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, possibly reducible and possibly ramified at other primes than $p$.

Let $K$ be the invariant field of the kernel of $\rho$. Let $\mathcal{P}$ be a prime in $K$ over $p$ and $D_\mathcal{P}$ its decomposition group. Denote the completion of $K$ with respect to $\mathcal{P}$ again by $K$. The we get the faithful representation

$$\rho : D_\mathcal{P} = Gal(K/Q_p) \to GL_2(\overline{\mathbb{F}}_p).$$
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(\mathbb{Q}/\mathbb{Q}) \to F_p^*\) be the cyclotomic character defined by $\sigma(\zeta_p) = \zeta_p^{\chi(\sigma)}$ for any $p$-th root of unity $\zeta_p$.

We say that $\rho$ 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 $\pi$ 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 $\pi$ 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 $\pi$ 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 $\rho$ to $I$ has the form $\left( \begin{array}{cc} \chi^b & * \\ 0 & \chi^a \end{array} \right)$ 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 $\left( \begin{array}{cc} 1 & * \\ 0 & 1 \end{array} \right)$.

**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
\[\left( \begin{array}{cc} 1 & * \\ 0 & 1 \end{array} \right).\]

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 $\rho$ to $D_p$ has the form $\left( \begin{array}{cc} \chi_1 & * \\ 0 & \chi_2 \end{array} \right)$. Here $\chi_1$ and $\chi_2$ are characters on $D_p$. The semisimplification of $\rho : D_p \to GL_2(\mathbb{F}_p)$ consists of the direct sum of $\chi_1$ and $\chi_2$. Its kernel is $I_w$ and thus we see that $D_p/I_w$ is abelian.
In particular, $K_1/Q_p$ is abelian. Since any abelian normal extension of $Q_p$ can be generated by roots of unity, we see that $K_0/Q_p$ 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 $\chi_1, \chi_2$ restricted to $I$ are powers $\chi^b, \chi^a$ of $\chi^d$, where $\gcd(a, b) = d$.

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

1. There exist two characters $\phi, \phi'$ 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.

\textbf{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 $K/Q_p$ 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 $\pi$, ramified of order $e$ over $p$. The index $e$ is not divisible by $p$. Please be warned, the $\pi$ we use here in this proof has in principle nothing to do with the $\pi$ we use elsewhere in these notes. The inertia group $I$ is now a cyclic group of order $e$, generated by an element we call $\sigma$. Let $F \in Gal(K_0/Q_p)$ be an element which is a lift of the Frobenius element of $Gal(K_0/Q_p)$. Then there exist $p$-adic units $\psi, \beta \in K_0$ such that

$$
\sigma(\pi) \equiv \psi\pi (\text{mod } \pi^2), \quad F(\pi) \equiv \beta\pi (\text{mod } \pi^2).
$$

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 $\rho$ to $I$ consists of a direct sum of two characters we call $\psi, \psi'$. Because $D_p$ is non-abelian, the characters $\psi, \psi'$ are distinct. Furthermore, conjugation of $\rho | I$ by $F$ interchanges the characters $\psi$ and $\psi'$. But we also have that $\rho(F^{-1} \circ \sigma \circ F) = \rho(\sigma)^p$.

Hence we conclude that $\psi^p = \psi'$ and $(\psi')^p = \psi$. 

\textbf{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 $0 \leq a \leq b \leq p-2$. We define $k(\rho) = 1+pa+b$. When $I_w$ is not trivial we take $0 \leq a \leq p-2$ and $1 \leq b \leq p-1$. 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 $\pi$ be again the uniformizer of $K/K_0$, then the action of $I$ can be described by a character $\psi$ with values in $\mathbb{F}_{p^2}$ via $\sigma : \pi (mod \; \pi^2) \mapsto \psi(\sigma)\pi (mod \; \pi^2)$. The characters $\phi,\phi'$ are powers of $\psi$. After interchanging $\phi,\phi'$ is necessary, we can find integers $a,b$ with $0 \leq a < b \leq p-1$ such that $\phi = \psi^{a+b\phi}$. 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 $\chi^e \otimes \rho$ we can shift the weight of $\rho$ 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 $2 \leq k \leq p$ and $k - 1 = b - a (mod \; p-1)$

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

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

$$v_p(D_{K/Q_p}) = \begin{cases} 1 + \frac{k-1}{p} - \frac{k-1+d}{(p-1)p^m} & \text{if } 2 \leq \tilde{k} \leq p \\ 2 + \frac{1}{(p-1)p} - \frac{2}{(p-1)p^m} & \text{if } \tilde{k} = p+1 \end{cases}$$

Comparing this with Tate’s result,

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

$$v_p(D_{K/Q_p}) \leq 2 + \frac{1}{p(p-1)} - \frac{2}{(p-1)p^m}.$$  

Application: take $p = 2$. Then $v_2(D_{K/Q_2}) \leq 5/2$. Assume that the representation representation $\rho$ of $G_Q$ is irreducible and unramified outside 2. For the discriminant $d_{K/Q}$ this implies

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

contradicting the Minkowski bound when $n > 400$ and the Odlyzko bound when $n \geq 8$. 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 $d_{K/Q}^{1/n} < 3^{7/3} < 13$. Via Odlyzko’s bounds we have a lower bound of 13 when $n \geq 48$. So, $n \leq 38$. But then the image $\rho(G_Q)$ is solvable and can be described explicitly. A case by case reduction then gives the result. \textbf{qed}
3 Proofs

Let $\mathcal{O}$ be the ring of integers in $K_1$. Then $\pi^d$ 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 $1 + \pi^d \alpha$ 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 $\alpha$ consists of the $p$-th roots of unity contained in $K_1$. So, if $d > 1$, the map $\alpha$ 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 $\alpha$ has kernel of order $p$ and $(U^{(1)})^p$ has index $p$ in $U^{(e+1)}$. \( \text{qed} \)

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 $NO^*_K$. Since $I_w$ is a $p$-group we can restrict $\phi$ 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 $\chi_0$. Then we have the conductor-discriminant relation

\[ [K : K_1]v_p(D_{K/K_1}) = \left( \sum_{\kappa \in I_w} f(\kappa) \right) v_p(\pi^d). \]
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 \(\kappa\). By the conductor-discriminant relation we now obtain

\[
v_p(D_{K/K_1} \leq \frac{1}{p^m} (p^m - 1)(e + 1)v_p(\pi^d).
\]

Together with \(v_p(D_{K_1/K_0} = 1 - 1/e\) and \(v_p(D_{K/Q_p}) = v_p(D_{K/K_1} + v_p(D_{K_1/K_0})\) we obtain

\[
v_p(D_{K/Q_p}) \leq 2 - (e + 1)/p^m.
\]

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 \(\pi^2 p\), \(p^m - 1\) have conductor dividing \(\pi p\) and the trivial character has trivial conductor. We get

\[
v_p(D_{K/K_1} \leq \frac{1}{p^m} ((p^m - p^m - 1)(1 + 2/e) + (p^m - 1)(1 + 1/e))
\]

from which

\[
v_p(D_{K/Q_p}) \leq 2 + \frac{1}{p(p - 1)} - \frac{1}{p^m - 1(p - 1)}
\]

follows immediately.

Proof of the Moon-Taguchi upper bound.

Let \(\phi : U^{(1)} \to I_w\) be the classfield map as before. In addition \(\phi\) 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

\[
\rho(\sigma) = \begin{pmatrix} \chi^a(\sigma) & * \\ 0 & \chi^b(\sigma) \end{pmatrix}, \quad \rho(\tau) = \begin{pmatrix} 1 & \psi \\ 0 & 1 \end{pmatrix}.
\]

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, \(\sigma(\pi^d) = \chi^d(\sigma) \pi^d(\mod \pi^{2d})\), we get

\[
\sigma(1 + u \pi^{di}) = (1 + u \chi^{di}(\sigma) \pi^{di})(\mod \pi^{d(i+1)}) = (1 + u \pi^{di}) \chi^{d(\sigma)}(\mod \pi^{d(i+1)}).
\]

Since \(\phi\) is \(I_t\)-equivariant, we conclude that \(\phi\) maps \(U^{(i)}/U^{(i+1)}\) to the trivial element if \(di \neq k - 1(\mod p - 1)\), i.e. \(i \neq (k - 1)/d(\mod e)\).
Suppose first that \((k-1)/d \neq 1 \text{ mod } e\). Then \(U(i)/U(i+1)\) has trivial image under \(\phi\) for 
\(i = (k-1)/d + 1, \ldots, e+1\). Since \(U(e+2)\) always has trivial image under \(\phi\) 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 
\[ v_p(D_{K/K_1}) \leq \frac{1}{(k-1)/d + 1} \cdot v_p(\pi^d). \]
This leads to 
\[ v_p(D_{K/Q_p}) \leq 1 + \frac{k-1}{p-1} - \frac{k-1 + d}{p^m(p-1)}. \]
Suppose now that \((k-1)/d = 1 \text{ mod } e\). Hence \(d = 1, e = p-1\) and \(k = 2 \text{ (mod } p-1)\). In this 
case both \(U(1)/U(2)\) and \(U(p)/U(p+1)\) may have non-trivial image under \(\phi\). By a result of Serre 
\(U(p)\) has trivial image if and only if \(K/K_1\) is ”nuisant ramié” 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. \(\text{qed}\)

4 References

[MT ] H.Moon, Y.Taguchi, Refinement of Tate’s discriminant bound and non-existence for 

[T ] J.Tate, The non-existence of certain Galois extensions of \(\mathbb{Q}\) unramified outside 2, 
