bounded by one on a wide open containing* $\{x : v(x) \geqq 0, v(x-14) < 4\}$. *This implies that the modulus of the congruence in Theorem* II1 *may be improved to* $2(k-k')$. (ii) *Theorem* II1 *implies the result of Hatada* [Ha], *that each eigenvalue of the Hecke operator* T_2 *acting on the space cusp forms of level* 1 *of any weight is divisible by* 8. (iii) *We have used the above techniques together with a Pari program set-up by Teitelbaum to show that the next smallest slope, after* 3, *of an overconvergent modular form of weight* 0 *and tame level* 1 *is* 7 *and the dimension of the space of such forms is* 1.

Index of Notation

We record here the first occurrence of a symbol after the introduction.

References

[A] Adolphson, A.: The U_p operator of Atkin on modular forms of level three, Illinois Journal of Math. **24**, 49 – 60 (1980)

[B] Bergman, G.: A weak nullstellensatz for valuations, Proc. of the AMS **28**, No. 1, 32–38 (71)

[BGR] Bosch, S., Güntzer, U., Remmert, R.: Non-Archimedian Analysis, Springer-Verlag, (1984)

[C-PSI] Coleman, R.: A p-adic Shimura isomorphism and p-adic periods of modular forms, Contemporary Mathematics **165**, 21–51 (1994)

[C-CO] Coleman, R.: Classical and overconvergent modular forms, Invent. math. **124**, 215–241 (1996)

[C-HCO] Coleman, R.: Classical and overconvergent modular forms of higher level (to appear)

[C-CPS] Coleman, R.: On the characteristic power series of the U-operator (to appear)

[CE] Coleman, R., Edixhoven, B.: The semi-simplicity of the U_p-operator on elliptic modular forms (to appear)

[D1] Dwork, B.: On Hecke polynomials, Invent. Math. **12**, 249–256 (1971)

[D2] Dwork, B.: The U_p operator of Atkin on modular forms of level 2 with growth conditions, SLN 350, 57 – 67 (1972)

[F-vP] Fresnel, J., van der Put, M.: Géométrie Analytique Rigide et Applications, Birkhäuser, Boston (1981)

[G-ApM] Gouvêa, F.: Arithmetic of p-adic modular forms, SLN 1304, (1988)

[G] Gouvêa, F.: Continuity Properties of Modular Forms

[GM-F] Gouvêa, F., Mazur, B.: Families of Modular Eigenforms, Mathematics of Computation, **58**, 793–805 (1992)

[GM-CS] Gouvêa, F.: On the characteristic power series of the U-operator, Annales de l'institut Fourier **43** 2, 301–312 (1993)

[Ha] Hatada: Eigenvalues of Hecke Operators on $SL(2,\mathbf{Z})$ Math. Ann. **239**, 75–96 (1979)

[H-IM] Hida, H.: Iwasawa modules attached to congruences of cusp forms, Ann. Sci. Ec. Norm. Sup., **19**, 231–273 (1986)

[H-GR] Hida, H.: Galois representations into $GL_2(\mathbf{Z}_p[[X]])$ attached to ordinary cusp forms, Invent. Math. **85**, 545–613 (1986)

[H-LE] Hida, H.: Elementary theory of L-functions and Eisenstein series, London Mathematical Society (1993)

[H-NO] Hida, H.: Nearly Ordinary Hecke Algebras and Galois Representations of Several Variables, JAMI Innaugural Conference Proceedings, (supplement to) Amer. J. Math. 115–134 (1990)

[Hj] Hijikata: Explicit formula of the traces of Hecke operators for $\Gamma_0(N)$, J. Math. Soc. of Japan, **26**, 56–82 (1974)

[I] Ihara, Y.: Hecke polynomials as congruence ζ functions in the elliptic modular case, Ann. of Math. 267–295 (1967)

[K-pMF] Katz, N.: P-adic properties of modular schemes and modular forms, Modular Functions of one Variable III, SLN 350, 69–190 (1972)

[K-pIE] Katz, N.: P-adic interpolation of real analytic Eisenstein series, Ann. of Math., **104**, 459–571 (1976)

[KM] Katz, N., Mazur, B.: Arithmetic Moduli of Elliptic Curves. Annals of Math. Stud. **108**, PUP (1985)

[Ko1] Koike, M: J. Fac. Sci. Univ. Tokyo 20 (1973), On p-adic properties of the Eichler–Selberg trace formula, Nagoya Math. J. **56**, (1975)

[Ko2] Koike, M.: On p-adic properties of the Eichler–Selberg trace formula II Nagoya Math. J. **64**, 87–96 (1976)

[L-A] Lang, S.: Algebra, Addison-Wesley (1993)

[L-CF] Lang, S.: Cyclotomic Fields, Springer-Verlag (1978)

[Li] Li, W.: Newforms and functional equations, Math. Ann. **212**, 285–315 (1975)

[MW] Mazur, B., Wiles, A.: Analogies between Function fields and number fields, Amer. J. Math. **105**, 507–521 (1983)

[M] Morita, Y.: Hecke polynomials of modular groups and congruence zeta functions of fiber varieties, J. Math. Soc. Japan **21**, 6127–637 (1969)

[Mi] Miyake, T.: Modular Forms, Springer (1989)

[S] Serre, J-P.: Endomorphismes complètements continues des espaces de Banach p-adiques, Publ. Math. I.H.E.S., **12**, 69–85 (1962)

[S-MZp] Serre, J.-P.: Formes modulaires et fonctiones zêta p-adiques, Modular Functions of one Variable III, SLN 350, 69–190 (1972)

[Sh] Shimura, G.: Introduction to the arithmetic theory of automorphic functions **11**, Publ. Math. Soc. Japan, Princ. Univ. Press (1971)

[SwD] Swinnerton–Dyer, H.P.F.: On l-adic representations and congruences of modular forms, Modular Functions of one Variable III, SLN 350, 1–55 (1972)

[T] Tate, J.: Rigid analytic spaces, Invent. Math. **12**, 257–289 (1971)

[W] Washington, L.: Introduction to Cyclotomic Fields, Springer-Verlag (1980)