

Uniform Distribution of Heegner Points

V. Vatsal

Dept. of Mathematics

University of British Columbia

vatsal@math.ubc.ca

August 12, 2003

1 INTRODUCTION

Let E be a (modular!) elliptic curve over \mathbf{Q} , of conductor N . Let K denote an imaginary quadratic field of discriminant D , with $(N, D) = 1$. If p is a prime, then there exists a unique \mathbf{Z}_p -extension K_∞/K such that $\text{Gal}(K/\mathbf{Q})$ acts nontrivially on $\text{Gal}(K_\infty/K)$. The field K_∞ is called the anticyclotomic \mathbf{Z}_p -extension of K . Let $E(K_\infty)$ denote the Mordell-Weil group of E over K_∞ . Then a fundamental conjecture of Mazur [Maz84] predicts that the size of $E(K_\infty)$ is controlled by the prime factorization of N in K . Equivalently, Mazur's conjecture relates the size of the Mordell-Weil group to the sign in the functional equation of certain L-series. The conjecture was verified by Greenberg, Rohrlich, and Rubin, in what Mazur calls the *exceptional* case, when E has complex multiplication by K . More generally, they settled the conjecture for certain abelian varieties with complex multiplication. For a discussion of this CM case, we refer the reader to [Gre85], [Roh84], and [Rub91].

Our goal in this paper is to treat the *generic* case, which occurs either when E has no CM, or when the field of complex multiplications is distinct from K . Under certain conditions on E and K , Mazur's conjecture predicts that the group $E(K_\infty)$ is finitely generated; our main result asserts that this is in fact the case, at least when p is an ordinary prime for E , or when the class number of K is prime to p .

The main new ingredient we introduce is that of equidistribution, following ideas used by Ferrero and Washington to study the cyclotomic μ -invariant. More precisely, we show that the Heegner points associated to definite quaternion algebras are *uniformly distributed* on the components of a certain curve X , and that the elements of a certain Galois group act *independently*, in a suitable sense. This, combined with a special value formula due to Gross, allows us to conclude that the special values of anticyclotomic L-functions are almost always nonzero, so that the statement about the Mordell-Weil groups follows from the machinery of Euler systems as developed by Bertolini and Darmon [BD97].

We caution the reader that the Heegner points considered here are not Heegner points in the classical sense. In particular, they do not give rise to a family of points on the Jacobian of a modular curve. The points we consider are sometimes referred to as Gross points, or special points (see [Gro87], or [BD96]).

To state the results more precisely, we need to introduce some notation. Let g denote a cuspidal newform of weight 2 on the group $\Gamma_0(N)$. Write the Fourier expansion of g as $g(z) = \sum a_n q^n$, so $a_n = a_n(g)$ is the eigenvalue of the Hecke operator T_n . Let χ denote any finite-order Hecke character of K , of conductor \mathfrak{f} . Then, for each place v of K , there exists an Euler factor $L_v(g, \chi, s)$. If v is finite and $v \nmid N \cdot \text{Norm}(\mathfrak{f})$, then we have the well-known formula

$$L_v(g, \chi, s) = \left(1 - a_v \chi_v(\pi_v) N v^{-s} + \chi_v^2(\pi_v) N v^{1-2s}\right)^{-1},$$

where Nv is the norm of v , and $\chi_v(\pi_v)$ denotes the local component of χ at v , evaluated on a v -adic uniformizer π_v . If v has residue characteristic q , and q is unramified in K , then the number a_v is either a_q or $a_q^2 - 2q$, depending on whether v is split or inert. If g has rational Fourier coefficients, corresponding to an elliptic curve, then $a_v - Nv - 1$ is simply the number of rational points of E over the residue field of K at v .

There is a similar but more complicated definition for the Euler factor at the remaining primes of K . At the infinite place of K , we have $L_\infty(g, \chi, s) = (2\pi)^{-2s} \Gamma(s)^2$. Then the L-series of g is defined by $L(g, \chi, s) = \prod_v L_v(g, \chi, s)$. It is well-known that $L(g, \chi, s)$ extends to an analytic function of $s \in \mathbf{C}$, and that it satisfies the functional equation

$$L(g, \chi, s) = e \cdot A^{s-1} \cdot L(g, \chi^{-1}, 2-s), \tag{1}$$

for suitable constants $e = e(\chi)$ and $A = A(\chi)$.

Recall that the character χ is called *anticyclotomic* if its conjugate under the action of $\text{Gal}(K/\mathbf{Q})$ is equal to its inverse. It is well-known that any such χ factors through $\text{Gal}(K_{\mathfrak{f}}/K)$, where $K_{\mathfrak{f}}$ is a suitable ring-class field with conductor \mathfrak{f} . In this case, it is possible to determine the sign $e(\chi)$ in the functional equation explicitly. Let χ_0 denote the trivial character of K , and let e_0 denote the corresponding sign in (1). Then one knows that

- $e_0 = -\chi_D(N)$, where χ_D is the quadratic Dirichlet character of K , and
- If χ is an anticyclotomic Hecke character of conductor prime to N , then $e(\chi) = e_0$.

It follows from the formulae above that if the conductor of χ is prime to N , then $e(\chi) = e_0$ is independent of χ . We will say that we are in the *definite* case if $e_0 = +1$, and that we are in the

indefinite case if not. For example, we are in the indefinite case if all primes dividing N are split in K ; this is the classical Heegner hypothesis. If N is squarefree, then we are in the definite case if and only if the number of prime factors of N that stay inert is odd. Note that it follows trivially that $L(g, \chi, 1) = 0$ for all χ , if we are in the indefinite case.

Now assume that χ is primitive of conductor \mathfrak{f} , with \mathfrak{f} prime to DN , and let $\rho(g, \chi)$ denote the order of vanishing of $L(g, \chi, s)$ at $s = 1$. The essence of Mazur's conjecture is that the number $\rho(g, \chi)$ is 'generically' determined by the sign e_0 . To state this more precisely, we let p denote a prime with $p \nmid ND$. For each integer $n \geq 0$, we let H_n denote the ring class field of K with conductor p^n . We put $H_\infty = \cup H_n$, and consider the behavior of $\rho(g, \chi)$ as χ varies over finite-order characters of $G_\infty = \text{Gal}(H_\infty/K)$. Then we have

CONJECTURE 1.1 (MAZUR) *Suppose that we are in the definite case. Then $L(g, \chi, 1) \neq 0$ for all but finitely many χ of conductor p^n . If we are in the indefinite case, then $L'(g, \chi, 1) \neq 0$ for all but finitely many χ .*

The Birch and Swinnerton-Dyer conjectures allow us to reformulate this in terms of Mordell-Weil groups. Thus, let E denote the abelian variety quotient of $J_0(N)$ associated to g by Shimura. If g has rational Fourier coefficients, then E is an elliptic curve as above. For a primitive character of conductor \mathfrak{f} , we let $E(K_{\mathfrak{f}})^\chi = e_\chi E(K_{\mathfrak{f}}) \subset E(K_{\mathfrak{f}}) \otimes \mathbf{Z}[\chi]$ denote the χ -isotypic part of the Mordell-Weil group of E over $K_{\mathfrak{f}}$, where $e_\chi = \sum \chi(\sigma)\sigma \in \mathbf{Z}[\chi][\text{Gal}(K_{\mathfrak{f}}/K)]$. We will write $r(g, \chi)$ for the dimension of $E(K_{\mathfrak{f}})^\chi \otimes \mathbf{C}$. Then the conjecture of Birch and Swinnerton-Dyer predicts that $\rho(g, \chi) = r(g, \chi)$. Letting χ vary over the anticyclotomic characters of p -power conductor as before, we are led therefore to the second part of Mazur's conjecture:

CONJECTURE 1.2 *Suppose that we are in the definite case. Then $E(K_{\mathfrak{f}})^\chi$ is finite for almost all χ , and $E(H_\infty)$ is finitely generated. If we are in the indefinite case, then $E(K_{\mathfrak{f}})^\chi$ has rank 1, for all but finitely many χ .*

It is now known, thanks to fundamental work of Bertolini and Darmon (see [BD97], Corollary D) that Conjecture 1.1 implies Conjecture 1.2, at least in the definite case. Indeed, we have

THEOREM 1.3 (BERTOLINI-DARMON) *Suppose that p is an odd prime. If $L(g, \chi, 1) \neq 0$, then $E(K_{\mathfrak{f}})^\chi$ is finite. The torsion subgroup of $E(H_n)$ is bounded as $n \rightarrow \infty$.*

Note that the notation here is slightly different from that appearing in Bertolini-Darmon. We are considering automorphic forms g that are unramified at p , while Bertolini and Darmon look at forms that are special at p . A translation between the two is given in the definitions immediately preceding

Corollary D of [BD97], and the discussion immediately following. Our p is one of the auxiliary primes ℓ_i , while p in Bertolini-Darmon denotes a prime where g is special, and which is inert in K .

In the indefinite case, the analogous implication would follow from [BD90], modulo a suitable generalization of the Gross-Zagier formula. (For the latter, see the recent work of Zhang [Zha01].) One should note as well that Mazur's original conjecture about the Mordell-Weil ranks referred only to the case where p is an *ordinary* prime for g . However, given the sign of the functional equations, and the work of Bertolini and Darmon, the present formulation seems to be reasonable.

To state our main result, we will write h_K for the class number of K . We factor

$$N = N^+ \cdot N^-,$$

where $(N^+, N^-) = 1$ and N^- is divisible only by primes that are inert in K .

THEOREM 1.4 *Suppose that g is a newform on $\Gamma_0(N)$, and that $K = \mathbf{Q}(\sqrt{-D})$ is an imaginary quadratic field. If the assumptions above hold for N , D , and p , then,*

- *If p is an ordinary prime for g , then $L(g, \chi, 1) \neq 0$ for all but finitely many χ of p -power conductor, and the Mordell-Weil group $E(H_\infty)$ is finitely generated.*
- *If p is a supersingular prime, then the same conclusion holds provided that the class number of K is prime to p .*
- *More generally, suppose that p is a supersingular prime and the class number h_K is divisible by p . Then, if χ_t is a tamely ramified character of order prime to p , we have $L(g, \chi_t \chi_w, 1) \neq 0$ for all but finitely many characters χ_w of $\text{Gal}(K_\infty/K)$, and the χ -isotypic part of $E(K_\infty(\chi_t))$ is finitely generated.*

As we have already remarked, the statements about the Mordell-Weil group follow from the non-vanishing of the appropriate L-values.

We note here that there are situations where $e(\chi) = 1$ even though our assumptions do not hold. For instance, it may happen that $e(\chi) = 1$ when p is a prime where g is special. For example, if we factor $N = N^+ \cdot N^-$ into a split and inert part as above, then if N^- is squarefree, has an even number of factors, and p is a prime dividing N^- , we will have $e(\chi) = 1$ for χ of conductor p^n , $n \geq 1$. The arguments of this paper may be generalized to this case in a straightforward manner, although the details and explicit formulae are unavoidably different.

Furthermore, we point out that we have made some restrictions in our results when p is a supersingular prime, and the class number of K is divisible by p . These restrictions can be lifted at the cost of some unpleasant technical complications.

We will clear up the above points in a separate work, in order to keep the present exposition simple.

Finally, it may also happen that the sign $e(\chi) = 1$ when N^- is not squarefree. We have chosen to avoid this case because it is not clear to us exactly what sort of Gross formula should hold in this situation. It seems possible that one could use the recent results of Zhang [Zha01] to clarify this case, and we hope to consider this in the future.

Sketch of the proof

The principal ingredient in the proof of the theorem is the study of Heegner points on definite quaternion algebras, or, more precisely, the distribution of these points on the components of a certain curve X . These curves X were introduced and studied by Gross [Gro87]. As we have already remarked, we will show that the Heegner points are *uniformly distributed* on the various components of X , and that the elements of a certain Galois group act independently (in a sense that we will make precise shortly). One deduces the nonvanishing of $L(g, \chi, 1)$ from this by invoking Gross' special value formula, which states that $L(g, \chi, 1)$ is essentially the height of a twisted Heegner point.

To elaborate on these ideas, let B denote the definite quaternion algebra over \mathbf{Q} which is ramified at the primes dividing N^- . This makes sense because we have assumed that N^- is square-free and divisible by an odd number of primes. Then Gross has associated to B a certain curve X . We refer the reader to §2 of the text for the precise definition. Here we merely recall that X is disconnected, being the union of genus zero curves Y_i defined over \mathbf{Q} , with one component for each conjugacy class of oriented Eichler orders of level N^+ .

In this framework, a Heegner point P of conductor p^n is a pair (f, R) , where $R \subset B$ is an Eichler order of level N^+ , and $f : K \rightarrow B$ is an embedding satisfying $f^{-1}(R) = O_n$, where O_n is the order of K with conductor p^n . We agree to identify pairs (f, R) and (f', R') if they are conjugate by an element of B^\times . In practice, one has to enlarge this definition to include orientations on O_n and R , but we will not concern ourselves with this detail here. The salient point is that $P = (f, R)$ determines a conjugacy class $[R]$ of Eichler orders of level N^+ . It can be shown that such a pair (f, R) determines a geometric point on the curve X which lies on the component corresponding to the class of R .

Write X_n for the set of Heegner points of conductor p^n . It turns out that there is a natural action of the Picard group $\text{Pic}(O_n)$ on the set X_n . Once one rigidifies the situation by imposing orientations, it becomes apparent that X_n is a homogenous space for $\text{Pic}(O_n)$, and that the group action is free and transitive. In particular, there are e_n Heegner points of conductor p^n , where e_n is the order of $\text{Pic}(O_n)$. To simplify matters, we shall assume for the purposes of this introduction that K has class number one. Then it is well-known that $e_n = cp^{n-1}$, for some simple constant c . On the other hand,

one knows also that there are only finitely many conjugacy classes of Eichler orders, and so it is natural to ask for the distribution of the elements of X_n in this finite set. If we write $\text{Cl}(B) = \text{Cl}(B, N^+)$ for the set of conjugacy classes of oriented orders of level N^+ , then our first result is a determination of the limiting distribution in $\text{Cl}(B)$ of the sets X_n . To state a precise formula, let us fix representatives R_i for the conjugacy classes of oriented Eichler orders. For each class $[R_i]$, we define a weight w_i by setting w_i equal to the order of R_i^\times . Then we define a mass number h by putting $h = \sum 1/w_i$, where the sum is taken over all the classes $[R_i]$. Let $\mu_i(n)$ denote the number of Heegner points $P = (f, R)$ of conductor p^n such that $R \in [R_i]$, and recall that e_n is the number of points of level p^n . Then our equidistribution result is the following

THEOREM 1.5 *We have*
$$\lim_{n \rightarrow \infty} \frac{\mu_i(n)}{e_n} = \frac{1}{w_i h} = \frac{1/w_i}{\sum 1/w_j}.$$

The proof of this theorem is fairly straightforward. Using strong approximation, one identifies the Heegner points with vertices on a certain tree T (the tree of $PGL_2(\mathbf{Q}_p)$), while $\text{Cl}(B)$ is identified with the vertices of a finite quotient graph \mathcal{G} . The Heegner points of conductor p^n are those which are reached by a certain kind of walk of length n on \mathcal{G} , and the theorem above then boils down to a standard sort of problem in the theory of random walks on graphs.

The independence of the Galois action is much more subtle, however. Since the bare statement may seem somewhat unmotivated, we will start by discussing Gross' special value formula, to show exactly what is needed.

Let $\text{Cl}(B)$ denote the set of conjugacy classes of Eichler orders of level N^+ , as above. As we have remarked, the set $\text{Cl}(B)$ is identified with the vertex set of a certain finite graph \mathcal{G} . Associated to this graph and the modular form g is a certain canonical function $\psi = \psi_g : \mathcal{G} \rightarrow \mathbf{R}$, defined via the Jacquet-Langlands correspondence as follows. Let π_p denote the local representation of $PGL_2(\mathbf{Q}_p)$ associated to g . Then π_p is an unramified principal series representation. The tree T is isomorphic to \tilde{G}/\tilde{K} , where $\tilde{G} = B_p^\times/\mathbf{Q}_p^\times = PGL_2(\mathbf{Q}_p)$ and $\tilde{K} = PGL_2(\mathbf{Z}_p)$ is a maximal compact subgroup. Then we put $\Gamma = R[1/p]^\times$, where R is any Eichler order of level N^+ . The representation π_p occurs in $L^2(\Gamma \backslash \tilde{G})$; the graph \mathcal{G} is identified with $\Gamma \backslash \tilde{G}/\tilde{K}$; and the function ψ is the unique \tilde{K} -fixed vector in $L^2(\Gamma \backslash \tilde{G})$ which generates the representation π_p . In particular, ψ is an eigenvector for the Hecke operator T_p . We may view ψ as a function on the Heegner points $P = (f, R)$ by defining $\psi(P) = \psi([R])$. Then, if χ is a primitive character of conductor p^n , Gross' formula states that

$$\frac{L(g, \chi, 1)}{\Omega_g} = \frac{1}{p^n} \sum_{P \in X_n} \sum_{\sigma \in G_n} \chi(\sigma) \psi(P^\sigma) \psi(P), \quad (2)$$

for a suitable period $\Omega_g \in R$. Here $G_n \cong \text{Pic}(O_n) = \text{Gal}(H_n/K)$, and P^σ denotes the conjugate point of P under the action of $\sigma \in G_n$. We refer the reader to §2 below for alternative formulations of

this identity.

We want to study the behavior of the L-value above for χ of conductor p^n , as $n \rightarrow \infty$. Then, since K has class number 1, the order e_n of G_n satisfies $e_n = cp^{n-1}$, where $c = (p \pm 1)/u$, and u is the order of O_K^\times . Also, $G_n = G_1 \times \Delta_n$, where G_1 has order c , and Δ_n is cyclic of order p^{n-1} . In particular, G_1 has order prime to p . We may then write $\chi = \chi_t \chi_w$, where χ_t is a tamely ramified character of G_1 , and χ_w is a ‘wild’ character of Δ_n . Let the tame character χ_t be fixed. According to well-known algebraicity results of Shimura, we will have $L(g, \chi, 1) \neq 0$ for all $\chi = \chi_t \chi_w$ of conductor p^n , as soon as *one* of the $L(g, \chi, 1) \neq 0$. This is because the algebraic parts of the $L(g, \chi, 1)$ are conjugate under the action of $\text{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, if χ_t is fixed and χ_w varies over wild characters of conductor p^n . Thus, following Rohrlich, we are led to consider the average

$$\frac{1}{p^{n-1}} \sum_{\chi_w} L(g, \chi, 1) / \Omega_g,$$

where we consider the tame part as fixed.

One can then insert the formula (2) and simplify the resulting expression, with the aim of evaluating the limit as $n \rightarrow \infty$. The exact formulae are complicated, owing to the fact that the primitive characters of conductor p^n are not a complete orthonormal set, but the general shape of the problem one encounters may be seen by considering just one of the terms that arises, namely

$$\frac{1}{p^{n-1}} \sum_{\tau \in G_1} \sum_{P \in X_n} \chi_t(\tau) \psi(P) \psi(P^\tau) \quad (3)$$

where the function ψ is as above. Observe here that the first sum is taken over $\tau \in G_1$, rather than over G_n . This comes from the fact that we are averaging over the wild characters χ_w of Δ_n . The point we would like to emphasize is that we need only consider the action of the fixed group G_1 rather than the whole group G_n . This turns out to be extremely useful.

Let us now examine the expression (3) in more detail, considering together the terms for which τ is fixed. There is a distinguished term, corresponding to $\tau = 1$, which looks like

$$\frac{1}{p^{n-1}} \sum_{P \in X_n} \psi(P)^2.$$

In view of Theorem 1.5, this is rather easy to evaluate. One finds quickly that the limit is proportional to the quantity $|\psi|^2$ defined by

$$|\psi|^2 = \frac{1}{h} \sum \psi([R_i])^2 / w_i.$$

In particular, this term is nonzero in the limit (since ψ is nonzero). It remains therefore to calculate the limits for $\tau \neq 1$. In this case, one has to consider

$$\frac{\chi_t(\tau)}{p^{n-1}} \sum_{P \in X_n} \psi(P) \psi(P^\tau), \quad (4)$$

and one would like to find the limit as $n \rightarrow \infty$. One can try to apply the same ideas as were used for the case where $\tau = 1$, but there is an evident problem: if one writes R and $R^\tau = R(P^\tau)$ for the Eichler orders coming from P and P^τ , one does not know, *a priori*, how to determine which pairs of conjugacy classes $[R], [R^\tau]$ of Eichler orders are obtained in this way, still less what the asymptotic frequency is. Note that the group $\text{Pic}(O_n)$ does *not* act on $\text{Cl}(B)$, and that R^τ is not determined in any obvious way by R . One would therefore like to say that *all* pairs C_1, C_2 of classes are obtained in this way, and with some predictable frequency. A different way of putting this is to ask whether there are any relations between the classes of R and R^τ .

It turns out that one can explicitly determine all such relations: they arise exclusively from elements $\tau \in G_0 \subset G_1$, where G_0 is a certain subgroup of G_1 coming from genus theory. We will not discuss the general case here, but to explain this phenomenon in the simplest case, recall that we are assuming in this introduction that $K = \mathbf{Q}(\sqrt{-D})$ has class number 1. This implies that D is a prime. Let $\mathfrak{D} = (\sqrt{-D})$ denote the unique ideal of O_K above D . Then \mathfrak{D} is principal in O_K , and defines the trivial element in the ideal class group $\text{Pic}(O_K)$. However, $\mathfrak{D} \cap O_n$ is *not* principal in the order O_n of conductor p^n , as soon as n becomes sufficiently large. Thus $\tau = \text{Frob}(\mathfrak{D})$ is a nontrivial element of $\text{Gal}(H_\infty/K)$. But $\mathfrak{D}^2 = (D)$ is principal, and generated by a rational integer, so that $\text{Frob}(\mathfrak{D})^2$ is trivial in $\text{Gal}(H_\infty/K)$. It follows that $\tau = \text{Frob}(\mathfrak{D})$ is an element of finite order in $\text{Gal}(H_\infty/K)$, and since G_1 is defined to be the torsion subgroup of $\text{Gal}(H_\infty/K)$, we find that $\tau \in G_1$.

For this particular $\tau = \text{Frob}(\mathfrak{D})$, it turns out that there are indeed relations between the classes of the points P and P^τ . We can compute the relations explicitly; they turn out to be given by the Hecke correspondence T_D . One can therefore compute the limit of (4), and it turns out that this is proportional to

$$\lim_{n \rightarrow \infty} \chi_t(\text{Frob}(\mathfrak{D})) \cdot \frac{a_D}{d+1} \cdot \frac{1}{p^{n-1}} \cdot \sum_{P \in X_n} \psi(P)^2 = \chi_t(\text{Frob}(\mathfrak{D})) \cdot \frac{a_D}{D+1} \cdot |\psi|^2,$$

where a_D is the D -th Fourier coefficient of our newform g , and $D+1$ is the degree of T_D (recall that D is prime). Putting together the contributions from $\tau = 1$ and $\tau = \text{Frob}(\mathfrak{D})$, we get a quantity which is proportional to

$$\left(1 + \chi_t(\text{Frob}(\mathfrak{D})) \cdot \frac{a_D}{D+1} \right) \cdot |\psi|^2,$$

which is nonzero because $\chi_t(\text{Frob}(\mathfrak{D})) = \pm 1$, and $a_D < d+1$ by the Weil bounds.

It remains to consider the remaining $\tau \in G_1$. It is here that we invoke deep theorems of Ratner [Rat95] on closures of unipotent flows on p -adic Lie groups, which imply that no relations exist between the classes of P and P^τ , unless $\tau \in G_0 = \{1, \text{Frob}(\mathfrak{D})\}$. One concludes from Ratner's theorems that if $\tau \notin G_0$, then all possible pairs C_1, C_2 of classes occur with the appropriate frequency, so that the limit in (4) is proportional to $(\sum \psi([R_i])/w_i)^2$. But the latter is zero, as follows easily from the fact that the function ψ is cuspidal. Thus we find that the Galois average is nonzero, for all $n \gg 0$.

We now want to explain in more detail our use of Ratner's theorem. Ratner's result, in its original form, is an extremely general statement about the nature of unipotent flows in Lie groups, which applies in fact to arbitrary products of real and p -adic groups (for possibly different p 's), but we will not discuss these generalities here. The p -adic versions of the results are stated below, as Theorems 4.6 and 4.13, and we refer the reader to the discussion there for details. In this introduction, we will concentrate on the following entirely concrete consequence:

PROPOSITION 1.6 *Let $G = SL_2(\mathbf{Q}_p)/\pm 1$, and let Γ and Γ' denote discrete and cocompact subgroups. Then, if Γ and Γ' are not commensurable, the product $\Gamma \cdot \Gamma' = \{\gamma \cdot \gamma' \mid \gamma \in \Gamma, \gamma' \in \Gamma'\}$ is dense in G .*

I am indebted to Dani and Raghunathan for showing me how to deduce this from Ratner's general theorem. It may be that the statement above admits a more elementary explanation, but we have not succeeded in finding any simple argument.

In any event, this proposition is used as follows. As we have already remarked, the function ψ factors through a certain quotient graph $\mathcal{G} = \Gamma \backslash \mathcal{T}$ of the tree \mathcal{T} . Each Heegner point P is identified with a vertex v on \mathcal{T} , and the class of the corresponding Eichler order R is computed as the image of v in \mathcal{G} . It turns out that if $\tau \in G_1$ and $P^\tau = (f^\tau, R^\tau)$, then the order R^τ is determined as the image of v in the quotient $\mathcal{G}' = \Gamma' \backslash \mathcal{T}$, for some suitable conjugate subgroup $\Gamma' = \Gamma_\tau$. Under the assumption that τ is not in the genus subgroup G_0 , we are able to verify that the groups Γ and Γ' are not commensurable. The proposition above then implies that $\Gamma \Gamma'$ is a dense subset of G , and it is therefore easy to find a vertex $v \in \mathcal{T}$ whose image in $\mathcal{G} \times \mathcal{G}'$ is any desired pair of vertices. We would like to point out here that, in practice, one needs to know not only that one can achieve any pair of vertices, but also that any such pair occurs with some predictable frequency. This kind of equidistribution is also a consequence of Ratner's theorems.

Proposition 1.6 above may be viewed as a non-abelian and p -adic version of the classical fact that multiples of an irrational number are everywhere dense on the torus. Indeed, the real number α is irrational if and only if the lattices \mathbf{Z} and $\alpha\mathbf{Z}$ are not commensurable in \mathbf{R} . Equivalently, if \mathbf{Z} and $\alpha\mathbf{Z}$ are not commensurable, then the sum $\mathbf{Z} + \alpha\mathbf{Z}$ is dense in the reals.

The reader who is familiar with the work of Ferrero and Washington on the cyclotomic μ -invariant

will doubtlessly have noticed parallels with the arguments sketched above. Namely, Ferrero and Washington are concerned with the distribution of the p -adic digits of the $(p - 1)$ -st roots of unity, and a key role is played in this by the fact that roots of unity, being irrational, generate uniformly distributed subgroups of the circle. Note also that the $(p - 1)$ -st roots of unity correspond, in the cyclotomic theory, to the tame Galois group G_1 . In retrospect this analogy is not surprising – in fact, it was the effort to generalize the methods of Ferrero and Washington to the anticyclotomic situation that was the motivation for the present work.

There are some evident extensions of the ideas introduced in this paper which we would like to mention. First, one might ask whether the analogy with Ferrero-Washington can be extended to yield information about the μ -invariant and special values modulo ℓ . This would then return information about the Tate-Shafarevich group of the abelian varieties. This question has been resolved, and will be presented in a forthcoming work [Vat].

Secondly, it is natural to investigate whether any of our methods can be extended to the indefinite situation, where the sign in the functional equation is -1 . While algebra and combinatorics seem to be inadequate tools for studying indefinite quaternion algebras, we may still obtain information about classical Heegner points and derivatives of L-functions by using the ‘sign-change’ phenomenon that was discovered and exploited by Bertolini and Darmon. Namely, one can use congruences modulo ℓ (the Jochnowitz congruences of [BD99]) to transfer information from the definite case to the indefinite one. Again, we refer the reader to [Vat] for a detailed discussion.

In conclusion, we would also like to make some comments about the special value formula that we use. The only published reference for this is the original paper of Gross [Gro87], where the case of an unramified character of a field with prime discriminant was treated. This formula was subsequently generalized to a certain extent in the thesis of Daghigh [Dag96], but that work has not yet appeared in print. The formula itself has been extensively quoted, however, and in this paper we will simply reproduce the version given by Bertolini and Darmon in the important paper [BD97]. However, this gap in the literature seems now to have been amply filled: very general Gross and Gross-Zagier formulae, valid for automorphic forms over totally real fields, have recently been announced by Zhang [Zha01].

It is a pleasure to thank the various people without whom this work would have been rendered much more difficult: W. Casselman, H. Darmon, and R. Greenberg. I am particularly indebted to S. Dani and M. S. Raghunathan for indicating the relevance of Ratner’s work, and to J. Friedman, who patiently explained the basic facts about graphs. Part of this research was conducted at the Tata Institute of Fundamental Research in December 1999, and I would like to thank the faculty and staff of the Institute for their hospitality and support.

Particular thanks are due to C. Cornut, for his careful examination of a preliminary version of the

manuscript, and for pointing out some errors in an early version of this work. I also owe to Cornut the idea of using degeneracy maps and Hecke operators to control the action of the genus subgroup. (See [Coro1] for a discussion in a somewhat different context.)

I am grateful to an anonymous referee for a number of improvements to the exposition of this work.

Finally, I would like to thank the organizers (G. Harder and N. Schappacher) and participants (especially Nimish Shah) of a workshop on ergodic theory and Heegner points, held at Oberwolfach in April 2001.

2 HEEGNER POINTS ON DEFINITE QUATERNION ALGEBRAS

In this section we briefly recall the definition and basic properties of the special points introduced by Gross. For details, we refer the reader to Gross' original article [Gro87] and to the papers of Bertolini-Darmon, especially [BD96], §2.

Anticyclotomic fields

Let K denote an imaginary quadratic field, and p an odd prime that is unramified in K . As in the introduction, we will write H_n for the ring class field of conductor p^n , and H_∞ for the compositum of all the H_n . Then we have

$$G_\infty = \text{Gal}(H_\infty/K) = G_1 \times \Delta_\infty$$

where $\Delta_\infty \cong \mathbf{Z}_p$ and G_1 is finite. If the class number of K is prime to p then we have $G_1 \cong \text{Gal}(H_1/K)$. But if the class number of K is divisible by p , then G_1 is harder to describe. In any case, let F denote the fixed field of Δ_∞ , and let F_n denote the fixed field of the unique subgroup of Δ_∞ of index p^{n-1} . We will write $G_n = \text{Gal}(F_n/K)$, so that $G_n = G_1 \times \Delta_n$, where Δ_n is cyclic of order p^{n-1} . The field F/K is a tamely ramified extension, and we will refer to G_1 as the *tame* subgroup. Note that the groups $\text{Pic}(O_n)$ and G_n will *not* in general be isomorphic, because the \mathbf{Z}_p -extension may have some unramified part. In particular, faithful characters of G_n may not be primitive of conductor p^n . What is true in general is that there exists a non-negative integer δ such that $G_{n+\delta} = \text{Pic}(O_n)$ for all n sufficiently large. This fact will cause some minor notational complications in the sequel, and the reader may wish simply to assume that the class number of K is prime to p , to avoid unrewarding distraction. In this paper, we will deal with characters of G_n , rather than primitive characters of conductor p^n .

Quaternion algebras

2.1 Let B denote the definite quaternion algebra of discriminant N^- . Let $R \subset B$ denote an Eichler order of level N^+ . Then an orientation on the order R is a choice of homomorphisms $\sigma_\ell : R \otimes \mathbf{F}_\ell \rightarrow k_\ell$ for each prime ℓ dividing N , where $k_\ell = \mathbf{F}_\ell$ if $\ell | N^+$ and $k_\ell = \mathbf{F}_{\ell^2}$ if $\ell | N^-$. (See [BD96], Sec. 1.1). The data of R together with the homomorphisms σ_ℓ is called an *oriented Eichler order*. If $b \in B^\times$, then $R' = bRb^{-1}$ is an Eichler order of level N^+ , and we equip R' with the orientation induced from R . Then a conjugacy class of oriented Eichler orders in B is an ordinary conjugacy class of Eichler orders, equipped with compatible orientations in this manner. Each ordinary conjugacy class gives rise to several oriented classes, according to the various possible orientations. Thus let $[R]$ denote the conjugacy class of oriented Eichler orders of level N^+ in B represented by R , and write $\text{Cl}(B)$ for the set of all such classes $[R]$. We will write h_B for the cardinality of $\text{Cl}(B)$. We will sometimes call the integer h_B the *class number* of B . It is equal to the number of left ideal classes of R in B . Note however that h_B and $\text{Cl}(B)$ both depend on N^+ .

It will be useful in the sequel to have an adelic description of the set $\text{Cl}(B)$. To this end, let $\hat{\mathbf{Z}}$ denote the profinite completion of \mathbf{Z} , and, for an algebra A , write \hat{A} for the adelization $A \otimes \hat{\mathbf{Z}}$. Then there is an isomorphism of sets

$$\text{Cl}(B) = B^\times \backslash \hat{B}^\times / \hat{R}^\times. \quad (5)$$

It follows from strong approximation in \hat{B} that the set $\text{Cl}(B)$ is finite (see [Vig80], pp 87–89).

The curves associated to definite quaternion algebras

In this section we follow [Gro87] and [BD96].

2.2 Let \mathbf{P} denote the conic curve defined over \mathbf{Q} by

$$\mathbf{P}(A) = \{x \in B \otimes A \mid x \neq 0, \text{Norm}(x) = \text{Tr}(x) = 0\} / A^\times,$$

for all \mathbf{Q} -algebras A . Then B^\times acts on \mathbf{P} by conjugation, and we have $\text{Aut}(\mathbf{P}) = B^\times$ as algebraic groups over \mathbf{Q} . The curve X associated to B is then defined by

$$X = B^\times \backslash (\mathbf{P} \times \hat{B} / \hat{R}).$$

Let g_1, g_2, \dots, g_r denote representatives of the double coset space $\text{Cl}(B) = B^\times \backslash \hat{B}^\times / \hat{R}^\times$. Then the groups $\Gamma_i = g_i \hat{R} g_i^{-1} \cap B^\times$ are finite subgroups of B^\times . The curves $Y_i = \Gamma_i \backslash \mathbf{P}$ are curves of genus zero, defined over \mathbf{Q} , and the curve X is given by

$$X = \cup_{i=1}^r Y_i.$$

The set $\mathbf{P}(K)$ of K -rational points on \mathbf{P} is identified with $\text{Hom}(K, B)$. Thus a K -rational point of X consists of a pair $(f, [R])$ where $f \in \text{Hom}(K, B)$, and $[R]$ is a class in $\text{Cl}(B)$. Finally, let M denote the Picard group $\text{Pic}(X)$. Since each component of X has genus zero, it is clear that M is the free \mathbf{Z} -module with basis e_1, e_2, \dots, e_r indexed by the set $\text{Cl}(B)$. Let w_i denote the order of the group Γ_i , for each i . Then we define a positive-definite inner product on M by specifying that

$$\langle e_i, e_j \rangle = w_i \delta_{ij}.$$

2.3 The module M defined above is equipped with an action of Hecke operators, as we now describe. Since $\mathbf{Q}^\times \subset B^\times$ acts trivially on \mathbf{P} , and since $\hat{\mathbf{Q}} = \mathbf{Q} \cdot \hat{\mathbf{Z}}$, because \mathbf{Q} has class number 1, the curve X can be written as

$$X = B^\times \backslash (\mathbf{P} \times \hat{\mathbf{Q}}^\times \backslash \hat{B}^\times / \hat{R}^\times).$$

Note that $\hat{\mathbf{Q}}^\times \backslash \hat{B}^\times / \hat{R}^\times$ breaks up as a product of local spaces $\mathbf{Q}_\ell^\times \backslash B_\ell^\times / R_\ell^\times$, where $R_\ell = R \otimes \mathbf{Z}_\ell$, for each prime ℓ . When ℓ does not divide N , one finds that

$$T_\ell = \mathbf{Q}_\ell^\times \backslash B_\ell^\times / R_\ell^\times = PGL_2(\mathbf{Q}_\ell) / PGL_2(\mathbf{Z}_\ell) \quad (6)$$

is an $\ell + 1$ -regular tree whose vertices correspond to the homothety classes of lattices in \mathbf{Q}_ℓ^2 (the Bruhat-Tits tree of $PGL_2(\mathbf{Q}_\ell)$). Equivalently, the vertices of T_ℓ correspond to the distinct maximal compact subgroups of $PGL_2(\mathbf{Q}_\ell)$. Two vertices v_1, v_2 of T_ℓ are connected by an edge if the corresponding homothety classes contain lattices L_1 and L_2 such that $L_1 \subset L_2$, and L_2/L_1 is cyclic of order ℓ . Given a vertex $x \in PGL_2(\mathbf{Q}_\ell) / PGL_2(\mathbf{Z}_\ell)$, the $\ell + 1$ neighbors may be written as $\eta_1 \cdot x, \dots, \eta_{\ell-1} \cdot x, \eta_\ell \cdot x$, where the η_i and η_ℓ are given by the formulae

$$\eta_i = \begin{pmatrix} 1 & i \\ 0 & \ell \end{pmatrix}, \quad 1 \leq i \leq \ell - 1, \quad \text{and} \quad \eta_\ell = \begin{pmatrix} \ell & 0 \\ 0 & 1 \end{pmatrix}. \quad (7)$$

Then the correspondence T_ℓ is defined by

$$T_\ell(x) = \sum_{i=0}^{\ell} \eta_i \cdot x.$$

Another way of describing the Hecke operator is to consider the double coset $U\eta_\ell U$, where $U = R_\ell^\times \cong GL_2(\mathbf{Z}_\ell)$; then the matrices η_i above satisfy $U\eta_\ell U = \cup \eta_i U$.

The correspondence extends to the product tree $\hat{\mathbf{Q}}^\times \backslash \hat{B}^\times / \hat{R}^\times$, and hence to the curve X . One obtains therefore an associated endomorphism, also denoted T_ℓ , of the module $M = \text{Pic}(X)$. When ℓ divides $N = N^+ \cdot N^-$, one can define an involution W_ℓ^\pm whose definition depends on whether ℓ

divides N^+ or N^- . For the exact definition, we refer the reader to [BD96], §1.5. In any case, we let \mathbf{T} denote the subalgebra (Hecke algebra) of $\text{End}_{\mathbf{Z}}(M)$ generated by the operators T_ℓ (for $\ell \nmid N$) and W_q^\pm (for $q|N$).

2.4 Now if g is any newform on $\Gamma_0(N)$, the Jacquet-Langlands correspondence gives rise to a homomorphism $\psi = \psi_g : \mathbf{T} \rightarrow \mathbf{R}$ such that $\psi(T_\ell) = a_\ell = a_\ell(g)$ for $(\ell, N) = 1$, and $\psi(W_q^\pm) = a_q$. According to a well-known theorem of Shimura, the numbers a_ℓ are all *real*. Let

$$M_g = M \otimes_{\mathbf{T}} \mathbf{R},$$

where the tensor product is taken with respect to the map $\psi_g : \mathbf{T} \rightarrow \mathbf{R}$ introduced above. One knows that M_g is a \mathbf{R} -vector space of dimension one. Fix an identification $M_g \cong \mathbf{R}$, or, equivalently, a generator $v \in M_g$. Then we may view ψ as an \mathbf{R} -valued function (also denoted by ψ) on M as follows. For each $m \in M$, we define $\psi(m)$ to be the image of $m \otimes 1$ under the composite $m \otimes 1 \in M \otimes_{\mathbf{T}} \mathbf{R} = M_g \cong \mathbf{R}$, where the latter isomorphism is induced by our fixed choice of a generator of M_g . Since M is the free module on the set $\text{Pic}(X) = \text{Cl}(B)$, we may apply ψ to the basis elements $[R] \in M$ to obtain a function $\psi : \text{Cl}(B) \rightarrow \mathbf{R}$. Choice of a different generator v of M_g in the definition has the effect of scaling ψ by a nonzero constant. We will regard the identification $M_g \cong \mathbf{R}$ and the function ψ as fixed once and for all.

For a given Eichler order R_i , we let w_i denote the order of R_i^\times . Let e_i denote the corresponding component in $\text{Pic}(X)$, and let e_0 denote the vector

$$e_0 = \sum_{[R_i] \in \text{Cl}(B)} \frac{1}{w_i} e_i \in \text{Pic}(X) \otimes \mathbf{Q}.$$

Then the lemma below states that the function ψ is cuspidal, while the vector e_0 is Eisenstein.

LEMMA 2.5 *With notations as above, we have*

- $\sum_{[R_i] \in \text{Cl}(B)} \psi([R_i])/w_i = 0$, and
- *the vector e_0 is an eigenvector for the Hecke operators. The eigenvalue of T_ℓ is $\ell + 1$, for any $\ell \nmid N$.*

We will give the proof of this lemma in §3.11.

Heegner Points

2.6 Let O denote any order of K . An orientation on O is a choice of homomorphisms $O_\ell \rightarrow k_\ell$ for each prime $\ell|N$, where k_ℓ is as in section 2.1. Fixing an orientation on the maximal order O_K induces orientations on the orders O_n of conductor p^n , for every $n \geq 1$. Then a Heegner point P of

conductor p^n consists of a pair $P = (f, R)$, where $f : K \rightarrow R$ is an oriented embedding such that $f(K) \cap R = f(O_n)$, where O_n is the order of K with conductor p^n , taken up to conjugation in B^\times . (An embedding $f(O) \subset R$ is said to be oriented if the orientations on R and O are compatible in the obvious manner.) We will write X_n for the set of Heegner points of level p^n . It is clear from the definition of the curve X that a Heegner point is a K -rational point on X (but we will not need this).

Given a point $P = (f, R) \in X_n$, the embedding f induces a map $\hat{K} \rightarrow \hat{B}$, whence an action of $\text{Pic}(O_n) \cong K^\times \backslash \hat{K}^\times / \hat{O}_n^\times$ on the set X_n . Explicitly, given an idele $\sigma \in K^\times \backslash \hat{K}^\times / \hat{O}_n^\times$, we define the point $P^\sigma = (f, R^\sigma)$ by letting

$$R^\sigma = B \cap f(\sigma) \hat{R} f(\sigma)^{-1}. \quad (8)$$

It is easily seen that P^σ is a well-defined Heegner point of conductor p^n . Furthermore, the group $\text{Pic}(O_n)$ acts *simply* and *transitively* on X_n (see [Gro87], page 133, and [BD96], Lemma 2.5). Thus there are e_n Heegner points of level p^n , where e_n is the order of the group $\text{Pic}(O_n)$. Note, however, that since the Heegner points are all rational over K , the action of $\text{Pic}(O_n)$ so defined is *not* the usual Galois action on geometric points.

2.7 More generally, if we are given a Heegner point $P = (f, R)$ of conductor p^n and an idele $\sigma \in \hat{B}$, we can form $R' = B \cap \sigma \hat{R} \sigma^{-1}$ as above. Then the pair (f, R') gives a Heegner point of some level and orientation. Applying this with suitably chosen σ as above leads to a description of the Hecke correspondence T_p and of its relationship with the Galois action. Thus, let σ denote one of the $p + 1$ elements

$$\eta_i = \begin{pmatrix} 1 & i \\ 0 & p \end{pmatrix}, \quad 1 \leq i \leq p - 1, \quad \text{and} \quad \eta_p = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \quad (9)$$

as before, where we view each matrix η_i, η_∞ as an element of \hat{B} with component one away from p and with p -component specified as above. Here $B_p = B \otimes \mathbf{Q}_p \cong M_2(\mathbf{Q}_p)$, and $R_p = R \otimes \mathbf{Z}_p$ is conjugate to the order $M_2(\mathbf{Z}_p)$. Let P_1, P_2, \dots, P_{p+1} denote the $p + 1$ Heegner points obtained from the prescription above. Then $T_p(P)$ is given by the formal sum

$$T_p(P) = P_1 + P_2 + \dots + P_{p+1}.$$

If P has conductor p^n with $n \geq 1$, then it can be shown that precisely one of the points P_1, \dots, P_{p+1} has conductor p^{n-1} , while the p others have conductor p^{n+1} . We will refer to the points P_1, \dots, P_{p+1} as the *neighbors* of P . The unique point \check{P} of conductor p^{n-1} is called the *predecessor* of P .

As for the remaining p points of conductor p^{n+1} , it can be shown that they are all conjugate under the action of the group $S_{n+1} = \ker(G_{n+1} \rightarrow G_n)$, so that we have the formal relationship

$$\sum_{\sigma \in S_n} P_*^\sigma = T_p(P) - \check{P}, \quad (10)$$

where P_* is any of the neighbors of P with conductor p^{n+1} . For proofs of all these assertions in a general context, we refer the reader to §2 of [BD96].

Each Heegner point $P = (f, R)$ will determine an element $[R] \in \text{Cl}(B)$. As we have remarked in the introduction, our main task in this paper is to study the distribution of the classes $[R]$ in $\text{Cl}(B)$ as P varies over the Heegner points of level p^n . To this end, we introduce the following convention: given a function ψ with domain $\text{Cl}(B)$, we will view ψ as a function on Heegner points by defining $\psi(P) = \psi([R])$, where $[R]$ is determined by P as above.

Heights and special values

2.8 Let χ be an anticyclotomic character of K , of conductor p^n . Then Gross has given a formula, in terms of the height pairing on the Heegner points, for the special value $L(g, \chi, 1)$ which we now proceed to recall.

Class field theory allows us to view χ as a function on integral ideals $\mathfrak{a} \subset O_K$. Let

$$\Theta_\chi = \sum_{\mathfrak{a}} \chi(\mathfrak{a}) q^{N\mathfrak{a}} = \sum a_n(\chi) q^n$$

denote the theta series associated to χ . Then Θ_χ is a modular form on $\Gamma_0(C_\chi)$, where $C_\chi = Dp^{2n}$. One has the fundamental identity

$$L(g, \chi, s) = \zeta(2s - 2) \sum_{(n, NC_\chi)=1} a_n(g) a_n(\chi) n^{-s} = D(g, \Theta_\chi, s),$$

where $g = \sum a_n(g) q^n$ is a newform on $\Gamma_0(N)$, and $D(g, \Theta_\chi, s)$ is the normalized Rankin-Selberg convolution of g and Θ .

With these notations, the Rankin-Selberg method and the Eichler trace formula lead to the main identity introduced by Gross (and subsequently generalized by Daghigh and then Zhang see [Gro87], [Dag96], and [Zha01]) which states that, if P is any Heegner point of conductor $p^n = \text{Conductor}(\chi)$, then

$$\left| \sum_{\sigma \in \text{Pic}(O_n)} \chi(\sigma) \psi(P^\sigma) \right|^2 = \frac{D(g, \Theta_\chi, 1)}{(g, g)} \cdot \frac{C_\chi^{1/2}}{C_g} = \frac{L(g, \chi, 1)}{(g, g)} \cdot \frac{C_\chi^{1/2}}{C_g} \in \overline{\mathbf{Q}}, \quad (11)$$

where C_χ is the level of Θ_χ , the real number (g, g) is the Petersson inner product, and ψ is the function defined above. The quantity C_g is an explicit fudge factor, depending on the identification $M_g \cong \mathbf{R}$ fixed in the definition of ψ . For our purposes, it will be enough to know that C_g is non-zero, and independent of χ .

The identity above implies that the left-hand-side of (11) is independent of the chosen point P . This may also be seen directly: if Q is another Heegner point of conductor p^n , then $Q = P^\tau$ for some

$\tau \in \text{Pic}(O_n)$, so that

$$\sum_{\sigma \in \text{Pic}(O_n)} \chi(\sigma) \psi(Q^\sigma) = \sum_{\sigma \in \text{Pic}(O_n)} \chi(\sigma) \psi(P^{\tau\sigma}) = \chi^{-1}(\tau) \sum_{\sigma \in \text{Pic}(O_n)} \chi(\sigma) \psi(P^\sigma),$$

and $\chi^{-1}(\tau)$ has absolute value 1.

Observe also that if χ is primitive of conductor p^n , then $C_\chi = Dp^{2n}$ and $\sqrt{C_\chi} = \sqrt{D}p^n$. Then let $\Omega = C_g(g, g)/D^{1/2}$, and define

$$L^{\text{al}}(g, \chi, 1) = \frac{L(g, \chi, 1)}{\Omega}. \quad (12)$$

Recall our notations from §2, especially the definition of the groups G_n and the relationship with $\text{Pic}(O_n)$. It will be useful to change notation here, and write X_n for the set $X_{n-\delta}$, so that X_n becomes a homogeneous space for G_n rather than $\text{Pic}(O_n)$. The elements of X_n are therefore Heegner points of level $p^{n-\delta}$. We use this convention in the following simple reformulation of (11), which will be useful in the sequel.

LEMMA 2.9 *Let χ denote a primitive character of G_n . Then we have*

$$L^{\text{al}}(g, \chi, 1) = \frac{1}{p^{n-\delta}} \sum_{P \in X_n} \sum_{\sigma \in G_n} \chi(\sigma) \psi(P^\sigma) \psi(P). \quad (13)$$

Proof. This follows easily from expanding out the left-hand side of (11) and using the transitivity of the Galois action on X_n . \square

2.10 Recall further that $G_n = G_1 \times \Delta_n$, where Δ_n is cyclic of order p^{n-1} . Given a character χ of G_n , we may decompose $\chi = \chi_t \cdot \chi_w$, where χ_t is a ‘tame’ character of G_1 , and χ_w is a character of Δ_n . (This notation is taken from Luo and Ramakrishnan, [LR97].) The character χ is faithful on Δ_n if and only if the ‘wild’ part χ_w has order p^{n-1} . Equivalently, χ is faithful if and only if χ_w is nontrivial on the subgroup $S_n \subset \Delta_n$ consisting of elements of order p . It is clear that there are $\phi(p^{n-1}) = (p-1)p^{n-2}$ such characters χ_w .

We are interested in studying the values $L(g, \chi, 1)$ as $\chi = \chi_t \chi_w$ varies over the characters of G_n . It will be convenient in the sequel to group together all characters with a fixed tame component χ_t . From now on, we consider a *fixed* character χ_t of G_1 , and, following the example of Rohrlich [Roh84], we introduce the average

$$L_n^{\text{av}}(f, \chi_t, 1) = \frac{1}{p^{n-2-\delta}} \sum_{\chi_w \in Y_n} L^{\text{al}}(g, \chi, 1), \quad (14)$$

where the sum is taken over the set Y_n of faithful characters χ_w of the group Δ_n . Thus χ is a primitive character of conductor $p^{n-\delta}$. If χ_t has order prime to p , then the fields $\mathbf{Q}(\chi_t)$ and $\mathbf{Q}(\chi_w)$ are

linearly disjoint, and it is clear that all the characters $\chi_t \chi_w$ with fixed χ_t are conjugate under the action of $\text{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. In view of this Galois conjugacy, it follows from well-known algebraicity results of Shimura that if $L(g, \chi, 1) \neq 0$ for some $\chi = \chi_t \chi'_w$, with $\chi'_w \in Y_n$, then in fact $L(g, \chi_t \chi_w, 1) \neq 0$ for all $\chi_w \in Y_n$.

We may now state our main result, which directly implies the theorem of the introduction.

THEOREM 2.11 *Let the hypotheses on N , D , and p be as in Theorem 1.4. Let χ_t be any character of G_1 . Then:*

- $\lim_{n \rightarrow \infty} L_n^{\text{av}}(g, \chi_t, 1)$ exists and is nonzero.
- For all $n \gg 0$, we there exists a character $\chi_w \in Y_n$ such that $L(g, \chi_t \chi_w, 1) \neq 0$.
- If the order of χ_t is prime to p , or if p is an ordinary prime for f , then $L(g, \chi_t \chi_w, 1) \neq 0$ for all χ_w in Y_n , as soon as n is sufficiently large.

It is clear that the first statement implies the second. In view of the preceding remarks, the second implies the third if χ_t has order prime to p . To complete the proof of Theorem 1.4, it suffices therefore to show how to eliminate the hypothesis on the order of χ_t if p is an ordinary prime. But this is easily achieved with the help of p -adic L-functions. As is well-known (see [BD96], for example) there exists a p -adic L-function in the ordinary case that interpolates the values $L(g, \chi_t \chi_w, 1)$, for fixed χ_t . Furthermore, the p -adic L-function is represented by an Iwasawa function in $O[\chi_t][[\Delta_\infty]]$. But a nonzero Iwasawa function has only finitely many zeroes, so that the third statement follows from the second.

It suffices therefore to verify the first assertion above. The proof will occupy the remainder of this paper.

Our first task is to transform the statement about L_n^{av} into one about Heegner points, by utilizing Lemma 2.9. We begin with a simple observation. Recall that the subgroup $S_n \subset \Delta_n$ was defined to be the subgroup of elements of order p . Write σ_0 for the identity element of Δ_n .

LEMMA 2.12 *Suppose that $n \geq 3$. Let s_n denote the function on Δ_n defined by $s_n(\sigma) = \sum_{\chi_w \in Y_n} \chi_w(\sigma)$. Then the following identity holds in the group algebra $\overline{\mathbf{Q}}[\Delta_n]$:*

$$\sum_{\sigma \in \Delta_n} s_n(\sigma) \cdot \sigma = p^{n-1} \cdot \sigma_0 - p^{n-2} \sum_{\sigma \in S_n} \sigma \tag{15}$$

Proof. Let ζ_r denote a primitive p^r -th root of unity, for each $r \geq 1$. Consider an element $\sigma \in \Delta_n$ of order p^t . Then it is clear that $s_n(\sigma) = \sum \chi_w(\sigma) = \text{Tr}(\zeta_t)$, if the trace is taken from $\mathbf{Q}(\zeta_{n-1})$ to \mathbf{Q} .

But it is well-known that the trace of ζ_t from $\mathbf{Q}(\zeta_t)$ to \mathbf{Q} is zero, if $t \geq 2$. Transitivity of the trace now implies that $s_n(\sigma) = 0$ if σ has order p^t , with $t \geq 2$.

It remains therefore to calculate $s_n(\sigma_0)$ and $s_n(\sigma)$, where $\sigma \neq \sigma_0$ lies in S_n . Obviously, $s_n(\sigma_0) = (p-1)p^{n-2} = p^{n-1} - p^{n-2}$, since there are $(p-1)p^{n-2}$ characters χ_w . On the other hand, one knows that if $\zeta = \zeta_1$ is a primitive p -th root of unity, then $\text{Tr}(\zeta) = -1$, if the trace is taken from $\mathbf{Q}(\zeta)$ to \mathbf{Q} . Since $\mathbf{Q}(\zeta_{n-1})$ has degree p^{n-2} over $\mathbf{Q}(\zeta)$, we find that $s_n(\sigma) = -p^{n-2}$, if σ has order p , and $\sigma \neq \sigma_0$. The lemma evidently follows. \square

2.13 Now, combining Lemma 2.9 and equation (14) we find that

$$p^{2n-2-2\delta} L_n^{\text{av}}(g, \chi_t, 1) = \sum_{\chi_w \in Y_n} \sum_{P \in X_n} \sum_{\sigma \in G_n} \chi(\sigma) \psi(P^\sigma) \psi(P).$$

Applying Lemma 2.12 to the sum on χ_w , this becomes

$$p^{2n-2-2\delta} L_n^{\text{av}}(g, \chi_t, 1) = p^{n-1} \sum_{P \in X_n} \sum_{\tau \in G_1} \chi_t(\tau) \psi(P^\tau) \psi(P) - p^{n-2} \sum_{P \in X_n} \sum_{\tau \in G_1} \sum_{\sigma \in S_n} \chi_t(\tau) \psi(P^{\sigma\tau}) \psi(P), \quad (16)$$

and

$$L_n^{\text{av}}(g, \chi_t, 1) = \frac{1}{p^{n-1-2\delta}} \sum_{P \in X_n} \sum_{\tau \in G_1} \chi_t(\tau) \psi(P^\tau) \psi(P) - \frac{1}{p^{n-2\delta}} \sum_{P \in X_n} \sum_{\tau \in G_1} \sum_{\sigma \in S_n} \chi_t(\tau) \psi(P^{\sigma\tau}) \psi(P). \quad (17)$$

We can simplify the second term in the expression above. Indeed, the formulae for the Hecke operators given in §2.7 imply that $\sum_{\sigma \in S_n} \psi(P^{\sigma\tau}) = a_p \psi(Q^\tau) - \psi(\check{Q}^\tau)$, where $Q \in X_{n-1}$ is the predecessor of P , and \check{Q} is the predecessor of Q . Thus we get

$$\begin{aligned} \sum_{P \in X_n} \sum_{\tau \in G_1} \sum_{\sigma \in S_n} \chi_t(\tau) \psi(P^{\sigma\tau}) \psi(P) &= \sum_{P \in X_n} \sum_{\tau \in G_1} \chi_t(\tau) \left(a_p \psi(Q^\tau) - \psi(\check{Q}^\tau) \right) \psi(P) \\ &= \sum_{Q \in X_{n-1}} \sum_{\tau \in G_1} \chi_t(\tau) \left(a_p \psi(Q^\tau) - \psi(\check{Q}^\tau) \right) \left(a_p \psi(Q) - \psi(\check{Q}) \right), \end{aligned} \quad (18)$$

where the second equality comes from the facts that, for $\sigma \in S_n$, all points P^σ give rise to the *same* points Q and \check{Q} , and that this association gives a bijection between distinct points $Q \in X_{n-1}$ and the S_n -orbits in X_n .

We end this section by stating the proposition upon which the proof of Theorem 2.11 depends. Recall the mass number h of the introduction, defined by putting

$$h = \sum_{[R] \in \text{Cl}(B)} 1/w_R,$$

where w_R is the order of R^\times . Then define the real number $|\psi|$ by

$$|\psi|^2 = \frac{1}{h} \sum \psi([R])^2 / w_R = \frac{1}{h} \sum |\psi([R])|^2 / w_R.$$

For each prime $q|D = \text{disc}(K)$, we let \mathfrak{q} denote the unique (ramified) prime of O_K above q . Then $\mathfrak{q}^2 = (q)$. Since q is a rational integer prime to p , the ideal (q) represents the trivial class in $\text{Pic}(O_n)$, for all n . On the other hand, \mathfrak{q} gives a nontrivial class in $\text{Pic}(O_n)$, as soon as n is sufficiently large. This is obvious if D is not prime, since \mathfrak{q} is not even principal in that case. If D is prime, then one need simply observe that O_n contains no element of norm D , for $n \gg 0$. Let $\text{Frob}(\mathfrak{q})$ denote the Frobenius of \mathfrak{q} in $\text{Gal}(H_\infty/K)$; then it follows that $\text{Frob}(\mathfrak{q})$ is an element of order two in G_1 . We put

$$E_D = \prod_{q|D} \left(1 + \chi_t(\text{Frob}(\mathfrak{q})) \frac{a_q}{q+1} \right).$$

Note that E_D is nonzero, since $\chi_t(\text{Frob}(\mathfrak{q})) = \pm 1$, and $|a_q| \leq 2\sqrt{q}$ by the Ramanujan bounds.

Finally, put $E_p = \left(a_p^2 \left(\frac{p-1}{p+1} \right) + 1 \right)$, and let the order e_n of G_n satisfy $e_n = c \cdot p^{n-\delta-1}$.

PROPOSITION 2.14 *Let the assumptions be as in Theorem 1.4. Then the following limit formulae hold:*

1. $\lim_{n \rightarrow \infty} \frac{1}{c p^{n-\delta-1}} \sum_{P \in X_n} \sum_{\tau \in G_1} \chi_t(\tau) \psi(P^\tau) \psi(P) = E_D \cdot |\psi|^2$, and
2. $\lim_{n \rightarrow \infty} \frac{1}{c p^{n-\delta-2}} \sum_{Q \in X_{n-1}} \sum_{\tau \in G_1} \chi_t(\tau) \left(a_p \psi(Q^\tau) - \psi(\check{Q}^\tau) \right) \left(a_p \psi(Q) - \psi(\check{Q}) \right) = E_p \cdot E_D \cdot |\psi|^2$.

Inserting the limits above into equation (17), and applying the Ramanujan bound, we conclude that $L_n^{\text{av}}(g, \chi_t, 1)$ is nonzero for all n sufficiently large.

3 DISCRETE GROUPS AND GRAPHS

We have reduced the proof of Theorem 2.11 to that of Proposition 2.14. In this section, we will develop a picture of the Heegner points which will enable us to reformulate Proposition 2.14 in terms of graph theory and trees. The material in this section is perforce somewhat descriptive; it may help the reader to look directly at the picture in §3.2.

Normal Forms

3.1 Our immediate object is to give a concrete answer to the following basic problem. Suppose we are given a Heegner point $P = (f, R)$ of conductor p^n , and a conjugate point $P^\sigma = (f^\sigma, R^\sigma)$, for $\sigma \in G_n$. Then R and R^σ determine conjugacy classes of oriented Eichler orders, and we would like to

give some simple prescription for determining the class of R^σ from the data of R and f and σ . The key turns out to be the notion of normal forms of Eichler orders, as in the work of Bertolini and Darmon (see [BD98], Sec. 2).

To explain this notion, we fix an oriented Eichler order R of level N^+ . This choice determines local orders R_ℓ , for each prime ℓ (including $\ell = p$). If R' is another Eichler order, one says that R' is in *normal form* if $R'_\ell = R_\ell$ for all primes $\ell \neq p$. Here we require that, if ℓ divides N , then the orientations on R_ℓ and R'_ℓ be equal. Note that we do *not* impose any condition on R'_p . In fact, it is clear from this definition that if R' is an Eichler order in normal form, then it is completely determined by the local order R'_p . Implicit in this definition is the choice of a fixed order R ; we will simply regard a choice of R as settled.

Now let $T = T_p$ denote the Bruhat-Tits tree of B_p . Thus T is a $p + 1$ regular tree, such that the vertices of T correspond to the maximal orders of B_p , while the oriented edges correspond to oriented Eichler orders of level p . Thus, if R' is in normal form, we may identify R' with a vertex v on the tree. We will see below that the set $\text{Cl}(B)$ may be identified with a certain finite quotient graph \mathcal{G}' of T . Each vertex of \mathcal{G}' will correspond to an element of $\text{Cl}(B)$. If \bar{v} is the image of v in \mathcal{G}' , the class $[R']$ is determined as the class corresponding to \bar{v} . Given Heegner points P and P^σ as above, we are led therefore to put the orders R and R^σ in normal form. We will see that, from this viewpoint, the action of the Galois group takes on a rather simple aspect.

Action of the local Galois group

3.2 Fix one of the h_K Heegner points $P = (f, R)$ of level 1, so that f is an embedding $K \rightarrow B$ with $f(K) \cap R = f(O_K)$. This choice of R determines local orders $R_\ell = R \otimes \mathbf{Z}_\ell \subset B_\ell$, for each prime ℓ . In particular, we obtain a maximal order $R_p \subset B_p$. The order R_p determines a vertex v_0 on T . It will be convenient to regard v_0 as determining an *origin* of the tree. This choice of origin also gives rise to an orientation on T , if we say that a vertex v is even or odd according to whether the distance from v to v_0 is even or odd, and require that a positively oriented edge always go from an even vertex to an odd one.

Now let v be any other vertex on T , corresponding to the order $R'_p \subset B_p$. Then define $\hat{R}' = R'_p \times \prod_{\ell \neq p} R_\ell$, and let $R' = B \cap \hat{R}'$. The pair $P' = (f, R')$ determines a Heegner point of conductor p^n , for some n . The conductor p^n of P' is closely related to the distance between v and v_0 on the tree T . The precise relationship depends on K and p , as follows.

Recall that G_1 denotes the tame subgroup of $\text{Gal}(H_\infty/K)$. Let $\Delta_0 \subset G_1$ denote the image of (tame) inertia. Thus Δ_0 has order $(p \pm 1)/u$, where the sign is \pm depending on whether p is split or not, and the integer u is equal to one half the order of O_K^\times . We simply write δ_0 to denote the order

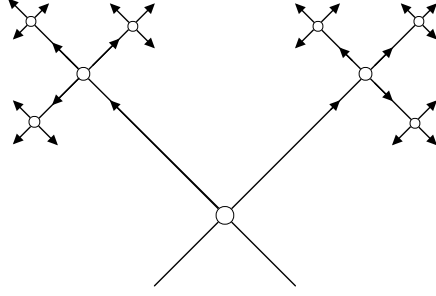


Figure 1: the tree when $p = 3$

of Δ_0 . Note that the tree T has degree $p + 1$, independently of p and K . In particular, the origin v_0 has $p + 1$ neighbours. One can check that of these $p + 1$ neighbors, either $p - 1$ or $p + 1$ give rise to Heegner points of level p , depending on whether p is split or not. Let v denote a neighbor of v_0 such that v has level p ; we will call v a *good neighbor*. If K is not $\mathbf{Q}(\sqrt{-1})$ or $\mathbf{Q}(\sqrt{-3})$, so that $u = 1$, then the Heegner points corresponding to good neighbors of v_0 are distinct modulo B^\times . If $K = \mathbf{Q}(\sqrt{-1})$ or $K = \mathbf{Q}(\sqrt{-3})$, then there is a small complication, arising from the presence of nontrivial units in O_K , that must be taken into account. In these cases, each of the $p \pm 1$ good neighbors of v_0 is equivalent to u others, and we fix, once and for all, a choice of representatives for the classes modulo B^\times . It is clear that such a set of representatives has cardinality $(p \pm 1)/u$. In order to keep the notation uniform, we will change terminology slightly and say that v is a good neighbor if it is in the fixed set of representatives modulo B^\times . Thus, in every case, the origin v_0 has precisely δ_0 good neighbors. Each good neighbor v determines an oriented edge from v_0 to v , which we will call a *good direction*.

With this in mind, consider a vertex $v = v_n$ at distance n from the origin. There is a unique path v_0, v_1, \dots, v_n leading from v_0 to v_n . We call v_n a *good vertex* if the walk from v_0 to v exits v_0 in a good direction. Equivalently, v is good if the vertex v_1 encountered at the first step of the walk is a good neighbor. Figure 1 above, drawn when $p = 3$ and $\delta_0 = 2$, may help the reader to visualize the situation. The good directions are marked by arrows.

It can be shown that if v is a good vertex, then the Heegner point determined by v has conductor p^n . It is clear that there are $\delta_0 \times p^{n-1}$ good vertices at distance n . Indeed, each such vertex is uniquely determined by the walk v_0, v_1, \dots, v_n , and there are δ_0 choices for v_1 , followed by p choices for each subsequent v_s . Write V'_n for the set of good vertices at distance $n - \delta$ and X'_n for the set of Heegner points so determined. Thus, if $v \in V'_n$, then the corresponding $P \in X'_n$ is determined by the pair (f, R') , where $R'_\ell = R_\ell$ for $\ell \neq p$, and R'_p is the order corresponding to v . Here the map $f : K \rightarrow B$

and the local orders R_p and R_ℓ are determined by the Heegner point $P = (f, R)$ corresponding to the origin v_0 , as fixed at the outset. In particular, the map f is independent of v .

3.3 We now consider the action of the local Galois group on the good Heegner points defined above. Thus write $U = U(K_p)$ for the local units in $O_K \otimes \mathbf{Z}_p$. Then $f(U) \subset R_p^\times$. Furthermore, class field theory gives a homomorphism $U \rightarrow G_n$ whose image is the inertia group I_n . Using the description of the Galois action from §2.6, one sees then that the group I_n acts on the Heegner points $P' \in X'_n$ according to the following prescription. Given $\sigma \in I_n$, one selects a local idele $\tilde{\sigma} \in U(K_p)$ such that the Artin symbol of $\tilde{\sigma}$ in I_n coincides with σ . Here we view $\tilde{\sigma}$ as an idele by giving it component one away from p . Now suppose that $P' = (f, R')$ as above, where the local order R'_p corresponds to the vertex $v \in T$. Then $P'^\sigma = (f, R'^\sigma)$, where R'^σ is determined by $\hat{R}'^\sigma = f(\tilde{\sigma})R'f(\tilde{\sigma}^{-1})$. Since the idele $\tilde{\sigma}$ is trivial away from p , one has simply to compute $f(\tilde{\sigma})R'_p f(\tilde{\sigma}^{-1})$. But it is easy to see that, in terms of the tree T , the order $f(\tilde{\sigma})R'_p f(\tilde{\sigma}^{-1})$ corresponds to the vertex $f(\tilde{\sigma})v$.

Thus, if we view the Heegner points $P \in X'_n$ as vertices on the tree, the action of I_n is by *translations*. This will be of the utmost importance in what follows. Note also that the inertia subgroup I_n admits a canonical decomposition $I_n = \Delta_0 \times \Delta'_n$, where Δ_0 is the maximal subgroup of order prime to p and $\Delta'_n \subset \Delta_n$ has index p^δ . Equivalently, Δ_0 is the image under the Artin map of the roots of unity in $U(K_p)$. Thus, if $\sigma \in \Delta_0$, then we can choose the corresponding idele $\tilde{\sigma}$ independently of n . This fact will also be important later.

The following proposition summarizes the basic facts about the good Heegner points and the local Galois action.

PROPOSITION 3.4 *There are $\delta_0 \times p^{n-1-\delta}$ good vertices at distance $n - \delta$ from v_0 , each corresponding to a distinct Heegner point $P \in X'_n$. If a vertex v is good, and corresponds to the Heegner point P' , then the vertex v^σ corresponding to the conjugate point P'^σ is $v^\sigma = f(\tilde{\sigma})v$. The set X'_n of good vertices is a simple homogeneous space for the action of the group I_n .*

Proof. Clear, in light of the preceding remarks. The fact that I_n acts simply transitively on X'_n comes from the fact that G_n acts simply transitively on X_n . \square

Action of the ideal class group

3.5 The preceding discussion gives a complete description of the action of G_n if the class group of K is trivial, since $G_n = I_n$ in this case. In order to treat the case of nontrivial class group, we need to enlarge the picture to include the h_K distinct choices of base point P of conductor 1. Roughly

speaking, one obtains several copies of the above picture, one for each of the h_K possible points of conductor 1, which are related by the action of the class group $\text{Cl}(K)$ of K .

Instead of working with $\text{Cl}(K)$ directly, it is actually more convenient to work with the group G_1 . So let $\tau \in G_1$, and fix an idele $\tilde{\tau}$ of K whose Artin symbol in $\text{Gal}(H_\infty/K)$ is equal to τ . Write

$$\tilde{\tau} = (t_\ell)_{\ell \neq p} \times t_p,$$

where t_* denotes the local component at the place $*$. If v is any vertex, we get a Heegner point $P' = (f, R')$ as before, by putting $\hat{R}' = R'_p \times \prod_{\ell \neq p} R_\ell$, where R'_p is the local order corresponding to v . Then the conjugate point P'^τ is given by the pair $P'^\tau = (f, S)$, where $\hat{S} = f(\tilde{\tau})\hat{R}'f(\tilde{\tau})^{-1}$. In order to put the order S into normal form, we select $b \in B^\times$ such that $b_\ell S_\ell b_\ell^{-1} = R_\ell$, for each $\ell \neq p$. Such a b exists by strong approximation in \hat{B} , since the local orders B_ℓ and S_ℓ are conjugate in B_ℓ . Note that the element b depends only on τ : it is enough to choose b such that $b_\ell t_\ell \in \mathbf{Q}_\ell^\times R_\ell^\times$, for $\ell \neq p$. Then the point P'^τ is represented by the pair

$$P'^\tau = (f', S') \tag{19}$$

where f' and S' are deduced from f and S via conjugation by the element $b \in B^\times$. The order S' is in normal form by definition.

Recall that the point P' was given by the pair (f, R') , where R' was in normal form, corresponding to the vertex v of the tree T . It follows from the definition of S' that the vertex corresponding to S'_p is $v^\tau = \tau_p v$, where $\tau_p = b_p f(t_p)$, and b_p and t_p are the local components in B_p of the elements b and $\tilde{\tau}$ respectively. Thus we find once again that the action of the Galois group on the Heegner points reduces simply to a translation. But the element τ_p does *not* in general fix the origin v_0 , as occurred in the case of the local Galois action. Furthermore, the embedding f has changed to f' (although this will not play an important role in the sequel).

For each $\tau \in G_1$, we obtain a vertex $v_0^\tau = \tau_p v_0$, and a collection of vertices $V_n'^\tau = \tau_p V_n'$ at distance $n - \delta$ from v_0^τ . We write $X_n'^\tau$ for the set of Heegner points of conductor $p^{n-\delta}$ of the form P'^τ , for $P' \in X_n'$. Then it is clear that the points $P'^\tau \in X_n'^\tau$ are represented by pairs (f', S') as above, where S' is in normal form, corresponding to the vertex $v' \in V_n'^\tau$, and the embedding f' depends only on τ , and is deduced from f by conjugation by the element $b = b_\tau \in B^\times$.

We summarize this discussion as

PROPOSITION 3.6 *Let $P = (f, R)$ denote a fixed Heegner point of level 1. Let $\tau \in G_1$, and fix an idele $\tilde{\tau} \in \hat{K}$ whose Artin symbol in $\text{Gal}(H_\infty/K) = G_1 \times \Delta_\infty$ is equal to τ . Let $P' = (f, R')$ denote a good Heegner point in X_n' , corresponding to the vertex v at distance $n - \delta$. Then:*

1. The conjugate point P'^τ may be represented by a pair (f', S') , where the order S' is in normal form, and S'_p corresponds to the vertex $v^\tau = \tau_p v$, for some fixed $\tau_p \in B_p$ which is independent of n and P' .
2. The element τ_p can be written as $\tau_p = b_p f(t_p)$, where t_p and b_p are the local components at p of $\tilde{\tau}$ and some $b \in B^\times$ respectively.

Let $G_0 \subset G_1$ denote the genus subgroup, namely, the subgroup generated by the elements $\text{Frob}(\mathfrak{q})$, where \mathfrak{q} is the unique prime of O_K above q , for each (ramified) rational prime q dividing $D = \text{disc}(K)$. Then the following innocuous proposition is actually rather important. It states that the action of $\tau \in G_1$ is rational if and only if $\tau \in G_0$.

PROPOSITION 3.7 *Let τ_p be as in Proposition 3.6. Then $\tau_p \in \mathbf{Q}_p^\times B^\times \subset B_p^\times \iff \tau \in G_0$.*

Proof. Suppose that $\tau_p \in \mathbf{Q}_p^\times B^\times \subset B_p$. We want to prove that $\tau \in G_0$. We begin with a series of reductions. Since H_∞/K is an anticyclotomic extension, elements of \mathbf{Q}_p^\times are local norms from H_∞ . Changing $\tilde{\tau}$ if necessary, we may assume that $\tau_p \in B^\times$. Since $\tau_p = b_p f(t_p)$, and b_p is clearly a \mathbf{Q} -rational point of B_p , we may even assume that $f(t_p)$ is a \mathbf{Q} -rational point. Finally, since $f(K)$ is a maximal commutative subalgebra of B , and since $f(t_p)$ commutes with $f(K)$, we can reduce to the case that $f(t_p)$ is in the image of K^\times .

Thus, recall that t_p is the local component of the idele $\tilde{\tau}$, where $\tilde{\tau}$ corresponds via class field theory to τ , and suppose that there exists $k \in K$ such that the local components of $\tilde{\tau}$ and k are equal in K_p^\times . Then $\tilde{\tau}k^{-1}$ has local component 1 at p . Furthermore, the Artin symbols of $\tilde{\tau}$ and $\tilde{\tau}k^{-1}$ are equal. Now let \mathfrak{a} denote the fractional ideal of K defined by $\tilde{\tau}k^{-1}$. Since $\tilde{\tau}k^{-1}$ is trivial at p , the ideal \mathfrak{a} is prime to p .

Since the extension H_∞/K is unramified away from p , the Artin symbol of \mathfrak{a} is well-defined in $\text{Gal}(H_\infty/K)$. From the definitions, we find that the Artin symbol of \mathfrak{a} is given by $\tau \in G_1$, which is an element of finite order m . Replacing \mathfrak{a} by $\mathfrak{b} = \mathfrak{a}^m$, we see that $\mathfrak{b} = (\beta)$ is principal, and has trivial Artin symbol in $\text{Gal}(H_\infty/K)$. Scaling β by a suitable rational integer, we may assume that β is even integral. Note that this does not change the Artin symbol.

Now, according to class field theory, the Artin map induces an isomorphism $\text{Pic}(O_n) \cong G_{n+\delta} = \text{Gal}(H_{n+\delta}/K)$. It follows that, for each $n \geq 1$, the ideal (β) has a generator β_n which is congruent to a rational integer modulo p^n . Since K is an imaginary quadratic field, the ideal (β) can have only finitely many generators (there are only finitely many units). Changing the generator β if needed, we may assume that β is congruent to a rational integer modulo p^n for infinitely many n . It follows that β lies in the closure of \mathbf{Z} in $K \otimes \mathbf{Z}_p$. Since $\beta \in O_K$, we conclude that $\beta \in \mathbf{Z}$. Thus we find that $\mathfrak{a}^m = (\beta)$,

where $\beta \in \mathbf{Z}$. Letting σ denote the nontrivial automorphism of K/\mathbf{Q} , we then find that

$$\left(\frac{\mathfrak{a}^\sigma}{\mathfrak{a}}\right)^m = \frac{(\mathfrak{a}^\sigma)^m}{\mathfrak{a}^m} = \frac{(\mathfrak{a}^m)^\sigma}{\mathfrak{a}^m} = \frac{\beta^\sigma}{\beta} = 1.$$

Thus $(\mathfrak{a}^\sigma/\mathfrak{a})^m$ is the unit ideal, and $\mathfrak{a}^\sigma = \mathfrak{a}$ by unique factorization. It follows from this that, if a split prime \mathfrak{q} divides \mathfrak{a} , then so does the conjugate \mathfrak{q}^σ . Thus \mathfrak{a} is divisible by $\mathfrak{q}\mathfrak{q}^\sigma = (q)$, where q is a rational prime. On the other hand, all inert primes in K are already principal, generated by rational integers. Thus $\mathfrak{a} = \mathfrak{a}' \cdot (b)$, where \mathfrak{a}' is a product of ramified primes, and (b) is principal and generated by a rational integer $b \in \mathbf{Z}$. Then (b) has trivial Artin symbol in G_n for all n . So the Artin symbol of \mathfrak{a} coincides with that of \mathfrak{a}' , which is a product of ramified primes, as required.

As for the converse, we have already remarked that if \mathfrak{q} is a prime of O_K which is ramified in K/\mathbf{Q} , then $\tau = \text{Frob}(\mathfrak{q})$ is an element of order two in G_1 . (See the discussion preceding Proposition 2.14.) By definition, the elements $\text{Frob}(\mathfrak{q})$ generate G_0 , and it suffices to show that $\tau_p \in B^\times$, for $\tau = \text{Frob}(\mathfrak{q})$. Thus let π denote a uniformizer of K at a prime $\mathfrak{q}|D$, and let $\tilde{\pi}$ denote the idele of K with component π at \mathfrak{q} , and with component 1 elsewhere. Then $\tilde{\pi}$ corresponds to $\text{Frob}(\mathfrak{q})$ under the Artin map. By strong approximation, there exists $b \in B^\times$ such that $bf(\tilde{\pi}) \in \mathbf{Q}_\ell^\times R_\ell^\times$, for all primes ℓ with $\ell \neq p$. It follows that we can take this b in the definition of τ_p , so that $\tau_p = b_p f(\tilde{\pi}_p)$. Since $\tilde{\pi}_p = 1$ by definition, we have $\tau_p = b_p \in B^\times \subset B_p$, as required. \square

3.8 We want to make some conventions about the genus subgroup G_0 , as this will be useful later. By definition, $G_0 \subset G_1$ is the subgroup generated by the elements $\text{Frob}(\mathfrak{q})$, where \mathfrak{q} runs over the primes of O_K that are ramified over \mathbf{Z} . Let r denote the number of such primes. Then it is clear that G_0 has cardinality 2^r . Indeed, it is clear that the order of G_0 is bounded from above by 2^r , since each Frobenius element is quadratic. On the other hand, it is well-known from genus theory that the image of G_0 in $\text{Pic}(O_K)$ has order 2^{r-1} , owing to the unique relation $\prod_{i=1}^r \mathfrak{q}_i = (\sqrt{D})$, which represents the trivial ideal class. Since (\sqrt{D}) is nontrivial in $\text{Pic}(O_n)$ for n sufficiently large, it follows that the various $\text{Frob}(\mathfrak{q}_i)$ are independent in $G_0 \subset G_1$. Thus, given $\tau \in G_0$, there exists a unique subset $I_\tau \subset \{q_1, q_2, \dots, q_r\}$ such that

$$\tau = \prod_{\mathfrak{q} \in I_\tau} \text{Frob}(\mathfrak{q}).$$

Conversely, any $I \subset \{q_1, q_2, \dots, q_r\}$ is of the form I_τ for a unique $\tau = \prod_{\mathfrak{q} \in I} \text{Frob}(\mathfrak{q})$. For $\tau \in G_0$ we may therefore define a squarefree integer $d = d_\tau$ by saying

$$d = d_\tau = \prod_{\mathfrak{q} \in I_\tau} q.$$

This notation will be helpful in the final section of this paper.

Finite quotients of the tree

In this section, we describe how to compute the conjugacy class of an Eichler order in normal form. The principal ideas come from graph theory, so we need to fix some notation, as in [God93].

3.9 A graph G consists of a set $V(G)$ of vertices of G , and a set $E(G)$ of edges. Each edge \vec{e} of G determines an ordered pair (v, w) of vertices of G , called the extremities of \vec{e} . The vertex v is called the origin of \vec{e} and the vertex w is called its terminus. The edge \vec{e} is said to join v to w . The set $E(G)$ is equipped with an involution $\vec{e} \mapsto \vec{e}'$, with the property that if \vec{e} joins v to w , then \vec{e}' joins w to v . A *loop* is an edge whose extremities coincide. A graph has *multiple edges* if distinct edges have the same origin and terminus. The *degree* of a vertex v is the number of edges originating from v . A *regular* graph is one which has neither loops nor multiple edges, and such that the degree $d = d(v)$ of any vertex v is independent of v . The integer d is called the degree of the graph. The graph G is called *connected* if for any pair (v, w) of vertices, there exists a sequence of edges leading from v to w .

Suppose G is a graph on a finite number m of vertices. Then the adjacency matrix $A(G)$ of G is the $m \times m$ matrix whose rows and columns are indexed by the vertices of G . The entry of $A(G)$ in the (v, w) -th position is number of edges from v to w . Then $A(G)$ is a symmetric matrix. It is well-known that the eigenvalues of $A(G)$ are real, and that $A(G)$ is diagonalizable. If G is regular and connected, of degree d , then the Frobenius-Perron theorem (see [God93]) states that the largest eigenvalue (in absolute value) of $A(G)$ is $\lambda = d$, and that this eigenvalue occurs with multiplicity one.

The graph G is called bipartite if $\lambda = -d$ is also an eigenvalue of $A(G)$. Geometrically, this means that the vertices of G are divided into two subsets, with every edge going from one subset to the other. Equivalently, to say that G is not bipartite means that there exists a circuit of odd length starting and ending at any given vertex (see [CDS80], Theorem 3.4).

3.10 Let $P = (f, R)$ denote the fixed Heegner point of conductor 1 selected above. Let $\Gamma' \subset B^\times$ denote the group $R[1/p]^\times$. Then it follows from strong approximation in \hat{B} that

$$\text{Cl}(B) = B^\times \backslash \hat{B}^\times / \hat{R}^\times = \Gamma' \backslash B_p^\times / \mathbf{Q}_p^\times R_p^\times.$$

Observe that $B_p^\times / \mathbf{Q}_p^\times R_p^\times \cong PGL_2(\mathbf{Q}_p) / PGL_2(\mathbf{Z}_p)$ is the vertex set of the tree T of B_p introduced in section 3.1. Thus we have identified the set $\text{Cl}(B)$ with the vertices of the quotient graph $\mathcal{G}' = \Gamma' \backslash T$. Note however that \mathcal{G} need not be a regular graph, because the group Γ' may have nontrivial elements of finite order. It will be convenient therefore to replace Γ' with a congruence subgroup Γ of finite index, such that Γ is torsion-free.

Thus let $M > 4$ denote a prime number with $(M, Nd_p) = 1$. We will assume that $M \equiv 3 \pmod{4}$ and that p is a quadratic residue modulo M . Then we let $\Gamma(M) \subset R[1/p]^\times$ denote the

subgroup of elements that are congruent to 1 modulo M , where we have fixed an isomorphism $R_M \cong M_2(\mathbf{Z}_M)$. We let $\Gamma \subset \tilde{G} = PGL_2(\mathbf{Q}_p)$ denote the projectivization of $\Gamma(M)$. Finally, we let $\Gamma_+ = \Gamma \cap PSL_2(\mathbf{Q}_p)$.

It follows from the fact that $M > 4$ that $\Gamma(M)$, and so Γ and Γ_+ , are torsion-free. Since we have assumed that $M \equiv 3 \pmod{4}$ and that p is a square modulo M , we find that p has odd order modulo M . By Lemma 7.3.9 of [Lub94], this implies that the quotient $\mathcal{G} = \Gamma(M) \backslash \mathcal{T} = \tilde{\Gamma} \backslash \mathcal{T}$ is *not* bipartite. We will use this fact later. On the other hand, it is clear that $\Gamma_+ \backslash \mathcal{T}$ is indeed bipartite. The inclusion $\Gamma(M) \subset R[1/p]^\times$ induces a finite-to-one map

$$\mathcal{G} = \Gamma(M) \backslash \mathcal{T} \rightarrow R[1/p]^\times \backslash \mathcal{T} = \text{Cl}(B).$$

We would like to view the function $\psi : \text{Cl}(B) \rightarrow \mathbf{R}$ as being defined on \mathcal{G} , via the projection above, and transfer calculations such as those appearing in Proposition 2.14 to the regular graph \mathcal{G} . However, there is a slight complication that arises from the presence of elliptic elements in $R[1/p]^\times$, which has to be navigated before we can proceed further. Thus, for a vertex $v \in \mathcal{T}$, we write $e = e_v$ to denote the order of the stabilizer in Γ' .

LEMMA 3.11 *The integer e_v is equal to the order of R_v^\times , where R_v is the Eichler order in normal form determined by v . If $\tilde{\psi} : \mathcal{G} \rightarrow \mathbf{R}$ is the pullback of ψ to \mathcal{G} , then $\sum_{v \in \mathcal{G}} \tilde{\psi}(v) = \sum_{[R_i] \in \text{Cl}(B)} \psi(R_i) / w_i = 0$, where w_i is the order of R_i^\times .*

Proof. The first assertion is clear. As for the second, let L^2 denote the finite-dimensional space of functions on the vertices of \mathcal{G} , equipped with the inner-product

$$(\phi_1, \phi_2) = \sum_{v \in \mathcal{G}} \phi_1(v) \bar{\phi}_2(v).$$

Define an operator ∇ on L^2 by saying

$$(\nabla \phi)(x) = \sum_{y \sim x} \phi(y),$$

for any $\phi \in L^2$, where the sum is taken over the neighbors of x . Then ∇ is a symmetric and positive-definite operator (see for example [Lub94], §4.2). By the definition of the Hecke operator T_p , we find that $\tilde{\psi}$ is an eigenfunction for ∇ , with eigenvalue $a_p = a_p(g)$. On the other hand, it is clear that the identity function $\mathbf{1}$ is also an eigenvector for ∇ , with eigenvalue $p + 1$. Then $a_p \neq p + 1$, since g is a cuspform. The proposition now follows from the fact that the vectors $\mathbf{1}$ and $\tilde{\psi}$ are orthogonal.

This argument also completes the proof of Lemma 2.5. Namely, Lemma 2.5 simply states the fact that the eigenvalue of $\mathbf{1}$ is $p + 1$, for any $p \nmid N$. \square

Random walks on graphs

3.12 We start by studying a general situation. Let G denote an arbitrary connected graph, assumed regular of degree d . We say that a walk on G has *no backtracking* if at no stage of the walk does one traverse an edge, and then immediately go back along the same edge in the opposite direction (see [Fri91], page 356). Given vertices x and y of G , we say that

$$x \sim_n y$$

if there is a length n walk without backtracking that starts at x and ends at y . (The length of a walk is the number of edges traversed.) It is clear that there are $d(d-1)^{n-1}$ walks without backtracking of length n departing from a given x . We will usually say ‘walk’ instead of ‘walk without backtracking,’ since these are only kinds of walks that we consider in the present work. If G is a tree, then $x \sim_n y$ if and only if x and y are distance n apart.

3.13 Now suppose that G is a finite graph on m vertices. Let a vertex $x \in G$ be fixed. Write W_n for the set of terminal vertices of all walks of length n that start from x , counted with the appropriate multiplicity. Then the set W_n has cardinality $d(d-1)^{n-1}$. The basic principle we will use is that, in this situation, the set W_n is uniformly distributed among the m vertices of G . Indeed, if $\mu_n(y)$ denotes the multiplicity of y in W_n , then

$$\frac{\mu_n(y)}{d(d-1)^{n-1}} \sim \frac{1}{m}, \quad (20)$$

as $n \rightarrow \infty$. In fact, one has the following result.

PROPOSITION 3.14 *Let ϕ be any complex valued function on G . Then, if G is not bipartite, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{d(d-1)^{n-1}} \sum_{y \in W_n} \phi(y) = \frac{1}{m} \sum_{x \in G} \phi(x). \quad (21)$$

Proof. This seems to be well-known amongst the graph-theorists. A proof may be found in [Fri91], Lemma 3.3, and Lemma 3.4. Note that in the special case where ϕ is the characteristic function of a vertex, the proposition reduces to the formula (20). \square

3.15 We will actually need a slightly more general statement than the one given above. As before, let a vertex $x \in G$ be fixed. Suppose we are given a collection of good neighbors of x , namely, vertices x_1, x_2, \dots, x_r of G adjacent to x . Each x_i determines an edge emanating from x , which we call a good direction. We will say that a walk (without backtracking, as always) is good if it exits from x in a good direction. Let $W_n = W_n(x)$ denote the set of terminal vertices of the good walks of length n ,

counted with multiplicity. Obviously, there are $r(d-1)^n$ such vertices. Then we have the following generalization of Proposition 3.14:

PROPOSITION 3.16 *Suppose that G is not bipartite. Let ϕ be any complex valued function on G . Then*

$$\lim_{n \rightarrow \infty} \frac{1}{r(d-1)^n} \sum_{y \in W_n} \phi(y) = \frac{1}{m} \sum_{x \in G} \phi(x). \quad (22)$$

Proof. We briefly sketch the method of deducing this from the previous proposition. By linearity, it is enough to consider the case where there is a single good direction from x , going, say, from x to y . Write $W_n(x)$ for the set of terminal vertices of good walks of length n from x , as above. For a vertex $a \in G$, we will write $W'_n(a)$ to denote the set of terminal vertices of all non-backtracking walks of length n , starting from a . Thus $W_n(x) \subset W'_n(x)$.

Let the function ϕ be given. Then write $T_n(\phi) = \sum_{v \in W_n(x)} \phi(v)$, and, for any $a \in G$, define $S_n(a, \phi) = \sum_{v \in W'_n(a)} \phi(v)$. Now let consider any walk of length n that exits x in a good direction. Such a walk reaches y at the first step, and then proceeds from y as a walk without backtracking. Thus each good walk of length from x determines a walk of length $n-1$ from y . The subset of walks from y that are obtained in this way is easily described: it is the set of walks that exit y in any direction other than the one leading back to x . Write $W_{n-1}(y, x)$ for the set of terminal vertices of the walks of length $n-1$ from y that simply return to x at the first step, and let $U_{n-1}(\phi) = \sum_{v \in W_{n-1}(y, x)} \phi(v)$. Then we have

$$T_n(\phi) = S_{n-1}(y, \phi) - U_{n-1}(\phi).$$

But each walk from y that returns to x in turn determines a walk of length $n-2$ from x that exits x in any direction other than the one back (again) to y . Thus we get

$$T_n(\phi) = S_{n-1}(y, \phi) - (S_{n-2}(x, \phi) - T_{n-2}(\phi)).$$

It follows that

$$T_n(\phi) = S_{n-1}(y, \phi) - S_{n-2}(x, \phi) + S_{n-3}(y, \phi) \cdots \pm S_1(z, \phi) \mp c_0,$$

where c_0 is some constant and $z = x$ or $z = y$, depending on whether n is odd or even. Thus,

$$\frac{T_n(\phi)}{(d-1)^{n-1}} = \frac{c_0}{(d-1)^{n-1}} + \sum_{j=1}^{n-1} (-1)^{n-1-j} \frac{S_j(z_j, \phi)}{(d-1)^{n-1}},$$

where $z_j = x$ or y according to the parity of j . By Proposition 3.14, the sum on the right converges absolutely as $j \rightarrow \infty$. Letting A denote the quantity $\frac{1}{m} \sum_{x \in G} \phi(x)$, we find that

$$\lim_{n \rightarrow \infty} \frac{T_n(\phi)}{(d-1)^{n-1}} = dA \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{(d-1)^j} = A,$$

as required. □

3.17 We can now complete the proof of Theorem 1.5 of the introduction, recalling at the outset the notations and definitions of the section on normal forms and Galois actions. In particular, we have a fixed base point P of level 1, giving an origin v_0 on the tree. Let X'_n denote the set of Heegner points corresponding to the good vertices at distance $n - \delta$. (Recall that δ denotes the nonnegative integer such that $G_{n+\delta} = \text{Pic}(O_n)$, for $n \gg 0$.) For $\tau \in G_1$, we get a different origin, and a different set $X'_n{}^\tau$ of good vertices at distance $n - \delta$. Then, projecting the set $X'_n{}^\tau$ onto the finite graph \mathcal{G} , we find from Proposition 3.16 that

$$\lim_{n \rightarrow \infty} \frac{1}{p^{n-1-\delta}} \sum_{P \in X'_n{}^\sigma} |\psi(P)|^2 = \frac{\delta_0}{m} \sum_{x \in \mathcal{G}} |\psi(x)|^2.$$

Here δ_0 denotes the order of the (tame) inertia subgroup in G_0 , as in §3.2. Summing this over all τ , we get

$$\lim_{n \rightarrow \infty} \frac{1}{p^{n-\delta-1}} \sum_{P \in X_n} |\psi(P)|^2 = \frac{c}{m} \sum_{x \in \mathcal{G}} |\psi(x)|^2 = \frac{c}{h} \sum_{R \in \text{Cl}(B)} \psi(R)^2 / w_R,$$

as asserted.

4 THE THEOREMS OF RATNER

In this section we will introduce the theorems of Ratner, and start collecting the consequences, in preparation for the proof of our main results in the next section.

Recall that our general aim is to determine the relations between the classes of Heegner points P and P^τ , for $\tau \in G_1$. The following theorem, which we shall deduce from Ratner's results, states that $[P]$ and $[P^\tau]$ are independent, as long as $\tau \notin G_0$. The proof will introduce some of the ingredients that will be needed later.

THEOREM 4.1 *Let C_1 and C_2 be arbitrary conjugacy classes of oriented Eichler orders of level N^+ , and let τ be an arbitrary nontrivial element of G_1 , such that $\tau \notin G_0$. Then there exists a Heegner point P of conductor p^n such that $P = (f_1, R_1)$ and $P^\tau = (f_2, R_2)$, where the order R_i represents the class C_i . Furthermore, one can choose the point P so that the integer n given by $\text{Cond}(P) = p^n$ has any desired parity.*

REMARK 4.2 The idea of the proof is simple. According to the description of the Galois action obtained earlier, it will suffice to find a vertex $v \in T$ corresponding to R_1 , such that the vertex $v^\tau = \tau_p v$ corresponds to R_2 . Note however that we may replace v by $\gamma_1 v$, for $\gamma_1 \in \Gamma$ without changing the image in $\text{Cl}(B)$. This has the effect of replacing $\tau_p v$ with $\tau_p \gamma_1 v$. One can then change $\tau_p v$ by translating

further by $\gamma_2 \in \Gamma$, to get the vertex $\gamma_2 \tau_p \gamma_1 v$. Thus we are led to prove that the double coset $\Gamma \tau_p \Gamma$ is a large subset of B_p^\times . For example, if one knew that $\Gamma \tau_p \Gamma$ were dense in B_p^\times , the theorem would follow immediately. This is in essence what we will do below.

4.3 It is clear that $\Gamma \tau_p \Gamma$ cannot be dense if $\Gamma \cap \tau_p \Gamma \tau_p^{-1}$ has finite index in each of Γ and $\tau_p \Gamma \tau_p^{-1}$. Before proceeding to Ratner's theorem, we therefore recall some simple facts about commensurability. Thus, recall that if H_1 and H_2 are subgroups of the group G , then we say that H_1 and H_2 are commensurable if $H_1 \cap H_2$ is of finite index in H_1 and H_2 . Given a subgroup H of G , the commensurator of H in G is the set of all $x \in G$ such that xHx^{-1} is commensurable with H . If H_1 and H_2 are commensurable, then they have the same commensurator. Applying these notions in our context, we have

PROPOSITION 4.4 *Let Γ be the subgroup fixed in (3.10). If $\alpha \in B_p^\times$ is such that $\Gamma_\alpha = \alpha \Gamma \alpha^{-1}$ is commensurable with Γ , then $\alpha \in \mathbf{Q}_p^\times B^\times$.*

Proof. See [Vig80], Corollaire 1.5, page 106. Note that Γ is actually the projectivization of $\Gamma(M) \subset R[1/p]^\times \subset B_p^\times$, but since conjugation by \mathbf{Q}_p^\times is trivial, the conjugate $\Gamma_\alpha = \alpha \Gamma \alpha^{-1}$ is well-defined for $\alpha \in B_p$.

4.5 We now want to state the theorems of Ratner that are crucial to our proof of Theorem 4.1. As a reference for this material, we refer the reader to the introduction of Ratner's paper [Rat95], and also to Chapter 1 of [Rag72] for the basic facts about lattices.

Let G be any p -adic Lie group and let Γ be a discrete subgroup. Then Γ is called a *lattice* in G if $\Gamma \backslash G$ admits a finite G -invariant measure. In particular, if Γ is discrete and cocompact, then it is a lattice in this sense. A subset $A \subset \Gamma \backslash G$ is called *homogeneous* if there is an $x \in \Gamma \backslash G$ and a closed subgroup $H \subset G$ such that $xHx^{-1} \cap \Gamma$ is a lattice in H , and $A = xH$. This implies in particular that A is closed in $\Gamma \backslash G$. An element $u \in G$ is called *ad-unipotent* if the automorphism $\text{Ad}(u)$ of the Lie algebra \mathfrak{G} of G is unipotent. In particular, an element that is unipotent in the usual sense is also ad-unipotent. Finally, a 1-parameter subgroup of G is an additive homomorphism $u : \mathbf{Q}_p \rightarrow G$. It is a fact that if G is a p -adic Lie group, then the image of any one-parameter subgroup consists of ad-unipotent elements. Ratner's deep result is then the following:

THEOREM 4.6 ([Rat95], THEOREM 2) *Let G and Γ be as above. Let $U \subset G$ be any subgroup generated by 1-parameter subgroups. Then, for any $x \in \Gamma \backslash G$, the closure of the orbit xU is homogeneous. Thus there exists a closed subgroup H of G such that $\overline{xU} = xH$ in $\Gamma \backslash G$.*

REMARK 4.7 In our application it will be important to have $U \subset H$. This is in fact true in general,

and seems to be well-known to the experts, but since this is not explicitly stated in Ratner's theorem, we give an elementary argument to reduce to this situation. Let $X \subset \Gamma \backslash H$ denote the closure \overline{xU} of xU . Then, by Ratner's theorem above, $X = xH$, so that the set X is stable under right translation by H . On the other hand the dense subset $xU \subset X$ is stable under U . Since X is closed, it follows that X is actually stable under U as well. Thus X is invariant under the subgroup generated algebraically by U and H . Since X is closed, we see that it is even invariant under the closed subgroup \tilde{H} generated by U and H . Then $x\tilde{H} \subset X$, and since $H \subset \tilde{H}$, we have $X = xH \subset x\tilde{H}$. It follows that $X = x\tilde{H}$, and $U \subset \tilde{H}$. Thus we may replace H by \tilde{H} if necessary to assume that $U \subset H$.

REMARK 4.8 We want to make one further observation about the group H appearing in the statement of Ratner's theorem. By the foregoing remarks, we may assume that $U \subset H$, and that the orbit of H is the closed set \overline{xU} . It follows that H may be taken to be the smallest subgroup of G such that $U \subset H$, and that the orbit of H is closed in $\Gamma \backslash G$. When Γ is a cocompact lattice, one knows that the orbit xH of H is closed in $\Gamma \backslash G$ if and only if $\Gamma \cap xHx^{-1}$ is a lattice in xHx^{-1} . (See [Rag72], Theorem 1.13.) Thus, if Γ is cocompact, we may characterize H as the smallest subgroup of G which contains U , and which is such that $\Gamma \cap xHx^{-1}$ is a lattice in xHx^{-1} . We will use this frequently below.

The proof of the following corollary was suggested by Dani and Raghunathan.

COROLLARY 4.9 *Let $G \subset B_p^\times$ denote the subgroup of elements with reduced norm 1, and let Γ be any cocompact lattice. Let $\alpha \in B_p^\times$ and set $\Gamma_\alpha = \alpha\Gamma\alpha^{-1}$. Then, if Γ and Γ_α are not commensurable, the set $\Gamma_\alpha \cdot \Gamma$ is dense in G .*

Proof. We apply Ratner's theorem to the product $G_* = G \times G$, the lattice $\Gamma_* = \Gamma_\alpha \times \Gamma \subset G_*$, the diagonal subgroup $\Delta = \{(g, g)\}$, for $g \in G$, and the point $x = 1$. Note that Δ is indeed ad-unipotent, for we have $\Delta(\mathbf{Q}_p) \cong SL_2(\mathbf{Q}_p)$, which is well-known to be generated by unipotent elements.

According to Ratner's theorem, the closure of the diagonal in $\Gamma_* \backslash G_*$ is the image of a closed subgroup H of G_* , with $\Delta \subset H$. Furthermore, $\Gamma_* \cap H$ is a lattice in H . Since $\Delta \cong G$ is the diagonal, we see that $\Gamma_* \cap \Delta$ is precisely $\Gamma_\alpha \cap \Gamma \subset G$. Under our hypotheses, the groups Γ and Γ_α are not commensurable, so that $\Gamma_\alpha \cap \Gamma$ has infinite index in at least one of Γ_α and Γ . In particular, $\Gamma_\alpha \cap \Gamma$ is not a lattice subgroup of G . It follows that H contains the diagonal with infinite index. Let $(h_1, h_2) \in H$ be such that $h_1/h_2 \neq \pm 1$. Then, since H contains the diagonal, we find that H contains an element $y = (h, 1)$ where $h \neq \pm 1$. Let $H_1 \subset H$ be the set of elements of the form $(h, 1)$. Then H_1 is normal in H . Conjugating by the diagonal, we see that the set $\{h \in G : (h, 1) \in H\}$ is an infinite normal subgroup of G . But $G = SL_2(\mathbf{Q}_p)$ has no nontrivial normal subgroups other than ± 1 . It follows that H contains $G \times 1$, and since H contains the diagonal, we find that $H = G_*$.

We conclude from this that the orbit of the diagonal Δ is dense in $\Gamma_* \backslash G_*$, so that the subset $(\gamma x, \gamma' x)$ is dense in $G \times G$. Considering the map $(p, q) \mapsto pq^{-1}$, of $G \times G$ to G , we see that $\Gamma_\alpha \cdot \Gamma$ is dense in G , as required. \square

REMARK 4.10 The proof above shows that the diagonal orbit Δ is dense in $\Gamma_* \backslash G_*$, for $G \cong PSL_2(\mathbf{Q}_p)$, and that the set $\Gamma_* \cdot \Delta$ is dense in $G \times G$. For later use, we will need a slight generalization, when PSL_2 is replaced by PGL_2 . Thus let \tilde{G} denote $B_p^\times / \mathbf{Q}_p^\times = \tilde{G} = PGL_2(\mathbf{Q}_p)$, and let $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ denote cocompact lattices in \tilde{G} . As before, we put $\tilde{\Gamma}_* = \tilde{\Gamma} \times \tilde{\Gamma}'$, and $\tilde{G}_* = \tilde{G} \times \tilde{G}$. Then we want to determine the closure of the orbit of the diagonal $\tilde{\Delta} \cong \tilde{G}$ in $\tilde{\Gamma}_* \backslash \tilde{G}_*$.

If $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ are commensurable, then $\tilde{\Delta} \cap \tilde{\Gamma}_* \cong \tilde{\Gamma} \cap \tilde{\Gamma}'$ is a lattice in $\tilde{\Delta}$, so that the diagonal orbit is closed. Suppose on the other hand that $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ fail to be commensurable. Let Γ and Γ' denote the intersections of $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ with $G = PSL_2$. Then Γ and Γ' are cocompact lattices in G , which are not commensurable. Thus if $\Delta \subset \tilde{\Delta}$ denotes the diagonal PSL_2 , the corollary above shows that the closure of $\tilde{\Gamma}_* \cdot \Delta$ contains $PSL_2 \times PSL_2$. Let $\tilde{H} \subset \tilde{G}_*$ denote the product group $\tilde{\Delta} \cdot (PSL_2 \times PSL_2)$. Then clearly $\tilde{\Delta} \subset \tilde{H}$, and \tilde{H} is contained in the closure of $\Gamma_* \cdot \tilde{\Delta}$. Since $\tilde{\Gamma}_* \cap \tilde{H}$ is obviously a lattice in \tilde{H} , it follows that the closure of the diagonal orbit coincides with the orbit of \tilde{H} .

4.11 We may now complete the proof of Theorem 4.1. Let $\tilde{G} = B_p / \mathbf{Q}_p^\times \cong PGL_2(\mathbf{Q}_p)$, and let $G \cong PSL_2(\mathbf{Q}_p) \subset \tilde{G}$ denote the image of the elements of norm 1. Let the lattice Γ be as in section 3.10, and let $\Gamma_+ = \Gamma \cap G$. Let τ_p denote the element of B_p^\times giving the action of τ . According to Proposition 3.7, τ_p is not in $\mathbf{Q}_p^\times B^\times$, which implies that Γ and $\tau_p^{-1} \Gamma \tau_p$ are not commensurable. Since Γ and Γ_+ are commensurable, we find that Γ_+ and $\Gamma'_+ = \tau_p^{-1} \Gamma_+ \tau_p = \tau_p^{-1} \Gamma \tau_p \cap G$ are not commensurable either. Applying the corollary above, we see that the set $\Gamma'_+ \Gamma_+$ is dense in G . If v is any vertex on the tree, it follows that the set of vertices of the form $\Gamma_+ \tau_p \Gamma_+ v$ is either equal to Gv , or to $\tau_p Gv$, depending on whether the image of τ_p in \tilde{G} is an element of G or not. In the first case, we get all those vertices that are at an even distance from v , and in the second case we get the vertices whose distance is odd.

We fix a base point v_0 on the tree, as in §3, corresponding to $P = (f, R)$ of level 1. We want to find a vertex v such that v and $\tau_p v$ represent Heegner points (f_i, R_i) , where the R_i are as in the statement of the theorem. Let the finite graph $\mathcal{G} = \Gamma \backslash T$ be as in section 3.10. Let $(x, y) \in \mathcal{G} \times \mathcal{G}$ be any pair of vertices. It then suffices to find a vertex $v \in T$ such that the image of $(\tau_p v, v) \in (\Gamma \times \Gamma) \backslash (T \times T) = \mathcal{G} \times \mathcal{G}$ is the given pair (x, y) .

Let (w, v) denote any lift of (x, y) to $T \times T$. It follows from our choice of the congruence group $\Gamma(M)$ in section 3.10, that the graph \mathcal{G} is not bipartite. We may assume therefore that the vertex v has any desired parity, and further, that the distance from v to w has any desired parity. Choosing the parities suitably, the above remarks imply that we can find $\gamma, \gamma' \in \Gamma_+$ such that $w = \gamma \tau_p \gamma' v$. Since

$\tau_p \gamma' v = \gamma^{-1} w$, we find that $(\tau_p v', v')$ has the required image in $G \times G$, where $v' = \gamma' v$. Note here that $\gamma' \in \Gamma_+$ lies in PSL_2 , and so does not disturb the parity conditions.

4.12 For the proof of Proposition 2.14, we need a more quantitative version of Ratner's theorem, namely the uniform distribution of certain one-parameter flows. To describe this, we proceed as follows. Let G denote a p -adic Lie group, and let $U = u(t) : \mathbf{Q}_p \rightarrow G$ be a one-parameter subgroup. If Γ denotes a lattice in G , and $x \in \Gamma \setminus G$, then Theorem 4.6 applies to U and x , and shows that the closure of xU in $\Gamma \setminus G$ is homogenous. Let H denote a closed subgroup such that $U \subset H$ and such that the closure X of xU is the orbit xH of H . Then $\Gamma \cap xHx^{-1}$ is a lattice in xHx^{-1} , and there is a unique H -invariant Borel measure μ on supported on X , normalized so that $\mu(X) = 1$. Ratner has shown that xU is uniformly distributed in X in the following precise sense. Let λ denote an additive Haar measure on \mathbf{Q}_p , and for $s \in \mathbf{R}$, let $F(s) \subset \mathbf{Q}_p$ denote the subset $\{x \in \mathbf{Q}_p : |x| \leq s\}$.

THEOREM 4.13 ([RAT95], THEOREM 3) *Let ϕ denote any bounded continuous function on X . Then we have*

$$\lim_{s \rightarrow \infty} \frac{1}{\lambda(F(s))} \int_{F(s)} \phi(xu(t)) d\lambda(t) = \int_X \phi(x) d\mu(x). \quad (23)$$

4.14 Let G be any p -adic Lie group and U a one-parameter subgroup as above. Let \mathfrak{G} denote the Lie algebra of G , and let $u \in \mathfrak{G}$ denote a nonzero vector tangent to U . Then u is nilpotent. We say that $a \in \mathfrak{G}$ is a *diagonal* for u if there is a nilpotent $u^* \in \mathfrak{G}$ such that

$$\text{ad}_{u^*}(u) = a, \quad \text{ad}_a(u) = -2u, \quad \text{ad}_a(u^*) = 2u^*.$$

The element u^* is called an *opposite* for u . Then u and u^* generate a Lie subalgebra $sl_2(u, u^*)$ of \mathfrak{G} isomorphic to $sl_2(\mathbf{Q}_p)$. The terminology is motivated by the example of $G = SL_2(\mathbf{Q}_p)$, when $\mathfrak{G} = sl_2$ consists of the matrices of trace zero. In this case, U can be taken to be the upper triangular unipotent subgroup, and the elements a, u^* , and u may be taken to be the matrices

$$a = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad u^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and } u = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Now let U, u, u^*, a be as above, and let $A \subset G$ denote a one-dimensional Lie subgroup of G normalizing U , whose tangent space is spanned by a . Then we say that a subgroup $A \subset G$ is a diagonal for U if there exists a one-parameter ad-unipotent subgroup U^* normalized by A which is tangent to u^* and such that $A \subset S$, where S is the subgroup of G generated by U and U^* . We require further that Ad_G maps A homomorphically onto the multiplicative 1-parameter subgroup of $\text{Ad}_G(S)$ with tangent vector a . For a discussion of this definition, we refer the reader to [Rat95], page 278. The existence and properties of such diagonal subgroups is a key ingredient in all of Ratner's theory. In the

example of $G = SL_2$, the opposite U^* is the lower triangular unipotent subgroup, and the diagonal A is the subgroup of diagonal matrices of determinant 1.

We would like to use these notions in the following very simple situation. Let G denote a p -adic Lie group, and $\Gamma \subset G$ a lattice. We suppose that there is a closed subgroup $S \subset G$ such that $S \cong PSL_2(\mathbf{Q}_p)$, and we fix such an identification. In particular, any unipotent subgroup U of PSL_2 gives a one-parameter subgroup of G . Then it is obvious that such a U admits a diagonal and an opposite. Namely, if we conjugate so that U becomes the (image of) the standard upper triangular unipotent subgroup, then A is the diagonal subgroup and U^* is the lower triangular subgroup, as above.

With these assumptions, let $x \in Y = \Gamma \backslash G$, and let X denote the closure of xU . Then X is the orbit of some closed $H \subset G$. Let μ be the normalized H -invariant measure on X , as above. Then Ratner has proved the following theorem.

THEOREM 4.15 ([RAT95], THEOREM 6, COROLLARY 4) *Suppose that the lattice $\Gamma \subset G$ is cocompact, and that the one parameter group U satisfies $U \subset S \subset G$, where $S \cong PSL_2(\mathbf{Q}_p)$. Then the measure μ is preserved by cSc^{-1} for some $c \in G$ such that c centralizes U .*

We note a simple consequence of this theorem. The measure μ is supported on the orbit of H . The fact that $S' = cSc^{-1}$ preserves the measure μ means that we may assume $S' \subset H$. We will use this observation below.

5 PROOF OF THE MAIN RESULTS

We may now complete the proof of Proposition 2.14. For the convenience of the reader, we repeat the statement here. Recall the definitions of the numbers h , c , and ψ : we have

$$h = \sum_{[R] \in \text{Cl}(B)} 1/w_R,$$

where w_R is the order of R^\times , and

$$|\psi|^2 = \frac{1}{h} \sum \psi([R])^2 / w_R = \frac{1}{h} \sum |\psi([R])|^2 / w_R.$$

We put $E_p = \left(a_p^2 \left(\frac{p+1}{p-1} \right) + 1 \right)$, and define $E_D = \prod_{q_i | D} (1 + \chi_t(\text{Frob}(q_i)) a_{q_i} / (q_i + 1))$. Finally, $c = e_n / p^{n-\delta-1}$, where e_n is the order of G_n .

PROPOSITION *The following limit formulae hold:*

$$\lim_{n \rightarrow \infty} \frac{1}{cp^{n-\delta-1}} \sum_{P \in X_n} \sum_{\tau \in G_1} \chi_t(\tau) \psi(P^\tau) \psi(P) = E_D \cdot |\psi|^2, \quad (24)$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{c p^{n-\delta-2}} \sum_{Q \in X_{n-1}} \sum_{\tau \in G_1} \chi_t(\tau) \left(a_p \psi(Q^\tau) - \psi(\check{Q}^\tau) \right) \left(a_p \psi(Q) - \psi(\check{Q}) \right) = E_p \cdot E_D \cdot |\psi|^2. \quad (25)$$

We will see below that the limits on the right-hand-side come from the terms on the left with $\tau \in G_0$, where $G_0 \subset G_1$ is the genus subgroup (the main terms) and that the terms for $\tau \notin G_0$ all tend to the value zero.

5.1 We begin by considering the terms coming from $\tau \in G_0$, in the first case of the proposition. In this case, we are required to calculate the asymptotic distribution of the vertices $(v, \tau_p v)$ in the product graph $\mathcal{G} \times \mathcal{G}$. Equivalently, we want the distribution of (v, v) in $\mathcal{G} \times \mathcal{G}_\tau$, where $\mathcal{G}_\tau = \Gamma_\tau \backslash \mathcal{T}$, and $\Gamma_\tau = \tau_p \Gamma \tau_p^{-1}$. Note that the map $v \mapsto (v, v) \mapsto \mathcal{G} \times \mathcal{G}_\tau$ factors through a finite quotient of the tree \mathcal{T} precisely when the groups Γ and Γ_τ are commensurable, which is the case if and only if $\tau \in G_0$. One can therefore compute the average when $\tau \in G_0$ by combinatorial means, as in the proof of Theorem 1.5 in the previous section. But it will be convenient to show how this may be deduced from Ratner's theorems instead, as this kind of argument is valid even when $\tau \notin G_0$, and also allows one to treat both limits of the proposition in a uniform manner. The reader will notice that the arguments below show that the main results of this paper may be proven without any recourse to the theory of finite graphs, by using Ratner's theorem and working with the tree instead. We have chosen to retain the graph theoretic arguments because they are natural in this context, and illustrate a special case of Ratner's theorem.

We consider therefore the following general situation. Let ϕ denote a bounded, continuous, real-valued function on $\tilde{G} = PGL_2(\mathbf{Q}_p) = B_p^\times / \mathbf{Q}_p^\times$, left-invariant under some discrete cocompact subgroup Γ , and right-invariant under some open compact subgroup \tilde{V} . Let $P = (f, R)$ denote a Heegner point of level 1, and let $P_1 = (f, R')$ denote a good Heegner point of level p , so that the predecessor of P_1 is P . Then, \tilde{V} will be one of the following:

- the full maximal compact $\tilde{K} = R_p^\times / \mathbf{Z}_p^\times$ stabilizing the vertex of \mathcal{T} determined by R_p , or
- the stabilizer of the oriented edge of the tree $\mathcal{T} = \tilde{G} / \tilde{K}$ going from P to P_1 .

These choices of \tilde{V} are motivated by the functions $\psi(P)$ and $a_p \psi(Q) - \psi(\check{Q})$ appearing in the proposition above, since ψ is a function of the vertices of \mathcal{T} , while $a_p \psi(Q) - \psi(\check{Q})$ is a function on the oriented edges of \mathcal{T} .

Now given any good Heegner point $P' = (f, R')$ of level p^n , for $n \geq 1$, we get a coset in \tilde{G} / \tilde{V} as follows. If \tilde{V} is the maximal compact \tilde{K} , then $\tilde{G} / \tilde{V} = \mathcal{T}$, and the coset determined by P' is simply the

corresponding vertex of T . If \tilde{V} is the stabilizer of an oriented edge, then the coset determined by P' is simply the edge whose origin is P' and whose terminus is the predecessor \check{P}' . In this way, we may view the function ϕ as being defined on the good Heegner points by putting $\phi(P') = \phi(g\tilde{V})$, where $g\tilde{V}$ is the coset determined by P' . (Recall that ϕ is right-invariant under \tilde{V} .)

Let $\tau_p \in \tilde{G}$ be given, and consider the function ζ on $\tilde{G} \times \tilde{G}$ defined by the formula $\zeta(g_1, g_2) = \phi(g_1)\phi(\tau_p g_2)$. Then ζ becomes a function on Heegner points by putting $\zeta(P') = \zeta(g, g)$, where g is a representative of the coset gV determined by P' . Our task in this situation is to compute the limit, as n tends to infinity, of the average $\frac{1}{p^{n-1}} \sum_{P \in X'_n} \zeta(P)$, where X'_n denotes the set of good Heegner points of level p^n . Note that it is not clear at the outset that this limit even exists.

To state the result, put $\Gamma' = \tau_p^{-1}\Gamma\tau_p$, so that ζ is a function on $\Gamma \times \Gamma' \backslash \tilde{G} \times \tilde{G}$. Let \tilde{X} denote the closure of the diagonal orbit $\tilde{\Delta}$ in $\Gamma_* \backslash \tilde{G}_*$, where $\Gamma_* = \Gamma \times \Gamma'$ and $\tilde{G}_* = \tilde{G} \times \tilde{G}$. Ratner's theorem (Theorem 4.6 above) states that \tilde{X} is the orbit of some group \tilde{H} , with $\tilde{\Delta} \subset \tilde{H} \subset \tilde{G}_*$. Since the orbit of \tilde{H} is closed, the subgroup $\Gamma_{\tilde{H}} = \Gamma_* \cap \tilde{H}$ is a lattice in \tilde{H} , and $\tilde{X} = \Gamma_{\tilde{H}} \backslash \tilde{H}$ carries the right Haar measure μ on \tilde{H} , normalized so that $\mu(\tilde{X}) = 1$. Then the key result we will use is the following

PROPOSITION 5.2 *Suppose that the graph $\Gamma \backslash T = \Gamma \backslash \tilde{G} / \tilde{K}$ is not bipartite. Then the following limit formula holds:*

$$\lim_{n \rightarrow \infty} \frac{1}{\delta_0 p^{n-\delta-1}} \sum_{P \in X'_n} \zeta(P) = \lim_{n \rightarrow \infty} \frac{1}{cp^{n-\delta-1}} \sum_{P \in X_n} \zeta(P) = \int_{\tilde{X}} \zeta(x) d\mu(x) = \int_{\Gamma_{\tilde{H}} \backslash \tilde{H}} \zeta(h) d\mu(h). \quad (26)$$

Here we remind the reader that X_n is the set of Heegner points of level $p^{n-\delta}$, and that the number c is defined so that $cp^{n-\delta-1} = \#X_n$, while X'_n denotes the set of good Heegner points at distance $n - \delta$ from the origin, and $\delta_0 p^{n-\delta-1}$ is the cardinality of X'_n . Thus the proposition states that the average of ζ over Heegner points of level p^n converges to the average of ζ over the closure of the diagonal orbit.

Let us admit for the moment the validity of this proposition, and show how to complete the proofs of our main results. The main point is to compute explicitly the integrals on the right of (26). Here there are two cases, depending on whether Γ and Γ' are commensurable or not. Indeed, the subgroup \tilde{H} of \tilde{G}_* satisfies $\tilde{\Delta} \subset \tilde{H} \subset \tilde{G}_*$, and by Remark 4.10, we must have either $\tilde{H} = \tilde{\Delta}$ or $\tilde{H} = \tilde{\Delta} \cdot (PSL_2 \times PSL_2)$. The first possibility occurs if and only if Γ and Γ' are commensurable.

Now we specialize to the cases of interest, where τ_p is the element of B_p^\times arising from $\tau \in G_1$ (see §3), the group Γ comes from a suitable congruence subgroup of $R[1/p]^\times$, and $\Gamma' = \Gamma_\tau = \tau_p^{-1}\Gamma\tau_p$.

We begin with the case where $\tau \in G_0$. We have $\tau \in G_0 \iff \tilde{H} = \tilde{\Delta}$, and in this case $\Gamma_{\tilde{H}} = \Gamma \cap \Gamma_\tau$ is a cocompact lattice in $\tilde{H} = \tilde{\Delta} \cong \tilde{G}$. Write $\Gamma(\tau)$ for $\Gamma_{\tilde{H}} = \Gamma \cap \Gamma_\tau$, to emphasize the dependence on τ . Let S_τ denote the finite set $\Gamma(\tau) \backslash \tilde{H} / \tilde{V}$, and let $s_\tau = \#S_\tau$; then since ζ is right- \tilde{V} -

invariant, the right hand side of (26) is simply

$$\int_{\Gamma_{\tilde{H}} \backslash \tilde{H}} \zeta(h) d\mu(h) = \frac{1}{s_\tau} \sum_{x \in S_\tau} \zeta(x). \quad (27)$$

We can make this more concrete, for the \tilde{V} and ζ at hand. To treat equation (24), we take \tilde{V} to be the maximal compact, and define $\zeta(g_1, g_2) = \psi(g_1)\psi(\tau_p g_2)$, where ψ is our original function on $G = \Gamma \backslash \mathcal{T} = \Gamma \backslash \tilde{G} / \tilde{V}$. To treat equation (25), we take \tilde{V} to be the stabilizer of an oriented edge of \mathcal{T} . In this case, let ϕ denote the function on oriented edges of \mathcal{T} defined by $\phi(\vec{e}) = a_p \psi(v) - \psi(w)$, where v, w are the origin and terminus of \vec{e} respectively. Identifying the set of oriented edges with the coset space \tilde{G} / \tilde{V} , we may view ϕ as a right- \tilde{V} -invariant function on \tilde{G} . Then we put $\zeta(g_1, g_2) = \phi(g_1)\phi(\tau_p g_2)$. Note that the hypothesis in Proposition 5.2 is satisfied, since by our choices the graph $G = \Gamma \backslash \mathcal{T}$ is not bipartite.

Let A_1 denote the right-hand side of (27), when $\tau = 1$, and let $S_1 = \Gamma \backslash \tilde{G} / \tilde{V}$. Let $\tau \in G_0$ be given. Let $\tau = \text{Frob}(q_1)\text{Frob}(q_2) \dots \text{Frob}(q_r)$, with $q_i | D$, as in section 3.8, and put $d = q_1 q_2 \dots q_r$. Then we claim that S_τ has cardinality $s_\tau = s_1 t_d$, where t_d is the degree of the d -th Hecke correspondence T_d , and that

$$A_d = \frac{1}{s_\tau} \sum_{x \in S_\tau} \zeta(x) = \frac{1}{s_1 t_d} \sum_{x \in S_\tau} \zeta(x) = \frac{a_d}{s_1 t_d} \sum_{x \in S_1} \zeta(x) = \frac{a_d}{t_d} A_1, \quad (28)$$

where a_d is the eigenvalue of T_d acting on the function ψ . To see this, it is most convenient to work adelicly. Recall therefore that Γ is a suitable congruence subgroup of $R[1/p]^\times$, of level M . We have assumed that M is relatively prime to NDp . Define the open compact subgroup V of \hat{B}^\times by saying $V = \prod V_\ell \subset \hat{R}^\times$, where the local component V_ℓ consists of elements in R_ℓ^\times that are congruent to 1 modulo M , at primes $\ell \neq p$, and setting $V_p \subset R_p$ to be the inverse image under the natural projection of $\tilde{V} \subset \tilde{K} = R_p^\times / \mathbf{Z}_p^\times$. Let $J = B^\times \cdot \hat{\mathbf{Q}}^\times \cdot B_p^\times \cdot V$. Then J is a normal subgroup of finite index in \hat{B}^\times (by strong approximation) and $S_1 = \Gamma \backslash \tilde{G} / \tilde{V} = B^\times \backslash J / V$. Note that V is a congruence subgroup, and that S_1 is endowed with a family of Hecke correspondences T_n , in the usual way. (See section 2.3 above, or [BD96], §2. A detailed discussion may also be found in §2 of [DT94]. We will recall the definition below.)

To relate S_τ and S_1 , we need to unwind the definitions from §3 above. Let $d = q_1 q_2 \dots q_r$ denote the factorization of d into distinct primes as above. Then each q_i is ramified in K , and we write π_i for a uniformizer at q_i . Now let $\tilde{\tau}$ denote the idele of the field K with component π_i at q_i , for $1 \leq i \leq r$, and with component 1 at other primes. Let $P = (f, R)$ be our fixed Heegner point of level 1. Then, according to Proposition 3.6, the local element $\tau_p \in \tilde{G}$ corresponding to τ is given by $\tau_p = b_p f(t_p)$, where t_p is the local component of $\tilde{\tau}$ at p , and b_p is the local component of some $b \in B^\times$, which is chosen such that the local component of $bf(\tilde{\tau})$ lies in $\mathbf{Q}_\ell^\times \hat{R}_\ell^\times$, for all primes $\ell \neq p$. For $\tau \in G_0$,

the explicit form of $\tilde{\tau}$ given above shows that the local component t_p is trivial, so that $\tau_p = b_p$ is an element of $B^\times \subset B_p$. Recalling the notation $\Gamma_\tau = \tau_p^{-1} \Gamma \tau_p$, we find from strong approximation that $\Gamma_\tau \backslash \tilde{G} / V_p = B^\times \backslash J / \hat{\mathbf{Q}}^\times V_\tau$, where V_τ is the open compact subgroup of \hat{B} given by $(V_\tau)_\ell = b_\ell^{-1} V_\ell b_\ell$, for $\ell \neq p$, and $(V_\tau)_p = V_p$, where $V = \prod V_\ell$ is the subgroup defined previously. By definition of the element b , we find that $V_\ell = (V_\tau)_\ell$ unless ℓ is one of the primes q_i appearing in the factorization of d . If $\ell = q_i$, then we have $(V_\tau)_{q_i} = f(\tilde{\pi}_i) V_{q_i} f(\tilde{\pi}_i)^{-1}$, where $\tilde{\pi}_i$ is the idele of K with component equal to π_i at q_i , and with component 1 at other places. Since q_i is ramified in K , we find that $f(\pi_i)$ has reduced norm q_i . Since the level M of V was assumed prime to D , the local component V_{q_i} is a maximal compact, and the double coset $V_{q_i} f(\tilde{\pi}_i) V_{q_i}$ is the one defining the usual Hecke operator T_{q_i} . Let $V(\tau) = V \cap V_\tau \subset V$. It follows that $S_\tau = B^\times \backslash J / \hat{\mathbf{Q}}^\times V(\tau)$. The inclusion of $V(\tau) \subset V$ induces a projection $S_\tau \rightarrow S_1$, and for given $x \in S_1$, the fiber above x is by definition the Hecke correspondence $T_d(x)$. The claim of equation (28) now follows readily, since the functions ψ and ϕ defining ζ are eigenfunctions for T_d with eigenvector a_d .

It follows that the terms for $\tau \in G_0$ in equations (24) and (25) each contribute a factor of $\chi_t(\tau) \frac{a_d}{t_d} A_1$ to the right hand side. Together, these give

$$\sum_{\tau \in G_0} \frac{a_d}{t_d} \chi_t(\tau) A_1 = A_1 \cdot \left(\sum_{\tau \in G_0} \frac{a_d}{t_d} \chi_t(\tau) \right) = A_1 \cdot \prod \left(1 + \frac{a_{q_i}}{t_{q_i}} \chi_t(\text{Frob}(q_i)) \right),$$

where the last equality follows from the fact that the primes q_i are distinct, so that the Hecke operators are multiplicative.

To complete our analysis of these terms, we need to compute the numbers A_1 in each of the two cases. In the situation of (24), we have already achieved this by combinatorial means above: we have

$$A_1 = |\psi|^2.$$

In the situation of (25), the set S_1 is simply the set of oriented edges of the finite graph G , and we have to compute the average value of ζ^2 , over all the edges \vec{e} of G :

$$\begin{aligned} \sum_{\vec{e}} (a_p \psi(x) - \psi(y))^2 &= \sum_{\vec{e}} \left(a_p^2 \psi(x)^2 + \psi(y)^2 - 2a_p \psi(x) \psi(y) \right) \\ &= a_p^2 (p+1) \sum_{x \in \mathcal{G}} \psi(x)^2 + (p+1) \sum_{y \in \mathcal{G}} \psi(y)^2 - 2a_p \sum_{\vec{e}} \psi(x) \psi(y) \\ &= a_p^2 (p+1) |\psi|^2 + (p+1) |\psi|^2 - 2a_p^2 |\psi|^2 = |\psi|^2 (p+1 + a_p^2 (p-1)). \end{aligned} \quad (29)$$

Since there are $m(p+1)$ oriented edges on G , where m is the number of vertices, we get

$$A_1 = |\psi|^2 \cdot E_p,$$

where E_p is as in the statement of the proposition. This completes our analysis of the terms with $\tau \in G_0$ in (24) and (25): we have shown that the contribution of these terms is exactly the quantity on the right-hand side.

It now remains to analyze the terms with $\tau \notin G_0$, and to show that these vanish in the limit. In this case the groups Γ and Γ_τ are not commensurable, and the closure of the diagonal orbit $\tilde{\Delta}$ is $\Gamma_H \backslash \tilde{H}$, where \tilde{H} is the subgroup $\tilde{H} \subset \tilde{G}_*$ given by $\tilde{H} = \tilde{\Delta} \cdot (PSL_2 \times PSL_2)$, and $\Gamma_{\tilde{H}} = \Gamma_* \cap \tilde{H}$. By hypothesis, the function ϕ in the definition of ζ is left-invariant under some $\gamma \in \Gamma$ such that γ interchanges the odd and even vertices on the tree \mathcal{T} . Then ζ is left-invariant under the element $(\gamma, 1)$, and right invariant under the compact subgroup $\tilde{V}_* = \tilde{V} \times \tilde{V}$. Using the fact that \tilde{V} is the stabilizer of either a vertex or an edge, one checks that the coset spaces $\tilde{G}/(PSL_2(\mathbf{Q}_p) \cdot \tilde{V})$ and $\tilde{G}_*/(\tilde{H} \cdot \tilde{V}_*)$ each has two elements, and that the nontrivial coset in the latter is represented by γ_* . Since ζ is left-invariant under γ_* and right invariant under \tilde{V}_* , it follows then that the average of ζ over the orbits \tilde{X} of \tilde{H} and the full $\Gamma_* \backslash \tilde{G}_*$ are equal. Thus, to prove that the integral over \tilde{X} is zero, it suffices to compute the integral over $\Gamma_* \backslash \tilde{G}_*$.

But this latter is easy. In the situation of (24), the function ζ is right-invariant by the maximal compact $\tilde{K} = PGL_2(\mathbf{Z}_p)$, and $\Gamma_* \backslash \tilde{G}_*/\tilde{K}_* = \mathcal{G} \times \mathcal{G}_\tau$, where $\tilde{K}_* = \tilde{K} \times \tilde{K}$. It follows that

$$\int_{\Gamma_* \backslash \tilde{G}_*} \zeta(x) d\mu(x) = \sum_{x,y} \psi(x) \psi(y),$$

where the sum is taken over pairs of vertices $(x, y) \in \mathcal{G} \times \mathcal{G}_\tau$. But clearly

$$\sum_{x,y} \psi(x) \psi(y) = \left(\sum_x \psi(x) \right) \left(\sum_y \psi(y) \right) = 0,$$

since ψ is orthogonal to the constant functions.

As for (25), ζ is right-invariant under the stabilizer \tilde{V} of an oriented edge. It follows that $\Gamma_* \backslash \tilde{G}_*/\tilde{V}_*$ consists of pairs (\vec{e}_1, \vec{e}_2) of oriented edges in \mathcal{G} and \mathcal{G}_τ . Then one simply has to show that $\sum_{\vec{e}} \zeta(\vec{e}) = 0$, where the sum is taken over all oriented edges of \mathcal{G} . But we have $\sum_{\vec{e}} \zeta(\vec{e}) = (p+1) \sum_{x \in \mathcal{G}} a_p \psi(x) + (p+1) \sum_{y \in \mathcal{G}} \psi(y) = 0$, by Lemma 3.11.

5.3 Thus we have reduced everything to proving Proposition 5.2. There are two possibilities for the function ζ on the left: we have $\zeta(P) = \phi(P)\phi(P^\tau)$, where either $\phi(P) = \psi(P)$, or $\phi(P) = a_p \psi(P) - \psi(\check{P})$. In the former case ζ arises from a function on the vertices of the tree, and in the latter from a function on the oriented edges. The argument is the same in each case, and we will endeavor in the sequel to keep the notation uniform. Thus let us consider the tree \mathcal{T} , together with a fixed base point given by a Heegner point $P = (f, R)$ of level 1. Let X'_n denote the set of good vertices

at distance $n - \delta$, corresponding to Heegner points of level $n - \delta$. Let $e'_n = \delta_0 p^{n-\delta-1}$ denote the cardinality of X'_n ; it is then enough to show that

$$\lim_{n \rightarrow \infty} \frac{1}{e'_n} \sum_{v \in X'_n} \zeta(P(v)) = A, \quad (30)$$

where we have written A to denote the right-most term of (26), and $P(v)$ is the Heegner point corresponding to v . Now, for each vertex v at distance n from the origin, we write \check{v} to denote the predecessor of v . Namely, \check{v} is the unique vertex at distance 1 from v and distance $n - 1$ from the origin. If v corresponds to the Heegner point P' , then \check{v} corresponds to the predecessor \check{P}' . We let $\vec{e}(v)$ denote the oriented edge with origin v and terminus \check{v} . We say that $\vec{e}(v)$ is at distance n if the origin v is at distance n from the origin, and we say that $\vec{e}(v)$ is a good edge if the origin v is a good vertex. Note that the definition of $\vec{e}(v)$ depends only on v and the choice of an origin on the tree.

Let us consider an arbitrary origin v_0 for T , and let us write Z_n for the set of objects (vertices or edges) at distance n from this v_0 . Thus if ζ comes from a function of vertices, we take Z_n to be the set of vertices at distance n , and if ζ comes from a function of edges, we take Z_n to be the set of oriented edges with origin at distance n and terminus at distance $n - 1$. Then Z_n has cardinality $e_n = (p + 1)p^{n-1}$. To prove (30), it is enough to show that

$$\lim_{n \rightarrow \infty} \frac{1}{e_n} \sum_{z \in Z_n} \zeta(z) = A, \quad (31)$$

for an arbitrary origin v_0 . Then the required average over the good objects at distance n follows as in the proof of Proposition 3.16.

We will first prove (31) when the limit is taken over the set $n = 2t$ of even integers. Recall that the tree T is given by $\tilde{G}/\tilde{K} = PGL_2(\mathbf{Q}_p)/PGL_2(\mathbf{Z}_p)$, where we choose the coordinates on PGL_2 so that $\tilde{K} = PGL_2(\mathbf{Z}_p)$ corresponds to a given origin v_0 . Then consider the image in T of the 1-parameter subgroup $U : t \rightarrow u(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, and let $v(t)$ denote the vertex determined by $u(t)$. Let $\vec{e}(t)$ denote the oriented edge determined by $v(t)$ and the choice of the origin coming from $\tilde{K} = v_0$. To keep the notation uniform, we will simply write $z(t) = v(t)$ or $\vec{e}(t)$ to denote the $z \in Z_n$ determined by $u(t)$ in this manner. Then $z(t)$ is at distance $2n$ from the origin, where $n = -\text{ord}_p(t)$. Furthermore, if $\tilde{K} = PSL_2(\mathbf{Z}_p)$, then we see that $\tilde{K}z(t)$ runs through the set of all objects $z \in Z_n$ at distance $2n$, and that the stabilizer \tilde{K}_t has index given by $e_{2n} = (p + 1)p^{2n-1}$. Thus, we find that

$$\frac{1}{e_{2n}} \sum_{z \in Z_n} \zeta(z) = \int_{\tilde{K}} \zeta(ku(t)) d\mu(k),$$

where μ is a Haar measure on \tilde{K} normalized so that $\mu(\tilde{K}) = 1$. Note that for fixed k , the map $t \mapsto ku(t)$ is another 1-parameter orbit.

Now let λ denote an additive Haar measure on \mathbf{Q}_p , normalized so that $\lambda(\mathbf{Z}_p) = 1$. For $n \geq 0$, let $F_n \subset \mathbf{Q}_p$ denote the set $\{t \in \mathbf{Q}_p : \text{ord}_p(t) \geq -n\}$. Then, according to Theorem 4.13, we have

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda(F_n)} \int_{F_n} \zeta(ku(t)) d\lambda(t) = \int_{X_k} \zeta(x) d\mu_k(x),$$

where X_k is the (homogeneous) closure of kU and μ_k is the U -invariant measure as in the statement of the theorem.

We claim that $\int_{X_k} \zeta(x) d\mu_k(x) = A$, for all but countably many k . This will then imply that

$$\int_{\tilde{K}} \left(\lim_{n \rightarrow \infty} \frac{1}{\lambda(F_n)} \int_{F_n} \zeta(ku(t)) d\lambda(t) \right) d\mu(k) = A.$$

Since \tilde{K} is compact, and the averages inside the limit are bounded, we may interchange the limits and the integral to obtain

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda(F_n)} \int_{F_n} \left(\int_{\tilde{K}} \zeta(ku(t)) dk \right) d\lambda(t) = A.$$

Setting $F'_n = \{t \in \mathbf{Q}_p | \text{ord}_p(t) = -n\}$, so that the image of F'_n consists of vertices at distance $2n$, one sees easily that

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda(F'_n)} \int_{F'_n} \left(\int_{\tilde{K}} \zeta(ku(t)) dk \right) d\lambda(t) = A,$$

which implies that $\frac{1}{e_{2n}} \sum_{d(v, v_0) = 2n} \zeta(v)$ converges to A as $n \rightarrow \infty$, as required.

To compute $\int_{X_k} \zeta(x) d\mu_k(x)$, we must first determine the size of the orbit closure X_k . To do this, recall that $\tilde{X} = \Gamma_{\tilde{H}} \backslash \tilde{H}$, where \tilde{H} is either the diagonal $\tilde{\Delta}$ or $\tilde{H} = \tilde{\Delta} \cdot (PSL_2 \times PSL_2)$, depending on whether Γ and Γ' are commensurable or not. Let $H = \Delta \subset \tilde{\Delta}$ in the former case, where $\Delta \subset \tilde{\Delta}$ is the diagonally embedded PSL_2 . In the latter case let $H = PSL_2 \times PSL_2$. In either case H is a closed subgroup of \tilde{H} , and we have $kU \subset kH \subset k\tilde{H} = \tilde{H}$, since $k \in \tilde{\Delta} \subset \tilde{H}$. Note that kH is closed in $\Gamma_* \backslash \tilde{G}_*$, since $\Gamma_* \cap H$ is a lattice in H . We will show that kU is dense in the orbit $Y_k = kH$, for all but finitely many k , so that $X_k = \overline{kU} = Y_k$.

But the orbit kH is isomorphic to $\Gamma_k \backslash kHk^{-1}$, where $\Gamma_k = \Gamma_* \cap kHk^{-1}$. Since H is normal in \tilde{H} , we have $kHk^{-1} \cong H$ and $\Gamma_k \cong \Gamma_H = \Gamma_* \cap H$. Under these identifications, the invariant measure μ on H corresponds to the measure μ_k on Y_k , where μ_k is obtained from μ via the automorphism of H given by conjugation by k . Using the fact that k lies in the compact subgroup \tilde{K} , we see easily that conjugation by k preserves the invariant measure on H , so that the average of ζ on the orbit kH is the same as the average over the identity orbit $Y_1 = X$. But the average of ζ on the identity orbit X of H

is equal to the average of ζ on the orbit \tilde{X} of \tilde{H} , because (as we have remarked several times already) ζ is invariant on the left by some $\gamma \in \Gamma$, which represents the nontrivial coset in $\tilde{G}/(PSL_2 \cdot \tilde{V})$. Thus, for k such that kU is dense in Y_k , we get $X_k = Y_k$, and that $\int_{X_k} \zeta(x) d\mu_k(x) = \int_X \zeta(x) d\mu(x) = A$, for all but countably many k , as asserted.

Thus it remains to show that kU is dense in kH , for all but countably many k . Here it is convenient to separate out two cases, depending on whether the diagonal orbit is closed or not. Consider first the case where the diagonal orbit is closed, and $X = \Gamma_H \backslash \Delta$. It is a general fact that if Γ in $PSL_2(\mathbf{Q}_p)$ is a cocompact lattice, then every orbit xU of a unipotent U is dense; we may see this in the present case by applying Theorem 4.15 to the group PSL_2 . Since every U has an opposite in PSL_2 , the theorem states precisely that the measure μ supported on the closure of kU is invariant under the full PSL_2 , so that the closure of kU contains the full orbit of PSL_2 .

In the second case, we have $H = PSL_2 \times PSL_2$. Note first of all that kU is dense in the orbit of kH if and only if kUk^{-1} is dense in the orbit of H . Now Theorem 4.15 shows that the closure of $U' = kUk^{-1}$ is invariant under the action of $S(U') = \{(g, ugu^{-1}) | g \in PSL_2\}$, for some $u \in U'$. If H' is a subgroup of H such that the closure of U' coincides with the orbit of H' , then we may assume that $S(U') \subset H'$. If the containment is strict, then, writing $u_* = (1, u)$, we find that the conjugate subgroup $u_*^{-1}H'u_*$ strictly contains the diagonal. But we have already seen (in the proof of Corollary 4.9 and in Remark 4.10) that the only subgroup of $PSL_2 \times PSL_2$ which strictly contains the diagonal is $PSL_2 \times PSL_2$ itself. This implies that $u_*^{-1}H'u_* = PSL_2 \times PSL_2 = H$ and $H' = H$. This implies that the orbit of U is dense in that of H , as asserted.

But if the containment is not strict, then we have $H = S(U')$, and the orbit of $S(U')$ is closed. But then $\Gamma \times \Gamma_\tau \cap S(U')$ must be a lattice in $S(U')$, and this happens if and only if the groups Γ and $u\tau_p\Gamma(u\tau_p)^{-1}$ are commensurable. But the commensurator of Γ in B_p is $\mathbf{Q}_p^\times \cdot B^\times$, and B^\times is countable. Since u is unipotent, it has determinant 1, and so it is clear that there are only countably many u such that $u\tau_p \in \mathbf{Q}_p^\times B^\times$. Since $\tau_p \notin \mathbf{Q}_p^\times B^\times$, any such u is nontrivial, and, as two unipotent subgroups of rank 1 that intersect non-trivially are necessarily equal, it follows that this can happen for only countably many $U' = kUk^{-1}$.

In summary, we have shown that $\lim_{n \rightarrow \infty} \frac{1}{e_{2n}} \sum_{d(v, v_0)=2n} \zeta(v) = A$, where v_0 is the origin of the tree. It remains to extend this to arbitrary v_0 , and to vertices at odd distance. This is relatively straightforward. If v is a vertex at even distance from the origin, then v is represented by $x \in PSL_2(\mathbf{Q}_p)$. One can then argue precisely as above, using Theorem 4.13, and considering 1-parameter flows that originate from x . Since the function ψ is invariant on the left by $\gamma \in R[1/p]^\times$, and the latter contains an element γ which interchanges odd and even vertices, so that $\lim_{n \rightarrow \infty} \frac{1}{e_{2n}} \sum_{d(v, v_0)=2n} \zeta(v) = A$

for arbitrary $v \in T$.

It remains finally to verify that

$$\lim_{n \rightarrow \infty} \frac{1}{e_{2n+1}} \sum_{d(v, v_0)=2n+1} \zeta(v) = A$$

for arbitrary $v_0 \in T$. But this may be checked as in the proof of (3.16). Indeed, if for any vertex v we define $T_n(v) = \sum_{d(v', v)=n} \zeta(v')$, and if we list the neighbors of v_0 as v_1, \dots, v_{p+1} , then we have the recurrence

$$T_{2n+1}(v) = \sum_{i=1}^{p+1} T_{2n}(v_i) - pT_{2n-1}(v),$$

from which the required average follows immediately. This completes the proof of the proposition.

REFERENCES

- [BD90] M. Bertolini and H. Darmon, *Kolyvagin's descent and Mordell-Weil groups over ring class fields*, J. Reine Angew. Math. **412** (1990), 63–74.
- [BD96] M. Bertolini and H. Darmon, *Heegner points on Mumford-Tate curves*, Invent. Math. **126** (1996), no. 3, 413–456.
- [BD97] M. Bertolini and H. Darmon, *A rigid analytic Gross-Zagier formula and arithmetic applications*, Ann. of Math. (2) **146** (1997), no. 1, 111–147, With an appendix by Bas Edixhoven.
- [BD98] M. Bertolini and H. Darmon, *Heegner points, p -adic L -functions, and the Cerednik-Drinfeld uniformization*, Invent. Math. **131** (1998), no. 3, 453–491.
- [BD99] M. Bertolini and H. Darmon, *Euler systems and Jochnowitz congruences*, Amer. J. Math. **121** (1999), no. 2, 259–281.
- [CDS80] D. Cvetković, M. Doob, and H. Sachs, *Spectra of graphs*, Academic Press Inc., New York, 1980.
- [Cor01] C. Cornut, *Mazur's conjecture for Heegner points*, to appear in Inv. Math., 2001.
- [Dag96] H. Daghigh, *Quaternion algebras*, Ph. D. thesis, McGill University, 1996.
- [DT94] F. Diamond and R. Taylor, *Non-optimal levels of mod ℓ modular representations*, Invent. Math. **115** (1994), 435–462.

- [Fri91] J. Friedman, *On the second eigenvalue and random walks in random d -regular graphs*, *Combinatorica* **11** (1991), no. 4, 331–362.
- [God93] C. Godsil, *Algebraic combinatorics*, Chapman & Hall, 1993.
- [Gre85] R. Greenberg, *On the critical values of Hecke L -functions for imaginary quadratic fields*, *Invent. Math.* **79** (1985), no. 1, 79–94.
- [Gro87] B. Gross, *Heights and the special values of L -series*, *Number Theory* (H. Kisilevsky and J. Labute, eds.), CMS Conference Proceedings, vol. 7, Amer. Math. Soc., 1987, pp. 115–189.
- [LR97] W. Luo and D. Ramakrishnan, *Determination of modular forms by twists of critical L -values*, *Invent. Math.* **130** (1997), 371–398.
- [Lub94] A. Lubotzky, *Discrete groups, expanding graphs and invariant measures*, Birkhäuser Verlag, Basel, 1994, With an appendix by J. Rogawski.
- [Maz84] B. Mazur, *Modular curves and arithmetic*, *Proceedings of the International Congress of Mathematicians, Vol. 1, 2* (Warsaw, 1983), PWN, 1984, pp. 185–211.
- [Rag72] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, New York, 1972, *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68*.
- [Rat95] M. Ratner, *Raghunathan’s conjectures for Cartesian products of real and p -adic Lie groups*, *Duke Math. J.* **77** (1995), no. 2, 275–382.
- [Roh84] D. E. Rohrlich, *On L -functions of elliptic curves and anticyclotomic towers*, *Invent. Math.* **75** (1984), no. 3, 383–408.
- [Rub91] K. Rubin, *The “main conjectures” of Iwasawa theory for imaginary quadratic fields*, *Invent. Math.* **103** (1991), no. 1, 25–68.
- [Vat] V. Vatsal, *Special values of anticyclotomic L -functions*, submitted to *Duke Math J.*
- [Vig80] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Springer Lecture Notes, vol. 800, Springer-Verlag, 1980.
- [Zhao01] S. Zhang, *Gross-Zagier formula for GL_2* , preprint, 2001.