Upper bounds for discriminants

Notes by F.Beukers

October 17, 2005

1 Introduction

Let p be a prime and

$$\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\overline{\mathbb{F}}_p)$$

a continuous, irreducible Galois representation unramified outside of p.
Then Moon and Taguchi show in [MT] that for the following pairs k, p no such representation exists:

$2 \leq p \leq 19$ and $k = 2, 3, 5, 7$

$2 \leq p \leq 7$ and $k = 4$.

For the following pairs k, p at most finitely many such representations exist:

$k = 3, 5$ and $p = 23, 29, 31$

$k = 7$ and $p = 23, 29$

Under the assumption of GRH we can find additional pairs. In 1973 (published in [T,1994]) Tate showed that no such representations with $p = 2$ exist and Serre, in the 1970's showed this for $p = 3$. Under assumption of GRH Brueggeman showed that no such representations for $p = 5$ exist.

2 Generalities

Let

$$\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\overline{\mathbb{F}}_p)$$

be a continuous Galois representation, possible reducible and possibly ramified at other primes than p.
Let K be the invariant field of the kernel op ρ. Let $\mathcal{P}$ be a prime in K over p and D_p its decomposition group. Denote the completion of K with respect to $\mathcal{P}$ again by K. The we get the faithful representation

$$\rho : D_p = Gal(K/\mathbb{Q}_p) \to GL_2(\overline{\mathbb{F}}_p).$$

Let K_0 be the maximal unramified extension of $\mathbb{Q}_p$ in K. Let K_1 be the maximal tamely ramified extension of $\mathbb{Q}_p$ in K. Hence

$$\mathbb{Q}_p \subset K_0 \subset K_1 \subset K.$$

We have the following Galois groups:

$I = Gal(K/K_0)$, the inertia group

$I_w = Gal(K/K_1)$, the wild ramification group

$I_t = I/I_w$, the tame ramification group.

We let $\chi : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{F}_p^*$ be the cyclotomic character defined by $\sigma(\zeta_p) = \zeta_p^{\chi(\sigma)}$ for any p-th root of unity ζ_p.
We say that ρ is finite at p if the extension K/K_1 can be generated by p-th roots of units in K_1.

Lemma 2.1 *The field $\mathbb{Q}_p(\zeta_p)$ contains an element π such that $\pi^{p-1} = -p$ and such that $\zeta_p - 1 \equiv \pi (\text{mod } \pi^2)$. Moreover, the Galois group $G_{\mathbb{Q}_p}$ acts on π via $\sigma : \pi (\text{mod } \pi^2) \mapsto \chi(\sigma)\pi (\text{mod } \pi^2)$.*

Lemma 2.2 *Suppose that I_w is non-trivial. Then*

1. *There exists a divisor d of $p-1$ such that $K_1 = K_0(\pi^d)$ where π is as in Lemma 2.1. The number $e = (p-1)/d$ is the ramification index of K_1/K_0.*

2. *The restriction of ρ to I has the form $\begin{pmatrix} \chi^b & * \\ 0 & \chi^a \end{pmatrix}$ where a, b are integers such that $\gcd(a, b, p-1) = d$.*

3. *The matrices in $\rho(I_w)$ are characterised by the shape $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$.*

Proof. The group $\rho(I_w)$ is a subgroup consisting of elements of order p^r for some r. It is an exercise to show that such a group is conjugate to a group of the form

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

Since the order of D_p/I_w is relatively prime to p, the elements of $\rho(I_w)$ are the only ones within this group.
Since D_p is a normaliser of the non-trivial group I_w, the restriction of ρ to D_p has the form $\begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix}$. Here χ_1 and χ_2 are characters on D_p. The semisimplification of $\rho : D_p \to GL_2(\overline{\mathbb{F}}_p)$ consists of the direct sum of χ_1 and χ_2. Its kernel is I_w and thus we see that D_p/I_w is abelian.

In particular, $K_1/\mathbb{Q}_p$ is abelian. Since any abelian normal extension of $\mathbb{Q}_p$ can be generated by roots of unity, we see that $K_0/\mathbb{Q}_p$ is generated by roots of unity whose order is prime to p and $K_1 \subset K_0(\zeta_p)$. By Lemma 2.1 there exists a number $d|p-1$ such that $K_1 = K_0(\pi^d)$. Consequently the ramification index e equals $(p-1)/d$.
Of course the characters χ_1, χ_2 restricted to I are powers χ^b, χ^a of χ^d, where $\gcd(a,b) = d$.
qed

Although not strictly necessary for our story, we recall what happens if I_w is trivial.

Lemma 2.3 *Suppose that I_w is trivial, i.e. K is tamely ramified over $\mathbb{Q}_p$. Then we have the following possibilities.*

1. *There exist two characters ϕ, ϕ' on I such that $\phi' = \phi^p, \phi = (\phi')^p$ and*
$$\rho\,|I = \begin{pmatrix} \phi & 0 \\ 0 & \phi' \end{pmatrix}.$$
Moreover, D_p is non-abelian in this case.

2. *There exist integers a, b such that*
$$\rho\,|I = \begin{pmatrix} \chi^b & 0 \\ 0 & \chi^a \end{pmatrix}.$$
Moreover, D_p is abelian in this case.

Proof. If the group D_p is abelian we can finish by the same arguments as in the previous Lemma. It then turns out that $K/\mathbb{Q}_p$ is generated by roots of unity and we are in the second case of our Lemma. So from now on we assume that D_p is non-abelian.
The extension $K = K_1/K_0$ is generated by a uniformiser π , ramified of order e over p. The index e is not divisible by p. Please be warned, the π we use here in this proof has in principle nothing to do with the π we use elsewhere in these notes. The inertia group I is now a cyclic group of order e, generated by an element we call σ. Let $F \in Gal(K/\mathbb{Q}_p)$ be an element which is a lift of the Frobenius element of $Gal(K_0/\mathbb{Q}_p)$. Then there exist p-adic units $\psi, \beta \in K_0$ such that
$$\sigma(\pi) \equiv \psi\pi (\text{mod } \pi^2), \qquad F(\pi) \equiv \beta\pi (\text{mod } \pi^2).$$
From this we deduce
$$(F \circ \sigma)(\pi) \equiv F(\psi\pi) \equiv \psi^p\beta\pi (\text{mod } \pi^2).$$
The latter is easily seen to be equal to $(\sigma^p \circ F)(\pi)(\text{mod } \pi^2)$. Hence $F \circ \sigma$ and $\sigma^p \circ F$ differ by an element from I_w which is the trivial group. We conclude that $F \circ \sigma = \sigma^p \circ F$.
Since I is cyclic the restriction of ρ to I consists of a direct sum of two characters we call ψ, ψ'. Because D_p is non-abelian, the characters ψ, ψ' are distinct. Furthermore, conjugation of $\rho|_I$ by F interchanges the characters ψ and ψ'. But we also have that $\rho(F^{-1} \circ \sigma \circ F) = \rho(\sigma)^p$. Hence we conclude that $\psi^p = \psi'$ and $(\psi')^p = \psi$. **qed**

We define the Serre-weight $k(\rho)$ as follows. First we deal with the case $\rho|_I = \begin{pmatrix} \chi^b & * \\ 0 & \chi^a \end{pmatrix}$. When I_w is trivial, we can interchange a, b if necessary so that we have $0 \le a \le b \le p-2$. We define $k(\rho) = 1 + pa + b$. When I_w is not trivial we take $0 \le a \le p-2$ and $1 \le b \le p-1$. When $b = a+1$ and $\chi^{-a} \otimes \rho$ is not finite at p, we set $k(\rho) = (a+1)(p+1)$ and $k(\rho) = 1 + p\min(a,b) + \max(a,b)$ in all other cases.
Secondly we deal with the case when $\rho|_I$ is a direct sum of two conjugate characters. Letting π be again the uniformizer of K/K_0, then the action of I can be described by a character ψ with values in $\mathbb{F}_{p^2}$ via $\sigma : \pi(\text{mod } \pi^2) \mapsto \psi(\sigma)\pi(\text{mod } \pi^2)$. The characters ϕ, ϕ' are powers of ψ. After interchanging ϕ, ϕ' is necessary, we can find integers a, b with $o \le a < b \le p-1$ such that $\phi = \psi^{a+pb}$. We set $k(\rho) = 1 + pa + b$.

Let us now turn back to the case when I_w is non-trivial. By taking tensor products $\chi^c \otimes \rho$ we can shift the weight of ρ by multiples of $p-1$. We do this in such a way that the new weight lies between 2 and $p+1$. We call this the reduced Serre-weight $\tilde{k}(\rho)$.
In the case when I_w is non-trivial it can be defined as follows. Let a, b as before and choose an integer k such that $2 \le k \le p$ and $k - 1 = b - a(\text{mod } p-1)$

$$\tilde{k} = \begin{cases} p+1 & \text{if } k = 2 \text{ and } \rho \otimes \chi^{-a} \text{ not finite} \\ k & \text{otherwise} \end{cases}$$

Theorem 2.4 (Moon, Taguchi) *Let $\mathcal{D}_{K/\mathbb{Q}_p}$ be the different of $K/\mathbb{Q}_p$ and define $v_p(p) = 1$. Let $d = \gcd(a, b, p-1)$. Then*

$$v_p(\mathcal{D}_{K/\mathbb{Q}_p}) = \begin{cases} 1 + \frac{\tilde{k}-1}{p-1} - \frac{\tilde{k}-1+d}{(p-1)p^m} & \text{if } 2 \le \tilde{k} \le p \\ 2 + \frac{1}{(p-1)p} - \frac{2}{(p-1)p^m} & \text{if } \tilde{k} = p+1 \end{cases}$$

Comparing this with Tate's result,

Theorem 2.5 (Tate) *With the same notations as before,*

$$v_p(\mathcal{D}_{K/\mathbb{Q}_p} \le 2 + \frac{1}{p(p-1)} - \frac{2}{(p-1)p^m}.$$

Application: take $p = 2$. Then $v_2(\mathcal{D}_{K/\mathbb{Q}_2}) \le 5/2$. Assume that the representation representation ρ of $G_{\mathbb{Q}}$ is irreducible and unramified outside 2. For the discriminant $d_{K/\mathbb{Q}}$ this implies

$$d_{K/\mathbb{Q}}^{1/n} \le 2^{5/2} < 5.66$$

contradicting the Minkowski bound when $n > 400$ and the Odlyzko bound when $n \ge 8$. A case by case reduction yields $n = 1$, the trivial representation.

Theorem 2.6 (Serre) *There are no irreducible continuous Galois representations, unramified outside $p = 3$.*

Proof Apply Tate's bound with $p = 3$ to get $d_{K/\mathbb{Q}}^{1/n} < 3^{7/3} < 13$. Via Odlyzko's bounds we have a lower bound of 13 when $n \ge 48$. So, $n \le 38$. But then the image $\rho(G_{\mathbb{Q}})$ is solvable and can be described explicitly. A case by case reduction then gives the result. **qed**

3 Proofs

Let $\mathcal{O}$ be the ring of integers in K_1. Then π^d is a generator of the ideal $\{x \in \mathcal{O} \mid |x|_p < 1\}$. We have $\mathcal{O} = \mathcal{O}_{K_0}[\pi^d]$. Recall that $e/(p-1)/d$.
The group of units in $\mathcal{O}$ is denoted by U. The group of units of the form $1+\pi^{di}\alpha$ with $\alpha \in \mathcal{O}$ is denoted by $U^{(i)}$. We have the filtration

$$U \supset U^{(1)} \supset U^{(2)} \supset \cdots \supset U^{(i)} \supset \cdots$$

Denote the p-th powers of the elements of $U^{(1)}$ by $(U^{(1)})^p$. Then we have,

$$U^{(e+2)} \subset (U^{(1)})^p \subset U^{(e+1)}.$$

More precisely,

Lemma 3.1 *We have*

$$U^{(e+2)} \subset (U^{(1)})^p \subset U^{(e+1)}.$$

If $d > 1$ then $(U^{(1)})^p = U^{(e+1)}$. If $d = 1$ then $(U^{(1)})^p$ has index p in $U^{(e+1)}$.

Proof. It is an exercise to show that $(U^{(1)})^p \subset U^{(e+1)}$ and $(U^{(2)})^p = U^{(e+2)}$. The first statement follows from this. Consider the p-th power map

$$\alpha : U^{(1)}/U^{(2)} \to (U^{(1)})^p/(U^{(2)})^p \subset U^{(e+1)}/U^{(e+2)}.$$

The kernel of α consists of the p-th roots of unity contained in K_1. So, if $d > 1$, the map α is a bijection and since the quotients $U^{(i)}/U^{(i+1)}$ all have the same cardinality, we conclude $(U^{(1)})^p = U^{(e+1)}$. When $d = 1$ and $K_1 = K_0(\zeta_p)$, the map α has kernel of order p and $(U^{(1)})^p$ has index p in $U^{(e+1)}$. **qed**

According to local classfield theory of the abelian extension K/K_1 we have a surjective classfield mapping

$$\phi : U \to I_w.$$

The kernel is precisely the norm group $N\mathcal{O}_K^*$. Since I_w is a p-group we can restrict ϕ to

$$\phi : U^{(1)} \to I_w.$$

Let $\kappa : I_w \to \mathbb{C}^*$ be a one-dimensional character. We define the conductor to be $\pi^{df(\kappa)}$ where

$$f(\kappa) = \min\{k \mid U^{(k)} \subset \ker(\kappa \circ \phi)\}.$$

In particular, $f(\chi_0) = 0$ for the trivial character χ_0. Then we have the conductor-discriminant relation

$$[K : K_1]v_p(\mathcal{D}_{K/K_1}) = \left(\sum_{\kappa \in \hat{I}_w} f(\kappa)\right) v_p(\pi^d).$$

Proof of Tate's theorem.
Notice that $(U^{(1)})^p \subset \ker(\kappa \circ \phi)$ for any character $\kappa : I_w \to \mathbb{C}^*$.
Suppose that $d > 1$. Then we have $U^{(e+1)} = (U^{(1)})^p$ and hence $f(\kappa) \leq e+1$ for all non-trivial characters κ. By the conductor-discriminant relation we now obtain

$$v_p(\mathcal{D}_{K/K_1} \leq \frac{1}{p^m}(p^m - 1)(e+1)v_p(\pi^d).$$

Together with $v_p(\mathcal{D}_{K_1/K_0} = 1 - 1/e$ and $v_p(\mathcal{D}_{K/\mathbb{Q}_p}) = v_p(\mathcal{D}_{K/K_1} + v_p(\mathcal{D}_{K_1/K_0}$ we obtain

$$v_p(\mathcal{D}_{K/\mathbb{Q}_p}) \leq 2 - (e+1)/p^m.$$

Suppose that $d = 1$ and $e = p-1$. Then $(U^{(1)})^p$ has index p in $U^{(e+1)}$. Of the p^m characters of I_w $p^m - p^{m-1}$ have conductor dividing $\pi^2 p$, $p^{m-1} - 1$ have conductor dividing πp and the trivial character has trivial conductor. We get

$$v_p(\mathcal{D}_{K/K_1}) \leq \frac{1}{p^m}\left((p^m - p^{m-1})(1 + 2/e) + (p^{m-1} - 1)(1 + 1/e)\right)$$

from which

$$v_p(\mathcal{D}_{K/\mathbb{Q}_p}) \leq 2 + \frac{1}{p(p-1)} - \frac{1}{p^{m-1}(p-1)}$$

follows immediately. **qed**

Proof of the Moon-Taguchi upper bound.
Let $\phi : U^{(1)} \to I_w$ be the classfield map as before. In addition ϕ is compatible with the action of I_t in the following sense

$$(\phi \circ \sigma)(u) = \sigma \phi(u) \sigma^{-1}$$

for all $\sigma \in I_t$. Suppose that $\tau \in I_w, \sigma \in I_t$ and

$$\rho(\sigma) = \begin{pmatrix} \chi^a(\sigma) & * \\ 0 & \chi^b(\sigma) \end{pmatrix} \qquad \rho(\tau) = \begin{pmatrix} 1 & \psi \\ 0 & 1 \end{pmatrix}.$$

then

$$\rho(\sigma\tau\sigma^{-1}) = \begin{pmatrix} 1 & \chi^{a-b}(\sigma)\psi \\ 0 & 1 \end{pmatrix} = \rho(\tau)^{\chi^{a-b}(\sigma)}.$$

Hence $\sigma\tau\sigma^{-1} = \tau^{\chi^{a-b}(\sigma)}$ for all $\sigma \in I_t$ and all $\tau \in I_w$.
Now consider the action of $\sigma \in I_t$ on $U^{(i)}/U^{(i+1)}$. Since, by Lemma 2.1, $\sigma(\pi^d) = \chi^d(\sigma)\pi^d (\mathrm{mod}\ \pi^{2d})$, we get

$$\begin{aligned} \sigma(1 + u\pi^{di}) &= (1 + u\chi^{di}(\sigma)\pi^{di}) (\mathrm{mod}\ \pi^{d(i+1)}) \\ &= (1 + u\pi^{di})^{\chi^{di}(\sigma)} (\mathrm{mod}\ \pi^{d(i+1)}) \end{aligned}$$

Since ϕ is I_t-equivariant, we conclude that ϕ maps $U^{(i)}/U^{(i+1)}$ to the trivial element if $di \neq k - 1 (\mathrm{mod}\ p-1)$, i.e. $i \neq (k-1)/d (\mathrm{mod}\ e)$.

Suppose first that $(k-1)/d \neq 1 (\mathrm{mod}\ e)$. Then $U^{(i)}/U^{(i+1)}$ has trivial image under ϕ for $i = (k-1)/d + 1, ..., e+1$. Since $U^{(e+2)}$ always has trivial image under ϕ we conclude that $f(\chi) \leq (k-1)/d + 1$ for all characters in $\hat{I}_w$. Application of the conductor-discriminant relation then gives us

$$v_p(\mathcal{D}_{K/K_1}) \leq (1 - p^{-m})((k-1)/d + 1) v_p(\pi^d).$$

This leads to

$$v_p(\mathcal{D}_{K/\mathbb{Q}_p}) \leq 1 + \frac{k-1}{p-1} - \frac{k-1+d}{p^m(p-1)}.$$

Suppose now that $(k-1)/d = 1 (\mathrm{mod}\ e)$. Hence $d = 1, e = p-1$ and $k = 2 (\mathrm{mod}\ p-1)$. In this case both $U^{(1)}/U^{(2)}$ and $U^{(p)}/U^{(p+1)}$ may have non-trivial image under ϕ. By a result of Serre $U^{(p)}$ has trivial image if and only if K/K_1 is "peu ramifié" if and only if the representation $\rho \otimes \chi^{-a}$ is finite. This, as remarked before, is equivalent to the case when K can be generated over K_1 by p-th roots of units in K_1. In this case we can proceed as before with $\tilde{k} = k = 2$. When $\tilde{k} = p+1$ we recover Tate's bound. **qed**

4 References

[MT] H.Moon, Y.Taguchi, Refinement of Tate's discriminant vound and non-existence for mod p galois representations., Documenta Math. Extra Kato Volume (2003), 641-654.

[T] J.Tate, The non-existence of certain Galois extensions of $\mathbb{Q}$ unramified outside 2, Contemp. Math. 174(1994), 153-156.

[S1] J.P.Serre, Note 229 on p710 of Serre's Collected Works, Volume 3.

[S2] J.P.Serre, Sur les représentations modualires de degré 2 de Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$, Duke Math.J. 54(1987), 179-230.