Author's Notes (2008)

(I) This is a reproduction of "On Congruence Monodromy Problems" Volume 1 (1968) and Volume 2 (1969) issued as Lecture Notes (No 1,2) from Department of Mathematics, Univeristy of Tokyo, which remained unpublished (except for a Russian translation [**11**] of Volume 1). When this reproduction was proposed, I was sure they were old, for old colleagues, but at the same time not so sure whether they did not contain anything new for the newcomers, because they had not been buried so peacefully. There had been several occasions when I felt I would have liked to share a piece of mathematics in its natural form whose understanding would have clarified my younger colleague's contemporary questions. And on some other occasions, I had to face with some basic misunderstandings[1]. So, I agreed that the main results (including those obtained later, which will be briefly reviewed below), some points of view, methods used, and related open problems may still deserve some attention.

I have then re-read the text carefully, and fortunately or unfortunately, found no errors other than small local ones. Aside from these corrections and some arrangements to unify the two Volumes, no changes have been made in the present reproduction. On the other hand, obviously, the author had been too nervous in giving all the details almost everywhere. So, for the possible readers of this reproduction I would suggest reading only the introductions and the outlines, and for details, only when necessary.

This reproduction was proposed by a colleague of mine, Takayuki Oda, who has also taken the labor to put the original typed text into TeX' text (mainly Volume 1). I wish to express my deep gratitude to him and to T.Ichikawa, M.Kaneko and H.Tsutsumi, who had helped a great deal in continuing and finishing this laborious task.

(II) For a given finite Galois extension K/k of global fields, knowing the global Artin L-functions $L(s,\chi,K/k)$ (for all irreducible characters χ of the Galois group) is one thing, knowing the Frobenius conjugacy class for each (unramified) prime of k is quite another. They are the same when k is the rational number field $\mathbf{Q}$, but *essentially different* for function fields over finite fields. This simple fact does not seem to be widely recognized even in the circle of number theorists, and since the justification of the present reproduction of old lecture notes will never be understood without this recognition, I started with this; now let me recall the reason.

From a global Artin L-function, what one can pick up as local data is, for each prime number p, just *the product of* Euler factors corresponding to those primes of k whose

[1]The biggest of which is that (even in the case of function fields) the Langlands correspondence should contain everything related to non-abelian classfield theory and hence our work merely gives its "examples" in disguise. The following (otherwise superfluous) subsection (II) is for some extra guide...

norms are some powers of p. When $k = \mathbf{Q}$, each p-factor can be picked up. But in case of function fields over $\mathbf{F}_q$, the situation is the other extreme; all norms are some powers of one and the same q. No single local factor is determined by the global L-function. Knowing the Artin-Weil polynomials for all χ does not help much in knowing the individual Frobenius elements. For a clearcut example, consider any non-trivial (e.g. A_5-) extension without constant extensions such that the genus of the overfield K is still 0. Then all L-functions for irreducible non-principal characters are identically equal to 1!

The situation is similar for L-functions associated with more general ℓ-adic representations of $\mathrm{Gal}(k^{sep}/k)$. Knowing their automorphy (by [7] [29]) does not help unless each local automorphic representation can be picked up from the global one.

(III) The "congruence monodromy problem" is aimed at the study of individual Frobenius elements for some special but natural systems of Galois extensions of function fields over finite fields arising from Shimura curves.

First let us briefly recall the elliptic modular case (Chapter 5, plus [15]). To begin with, recall that

$$\Delta = \{\delta \in SL_2(\mathbf{Z});\ \delta \equiv 1 (\mathrm{mod}\, 2)\}/\pm 1 \tag{1}$$

is the topological fundamental group of $\mathbf{P}^1(\mathbf{C}) - \{0, 1, \infty\}$. Its subgroups with finite indices correspond bijectively with the connected finite coverings of the complex projective line unramified outside $0, 1, \infty$. As in the main text, write $\mathbf{Z}^{(p)} = \mathbf{Z}[\frac{1}{p}]$, and for any prime $p \neq 2$ consider the group

$$\Gamma_p = \{\gamma \in SL_2(\mathbf{Z}^{(p)});\ \gamma \equiv 1 (\mathrm{mod}\, 2)\}/\pm 1. \tag{2}$$

Then Γ_p is, in the sense described below, "the arithmetic fundamental group" for the system of connected coverings of the projective line over $\mathbf{F}_{p^2}$ characterized by:

(i) it is unramified outside $0, 1, \infty$, and at most tamely ramified at $0, 1, \infty$,

(ii) all points of $\mathfrak{S}_p$ are decomposed completely.

Here, $\mathfrak{S}_p$ denotes the set of all supersingular λ-invariants of the elliptic curve $y^2 = x(x-1)(x-\lambda)$. They are the roots of the polynomial

$$P(\lambda) = \sum_{i=0}^{r} \binom{r}{i}^2 \lambda^i \qquad (r = (p-1)/2) \tag{3}$$

over $\mathbf{F}_p$ (all simple and contained in $\mathbf{F}_{p^2}$). For example, $\mathfrak{S}_3 = \{-1\}$.

It means, first, that the subgroups of Γ_p with finite indices correspond bijectively with the finite subcoverings in this system. In fact, we have the isomorphisms:

$$\begin{aligned} \Gamma_p \longrightarrow \widehat{\Gamma}_p &\simeq \{g \in \prod_{\ell \neq p} SL_2(\mathbf{Z}_\ell);\ g \equiv 1 (\mathrm{mod}\, 2)\}/\pm 1 \\ &\simeq \mathrm{Gal}(\mathfrak{R}/\mathbf{F}_{p^2}(\lambda)) = \mathrm{Gal}(\widehat{\mathfrak{R}}/\mathbf{F}_{p^2}(\lambda)), \end{aligned} \tag{4}$$

where $\widehat{\Gamma}_p$ denotes the profinite completion, $\mathfrak{R}$ is the composite of the fields of modular functions of all levels $\not\equiv 0 (\mathrm{mod}\, p)$ over $\mathbf{F}_{p^2}$, and $\widehat{\mathfrak{R}}$ is the total function field of the system characterized by (i)(ii). The first isomorphism expresses the congruence subgroup property of SL_2 over $\mathbf{Z}^{(p)}$ proved by J.Mennicke and J-P.Serre, the second is treated in Chapter

5, and the last equality holds because the Conjecture $\widehat{\mathfrak{R}} = \mathfrak{R}$ (Chapter 5 §30) was later solved [**15**]. This is the first aspect.

Secondly, Γ_p has another embedding as a discrete subgroup

$$\Gamma_p \longrightarrow G_{\mathbf{R}} \times G_p = PSL_2(\mathbf{R}) \times PSL_2(\mathbf{Q}_p), \tag{5}$$

which defines naturally the notion of "primitive elliptic" Γ_p-conjugacy classes (Chapter 1, §11). The main result of Chapter 5 asserts that the "positive" primitive elliptic Γ_p-conjugacy classes correspond bijectively with the set of $\mathbf{F}_{p^2}$-conjugacy classes of $\overline{\mathbf{F}}_p$-rational points of $\mathbf{P}^1 - \{0, 1, \infty\} - \mathfrak{S}_p$, via the Frobenius elements in (4).

From my point of view, this countable dense subgroup Γ_p is more remarkable compared to the full profinite Galois group $\widehat{\Gamma}_p$, because it is "just of the correct size" to enable us to pick up the Frobenius conjugacy classes; it explains why the conjugacy classes corresponding to $\mathfrak{S}_p$ are missing in Γ_p, and also allows another approach for study (Selberg type zeta functions, etc.). In this strong sense, Γ_p is "the arithmetic fundamental group".

It is, though, still mysterious. For example, is there a direct relation between the following two expressions?

$$\Gamma_p = \Delta *_{\Delta^0} \Delta', \tag{6}$$

$$\widehat{\Gamma}_p = \pi_1^{tame}(\mathbf{X})/\langle \mathfrak{S}_p \rangle. \tag{7}$$

Here, the first expression is the free product decomposition with amalgamated subgroup $\Delta^0 = \Delta \cap \Delta'$, where $\Delta' = \omega_p^{-1}\Delta\omega_p$ ($\omega_p = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$). Since Δ^0 is free of rank $p + 2$, this expresses Γ_p by 4 generators and $p + 2$ relations. In the second expression, $\mathbf{X}$ is the projective line minus $0, 1, \infty$ over $\mathbf{F}_{p^2}$, and π_1^{tame} is Grothendieck's tame fundamental group[2]. Thus, by (4), the quotient of $\pi_1^{tame}(\mathbf{X})$ by the normal subgroup $\langle \mathfrak{S}_p \rangle$ generated by the Frobenius elements above $\mathfrak{S}_p$ must coincide with $\widehat{\Gamma}_p$.

More basically, how should one understand the appearance of $\mathfrak{S}_p$ when we reduce the covering system modulo p? Its connection with the reduction mod p of the Schwarzian differential equation defining the uniformization of $\mathbf{P}^1 - \{0, 1, \infty\}$ over $\mathbf{C}$ (inspired by Igusa's differential equation satisfied by the polynomial $P(\lambda)$), is another subject of our study. Or can we start from curves over finite fields ? (For these, see (VI)(IX) below).

(IV) As stated in the General Introduction, the main aim of the intended series of Volumes was to solve the Main Conjectures (Conjectures 1,2,3) proposed there. But since then, instead of continuing on to Volumes 3,4,... etc., the subsequent related works by the author have been published one by one in some Journals or Proceedings. And later, these conjectures have been affirmatively solved, at least in the author's mind, when the final

[2]Let us briefly recall the basic known facts about $\pi_1^{tame}(\mathbf{X})$. Let $pr : \pi_1^{tame}(\mathbf{X}) \to \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_{p^2})$ denote the projection. Then, first, by Grothendieck, its kernel $\pi_1^{tame}(\mathbf{X} \otimes \overline{\mathbf{F}}_p)$ is a quotient of $\widehat{\Delta}$, about which not much is known other than that it has the same maximal "prime-to-p" quotient as $\widehat{\Delta}$. Secondly, by the work of A.Tamagawa [**41**] on the tame fundamental groups of affine hyperbolic curves over finite fields, the group-theoretic pair $(\pi_1^{tame}(\mathbf{X}), pr)$ in a sense contains all the information on the points of $\mathbf{X}$. But we still know very little about the difference between $\pi_1^{tame}(\mathbf{X} \otimes \overline{\mathbf{F}}_p)$ and $\widehat{\Delta}$, or about the Frobenius outer action on the former group.

connection was made at the time of the publication of Y.Morita's paper [36]. More precise account on what is proved and how, is explained in [25][26]. A brief review will be given below in (VII). But the proof obtained relies very much on Margulis' result [31] for the arithmeticity of discrete subgroups, and also on Shimura's work on Shimura curves over number fields and their reductions using the moduli of (highly structured) abelian varieties [39]. A more direct proof seems desirable.

Before going on to review later developments, let us recall the Main Conjectures stated in the General Introduction, under a slightly generalized form and with more precision, and now *as a theorem*, as it has been proved since then.

First, a generalization. Let $k_\mathfrak{p}$ be any finite extension of $\mathbf{Q}_p$. As for the group $G_\mathfrak{p}$, we shall now take

$$G_\mathfrak{p} = PGL_2^+(k_\mathfrak{p}) = \{g \in GL_2(k_\mathfrak{p});\ \mathrm{ord}_\mathfrak{p}\det(g) \equiv 0(\mathrm{mod}\,2)\}/k_\mathfrak{p}^\times,$$

and instead of assuming the image of the projection of Γ on $G_\mathfrak{p}$ to be dense, impose only that the topological closure of the projection image *contains* $PSL_2(k_\mathfrak{p})$. Let us call a discrete subgroup Γ of $G = G_\mathbf{R} \times G_\mathfrak{p}$ with finite volume quotient *irreducible lattice* when this condition is satisfied, in addition to the density of the projection image on $G_\mathbf{R}$. For our "over $\mathbf{F}_{q^2}$-results", this is the most natural choice.

Secondly, we shall make it more explicit on what choice of isomorphisms our basic functors will depend. We embed $\mathbf{Q}$ diagonally into the ring $\mathbf{C}\oplus\bar{k}_\mathfrak{p}$ ($\bar{k}_\mathfrak{p}$: an algebraic closure of $k_\mathfrak{p}$), and fix *an algebraic closure* $\overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$ *of* $\mathbf{Q}$ in this ring. This is equivalent with fixing an isomorphism between the algebraic closure of $\mathbf{Q}$ in $\mathbf{C}$ and that in $\bar{k}_\mathfrak{p}$. Moreover, we assume that $\overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$ can be so chosen as to contain $t(\gamma) = (t(\gamma_\mathbf{R}), t(\gamma_\mathfrak{p}))$ for all $\gamma \in \Gamma$. Here, for any $g \in PGL_2$ over any field, we put $t(g) = (\mathrm{tr}(\tilde{g}))^2/\det(\tilde{g})$, $\tilde{g}$ being a representative of g in GL_2. Before Margulis' arithmeticity result, this assumption was one of the first goals to prove (Chapter 3, §§12-13). But by his arithmeticity, Γ must be commensurable with a group arising from a quaternion algebra B over a totally real number field F, and hence this condition is obviously satisfied. Note that $F = \mathbf{Q}(t(\gamma); \gamma \in \Gamma)$ (cf. Chapter 4, esp. §4).

Given such a choice of $\overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$, we define the *positivity* of an element γ of Γ_z (the stabilizer of $z \in \mathfrak{H}$ in Γ; cf. Chapter 1, §3), as follows. Since γ fixes z, its action on the tangent space at z is a scalar multiplication. Call this scalar $\lambda_\mathbf{R}$. It is a root of $\lambda_\mathbf{R} + \lambda_\mathbf{R}^{-1} + 2 = t(\gamma_\mathbf{R})$. Define $\lambda_\mathfrak{p}$ by the condition $(\lambda_\mathbf{R}, \lambda_\mathfrak{p}) \in \overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$. We call $\gamma \in \Gamma_z$ positive when $\mathrm{ord}_\mathfrak{p}\lambda_\mathfrak{p} < 0$. This is in accordance with the definition of positive generator in Chapter 5 Part 2 §23. Finally, for any subring $A_\mathfrak{p} \subseteq \bar{k}_\mathfrak{p}$, we write

$$A_\mathfrak{p}^{alg} = \overline{\mathbf{Q}}_{\infty,\mathfrak{p}} \cap (\mathbf{C} \oplus A_\mathfrak{p}).$$

Thus, $A_\mathfrak{p}^{alg}$-schemes are equipped with the base changes $\otimes\mathbf{C}$, $\otimes A_\mathfrak{p}$.

Main Theorem *To each torsion-free cocompact irreducible lattice* Γ *of* $G = G_\mathbf{R}\times G_\mathfrak{p}$, *one can associate a complete smooth geometrically irreducible curve* $\mathbf{X}$ *over* $\mathbf{F}_{q^2}$ ($q = N(\mathfrak{p})$) *of genus* $g \geq 2$, *together with a set* $\mathfrak{S}$ *consisting of* $(q-1)(g-1)$ $\mathbf{F}_{q^2}$*-rational points of* $\mathbf{X}$, *such that if we denote by* $\wp(\Gamma)$ (*resp.* $\wp(\mathbf{X})$) *the set of all positive primitive elliptic* Γ*-conjugacy classes* (*resp. that of* $\mathbf{F}_{q^2}$*-conjugacy classes of* $\overline{\mathbf{F}}_q$*-rational points of* $\mathbf{X} - \mathfrak{S}$),

then:

(MT-1) *There is a canonical bijection*

$$\wp(\Gamma) \approx \wp(\mathbf{X}). \tag{8}$$

(MT-2) *There is a canonical isomorphism (determined up to conjugacy)*

$$\widehat{\Gamma} \simeq \mathrm{Gal}(\widehat{\mathfrak{R}}/K) \tag{9}$$

from the profinite completion of Γ *to the Galois group* $\mathrm{Gal}(\widehat{\mathfrak{R}}/K)$, *where* K *is the function field of* $\mathbf{X}$ *over* $\mathbf{F}_{q^2}$ *and* $\widehat{\mathfrak{R}}$ *is the maximal unramified Galois extension of* K *in which all points of* $\mathfrak{S}$ *are decomposed completely.*

(MT-3) *For each element of* $\wp(\mathbf{X})$, *its Frobenius conjugacy class in* $\mathrm{Gal}(\widehat{\mathfrak{R}}/K)$ *is given by the image of the corresponding element of* $\wp(\Gamma)$.

I shall review briefly how they have been proved, together with other related results, open problems, and give additional references.

(V) The main lines of investigations pursued after these Volumes will now be summarized. Here, brevity is preferred; for precise account, cf. the referred articles. For (V), cf. **[20][21][22]**. We assume, for simplicity here, that Γ is cocompact and torsion-free.

First, recall that $G_{\mathfrak{p}}$ is the free product of two mutually adjacent maximal compact subgroups V, V' with amalgamated subgroup $V^0 = V \cap V'$, and that Γ can be regarded as that of two corresponding discrete subgroups Δ, Δ' of $G_{\mathbf{R}} = PSL_2(\mathbf{R})$ with amalgamated subgroup $\Delta^0 = \Delta \cap \Delta'$. (Chap 2, §28). Hence each Γ determines a system

$$\mathcal{X}_{\mathbf{C}} = \{X_{\mathbf{C}} \leftarrow X^0_{\mathbf{C}} \rightarrow X'_{\mathbf{C}}\} \tag{10}$$

of three compact Riemann surfaces. Subgroups of Γ with finite indices correspond bijectively with the systems of "unramified coverings of $\mathcal{X}_{\mathbf{C}}$".

On the other hand, for a given (proper smooth geometrically irreducible) curve $\mathbf{X}$ over $\mathbf{F}_{q^2}$ and a non-empty subset $\mathfrak{S}$ of $\mathbf{X}(\mathbf{F}_{q^2})$, we can associate a system

$$\mathcal{X}_q = \{\mathbf{X} \leftarrow \mathbf{X}^0 \rightarrow \mathbf{X}'\}, \tag{11}$$

where $\mathbf{X}'$ is the conjugate of $\mathbf{X}$ over $\mathbf{F}_q$, $\mathbf{X}^0$ consists of two irreducible components $\Pi \overset{pr_1}{\simeq} \mathbf{X}$ and $\Pi' \overset{pr_2}{\simeq} \mathbf{X}'$ (pr_i: the projections), the image on $\mathbf{X} \times \mathbf{X}'$ of Π (resp. Π') is the graph of the q-th power morphism $\mathbf{X} \rightarrow \mathbf{X}'$ (resp. $\mathbf{X}' \rightarrow \mathbf{X}$), and finally, $pr_1(\Pi \cap \Pi') = \mathfrak{S}$. Note that this is possible because $\mathfrak{S} \subseteq \mathbf{X}(\mathbf{F}_{q^2})$, and that Π, Π' meet transversally at each intersection. A key point of this construction is that, finite unramified coverings of $\mathbf{X}$ in which all points of $\mathfrak{S}$ decompose completely correspond bijectively with "finite etale coverings of $\mathcal{X}_q$". The system $\mathcal{X}_q$, in a sense, *geometrically* realizes the arithmetic condition of complete decompositions above $\mathfrak{S}$.

Thus, in order to compare the coverings of $\mathcal{X}_{\mathbf{C}}$ and $\mathcal{X}_q$, which is the main content of (MT-2), the crucial point is to study (the existence and consequences of) a system

$$\mathcal{X} = \{\mathfrak{X} \leftarrow \mathfrak{X}^0 \rightarrow \mathfrak{X}'\} \tag{12}$$

of proper flat $O_{\mathfrak{p}}^{(2)}$-schemes whose reduction mod $\mathfrak{p}$ gives $\mathcal{X}_q$ and whose suitable base change $\otimes\mathbf{C}$ gives the complex system $\mathcal{X}_{\mathbf{C}}$. Here, $O_{\mathfrak{p}}^{(2)}$ denotes the unique unramified quadratic extension of the ring $O_{\mathfrak{p}}$ of integers of $k_{\mathfrak{p}}$. Each member of $\mathcal{X}$ is assumed to be integral and normal, and $\mathfrak{X}, \mathfrak{X}'$ moreover smooth. It is also natural to impose the additional condition that $\mathcal{X}$ be *symmetric*, i.e., its transpose is conjugate to itself over $O_{\mathfrak{p}}$.

First, in (VI) below, we shall discuss *consequences* of existence of the connecting system $\mathcal{X}$. I recall that when this work was presented (around 1975), the major reaction was "well, if you assume this, then no wonder such consequences can be derived". This is in a sense correct; I was able to prove them. But compare with the single curve case,

$$\{X_{\mathbf{C}}\} \quad \text{and} \quad \{\mathbf{X}\} \qquad \textit{connected by} \qquad \{\mathfrak{X}\} \quad \text{over } O_{\mathfrak{p}}.$$

Grothendieck theory of fundamental groups tells us that the category of etale coverings of the latter two schemes are mutually equivalent, but as for the comparison with the etale coverings of $X_{\mathbf{C}}$, we need Abhyankar's lemma. Even if we do not care about the constant rings, the best comparison theorem is for those coverings whose (Galois-)degrees are not divisible by the residue characteristic. A great merit here is that it is very general. In our case, Shimura curves are very rare, especially if one fixes the genus (cf. **[32][40][33]**), so, for many researchers, our object of study may appear to be too special. But this is more like the study of Hilbert modular surfaces each of which is isolated and does not have moduli. And a merit of considering the three-curve systems is that we have a *complete* comparison theorem (The "second Galois theory" described below). Other arithmetic results also require codimension 2 geometric considerations, combined with some specific group theory arising from the decomposition $\mathbf{X}^0 = \Pi \cup \Pi'$. I hope this will now be better accepted by the readers.

(VI) **Study starting from any given lifting $\mathcal{X}$ of a system $\mathcal{X}_q$.** Here, we *assume* the existence of $\mathcal{X}$ whose reduction is of the form $\mathcal{X}_q$. We do not assume that its base change $\otimes\mathbf{C}$ corresponds to some Γ. In particular, included here is the case where the projections $X_{\mathbf{C}}^0 \to X_{\mathbf{C}}$ and $X_{\mathbf{C}}^0 \to X_{\mathbf{C}}'$ involve ramifications. The basic references are **[20][21][22]** ([**21**] supplies the remaining details for [**20**]; [**22**] is for "the second Galois theory" in the ramified case)).

[**The first Galois theory**] This is the arithmetic Galois theory "mod $\mathfrak{p}$" arising from two projections in $\mathcal{X}$ indicated by horizontal arrows. Denote by $K, K' \subset K^0$ the function fields of $\mathfrak{X}, \mathfrak{X}', \mathfrak{X}^0$ respectively. They are function fields of one variable over the unique unramified quadratic extension $k_{\mathfrak{p}}^{(2)}$ of $k_{\mathfrak{p}}$. Let L be the smallest Galois extension of K^0 which is Galois over both K, K'. It is an infinite extension. It does not follow from the assumptions, nor is it additionally assumed, that only finitely many prime divisors of K are ramified in L. Let V, V', V^0 be the Galois groups of L over K, K', K^0, respectively, and call $G_{\mathfrak{p}}^{\star}$ the subgroup of the automorphism group of L generated by V, V'. When it corresponds to some Γ, L is essentially the $G_{\mathfrak{p}}$-field studied in Chapter 2. (To be more precise, $\overline{\Gamma}_{\mathfrak{p}}$ (the topological closure) acts trivially over the constant field of L, and the quotient $G_{\mathfrak{p}}/\overline{\Gamma}_{\mathfrak{p}}$ corresponds to the constant field extension in L.) In the general situation,

$G_{\mathfrak{p}}^{\star}$ is still the free product of V, V' with amalgamated subgroup V^0, and the union of the coset spaces $(V \setminus G_{\mathfrak{p}}^{\star}) \sqcup (V' \setminus G_{\mathfrak{p}}^{\star})$ has a structure of a (two-colored) tree which is the same as the tree $\mathcal{T}_{\mathfrak{p}}$ associated with $PGL_2^{+}(k_{\mathfrak{p}})$. Thus, $G_{\mathfrak{p}}^{\star}$ can be regarded as a subgroup of the automorphism group of $\mathcal{T}_{\mathfrak{p}}$.

Now consider the $G_{\mathfrak{p}}^{\star}$-orbit space in the set of all those places of $L/k_{\mathfrak{p}}$ whose stabilizer in $G_{\mathfrak{p}}^{\star}$ is "big" (the condition [A] in [**20**]§3.3). Then (*ibid.* Th. 3.4.1)

Theorem A *The reduction mod* $\mathfrak{p}$ *induces a bijection between this orbit space and the set of all* $\mathbf{F}_{q^2}$*-conjugacy classes of* $\overline{\mathbf{F}}_q$*-rational points of* $\mathbf{X} - \mathfrak{S}$.

A basis for this proof is the construction of the *canonical lifting* of each $\mathbf{F}_{q^{2n}}$-rational point of $\mathbf{X} - \mathfrak{S}$ to a $k_{\mathfrak{p}}^{(2n)}$-rational point of $\mathfrak{X} \otimes k_{\mathfrak{p}}^{(2n)}$, where $k_{\mathfrak{p}}^{(2n)}$ denotes the unique unramified extension of $k_{\mathfrak{p}}$ of degree $2n$ ($n \geq 1$). Just a word to explain the geometric construction of the canonical lifting. For $n = 1$, let $\mathfrak{X}^{00}$ be the image of $\mathfrak{X}^0$ on the fiber product $\mathfrak{X} \times \mathfrak{X}'$. Then the canonical lifting of an $\mathbf{F}_{q^2}$-rational point $\mathbf{x}$ of $\mathbf{X} - \mathfrak{S}$ is simply the first projection of the unique ordinary double point on the general fiber of $\mathfrak{X}^{00}$ that lifts $(\mathbf{x}, \mathbf{x}^q)$. For $n > 1$, we use higher iterations of the algebraic correspondence $\mathfrak{X}^{00}$ and use the unique liftability of an ordinary double point on the special fiber that is not normal on the total $O_{\mathfrak{p}}$-scheme, to an ordinary double point on the general fiber. This theorem will be a basis for the proof of (MT-1), and a certain stronger form, for that of (MT-3).

[**The second Galois theory**] Denote by $\Delta^*, (\Delta^*)', (\Delta^*)^0$ the topological fundamental groups of the compact Riemann surfaces obtained from $\mathfrak{X}$, $\mathfrak{X}'$, $\mathfrak{X}^0$ by any given base change $\otimes\mathbf{C}$ (and w.r.t. a compatible set of base points), so that we have canonical homomorphisms $(\Delta^*)^0 \to \Delta^*$, $(\Delta^*)^0 \to (\Delta^*)'$. Call Γ^* the free product of $\Delta^*, (\Delta^*)'$ with amalgamation defined by these two homomorphisms. When it corresponds to some Γ (cocompact, torsion-free), then $\Gamma^* = \Gamma$. In general, we have (cf [**20**][**21**][**22**])

Theorem B *Subgroups of* Γ^* *with finite indices are categorically equivalent with the connected etale coverings of each of* $\mathcal{X}$, $\mathcal{X}_q$ *and* $\mathcal{X}_{\mathbf{C}}$.

A basis for this proof is a "Frobenius-criterion" for good reduction of unramified coverings [**18**](a joint work with H. Miki). This theorem will be basic in proving (MT-2).

[**The unramified case**] In general, we have

$$|\mathfrak{S}| \geq (q-1)(g-1) \tag{13}$$

(g : the genus of $\mathbf{X}$). The equality holds if and only if the two projections in $\mathcal{X}$ are both unramified on the general fiber. We call the system "unramified", when this is satisfied. In this case, the homomorphisms $\Delta^*, (\Delta^*)' \to \Gamma^*$ are injective and Γ^* can be regarded as a discrete subgroup of the product $G_{\mathbf{R}} \times G_{\mathfrak{p}}^{\star}$;

$$\Gamma^* \subset G_{\mathbf{R}} \times G_{\mathfrak{p}}^{\star}. \tag{14}$$

In this unramified case, (a suitably rephrased) Main Theorem is valid (cf. [**20**][**21**]). In particular,

Theorem C *The above Main Theorem is valid if there exists a system $\mathcal{X}$ over $(O_{\mathfrak{p}}^{(2)})^{alg}$ connecting the given $\mathcal{X}_{\mathbf{C}}$ with $\mathcal{X}_q$.*

[The ramified case] In this case, the group Γ^* thus constructed can be a finite group. If one starts from an irreducible lattice Γ of G which is torsion-free but not cocompact, and build a system $\mathcal{X}_{\mathbf{C}}$, then passing to the groups with the asterisk means dividing by the normal subgroup generated by all parabolic elements; hence Γ^* can be much smaller than the original group Γ. For example, let N be an integer with $N > 1, (N, p) = 1$, and take Γ to be the principal congruence subgroup of $PSL_2(\mathbf{Z}^{(p)})$ of level N. Then it has three corresponding systems $\mathcal{X}$, $\mathcal{X}_{\mathbf{C}}$, $\mathcal{X}_p$ (a symmetric form of the Kronecker congruence relation), and in this case, it turns out that $\Gamma^* = (1)$ (cf. [**15**])[3] This was the key to the affirmative solution of "Conjecture Γ^*" proposed in Chapter 5 Part 2 (cf. the arguments given in §33).

[The associated differential ω] The basic reference is [**17**] (cf. also [**12**][**27**]). The special fiber $\mathbf{X}$ of $\mathfrak{X}$ defines a discrete valuation $\mathfrak{p}_K$ of the function field K, and Π, Π' also define those of K^0 extending $\mathfrak{p}_K$. Consider the $\mathfrak{p}_K$-adic completion $\widetilde{K}$ of K. Then the component Π induces an endomorphism σ of $\widetilde{K}$ which lifts the q-th power map of the residue field. Let $\widetilde{K}^{ur}$ denote the completion of the maximal unramified extension of $\widetilde{K}$, and σ^{ur} denote the unique extension of σ to an endomorphism of $\widetilde{K}^{ur}$ that induces the q-th power map of the residue field.

Theorem D (i) *There exists a differential ω of $\widetilde{K}^{ur}$, unique up to constant multiples, which is $\mathfrak{p}$-integral and whose reduction mod $\mathfrak{p}$ is not identically zero, satisfying*

$$\omega^{\sigma^{ur}}/\omega \in k_{\mathfrak{p}}^{\times}. \tag{15}$$

(ii) *In the unramified case, let S be the canonical S-operator of L (cf. Chapter 2, Part 3B) restricted to K, and S^{ur} denote the unique S-operator of $\widetilde{K}^{ur}$ that extends S. Then S^{ur} is an inner S-operator with respect to ω; namely,*

$$S^{ur}\langle\omega\rangle = 0. \tag{16}$$

A basic numerical invariant is $\nu := \mathrm{ord}_{\mathfrak{p}_K}(\xi^{\sigma^{ur}}/\xi)$, where ξ is any non-zero differential on $\widetilde{K}^{ur}$. This is clearly independent of the choice of ξ, and is also equal to the different-exponent of the valuation defined by Π in the extension K^0/K'. This ν is also related to a codimension 2 invariant. The scheme $\mathfrak{X}^0$ being normal, a formal local equation for $\mathfrak{X}^0$ at each double point $P \in \Pi \cap \Pi'$ on the special fiber can be written as $XY = \pi^{\mu_P}$ (π: a prime element) with some $\mu_P > 0$. We have $\mu_P \geq \nu$ for any P, and the equality holds if the projections are unramified on every generalization of P on the general fiber. In particular, in the unramified case, $\mu_P = \nu$ holds for any P; hence $\mathfrak{X}^0$ is regular if and only if $\nu = 1$; cf [**17**]. On the other hand, $0 < \nu \leq \mathrm{ord}_{\mathfrak{p}} q$; hence $\nu = 1$ when $O_{\mathfrak{p}} = \mathbf{Z}_p$.

[3]Incidentally, this (Lemma 3.2 of [**15**]) was later used for "lowering the levels" in Wiles [**43**] .

Now, since σ^{ur} commutes with each element of $\mathrm{Aut}(\widetilde{K}^{ur}/\widetilde{K}) \simeq \mathrm{Gal}(\mathbf{K}^{sep}/\mathbf{K})$, where $\mathbf{K} = \mathbf{F}_{q^2}(\mathbf{X})$, this Galois group acts on ω by scalar multiplications, hence a character

$$\chi : \mathrm{Gal}(\mathbf{K}^{sep}/\mathbf{K}) \longmapsto (O_{\mathfrak{p}}^{ur})^{\times}, \tag{17}$$

into the $\mathfrak{p}$-adic unit group $(O_{\mathfrak{p}}^{ur})^{\times}$ is induced. This character χ is unramified outside $\mathfrak{S}$.

In the unramified case, we can show easily that S^{ur} is the unique σ^{ur}-invariant S-operator and hence must coincide with the inner S-operator with respect to ω. This *associated differential* plays a central role in the process of construction of a system $\mathcal{X}$ from $\mathcal{X}_q$. The $(q-1)$st tensor power of the reduction mod $\mathfrak{p}$ of ω, called here $\omega_{\mathfrak{p}}$, is rational over $\mathbf{X}$, and has $2\sum_{P\in\mathfrak{S}} P$ as its divisor; in short,

$$(\omega_{\mathfrak{p}}^{\otimes(q-1)}) = 2\mathfrak{S}. \tag{18}$$

In some cases, this differential can be calculated by using the Schwarzian differential equation mod $\mathfrak{p}$, and hence also their zero set $\mathfrak{S}$ can be computed. This is a generalization of (3). The differential $\omega_{\mathfrak{p}}$ is essentially related to the first infinitesimal lifting of $\mathcal{X}_q$ (see (IX) below).

(VII) **How our Main Theorem has been proved** (The basic references are **[26][25]**.)

Theorem C in (VI) reduced our Main Theorem to the existence of a system $\mathcal{X}$ over $(O_{\mathfrak{p}}^{(2)})^{alg}$ whose base change $\otimes\mathbf{C}$ corresponds to the system $\mathcal{X}_{\mathbf{C}}$ induced from Γ. By second Galois theory, if the existence is shown for one group Γ, then it holds for any $\Gamma' \subset \Gamma$ with finite index. M.Ohta's argument in **[37]** shows that one can go the other way, too; i.e., one can pass over from a result for one Γ to any $\Gamma' \triangleright \Gamma$ with finite index. Thus, it suffices to show the existence for one Γ from each commensurability class.

I tried to find a more direct proof, but the way this has been settled is as follows. For the case of Shimura curves (cf. Chapter 4 of this Volume), Shimura's congruence relation for almost all $\mathfrak{p}$ **[39]** already shows the existence of $\mathcal{X}$ for almost all $\mathfrak{p}$ starting from discrete groups over $\mathbf{Z}$. Then Y.Morita proved the existence for individual primes $\mathfrak{p}$ not dividing the discriminant of the quaternion algebra B, by combining Shimura theory with our computation of $\zeta_{\Gamma}(u)$ given in Theorem 1 of Chapter 1, Part 1 (cf. Y.Morita **[36]**; esp. §1.2; Remark right below Main Theorem 2 [4]). He also gave proofs of some parts of (MT-1)(MT-3). And by the celebrated Margulis' arithmeticity theorem **[31][32]**, every Γ is arithmetic. We do not know whether the congruence subgroup property holds also for the quarternion modular groups over "$\mathfrak{o}_F^{(\mathfrak{p})}$", but this does not matter, because our second Galois theory is general. For more details, cf. **[20]**§6 and also **[25]** §§4-5.

A direct proof may be obtained in the following way. Let us assume the existence of $\overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$ satisfying the condition stated in (IV). Let γ be any negative (resp. positive) element in the stabilizer Γ_z. Then its action on the tangent space at z reduces mod $\mathfrak{p}$ to 0 (resp. ∞). This should imply the *inseparability* at "$z(\mathrm{mod}\,\mathfrak{p})$" of one of the two projections of

[4]The case when the quaternion algebra B is over $F = \mathbf{Q}$ goes back to his Master's thesis (University of Tokyo, 1970). I also understand that the general case was done around 1973 while he was staying at IAS, Princeton. This general case is much more difficult, because B is *not* totally indefinite and the relation with the moduli problem is more complicated. Cf. also Langlands **[30]** (totally indefinite case, but includes higher dimensional cases, more adelic and not using discrete subgroups over $\mathbf{Z}^{(p)}$).

the correspondence defined by the system $\mathcal{X}_{\mathbf{C}}$, where the base field is suitably lowered to a number field. To use this more systematically, observe that the action of each of V, V' on the p-adic projective line is transitive while that of V^0 has two distinct orbits. Use this orbital decomposition for the "p-part" of z. This aims at a non-abelian analogue of the method used in Shimura-Taniyama theory for proving the prime ideal decomposition of the Frobenius endomorphism. There, a crucial point lies on looking at the action of the "π-multiplication" on holomorphic differentials, reducing it modulo a prime and showing that the reduced endomorphism must be inseparable. In our case, I started with a minimal regular model of the base curve, and tried to proceed analogously, appealing also to the purity of branch locus, but was able to obtain only a partial result. With modern tools at hand, someone else may be able to cultivate this method fully.

(VIII) **An additional conjecture** (cf. [25]) First, let us note the following. If $\mathcal{X}$ is any system (12) over $O_{\mathfrak{p}}^{(2)}$ for some p-adic field $k_{\mathfrak{p}}$ which reduces mod p to the system $\mathcal{X}_q$ ($q = \mathrm{N}(\mathfrak{p})$), and $k'_{\mathfrak{p}}/k_{\mathfrak{p}}$ is any totally ramified extension, then the corresponding base change of $\mathcal{X}$ also reduces to the same $\mathcal{X}_q$. What would characterize the *minimal* choice of the base ring is the *regularity* of the middle scheme $\mathfrak{X}^0$, which is equivalent to the condition $\nu = 1$ when $\mathcal{X}$ is of unramified type. Note that the other two schemes are smooth and hence always regular. The following conjecture still remains open.

Conjecture *Torsion-free cocompact irreducible lattices Γ of $G = G_{\mathbf{R}} \times G_{\mathfrak{p}}$, and unramified symmetric systems $\mathcal{X}$, consisting of regular schemes over $O_{\mathfrak{p}}^{(2)}$ whose reduction is of the form $\mathcal{X}_q$, are categorically equivalent. The equivalence functor depends on the choice of $\overline{\mathbf{Q}}_{\infty,\mathfrak{p}}$.*

Starting from $\mathcal{X}$ we have constructed Γ, as explained in (VI) except that we have not shown $G_{\mathfrak{p}}^{\star} \subseteq PGL_2^{+}(k_{\mathfrak{p}})$. We know $G_{\mathfrak{p}}^{\star} \subseteq \mathrm{Aut}(\mathcal{T}_{\mathfrak{p}})$, but in general without the unramifiedness condition, the image can be essentially bigger than $PGL_2^{+}(k_{\mathfrak{p}})$. The problem here is to find a natural p-adic projective linear representation of the group $G_{\mathfrak{p}}^{\star}$ in the unramified case. Of course, using arithmeticity of Γ, one should be able to say much more about the image, but we want to find a natural representation. The naming of the title "congruence monodromy problem" was initially motivated by the thought that this is probable. (The arithmeticity was then "half-suspected".) The study of canonical S-operator and its connection with ω may give some hint.

Starting from Γ, as explained in (VII), $\mathcal{X}$ was constructed, except that we do not know in general whether the schemes are regular. Only when $k_{\mathfrak{p}} = \mathbf{Q}_p$, we must have $\nu = 1$ and hence the regularity is automatically satisfied.

(IX) **Which systems $\mathcal{X}_q$ are liftable to $\mathcal{X}$?** ([19][23]) Starting from a given system $\mathcal{X}_q$, or equivalently, starting from a pair $(\mathbf{X}, \mathfrak{S})$, we ask ourselves whether there exists a system $\mathcal{X}$ over some (or given) p-adic ring of integers, or more generally, ask what can be said about the existence and the main properties of "the universal deformation" of $\mathcal{X}_q$.

In [19], we studied each infinitesimal step of the lifting, and have shown that the associated differential plays a fundamental role. In particular, assume $O_{\mathfrak{p}} = \mathbf{Z}_p$. Then, for

a given lifting

$$\{\mathfrak{X}_n \leftarrow \mathfrak{X}_n^0 \rightarrow \mathfrak{X}_n'\}$$

of X_q over $O_{\mathfrak{p}}^{(2)}/\mathfrak{p}^n$ $(n \geq 1)$, we have proved that its liftings mod $\mathfrak{p}^{n+1}$ correspond bijectively with the differentials ω_n on some (generically unramified) covering of $\mathfrak{X}_n$ satisfying some local conditions above each point of $\mathfrak{S}$. The local condition is an obvious necessary condition. At first I did not understand how to patch local liftings together, but then I realized that the associated differential is *the* natural "globe" on which this patch work should be done. Here, of course, local or global refers to the geometic part and not to the base ring. In particular, the first infinitesimal step is well-understood, including the case of the liftings to $\mathbf{F}_p[\epsilon]$ $(\epsilon^2 = 0)$.

In [**23**], the existence of the universal deformation of X_q over a complete noetherian local $W(\mathbf{F}_q)$-algebra R is proved and some properties of R is studied. For example, corresponding to (13), we have $\dim R = 0$ if $|\mathfrak{S}| < (q-1)(g-1)$, and on the other hand, if $|\mathfrak{S}| \geq 2(q-1)(g-1)$, then $\dim R \geq 1$, which implies that either X_q has a lifting over $O_{\mathfrak{p}}^{(2)}$ for some $k_{\mathfrak{p}}$, or has a non-trivial deformation over $\mathbf{F}_{q^2}[[t]]$. Some sufficient conditions for the liftability to $\mathbf{Z}_p$, uniqueness of symmetric liftings in the unramified case, and some numerical examples are also given in these articles. (For the case $g = 2$, see also [**10**].)

But we had only started; I hope that it will be continued on by someone.

[**Problems**] (i) Can one make use of the $\mathfrak{p}$-adic character χ described above in the study of liftings? (ii) Can one characterize χ classfield theoretically in the elliptic modular case? (In this case, the kernel of χ corresponds to Igusa's modular tower, and the differential ω, to $\mathrm{d}\log q$, where q is "Tate's q = Dwork's q".)

(X) Additional Comments

[**Curves over $\mathbf{F}_{q^2}$ with many rational points**] ([**15**][**20**][**24**][**25**]; see also Tsfasman-Vladut-Zink [**42**] and Drinfeld-Vladut [**8**]) [5] One of my starting points (1965) was the following observation. The reduction mod $\mathfrak{p}$ of a typical Shimura curve has a "$\mathfrak{p}$-canonical" model over $\mathbf{F}_{q^2}$ $(q = \mathrm{N}(\mathfrak{p}))$ that contains $(q-1)(g-1)$ special rational points. Since this is for the whole system, these points should decompose completely in the system of higher level coverings. This was observed first from the computation of the zeta function of Γ. Earlier, after having worked with the discrete subgroups of $PGL_2(k_{\mathfrak{p}})$, I was looking for discrete subgroups whose Selberg type zeta function is closer to a congruence zeta function of a curve. The extra factor $(1-u)^{(q-1)(g-1)}$ of $\zeta_\Gamma(u)$ suggested the existence of these special points[6]. As is well-known, the discovery of Goppa codes has drawn a great deal

[5] Here, only the older references are given. This is mainly because the later important papers on this subject are (numerous and) well-known, but also because of frequent confusions related to less-known "history". In terms of the quantity "$A(q) = \limsup_{C/\mathbf{F}_q} N(C)/g(C)$", the historical order is: first, the inequality $A(q) \geq \sqrt{q} - 1$ (q: square) was recognized and proved, then $A(q) \leq \sqrt{2q}$ (general q; [**24**]... which showed that it can be essentially better than the Weil bound), and then, by refining this idea, a decisive inequality $A(q) \leq \sqrt{q} - 1$ (general q) was proved [**8**].

[6] As is pointed out in my earlier article ([15] in the (main) Bibliography), the exponent $(q-1)(g-1)$ is the multiplicity in $L^2(G/\Gamma)$ of the tensor product of the "special representations" of $G_{\mathbf{R}}$ and of $G_{\mathfrak{p}}$. Further study of the zeta function of Γ using spectral decomposition has been pursued by J.P.Labesse [**28**].

of attention to curves with many rational points, and then significant progress has been made by several mathematicians.

[**Changing quaternions**] An unexpected development of our study of $G_{\mathfrak{p}}$-fields was the work of Čerednik [2] [3] on the comparison of Shimura curves at a prime $\mathfrak{q}$ dividing the discriminant of B/F with Mumford curves associated with discrete subgroups of $PGL_2(F_{\mathfrak{q}})$. The latter arises from the quaternion algebra B' obtained from B by replacing the ramified prime $\mathfrak{q}$ by the unique unramified archimedean prime.

[**Changing levels, Modularity**] Another unexpected direction of the use of a lemma in [15] (lemma 3.2) is in Ribet [38] and Wiles [43]. A conjectural (representation theoretic) generalization and its important consequences are given in Clozel-Harris-Taylor [4] (cf. also Diamond-Taylor [5]).

[**More recent new developments**] Among other recent works in which I was able to see some (internal) connections, I would like to mention, especially:

The work [6] of Darmon,

Mochizuki's theory (cf. e.g. [34][35]).

Cf. also, [1][9][41].

[**Left to the future....**] The most crucial question would be: "Is there any *other* system that also possesses an arithmetic fundamental group in such a strong sense as above?". Especially, one asks:

"Does there exist an infinite Galois extension $\mathfrak{K}/K$ over a number field and a discrete group Γ equipped with two embeddings:

$$i : \Gamma \longrightarrow \widehat{\Gamma} \simeq \mathrm{Gal}(\mathfrak{K}/K),$$

$$j : \Gamma \longrightarrow G \supset T \simeq (\mathbf{R}^+)^{\times},$$

such that G is some locally compact group, $j(\Gamma)$ is a lattice in G, and that some analogy of the case $G = PSL_2(\mathbf{R}) \times PGL_2^+(k_{\mathfrak{p}})$ holds when $PSO_2(\mathbf{R}) \times PGL_2^+(k_{\mathfrak{p}})$ is replaced by T?"

This would in particular imply that the Dedekind zeta function $\zeta_K(s)$ of K is essentially equal to the Selberg type zeta function $\zeta_{\Gamma,T}(s)$. The extension $\mathfrak{K}/K$ must be almost unramified, and the Galois group must be essentially non-abelian, because only the prime powers should possess Frobenius elements. In order to understand the zeta functions having the Euler product, such as $\zeta_K(s)$, better, it is of course indispensable to find a structure on the set of all prime powers of K. It was encouraging that in our case, the prime powers are those Γ-conjugacy classes that are conjugate in G to an element of T.

This has been my dream for 40 years, but has not been realized at all. Some colleagues may have heard about this all too often. But I hope it will keep on living modestly somewhere in someone's mind...

2008 Summer Yasutaka Ihara