Travaux mathématiques

-

Special issue
based on the
Winter School on Galois Theory
Luxembourg, February 2012
Volume II

-

Editors
Sara Arias-de-Reyna
Lior Bary-Soroker
Gabor Wiese

-

Vol. XXIII, 2013

UNIVERSITÉ DU LUXEMBOURG

Faculté des Sciences,
de la Technologie
et de la Communication

Travaux mathématiques

Volume XXIII, 2013

Published by the
UNIVERSITY OF LUXEMBOURG

ISBN: 978-2-919940-09-7
ISSN: 1024-1833

UNIVERSITY OF LUXEMBOURG
Campus Kirchberg
Mathematics Research Unit
6, rue Richard Coudenhove-Kalergi
L-1359 Luxembourg
Grand-Duchy of Luxembourg

Phone: (+352) 466 644 6376
Fax: (+352) 466 644 6377
Email: carine.molitor@uni.lu,
norbert.poncin@uni.lu

Travaux mathématiques

Presentation

The journal *Travaux mathématiques* is published by the Mathematics Research Unit of the University of Luxembourg. Even though the main focus of the journal is on original research articles, surveys and historical studies are also welcome.

Editors-in-Chief

Address:
University of Luxembourg
Mathematics Research Unit
6, rue Coudenhove-Kalergi
L-1359 Luxembourg
Grand-Duchy of Luxembourg
Email: carine.molitor@uni.lu or *norbert.poncin@uni.lu*

Associate Editors

Editorial Board

Travaux mathématiques

Special issue
based on the
Winter School on Galois Theory
Luxembourg, February 2012
Volume II

Editors
Sara Arias-de-Reyna
Lior Bary-Soroker
Gabor Wiese

Vol. XXIII, 2013

Faculté des Sciences,
de la Technologie
et de la Communication

Volume II of the
Proceedings of the
Winter School on

Galois Theory

15 – 24 February 2012

Université du Luxembourg
Luxembourg

Edited by

Sara Arias-de-Reyna
Lior Bary-Soroker
Gabor Wiese

Proceedings of the Winter School on Galois Theory (Volume II)
Université du Luxembourg, 15 – 24 February 2012
Luxembourg

Organisers

Sara Arias-de-Reyna
Lior Bary-Soroker
Gabor Wiese

Sponsors

Fonds National de la Recherche Luxembourg, Luxembourg
Foundation Compositio Mathematica, The Netherlands
DFG-Sonderforschungsbereich TRR 45, Germany
University of Luxembourg, Mathematics Research Unit, Luxembourg

Preface

The work of Evariste Galois (1811 – 1832), whose 200th birthday was recently celebrated, is at the origin of a theory that plays a key role in many mathematical disciplines, such as number theory, algebra, topology, and geometry.

From 15 until 24 February 2012 we organised a Winter School on Galois Theory at the University of Luxembourg, which attracted almost 100 participants. In the winter school two very important topics were united: on the one hand, in recent years there have been tremendous developments in number theory, such as the proofs of two famous conjectures: Serre's Modularity Conjecture and the Sato-Tate Conjecture. It is becoming clear that these are only some instances of a bigger picture encompassed in new directions in the Langlands programme. *Galois representations* are at the core of all of this.

On the other hand, in *Galois field theory*, motivated by the famous inverse Galois problem first considered by Hilbert, the so-called regular inverse Galois problem could be solved over various classes of fields. This has been done by developing "patching", an algebraic analog of the "cut-and-paste" constructions in complex geometry. The method of patching has been further developed and expanded and achieved results both in Galois theory and algebra, e.g., in differential Galois theory, division algebras and quadratic forms.

The communities of people working on Galois representations and those working in Galois field theory currently only seem to be insufficiently interacting, although for a large part they are studying the same kind of objects, namely absolute Galois groups. One instance, where the possible connection was successfully exploited, is in the proof of potential modularity: a classical result of field arithmetic on the field of all totally real numbers is essential.

Both groups use their own very advanced sets of tools. On the one hand, at the winter school the statement of Serre's Modularity Conjecture, some of the tools involved in the proof of it, as well as some generalisations and connected open problems were presented. On the other hand, several of the patching techniques and various applications were explained.

The winter school consisted of three 'preparatory days', whose content was accessible to Master students, whereas the main week of the school addressed mainly PhD students or young postdocs. The current Volume II unites four sets of lecture notes from the main week, two concerning Galois representations, two about patching, and one research article.

Gebhard Böckle's contribution is a quite comprehensive survey on Galois representations, touching on as diverse subjects as class field theory, Galois representations attached to automorphic forms, Hilbert modular forms, compatible systems of Galois representations, Weil-Deligne representations, p-adic Hodge theory, the Fontaine-Mazur Conjecture, the Sato-Tate Conjecture, modularity proofs and Fermat's Last Theorem, and many more. The survey focusses on the key ideas and the long list of recommended references enables the reader to pursue himself/herself any of the mentioned topics in greater depth.

Michael Schein's notes sketch the proof due to Khare and Wintenberger (building on the work of many mathematicians) of one of the major theorems in arithmetic algebraic geometry in recent years, namely Serre's Modularity Conjecture. It states a precise parametrisation of all Galois representations of a certain type through modular forms. The notes also mention generalisations of Serre's Modularity Conjecture to totally real fields and even contain a section on the mod-p Langlands programme, which provides a natural framework for this kind of questions.

Moshe Jarden's notes are based on his book on algebraic patching. They develop the method of algebraic patching from scratch and give applications in contemporary Galois theory. The notes start from the patching axioms, proceed to examples of families of fields satisfying the axioms, like complete fields, fraction fields of power series rings, and more generally ample fields, and finish with the solution of the regular inverse Galois problem over such fields, and more generally the solution of split embedding problems. The notes culminate in a proof of a theorem of Pop saying that a Hilbertian pseudo algebraically closed field has ω-free absolute Galois group.

David Harbater's notes are complementary to Jarden's notes, and describe recent applications of patching in other aspects of algebra, for example: differential algebra, local-global principles, quadratic forms, and more. The notes (as well as the lectures) focus on the big picture and on providing the reader with intuition. The notes also contain a historical discussion on the development of patching, surveying the connection between the different patching methods.

In their research article Wulf-Dieter Geyer and Moshe Jarden present a theorem of Kollár on the density property of valued pseudo algebraically closed fields and a theorem of Abraham Robinson on the model completeness of the theory of algebraically closed non-trivially valued fields. Then they prove that the theory T of nontrivial valued fields in an appropriate first order language has a model completion $\tilde{T}$ and give an explicit algebraic presentation of $\tilde{T}$.

We wish to express our sincere gratitude to all speakers at the winter school for their excellent lectures, and for having taken a lot of time and care in elaborating their notes for these proceedings.

May the readers enjoy them and may they serve well!

The Editors

Sara Arias-de-Reyna
Lior Bary-Soroker
Gabor Wiese

July 2013

Travaux mathématiques, Volume 23 (2013), 5–35, © by the author

Galois representations

by Gebhard Böckle

Abstract

The present notes develop some basic language on Galois representations and strictly compatible families of such. As guiding examples we use elliptic curves, (Hilbert) cusp forms and the (partially conjectural) relation between the two. The notion of strict compatibility in families at ramified primes is used to motivate Weil-Deligne representations and p-adic Hodge theory à la Fontaine. The notes end with a rough sketch of how to use the concepts presented to give a proof of the modularity of elliptic curves over $\mathbb{Q}$. The emphasis throughout the notes is to explain the interrelations and usefulness of the concepts covered. Proofs are mostly omitted, but many references to the (vast) literature are given.

These notes are a slightly expanded version of parts of a lecture series at the Luxembourg Winter School 2012, organized by Gabor Wiese, Lior Bary-Soroker and Sara Arias-de-Reyna. The notes claim no originality.

MSC (2010): 11F80, 11F33, 11F41, 11F85, 11G05, 11S20, 11S37.

Contents

1 Introduction

The main focus of the present notes are Galois representations, which are an important tool in modern number theory: Galois representations provide a convenient and concrete way to represent a geometric object over a number field. Via these representations, the geometric object tells us much about quotients of the absolute Galois group of its underlying number field. Moreover Galois representations provide one with a tool to compare geometric objects: in the simplest form, such a comparison could be that two objects are related if they have isomorphic Galois representations. Already this very basic notion has deep consequences, as we shall see.

These notes do not wish to give an axiomatic treatment. They will not explain what a general "geometric object" (a motive) could mean. They focus on examples, mainly those of elliptic curves and (Hilbert) modular forms. Each example will provide us with a family of Galois representations, one member for each of the primes of the coefficient field of the object. This motivates us not only to describe the individual members, but also to place the families within the setting of compatible systems of Galois representations. We proceed in several steps:

Section 2 locates Galois representations within number theory, indicates some applications of Galois representation from automorphic forms, and fixes some notation. In Section 3 we describe, following our two main examples, a notion of compatible system that only uses unramified primes. This is enough to characterize the representation up to isomorphism. In practice it is also very useful to have a notion of compatibility at the ramified primes. This requires Weil-Deligne representations, which we introduce in Section 4. After this, the only primes that are not included into a notion of compatibility are those that divide the residue characteristic of the coefficient field of the Galois representation. This needs p-adic Hodge theory, of which we give some flavor in Section 5. The following Section 6 surveys some standard predictions of the Fontaine-Mazur conjectures which try to answer the question which Galois representations one can hope to find in geometry. The final Section 7 indicates the role of Galois representations in modularity proofs of elliptic curves – such as the proof of Fermat's Last Theorem by Wiles and Taylor-Wiles – a theme repeatedly discussed throughout the text. The material is supplemented by an extensive bibliography, to which we give detailed references in the main text.

Acknowledgements: Many thanks go to Sara Arias-de-Reyna, Hendrik Verhoek, Gabor Wiese for comments, corrections and suggestions. I would also like to thank the referee for many corrections and helpful suggestions.

2 Number theory and Galois representations

2.1 The absolute Galois group of a number field

Basic question: Describe all Galois extension of a number field of a certain type. **But** this is too difficult!

Much of current day number theory is concerned with understanding extensions E/F of a number field F and their ramification properties. In applications one is mostly concerned with the case that E/F is Galois. Since the absolute Galois group $G_F = \mathrm{Gal}(\bar{F}/F)$ is profinite, it suffices to understand all finite Galois extensions, although it is often useful to consider profinite extensions. To understand the ramification properties it is also important to understand the absolute Galois group of a local field. This is considerably simpler and for $p > 2$ it is actually solved by Koch and Jannsen-Wingberg, see [JW82].

2.2 Abelian Class Field Theory versus the Langlands program

The first main success in understanding Galois extensions of number fields is abelian class field theory. It gives a complete classification of all abelian extensions and their ramification properties.

Abelian class field theory by itself is rarely constructive. Over $\mathbb{Q}$ the cyclotomic extensions generate $\mathbb{Q}^{\mathrm{ab}}$. Over CM fields one can consider CM abelian varieties and the Galois representations attached to their torsion points. The theory over quadratic imaginary fields can be found in [Si91]. For the general CM case, see [Sh98]. Beyond this little is known. For certain conjectural constructions, see [Da04].

Beginning in the late 1960's, mainly due to Langlands a new approach developed. The abelian case was considered as the GL_1 case. Langlands idea was to use automorphic forms and representations to develop a class field theory for GL_n. Automorphic representations for a reductive group G over F should give rise to Galois representations into the dual group of G and in a vague sense all such Galois representations should come from automorphic representations and thus from modular forms. In fact to get modular forms, one needs to require some algebraicity of the automorphic representations. But this would get us too far from the topic. The so far most successful case is the group GL_2 over totally real fields. Other cases beyond these lectures are unitary groups (i.e., inner forms of GL_n) and symplectic groups. References are [AEK03], [BC09], [La03].

The Langlands program is constructive whenever one has an algebraic theory of automorphic forms. This seems to lead to a similar constraint as above: The automorphic forms have to be defined over a totally real or a CM field.

Another important source of Galois representations is étale cohomology. In fact, in all cases above, one uses étale cohomology to construct Galois representations from some geometric data. A priori étale cohomology seems to give much more representations than the theory of automorphic forms. But in the end, one might hope that all geometric semisimple Galois representations come from automorphic forms; see [Bel09].

For time constraints, I shall say nothing about the automorphic side; local Langlands correspondence, compatibilities, etc.

2.3 Applications of Galois representations from automorphic forms:

Much of the following developments go back to Wiles work on Fermat's Last theorem (FLT):

(a) The Taniyama-Shimura conjecture, which is a theorem by Breuil-Conrad-Diamond-Taylor [BCDT], states that every elliptic curve over $\mathbb{Q}$ is modular.

(b) By results of Frey, Ribet and Serre [Ri90, Se87](proved in the 80's) this implies FLT.

(c) Wiles proof of FLT established sufficiently many cases of Taniyama-Shimura to deduce FLT; [Wi95, TW95], of which good surveys are [DDT97] and [CSS].

(d) For elliptic curves A over an arbitrary totally real number field F, it follows that if they are modular, then their L-function has an entire continuation to the complex plane. Using potential modularity, Taylor has proved results that show in many cases that the L-function of A/F has a meromorphic continuation to $\mathbb{C}$. See for instance [Sn09, Tay06].

(e) If one could show that $\operatorname{Sym}^n A$ is modular for all $n \in \mathbb{N}$ and all elliptic curves A/F without CM, then the Sato-Tate conjecture on the deviation of $\#A(k_v)$ from $\#k_v + 1$ follows for A. One cannot quite show that. But one can prove potential versions of this which suffice to prove the conjecture: [BLGG, CHT, HSBT, Tay08].

Here is the heuristic for the Sato-Tate distribution: Let X be the set of conjugacy classes of elements in $\mathrm{SU}_2(\mathbb{C})$. A representative of a conjugacy class is of the form $\left(\begin{smallmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{smallmatrix}\right)$. Therefore conjugacy classes can be considered as elements $\theta \in \mathbb{R}/\pi\mathbb{Z}$. Consider

$$\mathrm{SU}_2(\mathbb{C}) \to X \cong [0, \pi] \overset{-\cos}{\cong} [-1, 1].$$

The pushforward of the measure of equi-distribution on $\mathrm{SU}_2(\mathbb{C})$ yields the measure with distribution $\frac{2}{\pi} \sin^2 \theta \mathrm{d}\theta$ on $[0, \pi]$, or $\frac{2}{\pi}\sqrt{1 - t^2}\mathrm{d}t$ on $[-1, 1]$.

The ST-conjecture asserts that for $\ell \to \infty$ (or $\lambda \to \infty$), the distribution of the numbers

$$\frac{a_\ell(A)}{2\sqrt{\ell}} \in [-1, 1]$$

(for an elliptic curve A without complex multiplication), and of the numbers

$$\frac{a_\lambda(f)}{2\lambda^{(k-1)/2}}$$

(for a modular form f of weight k and without complex multiplication) converges to a measure on $[-1, 1]$ with the ST-distribution

$$\frac{2}{\pi}\sqrt{1-t^2}\mathrm{d}t.$$

In one lecture of the course notes [Ha07], by Michael Harris, on the proof of the Sato-Tate conjecture by Clozel, Harris, Shepherd-Barron and Taylor, the relation between the Sato-Tate conjecture and the meromorphy of the L-function of $\operatorname{Sym}^n A$, $n \in \mathbb{N}$ is explained.

2.4 Some notation

(a) For a number field F fix an algebraic closure $\bar{F} = \bar{\mathbb{Q}}$.

(b) For a place v, i.e., an equivalence class of norms on F, let F_v be the completion of F at v and fix an algebraic closure $\bar{F}_v$.

(c) If v is non-archimedean, define $\mathcal{O}_v$ as the ring of integers of F_v, π_v as a uniformizer, k_v as the residue field at v and $q_v := \# k_v$ is the order of k_v.

(d) Fix an embedding (i.e., an F-algebra homomorphism) $\bar{F} \hookrightarrow \bar{F}_v$. This yields a homomorphism of Galois groups $G_v := G_{F_v} \to G_F$ (known to be injective) from the diagram

$$\begin{tikzcd}
\bar{F} \arrow[r] \arrow[d, dash, "\mathrm{Gal}(\bar{F}/F)=:G_F"'] & \bar{F}_v \arrow[d, dash, "G_{F_v}=:G_v"] \\
F \arrow[r] & F_v
\end{tikzcd}$$

(e) For a set of places S of F, denote by $G_{F,S}$ the quotient of G_F that is the Galois group of the maximal extension of F in $\bar{F}$ unramified outside S.

(f) For v a place not in S, fix a Frobenius automorphism $\mathrm{Frob}_v \in G_{F,S}$ which is unique up to conjugation. (The kernel of $G_v \to G_{F,S}$ contains the inertia subgroup of G_v and $G_v/I_v \cong G_{k_v}$ which in turn is generated by the Frobenius. We take the geometric Frobenius.)

3 L-functions and Galois representations

3.1 Elliptic curves

Let A be an elliptic curve over a number field F.

Let N be the conductor of A. It is defined as a product of local conductors.

The latter are 1 at every place of good reduction of A. If A does not have good reduction at v, then the prime corresponding to v divides N.

Definition 3.1. (a) For a prime ℓ, denote by $\mathrm{Ta}_\ell(A)$ the G_F representation on $\varprojlim A[\ell^n](\bar{F})$ and write $V_\ell(A)$ for $\mathrm{Ta}_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$.

(b) The L-factor of A at v is $L_v(A, T) := \det(1 - T\mathrm{Frob}_v | V_\ell(A)^{I_v})^{-1}$

(c) The L-function of A is

$$L(A, s) := \prod_{v \,\text{finite}} L_v(A, q_v^{-s}) \text{ for } \Re(s) \gg 0.$$

It is not clear that the L-factors are independent of ℓ. So in principle ℓ should occur in the above notation. However part (b) of the following result clarifies this problem and shows the independence of ℓ, assuming that $v \nmid \ell$.

Theorem 3.2 (see [Si85], [Si91]). *(a) If $v \nmid \ell$, then the representation $V_\ell(A)$ is ramified at v if and only if v divides the conductor N (Theorem of Néron-Ogg-Shafarevich).*

(b) If $v \not| N\ell$, then $L_v(A, T)^{-1} = 1 - a_v(A)T + q_vT^2$ where $L_v(A, 1)^{-1} \overset{!}{=} \#A(k_v)$ defines $a_v(A)$.

(c) If v divides N but not ℓ, then $L_v(A, T)$ can be computed by Tate's Algorithm.

(d) The L-function defined above converges for all $s \in \mathbb{C}$ with $\Re(s) > 3/2$.

(e) The representation $V_\ell(A)$ is semisimple.

3.2 On traces and characteristic polynomials

Lemma 3.3. *Let Π be a profinite group, let $\mathcal{F} \subset \Pi$ be a subset such that Π is the topological closure of the conjugacy classes of $\mathcal{F}$, and fix a positive integer n. Then any continuous* semi-simple *representation $\rho\colon \Pi \to \mathrm{GL}_n(L)$ where $L \in \{\mathbb{C}, \overline{\mathbb{F}_\ell}, \overline{\mathbb{Q}_\ell}\}$ is uniquely determined by the characteristic polynomials* $\mathrm{charpol}(\rho(g)) \in L[T]$ *for $g \in \mathcal{F}$.*

For $\mathbb{C}$ this is classical representation theory, for $\overline{\mathbb{F}_\ell}$ this follows from the theorem of Brauer-Nesbitt, see [CR62, 30.16]. A proof for $\overline{\mathbb{Q}_\ell}$ is in [Tay91].

Theorem 3.4. *Given $E/\mathbb{Q}_\ell$ finite and V a finite dimensional continuous linear G_F representation over E. Let k_E be the residue field of E, i.e. $k_E = \mathcal{O}_E/\mathfrak{m}_E$.*

(a) If V is semisimple, then there is a set S of density zero outside of which V is unramified.

(b) In the situation of (a), one has $\det(1 - T\mathrm{Frob}_v|V) \in \mathcal{O}_E[T]$ *for all $v \notin S$ and ρ is completely determined by these characteristic polynomials, or even the traces of* $\rho(\mathrm{Frob}_v)$, *$v \notin S$.*

(c) There exists a unique continuous semisimple G_F-representation $\bar{V}$ with

$$\det(1 - T\mathrm{Frob}_v|\bar{V}) \equiv \det(1 - T\mathrm{Frob}_v|V) \mod \mathfrak{m}_E \text{ in } k_E[T] \quad \forall v \notin S.$$

Proof. (a) See [KR01]. (b) Follows from the existence of a G_F-stable lattice (which is deduced from the compactness of G_F) and Lemma 3.3. (c) One reduces the lattice from (b), semisimplifies the reduction and applies Lemma 3.3. □

Corollary 3.5. *Let A/F be an elliptic curve. Then $V_\ell(A)$ is completely characterized by the condition* $\mathrm{Trace}(\mathrm{Frob}_v|V_\ell(A)) = a_v(A)$ *for all finite places v not dividing $N\ell$*

3.3 Hilbert modular forms

The following definitions are from [Tay89]. General texts on Hilbert modular forms are [BGHZ], [Fr90], [Go02], [Hi06] and references therein.

Let F be a totally real number field. Let I be the set of embeddings $F \hookrightarrow \mathbb{R}$. Denote by $\mathbb{A}_F$ the adele ring of F. Write $\mathbb{A}_F = \mathbb{A}_f \times \mathbb{A}_\infty$ for the decomposition into the finite and infinite adeles. Fix $k = (k_\tau) \in \mathbb{Z}^I$ such that $k_\tau \geq 2$ for each component. and suppose that all components have the same parity.[1] Set $t = (1, \ldots, 1) \in \mathbb{Z}^I$, and set $m = k - 2t$. Also choose $v \in \mathbb{Z}^I$ such that each $v_\tau \geq 0$, some $v_\tau = 0$ and $m + 2v = \mu t$ for some $\mu \in \mathbb{Z}_{\geq 0}$. Set $\mathfrak{h} := \{z \in \mathbb{C} \mid \Im(z) > 0\}$.

For $f\colon \mathrm{GL}_2(\mathbb{A}_F) \to \mathbb{C}$ and $u = u_f u_\infty \in G_f \times G_\infty = \mathrm{GL}_2(\mathbb{A}_F)$ define

$$(f|_k u)(x) := j(u_\infty, \underline{i})^{-k} \det(u_\infty)^{v+k-t} f(xu^{-1})$$

where:

- $\underline{i} = (\sqrt{-1}, \ldots, \sqrt{-1}) \in \mathfrak{h}^I$;
- $j\colon G_\infty \times \mathfrak{h}^I \to \big(\mathbb{C}^*\big)^I, \Big(\left(\begin{smallmatrix} a_\tau & b_\tau \\ c_\tau & d_\tau \end{smallmatrix}\right)_{\tau \in I}, \big(z_\tau\big)_{\tau \in I}\Big) \mapsto (c_\tau z_\tau + d_\tau)_{\tau \in I}$;
- $(\alpha_\tau)^{(n_\tau)} := \prod_{\tau \in I} \alpha_\tau^{n_\tau}$ for $(\alpha_\tau) \in (\mathbb{C}^*)^I$ and $(n_\tau) \in \mathbb{Z}^I$.

Definition 3.6. For $U \subset G_f$ a compact open subgroup one defines the space of Hilbert modular cusp forms $S_k(U)$ of level U and weight k to be the set of functions $f\colon \mathrm{GL}_2(F)\backslash\mathrm{GL}_2(\mathbb{A}_F) \to \mathbb{C}$ satisfying the following conditions:

(a) $f|_k u = f$ for all $u \in UZ_\infty$ where $Z_\infty = \big(\mathbb{R}^* \cdot \mathrm{SO}_2(\mathbb{R})\big)^I \subset G_\infty$;

(b) for all $x \in G_f$, the function $f_x\colon \mathfrak{h}^I \to \mathbb{C}$ defined by

$$uz_0 \mapsto j(u_\infty, z_0)^k \det(u_\infty)^{t-v-k} f(xu)$$

for $u \in G_\infty$ is well-defined and holomorphic;

[1] In [Hi06], Hida explains after formula (2.3.9) why without the parity condition, the space of Hilbert modular forms is zero – he uses a different but equivalent formalism, in which this statement can be formulated more meaningfully.

(c) $\int_{\mathbb{A}_F/F} f\left(\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right)x\right)\mathrm{d}a = 0$ for all $x \in \mathrm{GL}_2(\mathbb{A}_F)$ and $\mathrm{d}a$ an additive Haar measure on $\mathbb{A}_F/F$.

Depending on the choice of level, one can also talk about a nebentype character. In any case, up to conjugation of U, there is a largest ideal N of $\mathcal{O}_F$ such that $U \supset \{g \in G_f \mid g \equiv 1 \pmod{N}\}$. We call N the level.

Definition 3.7. For $x \in G_f$ one defines the Hecke operator $[UxU]$ for a function $f\colon \mathrm{GL}_2(\mathbb{A}_F) \to \mathbb{C}$ as

$$[UxU]f := \sum_i f|_{x_i}$$

where $UxU = \coprod_i Ux_i$. This assignment is well defined for $f \in S_k(U)$ and defines an endomorphism in $\mathrm{End}_{\mathbb{C}}(S_k(U))$.

In the special case $x = \left(\begin{smallmatrix} \pi_v & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with v not a divisor of N, one abbreviates $T_v := [UxU]$.

For $x = \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha \end{smallmatrix}\right)$ with $\alpha = (\pi_v^{\mathrm{ord}_v(\mathfrak{a})})_v$ for a fractional ideal $\mathfrak{a}$ of F prime to N one calls $S_{\mathfrak{a}}$ the diamond operator for $\mathfrak{a}$. If $\mathfrak{a}$ is a prime ideal corresponding to the place v we simply write S_v.

Moreover, we define $\mathbb{T}_k(U)$ as the $\mathbb{Z}$-subalgebra of $\mathrm{End}_{\mathbb{C}}(S_k(U))$ generated by the T_v and S_v for all $v \nmid N$.

A cusp form f of weight k and level U is called an eigenform for $\mathbb{T}_k(U)$ if it is a simultaneous eigenvector for all T_v, S_v with $v \nmid N$. The eigenvalues are denoted by $a_v(f)$ and $\chi_v(f)$, respectively.

Since $\mathbb{T}_k(U)$ is commutative, eigenforms exist.

Theorem 3.8. *Let $f \in S_k(U)$ be a Hecke eigenform and write $a_v(f)$ for the eigenvalue under T_v for all v not dividing the level N of U.*

(a) The coefficient field *$E_f := \mathbb{Q}(a_v(f) \mid v \nmid N)$ is a finite extension of $\mathbb{Q}$. All $a_v(f)$, $v \nmid N$, are integral.*

(b) For any prime λ of the ring of integers $\mathcal{O}_f$ of E_f, there exists a unique[2] *continuous representation $V_\lambda(f)$ (isomorphic to $E_{f,\lambda}^2$), say*

$$\rho_{f,\lambda}\colon G_F \to \mathrm{GL}_2(E_{f,\lambda})$$

which is unramified outside $N\ell$ and satisfies

$$\det(1 - T\mathrm{Frob}_v|V_\lambda(f)) = 1 - a_v(f)T + \chi_v(f)q_vT^2 \textit{ for all } v \nmid N\ell$$

where ℓ is the rational prime under λ and $E_{f,\lambda}$ is the completion of E_f at λ.

[2] See Remark 3.3 and Theorem 3.10(d).

Part (a) is due to Shimura; see [Sh71, Thm. 3.48] and [Hi06, (4.3.7)]. Part (b) follows from work of Eichler-Shimura, Deligne, Ohta, Carayol, Taylor and Blasius-Rogawski. If $[F:\mathbb{Q}]$ is odd or if f has a supercuspidal prime, then the construction takes place in the étale cohomology of a Shimura curve. Taylor's argument for $[F:\mathbb{Q}]$ even is via congruences and the Jacquet-Langlands correspondence for GL_2. See [Ca86, Tay89].

Definition 3.9. Let f be a cuspidal Hecke eigenform as in the previous theorem.

(a) The L-factor of f at v is $L_v(f,T) := \det(1 - T\mathrm{Frob}_v|V_\lambda(f)^{I_v})^{-1}$ where λ is any place of E_f such that v and λ have different residue characteristics – see Theorem 3.10(a).

(b) The L-function of f is

$$L(f,s) := \prod L_v(f, q_v^{-s}) \text{ for } \Re(s) \gg 0.$$

Theorem 3.10. *Let* $k_0 := \max\{k_\tau \mid \tau \in I\}$. *Then*

(a) The local L-factors are independent of the choice of λ *(as long as* v *and* λ *are not above the same rational prime* ℓ*).*

(b) The poles of the local L-factor at $v \nmid N$ *are algebraic integers of absolute value* $q_v^{(k_0-1)/2}$ *(under any complex embedding), i.e. Weil numbers.*

(c) The L-function defined above converges for all $s \in \mathbb{C}$ *with* $\Re(s) > \frac{k_0+1}{2}$. *It has an entire continuation to the complex plane and a functional equation* $\Lambda(f, k_0 - s) = \Lambda(f,s)$ *where* Λ *is obtained from* L *by multiplication by suitable L-factors at* ∞ *– see the references.*

(d) The representation $V_\lambda(f)$ *is irreducible.*

References: [RT11], [Sk09], [Bl06], [Sai11, Thm 2], [Ri85] for part (d), case of elliptic modular forms, [Tay97] for (d) for Hilbert modular forms.

Remark 3.11. The idea of [Ri85] is as follows: If the representation is reducible get $\varepsilon_\ell^a\ \varepsilon_\ell^b$ on the diagonal up to finite order with product ε_ℓ^{k-1} up to finite order (by CFT). The Ramanujan-Petersson bound yields $2a = 2b = k-1$. Growth of L-function at $s = k$ gives two contradictory bounds (cusp form versus Eisenstein series.)

3.4 Compatible systems of Galois representations I (see [Se68])

Definition 3.12 (Weakly compatible system)**.** Let E be a number field and $\mathcal{P}$ its set of finite places. Let S_λ consist of the places v of F such that v and λ lie over the same rational prime ℓ. A family of n-dimensional continuous Galois representations $\big(\rho_\lambda\colon G_F \to \mathrm{GL}_n(E_\lambda)\big)_{\lambda\in\mathcal{P}}$ is an *E-rational weakly compatibly system (with finite ramification set S)* if

(a) for all $\lambda \in \mathcal{P}$, the representation ρ_λ is unramified outside $S \cup S_\lambda$;

(b) for all finite places v of F not in S there exists a polynomial $p_v(T) \in E[T]$ such that

$$p_v(T) = \det(1 - T\rho_\lambda(\mathrm{Frob}_v)) \in E_\lambda[T] \quad \forall\lambda \text{ such that } v \notin S_\lambda,$$

where E is canonically a subfield of E_λ, its completion at λ.

Example 3.13. Let μ_{ℓ^∞} denote the set of ℓ-power roots of unity in $\bar{F}$ for some number field F. It is clearly stable under G_F. Define the ℓ-adic cyclotomic character $\varepsilon_\ell\colon G_F \to \mathrm{Aut}(\mu_{\ell^\infty}) \cong \mathbb{Z}_\ell^*$ by $g \mapsto (\zeta_{\ell^n} \mapsto \zeta_{\ell^n}^{\varepsilon_\ell(g)})$ for all n. (This is independent of the choice of roots (ζ_{ℓ^n}).) Then $(\varepsilon_\ell)_\ell$ forms a compatibly $\mathbb{Q}$-rational system with ramification set $S = \emptyset$.

Extending Definition 3.1, we shall write $V_\ell(A)$ for the rational ℓ-adic Tate module of any abelian variety A.

Theorem 3.14. *(a) If A is an abelian variety over a number field F, then the representations $V_\ell(A)$ form a $\mathbb{Q}$-rational weakly compatible system with ramification set, the set of places of F where A has bad reduction.*

(b) If f is a Hilbert modular form over a totally real field F , then the representations $V_\lambda(f)$ form an E_f-rational weakly compatible system with ramification set, the set of places of F dividing the (minimal) level N of f.

Part (a) is [ST68, Thm. 3]. Part (b) is immediate from Theorem 3.8.

Remark 3.15. Now we can meaningfully talk about the symmetric powers of an elliptic curve A/F. Namely we can mean by this the weakly compatible $\mathbb{Q}$-rational family $(\mathrm{Sym}^n V_\ell(A))$ of representations of dimension $n + 1$.

3.5 What does it mean for an elliptic curve A to be modular?

- A and f have the same L-function.
  It is elementary to see that this implies $a_v(A) = a_v(f)$ for all places of F at which A has good reduction and which do not divide the level.
  Converting this into a statement about Galois representations it follows that $V_\ell(A) = V_\ell(f)$ (and $E_f = \mathbb{Q}$). Remembering that we defined L-factors at bad (or at all places) via the Galois representation (and inertia invariants), it follows that we must have equality of the remaining L-factors.

- The ℓ-adic Tate module $V_\ell(A)$ of A and the ℓ-adic Galois representation $V_\ell(f)$ attached to f are isomorphic for one ℓ (or all ℓ).

- Even better: From f one can (often, and always if $[F : \mathbb{Q}]$ is odd) construct an elliptic curve A_f (in some jacobian of a modular/Shimura curve). Then A being modular means that A is isogenous to some A_f.

The three notions above are equivalent, provided they all make sense – note that A_f is not always defined. The proof is an instructive exercise. The following remarks may be helpful.

Remark 3.16. (a) If two elliptic curves A and A' are isogenous over F, then they have the same Galois representation and thus the same L-function. Consider the isogeny $0 \to K \to A \to A' \to 0$ with finite kernel K. Passing to ℓ^∞-torsion points and Tate modules, we deduce

$$0 \to K[\ell^\infty] \to \mathrm{Ta}_\ell(A) \to \mathrm{Ta}_\ell(A') \to 0.$$

Tensoring with $\mathbb{Q}_\ell$ over $\mathbb{Z}_\ell$, the left hand term disappears and the other two become isomorphic.

(b) It is a deep theorem due to Faltings [Fa82], the semisimplicity conjecture of Tate, that shows that for Galois representations of abelian varieties A, A' over a number field F we have that A and A' are isogenous if and only if they have isomorphic Galois representations. Faltings proves that the following natural homomorphism is an isomorphism:

$$\underbrace{\mathrm{Hom}(A, A') \otimes_{\mathbb{Z}} \mathbb{Q}_\ell}_{\text{rational isog. } A\to A'} \longrightarrow \underbrace{\mathrm{Hom}_{\mathbb{Q}_\ell[G_F]}(V_\ell(A), V_\ell(A'))}_{G_F-\text{equiv. homom.}}.$$

4 Weil-Deligne representations

The following is based on notes by T. Gee from the 2011 Winter School in Postech, Korea. The ultimate source for this is the important article [Tat79].

Question: Can one define a refined notion of compatible system that also takes ramified primes into account?
Answer: Yes, by introducing Weil-Deligne representations.
Question: Can one refine the notion of compatible system even further to also include primes above the residue characteristic of the representation?
Answer: Yes (later), via Fontaine's p-adic Hodge theory – and again Weil-Deligne representations.

4.1 Galois representations of local fields

The Weil group

Let $K/\mathbb{Q}_p$ be a finite extension, $\mathcal{O}_K$ its ring of integers, π_K its uniformizer, $k = k_K$ its residue field of cardinality q_K. Let $\mathrm{v}_K \colon K^* \to \mathbb{Z}$ be the normalized additive valuation and $|\,|_K$ the multiplicative valuation with $|\pi_K|_K = q_K^{-1}$.

Every element g of $G_K = \mathrm{Gal}(\bar{K}/K)$ preserves $\mathcal{O}_{\bar{K}}$ and induces an automorphism of the residue field $\bar{k}$ of $\mathcal{O}_{\bar{K}}$. The kernel of the induced homomorphism $G_K \to G_k$ is the inertia subgroup I_K.

Denote by Frob_K the canonical topological generator of G_k, the geometric Frobenius, i.e., the inverse automorphism to $\bar{k} \to \bar{k} : x \mapsto x^{q_K}$. Pullback of $1 \to I_K \to G_K \to G_k \to 1$ along $\langle \mathrm{Frob}_K \rangle \to G_k$ defines the Weil group W_K in the s.e.s.

$$0 \to I_K \to W_K \to \langle \mathrm{Frob}_K \rangle \to 0.$$

Here W_K is a topological group by taking the neighborhoods of I_K (under the profinite topology of G_K) as a neighborhood basis of the identity.

The inertia subgroup

Define K^{nr} as $\bar{K}^{I_K}$. Then $K^{\mathrm{nr}} = \cup_n K(\zeta_{p^{n!}-1})$ and $\mathrm{Gal}(K^{\mathrm{nr}}/K) \to G_k$ is an isom.
Define $K^{\mathrm{tame}} := \cup_{\gcd(n,p)=1} K^{\mathrm{nr}}(\pi^{1/n})$. By Kummer theory K^{tame}/K is Galois and

$$\mathrm{Gal}(K^{\mathrm{tame}}/K) \cong \mathrm{Gal}(K^{\mathrm{tame}}/K^{\mathrm{nr}}) \rtimes \mathrm{Gal}(K^{\mathrm{nr}}/K) \cong \hat{\mathbb{Z}}' \rtimes \hat{\mathbb{Z}}$$

with $\hat{\mathbb{Z}}' = \prod_{\ell \neq p} \mathbb{Z}_\ell$.

Kummer theory says that $g \in \mathrm{Gal}(K^{\mathrm{tame}}/K^{\mathrm{nr}})$ maps to $\hat{\mathbb{Z}}'$. Explicitly: choose a compatible system of roots of unity $\zeta := (\zeta_n)_{n\,\mathrm{prime\,to}\,p} \subset K^{\mathrm{nr}}$. Define for $g \in I_K$ a sequence $t_n(g)$ in the inverse limit $\hat{\mathbb{Z}}'$ by $g(\pi^{1/n})/\pi^{1/n} = \zeta_n^{t_n(g)}$. The t_n define a surjective homomorphism $t_\zeta : I_K \to \hat{\mathbb{Z}}'$. Denote by $t_{\zeta,\ell}$ the composite of t_ζ with the projection $\hat{\mathbb{Z}}' \to \mathbb{Z}_\ell$. The kernel of $G_{K^{\mathrm{nr}}} \to \mathrm{Gal}(K^{\mathrm{tame}}/K^{\mathrm{nr}})$ is the wild ramification subgroup P_K. It is the pro-p-Sylow subgroup of I_K.

Let $\xi_K \colon W_K \to \langle \mathrm{Frob}_K \rangle \to \mathbb{Z}$ be character defined by $\mathrm{Frob}_K \mapsto 1$.Then one has

$$t_\zeta(g\tau g^{-1}) = q_v^{-\xi_K(g)} t_\zeta(\tau) \tag{4.1}$$

for $\tau \in I_K$ and $g \in W_K$.

Theorem 4.1 (Main Thm of local CFT)**.** *Let W_K^{ab} denote the group $W_K/\overline{[W_K, W_K]}$. Then there is a unique system of isomorphisms (for all extensions of $\mathbb{Q}_p$)*

$$\mathrm{Art}_K : K^* \to W_K^{\mathrm{ab}}$$

such that

(a) if K'/K is a finite extension and $\pi_{K'/K} : W_{K'}^{\mathrm{ab}} \overset{\mathrm{can}}{\to} W_K^{\mathrm{ab}}$, then $\pi_{K'/K} \circ \mathrm{Art}_{K'} = \mathrm{Art}_K \circ N_{K'/K}$,

(b) and we have a commutative square

$$\begin{tikzcd}
K^* \arrow[r, "\mathrm{Art}_K", "\sim"'] \arrow[d, "\mathrm{v}_K"'] & W_K^{\mathrm{ab}} \arrow[d, "\mathrm{can}\colon g\mapsto\bar{g}"] \arrow[dl, "\xi_K"] \\
\mathbb{Z} \arrow[r, "a\mapsto\mathrm{Frob}_K^a"'] & \langle \mathrm{Frob}_K \rangle
\end{tikzcd}$$

Definition 4.2 (Weil and Weil-Deligne representations)**.** Let L be a field of characteristic zero.

A representation of W_K over a field L (on a finite dimensional vector space over L) is a representation which is continuous with respect to the discrete topology on L and the one defined above for W_K.

A Weil-Deligne representation of W_K on a finite dimensional L-vector space V is a pair (r, N) (or a triple $(V; r; N)$) where r is a representation of W_K on V and N is in $\mathrm{End}(V)$ such that for all $\sigma \in W_K$ one has the following analog of (4.1)

$$r(\sigma) N r(\sigma)^{-1} = q_v^{-\xi_K(\sigma)} N. \tag{4.2}$$

Remark 4.3. For r as above the image $r(I_K)$ is finite. Moreover by considering the eigenvalues of N it easily follows that N is nilpotent. Finally, the relation (4.2) is equivalent to $r(\sigma)N = Nr(\sigma)$ for all $\sigma \in I_K$ and $r(\phi)Nr(\phi)^{-1} = q_K^{-1}N$ for $\phi \in W_K$ a lift of the geometric Frobenius Frob_K.

The conductor of a WD-representation is

$$c(r, N) := c(r) + \dim V^{I_K} - \dim(\mathrm{Ker}(N : V \to V)^{I_K})$$

where $c(r)$ is the usual Artin conductor of a discrete representation in characteristic zero. The Artin conductor of $c(r)$ can be defined as the Artin conductor $c(r')$ of the finite image representation r' from part (f) of the following exercise – r' is a twist of r by an unramified character.

Exercise 4.4. (a) For a representation $(V; r)$ of W_K and $m \geq 1$, define $\mathrm{Sp}_m(r)$ as the triple

$$\Big(\bigoplus_{i=1,\ldots,m} V, \bigoplus_{i=1,\ldots,m} r \cdot |\mathrm{Art}_K^{-1}|_K^{m-i}, N \Big)$$

with N restricted to the i-th component V that is acted on by $r \cdot |\mathrm{Art}_K^{-1}|_K^{i}$ being the isomorphism to the $i+1$-th component acted on by $r \cdot |\mathrm{Art}_K^{-1}|_K^{i+1}$. Then this defines a WD-representation.

(b) Every WD representation is isomorphic to a direct sum of representations $\mathrm{Sp}_m(r)$.

(c) If $(r; V; N)$ is a WD representation of W_K and K'/K is a finite extension, then the restriction $(r|_{G_{K'}}; V; N)$ is a WD representation of $W_{K'}$.

(d) If r is a representation of W_K, then there exists a finite index subgroup H such that $r(H)$ lies in $Z(r(W_K))$. In particular, the projective representation induced from r has finite image.

(e) There exists a representation r' of G_K (of finite image) such that r and $r'|_{W_K}$ have the same projective image, and in particular any Weil representation is a twist of a representation of G_K (of finite order) by character of W_K. (Hint: [Se77, Cor. of Thm. 4].)

(f) There exists a representation r' of G_K (of finite image) and an unramified character χ of W_K such that $r = \chi \otimes r'$.

(g) Let σ be in $W_K \setminus I_K$. Then for any $\tau \in W_K$ there exist n in $\mathbb{Z}$ and $m \in \mathbb{Z}_{>0}$ such that $r(\sigma^m) = r(\tau^n)$.

(h) For a representation r of W_K the following conditions are equivalent: (a) r is semisimple. (b) $r(\sigma)$ is semisimple for all $\sigma \in W_K$. (c) $r(\sigma)$ is semisimple for some $\sigma \notin I_K$.

(i) If $(r; N)$ is a Weil-Deligne representation of W_K, then $(r, N)^{F-\mathrm{ss}} := (r^{\mathrm{ss}}, N)$ is a WD-representation of W_K.

Note that by r^{ss} we mean the following semisimplification: suppose the α is any automorphism of a vector space over a field L. Then α can be written in a unique way as $\alpha = \alpha^{\mathrm{ss}} \cdot \alpha^{\mathrm{unip}}$ for commuting endomorphisms α^{ss} and α^{unip} such that α^{ss} is semisimple, i.e., it is diagonalizable over L^{alg}, and α^{unip} is unipotent, i.e., all of its eigenvalues are one. Now one defines r^{ss} to mean that for any $g \in W_K$, one sets $r^{\mathrm{ss}}(g) := (r(g))^{\mathrm{ss}}$ in the sense just described. Note that since I_K has finite image under r, all elements in $r(I_K)$ are semisimple.

Definition 4.5. A WD-representation (r, N) is Frobenius semisimple if r is semisimple. (i.e. $r(\phi)$ is semisimple as an endom.)

Definition 4.6. Let L be an algebraic extension of $\mathbb{Q}_\ell$ with $\ell \neq p$.

(a) $A \in GL_n(L)$ is bounded if $\det A$ lies in $\mathcal{O}_L^*$ and $\det(1 - TA)$ in $\mathcal{O}_L[T]$.

(b) A representation r of W_K is bounded if $r(\sigma)$ is bounded for all σ in W_K

Remark 4.7. (i) In (a), the matrix A is bounded if it stabilizes an $\mathcal{O}_L$ lattice in L^n.

(ii) In (b), the representation r is bounded if and only if $r(\sigma)$ is bounded for some $\sigma \notin I_K$.

Theorem 4.8 (Grothendieck's Monodromy Theorem, [Tat79, Cor. (4.2.2)])**.** *Suppose $l \neq p$, $K/\mathbb{Q}_p$ is finite, $L/\mathbb{Q}_\ell$ is finite and V is a finite dimensional L-vector space. Fix a lift $\phi \in W_K$ of Frob_K and a compatible system $\zeta = (\zeta_n)$ of roots of unity in K^{alg}. (This defines a unique $t_{\zeta,\ell}\colon I_K \to \mathbb{Z}_\ell$ for all $\ell \neq p$.)*

For any continuous representation $\rho : G_K \to \mathrm{GL}(V)$, there exists a finite extension K' of K such that $\rho(I_{K'}) \equiv 1 \mod 2\ell$ for an $\mathcal{O}_L$-lattice of V stabilized by G_K and there exists a unique nilpotent endomorphism N of V such that for all $\sigma \in I_{K'}$ one has $r(\sigma) = \exp(Nt_{\zeta,\ell}(\sigma))$.

Moreover if $r : W_K \to \mathrm{GL}(V)$ is defined by

$$r(\sigma) = \rho(\sigma)\exp(Nt_{\zeta,\ell}(\phi^{-\xi_K(\sigma)}\sigma)),$$

then $(r, N) =: \mathrm{WD}(\rho)$ defines a WD-representation of W_K. The functor $\mathrm{WD} = \mathrm{WD}_{\phi,\zeta}$ defines an equivalence of categories from continuous representations ρ to bounded WD-representations (r, N).

Finally for any choices (ϕ, ζ) and (ϕ', ζ') there is a natural isomorphism

$$\mathrm{WD}_{\phi,\zeta} \to \mathrm{WD}_{\phi',\zeta'}\,.$$

Proof. Exercise: The main tool needed is the existence of an ℓ-adic logarithm. This is ensured by the condition that ρ on $I_{K'}$ has pro-ℓ image and that the matrices of this image are congruent to $1 \mod 2\ell$. Then the usual series for the log converges. □

Remark 4.9. Suppose $\rho : G_K \to \mathrm{Aut}(V)$ is unramified. Then $N = 0$ and $r(I_K) = \{1\}$ for $(r, N) = \mathrm{WD}(V)$. Thus r is completely determined by $\rho(\phi)$ for a lift ϕ of Frob_K. In other words, $\mathrm{WD}(\rho)$ depends on the conjugacy class of $\rho(\phi)$, i.e., its rational canonical form. If one passes to $\mathrm{WD}(\rho)^{F-\mathrm{ss}}$, then the isomorphism type of the latter is completely determined by the characteristic polynomial $\det(1 - T\mathrm{Frob}_K|V)$.

4.2 Compatible systems II

Definition 4.10 (Strictly compatible system)**.** Let E be a number field and $\mathcal{P}$ its set of finite places. For $\lambda \in \mathcal{P}$ let S_λ denote the set of places v of a number field F such that v and λ lie over the same rational prime ℓ. A family of n-dimensional continuous Galois representations $(V_\lambda)_{\lambda\in\mathcal{P}}$ of G_F is an *E-rational strongly compatibly system (with finite ramification set S)* if

(a) for all $\lambda \in \mathcal{P}$, the representation V_λ is unramified outside $S \cup S_\lambda$;

(b) for all finite places v of F not in S there exists a polynomial $p_v(T) \in E[T]$ such that

$$p_v(T) = \det(1 - T\mathrm{Frob}_v|V_\lambda) \in E_\lambda[T] \quad \forall\lambda \text{ such that } v \notin S_\lambda;$$

(c) for all finite places v in S there exists an Frobenius semisimple WD-representation (r_v, N_v) of F_v such that

$$\mathrm{WD}(V_\lambda|_{G_{F_v}})^{F-\mathrm{ss}} = (r_v, N_v) \quad \forall\lambda \text{ such that } v \notin S_\lambda.$$

Conjecture 4.11 (Fontaine, Serre, Deligne)**.** *If V is a representation that occurs in the ℓ-adic étale cohomology of a smooth proper variety over a local field, then its associated Weil-Deligne representation is Frobenius semisimple. Moreover the Weil-Deligne representation is independent of ℓ. See [Tat79], [Fo94c, Section 2.4.], [Se91, §§11,12].*

Theorem 4.12 (Carayol, Eichler-Shimura, Langlands, Deligne, see [Ca86])**.** *For v a Hilbert modular eigenform, the family $(V_\lambda(f))$ is a strongly compatible E_f-rational system;*

Theorem 4.13. *Suppose A/F is an abelian variety. Then $(V_p(A))$ is a strongly compatible system.*

I could not find a reference for Theorem 4.13 as stated. Strict compatibility at places of good reduction is standard. The case of semistable reduction can be deduced from the thesis of A. Laskar from 2011 (Strasbourg). According to private communication with Fontaine, the result as stated has the status of a theorem. It can be deduced from Raynaud's rigid analytic models for abelian varieties with semistable reduction. Apparently the finite Galois action of a Galois extension over which semistable or good reduction is acquired poses no problems. But a reference for the general case seems to be absent from the literature.

5 How to deal with primes above ℓ?

5.1 An example

Given A/F an elliptic curve, ℓ a rational prime, v a place of F above ℓ and with residue field k_v. What do we know about $V_\ell(A)$ restricted to a decomposition group at v?

- **good reduction at** v**.** If A/F_v has good reduction, then A/k_v is an elliptic curve and there is a short exact sequence

  $$0 \to A^0[\ell^\infty](\bar{F}_v) \to A[\ell^\infty](\bar{F}_v) \to A[\ell^\infty](\bar{k}_v) \to 0$$

  where $A^0[\ell^\infty]](\bar{F}_v)$ is given by a formal group of dimension 1 and height 1 (ordinary case) or height 2 (supersingular case) and $A[\ell^\infty](\bar{k}_v)$ is either isomorphic to $\mathbb{Q}_\ell/\mathbb{Z}_\ell$ if A/F_v is ordinary or trivial if A/F_v is supersingular.

- **ordinary subcase.** Here $V_\ell(A)|_{G_v}$ is an extension of two 1-dimensional representations and thus of the form

  $$\rho_v : G_v \to \mathrm{GL}_2(\mathbb{Q}_\ell) : g \mapsto \begin{pmatrix} \varepsilon_\ell(g)\chi(g) & * \\ 0 & \chi^{-1}(g) \end{pmatrix}$$

  with respect to a suitable basis and where ε_ℓ is the ℓ-adic cyclotomic character and χ is an unramified character. Due to the Weil-paring, the determinant must be ε_ℓ. Note the ε_ℓ is infinitely wildly ramified – so that there is no associated WD-representation.

- **supersingular subcase.** Now $V_\ell(A)|_{G_v}$ need not have a filtration. If $\mathrm{End}_{\bar{F}_v}(A/F_v)$ is 2-dimensional then the formal group is given by two conjugate Lubin-Tate characters and easy to describe via local class field theory. In general, the representation is absolutely irreducible and remains so over any finite index subgroup H of G_v. The mod ℓ reduction is rather special as can be seen from analyzing the ℓ-torsion group via the formal group law of A/F_v. The representation is infinitely ramified and again there is no direct way to get a WD representation.

Exercise 5.1. Suppose A has good reduction and π denotes the Frobenius endomorphism on A/k_v, so that π satisfies the quadratic polynomial $p_v(T) = T^2 - a_vT + q_v$ with integer coefficients, where $\#A(k_v) = q_v - a_v(a) + 1$. Let $\alpha, \beta \in \overline{\mathbb{Z}}$ be the roots of p_v so that $v_{q_v}(\alpha\beta) = 1$. It is also standard that $\#A(k_v^n) = q_v^n - \alpha^n - \beta^n + 1$ for k_v^n the unique extension of the finite field k_v of degree n. Show that A/k_v is supersingular if and only if $v_{q_v}(\alpha), v_{q_v}(\beta) > 0$, if and only if $v_{q_v}(\alpha) = v_{q_v}(\beta) = 1/2$. If A_v is ordinary, then without loss of generality $v_{q_v}(\alpha) = 0$. Show that $\chi^{-1}(\mathrm{Frob}_v)$ acts on the p^∞ torsion points in the same way as π and that $\chi^{-1}(\mathrm{Frob}_v) = \alpha$.

- **semistable reduction at v (here only split multiplicative reduction).** Here one uses the Tate curve of A/F_v. It shows that the Galois action on the ℓ^∞-torsion points is given by the Galois action on $\bar{F}_v^*/q^{\mathbb{Z}}$ for $q \in F_v^*$ an element of valuation strictly less than one. The ℓ^∞ torsion points are given by the set $\{q^{i/\ell^n}\zeta_{\ell^n}^j \mid i, j \in \mathbb{Z}\}$. This describes an infinite Kummer extension of F_v. The corresponding Galois extension is of the form

  $$\rho_v : G_v \to \mathrm{GL}_2(\mathbb{Q}_\ell) : g \mapsto \begin{pmatrix} \varepsilon_\ell(g) & * \\ 0 & 1 \end{pmatrix}$$

- **potentially good or potentially semistable reduction (or non-split multiplicative reduction).**
  The representation is as above after one restricts G_{F_v} to a suitable open subgroup (coming from the field over which good or semistable reduction is acquired).

Comparison to ℓ'-adic data.
If one looks at the ℓ'-power torsion, then the above cases correspond to unramified, unramified, semistable ($N \neq 0$) WD-representations or the general case of a WD-representation.

For the representations that arise from modular forms one has a similar behavior. The solution of the puzzle:

5.2 Fontaine's mysterious functors

References are [Ber02, Ber10, Ber12], [CB09], [Fo94a, Fo94b], [FO09], [GM09].

Let K be a finite extension of $\mathbb{Q}_p$ and K_0 the subfield of K that is maximal unramified over $\mathbb{Q}_p$. Fontaine defines functors D_*, $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{cris}, \mathrm{st}\}$ from

$$\Big\{\rho\colon G_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p) \mid \rho \text{ is cont.}\Big\}$$

to modules over $K_* \otimes_{\mathbb{Q}_\ell} \overline{\mathbb{Q}_\ell}$ for $K_* = K$, K, K_0, or K_0, respectively, with some additional structure coming from some rings B_* of Fontaine with $B_{\text{cris}} \subset B_{\text{st}} \subset B_{\text{dR}}$ and B_{HT} is the graded ring associated to B_{dR}. The structures are a continuous action of G_K and

(a) a graduation,

(b) a filtration,

(c) a filtration and a semilinear endomorphism ϕ (Frobenius)

(d) a filtration and two endomorphisms ϕ, N – for more, see 5.3.

The rings B_* also satisfy $B_*^{G_K} = K_*$. One calls a representation ρ to be $*$ (= Hodge-Tate, de Rham, crystalline, semistable) if

$$(V(\rho)) \otimes_{\mathbb{Q}_p} B_*)^{G_K} \text{ is free over } K_* \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p} \text{ of rank equal to } \dim_{\overline{\mathbb{Q}_p}} V(\rho).$$

In all cases one has $\leq$ for the rank.

The ring B_{HT} attaches the HT-weights (in $\mathbb{Z}$) to a representation. The ring B_{dR} carries a $\mathbb{Z}$-graded filtration. The property of being Hodge-Tate or de Rham is invariant under finite extensions E/K. The property of being cris and st are not invariant under such extensions. So one also has the notions of *potentially crystalline* and *potentially semistable*; for instance ρ is potentially crystalline if there exists a finite extension E/K such that $\rho|_{G_E}$ is crystalline. The implications for a representation V are described in the following diagram

$$\begin{tikzcd}
V \text{ is cris} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & V \text{ is semist.} \arrow[d, Rightarrow] & & \\
V \text{ is pot. cris} \arrow[r, Rightarrow] & V \text{ is pot. semist.} \arrow[r, Leftrightarrow, "\text{A−B−K−M}"] & V \text{ is dR} \arrow[r, Rightarrow] & V \text{ is HT.}
\end{tikzcd}$$

The proof of Fontaine's conjecture that the notions de Rham and potentially semistable are equivalent, due to A=I. André, B=L. Berger, K=K. Kedlaya, M=Z.

Mebkhout, is described by Colmez in [Co03]. Berger proves that Crew's conjecture is equivalent to the conjecture of Fontaine, and the three other authors, independently, give a proof of Crew's conjecture.

5.3 The WD-representation of a potentially semistable $V = V(\rho)$ of rank d (after Fontaine)

Let K, V be as above and suppose K'/K is a finite Galois extension such that $V|_{G'_K}$ is semistable. Then $D_{\mathrm{st}}(V|_{G_{K'}}) = (D, \phi, N, \mathrm{Fil}^i)$ where

(a) D is a free $K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$ module of rank d.

(b) $\phi_D \colon D \to D$ is $\sigma \otimes \mathrm{id}$ linear with $\sigma \colon K'_0 \to K'_0$ the Frobenius in $\mathrm{Gal}(K'_0/\mathbb{Q}_p)$,

(c) $N_D \colon D \to D$ an $K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$-linear endomorphism such that $N_D\phi_D = p\phi_D N_D$,

(d) a decreasing, separating, exhausting filtration Fil^i of $(D \otimes_{K'_0} K')$,

and the quadruple is equipped with an action of $\mathrm{Gal}(K'/K)$, i.e., an action

$$\mathrm{Gal}(K'/K) \longrightarrow \mathrm{Aut}(D, \phi_D, N_D, \mathrm{Fil}^\bullet).$$

If the representation is semistable, one has $K' = K$, if it is potentially crystalline, then $N = 0$ (and vice versa).

Remark 5.2. The quadruple $(D, \phi_D, N_D, \mathrm{Fil}^i)$ is weakly admissible - which is a characterization of the image of the mysterious functor; it means that $t_H(D) = t_N(D)$ and $t_H(D') \leq t_N(D')$ for all stable (but not necessarily free) subobjects D' of D. Here $t_H(D)$ is the index of the unique jump of the filtration of the induced filtration $\mathrm{Fil}^i \bigwedge^d D$ and t_N is the slope of $\bigwedge^d \phi_D$.

The associated WD-representation to ρ is basically obtained by forgetting the filtration, see [GM09] and [Sav05, Def. 2.15] for $K = \mathbb{Q}_p$: Observe that

$$K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p} = \bigoplus_{\tau \colon K'_0 \to \overline{\mathbb{Q}_p}} \overline{\mathbb{Q}_p}.$$

Correspondingly one has

$$D = D \otimes_{K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}} K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p} = \bigoplus_{\tau \colon K'_0 \to \overline{\mathbb{Q}_p}} D_\tau$$

with suitable components D_τ of D. One verifies that $\phi_D^{[K'_0:\mathbb{Q}_p]}$ induces an isomorphism of the D_τ since it is $K'_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$-linear. Fix τ_0 among the τ.

Definition 5.3. Let K, V, K', D be as above and let $\phi \in W_K$ be a lift of Frob_K. Then the WD-representation $\mathrm{WD}(V)$ is the triple (U, r, N) where

(a) U is the $\overline{\mathbb{Q}_p}$ vector space D_{τ_0},

(b) $r\colon W_K \to \mathrm{Aut}_{\overline{\mathbb{Q}_p}}(U)$ is the Weil representation determined by

  (i) defining $r|_{I_K}$ as the restriction of the $K_0 \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$-linear action $I_K \to \mathrm{Gal}(K'/K)$ on D to the invariant subspace D_{τ_0},

  (ii) and having ϕ act as $\phi_2^{-1}\phi_1$ on D_{τ_0} where ϕ_1 is the action of ϕ via $\mathrm{Gal}(K'/K)$ and ϕ_2 is the action $\phi_D^{[K_0:\mathbb{Q}_p]}$ both times on D and then restricted to the invariant subspace D_{τ_0},

(c) N is the restriction of the endomorphism N_D from D to its invariant subspace D_{τ_0}.

In the case where $K = \mathbb{Q}_p$ and K'/K is totally ramified, so that $K_0' = K_0 = \mathbb{Q}_p$, the vector space U is simply equal to D and ϕ is the inverse of ϕ_D (and $N = N_D$).

Note: The eigenvalues of $r(\phi)$ to the power $[K_0' : K_0]$ are those of $\phi_D^{-[K_0':\mathbb{Q}_p]}$. Hence their p-adic valuations are up to the scalar $[K_0 : \mathbb{Q}_p]$ the slopes (= p-adic valuations) of the eigenvalues of ϕ_D. In particular they are typically not units, and thus r is typically **unbounded**. This is therefore different from the case where K and L have different residue characteristic.

5.4 Continuation of Example 5.1 of an elliptic curve A/F_v

(a) If A has ordinary reduction, one has a short exact sequence

$$0 \to D_{\mathrm{cris}}(\mathbb{Z}_\ell(\varepsilon_\ell\chi)) \to D_{\mathrm{cris}}(V_\ell(A)) \to D_{\mathrm{cris}}(\mathbb{Z}_\ell(\chi^{-1})) \to 0$$

where the outer modules have underlying D of rank 1. Thus $D_{\mathrm{cris}}(V_\ell(A))$ is reducible.

(b) If A has supersingular reduction, then either $D_{\mathrm{cris}}(V_\ell(A))$ is a simple object and remains so after base change to any finite extension E/F_v, or there is an extension E/F_v of degree at most 2 over which $D_{\mathrm{cris}}(V_\ell(A))$ becomes the sum of two 1-dimensional subobjects.

(c) If A has semistable but non-good reduction, then $N = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $D_{\mathrm{st}}(V_\ell(A))$ has rank 2 and is reducible while $D_{\mathrm{cris}}(V_\ell(A))$ has rank 1 only; see [Ber02, p. 18].

5.5 Compatible systems III

Definition 5.4 (Strictly compatible system, strong sense)**.** Let E be a number field and $\mathcal{P}$ its set of finite places. For $\lambda \in \mathcal{P}$, and F a number field, let S_λ denote the places v of S such that v and λ lie over the same rational prime ℓ.

A family of n-dimensional continuous Galois representations $(V_\lambda)_{\lambda\in\mathcal{P}}$ of G_F is an *E-rational strictly compatibly system in the strong sense, (with finite ramification set S)* if

(a) for all $\lambda \in \mathcal{P}$, the representation V_λ is unramified outside $S \cup S_\lambda$ and potentially semistable at the places in S_λ;

(b) for all finite places v of F there exists a Frobenius semisimple WD-representation (r_v, N_v) of F_v such that

$$\mathrm{WD}(V_\lambda|_{G_{F_v}})^{F-\mathrm{ss}} = (r_v, N_v) \quad \forall \lambda \in \mathcal{P}.$$

Remark 5.5. As remarked earlier, if v is not in S, then (r_v, N_v) is simply $(\rho(\mathrm{Frob}_v)^{\mathrm{ss}}, 0)$ – this is also true for places v in $S_\lambda \setminus S$ (so that V_λ is crystalline locally at v). In particular, at $v \notin S \cup S_\lambda$ it suffices to require a compatibility of characteristic polynomials

$$p_v(T) = \det(1 - T\mathrm{Frob}_v|V_\lambda) \in E_\lambda[T] \quad \forall \lambda \text{ such that } v \notin S_\lambda;$$

Remark 5.6. One can add further conditions on compatible systems: For instance a purity of weight condition: All roots of all characteristic polynomials are Weil-number of the same weight. Etc.

Theorem 5.7 (Faltings, see [Fa87]). *For a cuspidal Hecke eigenform f the representations $(V_\lambda(f))_\lambda$ are semistable of HT-weights* $(0, k-1)$.

Remark 5.8. The HT weights of $V_\ell(A)$ for an abelian variety A over a number field F are 0 and 1 equally distributed over all places v of F above ℓ and all embeddings $F_v \hookrightarrow \overline{\mathbb{Q}_\ell}$. A proof in other language is contained in [Tat67].

Theorem 5.9 (T. Saito, see [Sai11]). *For a cuspidal Hilbert modular Hecke eigenform f the system $(V_\lambda(f))_\lambda$ is strictly compatible in the strong sense.*

Question 5.10. *Is it known that for an abelian variety A over number field F the family $(V_\ell(A))_\ell$ is strictly compatible in the strong sense? (The HT-weights are* 0 *and* 1 *each with multiplicity* $\dim A$*.)*

5.6 A refinement: Fontaine Laffaille theory

The above is not the end of the story of understanding ℓ-adic Galois representations of ℓ-adic fields. One also needs *integral information* on $V(\rho)$. Fontaine's theory above only contributes to ℓ-adic information, but does not help if one studies the mod ℓ-reduction. One theory that achieves this is due to Fontaine and Laffaille from [FL82].

Let $K_0/\mathbb{Q}_p$ be unramified (there are slight extensions which shall not bother us) with ring of integers W, the ring of Witt vectors of the residue field of K_0.

Definition 5.11. A *strongly divisible module* is a free W-module M of finite type equipped with a decreasing filtration by sub-W-modules $(\mathrm{Fil}^i M)_{i\in\mathbb{Z}}$ such that $\mathrm{Fil}^0 M = M$, $\mathrm{Fil}^i M = 0$ for $i \gg 0$, each subquotient $M/Fil^i M$ free over W and one has a semilinear, i.e., a σ-linear, endomorphism $\phi\colon M \to M$ such that $\phi(\mathrm{Fil}^i M) \subset p^i \mathrm{Fil}^i M$ and $M = \sum_{i\geq 0} p^{-i}\phi(\mathrm{Fil}^i M)$.

Theorem 5.12 (Fontaine-Laffaille). *If M is a strongly divisible module, then $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ is admissible, i.e., equal to $D_{\mathrm{cris}}(V)$ for some $\mathrm{rank}_W M$ dimensional (crystalline) representation of G_{K_0}.*
Suppose conversely that V is a crystalline representation of G_{K_0} of HT weights between 0 and $p-2$ and with $\mathbb{Q}_p$-coefficients. Then $D_{\mathrm{cris}}(V)$ contains a strongly divisible module. (Hard part.)

Let A be a W-algebra. (e.g. W or W/pW or a finite W-algebra, or ...)

Definition 5.13. A *Fontaine-Laffaille module over A* is a finitely generated $W \otimes A$-module M equipped with a decreasing filtration by sub-$W \otimes A$-modules $(\mathrm{Fil}^i M)_{i\in\mathbb{Z}}$ such that $\mathrm{Fil}^0 M = M$, $\mathrm{Fil}^i M = 0$ for $i \geq p-1$, and with a semilinear, i.e., a $\sigma \otimes \mathrm{id}_A$-linear, endomorphism $\phi^i\colon \mathrm{Fil}^i M \to M$ such that $\phi|_{\mathrm{Fil}^i M} = \phi^i$ and $M = \sum_{i\geq 0} \phi^i(\mathrm{Fil}^i M)$.

Define a torsion crystalline representation of weight k ($k \in \mathbb{N}$) to be any finite representation of G_{K_0} that can be written as T/T' where T is a Galois stable lattice in a crystalline representation of G_{K_0} with Hodge-Tate weights in $\{0, \ldots, k\}$ and $T' \subset T$ is a Galois stable sublattice. This yields a category of crystalline torsion modules.

Theorem 5.14 (Fontaine-Laffaille). *Suppose A is a finite W-algebra. If M is a Fontaine-Laffaille-module over A, then under the integral version T_{cris} of D_{cris} the module M is the image of the crystalline torsion A-module.*
Conversely, if M is a torsion crystalline representation over A of weight $k \leq p-2$, then the image under T_{cris} is a Fontaine-Laffaille-module M over A.

Special case (we still assume that $K_0/\mathbb{Q}_p$ is unramified):

Theorem 5.15 (Fontaine-Laffaille + Raynaud). *Let $\mathrm{MF}^1_{\mathrm{tor}}$ be the category of Fontaine-Laffaille torsion modules over W with $\mathrm{Fil}^2 = 0$ and suppose $p > 2$. Then there are equivalences of abelian categories:*

$$\mathrm{MF}^1_{\mathrm{tor}} \xrightarrow[\mathrm{FL}]{\cong} \{\text{ finite flat group schemes}/W\} \xrightarrow[\mathrm{Ray.}]{\cong} \{\text{ flat repns. of } G_{K_0}\}.$$

The latter applies in particular to modular forms of weigh 2 and abelian varieties (with level prime to p or conductor not divisible by p, respectively). References are [Ra74] and [FL82].

It is possible to describe finite flat group schemes over finite extensions $K/\mathbb{Q}_p$. For this rather deep theory, we refer to [Ki09b].

6 The Fontaine Mazur conjecture

The Fontaine-Mazur conjecture is the following statement:

Conjecture 6.1 ([FM95]). *Let F be a number field and S a finite set of places of F. Suppose that $\rho\colon G_{F,S} \to \mathrm{GL}_n(\overline{\mathbb{Q}_\ell})$ is continuous and irreducible and that $\rho|_{G_v}$ is de Rham for all $v|\ell$. Then ρ* comes from geometry*, i.e., there exists a smooth projective variety X over F such that ρ is a subquotient of some ℓ-adic cohomology $H^i(X_{\bar F}, \mathbb{Q}_\ell)$.*

In special cases, the conjecture can be phrased in a more concrete form, in the sense that a recipe is given where in geometry one can find the representation:

Conjecture 6.2. *Let S be a finite set of places of $\mathbb{Q}$. Suppose that $\rho\colon G_{\mathbb{Q},S} \to \mathrm{GL}_2(\overline{\mathbb{Q}_\ell})$ is continuous and irreducible. Suppose further that $\rho|_{G_p}$ is de Rham with HT weights $0, w \geq 1$. Then there exists an elliptic cuspidal Hecke eigenform f and a place λ of the coefficient field E_f of f such that $V(\rho) \cong V_\lambda(f)$.*

Under the hypothesis that ρ is odd, conjecture 6.2 is proved in the majority of cases by Emerton and by Kisin, independently, in [Em11] and [Ki09a]. The oddness does follow from the Fontaine-Mazur conjecture, but so far it remains a technical hypothesis in basically all arguments.

A second special case is the following:

Conjecture 6.3. *Let F be a totally real number field. Suppose $\rho\colon G_{F,S} \to \mathrm{GL}_2(\overline{\mathbb{Q}_\ell})$ is continuous, irreducible and of weight 2, i.e., it is de Rham at all places above ℓ with HT-weights $(0,1)$ for all places $v|\ell$ and all embeddings $F_v \to \overline{\mathbb{Q}_\ell}$. Then ρ arises from a Hilbert modular form of weight $(2,\ldots,2)$.*

Conjecture 6.3 in the form stated is completely open – even if one assumes that ρ is totally odd. All known results impose severe restrictions on the ramification at places above ℓ. A weaker assertion is given by the potential modularity results of Taylor. We follow the cases given in [Sn09]:

Theorem 6.4 (Taylor,Dieulefait (version of [Sn09])). *Suppose ρ is as in conjecture 6.3 and that its reduction mod ℓ is* absolutely irreducible*, that $\ell \geq 3$ and that for $\ell = 5$ some extra hypotheses are satisfied. Then*

(a) the L-function $L(\rho, s)$ has a meromorphic continuation to $\mathbb{C}$ and the expected functional equation;

(b) there exists a strictly compatibly system (in the strong sense) $(\rho_\lambda\colon G_{F,S} \to \mathrm{GL}_n(E_\lambda))_\lambda$ for some number field E and a place λ_0 such that $\rho_{\lambda_0} = \rho$.

Note that the theorem applies in particular to elliptic curves over F. The proof, as a combination of Brauer's theorem on characters and the potential modularity result by Taylor; both ideas stem from [Tay06]. Dieulefait in [Die07] observed how to deduce (b) from Taylor's results. An excellent survey of potential modularity and modularity lifting theorems is also [Bu10].

7 How to prove that $A/\mathbb{Q}$ is modular?

- Let N be the conductor of the elliptic curve A.
- We search for a cuspidal Hecke eigenform of weight 2 because A has HT-weights $\{0,1\}$.
- Choose a prime $p > 2$ (for simplicity) not dividing N for which $\overline{A[p]}$ is *absolutely irreducible* and *modular*; the later means that $A[p] \cong \overline{V_\wp(f)}$ for some modular form f (any weight and level). [By Taylor, potentially, this is always possible. Wiles in [Wi95] uses $p = 3$ (but has to deal with problems coming from $p|N$) and results of Langlands and Tunnel.]
- By the weight and level part of Serre's conjecture: we can assume that f has level dividing N and weight equal to 2.
- Consider the Hecke-algebra $\mathbb{T}_2(\Gamma_0(N))$ over $\mathbb{Z}_p$ acting on $S_2(\Gamma_0(N))$ generated by T_q and S_q for primes q not dividing Np. Each Hecke-eigensystem of some eigenform $g \in S_2(\Gamma_0(N))$ yields a ring homomorphism $\mathbb{T}_2(\Gamma_0(N)) \to \overline{\mathbb{F}_p} : T_v \mapsto a_v(f) \mod \mathfrak{m}_{\overline{\mathbb{Z}_p}}$. Let $\mathfrak{m}_{\bar\rho}$ be the kernel for our fixed f. Then the localization $\mathbb{T}_{\bar\rho} := \mathbb{T}_2(\Gamma_0(N))_{\mathfrak{m}_{\bar\rho}}$ is a local ring.

Proposition 7.1. *(a)* $\mathrm{Hom}_{\mathbf{Ri}}(\mathbb{T}_{\bar\rho}, \overline{\mathbb{Q}_p})$ *is in bijection with the set of cuspidal Hecke newforms of level dividing* N *and weight* 2 *whose associated mod p Galois representation is isomorphic to* $\bar\rho$*. (Mod p means mod* λ *for some* λ *over p.)*

(b) There exists a Galois representation $\rho^{\mathrm{mod}} \colon G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{T}_{\bar\rho})$ *characterized completely by*

$$\mathrm{Trace}\,\rho(\mathrm{Frob}_v) = T_v \ \textit{(and } \det\rho(\mathrm{Frob}_v) = q_v\textit{)} \quad \textit{for all } v \not| Np.$$

The proposition says that any Galois representation that arises from a level N form g of weight 2 and that is congruent to f mod p is obtained from ρ^{mod} by a ring homomorphism $\mathbb{T}_{\bar\rho} \to \overline{\mathbb{Q}_p}$. Let $M_{\bar\rho}$ denote the set of these modular forms.

Proof. Part (a) is an elementary exercise. For (b), let $\mathcal{O}$ be the ring of integers of a finite extension K of $\mathbb{Q}_p$ that contains all coefficients of all $g \in M_{\bar\rho}$. Then there is a representation

$$\widetilde{\rho} \colon G_{\mathbb{Q}} \to \mathrm{GL}_2(\prod_{g \in M_{\bar\rho}} \mathcal{O})$$

with $\widetilde{\rho}(\mathrm{Frob}_v) = (a_v(g))_{g \in M_{\bar\rho}}$. Observe that $\mathbb{T}_{\bar\rho}$ is a natural subring via the homomorphisms from (a) of $\prod_{g \in M_{\bar\rho}} \mathcal{O}$ which contains all traces $\mathrm{Trace}(\widetilde{\rho}(\mathrm{Frob}_v)) = T_v$. (under the various embeddings, T_v maps to the tuple $(a_v(g))_{g \in M_{\bar\rho}}$.) Now there is a theorem of Carayol [Ca94] and independently Serre that uses the hypothesis:

(i) $\mathbb{T}_{\bar\rho}$ is a complete noetherian local ring with finite residue field, (ii) the ring $\prod_{g\in M_{\bar\rho}}\mathcal{O}$ is semilocal with finite residue fields and it contains $\mathbb{T}_{\bar\rho}$, (iii) all traces of the representation $\widetilde{\rho}$ belong to the subring $\mathbb{T}_{\bar\rho}$, (iv) $\bar\rho$ is absolutely irreducible. Then $\widetilde{\rho}$ is already defined over $\mathbb{T}_{\bar\rho}$, and this is exactly the claim of (b) □

So now that we have a universal Galois representation for cusp forms of level N, weight 2 and fixed $\bar\rho$, the idea is to compare it to a Galois theoretically defined universal representation, that is hopefully of the same kind but defined abstractly and made so that it "contains" $V_p(A)$.

Define ($R = R_{\bar\rho,N,2}$, ρ_R) as the "universal deformation ring" that parameterizes all 2-dimensional p-adic Galois representation whose mod p reduction is $A[p]$, which at primes $q \mid N$ have the same WD-representation type as A, which at p are crystalline of weights 0,1 (flat)[3], and which have the same determinant as $V_p(A)$ and are unramified at all primes not dividing Np. (By the Weil-paring the latter determinant is a finite twist of the cyclotomic character.)
Such a ring was first defined and studied by Mazur. The main method to relate such rings to Hecke algebras is due to Wiles.
From the universality of R and the construction of $\mathbb{T}_{\bar\rho}$ one deduces (elementary):

Proposition 7.2. *There exists a unique surjective homomorphism $\alpha\colon R \to \mathbb{T}_{\bar\rho}$ such that*

$$\rho^{\mathrm{mod}} \overset{\text{conj.}}{\sim} \alpha\circ\rho_R.$$

In the above situation one has the following result:

Main Theorem 7.1 ([Wi95, TW95, BCDT])**.** *The map $\alpha\colon R \to \mathbb{T}_{\bar\rho}$ is an isomorphism.*

There are many refinements by Clozel, Harris, Khare-Wintenberger, Kisin, Taylor and "his" school and others, of which some are listed in the bibliography, e.g. [BLGG, CHT, KW09, Ki09a, Ki09b, Tay08].

References

[AEK03] *Harmonic Analysis, the Trace Formula and Shimura Varieties*, eds. J. Arthur, D. Ellwood, R. Kottwitz, Proceedings of the Clay Mathematics Institute 2003 Summer School, The Fields Institute, Toronto, Canada, June 2–27, 2003, Clay Mathematics Proceedings, Vol. 4

[BLGG] T. Barnet-Lamb, T. Gee and D. Geraghty, *The Sato-Tate conjecture for Hilbert modular forms.* J. Amer. Math. Soc. **24** (2011), no. 2, 411–469.

[3] Remember that for simplicity we assumed that p does not divide the conductor N of A, so that $V_\ell(A)$ itself is crystalline at p of HT-weights 0,1.

[Bel09] J. Bellaïche, *Automorphic forms for unitary groups and Galois representations, Eigenvarieties of unitary groups.* Three lectures at the Clay Mathematical Institute Summer School, Honolulu, Hawaii, 2009.
`http://people.brandeis.edu/~jbellaic/AutomorphicHawaii3.pdf`

[BC09] J. Bellaïche, G. Chenevier, *Families of Galois representations and higher rank Selmer groups*, Astérisque 324 , Soc. Math. France, 2009.

[Ber02] L. Berger, *An introduction to the theory of p-adic representations*, arXiv:math.NT/0210184v1

[Ber10] L. Berger, *Galois representations and* (ϕ, Γ)*-modules*, perso.ens-lyon.fr/laurent.berger/ihp2010.php

[Ber12] L. Berger, *On p-adic Galois representations*, perso.ens-lyon.fr/laurent.berger/autrestextes.php

[Bl06] D. Blasius, *Hilbert modular forms and the Ramanujan conjecture.* Noncommutative geometry and number theory, Aspects Math., E37, Vieweg, 2006, 35–56. see
`http://www.math.ucla.edu/ blasius/papers.html`

[BCDT] C. Breuil, B. Conrad, F. Diamond and R. Taylor, *On the modularity of elliptic curves over* $\mathbb{Q}$*: wild* 3*-adic exercises*, J. Amer. Math. Soc. **14** (2001), 843–939.

[BGHZ] J.-H. Bruinier, G. van der Geer, G. Harder and D. Zagier, *The 1-2-3 of modular forms.* Lectures from the Summer School on Modular Forms and their Applications held in Nordfjordeid, June 2004. Edited by Kristian Ranestad. Universitext. Springer-Verlag, Berlin, 2008.

[La03] D. Bump, J. W. Cogdell, E. de Shalit, D. Gaitsgory, E. Kowalski, S. S. Kudla, *An introduction to the Langlands program*, Lectures presented at the Hebrew University of Jerusalem, Jerusalem, March 12–16, 2001. Edited by J. Bernstein and S. Gelbart. Birkhäuser Boston, Inc., Boston, MA, 2003.

[Bu10] K. Buzzard, *Potential modularity—a survey.* arXiv:1101.0097v2

[Ca86] H. Carayol, *Sur les représentations* ℓ*-adiques associées aux formes modulaires de Hilbert.* Ann. Sc. École Norm. Sup. (4) **19** (1986), no. 3, 409–468.

[Ca94] H. Carayol, *Formes modulaires et représentations galoisiennes à valeurs dans un anneau local complet.* Contemp. Math., **165** (1994), 213–237.

[CHT] L. Clozel, M. Harris and R. Taylor, *Automorphy for some* ℓ*-adic lifts of automorphic mod l Galois representations*, Pub. Math. IHES **108** (2008), 1–181.

[Co03] P. Colmez, *Les conjectures de monodromie p-adiques.* Sém. Bourbaki, Astérisque **290** (2003), Exp. No. 897, 53–101.

[CB09] B. Conrad and O. Brinon, *CMI Summer School Notes on p-adic Hodge Theory*,
`http://math.stanford.edu/ conrad/papers/notes.pdf`

[CSS] *Modular forms and Fermat's last theorem.* (Boston, MA, August 9–18, 1995). Ed. G. Cornell, J. H. Silverman, G. Stevens, Springer-Verlag, New York, 1997.

[CR62] C.W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Wiley Interscience, New York, 1962.

[Da04] H. Darmon, *Rational points on modular elliptic curves.* CBMS Regional Conf. Series in Math. **101**. AMS, Providence, RI, 2004.

[DDT97] H. Darmon, F. Diamond and R. Taylor, *Fermat's last theorem.* in Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993), 2–140, Int. Press, Cambridge, MA, 1997.

[Die07] L. Dieulefait, *Uniform behavior of families of Galois representations on Siegel modular forms and the endoscopy conjecture*, Bol. Soc. Mat. Mexicana (3) **13** (2007), no. 2, 243–253.

[Em11] M. Emerton, *Local-global compatibility in the p-adic Langlands programme for* $\mathrm{GL}_2/\mathbb{Q}$, preprint 2011, see
`http://math.uchicago.edu/ emerton/preprints.html`

[Fa87] G. Faltings, *Hodge-Tate structures and modular forms.* Math. Ann. **278** (1987), 133–149.

[Fa82] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), 349–366.

[Fo94a] J.-M. Fontaine. *Le corps des périodes p-adiques.* With an appendix by Pierre Colmez. Périodes p-adiques (Bures-sur-Yvette, 1988). Astérisque **223** (1994), 59–111.

[Fo94b] J.-M. Fontaine. *Représentations p-adiques semi-stables.* With an appendix by Pierre Colmez. Périodes p-adiques (Bures-sur-Yvette, 1988). Astérisque **223** (1994), 113–184.

[Fo94c] J.-M. Fontaine. *Représentations ℓ-adiques potentiellement semi-stables.* In Périodes p-adiques, Astérisque, **223** (1994), 321–347.

[FL82] J.-M. Fontaine and G. Laffaille, *Construction de représentations p-adiques*, Ann. Sci. ENS **15** (1982), 547–608.

[FO09] J.-M. Fontaine and Y. Ouyang, *Theory of p-adic Galois Representations* staff.ustc.edu.cn/~yiouyang/galoisrep.pdf

[FM95] J.-M. Fontaine and B. Mazur, *Geometric Galois representations. Elliptic curves, modular forms, & Fermat's last theorem.* (Hong Kong, 1993), 41–78, Ser. Number Theory **I**, Int. Press, Cambridge, MA, 1995.

[Fr90] E. Freitag, *Hilbert modular forms.* Springer-Verlag, Berlin, 1990.

[GM09] E. Ghate and A. Mezard, *Filtered modules with coefficients.* Trans. Amer. Math. Soc. **361** (2009), no. 5, 2243–2261.

[Go02] E. Z. Goren, *Lectures on Hilbert modular varieties and modular forms.* With the assistance of Marc-Hubert Nicole. CRM Monograph Series **14**. AMS, Providence, RI, 2002.

[Ha07] M. Harris, *The Sato Tate conjecture: Analytic Arguments*, see `people.math.jussieu.fr/~harris/SatoTate/index.html`

[Hi06] H. Hida, *Hilbert modular forms and Iwasawa Theory*, Oxford Math. Monographs, Oxford University Press, Oxford, 2006.

[HSBT] M. Harris, N. Shepherd-Barron, R. Taylor, *A family of Calabi-Yau varieties and potential automorphy*, Preprint, 2006.

[JW82] U. Jannsen and K. Wingberg, *Die Struktur der absoluten Galoisgruppe p-adischer Zahlkörper.* Invent. Math. **70** (1982), no. 1, 71–98.

[KR01] C. Khare and C.S. Rajan, *The density of ramified primes in semisimple p-adic Galois representations.* Internat. Math. Res. Notices 2001, no. 12, 601–607.

[KW09] C. Khare and J.-P. Wintenberger, *Serre's modularity conjecture (I) and (II)*, Invent. math. **178** (2009), 485–504 and 505–586

[Ki09a] M. Kisin, *The Fontaine-Mazur conjecture for* GL_2, J. Amer. Math. Soc. **22** (2009), 641–690.

[Ki09b] M. Kisin, *Moduli of finite flat group schemes and modularity*, Ann. of Math. **170** (2009), 1085–1180.

[RT11] A. Raghuram and N. Tanabe, *Notes on the arithmetic of Hilbert modular forms.* arXiv:1102.1864v1

[Ra74] M. Raynaud, *Schémas en groupes de type* $(p, \ldots, p)$, Bull. SMF **102** (1974), 241–280.

[Ri85] K.A. Ribet *On ℓ-adic representations attached to modular forms II.* Glasgow Math. J. **27** (1985), 185–194.

[Ri90] K.A. Ribet, *On modular representations of* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ *arising from modular forms*, Inv. Math. **100** (1990), 431–476.

[Sai11] T. Saito, *Hilbert modular forms and p-adic Hodge theory.* Compos. Math. **145** (2009), no. 5, 1081–1113.

[Sav05] D. Savitt, *On a conjecture of Conrad, Diamond, and Taylor.* Duke Math. J. **128** (2005), no. 1, 141–197.

[Se68] J.-P. Serre, *Abelian ℓ-adic Galois representations and elliptic curves.* W.A. Benjamin, Inc, New York, 1968.

[Se77] J.-P. Serre, *Modular forms of weight one and Galois representations.* Proc. Durham Symp. on Algebraic Number Fields Acad. Press, London, 1977, 193–267.

[Se87] J.-P. Serre, *Sur les représentations modulaires de degré 2 de* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Duke Math. J. **54** (1987), no. 1, 179–230.

[Se91] J.-P. Serre, *Propriétés conjecturales des groupes de Galois motiviques et des représentations ℓ-adiques.* In Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math. **55**, AMS Providence, RI, 1994, 377–400.

[ST68] J.-P. Serre and J. Tate *Good Reduction of Abelian Varieties.* Ann. Math. **88** (1968), no. 3, 492–517.

[Sh71] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Forms*, Publ. Math. Soc. Japan **11**, Princeton University Press, Princeton, NJ, 1971.

[Sh98] G. Shimura, *Abelian Varieties with Complex Multiplication and Modular Functions*, Princeton University Press, Princeton, NJ, 1998.

[Si85] J. Silverman, *The Arithmetic of Elliptic Curves*, Springer GTM **106** (1985), Springer New York.

[Si91] J. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer GTM **151** (1991), Springer New York.

[Sk09] C. Skinner, *A note on the p-adic Galois representations attached to Hilbert modular forms*, Documenta Math. **14** (2009) 241–258.

[Sn09] A. Snowden, *On two dimensional weight two odd representations of totally real fields*, arXiv:0905.4266v1

[Tat67] J. Tate, *p-divisible groups*, in "Proc. Conf. on Local Fields" (Driebergen), Springer-Verlag, 1967, 158–183.

[Tat79] J. Tate: *Number Theoretic Background*, Proceedings Symp. Pure Math. **33** (1979), 3–26.

[Tay89] R. Taylor, *On Galois representations associated to Hilbert modular forms.* Invent. Math. **98** (1989), no. 2, 265–280.

[Tay91] R. Taylor, *Galois representations associated to Siegel modular forms of low weight*, Duke Math. J. **63** (1991), p. 281–332.

[Tay97] R. Taylor, *On Galois representations associated to Hilbert modular forms II.* in Elliptic Curves, Modular Forms & Fermat's Last Theorem (Hong Kong, 1993), J. Coates and S.-T. Yau, eds., International Press, 1997, 333–340.

[Tay06] R. Taylor, *On the meromorphic continuation of degree two L-functions*, Doc. Math. (2006), Extra Vol., 729–779.

[Tay08] R. Taylor, *Automorphy for some ℓ-adic lifts of automorphic mod l Galois representations. II*, Pub. Math. IHES **108** (2008), 183–239.

[TW95] R. Taylor and A. Wiles,*Ring-theoretic properties of certain Hecke algebras.* Ann. Math. (2), **141** (1995), no. 3, 553–572.

[Wi95] A. Wiles, *Modular elliptic curves and Fermat's last theorem.* Ann. Math. (2), **141** (1995), no. 3, 443–551.

Gebhard Böckle
Computational Arithmetic Geometry IWR (Interdisciplinary Center for Scientific Computing), University of Heidelberg
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany
gebhard.boeckle@iwr.uni-heidelberg.de

Travaux mathématiques, Volume 23 (2013), 37–86, © by the author

Patching in Algebra

by David Harbater

Abstract

Patching methods have been used in algebra to perform constructions in Galois theory and related areas. More recently, this approach has also been used to obtain local-global principles over function fields of arithmetic curves. These in turn have applications to structures such as quadratic forms and central simple algebras. This article surveys these developments as well as giving background and examples.

MSC (2010): 11E72, 13J10, 14H25 (Primary) 12E30, 14G05, 20G15 (Secondary).

Contents

1 Introduction and motivation

This article corresponds to the lecture series given by the author at the Luxembourg Winter School, on the topic of patching in algebra.

Those lectures were complementary to the ones delivered by Moshe Jarden. His lectures focused on patching in Galois theory, whereas these focused more on applications of patching in other aspects of algebra, including variants of Galois theory.

This article is also complementary to the author's long manuscript [Hrb03] on patching and Galois theory. That manuscript focused on the approaches of formal and rigid patching. In contrast, the Luxembourg lectures of Moshe Jarden focused on the approach of algebraic patching, and the lectures presented here focused on an approach called patching over fields.

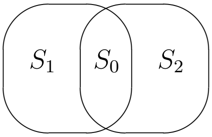
The lectures presented here were informal in tone, emphasizing the ideas and intuition, providing pictures, and sketching proofs in key cases, rather than seeking generality and completeness. A similar approach is taken here. The presentation draws heavily on [HH10], [HHK09], and [HHK11a], co-authored with Julia Hartmann and with Daniel Krashen. The reader is referred to those manuscripts for more detail.

In the lectures, and in this manuscript, we describe patching for vector spaces and related algebraic objects over function fields of curves that are defined over complete discretely valued fields such as $k((t))$. We give applications to inverse Galois theory and related questions. After describing variants on the patching set-up to allow greater generality and flexibility, we consider the problem of patching torsors, which are useful in classifying many other types of algebraic objects. We then build on the results about torsors to study the notion of local-global principles. Such principles, which are analogous to classical results about global fields, can be regarded in a sense as a complement to patching (studying given global objects by looking locally, rather than constructing global objects by using local ones). Applications are then given to quadratic forms and central simple algebras.

First, we begin with a discussion of the historical background and context of patching.

1.1 Patching in analysis

The idea for patching originated in analysis in the nineteenth century, to construct global objects from local objects; i.e. from objects that are defined on subsets and agree on overlaps. A typical situation is represented by this picture, where objects over S_1 and S_2 that agree over $S_0 = S_1 \cap S_2$ are patched to yield an object over $S = S_1 \cup S_2$:



$S = S_1 \cup S_2$

In the analytic situation, the sets S_i could be open subsets of a complex manifold. The objects to be patched might be vector bundles, or perhaps finite extensions of the fields of meromorphic functions on the sets S_i.

An early application of this approach concerned Hilbert's 21st problem, on the existence of a linear linear differential equation having a prescribed monodromy group. See Section 3.2 below for a further discussion of this, including more recent algebraic generalizations.

1.2 From analysis to algebra

Serre's paper GAGA [Ser56] made it possible to pass between complex algebraic geometry (in which one uses classical metric open sets) and algebraic geometry with respect to the Zariski topology. In particular, it showed that the cohomology of a coherent sheaf on a complex projective algebraic variety agrees with the cohomology of the induced sheaf in the analytic topology. As a result, one can, for example, cite results in the text by Griffiths-Harris ([GH78]) in the context of the text by Hartshorne ([Hts77]).

Serre's result can be viewed as a form of patching that passes from analysis to algebra. Namely, one can cover a complex algebraic variety V by metric open sets S_i. GAGA says, for example, that giving holomorphic differential forms ω_i on S_i for all i, which agree on the overlaps $S_i \cap S_j$, is equivalent to giving a *regular* (algebraic) differential form ω on V.

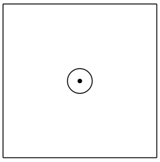
For a more detailed discussion of GAGA from the perspective of patching, see Section 2 of [Hrb03].

1.3 Patching in algebra

The main difficulty in carrying over the patching approach to a more purely algebraic situation is that non-empty Zariski open subsets are very large — in fact, in an irreducible variety, they are dense, with the complement being of lower dimension. Thus, for example, if the objects to be patched are finite extensions of the fields F_i of rational functions on the sets $S_i \subset S$, then the fields F_i are just isomorphic to the field F of rational functions on the variety S itself, and the procedure would not yield anything new. In order to be able to construct new extensions of the field F (e.g. for purposes of Galois theory), one would need smaller sets S_i to consider, whose function fields would strictly contain F.

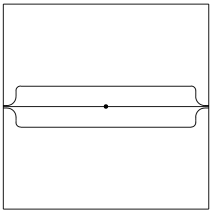
Ironically, a way of dealing with the limitation of the Zariski topology is due to Zariski himself. His approach used completions. As a first example, consider the affine (x, t)-plane $\mathbb{A}^2_k = \mathrm{Spec}(k[x, t])$ over a field k. Viewing formal power series as "functions defined near the origin," we can take the ring $k[[x, t]]$ and its spectrum $\mathrm{Spec}(k[[x, t]])$, viewing this as a small neighborhood of the origin. Here the power series ring $k[[x, t]]$ is the completion of the ring $k[x, t]$ at the maximal ideal $\mathfrak{m} = (x, y)$ corresponding to the origin.

The corresponding picture is:



The function field of $\mathrm{Spec}(k[[x, t]])$ is the fraction field $k((x, t))$ of $k[[x, t]]$, called the field of Laurent series in x, t over k. (Note, though, that this fraction field is not complete. Also, the elements in this field are not necessarily series in x and t; e.g. $1/(x + t)$.)

As a more sophisticated example, we can consider a "small neighborhood" of the x-axis in $\mathbb{A}^2_k$. The x-axis is defined by the ideal (t) in $k[x, t]$; and motivated by the previous example we can take the t-adic completion of $k[x, t]$. This is $\lim\limits_{n\to\infty} k[x, t]/(t^n) = k[x][[t]]$. Its spectrum can be viewed as a "tubular neighborhood" of the x-axis in $\mathbb{A}^2_k$, made up by a union of small neighborhoods of all the closed points on the x-axis. Intuitively, this tubular neighborhood "pinches down" near the point at infinity, since that point is not on the (affine) x-axis.



Note, for example, that the element $x - t$ defines a curve in the affine (x, t)-plane that meets this neighborhood non-trivially, intersecting the x-axis at the origin; and correspondingly, $(x - t)$ is a proper ideal in the ring $k[x][[t]]$. On the other hand, the element $1 - xt$ defines a curve in the affine plane that does not meet this tubular neighborhood, since it approaches the x-axis at $x = \infty$; and indeed, $(1 - xt)$ is the unit ideal in $k[x][[t]]$.

1.4 Formal patching

Grothendieck developed the above idea into the theory of formal schemes. Using that theory, one can do *formal patching.* In this situation, one can regard a ring like $k[x][[t]]$ as analogous to the ring of holomorphic functions on a complex metric open set; and by giving objects over such rings one can obtain a more global object.

The key theorem is a analog of Serre's GAGA in the context of formal schemes. It was referred to as GFGA in Grothendieck's paper [Gro59] of the same name, to emphasize the parallel and to explain that it related the formal and algebraic contexts, just as Serre's result related the analytic and algebraic contexts. This result later came to be known as Grothendieck's Existence Theorem. See [Gro61, Corollaire 5.1.6].

Beginning the 1980's, I used this approach in Galois theory, e.g. to realize all finite groups as Galois groups over fields such as $K(x)$ where K is the fraction field of a complete local domain that is itself not a field (e.g. $K = \mathbb{Q}_p$ for some prime p, or $K = k((t))$ for some field k). See [Hrb87]. Other results included the freeness of the absolute Galois group of $k(x)$ for k an arbitrary algebraically closed field ([Hrb95]), and the proof of Abhyankar's Conjecture on the Galois groups of étale covers of affine curves in characteristic p ([Hrb94]). See Section 1.5 below for a further discussion of such results and related work of others. Also see Section 3 of [Hrb03].

1.5 Rigid patching

Another form of patching is based on Tate's theory of rigid analytic spaces (see [Tat71]). This theory is modeled on that of complex analytic spaces, but with differences to account for the fact that the topology induced by an ideal is totally disconnected. These differences allow analytic continuation to be unique, as when working over $\mathbb{C}$, thereby leading to a "rigid" structure. (In contrast, a "floppy" structure that would result from using a more naive construction in this totally disconnected context.)

Rigid analytic spaces are phrased in terms of convergent power series in a non-archimedean metric. For example, for the affine x-line over $K = k((t))$, the ring $K\{x\}$ of power series in $K[[x]]$ that are convergent on the closed t-adic unit disc turns out to be the same as the ring $k[x][[t]][t^{-1}]$, which is a localization of the ring considered above. In fact, Raynaud later reinterpreted rigid analytic spaces in terms of formal schemes, roughly giving a dictionary between the two frameworks in the case that K is a complete discretely valued field. (See [Ray74].)

In the early 1990's, Serre suggested that results in Galois theory could also be obtained via rigid patching. Using this approach, results in [Hrb87] were then reproved by Q. Liu in [Liu95]. Later, this approach was used to prove many other results related to Galois theory. In [Ray94], M. Raynaud used rigid patching (and other methods) to prove Abhyankar's Conjecture in the case of the affine line;

this was one of the ingredients in the proof of the general case in [Hrb94]. F. Pop proved a number of results on the structure of absolute Galois groups of function fields in a series of papers including [Pop94] and [Pop95], the latter of which gave a rigid proof of the result proved at the same time in [Hrb95] using formal patching. See Section 4 of [Hrb03] for more about the use of rigid patching in Galois theory.

1.6 Algebraic patching

In the mid to late 1990's, another framework for patching was established, in work of D. Haran, M. Jarden and H. Völklein. This framework, called *algebraic patching*, avoids the machinery of Grothendieck and Tate, and is designed to isolate what is needed for applications to Galois theory. As the name indicates, the approach avoids mention of geometric objects, preferring to focus on the rings and fields involved. Like rigid patching, it relies on convergent power series with respect to a non-archimedean metric, but without mention of rigid analytic spaces. Key ingredients include versions of Cartan's Lemma and the Weierstrass Preparation Theorem.

Some of the results that had been shown by formal or rigid methods were reproven in this framework, as well as additional results concerning the Galois theory of function fields of curves. See [HV96] and [HJ98] for early papers in this direction. See also the lecture notes of Moshe Jarden at the Luxembourg Winter School, and also the volumes [Jar11] and [Völ96].

1.7 Patching over fields

Beginning in 2006, a framework of *patching over fields* was developed by Julia Hartmann and myself in [HH10], for the purpose of making patching more applicable to other algebraic contexts, and also to avoid heavy machinery. It uses and emphasizes the fraction fields of the rings that appear in formal and rigid approaches. In this way, the spaces that are used in those earlier approaches, and sheaves of functions on those spaces, are replaced by fields and vector spaces over them. As a consequence, this approach is more elementary in nature, but more general in application, since it permits uses in situations in which the objects are inherently defined over fields rather than rings or spaces.

Like algebraic patching, it relies on a form of Cartan's Lemma, and it uses a form of Weierstrass Preparation. But like formal patching, it relies on adic completions of rings (e.g. formal power series) and their fraction fields, rather than on convergent power series.

This approach is being used in work on quadratic forms, central simple algebras, and analogs of Galois theory (Galois theory of differential equations and Galois theory of division algebras). See in particular [HHK09] and [HHK11a], written jointly with Julia Hartmann and Daniel Krashen.

The remainder of this lecture series discusses this framework and some of its uses.

2 Patching algebraic structures

We begin by describing patching over fields in a basic situation: where the objects being patched are finite dimensional vector spaces. Afterwards we turn to other types of algebraic structures.

2.1 Patching vector spaces over fields

Motivated by the ideas in Section 1, we consider four fields F, F_1, F_2, F_0 that fit into a commutative diagram

(2.1)

$$\begin{tikzcd}
& F_0 & \\
F_1 \arrow[ur, dash] & & F_2 \arrow[ul, dash] \\
& F \arrow[ul, dash] \arrow[ur, dash] &
\end{tikzcd}$$

where the lines represent inclusions. We assume here that F is the intersection of F_1 and F_2 inside the common overfield F_0.

As an example, we may consider the situation discussed before: a space S is covered by two subsets S_1, S_2; and S_0 is the intersection of $S_1 \cap S_2$. In this context we take F to be the field of (rational or meromorphic) functions on S and we let F_i be the field of functions on S_i, for $i = 0, 1, 2$. Given some structure over S_1 and over S_2, together with an isomorphism between their restrictions to S_0, we wish to "patch" them together in order to obtain a structure over the full space S that induces the given structures compatibly. From the algebraic point of view, given structures over F_1, F_2 with an isomorphism between the structures they induce over F_0, we want to show that there is a unique structure over F that induces them compatibly.

To make this more precise, we need to interpret the notion of "structure". As a first case, take the structures to be finite dimensional vector spaces. Given a field E, let $\mathrm{Vect}(E)$ be the category of finite dimensional vector spaces over E. If $E \subseteq E'$, there is a functor $\mathrm{Vect}(E) \to \mathrm{Vect}(E')$ given by $V \mapsto V \otimes_E E'$. For a diagram of four fields as in (2.1) above, there is a base change functor

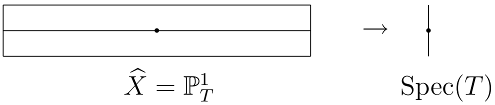
$$\Phi : \mathrm{Vect}(F) \to \mathrm{Vect}(F_1) \times_{\mathrm{Vect}(F_0)} \mathrm{Vect}(F_2), \tag{2.2}$$

where the objects in the right hand category consist of triples (V_1, V_2, μ), with V_1 in $\mathrm{Vect}(F_i)$ and with μ an isomorphism of F_0-vector spaces $V_1 \otimes_{F_1} F_0 \to V_2 \otimes_{F_1} F_0$. (This is roughly a fiber product of categories. Actually, because of the choice of μ, it is not exactly a fiber product, but rather a "2-fiber product".) Objects in this category will be called *patching problems.* The desired patching assertion is that the functor Φ is an equivalence of categories. In particular, this says that every patching problem has a *solution*; i.e. there is an object in $\mathrm{Vect}(F)$ that induces it, up to isomorphism. While we refer here to finite dimensional vector spaces, we will also later consider the corresponding situation for other interpretations of "structure"; e.g. for finite dimensional associative algebras, or Galois extensions with a given group, etc.

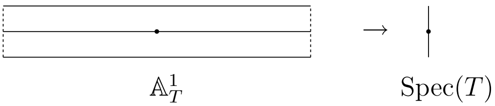
To be able to obtain an equivalence of categories as above, we need to choose the fields F_i appropriately. If we work classically (e.g. analytically, say with algebraic varieties over $\mathbb{R}$ or $\mathbb{C}$), we can take the sets S_i to be metric open sets, and take F_i to be the field of meromorphic or rational functions on S_i. But if we work with varieties over more general fields, there is the problem that the Zariski topology is too coarse: all open sets in an irreducible variety have the *same* function field. Our approach will be to consider varieties over fields K which (like $\mathbb{R}$ and $\mathbb{C}$) are complete with respect to a metric. Namely, we can take a base field K that is a complete discretely valued field such as $k((t))$ or $\mathbb{Q}_p$; i.e. the fraction field of a complete discrete valuation ring T (in these two examples, $k[[t]]$ and $\mathbb{Z}_p$).

2.2 Basic example: the line

As an example, let $T = k[[t]]$ and $K = k((t))$, and consider the T-curve $\widehat{X} = \mathbb{P}^1_T$, the projective x-line over T. There is then a structure morphism $\widehat{X} \to \mathrm{Spec}(T)$, where $\mathrm{Spec}(T)$ consists of two points: a closed point corresponding to the maximal ideal (t) of T, and the generic point corresponding to the zero ideal of T. The closed point is a copy of $\mathrm{Spec}(k)$, and the generic (open) point of $\mathrm{Spec}(T)$ is a copy of $\mathrm{Spec}(K)$. We may view $\mathrm{Spec}(T)$ as a "small neighborhood" of the origin in the t-line over k. The corresponding picture is:



Here the fiber X of $\widehat{X}$ over the closed point of $\mathrm{Spec}(T)$ is a copy of $\mathbb{P}^1_k$, and the fiber over the generic point of $\mathrm{Spec}(T)$ is a copy of $\mathbb{P}^1_K$. The affine T-line $\mathbb{A}^1_T = \mathrm{Spec}(k[[t]][x])$ is a Zariski open subset of $\widehat{X}$, and its fibers over the closed and generic points of $\mathrm{Spec}(T)$ are the affine lines over k and K, respectively.



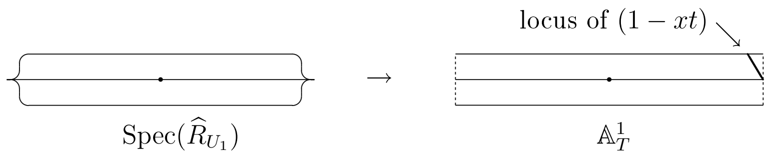
The field of rational functions F on $\widehat{X}$ is the same as the field of rational functions on this open subset, viz. $k((t))(x)$ (this being the fraction field of $k[[t]][x]$).

We can instead consider the affine x-line over the base ring $T = \mathbb{Z}_p$ of p-adic integers, with uniformizer $t := p$. In this situation the residue field k is $\mathbb{F}_p$ and the fraction field K is $\mathbb{Q}_p$. The schematic pictures look the same as above, and the above discussion carries over to this case, with $k[[t]][x]$ replaced by $\mathbb{Z}_p[x]$ and $k((t))(x)$ replaced by $\mathbb{Q}_p(x)$.

2.3 Subsets and overfields

In general, consider a proper smooth curve $\widehat{X}$ over $\mathrm{Spec}(T)$, with T a complete discrete valuation ring having uniformizer t, and let F be the field of rational functions on $\widehat{X}$. Let $X \subset \widehat{X}$ be the closed fiber. For any subset $U \subset X$ that does not contain all the closed points of X, let R_U be the ring of rational functions on $\widehat{X}$ that restrict to rational functions on X and which are regular at the points of U. Let $\widehat{R}_U$ be the t-adic completion of R_U. Note that if U contains at least one closed point of X, then R_U and $\widehat{R}_U$ are two-dimensional domains; whereas in the case $U = \varnothing$, the rings $R_\varnothing$ and $\widehat{R}_\varnothing$ are discrete valuation rings. With U as above, the fraction field F_U of $\widehat{R}_U$ is an overfield of F. We also write $F_X = F$.

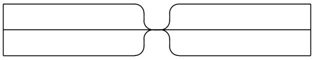
In the above example of the projective T-line, we can take the open set $U_1 = \mathbb{A}^1_k = \mathrm{Spec}(k[x])$, the affine x-line over the residue field k of T. This is the complement of the point $x = \infty$ in $X = \mathbb{P}^1_k$. In the case that $T = k[[t]]$, the ring R_{U_1} is the subring of $F = k((t))(x)$ consisting of rational functions f/g where $f, g \in k[[t]][x]$ such that g does not vanish anywhere on U_1. That is, the reduction of g modulo t is a unit in $k[x]$, or equivalently a non-zero constant in k. The t-adic completion $\widehat{R}_{U_1}$ is given by $k[x][[t]]$. Observe that this ring strictly contains $k[[t]][x]$, the ring of regular functions on $\mathbb{A}^1_T$. For example, $\sum_0^\infty x^i t^i$ is contained in the former, but not the latter. This inclusion of rings corresponds to a morphism of their spectra in the other direction, viz. $\mathrm{Spec}(\widehat{R}_{U_1}) \to \mathbb{A}^1_T$. Note that $\sum_0^\infty x^i t^i = (1 - xt)^{-1} \in k[x][[t]]$, and so $(1 - xt)$ is the unit ideal in $k[x][[t]]$, and does not correspond to a point of $\mathrm{Spec}(\widehat{R}_{U_1})$. But $(1 - xt)$ does generate a prime ideal in $k[[t]][x]$, and does define a non-empty closed set in $\mathrm{Spec}(k[[t]][x]) = \mathbb{A}^1_T$. (Cf. the discussion in Section 1.3.)



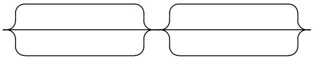
Similarly, the fraction field $F = k((t))(x)$ of $k[[t]][x]$ is strictly contained in the fraction field F_{U_1} of $\widehat{R}_{U_1} = k[x][[t]]$, which is in fact transcendental over F. Intuitively, we can regard $\mathrm{Spec}(\widehat{R}_{U_1})$ as an analytic open subset of the algebraic

(Zariski) open subset $\mathbb{A}^1_T \subset \widehat{X}$; and regard $k[x][[t]]$ and its fraction field as consisting of the holomorphic and meromorphic functions on this set.

In the above example, we can also consider the open set $U_2 \subset X$ given by the complement of the point $x = 0$ on $X = \mathbb{P}^1_k$. This is another copy of the affine line over k, with ring of functions $k[x^{-1}]$. We then obtain $\widehat{R}_{U_2} = k[x^{-1}][[t]]$, with fraction field F_{U_2}. Let $U_0 = U_1 \cap U_2$; this is the complement of the two points $x = 0, \infty$ in X. We then have $\widehat{R}_0 = k[x, x^{-1}][[t]]$, with fraction field F_{U_0}.



$\mathrm{Spec}(\widehat{R}_{U_2})$



$\mathrm{Spec}(\widehat{R}_{U_0})$

If we take U to be the empty subset of X (or equivalently, the subset consisting just of the generic point of X), then $\widehat{R}_U = \widehat{R}_\varnothing = k(x)[[t]]$, which is a complete discrete valuation ring (unlike the rings above, which were two dimensional). Its quotient field is $F_\varnothing = k(x)((t))$.

As in Section 2.2, we can instead consider the analogous case of the x-line over $T = \mathbb{Z}_p$. The schematic pictures are again the same as in the power series case, though the rings are a bit more awkward to write down explicitly. With U_1 the affine line as above, the ring $\widehat{R}_{U_1}$ is the p-adic completion of $\mathbb{Z}_p[x]$, i.e. $\varprojlim \mathbb{Z}_p[x]/(p^n)$. For example $1 - px$ is a unit in this ring, with inverse $\sum_0^\infty p^i x^i$. On the other hand, the elements x, $x - p$, and p each define proper principal ideals, corresponding to curves in $\mathrm{Spec}(\widehat{R}_{U_1})$. Similarly, if we let U_2 be the complement of ∞ in $\mathbb{P}^1_T$ and let U_0 be the complement of the two points $0, \infty$ in $\mathbb{P}^1_T$, then $\widehat{R}_{U_2}$ and $\widehat{R}_{U_0}$ are the p-adic completions of $\mathbb{Z}_p[x^{-1}]$ and $\mathbb{Z}_p[x, x^{-1}]$. Note also that in this situation, $\widehat{R}_\varnothing$ is the p-adic completion of the discrete valuation ring $\mathbb{Z}_p[x]_{(p)}$. Its fraction field $F_\varnothing$ is the same as the completion of the field $\mathbb{Q}_p(x)$ with respect to the Gauss valuation; i.e. with respect to the metric induced on this field by the ideal $(p) \subset \mathbb{Z}_p[x]_{(p)}$.

2.4 Solutions to patching problems

As the above pictures suggest, we can regard $\widehat{X}$ as covered by the "analytic open sets" $S_1 := \mathrm{Spec}(\widehat{R}_{U_1})$ and $S_2 := \mathrm{Spec}(\widehat{R}_{U_2})$, with $S_0 := \mathrm{Spec}(\widehat{R}_{U_0})$ as the intersection of these sets. In fact, this can be made precise: the natural morphisms $S_0 \to S_i \to \widehat{X}$ $(i = 1, 2)$ define injections on the underlying sets of points, with the images of S_1 and S_2 covering $\widehat{X}$, and with the image of S_0 in $\widehat{X}$ being the intersection of those two images. Moreover, as we discuss below in Section 5, the fields F and $F_i := F_{U_i}$ $(i = 0, 1, 2)$ form a diagram of fields as in (2.1); and the corresponding functor Φ as in (2.2) turns out to be an equivalence of categories.

More generally, let $\widehat{X}$ be any smooth projective T-curve with function field F and closed fiber X, and let U_1, U_2 be subsets of X with $U_1 \cup U_2 = X$. Suppose that neither U_i contains all the closed points of X. Write $U_0 = U_1 \cap U_2$ and $F_i := F_{U_i}$. Then the base change functor $\Phi : \mathrm{Vect}(F) \to \mathrm{Vect}(F_1) \times_{\mathrm{Vect}(F_0)} \mathrm{Vect}(F_2)$ is an equivalence of categories. Thus given finite dimensional vector spaces V_i over F_i ($i = 1, 2$) and an F_0-isomorphism between the F_0-vector spaces $V_i \otimes_{F_i} F_0$ that they induce, there is a unique finite dimensional F-vector space V inducing them *compatibly*. Moreover, if we identify V_i with its isomorphic image in $V_i \otimes_{F_i} F_0$, and if we identify $V_1 \otimes_{F_1} F_0$ with $V_0 := V_2 \otimes_{F_2} F_0$ via the isomorphism μ, then we get that V is equal to $V_1 \cap V_2$ inside V_0. In particular, if we let $n = \dim_{F_1}(V_1)$ (which is equal to $\dim_{F_2}(V_2)$), then $\dim_F(V_1 \cap V_2)$ is necessarily equal to n. See Theorem 5.3 below.

2.5 Patching for other objects

While the above specifically concerns finite dimensional vector spaces, many algebraic objects over a field consist of a finite dimensional vector space together with additional structure that is given by maps such that certain diagrams commute.

For example, consider finite dimensional associative algebras A over F. To give A is to give a finite dimensional F-vector space A together with an F-vector space homomorphism $A \otimes_F A \to A$ such that a certain diagram commutes (corresponding to the associative law). The above equivalence of categories for vector spaces (which preserves tensor products), together with the assertion that $V = V_1 \cap V_2$, yields the corresponding equivalence of categories for finite dimensional associative algebras, again with inverse given by intersection.

As explained in [HH10, Section 7], some other examples of algebraic objects for which equivalences of categories follow in this manner are these: finite dimensional associative F-algebras with identity; finite dimensional separable F-algebras; central simple F-algebras (i.e. the center is F and there are no non-trivial two-sided ideals); differential F-modules (i.e. finite dimensional F-vector spaces together with a derivation over F); separable F-algebras (i.e. products of finitely many finite separable field extensions of F); and G-Galois (commutative) F-algebras (for some finite group G).

An object of this last sort is by definition a separable F-algebra A of dimension equal to $|G|$ whose fixed field under G is F. These are of the form $\prod E$, a product of copies of a Galois field extension E/F of Galois group $H \subseteq G$, indexed by the cosets of H in G. If $H = G$, this is the same as a G-Galois field extension. At the other extreme, if $H = 1$, then A is a product of copies of F indexed by and permuted by G; this is called a *trivial* G-Galois F-algebra. If E/F is any G-Galois F-algebra, then $E_{\widetilde{F}} := E \otimes_F \widetilde{F}$ is a G-Galois $\widetilde{F}$-algebra for any field extension $\widetilde{F}/F$; and if $\widetilde{F}/F$ is sufficiently large (e.g. the separable closure of F), then the G-Galois $\widetilde{F}$-algebra $E_{\widetilde{F}}$ is trivial.

3 Applications of patching

Using patching for algebraic structures as discussed above, one can obtain applications of various sorts. For example, using the category of G-Galois F-algebras, one can obtain applications of patching to Galois theory. Using the category of differential modules, one can obtain applications to differential Galois theory (the Galois theory of linear differential equations). And using central simple algebras, one can obtain an application to a division algebra analog of Galois theory. We discuss these in turn below.

3.1 Applications to Galois theory

Suppose that we have a finite group G that we wish to realize as a Galois group over F, the function field of a curve over a complete discretely valued field. We can proceed inductively, generating G by two strictly smaller subgroups $H_1, H_2 \subset G$. Suppose that each H_i is the Galois group of a Galois field extension E_i/F_i, where the fields F_i are overfields of F as in (2.1). Suppose in addition that such extensions can be found for which $E_i \otimes_{F_i} F_0$ is a trivial H_i-Galois F_0-algebra. Taking a product of finitely many copies of E_i, indexed by the cosets of H_i in G, we obtain G-Galois F_i-algebras A_i for $i = 1, 2$. But note that A_i is *not* a field, because the index $(G : H_i)$ is greater than one. Nevertheless, we can use these algebras to obtain the desired field extension over F, using the composition μ of isomorphisms $A_1 \otimes_{F_1} F_0 \xrightarrow{\sim} \prod_G F_0 \xrightarrow{\sim} A_2 \otimes_{F_2} F_0$ arising from the triviality of A_i over F_0. Namely, the algebras A_i together with this isomorphism define a patching problem for G-Galois algebras, i.e. an element of the category $G\mathrm{Alg}(F_1) \times_{G\mathrm{Alg}(F_0)} G\mathrm{Alg}(F_2)$, where GAlg denotes G-Galois algebras. By the equivalence of categories Φ, this is induced by an object A in $G\mathrm{Alg}(F)$, i.e. by a G-Galois F-algebra. Using the fact that H_1, H_2 generate G, together with the choice of patching isomorphism μ as above, one can show ([HH10, Section 7]) that A is actually a G-Galois *field* extension of F!

This strategy explains how to realize individual finite groups as Galois groups over the field F, provided that the subgroups can themselves be realized. Since every finite group is generated by cyclic subgroups (in fact, even by cyclic subgroups of prime power order), it is sufficient to realize those, subject to the condition of being trivial over F_0. This is easy to do by Kummer theory if F has characteristic zero and contains all roots of unity. More generally this strategy works provided that G contains primitive n-th roots of unity ζ_n, where n ranges over the orders of the cyclic generating subgroups. Without this condition, if n is not divisible by the characteristic of F, one can construct a suitable n-cyclic Galois extension of $F_i(\zeta_n)$ by Kummer theory such that it is induced by some n-cyclic Galois extension of F_i. Finally, if $\mathrm{char}(F) = p$, then Artin-Schreier theory and Witt vectors can be used to construct appropriate cyclic extensions of p-power order. As a result, it follows that every finite group is a Galois group over F. See [Hrb87], where the details of the construction of these cyclic building blocks is given, though in the context of formal patching.

One can go further, by studying the absolute Galois group of F, viz. the Galois group $\mathrm{Gal}(F) := \mathrm{Gal}(F^{\mathrm{sep}}/F)$ over F of the separable closure F^{sep} of F. This profinite group is the inverse limit of the Galois groups of the finite Galois extensions of F. So to understand the group $\mathrm{Gal}(F)$, it suffices to find all the finite Galois groups over F and how they fit together in an inverse system (corresponding to the direct system of finite Galois extensions of F). For this, one wants to know when one can solve *embedding problems*. From the perspective of fields, the question is this: Given a finite group G and a quotient group H, and given an H-Galois field extension E/F, can E be embedded in a G-Galois field extension of F? Reinterpreting this in terms of groups, it asks: Given a surjection $\pi : G \to H$ of finite groups, and a surjection $f : \mathrm{Gal}(F) \to H$, is there a surjection $\widetilde{f} : \mathrm{Gal}(F) \to G$ such that $\pi\widetilde{f} = f$? If the answer is always yes, and if $\mathrm{Gal}(F)$ is countably generated as a profinite group, then $\mathrm{Gal}(F)$ must be a free profinite group, by a theorem of Iwasawa ([Iwa53, p.567]). A generalization of this theorem, due to Melnikov and Chatzidakis, handles the uncountable case [Jar95, Lemma 2.1].

Extending the patching strategy to this context has made it possible to show that many embedding problems can be solved; and using those results, it has been shown that the absolute Galois group of $k(x)$ is free if k is an algebraically closed field of arbitrary characteristic. (See [Hrb95], [Pop95], [HJ00], which respectively describe this proof in terms of formal, rigid and algebraic patching. This theorem had previously been shown just in characteristic zero, by relating k to $\mathbb{C}$; see [Dou64].) Note that $k(x)$ is *not* a field of the form F considered above, though $k((t))(x)$ is of that form. But it is possible to use results about embedding problems over $k((t))(x)$ to obtain such results over the field $k(x)$, by passing to a subfield of $k((t))(x)$ of finite type and then using that a variety over an algebraically closed field k has k-points. For more about this direction, see Section 5 of [Hrb03] and also the Luxembourg notes of Moshe Jarden.

3.2 Applications to differential algebra

Here we consider patching in the context of a *differential field*, i.e. a field F together with a derivation ∂. We assume that $\mathrm{char}(F) = 0$, to avoid the unpleasant situation in which $\partial(x^p) = 0$. The *field of constants* in F is the subfield C of elements $c \in F$ such that $\partial(c) = 0$. (For example, in the differential field $F = K(x)$ with derivation $\partial/\partial x$, the constant field is K, using $\mathrm{char}(F) = 0$.) A *differential module* M over F is a finite dimensional F-vector space together with an F-derivation ∂_M on M. That is, $\partial : M \to M$ is an additive map such that $\partial_M(fm) = \partial(f)m + f\partial_M(m)$ for $f \in F$ and $m \in M$.

We consider in particular a field F as before, i.e. the field of rational functions on a smooth projective curve $\widehat{X}$ over a complete discretely valued field T. Let U_1, U_2 be subsets of the closed fiber X, such that $X = U_1 \cup U_2$ and neither U_i contains all the closed points of X, and let $U_0 = U_1 \cap U_2$. Write $F_i = F_{U_i}$ for each i. Once we give F the structure of a differential field whose constant field is the fraction field K of T (e.g. taking $\partial = \partial/\partial x$ if $\widehat{X} = \mathbb{P}^1_T$), the overfields F_i obtain compatible structures as differential fields.

Write $\mathrm{DiffMod}(E)$ for the category of differential modules over a field E. As noted in Section 2.5, in the above context the functor

$$\Phi : \mathrm{DiffMod}(F) \to \mathrm{DiffMod}(F_1) \times_{\mathrm{DiffMod}(F_0)} \mathrm{DiffMod}(F_2),$$

is an equivalence of categories, and in particular every patching problem for differential modules has a solution.

This can be used to prove the analog of the inverse Galois problem in the context of differential algebra. Whereas ordinary Galois theory originated in the study of polynomial equations, differential Galois theory over a differential field F originated in the study of linear differential equations, which give rise to differential modules. Given such an equation (or the associated differential module), there is an analog of the splitting field of a polynomial, called the associated *Picard-Vessiot extension*. This is a differential field extension E/F that is generated by the solutions to the differential equation, and such that the field of constants of E is the same as the field of constants of F. It is known that a Picard-Vessiot extension E exists and is unique if the constant field C is algebraically closed (unlike for the fields F we have been considering above); for more general fields there are additional subtleties. The associated *differential Galois group* is the automorphism group of E/F as an extension of differential fields. This is a *linear algebraic group*, i.e. a (smooth) Zariski closed subgroup of $\mathrm{GL}_{n,C}$. See [MP03, Section 2] for more details.

The inverse differential Galois problem asks whether every linear algebraic group over C is a differential Galois group over F. The case that $C = \mathbb{C}$ was proven in [TT79], building on classical work of Plemelj on Hilbert's 21st problem concerning the realization of groups as monodromy groups of linear differential equations. In [Hrt05], it was shown that if $F = C(x)$ with C any algebraically closed of characteristic zero, then every linear algebraic group over C is the differential Galois group of some differential module. Using patching of differential modules, it can be shown that the same is true if F is instead a field as considered in the general discussion above, viz. the function field of a curve over a complete discretely valued field K of characteristic zero. (See [Hrt07].) Using this, it is possible to obtain another proof of the result for function fields over an algebraically closed field of characteristic zero.

In connection with the above historical comments, it is worth mentioning that the classical work of Plemelj, Birkhoff and others in this area relied on analytic patching methods (e.g. see [Bir17]). A key idea was the use of matrix factorization. While the term "Riemann-Hilbert problem" was initially used to refer to versions of Hilbert's 21st problem, it has also come to mean problems related to matrix factorization in an analytic context. In other contexts involving patching, matrix factorization also plays a key role for related reasons, often under the name "Cartan's Lemma". See [Hrb03, Section 2.2] for a discussion of its use in Serre's GAGA; [Hrb84, Section 2] for its use in formal patching; the Luxembourg lectures of Moshe Jarden for its use in algebraic patching; and see Section 4 below for its use in patching over fields.

3.3 Application to Galois theory of division algebras

Given a field F, a (central) *division algebra* A over F is a finite dimensional associative F-algebra with identity such that A is a division ring and the center of A is equal to F. Two examples are the Hamilton quaternion algebra $\mathbb{H}$ over the field $\mathbb{R}$, and the $n \times n$ matrix algebra $\mathrm{Mat}_n(F)$ over any field F. The F-dimension of A is necessarily a square d^2, and d is called the *degree* of A as an F-algebra. Every subfield E of A that contains F satisfies $[E : F] \le d$, and moreover there exists such a subfield E with $[E : F] = d$. Such a field E is called *maximal*. If such a maximal subfield $E \subseteq A$ is Galois over F, say with Galois group G, then the algebra structure of A can be described rather explicitly, and in these terms A is called a *crossed-product F-algebra* for G. (For example, see [Pie82].)

In [Sch68], Schacher defined a finite group G to be *admissible* over a field F if there is an F-division algebra A that contains a maximal subfield E that is Galois over F with group G. Equivalently, the condition is that there is a crossed-product F-algebra A for G. One can then ask the following analog of the inverse Galois problem: Given a field F, which finite groups are admissible over F?

Unlike the case of the usual inverse Galois problem, it is known that *not* every finite group is admissible over $\mathbb{Q}$. In [Sch68], Schacher showed that a necessary condition for a group G to be admissible over $\mathbb{Q}$ is that every Sylow subgroup of G is metacyclic (i.e. the extension of a cyclic group by a cyclic group). He conjectured that the converse should be true; but this remains open, although it is known in the case that the group is solvable ([Son83]). Work on this problem has also been done more generally for global fields, i.e. fields F that are finite extensions of either $\mathbb{Q}$ or $\mathbb{F}_p(x)$ for some prime p (e.g. see [Sch68, Corollary 10.3] in the function field case).

Using patching, the admissibility problem can be studied in the case that the field F is of the type that we have been considering in this manuscript, i.e. a finitely generated field of transcendence degree one over a complete discretely valued field K. In particular, in the case that $K = \mathbb{C}((t))$, it was shown in [HHK11] via patching over fields that a finite group G is admissible over F if and only if every Sylow subgroup of G is abelian metacyclic (or equivalently, a direct product of two cyclic groups). Here, the abelian condition is related to the fact that this choice of K contains all the roots of unity.

4 Criterion for patching

We return to a more general situation, as at the beginning of Section 2, with four fields forming a diagram of inclusions:

$$\begin{tikzcd}
 & F_0 & \\
F_1 \arrow[ur, dash] & & F_2 \arrow[ul, dash] \\
 & F \arrow[ul, dash] \arrow[ur, dash] &
\end{tikzcd}$$

We will give a criterion for patching to hold for vector spaces over these fields, for use in obtaining the type of results discussed above. That is, we give a necessary and sufficient condition for the functor

$$\Phi : \mathrm{Vect}(F) \to \mathrm{Vect}(F_1) \times_{\mathrm{Vect}(F_0)} \mathrm{Vect}(F_2),$$

to be an equivalence of categories.

To state this criterion, we consider two conditions:

Condition 4.1. *Factorization property ("Cartan's Lemma"): For every $n \geq 1$, every matrix $A_0 \in \mathrm{GL}_n(F_0)$ can be factored as $A_1 A_2$ with $A_i \in \mathrm{GL}_n(F_i)$.*

Condition 4.2. *Intersection property: $F = F_1 \cap F_2 \subseteq F_0$.*

As at the beginning of Section 2, Condition 4.2 is intuitively plausible if we think of F as the field of rational functions on a space S; F_1, F_2 as the fields of functions on two subspaces S_1, S_2 with $S_1 \cup S_2 = S$; and F_0 as the field of functions on $S_0 = S_1 \cap S_2$. Note also that one may write the factorization with one of the factors A_i replaced by A_i^{-1}; sometimes this is a more natural way to write it.

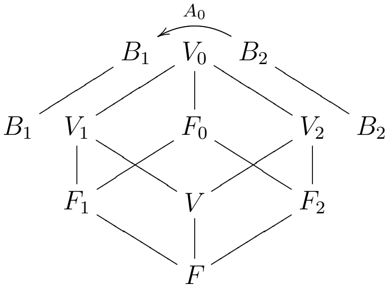
Theorem 4.3. *Φ is an equivalence of categories if and only if Conditions 4.1 and 4.2 hold.*

Proof. For the reverse direction, the key assertion to show is that Conditions 4.1 and 4.2 imply that every patching problem has a solution. To do this, suppose that we are given a vector space patching problem $\mathcal{V}$ for fields satisfying these conditions. This patching problem consists of finite dimensional F_i-vector spaces V_i for $i = 1, 2$, of a common dimension n, together with an isomorphism

$$\mu : V_1 \otimes_{F_1} F_0 \to V_2 \otimes_{F_1} F_0 =: V_0$$

of F_0-vector spaces. We may then identify $V_1 \otimes_{F_1} F_0$ with V_0 via μ, and we can thus view V_1, V_2 as subsets of V_0. With respect to these inclusions, we will show that the intersection $V := V_1 \cap V_2$, viewed as a vector space over F, is a solution to the given patching problem.

To do this, we will find a common basis B for V_1, V_2 over the fields F_1, F_2 respectively, and will show that V is the F-span of B. First, let B_i be any F_i-basis for V_i, for $i = 1, 2$. Thus B_1, B_2 are both F_0-bases of V_0, and there is a transition matrix $A_0 \in \mathrm{GL}_n(F_0)$ between them, satisfying $B_1 = A_0(B_2)$.



Next, by Condition 4.1, we may write $A_0 = A_1 A_2$ with $A_i \in \mathrm{GL}_n(F_i)$, $i = 1, 2$. Let $B = A_2(B_2) = A_1^{-1}(B_1)$. Then B is also an F_i-basis of V_i, for $i = 0, 1, 2$. The intersection $V = V_1 \cap V_2$ is thus the F-span of the common basis B, since $F = F_1 \cap F_2$ by Condition 4.2.

The F-vector space V therefore induces V_1 and V_2 with respect to the above inclusions of V_1, V_2 into V_0. This shows that V, together with the corresponding isomorphisms, provides a solution to the patching problem. Equivalently, the functor Φ is essentially surjective, i.e. surjective on isomorphism classes.

To complete the proof in this direction, note that Condition 4.2 yields the short exact sequence

$$0 \to F \xrightarrow{\Delta} F_1 \times F_2 \xrightarrow{-} F_0 \to 0 \tag{4.1}$$

of F-vector spaces, where Δ is the diagonal inclusion and $-$ is the subtraction map. Tensoring over F with $\mathrm{Hom}(V, W)$ yields the exact sequence

$$0 \to \mathrm{Hom}_F(V, W) \xrightarrow{\Delta} \mathrm{Hom}_{F_1}(V_1, W_1) \times \mathrm{Hom}_{F_2}(V_2, W_2) \xrightarrow{-} \mathrm{Hom}_{F_0}(V_0, W_0) \to 0,$$

using that $F_i \otimes_F \mathrm{Hom}_F(V, W) = \mathrm{Hom}_{F_i}(V_i, W_i)$. This says that that the natural map $\mathrm{Hom}(V, W) \to \mathrm{Hom}(\Phi(V), \Phi(W))$ is a bijection, i.e. that the functor Φ is fully faithful. Thus it is an equivalence of categories, completing the proof of the reverse direction.

For the forward direction, given a quadruple of fields as above, note that each $A_0 \in \mathrm{GL}_n(F_0)$ defines a patching problem with A_0 the transition matrix between given bases of the vector spaces over F_1, F_2. A basis for a solution to this patching problem yields a factorization of A_0. So Condition 4.1 holds. Now let $F' := F_1 \cap F_2$; thus $F \subseteq F'$. The reverse direction of the theorem (proven above) implies that the base change functor $\mathrm{Vect}(F) \to \mathrm{Vect}(F')$ is an equivalence of categories; and this implies that $F = F' = F_1 \cap F_2$, giving Condition 4.2. □

5 Satisfying the patching criterion

We would like to use Theorem 4.3 to show that patching holds for quadruples of fields as in Section 2.4. That is, F is the function field of a smooth projective T-curve $\widehat{X}$ having closed fiber X, where T is a complete discrete valuation ring with fraction field K. We take subsets $U_1, U_2 \subset X$, neither containing all the closed points of X; and we write $U = U_1 \cup U_2$ and $U_0 = U_1 \cap U_2$. Consider the fraction fields $F_i := F_{U_i}$ of $\widehat{R}_i := \widehat{R}_{U_i}$. By Theorem 4.3, proving that patching holds for the fields $F_U \subset F_1, F_2 \subset F_0$ is equivalent to proving that these fields satisfy Conditions 4.1 and 4.2. (Recall that $F_U = F$ by definition, if $U = X$.) We will describe the proof of patching in explicit examples, for simplicity of exposition. For full proofs, see [HH10, Section 4].

We begin with the factorization property, and illustrate it for the example of the quadruple considered in Sections 2.3 and 2.4 above. That is, we take $T = k[[t]]$ and $K = k((t))$; we let $\widehat{X}$ be the projective line $\mathbb{P}^1_T$, with function field $F = k((t))(x)$; and we take U_1, U_2, U_0 to be the complements in $\mathbb{P}^1_k$ of the sets $\{\infty\}, \{0\}, \{0, \infty\}$, respectively. Thus $\widehat{R}_1 = k[x][[t]]$, $\widehat{R}_2 = k[x^{-1}][[t]]$, and $\widehat{R}_0 = k[x, x^{-1}][[t]]$.

We first explain why Condition 4.1 holds in the key case in which

$$A_0 \in \mathrm{GL}_n(\widehat{R}_0) \text{ and } A_0 \equiv I \mod t. \tag{5.1}$$

In this situation we will obtain that in fact $A_0 = A_1 A_2$ where $A_i \in \mathrm{GL}_n(\widehat{R}_i)$ for $i = 1, 2$. We will do this by constructing A_1, A_2 modulo t^j inductively on j, thereby finding the successive coefficients of the powers of t in the entries of the matrices A_i.

To start the inductive process, let A_1, A_2 be congruent to I modulo t. To do the inductive step, we use that every element in $k[x, x^{-1}]$ is the sum of elements in $k[x]$ and $k[x^{-1}]$. As an example, take $n = 1$, and consider the 1×1 matrix

$$A_0 = (1 + (x + 1 + x^{-1})t) \in \mathrm{GL}_1(\widehat{R}_0) = \widehat{R}_0^\times.$$

Modulo t, the factorization is just $1 \cdot 1$. For the factorization modulo t^2, we write $x + 1 + x^{-1}$ as the sum of $x + 1 \in k[x]$ and $x^{-1} \in k[x^{-1}]$, and obtain

$$A_0 \equiv (1 + (x + 1)t)(1 + x^{-1}t) \mod t^2.$$

The discrepancy between the left and right hand sides is $(1 + x^{-1})t^2$, and so at the next step we write

$$A_0 \equiv (1 + (x + 1)t - t^2))(1 + x^{-1}t - x^{-1}t^2) \mod t^3.$$

We continue in this way, and in the limit we get the desired matrices A_i. The same strategy handles the $n \times n$ case, again provided that A_0 is in the above key case.

This approach can be also used if we take two more general sets U_1, U_2 whose union is $\mathbb{P}^1_k$, neither of which contains all the closed points of $\mathbb{P}^1_k$. The inductive procedure as above still works in the key case (5.1), since again we can write every regular function on $U_0 := U_1 \cap U_2$ as a sum of regular functions on U_1 and U_2, because of the partial fractions decomposition. For example, if U_1 is the complement of the point $x = 1$ and U_2 is the complement of $x = -1$, and if $A_0 = (1 + \frac{1}{x^2-1}t)$, then we can write $\frac{1}{(x-1)(x+1)} = \frac{1}{2(x-1)} - \frac{1}{2(x+1)}$, and proceed as before in order to factor A_0. (This works even if U_1, U_2 are not necessarily Zariski open; e.g. if U_1 is a closed point of $\mathbb{P}^1_k$ and U_2 is its complement.) In this way, we see that Condition 4.1 holds in the key case (5.1). The approach also works for a more general choice of complete discrete valuation ring T.

Before turning to the general case, we observe an important consequence of Condition 4.1 under (5.1), taking $n = 1$:

Theorem 5.1 (Weierstrass Preparation Theorem)**.** *If U is an open subset of the closed fiber X, then every $f \in F_U$ can be written in the form $f = au$ with $a \in F$ and $u \in \widehat{R}_U^\times$.*

Proof. For simplicity of exposition, we explain the proof in the special case that $T = k[[t]]$, with $\widehat{X} = \mathbb{P}^1_T$ and $U = \mathbb{A}^1_k$, the affine x-line over k. Since F_U is the fraction field of $\widehat{R}_U$, we may assume that f lies in $\widehat{R}_U = k[x][[t]]$. Let $f_0 \in k[x]$ be the constant term of f, viewing f as a power series in t. Thus $f/f_0 \in k(x)((t))$, and $f/f_0 \equiv 1 \mod t$. Writing $U_1 = \{\infty\}$ and $U_2 = U$, Condition 4.1 under (5.1) with $n = 1$ asserts that $f/f_0 = f_1 f_2$ for some $f_i \in \mathrm{GL}_1(R_i) = R_i^\times$ for $i = 1, 2$. But $f_0 f_1 = f f_2^{-1}$, where the left hand side lies in $\widehat{R}_1[x] = k[x^{-1}]_{(x^{-1})}[[t]][x]$ and the right hand side lies in $\widehat{R}_2 = k[x][[t]]$. Since the left hand side has bounded degree in x, and the right hand side has no x^{-1} terms, this common element lies in $k[[t]][x] \subset F$. So we may take $a = f_0 f_1$ and $u = f_2$. □

In the case that $\widehat{X} = \mathbb{P}^1_T$, this is equivalent to the classical Weierstrass Preparation Theorem (see [Bou72], Proposition VII.3.8.6).

Corollary 5.2. *Let U_1, U_2 be open subsets of the closed fiber X, let $U_0 = U_1 \cap U_2$, and write $U = U_1 \cup U_2$ and $F_i = F_{U_i}$. Then $F_U = F_1 \cap F_2 \subseteq F_0$.*

Proof. For simplicity, we take the special case of $T = k[[t]]$; $\widehat{X} = \mathbb{P}^1_T$; U_1 is the complement of the point 0; and U_2 is the complement of the point ∞. Here $U_i = \mathrm{Spec}(\widehat{R}_i)$ with $\widehat{R}_1 = k[x^{-1}][[t]]$ and $\widehat{R}_2 = k[x][[t]]$. Write $R' = \widehat{R}_1[x] \cap \widehat{R}_2$. Thus $R' = k[[t]][x]$, with fraction field equal to F.

Say $f \in F_1 \cap F_2$. By Theorem 5.1, we may write $f = f_1 u_1 = f_2 u_2$ with $f_i \in F$ and $u_i \in \widehat{R}_i^\times$. Since F is the fraction field of R', we can write $f_i = a_i/b_i$ with $a_i, b_i \in R'$. So $f = a_1 u_1/b_1 = a_2 u_2/b_2$. The common element $a_1 b_2 u_1 = a_2 b_1 u_2$ lies in R', since the left hand side lies in $R' \widehat{R}_1^\times \subseteq \widehat{R}_1[x]$ and the right hand side lies in $R' \widehat{R}_2 = \widehat{R}_2$. Therefore $f = a_1 b_2 u_1 / b_1 b_2$ lies in F, being the ratio of two elements of R'. □

Using this, we obtain that the criterion of Section 4 is satisfied for our fields, and hence patching holds.

Theorem 5.3. *Let U_1, U_2 be subsets of the closed fiber X of $\widehat{X}$, and write $U = U_1 \cup U_2$ and $U_0 = U_1 \cap U_2$. Write $F_i = F_{U_i}$. Then the four fields $F_U \subseteq F_1, F_2 \subseteq F_0$ satisfy Conditions 4.1 and 4.2, and hence the base change functor Φ of* (2.2) *is an equivalence of categories. That is, patching holds for finite dimensional vector spaces over these fields.*

Proof. Condition 4.2 follows from Corollary 5.2, and the last part of the assertion follows from the first part together with Theorem 4.3. So it remains to show that Condition 4.1 holds.

For this, one first reduces to the case that U_1 and U_2 are disjoint. This is done by applying the disjoint case to an invertible matrix over $F_0 \subseteq F_\varnothing$, with respect to the fields F_1 and F_2', where $F_2' = F_{U_2'}$ with U_2' the complement of U_0 in U_2. One checks that the second factor is an invertible matrix over F_2, by applying Corollary 5.2 to get $F_0 \cap F_2' = F_2$.

In the disjoint case, we are given $A_0 \in \mathrm{GL}_n(F_\varnothing)$; after multiplying by a power of t, we may assume that A_0 is a matrix over the complete discrete valuation ring $\widehat{R}_\varnothing$. By the t-adic density of $R_\varnothing \subseteq F \subseteq F_U$ in $\widehat{R}_\varnothing$, there is a matrix C over $R_\varnothing$ such that $A_0 C \equiv I \mod t$. We then conclude by using that Condition 4.1 holds in the Key Case (5.1). □

As in Section 2.5, in our geometric situation we then have patching for many other algebraic objects, which consist of finite dimensional vector spaces with additional structure. And as in Section 3, various applications then follow.

6 Variants on the patching set-up

6.1 Using more than two open sets

Rather than covering a subset U of the closed fiber X with just two proper subsets, it is possible to use a larger number of subsets. This can be useful in constructions, e.g. in the situations described in Section 3. Namely, suppose that $U \subseteq X$ is the union of subsets U_i, for $i = 1, \ldots, n$. For simplicity we assume that all double intersections $U_i \cap U_j$ are equal to a common set U_0, for $i \neq j$. Suppose we are given finite dimensional vector spaces V_i over each $F_i := F_{U_i}$ together with isomorphisms $\nu_i : V_i \otimes_{F_i} F_0 \to V_0$ for $i = 1, \ldots, n$. Then the arguments in Section 5 can be generalized to show that there is a unique choice of a finite dimensional F_U-vector space V together with isomorphisms $\alpha_i : V \otimes_{F_U} F_i \to V_i$ for $i = 0, \ldots, n$ such that $\nu_i \circ (\alpha_i \otimes F_0) = \alpha_0$ for $i = 1, \ldots, n$. Moreover this defines an equivalence between the category $\mathrm{Vect}(F_U)$ and the category of patching problems that consist of data $(V_i, \nu_i)_i$ as above. This equivalence of categories can be proven from the equivalence with two open sets via induction. See [HH10, Theorem 4.14] for details.

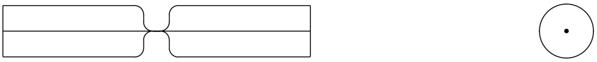
Since the above functor is an equivalence of categories and preserves tensor products, it carries over from finite dimensional vector spaces to other categories of objects that consist of such vector spaces together with additional structure, as in the situation of Section 2.5. As a result, applications can be obtained in a single step, rather than building up object inductively as in Section 3.1. In that situation, a finite group G can be generated by cyclic subgroups G_i for $i = 1, \dots, n$. By covering X by n open subsets U_i as above, and building a G_i-Galois extension of F_i for each i (e.g. by Kummer theory), one can then obtain a G-Galois extension of F. The agreement over F_0 can be achieved by choosing the cyclic extensions of the field F_i each to induce trivial extensions over the field F_0.

6.2 Using a point and a complement

Additional flexibility in constructions can be obtained by allowing patches that come from points, rather than from subsets of the closed fiber X.

For any point $P \in X$, let R_P be the local ring of $\widehat{X}$ at P, and let $\widehat{R}_P$ be the completion of this local ring at its maximal ideal $\mathfrak{m}_P$. This is a domain, and its fraction field will be denoted by F_P. (Note that although $R_P = R_{\{P\}}$, the completions $\widehat{R}_P$ and $\widehat{R}_{\{P\}}$ are different, the former being the $\mathfrak{m}_P$-adic completion and the latter being the t-adic completion. Thus F_P and $F_{\{P\}}$ are also different, with the former containing the latter.)

For example, if $T = k[[t]]$ and $\widehat{X} = \mathbb{P}^1_T$, we may consider the point P where $x = t = 0$. Then $\widehat{R}_P = k[[x, t]]$ and F_P is the fraction field $k((x, t))$ of $k[[x, t]]$. Taking U to be the complement of P in the closed fiber $X = \mathbb{P}^1_k$, so that $\widehat{R}_U = k[x^{-1}][[t]]$, we obtain the following picture:



Spec($\widehat{R}_U$) Spec($\widehat{R}_P$)

Patching can then be carried out using the fields F_P and F_U in place of the two fields F_{U_1} and F_{U_2} of Section 2. To do this, we also need to take an appropriate overfield of F_U and F_P, which will take the place of the field F_{U_0} in Section 2. In the above example, we will take the field $F_0 := k((x))((t))$, which is the fraction field of the domain $\widehat{R}_0 := k((x))[[t]]$. Note that $\widehat{R}_0$ is t-adically complete, and contains the rings $\widehat{R}_U$ and $\widehat{R}_P$; and similarly F_0 contains F_U and F_P. This choice of $\widehat{R}_0$ makes sense intuitively, because Spec(U) is the complement of the point $(x = 0)$ in X, while Spec$\big(k[[x]]\big)$ can be viewed as a small neighborhood of this point in X, so that $\mathrm{Spec}\big(k((x))\big)$ can be viewed as the corresponding punctured neighborhood (since $k((x)) = k[[x]][x^{-1}]$).

More generally, if T is a complete discrete valuation ring and $\widehat{X}$ is a smooth projective T-curve, then we may pick a closed point P on the closed fiber X of $\widehat{X}$, and consider the ring $\widehat{R}_P$ and its fraction field F_P. The complement U of P in X is a smooth affine curve, and we may also consider the ring $\widehat{R}_U$ and its fraction field F_U. Let $\pi \in R_P$ be an element whose reduction is a uniformizer of the complete local ring of X at P. We then take $\widehat{R}_0$ to be the t-adic completion of $R_P[\pi^{-1}]$, and F_0 to be its fraction field.

In this situation, the strategy of Section 4 can be carried over, to show that Conditions 4.1 and 4.2 hold. As a result, the corresponding base change functor

$$\Phi : \mathrm{Vect}(F) \to \mathrm{Vect}(F_U) \times_{\mathrm{Vect}(F_0)} \mathrm{Vect}(F_P),$$

is an equivalence of categories. (For details, see [HH10, Theorem 5.9]. There is also a generalization that allows more than one point P; see [HH10, Theorem 5.10].)

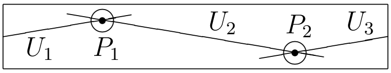
This approach is useful for certain applications, such as split embedding problems in Galois theory. In such a problem, one is given a surjection of finite groups $f : G \to H$ together with a section $s : H \to G$ of f, and also an H-Galois field extension E/F. The problem is then to embed E into a G-Galois field extension E'/F, such that the Galois correspondence associates the inclusion $E \hookrightarrow E'$ to the surjection f. (See the discussion in Section 3.1, where no splitting condition was assumed.) The difficulty with using just the fields F_U is that the extension E/F does not in general become trivial over any F_U. But E/F will become trivial over F_P if P is unramified and split in E/F. This approach has been used to obtain a variety of results about embedding problems in Galois theory, and those in turn have been used to obtain information about the structure of absolute Galois groups (e.g. as discussed in Sections 1.4–1.6 above).

6.3 Patching over singular curves

Until now, we have considered smooth projective T-curves $\widehat{X}$. But given a function field F of transcendence degree one over the fraction field K of a complete discrete valuation ring T, there need not be a smooth model $\widehat{X}$ of F over T. That is, there need not be a smooth projective T-curve $\widehat{X}$ whose function field is the K-algebra F. For greater generality, we will now permit projective T-curves $\widehat{X}$ that are merely assumed to be *normal* as schemes, i.e. the local rings are integrally closed domains. Given any function field F of transcendence degree one over K, one can easily obtain a normal projective model $\widehat{X}$, e.g. by writing F as a finite extension of $K(x)$, and then taking the normalization of $\mathbb{P}^1_T$ in F. (In fact, by [Abh69] and [Lip75], for each such F there even exist regular projective models $\widehat{X}$ over T, i.e. ones for which every local ring $\mathcal{O}_{\widehat{X},P}$ is a regular local ring. These can be obtained by starting with any projective model and then applying a suitable combination of normalization and blowing up. Such regular models, however, still need not be smooth, the latter condition being equivalent to the closed fiber being smooth over the residue field k of T.)

Typically, the closed fiber X of a normal projective model $\widehat{X}$ of F over T will have several irreducible components, which will meet at several closed points. Let $\mathcal{P}$ be a non-empty finite set of closed points of X that contains all of these intersection points, and let $\mathcal{U}$ be the set of irreducible components of the complement of $\mathcal{P}$ in X. For each $P \in \mathcal{P}$ we can consider the complete local ring $\widehat{R}_P$ of $\widehat{X}$ at P, and its fraction field F_P. For each $U \in \mathcal{U}$ we can consider the t-adic completion $\widehat{R}_U$ of the ring R_U of rational functions that are regular on U; and take F_U to be the fraction field of this domain. We then have a set-up that combines the situations of the previous two subsections, using fields both of the types F_U and F_P, and typically using more than two fields in total.

For example, the following picture illustrates a choice of $\widehat{X}$ in which the closed fiber X has irreducible components X_1, X_2, X_3, where X_2 meets each of the other two components at a single point (P_1, P_2 respectively). The open subsets U_1, U_2, U_3 of X are the connected components of $X \smallsetminus \{P_1, P_2\}$, with U_i being a Zariski open dense subset of X_i. Thus U_1 (resp. U_2, resp. U_3) is the complement of P_1 (resp. P_1 and P_2, resp. P_2) in X_1 (resp. X_2, resp. X_3). The picture also illustrates the spectrum of $\widehat{R}_{P_i}$ for $i = 1, 2$, as a small neighborhood of P_i. (The spectrum of $\widehat{R}_{U_i}$, for $i = 1, 2, 3$, could similarly be illustrated as a neighborhood of U_i in $\widehat{X}$; but for visual clarity these spectra are not shown.)



As in Section 6.2, we need to define overfields for the fields F_P ($P \in \mathcal{P}$) and F_U ($U \in \mathcal{U}$). But unlike Section 6.1, there is no common overfield for all these fields. Instead, we define an overfield for F_U and F_P only if P and U are incident; i.e. if P is a point in the closure of U.

Thus in the above example, there will be four overfields, arising the pairs (U_1, P_1), (U_2, P_1), (U_2, P_2), (U_3, P_2). The overfield arising from a pair (U_i, P_j) will contain the fields F_{U_i} and F_{P_j}.

Moreover, this will be done in a way that if the closed fiber is smooth (and therefore irreducible) and there is just one point P chosen, the resulting overfield will be the same as the overfield F_0 that was considered in Section 6.1. Finally, and most crucially, the fields F_U, F_P, and these overfields will satisfy a generalization of Conditions 4.1 and 4.2; and as a result, the associated base change functor Φ (as in (2.2)) will be an equivalence of categories.

To define these overfields, we use the notion of *branches*. To illustrate, consider the affine curve C in the x, y-plane given by $y^2 = x^3 + x^2$. This is shaped like the letter α, with the node at the origin O. The curve C is irreducible, and so the local ring $\mathcal{O}_{C,O}$ is a domain (with fraction field equal to the function field of C). But the completion $\widehat{\mathcal{O}}_{C,O}$ of the local ring is not a domain. Explicitly, the completion is isomorphic to the ring $k[[x,z]]/(z^2-x^2)$, via the isomorphism taking y to zf, where $f=(1+x)^{1/2}\in k[[x]]$. Geometrically, the spectrum of $\widehat{\mathcal{O}}_{C,O}$ has two irreducible components, corresponding to the two "branches" of C at O. They are respectively defined by the two minimal primes $\wp_1=(z-x)$, $\wp_2=(z+x)$ of $\widehat{\mathcal{O}}_{C,O}$. This observation can be used to motivate the formal definition: a *branch* of a variety V at a point P is a minimal prime of $\widehat{\mathcal{O}}_{V,P}$.

Returning to our situation, for each $P\in\mathcal{P}$ and for each branch of X at P, we wish to associate an overfield of F_P; this should also be an overfield of F_U, where $U\in\mathcal{U}$ is the unique element on whose closure $\bar{U}$ the branch lies (i.e. such that it is a branch of $\bar{U}$ at P). To illustrate, in the example with the reduced closed fiber pictured above, there will be two branches of X at P_1 (along the closures of U_1 and U_2) and two at P_2 (along the closures of U_2 and U_3); these four branches will correspond to the four desired overfields.

For $P\in\mathcal{P}$, the inclusion $X\to\widehat{X}$ induces a surjection $\widehat{R}_P=\widehat{\mathcal{O}}_{\widehat{X},P}\to\widehat{\mathcal{O}}_{X,P}$ whose kernel is the radical of (t). By taking inverse images under this surjection, the minimal primes of $\widehat{\mathcal{O}}_{X,P}$ can then be identified with the height one primes of $\widehat{R}_P$ that contain t. Thus we may regard each of these height one primes $\wp$ as a *branch* of X at P. Viewed as a point of $\operatorname{Spec}(\widehat{R}_P)$, the image of $\wp$ in $\widehat{X}$ lies on X, and in fact is the generic point of an irreducible component of X. We regard this as the component on which the branch lies; and the unique $U\in\mathcal{U}$ that is contained in this component is an open set that is incident to P. In this situation, we can now define an overfield $F_\wp$ of F_U and F_P, associated to this branch.

Namely, for $\wp$ a height one prime of $\widehat{R}_P$ that contains t, let $R_\wp$ be the local ring of $\widehat{R}_P$ at $\wp$. This is a discrete valuation ring, with t lying in its maximal ideal. Let $\widehat{R}_\wp$ be its completion; this is a complete discrete valuation ring, whose fraction field will be denoted by $F_\wp$. This is the desired overfield of F_U and F_P, where $\wp$ is a branch on the closure of U.

We then have a finite inverse system of fields $F_U, F_P, F_\wp$, indexed by the disjoint union $\mathcal{U}\sqcup\mathcal{P}\sqcup\mathcal{B}$, where $\mathcal{B}$ is the set of branches of X at the points of $\mathcal{P}$. Here we have inclusions of F_U and F_P into $F_\wp$ if $\wp$ is a branch of X at the point P lying on the closure of U. In the special case of Section 6.2, there is just one branch $\wp$ to consider, viz. the unique branch of X at P; this lies on U. The associated field $\wp$ is then the same as the field F_0 considered in that section. As another example, suppose that the closed fiber X is irreducible but singular, and is isomorphic to the nodal curve C discussed above. Taking $\mathcal{P}$ to consist just of the nodal point P, the set $\mathcal{U}$ will then consist just of the complement $U:=X\smallsetminus P$. But there will be *two* branches $\wp_1,\wp_2$ of X at P, both lying on the closure of U. Thus in this case we have two overfields $F_{\wp_1}, F_{\wp_2}$ in the inverse system, and we will need to consider both.

In general in the above situation, it can be shown that the inverse limit of the fields $F_U, F_P, F_\wp$ is just the field F, which is a subfield of each of them. In the special case of Section 6.2, this is another way of asserting that F is the intersection of F_U and F_P inside $F_0 = F_\wp$. Thus this *inverse limit property* is a way of generalizing the intersection Condition 4.2 to the case of singular curves, where there may be multiple patching fields F_U, F_P and multiple overfields $F_\wp$. (See [HH10, Proposition 6.3] for details. There this was shown by reducing to the case of $\mathbb{P}^1_T$ with $\mathcal{P} = \{\infty\}$, by choosing a finite morphism $\widehat{X} \to \mathbb{P}^1_T$.)

The factorization Condition 4.1 can also be generalized to the current situation. The condition then becomes a *simultaneous factorization property*. That is, given elements $A_\wp \in \mathrm{GL}_n(F_\wp)$ for each $\wp \in \mathcal{B}$, there exist elements $A_U \in \mathrm{GL}_n(F_U)$ for each $U \in \mathcal{U}$ and elements $A_P \in \mathrm{GL}_n(F_P)$ for each $P \in \mathcal{P}$, satisfying the following condition: For each triple $U, P, \wp$ with $\wp$ a branch at P lying on the closure of U, the identity $A_\wp = A_U A_P$ holds in $\mathrm{GL}_n(F_\wp)$. Here we regard F_U, F_P as subfields of $F_\wp$. In the above situation, this property always holds. (Again, this can be shown by reducing to the case of $\mathbb{P}^1_T$ with $\mathcal{P} = \{\infty\}$, which was discussed in Section 6.2. A stronger result was proven this way in [HHK09, Theorem 3.6].)

Consider, for example, the model $\widehat{X}$ displayed in the picture above. There are four branches $\wp_1, \wp_2, \wp_3, \wp_4$, which are respectively associated to the four pairs (U_1, P_1), (U_2, P_1), (U_2, P_2), (U_3, P_2). Suppose we are given elements $A_{\wp_i} \in \mathrm{GL}_n(F_{\wp_i})$ for $i = 1, 2, 3, 4$. The simultaneous factorization property asserts that there exist elements $A_{U_j} \in \mathrm{GL}_n(F_{U_j})$ for $j = 1, 2, 3$, and elements $A_{P_\ell} \in \mathrm{GL}_n(F_{P_\ell})$ for $\ell = 1, 2$, such that $A_{\wp_1} = A_{U_1} A_{P_1}$, $A_{\wp_2} = A_{U_2} A_{P_1}$, $A_{\wp_3} = A_{U_2} A_{P_2}$, and $A_{\wp_4} = A_{U_3} A_{P_2}$, in the respective groups $\mathrm{GL}_n(F_{\wp_i})$.

The analog of Theorem 4.3 also holds in the current situation; i.e. the base change functor is an equivalence of categories if and only if the simultaneous factorization and inverse limit properties hold. (This can be shown by replacing the fields F_1, F_2, F_0 in Theorem 4.3 by the F-algebras $\prod F_P, \prod F_U, \prod F_\wp$, and proceeding as before.) As a consequence, patching holds for finite dimensional vector spaces in this situation. More precisely, we have the following (see also [HH10, Theorem 6.4]):

Theorem 6.1. *Given a projective normal curve $\widehat{X}$ over T, with $\mathcal{P}, \mathcal{U}, \mathcal{B}$ as above, the associated base change functor*

$$\Phi : \mathrm{Vect}(F) \to \varprojlim \mathrm{Vect}(F_\xi)$$

is an equivalence of categories, where ξ ranges over $\mathcal{P} \sqcup \mathcal{U} \sqcup \mathcal{B}$.

Thus if we are given finite dimensional vector spaces over the fields F_U and F_P, together with isomorphisms between the vector spaces that they induce over the overlap fields $F_\wp$, then there is a unique vector space over F that induces all of them compatibly; and this is an equivalence of categories. This assertion automatically carries over to other algebraic objects that consist of finite dimensional vector spaces with additional structure, as in Section 2.5.

7 Patching torsors

Until now, we have been patching structures that consist of finite dimensional vector spaces with additional structure (i.e. such that some diagrams commute). For example, G-Galois F-algebras are finite dimensional because the group G is finite. But what if we allow *infinite* dimensional vector spaces? For example, can we generalize G-Galois F-algebras to infinite groups G, such as matrix groups? To consider this, we introduce the notion of torsors.

7.1 Introduction to torsors

Say G is a linear algebraic group over a field F, i.e. a smooth Zariski closed subgroup of GL_n for some n. Here G need not be connected, and in particular all ordinary finite groups G are linear algebraic groups (as groups of permutation matrices). A *G-torsor* over F is a principal homogeneous G-space over F. That is, it is an F-variety H together with a (right) G-action $\alpha : H \times G \to H$ that is simply transitive. To be more precise, this simple transitivity property asserts that the morphism $(\mathrm{pr}_1, \alpha) : H \times G \to H \times H$ is an isomorphism of F-schemes, where pr_1 is the first projection map. So intuitively, given two points of H, there is a unique element of G taking one to the other.

A G-torsor H is *trivial* if it is F-isomorphic to G, with the G-action being given by right multiplication, with respect to this isomorphism. Note that a G-torsor H is trivial if and only if it has an F-point. Namely, in the forward direction, the identity element of G corresponds to an F-point of H; and in the reverse direction, by sending an F-point of H to the identity of G we obtain a unique isomorphism $H \to G$ that is compatible with the right G-actions. Thus if F is algebraically closed, then every G-torsor is trivial, since it automatically has an F-point.

Torsors are important in part because they classify various algebraic objects. For example, torsors for the orthogonal group $\mathrm{O}(n)$ classify quadratic forms in n variables over F, provided that $\mathrm{char}(F) \neq 2$ (see [KMRT98, (29.28)]).

In the case that G is a finite group in the classical sense, if we regard G as a torsor over F, then G-torsors are of the form $\mathrm{Spec}(A)$, where A is a G-Galois F-algebra. (Note that G acts on A on the left, since it acts on the torsor on the right.) So in this case, the theory of G-torsors is just the theory of G-Galois F-algebras. At one extreme, if A is a G-Galois *field* extension E/F, then $\mathrm{Spec}(A)$ consists of just one point, viz. $\mathrm{Spec}(E)$. At the other extreme, if A is a product of copies of F indexed by the elements of G, then $\mathrm{Spec}(A)$ consists of $|G|$ copies of $\mathrm{Spec}(F)$, again indexed by G. This is a *trivial* G-torsor.

7.2 Torsors and cohomology

It is convenient to study torsors in terms of Galois cohomology. Say G is a linear algebraic group over a field F. Consider maps $\chi : \mathrm{Gal}(F) \to G(F^{\mathrm{sep}})$ such that

$$\chi(\sigma\tau) = \chi(\sigma)\sigma(\chi(\tau)) \text{ for all } \sigma, \tau \in \mathrm{Gal}(F).$$

Such maps χ are called 1-*cocycles* for $\mathrm{Gal}(F)$ in $G(F^{\mathrm{sep}})$. Two 1-cocycles χ, χ' are called *cohomologous* if there is a $g \in G(F^{\mathrm{sep}})$ such that

$$\chi(\sigma) = g^{-1}\chi'(\sigma)\sigma(g) \text{ for all } \sigma \in \mathrm{Gal}(F).$$

The set of cohomology classes of 1-cocycles is denoted by $H^1(F, G)$. This is the *first Galois cohomology set* of F with coefficients in G. It is a group if G is commutative; otherwise it is just a pointed set (whose distinguished element corresponds to the trivial cocycle χ).

The key point is that there is a natural bijection

$$(7.1) \qquad \{\text{isomorphism classes of } G\text{−torsors over } F\} \;\leftrightarrow\; H^1(F, G).$$

Namely, given a G-torsor H, pick a point $P \in H(F^{\mathrm{sep}})$. For each $\sigma \in \mathrm{Gal}(F)$, we have the two points $P, \sigma(P) \in H(F^{\mathrm{sep}})$. By the torsor property, there is a unique element $g \in G(F^{\mathrm{sep}})$ such that $\sigma(P) = Pg$. Write $\chi(\sigma) = g$. Then the map $\chi : \mathrm{Gal}(F) \to G(F^{\mathrm{sep}})$ is a 1-cocycle. It depends on the choice of P; but it is straightforward to check that changing P does not change the equivalence class of χ. Thus we have a well-defined element $[\chi]$ of $H^1(F, G)$ that is associated to the G-torsor H. One can then check that the association $H \mapsto [\chi]$ defines a bijection as in (7.1). (Note also that this bijection parallels the fact in topology that the principal G-bundles over a space X are classified by $H^1(X, G)$.)

One can also define $H^n(F, G)$ for other values of n, though for $n > 1$ one requires G to be commutative. In particular, $H^0(F, G)$ is just $G(F)$. Moreover $H^0(F, E/J)$ can also be defined for any subgroup $J \subseteq E$ of a group E; this is the set of $\mathrm{Gal}(F)$-invariant cosets $eJ_{F^{\mathrm{sep}}}$ of $J_{F^{\mathrm{sep}}}$, with $e \in E(F^{\mathrm{sep}})$.

See [Ser00, I, Section 5] for more about non-abelian Galois cohomology.

7.3 Torsors and patching

We now return to the question of patching. As in Section 3.1, it is possible in the context of our fields to patch G-Galois F-algebras, or equivalently G-torsors, where G is a finite group. More generally, let G be a linear algebraic group over F. If G is infinite, then any G-torsor is of the form $H = \mathrm{Spec}(A)$, where A is an F-algebra that is *infinite* dimensional over F. It turns out that it is still possible to patch G-torsors!

To see this, we first relate torsors to matrices. View $G \subseteq \mathrm{GL}_n$ over F. For $h \in \mathrm{GL}_n(F^{\mathrm{sep}})$, consider the translate (or equivalently, coset) $hG \subseteq \mathrm{GL}_n$. If

$h \in \mathrm{GL}_n(F)$, then via left multiplication by h (which is an F-isomorphism), hG is isomorphic to the trivial G-torsor $hG = G$. If $h \notin \mathrm{GL}_n(F)$, then the translate hG need not be defined over F; but if it is, then it defines a G-torsor that will in general be non-trivial. To say that hG is defined over F is equivalent to the condition that for every $\sigma \in \mathrm{Gal}(F)$, $hG = (hG)^\sigma = h^\sigma G$; i.e. that $h^\sigma \in hG$.

As an example, let $F = \mathbb{R}$, let G be the orthogonal group $\mathrm{O}(2)$, and let h be the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right)$. Then hG defines a non-trivial G-torsor over $\mathbb{R}$.

The above construction in fact gives rise to *all* torsors; i.e. every G-torsor over F arises from a translate hG as above. This follows from the exact sequence

$$1 \to H^0(F, G) \to H^0(F, \mathrm{GL}_n) \to H^0(F, \mathrm{GL}_n/G) \to H^1(F, G) \to H^1(F, \mathrm{GL}_n) \tag{7.2}$$

(see [Ser00, I.5.4, Proposition 36]) and the fact that $H^1(F, \mathrm{GL}_n)$ is trivial by Hilbert's Theorem 90 ([KMRT98, Theorem 9.2]). In fact this shows that there is a bijection of pointed sets $\mathrm{GL}_n(F)\backslash H^0(F, \mathrm{GL}_n/G) \to H^1(F, G)$, which classifies G-torsors over F.

So we can study torsors by studying matrices. Using this approach, one can show in an abstract context:

Theorem 7.1. *Given a finite inverse system of fields, if patching holds for finite dimensional vector spaces, then patching also holds for G-torsors, for all linear algebraic groups G.*

That is, if F is the inverse limit of a finite inverse system of fields $(F_\xi)_{\xi \in I}$, and if the base change functor

$$\Phi : \mathrm{Vect}(F) \to \varprojlim \mathrm{Vect}(F_\xi)$$

on vector spaces is an equivalence of categories, then so is the base change functor

$$\Phi_G : G\mathrm{Tors}(F) \to \varprojlim G\mathrm{Tors}(F_\xi)$$

for any linear algebraic group G over F, where GTors denotes the category of G-torsors over the given field.

We explain the proof in the basic context of four fields as in (2.1) and (2.2). Suppose that these fields $F \subseteq F_1, F_2 \subseteq F_0$ satisfy patching for finite dimensional vector spaces. By Theorem 4.3, the factorization Condition 4.1 and the intersection Condition 4.2 hold for these fields. Suppose we are given G-torsors H_i over F_i for $i = 1, 2$, together with an isomorphism over F_0. By the above discussion, we may write $H_i = h_i G$, with $h_i \in \mathrm{GL}_n(F_i^{\mathrm{sep}})$; and we are given an F_0-isomorphism $h_2 G \to h_1 G$. This map is defined by left multiplication by some $g_0 \in \mathrm{GL}_n(F_0)$. Thus $h_1 G = g_0 h_2 G$. Applying Condition 4.1 to g_0, we obtain elements $g_i \in \mathrm{GL}_n(F_i)$ for $i = 1, 2$ such that $g_1^{-1} g_2 = g_0$. Let $h_i' = g_i h_i \in \mathrm{GL}_n(F_i^{\mathrm{sep}})$ for $i = 1, 2$. Thus $h_1' G = h_2' G$ over F_0. That is, the translates $h_i' G$, for $i = 1, 2$, define the *same* F_0-point on the quotient GL_n/G, which is an F-variety (e.g. see [Bor91, Theorem II.6.8]). This point of GL_n/G is then defined over both F_1 and F_2, and hence over F by Condition 4.2. The point thus defines an element of $H^0(F, \mathrm{GL}_n/G)$, and hence a G-torsor defined over F, by the above exact sequence (7.2).

Essentially the same argument holds for a more complicated finite inverse system of fields. We thus obtain that patching for finite dimensional vector spaces implies patching for torsors.

In the context of one-variable function fields F over a complete discretely valued field K, with a normal projective model $\widehat{X}$ and sets $\mathcal{P}, \mathcal{U}, \mathcal{B}$ as before, patching for finite dimensional vector spaces holds by Theorem 6.1. Hence Theorem 7.1 yields:

Corollary 7.2. *In the situation of Section 6.3, patching for G-torsors holds with respect to the index set $\mathcal{P} \sqcup \mathcal{U} \sqcup \mathcal{B}$, for any linear algebraic group G over F.*

8 Local-global principles

Patching for torsors can be used to study local-global principles for algebraic objects. For example, using torsors under the orthogonal group $\mathrm{O}(n)$, local-global principles can be obtained for quadratic forms.

The most classical local-global principle, in fact, concerns quadratic forms, though over a different type of field. The Hasse-Minkowski theorem states that a quadratic form over $\mathbb{Q}$ is isotropic (i.e. has a non-trivial zero) if and only if it is isotropic over each field $\mathbb{Q}_p$ and also over $\mathbb{R}$. More generally, the analogous assertion holds for any global field K, with respect to its completions K_v with respect to the absolute values v on K. In the case of equal characteristic global fields (i.e. function fields of curves over a finite field), this is equivalent to taking the completions at the closed points of the associated smooth projective curve.

A related classical local-global principle is the theorem of Albert, Brauer, Hasse, and Noether. That theorem concerns central simple algebras over global fields. It says that such an algebra is split (i.e. is isomorphic to a matrix algebra over the given field) if and only it is split over each completion.

Local-global principles can typically be rephrased in terms of the existence of rational points on varieties. For example, in the context of the Hasse-Minkowski theorem, let q be a quadratic form over a global field K. If q is a form in n variables, then it defines a quadric hypersurface Q in $\mathbb{P}^{n-1}_K$; and q is isotropic over K if and only if Q has a K-point. From this point of view, local-global principles assert that a variety has a point over K if and only if it has a point over each completion of K.

To go beyond the context of global fields, we broaden the notion of a local-global principle: it is an assertion that a given property (such as the presence of points on varieties) holds over a given field F if and only if it holds over each of a given set of overfields F_ξ of F, where ξ ranges over some index set I.

In the situation of torsors, we can consider the following local-global principle: A G-torsor H over F is trivial (or equivalently, has an F-point) if and only if the induced torsor $H_\xi := H \times_F F_\xi$ is trivial over F_ξ for each $\xi \in I$. This assertion can rephrased in terms of Galois cohomology, since the G-torsors over F are classified by $H^1(F, G)$. Namely, there is a natural *local-global map* $H^1(F, G) \to \prod_\xi H^1(F_\xi, G)$ of pointed sets (or of groups, if G is commutative). The local-global principle then states that the kernel of this map is trivial; i.e. that only the trivial element of $H^1(F, G)$ maps to the trivial element in the product. This need not always hold, however; and a related problem is then to determine the kernel of this map, and in particular to determine whether it is finite (even if not necessarily trivial).

An important classical example considers a global field K and its completions K_v, along with an abelian variety A over K (e.g. an elliptic curve over K). In this case the local-global map is a group homomorphism, and its kernel is a group, called the *Tate-Shafarevich group* $Ш(K, A)$. A major open question is whether its order is necessarily finite. In fact, its order has been conjectured in the case of elliptic curves, in terms of special values of L-functions; this is the conjecture of Birch and Swinnerton-Dyer.

In the case that G is a linear algebraic group over a number field K, it was shown by Borel and Serre ([BS64]) that $Ш(K, G)$ is finite. (Here we again take the local-global map with respect to the set of absolute values on K.) For the analogous problem in the function field case of global fields, the finiteness of $Ш(K, G)$ was shown by Brian Conrad ([Con12]).

9 Local-global principles in the patching context

In our context, we will take F to be the function field of a projective normal curve $\widehat{X}$ over a complete discrete valuation ring T, and we will let F_ξ range over a finite set of overfields corresponding to patches F_U and F_P as in Section 6.3. Note that this involves using just a *finite* collection of overfields, unlike the classical situation in which infinitely many overfields are considered. The local-global principle for torsors then says that a G-torsor H over F is trivial if and only if H_ξ is trivial over F_ξ for each $\xi \in \mathcal{P} \sqcup \mathcal{U}$. Equivalently, it says that the kernel of the local-global map

$$H^1(F, G) \to \prod_{P \in \mathcal{P}} H^1(F_P, G) \times \prod_{U \in \mathcal{U}} H^1(F_U, G)$$

is trivial. The kernel of this map will be denoted by $Ш_{\widehat{X},\mathcal{P}}(F, G)$, where $\mathcal{P}$ determines $\mathcal{U}$. If the model $\widehat{X}$ of F is understood, we will simply write $Ш_{\mathcal{P}}(F, G)$.

For example, suppose that $F = K(x)$, where as before K is the fraction field of T, and take $\widehat{X} = \mathbb{P}^1_T$. Let $\mathcal{P}$ consist just of the point $P = \infty$ on the closed fiber $X = \mathbb{P}^1_k$, so that $\mathcal{U}$ consists just of the single open set $U = \mathbb{A}^1_k$. In this case $Ш_{\mathcal{P}}(F, G)$ is the kernel of $H^1(F, G) \to H^1(F_P, G) \times H^1(F_U, G)$. Let $\wp$ be the unique branch of X at P. Using the fact that patching holds for finite dimensional vector spaces in this context, one then obtains the following Mayer-Vietoris type exact sequence (see [HHK11a, Theorem 3.5]):

Theorem 9.1. *There is an exact sequence of pointed sets*

$$\begin{tikzcd}
1 \arrow[r] & H^0(F,G) \arrow[r] & H^0(F_P,G) \times H^0(F_U,G) \arrow[r] & H^0(F_\wp,G) \arrow[dll, "\delta"] \\
 & H^1(F,G) \arrow[r] & H^1(F_P,G) \times H^1(F_U,G) \arrow[r, shift left] \arrow[r, shift right] & H^1(F_\wp,G).
\end{tikzcd}$$

Proof. Define the maps on $H^i(F, G)$ to be the diagonal inclusions. Define the last arrow on the first line as the quotient $i_U^{-1} i_P$ of the inclusion maps from the factors in the middle term to the group $H^0(F_\wp, G)$. (This quotient map is not in general a group homomorphism.) The two arrows at the end of second line are the maps arising from the inclusions of F_P, F_U into $F_\wp$. Exactness at the middle term of that line means that the equalizer of these two maps is equal to the image of $H^1(F, G)$. (If G is commutative then $H^1(F_\wp, G)$ is a group and we could take the corresponding quotient instead of using a double arrow.)

The coboundary map δ is defined as follows: Take trivial G-torsors H_P, H_U over F_P, F_U with rational points x_P, x_U. Given $g_\wp \in H^0(F_\wp, G)$, we have an isomorphism

$$H_P \times_{F_P} F_\wp \to H_U \times_{F_U} F_\wp$$

of trivial G-torsors over $F_\wp$, taking x_P to $x_U g_\wp$. This defines a patching problem for G-torsors. By Corollary 7.2, there is a solution H to this patching problem, viz. a G-torsor H over F corresponding to an element of $H^1(F, G)$. This element is then defined to be $\delta(g_\wp)$.

With the above maps, one can then check that the sequence is exact. □

In the example of the line, exactness implies that $Ш_{\mathcal{P}}(F, G)$ is the cokernel of the last map of H^0 terms, i.e. $G(F_P) \times G(F_U) \to G(F_\wp)$, given by $(g_P, g_U) \mapsto g_U^{-1} g_P$, where we regard $G(F_P), G(F_U)$ as contained in $G(F_\wp)$. But the surjectivity of this map is equivalent to the factorization property for the group G. (Note that until now we have considered factorization only for the groups GL_n, not for other linear algebraic groups G.) Hence the local-global principle for G-torsors is equivalent to the factorization property for the group G.

While the above is for the example of the line and with $\mathcal{P} = \{\infty\}$, Theorem 9.1 carries over to arbitrary normal T-curves $\widehat{X}$ together with $\mathcal{P} \subset X$. (See [HHK11a, Theorem 3.5].) There the middle and last terms on each line of the Mayer-Vietoris exact sequence are replaced by products, where $P, U, \wp$ respectively range over $\mathcal{P}$, $\mathcal{U}$, $\mathcal{B}$. Just as the simultaneous factorization property for GL_n was considered in Section 6.3, one can also consider this property for an arbitrary linear algebraic group G. We then obtain:

Corollary 9.2. *In the situation of Section 6.3, if G is a linear algebraic group over F, then the local-global principle for G-torsors is equivalent to the simultaneous factorization property for the group G.*

10 Obstructions to local-global principles

In the previous section, local-global principles for torsors in the context of patches were reformulated in terms of factorization. Using this, we can now examine when such principles hold, and whether the obstruction $Ш_{\mathcal{P}}(F, G)$ is finite even if not necessarily trivial.

10.1 Case of rational connected groups

As we discuss below, simultaneous factorization holds for G, and hence $Ш_{\mathcal{P}}(F, G)$ vanishes, provided that the group G is connected and rational (i.e. rational as an F-variety, meaning that it is birationally isomorphic to $\mathbb{A}^n_F$ for some n). For example, the special orthogonal group $\mathrm{SO}(n)$ is a connected rational group, by the Cayley parametrization. In fact many connected linear algebraic groups are known to be rational; and over an *algebraically closed field* it is known that every linear algebraic group is rational.

We describe the idea of the proof, and to simplify the discussion we again return to the example of the projective line and $\mathcal{P} = \{\infty\}$. As in the case of proving factorization for the group GL_n, we will construct the factors inductively, modulo successive powers of the uniformizer t of T. But unlike in the GL_n case, we have to be careful to ensure that the limit of these inductive sequences of mod t^i factors will itself lie in the group G. (This was automatic for GL_n, since any matrix that is congruent to the identity modulo t will lie in GL_n.) To do this, we will use the rationality of the group G.

Specifically, since G is a connected rational variety, say of dimension m, there are Zariski open neighborhoods $N \subseteq G$ and $N' \subseteq \mathbb{A}^m$ of $1 \in G$ and $0 \in \mathbb{A}^m$, respectively, together with an isomorphism of F-varieties $\eta : N \to N'$. The multiplication map $\mu : G \times G \to G$ on G, restricted to an open neighborhood of $(1, 1)$ in $N \times N$, corresponds under η to a map $f : \tilde{N}' \to N'$, where $\tilde{N}'$ is an open dense subset of $N' \times N'$ that contains the origin. By an inductive construction that generalizes the factorization construction for GL_n, we can show that for some open neighborhood M' of the origin in N', every $B_0 \in M'(F_\wp)$ is of the form $f(B_U, B_P)$, for some $B_U \in N'(F_U)$ and $B_P \in N'(F_P)$ with $(B_U, B_P) \in \tilde{N}'$. Now take any $A_0 \in G(F_\wp)$ that we wish to factor. After translation, we may assume that $A_0 \in N(F_\wp)$ and that $B_0 := \eta(A_0) \in M'(F_\wp)$. Taking B_U, B_P as above, the elements $A_U := \eta^{-1}(B_U) \in G(F_U)$ and $A_P := \eta^{-1}(B_P) \in G(F_P)$ then satisfy $A_U A_P = A_0$, as desired. (See [HHK09], Theorems 3.2 and 3.4, for more details.)

In the more general situation, where $\mathcal{P}$, $\mathcal{U}$, and $\mathcal{B}$ can each have more than one element, one can generalize this and prove *simultaneous* factorization for collections of elements of $G(F_\wp)$, $\wp \in \mathcal{B}$. (See [HHK09], Theorem 3.6.) By Corollary 9.2, it follows that the local-global principle holds for G-torsors, or equivalently $\text{Ш}_{\mathcal{P}}(F, G)$, where G is any rational connected linear algebraic group over F.

10.2 Case of finite groups

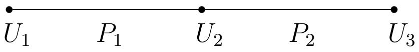
Above, the groups considered were connected. This leaves open the question of what happens if the group is disconnected. For example, suppose that G is an ordinary finite group, viewed as a finite linear algebraic group. As discussed in Section 7, G-torsors over F are then just the spectra of G-Galois F-algebras. For such an F-algebra A, we can take the normalization of $\widehat{X}$ in A, and obtain a G-Galois branched cover $\widehat{Y} \to \widehat{X}$ whose corresponding extension of F is A. (Here, the *normalization* is obtained by taking the integral closure S of R in A for every Zariski affine open subset $\operatorname{Spec}(R) \subset \widehat{X}$; and then taking $\widehat{Y}$ to be the union of the affine schemes $\operatorname{Spec}(S)$. More formally, $\widehat{Y} = \underline{\operatorname{Spec}}(\mathcal{S})$, where $\mathcal{S}$ is the integral closure of the structure sheaf $\mathcal{O}_{\widehat{X}}$ in A.) Here $\widehat{Y}$ is connected (in fact irreducible) if and only if A is a field.

Thus $H^1(F, G)$ is in bijection with the set of isomorphism classes of G-Galois branched covers of $\widehat{X}$. The subset $\text{Ш}_{\mathcal{P}}(F, G) \subseteq H^1(F, G)$ corresponds to those covers that are trivial over each F_P and each F_U. Covers of that type, which we call *split covers*, are necessarily unramified over $\widehat{X}$, since they are unramified (in fact, trivial) locally.

As we will discuss, the local-global principle for G can fail in this situation. But the obstruction $\text{Ш}_{\mathcal{P}}(F, G)$ is finite, and can be computed in terms of a graph Γ that is associated to the closed fiber X of $\widehat{X}$. Namely, Γ can be regarded as a topological space, and $\text{Ш}_{\mathcal{P}}(F, G)$ can be described in terms of its fundamental group.

This graph, called the *reduction graph*, is easiest to describe in the special case that there are exactly two branches of X at each point of $\mathcal{P}$, and that these two branches lie on distinct irreducible components of X. In this case, we can use the definition given in [DM69, p. 86]: the vertices of the graph correspond to the irreducible components of X, the edges correspond to the (singular) points that lie on two distinct components, and the vertices of a given edge correspond to the two components on which the corresponding point lies.

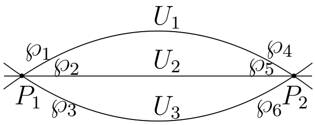
For example, the graph associated to the configuration pictured in Section 6.3 is as follows:



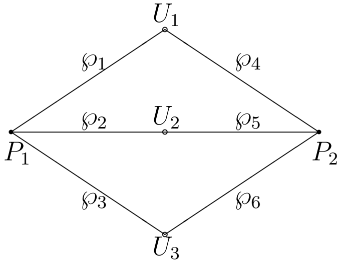
In the case that more than two components of X can meet at a point, the above definition does not apply. But one can instead consider a modified version, which in the above special case gives the barycentric subdivision of the graph as described above. This graph will then be homotopic to the graph above, and that will suffice for our purposes (for which only the homotopy class of the graph will be relevant).

Namely, we construct a bipartite graph, i.e. a graph whose vertex set $\mathcal{V}$ is partitioned into two subsets $\mathcal{V}_1, \mathcal{V}_2$ such that every edge connects a vertex in $\mathcal{V}_1$ to one in $\mathcal{V}_2$. Choose a non-empty finite subset $\mathcal{P}$ of X that contains all the points where the closed fiber X has more than one branch. (This contains in particular all the points where distinct irreducible components of X meet.) Let $\mathcal{U}$ consist of the connected components of $X \smallsetminus \mathcal{P}$; its elements U are in bijection with the irreducible components of X, by taking Zariski closures. We also obtain a set $\mathcal{B}$, consisting of the branches of X at the points of $\mathcal{P}$; each branch lies on the closure of a unique $U \in \mathcal{U}$. For the bipartite graph, let $\mathcal{V}_1 = \mathcal{P}$ and $\mathcal{V}_2 = \mathcal{U}$. The edges of the graph correspond to the branches of X at the points of $\mathcal{P}$, with the vertices of a branch corresponding to the associated point $P \in \mathcal{P}$ and element $U \in \mathcal{U}$.

As an example, consider the situation of a closed fiber X consisting of three components X_1, X_2, X_3 that all meet at two points P_1, P_2. Thus $U_i = X_i \smallsetminus \{P_1, P_2\}$. There are six branches $\wp_1, \dots, \wp_6$ on X at the points of $\mathcal{P}$, one for each pair (P_i, U_j). This configuration can be pictured as follows:



The associated bipartite graph is then as follows, where the five vertices correspond to $\mathcal{P} \sqcup \mathcal{U}$ and the six edges correspond to the six branches $\wp_i$.



In general, if we enlarge $\mathcal{P}$ by adding an additional closed point (and shrinking the corresponding set $U \in \mathcal{U}$ by deleting that point), then the homotopy class of the graph is unchanged. The same is true if we blow up $\widehat{X}$ at a regular point $P \in \mathcal{P}$.

Coming back to local-global principles, consider a finite group G, which we can also consider as a finite torsor over F. An element of $\text{Ш}_{\mathcal{P}}(F, G)$ induces a G-Galois covering space of the associated reduction graph Γ, as follows. This covering space is locally trivial over a neighborhood of each vertex, including the adjacent open edges. The local covers are glued over the overlaps, according to the patching data of the given element of $\text{Ш}_{\mathcal{P}}(F, G) \subseteq H^1(F, G)$. This data consists of the given $F_\wp$-isomorphisms of the trivial G-torsors over F_U and F_P that are obtained via base change from the given torsor over F. (These base changes are trivial G-torsors because the given torsor lies in $\text{Ш}_{\mathcal{P}}(F, G)$.) Conversely, a covering space of Γ induces an element of $\text{Ш}_{\mathcal{P}}(F, G)$, by patching.

Thus the elements of $\text{Ш}_{\mathcal{P}}(F, G)$ are classified by $\text{Hom}(\pi_1(\Gamma), G)/\sim$, where $\sim$ is the equivalence relation given by conjugating maps by elements of G. This is because this set classifies the (possibly disconnected) G-Galois covering spaces of the graph Γ. (If we instead classified *pointed* G-Galois covering spaces of Γ, we would not need to mod out by this equivalence relation.)

Given $\widehat{X}$ and $\mathcal{P}$, we can find Γ explicitly, and then compute

$$\text{Ш}_{\mathcal{P}}(F, G) = \text{Hom}(\pi_1(\Gamma), G)/\sim,$$

which is a finite set. Namely, $\pi_1(\Gamma)$ is a free group of some finite rank r, where r is the "number of loops in Γ". We can then identify $\text{Hom}(\pi_1(\Gamma), G)$ with G^r. Thus $\text{Ш}_{\mathcal{P}}(F, G)$ can be identified with G^r/G, where the action of G on G^r is by uniform conjugation.

Note that this description of $\text{Ш}_{\mathcal{P}}(F, G)$ is independent of the choice of $\mathcal{P}$ and of $\widehat{X}$ (which can be varied by blowing up), since the homotopy class of Γ is independent of these choices. Note also, for $G \neq 1$, that $\text{Ш}_{\mathcal{P}}(F, G)$ vanishes (or equivalently, the local-global principle holds) if and only if Γ is a tree.

10.3 Disconnected rational groups

We have considered the obstructions $Ш_{\mathcal{P}}(F,G)$ to local-global principles in the special cases that G is a rational connected group, or if G arises from a finite group. Here we turn to linear algebraic groups such as $\mathrm{O}(n)$ that are at neither of those two extremes.

We say that a (not necessarily connected) linear algebraic group G is *rational* if each connected component of the F-variety G is a rational F-variety. An equivalent characterization is that the identity component G^0 of G is a rational connected group and that each connected component of G has an F-point. Another equivalent characterization is that G^0 is a rational connected group and the quotient $\bar{G} := G/G^0$ is a finite constant group scheme (i.e. arises from an ordinary finite group) such that $G(F) \to \bar{G}(F)$ is surjective.

For such a group G, we have a short exact sequence

$$1 \to G^0 \to G \to \bar{G} \to 1$$

of groups, with G^0 connected and $\bar{G}$ a finite constant group. Associated to this is a long exact cohomology sequence, which involves just H^0 and H^1 if G is not commutative (see [Ser00, I.5.5, Proposition 38]). Combining that with the Mayer-Vietoris sequence in Theorem 9.1, one obtains a short exact sequence of pointed sets (and of groups, if G is commutative):

$$1 \to Ш_{\mathcal{P}}(F,G^0) \to Ш_{\mathcal{P}}(F,G) \to Ш_{\mathcal{P}}(F,\bar{G}) \to 1.$$

(See Corollary 2.6 of [HHK11a].) Here the first term vanishes, as discussed in Section 10.1; and the third term is given by $\mathrm{Hom}(\pi_1(\Gamma),\bar{G})/\sim)$, as in Section 10.2. We then obtain the following description of the obstruction $Ш_{\mathcal{P}}(F,G)$ to the local-global principle for G-torsors (where some extra work is needed if G is not commutative):

Theorem 10.1. *Let $\widehat{X}$ be a normal projective model of a field F of transcendence degree one over a complete discretely valued field. Let $\mathcal{P}$ be a non-empty finite subset of the closed fiber X that contains all the points where X has more than one branch. Let G be a rational linear algebraic group over F, and write $\bar{G} = G/G^0$. Then*

$$Ш_{\mathcal{P}}(F,G) = \mathrm{Hom}(\pi_1(\Gamma),\bar{G})/\sim .$$

As a consequence, $Ш_{\mathcal{P}}(F,G)$ is finite, and its order can be explicitly computed, for G rational. Moreover we obtain the precise condition for the local-global principle to hold:

Corollary 10.2. *With F as above and G a rational linear algebraic group over F, the local-global principle for G-torsors holds if and only if either G is connected or the reduction graph is a tree.*

Namely, these are precisely the conditions under which $\mathrm{Hom}(\pi_1(\Gamma), \bar{G})$, and hence also $\mathrm{Hom}(\pi_1(\Gamma), \bar{G})/\sim$, is trivial.

The above discussion answers the question of whether there is a local-global principle, and more generally what the obstruction is to such a principle, in the case of *torsors* for rational groups. But the question remains about local-global principles for other homogeneous spaces for such groups.

To be more precise, consider an F-variety H together with a right G-action $\alpha : H \times G \to H$. We say that *$G$ acts transitively on H* if for every field extension E/F, the action of $G(E)$ on $H(E)$ is transitive. Every torsor has this property; but not conversely, since transitive actions can have stabilizers.

We can then show:

Proposition 10.3. *Let G be a linear algebraic group over a field F as above. If the local-global principle holds for G-torsors, then it also holds for* **all** *F-varieties H on which G acts transitively. Equivalently, the simultaneous factorization condition for G implies that local-global principles hold for all transitive G-varieties.*

Proof. The first part follows from the second part together with Corollary 9.2. Concerning the second part, for simplicity, we sketch the proof in the simple case where we have just one U and one P. More abstractly, we have four fields $F \subseteq F_1, F_2 \subseteq F_0$ satisfying Conditions 4.1 and 4.2 (the latter also being satisfied here, as discussed in Section 6).

Say that we have points $P_1 \in H(F_1)$ and $P_2 \in H(F_2)$. We wish to find a point $P \in H(F)$. Viewing $P_i \in H(F_0)$, the transitivity property implies that there exists some $A_0 \in G(F_0)$ such that $P_1 A_0 = P_2$. By the factorization condition, there exist $A_1 \in G(F_1)$ and $A_2 \in G(F_2)$ such that $A_0 = A_1 A_2$. Let $P_1' = P_1 A_1 \in H(F_1)$ and let $P_2' = P_2 A_2^{-1} \in H(F_2)$. Then P_1', P_2' define the same point in $H(F_0)$. Since $F_1 \cap F_2 = F$, it follows that this common point P is defined over F, i.e. lies in $H(F)$, as desired. □

11 Applications of local-global principles

11.1 Applications to quadratic forms

As before, let F be the function field of a curve over a complete discretely valued field K, and let $\widehat{X}$ be a normal projective model of F over the valuation ring T of K. Suppose now that $\mathrm{char}(F) \neq 2$. If q is a quadratic form over F, then after a change of variables it can be diagonalized, i.e. written as $q = \sum_{i=1}^n a_i z_i^2$. We assume that q is regular, i.e. each $a_i \neq 0$. As in Section 8, q defines a quadric hypersurface Q in $\mathbb{P}_K^{n-1}$; and q is isotropic over F if and only if Q has an F-point.

In this situation, we may choose a non-empty finite subset $\mathcal{P} \subset X$ of the closed fiber with properties as before, and let $\mathcal{U}$ be the set of connected components of the complement. The local-global principle for q over F with respect to the overfields

F_ξ, for $\xi \in \mathcal{P} \sqcup \mathcal{U}$, then asserts that q is isotropic over F if and only if it is isotropic over each F_ξ; or equivalently,

$$Q(F_\xi) \neq \varnothing \text{ for all } \xi \in \mathcal{P} \sqcup \mathcal{U} \ \Leftrightarrow\ Q(F) \neq \varnothing.$$

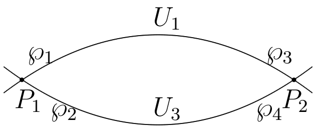
The question is whether this principle holds.

Let $\mathrm{O}(q)$ be the orthogonal group associated to the quadratic form q; i.e. the subgroup of GL_n that preserves the form q. The projective hypersurface Q is a homogeneous space for the group $\mathrm{O}(q)$, though not a torsor for this group. This group is rational, by the Cayley parametrization ([KMRT98], p. 201, Exercise 9), but not connected. In fact, it has two connected components, with the identity component being $\mathrm{SO}(q)$.

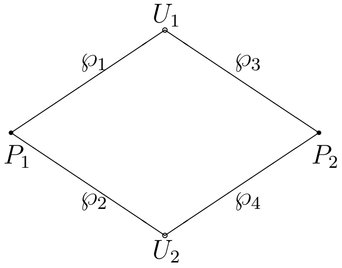
Note that $\dim(Q) = n-2$. In particular, Q is of dimension zero if q is a binary quadratic form. In that case, Q can consist of two points, and be disconnected (e.g. if $q = x^2 - y^2$). But if $n > 2$, then Q is connected, and hence the connected subgroup $\mathrm{SO}(q)$ also acts transitively on Q. Since the group $\mathrm{SO}(q)$ is both rational and connected, the local-global principle holds for Q, using Corollary 10.2 and Proposition 10.3. Rephasing this, we have:

Theorem 11.1. *The local-global principle for isotropy of quadratic forms over F, with respect to patches F_ξ, holds for forms of dimension $n > 2$.*

This result is analogous to the classical Hasse-Minkowski theorem for quadratic forms over global fields (see Section 8). But unlike that situation, here there can be an exception, in the case of binary forms. And in fact, there really do exist examples of binary forms in which the principle does not hold. By Corollary 10.2, any such example must involve a field F for which the reduction graph associated to a model is not a tree. The simplest case of this is a Tate curve, where the general fiber is a genus one curve over K, and the closed fiber consists of one or more projective lines whose crossings form a "loop". One such possibility consists of two projective lines that cross each other at two points. The closed fiber then looks like



and the reduction graph looks like



which is not a tree. An explicit example of this situation is the double cover $\widehat{X}$ of $\mathbb{P}^1_T$ given in affine coordinates by $y^2 = x(x-t)(1-xt)$, with P_1, P_2 being the points $x = 0, \infty$ on the closed fiber ($t = 0$). The form $q = x(x-t)z_1^2 - z_2^2$ is locally isotropic but not isotropic over the function field of $\widehat{X}$. (This phenomenon was observed by J.-L. Colliot-Thélène, based on [Sai83, Example 2.7]. See [CPS12, Remark 4.4] and [HHK09, Example 4.4] for Tate curve examples with an irreducible closed fiber.)

There is another way to understand the fact that the local global principle will hold for *all* quadratic forms over F (without any restriction on dimension) if and only if the associated reduction graph Γ is a tree. This concerns the *Witt group* of the field F. This group $W(F)$ is defined to be the set of equivalence classes of quadratic forms, with two forms being considered equivalent if they differ by a hyperbolic form $\sum_{i=1}^m (x_i^2 - y_i^2)$, up to a change of variables. These equivalence classes form a group under orthogonal direct sum, i.e. adding representative forms in disjoint sets of variables. The local-global principle for the Witt group would assert that $W(F)$ is trivial if and only if each group $W(F_\xi)$ is trivial. As a result of the two-dimensional exception in Theorem 11.1, this will not always hold, due to the presence of two-dimensional forms that become isotropic over each F_ξ but are not isotropic over F. (A two-dimensional form is hyperbolic if and only if it is isotropic.) Instead, the obstruction to this principle can be found explicitly, using the results described above:

Proposition 11.2. *The kernel of the local-global map*

$$W(F) \to \prod_{\xi \in \mathcal{U} \sqcup \mathcal{P}} W(F_\xi)$$

on Witt groups is $\mathrm{Hom}(\pi_1(\Gamma), \mathbb{Z}/2)$*, where* Γ *is the reduction graph of any regular projective model of* F*. Hence the local-global principle for Witt groups holds if and only if* Γ *is a tree.*

The key ingredients in the proof are Theorem 11.1 and Theorem 10.1 with G an orthogonal group. For more details, see [HHK11a, Theorem 9.6] (where the assertion was phrased somewhat differently).

A numerical application of the above ideas is motivated by classical quadratic form theory over global fields. Namely, the Hasse-Minkowski theorem implies the following theorem of Meyer: If q is a quadratic form over $\mathbb{Q}$ of dimension greater than four, and if q is not positive or negative definite, then q is isotropic over $\mathbb{Q}$. (See [Ser73, Section IV.3.2].) More generally, this assertion holds over any global field. Note that in the function field case, no forms are positive or negative definite, and the assertion simplifies, to say that every quadratic form in more than four variables is isotropic.

In the situation of the fields F that we have been considering, i.e. one-variable function fields over a complete discretely valued field K, Theorem 11.1 can similarly be used to obtain an analogous result concerning isotropy of forms in "too many variables." (Since we are in the function field case, there is no issue of being positive or negative definite.) In particular, if $K = \mathbb{Q}_p$ with $p \neq 2$, and if q is a quadratic form over F of dimension greater than 8, then q is isotropic over F. And to give another example, if $K = \mathbb{Q}_p((t))$ with $p \neq 2$, and if q is a quadratic form over F of dimension greater than 16, then q is isotropic over F. These and other related results (e.g. for fields of the form F_ξ with $K = \mathbb{Q}_p$) were shown in [HHK09, Section 4] using the above methods.

Until the 1990's, it was not known if there were results of this sort, even in the case of $K = \mathbb{Q}_p$. The above result in that case was first shown by a different approach in [PS10]. Another proof, in [Lee12], later showed that that result holds for $K = \mathbb{Q}_2$ as well. Building on [HHK09, Section 4] and [HHK11a], it has been recently shown more generally in [PS13] that if K is a complete discretely valued field of characteristic zero whose residue field is a perfect field of characteristic two, then any quadratic form of dimension greater than 8 must be isotropic over F. These results, however, were obtained by proofs that did not apply to such cases as $K = \mathbb{Q}_p((t))$.

11.2 Applications to central simple algebras

By a similar approach, local-global principles can be obtained for central simple algebras over our field F. We begin by reviewing some background; see also [Pie82].

Recall that Wedderburn's Theorem states that every (finite dimensional) central simple algebra A over a field K is of the form $\mathrm{Mat}_n(D)$, where D is a (central) division algebra over K. Moreover the integer n is uniquely determined by A, and D is unique up to isomorphism. The *index* of A is the degree of the division algebra D (see Section 3.3); equivalently, it is the minimal value of $[E : F]$ where E/F is a field extension such that $A_E := A \otimes_F E$ is split over E (i.e. a matrix algebra).

One says that two central simple F-algebras A, A' are *Brauer equivalent* if the associated division algebras are isomorphic. The set of Brauer equivalence classes form a group under tensor product, called the *Brauer group* $\mathrm{Br}(K)$ of

K. By the above, its elements are in bijection with isomorphism classes of K-division algebras. One says that A is *split* if its Brauer class is the trivial class; i.e. $A = \mathrm{Mat}_n(K)$ for some n. This is equivalent to the condition $\mathrm{ind}(A) = 1$. As mentioned in Section 8, the classical theorem of Albert, Brauer, Hasse, and Noether says that if K is a *global* field and if A is a central simple K-algebra, then A is split over K if and only if $A_v := A \otimes_K K_v$ is split over K_v for every completion K_v of K.

Analogously in our situation, with F, $\widehat{X}$, $\mathcal{P}$, and $\mathcal{U}$ as before, we can show the following local-global principle for central simple algebras ([HHK09], Theorem 5.1):

Theorem 11.3. *A central simple F-algebra A is split over F if and only if $A_\xi := A \otimes_F F_\xi$ is split over F_ξ for every $\xi \in \mathcal{P} \sqcup \mathcal{U}$. In fact even more is true:* $\mathrm{ind}(A) = \mathrm{lcm}_{\xi \in \mathcal{P} \sqcup \mathcal{U}} \mathrm{ind}(A_\xi)$.

The proof parallels that of Theorem 11.1. But instead of using the rational connected group $\mathrm{SO}(q)$, we use the group $\mathrm{GL}_1(A)$. If the degree of A is d (i.e. $\dim_F(A) = d^2$), then $\mathrm{GL}_1(A)$ is a Zariski open subset of $\mathbb{A}_F^{d^2}$, and hence it is a rational and connected F-group. There are canonically defined varieties $\mathrm{SB}_i(A)$, known as *generalized Severi-Brauer varieties* of A, on which $\mathrm{GL}_1(A)$ acts transitively. (See [VdB88], p. 334, and [See99], Theorem 3.6.) Using these in place of the hypersurface Q, the argument in the quadratic form case carries over to provide the desired local-global principle for central simple F-algebras. Note that here, unlike in the quadratic forms situation, there is no exception to the principle. That is because here we consider varieties on which a rational connected group acts transitively, whereas in the quadratic forms case the connectivity property can fail for $n = 2$.

As in the case of quadratic forms, the local-global principle for central simple algebras can be used to obtain the values of numerical invariants associated to F that concern the behavior of central simple algebras. If A is a central simple F-algebra, then the Brauer class of A has finite order in $\mathrm{Br}(F)$; this is called the *period* of A. Regardless of the ground field, the period always divides the index, and moreover those two integers are divisible by precisely the same set of primes. Thus for every A there is an integer e such that $\mathrm{ind}(A)$ divides $\mathrm{per}(A)^e$. An important question is whether there is a value of e that works for *all* A (or at least, all A whose period is not divisible by the relevant characteristic). Paralleling the argument in the case of quadratic forms, such uniform values of e can be found in many cases. For example, if $K = \mathbb{Q}_p((t))$ then for algebras of period not divisible by p, we have $e = 2$ for A over K, and $e = 3$ for A over F. As in the situation of quadratic forms, there are also similar results over the fields F_P and F_U. See [HHK09, Section 5], where the above is carried out. See [Lie11] for a different proof in the case of the field F. See also [PS13] for recent results in the case that K is of mixed characteristic $(0, p)$ and the period of the algebra A is a power of p.

12 Complements

12.1 Other local-global set-ups

Above, we have considered local-global principles for one-variable function fields F over a complete discretely valued field K. These have been phrased in terms of a finite set of overfields F_ξ of F, corresponding to a choice of patches on a normal projective model $\widehat{X}$ of F over the valuation ring T of K. While inspired by classical local-global principles for global fields, this set-up is not quite analogous, since in the classical case one takes infinitely many overfields, corresponding to all the completions of the global field. Also, the classical local-global principles do not depend on making a choice, unlike our situation above, where we choose a non-empty finite subset $\mathcal{P}$ of the closed fiber X of $\widehat{X}$, in order to define our set of overfields.

The above framework can be modified, however, to be closer in spirit to the classical situation. We describe two ways to do this.

The first of these begins with a normal model $\widehat{X}$ of F, and considers the set of *all* the fields F_P for $P \in X$. This includes not only the infinitely many closed points of X, but also the finitely many non-closed points of X. These latter points are the generic points η of the (finitely many) irreducible components U of X. (Here F_η is the fraction field of the completion $\widehat{R}_\eta$ of the local ring $R_\eta = \mathcal{O}_{\widehat{X},\eta}$ of $\widehat{X}$ at η.)

In this situation, unlike our prior set-up above, we do not have overfields that would play the role of the fields $F_\wp$. So we cannot ask for patching to hold. But we can still ask for local-global principles. In the case of G-torsors, for G a linear algebraic group over F, this says that the local-global map

$$H^1(F, G) \to \prod_{P \in X} H^1(F_P, G)$$

is injective. Let $Ш_X(F, G)$ denote the kernel of this map (with the model $\widehat{X}$ being understood). It turns out that if G is rational over F, then $Ш_X(F, G)$ is naturally isomorphic to $Ш_{\mathcal{P}}(F, G)$, for any choice of a finite non-empty subset $\mathcal{P}$ of X that contains all the points where X has more than one branch ([HHK11a, Corollary 5.6]). Thus the obstruction $Ш_{\mathcal{P}}(F, G)$ is more canonical than it had appeared to be, as it depends only on the model $\widehat{X}$. Moreover $Ш_X(F, G)$ can therefore be identified with $\mathrm{Hom}(\pi_1(\Gamma), \bar{G})/\sim$, where the reduction graph Γ is taken with respect to any choice of $\mathcal{P}$ as above (and where $\bar{G} = G/G^0$ as before).

The second modification instead considers the set Ω_F of (equivalence classes of) discrete valuations v on the field F, just as in the classical case of global function fields one considers the set of discrete valuations. For G a linear algebraic group over F, let $Ш(F, G)$ be the kernel of the local-global map

$$H^1(F, G) \to \prod_{v \in \Omega_F} H^1(F_v, G).$$

Then $Ш(F, G)$ is naturally contained in $Ш_X(F, G)$ ([HHK11a, Proposition 8.2]), and the question is whether they are equal. In general this is unknown, but it is known in several cases (see [HHK11a, Theorem 8.10]), e.g. if G is rational and the residue field k of T is algebraically closed of characteristic zero. Moreover the local-global principle, in this sense, is known to hold for quadratic forms of dimension greater than two provided that the residue field k is not of characteristic two (see [CPS12], Theorem 3.1), thereby carrying over Theorem 11.1 to this situation. Moreover the local-global principle for central simple algebras, given in the first part of Theorem 11.3, also carries over; see [CPS12, Theorem 4.3(ii)] and [HHK11a, Theorem 9.13]. The second part of Theorem 11.3 also has an analog for discrete valuations, at least in the presence of sufficiently many roots of unity; see [RS13, Theorem 2], whose proof relies on Theorem 11.3.

12.2 Non-rational groups

Although our patching results for torsors do not require that the linear algebraic group is rational, the proofs of the results above concerning local-global principles do require that. There is then the question of whether such results hold more generally. In particular, there is the question of whether local-global principles hold for connected linear algebraic groups that are not rational.

In certain cases where a connected linear algebraic group G over F is not known to be rational, local-global principles have been shown. In particular, this was done for groups of type G_2 in [HHK11a, Example 9.4], by using local-global principles for quadratic forms. This has also been done for various other types of groups by combining cohomological invariants with local-global principles for higher Galois cohomology; see [CPS12, Section 5], [Hu12], and [HHK12, Section 4]. Also, in [Kra10], it was shown that local-global principles hold for connected linear algebraic groups that are *retract rational*, a condition that is strictly weaker than being rational.

These results suggest the possibility that local-global principles might hold for all connected linear algebraic groups over our fields F. But in fact, this is not the case. In [CPS13], examples were obtained of a connected linear algebraic group G over a field F as above, such that the local-global principle for G-torsors fails. In fact, it fails in each of the three settings discussed above: with respect to patches, point on the closed fiber, and discrete valuations.

In light of this, it would be very interesting to find a necessary and sufficient condition on connected linear algebraic groups G over F for local-global principles to hold.

References

[Abh69] Shreeram S. Abhankar. *Resolution of singularities of algebraic surfaces.* In *1969 Algebraic Geometry (Internat. Colloq., Tata Inst. Fund. Res., Bombay, 1968)*, pp. 1–11, Oxford Univ. Press, London.

[Bir17] George David Birkhoff. *A theorem on matrices of analytic functions*. Math. Ann. 74 (1917), 240–251.

[Bor91] Armand Borel. *Linear Algebraic Groups*, second edition. Graduate Texts in Mathematics, volume 126. Springer-Verlag, New York, Berlin and Heidelberg, 1991.

[BS64] Armand Borel and Jean-Pierre Serre. *Théorèmes de finitude en cohomologie galoisienne.* Comment. Math. Helv. **39** (1964), 111–164.

[Bou72] Nicolas Bourbaki. *Commutative Algebra.* Addison-Wesley Publishing Co., 1972.

[CPS12] Jean-Louis Colliot-Thélène, R. Parimala, and V. Suresh. *Patching and local-global principles for homogeneous spaces over function fields of p-adic curves.* Comment. Math. Helv. **87** (2012), 1011–1033.

[CPS13] Jean-Louis Colliot-Thélène, R. Parimala, and V. Suresh. *Lois de réciprocité supérieures et points rationnels*. 2013 manuscript. Available at arXiv:math/1302.2377.

[Con12] Brian Conrad. *Finiteness theorems for algebraic groups over function fields*. Compos. Math. **148** (2012), 555–639.

[DM69] Pierre Deligne and David Mumford. *The irreducibility of the space of curves of given genus*. Publ. Math. IHES, vol. 36, 75–109, 1969.

[Dou64] Adrien Douady. *Détermination d'un groupe de Galois*. C.R. Acad. Sci. Paris **258** (1964), 5305–5308.

[Fr95] Michael Fried, ed. *Recent Developments in the Inverse Galois Problem*, AMS Contemporary Math. Series, vol. 186, 1995.

[GH78] Phillip Griffiths, Joseph Harris. *Principles of algebraic geometry*. Pure and Applied Mathematics. Wiley-Interscience, New York, 1978.

[Gro59] Alexander Grothendieck. *Géométrie formelle et géométrie algebrique.* Sem. Bourbaki **182** (1959), 1–28.

[Gro61] Alexander Grothendieck. *Élements de géométrie algebrique, III.* Publ. Math. IHES, vol. 11, 1961.

[HJ98] Dan Haran and Moshe Jarden. *Regular split embedding problems over function fields of one variable over ample fields.* J. Algebra **208** (1998), 147–164.

[HJ00] Dan Haran and Moshe Jarden. *The absolute Galois group of $C(x)$.* Pacific J. Math. **196** (2000), 445–459.

[HV96] Dan Haran and Helmut Völklein. *Galois groups over complete valued fields.* Israel J. Math. **93** (1996), 9–27.

[Hrb84] David Harbater. *Convergent arithmetic power series.* Amer. J. Math. **106** (1984), 801–846.

[Hrb87] David Harbater. *Galois covers of the affine line.* In: *Number Theory: New York, 1984-85.* Springer LNM, vol. 1240 (1987), pp. 165–195.

[Hrb94] David Harbater. *Abhyankar's conjecture on Galois groups over curves.* Inventiones Math., **117** (1994), 1–25.

[Hrb95] David Harbater. *Fundamental groups and embedding problems in characteristic p.* In [Fr95], pp. 353–369.

[Hrb03] David Harbater. *Patching and Galois theory.* In [Sch03], pp.313–424.

[HH10] David Harbater and Julia Hartmann. *Patching over fields.* Israel J. Math. **176** (2010), 61–107.

[HHK09] David Harbater, Julia Hartmann, and Daniel Krashen. *Applications of Patching to Quadratic Forms and Central Simple Algebras.* Inventiones Math. **178** (2009), 231–263.

[HHK11] David Harbater, Julia Hartmann, and Daniel Krashen. *Patching subfields of division algebras.* Transactions of the AMS, **363** (2011), 3335–3349.

[HHK11a] David Harbater, Julia Hartmann, and Daniel Krashen. *Local–global principles for torsors over arithmetic curves.* 2011 manuscript. Available at arXiv:math/1108.3323.

[HHK12] David Harbater, Julia Hartmann, and Daniel Krashen. *Local–global principles for Galois cohomology.* 2012 manuscript. To appear in Commentarii Mathematici Helvetici. Available at arXiv:math/1208.6359.

[Hrt05] Julia Hartmann. *On the inverse problem in differential Galois theory.* J. Reine Angew. Math. **586** (2005), 21–44.

[Hrt07] Julia Hartmann. *Patching and differential Galois groups.* In: Arithmetic and Differential Galois Groups, Oberwolfach reports vol. 4, no.2, EMS 2007.

[Hts77] Robin Hartshorne. *Algebraic geometry.* Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York - Heidelberg, 1977.

[Hu12] Yong Hu. *Hasse principle for simply connected groups over function fields of surfaces.* 2012 manuscript. Available at arXiv:math/1203.1075.

[Iwa53] Kenkichi Iwasawa. *On solvable extensions of algebraic number fields.* Annals of Math. **58** (1953), 548–572.

[Jar95] Moshe Jarden. *On free profinite groups of uncountable rank.* In [Fr95], pp. 371–383.

[Jar11] Moshe Jarden. *Algebraic patching.* Springer Monographs in Mathematics. Springer, Heidelberg, 2011.

[KMRT98] Max-Albert Knus, Alexander S. Merkurjev, Markus Rost, and Jean-Pierre Tignol. *The Book of Involutions.* American Mathematical Society, Providence, RI, 1998.

[Kra10] Daniel Krashen. *Field patching, factorization and local-global principles* (S. Garibaldi, R. Sujatha, and V. Suresh, eds.). In: *Quadratic forms, linear algebraic groups, and cohomology.* Developments in Mathematics, Vol. 18, pp. 57–82, Springer, 2010.

[Lee12] David Leep. *The u-invariant of p-adic function fields.* J. Reine Angew. Math., DOI 10.1515/crelle.2012.029, 2012.

[Lie11] Max Lieblich. *Period and index in the Brauer group of an arithmetic surface*, with an appendix by Daniel Krashen. J. Reine Angew. Math. **659** (2011), 1–41.

[Lip75] Joseph Lipman. *Introduction to resolution of singularities.* In *Algebraic geometry (Proc. Sympos. Pure Math., Vol. 29, Humboldt State Univ., Arcata, Calif., 1974)*, pages 187–230. Amer. Math. Soc., Providence, R.I., 1975.

[Liu95] Qing Liu. *Tout groupe finit est un groupe de Galois sur* $\mathbb{Q}_p(T)$, d'aprés Harbater. In [Fr95], pp. 261–265.

[MP03] B. Heinrich Matzat and Marius van der Put. *Constructive differential Galois theory.* In [Sch03], 425–467.

[PS10] R. Parimala and V. Suresh. *The u-invariant of the function fields of p-adic curves.* Annals of Math. **172** (2010), 1391–1405.

[PS13] R. Parimala and V. Suresh. *Period-index and u-invariant questions for function fields over complete discretely valued fields.* 2013 manuscript. Available at arXiv:math/1304.2214.

[Pie82] Richard S. Pierce. *Associative algebras.* Graduate Texts in Mathematics, No. 88. Springer-Verlag, New York - Heidelberg - Berlin, 1980.

[Pop94] Florian Pop. *Half Riemann Existence Theorem with Galois action.* In: *Algebra and number theory*, G. Frey and J. Ritter, eds., de Gruyter Proceedings in Math. (1994), 1–26.

[Pop95] Florian Pop. *Étale Galois covers of affine smooth curves.* Inventiones Math. **120** (1995), 555–578.

[Ray74] Michel Raynaud. *Géométrie analytique rigide d'aprés Tate, Kiehl,....* Bull. Soc. Math. France, Mémoire 39–40 (1974), 319–327.

[Ray94] Michel Raynaud. *Revêtements de la droite affine en caractéristique $p > 0$ et conjecture d'Abhyankar.* Inventiones Math. **116** (1994), 425–462.

[RS13] B. Surendranath Reddy, V. Suresh. *Admissibility of groups over function fields of p-adic curves.* Adv. Math. **237** (2013), 316–330.

[Sai83] Shuji Saito, *Class field theory for curves over local fields.* J. Number Theory **21** (1985), 44–80.

[Sch68] Murray M. Schacher. *Subfields of division rings. I.* J. Algebra **9** (1968), 451–477.

[Sch03] Leila Schneps, ed. *Galois groups and fundamental groups.* MSRI Publications series, vol. 41, Cambridge University Press, 2003.

[See99] George F. Seelinger. *Brauer-Severi schemes of finitely generated algebras.* Israel J. of Math. **111** (1999), 321–337.

[Ser56] Jean-Pierre Serre. *Géométrie algébrique et géométrie analytique.* Ann. Inst. Fourier, Grenoble **6** (1956), 1–42.

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic.* Graduate Texts in Mathematics, No. 7. Springer-Verlag, New York - Heidelberg - Berlin, 1973.

[Ser00] Jean-Pierre Serre. *Galois Cohomology.* Springer Monographs in Mathematics, Berlin, Heidelberg and New York, 2000.

[Son83] Jack Sonn. *$\mathbb{Q}$-admissibility of solvable groups.* J. Algebra **84** (1983), 411–419.

[Tat71] John Tate. *Rigid analytic spaces.* Inventiones Math. **12** (1971), 257–289.

[TT79] Carol Tretkoff and Marvin Tretkoff. *Solution of the inverse problem of differential Galois theory in the classical case.* Amer. J. Math. **101** (1979), 1327–1332.

[VdB88] Michel Van den Bergh. *The Brauer-Severi scheme of the trace ring of generic matrices.* In: *Perspectives in ring theory (Antwerp, 1987)*, pp. 333–338, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 233, Kluwer Acad. Publ., Dordrecht, 1988.

[Völ96] Helmut Völklein. *Groups as Galois groups. An introduction.* Cambridge Studies in Advanced Mathematics, vol. 53. Cambridge University Press, Cambridge, 1996.

David Harbater
Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104-6395, USA
email: harbater@math.upenn.edu

The author was supported in part by NSF grant DMS-0901164.

Travaux mathématiques, Volume 23 (2013), 87–137, © by the author

Algebraic Patching[1]

by Moshe Jarden

Abstract

The most effective challenge to the inverse problem of Galois theory has been Hilbert Irreducibility Theorem. Indeed, one may use both arithmetic and geometry in order to realize finite groups over $\mathbb{Q}(t)$. Once this has been successfully done for a finite group G, HIT yields many specializations of t to elements of $\mathbb{Q}$ that lead to a realization of G over $\mathbb{Q}$. However, the realization of G over $\mathbb{Q}(t)$ usually requires the existence of rational point on a certain algebraic variety defined over $\mathbb{Q}$. Unfortunately, one can not always guarantee the existence of such a point, so the inverse Galois problem over $\mathbb{Q}$ is still wide open.

The only known class of fields for which points that lead to realization of all finite groups exist is that of "ample fields". A field K is said to be **ample** if every absolutely irreducible curve C defined over K with a K-rational simple point has infinitely many K-rational points. Among others, PAC fields, Henselian fields, and real closed fields are ample. Using a method called "algebraic patching" we will prove that if K is an ample field, then every finite split embedding problem over $K(t)$ is properly solvable. In particular, if K is countable and algebraically closed, this implies that $\mathrm{Gal}(K)$ is isomorphic to the free profinite group $\hat{F}_\omega$ on countably many generators. Also, if K is PAC and countable, then K is Hilbertian if and only if $\mathrm{Gal}(K) \cong \hat{F}_\omega$.

MSC (2010): 12E30.

[1] For more details, including exact references, see "Algebraic Patching", Springer 2011, by Moshe Jarden.

Contents

Introduction

The ultimate goal of Galois theory is to describe the structure of the absolute Galois group $\mathrm{Gal}(\mathbb{Q})$ of $\mathbb{Q}$. This structure will be specified as soon as we know which finite embedding problems can be solved over $\mathbb{Q}$. If every finite Frattini embedding problem and every finite split embedding problem is solvable, then every embedding problem is solvable [9, Prop. 22.5.8]. However, not every finite Frattini problem over $\mathbb{Q}$ can be solved. For example,

$$(\mathrm{Gal}(\mathbb{Q}) \to \mathrm{Gal}(\mathbb{Q}(\sqrt{-1})/\mathbb{Q}),\ \mathbb{Z}/4\mathbb{Z} \to \mathrm{Gal}(\mathbb{Q}(\sqrt{-1})/\mathbb{Q}))$$

is an unsolvable Frattini embedding problem. So, one may ask:

Problem A. Is every finite split embedding problem over $\mathbb{Q}$ solvable?

More generally, one would like to know:

Problem B. Let K be a Hilbertian field. Is every finite split embedding problem over K solvable?

An affirmative answer to Problem B will follow from an affirmative answer to the problem for the subfamily of Hilbertian fields consisting of all rational function fields:

Problem C (Débes–Deschamps)**.** Let K be a field and x a variable. Is every finite split embedding problem over $K(x)$ solvable?

1 Ample Fields

The most significant development around Problem C is its affirmative solution for ample fields K. This family includes two subfamilies that seemed to have nothing in common: PAC fields and Henselian fields. Indeed, if K is a PAC field and v is a valuation of K, then the Henselization K_v of K at v is the separable closure K_s of K. In particular, if a PAC field is not separably closed, then it is not Henselian.

Florian Pop made a surprising yet simple and useful observation that both PAC fields and Henselian fields are existentially closed in the fields of formal power series over them. This property is one of several equivalent definitions of an ample field.

Proposition 1.1. *The following conditions on a field K are equivalent:*

(a) For each absolutely irreducible polynomial $f \in K[X,Y]$, the existence of a point $(a,b) \in K^2$ such that $f(a,b) = 0$ and $\frac{\partial f}{\partial Y}(a,b) \neq 0$ implies the existence of infinitely many such points.

(b) *Every absolutely irreducible K-curve C with a simple K-rational point has infinitely many K-rational points.*

(c) *If an absolutely irreducible K-variety V has a simple K-rational point, then $V(K)$ is Zariski-dense in V.*

(d) *Every function field of one variable over K that has a K-rational place has infinitely many K-rational places.*

(e) *K is existentially closed in each Henselian closure $K(t)^h$ of $K(t)$ with respect to the t-adic valuation.*

(f) *K is existentially closed in $K((t))$.*

Proof. We only prove the implication "(f) $\Longrightarrow$ (a)" and refer the reader for the other implications to [12, Lemma 5.3.1].

Inductively suppose there exist $(a_i, b_i) \in K^2$, $i = 1, \dots, n$, such that $f(a_i, b_i) = 0$ and $a_1, \dots, a_n$ are distinct. We choose $a' \in K[[t]]$, t-adically close to a such that $a' \neq a_i$, $i = 1, \dots, n$. Then $f(a', b)$ is t-adically close to 0 and $\frac{\partial f}{\partial Y}(a', b) \neq 0$. Since $K((t))$ is Henselian, there exists $b' \in K[[t]]$ such that $f(a', b') = 0$ and $\frac{\partial f}{\partial Y}(a', b') \neq 0$. Since K is existentially closed in $K((t))$, there exists $a_{n+1}, b_{n+1} \in K$ such that $f(a_{n+1}, b_{n+1}) = 0$ and $a_{n+1} \neq a_1, \dots, a_n$. This concludes the induction. □

Corollary 1.2. *Every ample field is infinite.*

It is possible to strengthen Condition (b) of Proposition 1.1 considerably.

Lemma 1.3 (Fehm [7, Lemma 4]). *Let K be an ample field, C an absolutely irreducible curve defined over K with a simple K-rational point, and $\phi\colon C \to C'$ a separable dominant K-rational map to an affine curve $C' \subseteq \mathbb{A}^n$ defined over K. Then, for every proper subfield K_0 of K,* $\mathrm{card}(\phi(C(K)) \smallsetminus \mathbb{A}^n(K_0)) = \mathrm{card}(K)$.

The proof uses among others a trick of Jochen Koenigsmann that Pop applied to prove Proposition 1.4(b) below.

Proposition 1.4. *Let K be an ample field, V an absolutely irreducible variety defined over K with a K-rational simple point, and K_0 a subfield of K. Then:*

(a) $K = K_0(V(K))$.

(b) $\mathrm{card}(V(K)) = \mathrm{card}(K)$.

Proposition 1.5 (Pop [16, Prop. 1.2]). *Every algebraic extension of an ample field is ample.*

Problem 1.6. Let L/K be a finite separable extension such that L is ample. Is K ample?

2 Examples of Ample Fields

The properties causing a field K to be ample vary from diophantine, arithmetic, to Galois theoretic.

(a) PAC fields, in particular, algebraically closed fields.

(b) Henselian fields.

More generally, we say that a pair $(A, \mathfrak{a})$ consisting of a domain A and a nonzero ideal $\mathfrak{a}$ of A is **Henselian** if for each $f \in A[X]$ satisfying

$$f(0) \equiv 0 \mod \mathfrak{a} \quad \text{and} \quad f'(0) \text{ is a unit} \quad \mod \mathfrak{a}$$

there exists $x \in \mathfrak{a}$ such that $f(x) = 0$.

Pop [17, Thm. 1.1] has observed that the proof that Henselian fields are ample can be adjusted to a proof that if $(A, \mathfrak{a})$ is a Henselian pair, then $\mathrm{Quot}(A)$ is ample.

(c) If A is complete with respect to a nonzero ideal $\mathfrak{a}$, then $(A, \mathfrak{a})$ is a Henselian pair, hence $\mathrm{Quot}(A)$ is ample.

For example, $K((X_1, \dots, X_n))$ is ample for every $n \geq 1$ and any field K. So is, for example, the field $\mathrm{Quot}(\mathbb{Z}[[X_1, \dots, X_n]])$. Note that if $n \geq 2$, then $F = K((X_1, \dots, X_n))$ is Hilbertian (by Weissauer [9, Thm. 13.9.1]), hence F is not Henselian (by Geyer [9, Lemma 15.5.4]), even though the ring $K[[X_1, \dots, X_n]]$ is complete and therefore Henselian.

(d) Real closed fields.

(e) Fields satisfying a local global principle.

Let K be a field and $\mathcal{K}$ be a family of field extensions of K. We say that K is P$\mathcal{K}$C (or also that K satisfies a **local global principle** with respect to $\mathcal{K}$) if every nonempty absolutely irreducible variety defined over K with a simple $\bar{K}$-rational point for each $\bar{K} \in \mathcal{K}$ has a K-rational point. In this case, if each $\bar{K} \in \mathcal{K}$ is ample, then K is also ample.

For example, let K be a countable Hilbertian field and S a finite set of **local primes** of K. Thus, each $\mathfrak{p} \in S$ is an equivalence class of absolute values whose completion $\hat{K}_\mathfrak{p}$ is a local field. Let $K_\mathfrak{p} = K_s \cap \hat{K}_\mathfrak{p}$. Consider also an e-tuple $\sigma = (\sigma_1, \dots, \sigma_e)$ taken in random in $\mathrm{Gal}(K)^e$ (with respect to the Haar measure). Let $K_s(\sigma)$ be the fixed field in K_s of $\sigma_1, \dots, \sigma_e$ and let $K_s[\sigma]$ be the maximal Galois extension of K in $K_s(\sigma)$. Then the field

$$K_{\mathrm{tot,S}}[\sigma] = K_s[\sigma] \cap \bigcap_{\mathfrak{p} \in S} \bigcap_{\rho \in \mathrm{Gal}(K)} K_\mathfrak{p}^\rho$$

is P$\mathcal{K}$C with $\mathcal{K} = \{K_{\mathfrak{p}}^{\rho} \mid \mathfrak{p} \in S,\ \rho \in \mathrm{Gal}(K)\}$. Since each of the fields $K_{\mathfrak{p}}^{\sigma}$ is Henselian or real closed, it is ample (by (b) and (d) above). Hence, $K_{\mathrm{tot,S}}[\sigma]$ is ample (Geyer-Jarden [12, Example 5.6.6]).

(f) Fields with a pro-p absolute Galois group (Colliot-Thélène [2], Jarden [11]).

Problem 2.1. Let K be a field such that the order of $\mathrm{Gal}(K)$ is divisible by only finitely many prime numbers. Is K ample?

3 Finite Split Embedding Problems

The raison d'etre of ample fields is that they are the only known fields for which Problem C has an affirmative answer.

Theorem 3.1 (Pop [Main Thm. A], Haran-Jarden [3, Thm. A])**.** *Let K be an ample field, L a finite Galois extension of K, and x a variable. Suppose* $\mathrm{Gal}(L/K)$ *acts on a finite group H. Then $K(x)$ has a Galois extension F that contains L and there is a commutative diagram*

$$\begin{tikzcd}
 & \mathrm{Gal}(F/K(x)) \arrow[dl, "\gamma"'] \arrow[d, "\mathrm{res}"] \\
\mathrm{Gal}(L/K) \ltimes H \arrow[r, "\alpha"] & \mathrm{Gal}(L/K)
\end{tikzcd}$$

in which α is the projection on the first component and γ is an isomorphism.

The proof of Theorem 3.1 is carried out in two steps. First one solves the corresponding embedding problem over the field $\hat{K} = K((t))$ using patching. Then one reduces the solution obtained over $\hat{K}(x)$ to a solution over $K(x)$, using that K is existentially closed in $\hat{K}$.

The most striking application of Theorem 3.1 is a solution of a problem of Field Arithmetic that stayed open for a long time:

Theorem 3.2 ([12, Thm. 5.10.2])**.** *Every PAC Hilbertian field K is ω-free (that is, every finite embedding problem over K is solvable). In particular, if K is countable, then* $\mathrm{Gal}(K)$ *is isomorphic to the free profinite group $\hat{F}_{\omega}$ on countably many generators.*

For the next application we need an improvement of Theorem 3.1.

Theorem 3.3 (Harbater-Stevenson [10, Thm. 4.3], Pop [17, Thm. 1.2], Jarden [12, Prop. 8.6.3])**.** *Let K be an ample field and x a variable. Then every finite split embedding problem over $K(x)$ has as many solutions as the cardinality of K.*

In particular, this theorem applies when K is algebraically closed. Since $\mathrm{Gal}(K(x))$ is projective, we get the following generalization of a theorem that was proved in characteristic 0 using Riemann existence theorem.

Corollary 3.4 (Harbater [6, Cor. 4.2], Pop [15, Geometric (SC)], Haran-Jarden [4, Main Theorem])**.** *Let K be an algebraically closed field of cardinality m. Then* $\mathrm{Gal}(K(x)) \cong \hat{F}_m$*.*

Actually, Theorem 3.3 was proved in a stronger form, in which $K(x)$ is replaced by an arbitrary function field E of one variable over K and the solution field is regular over the field of constants of E.

Harbater-Stevenson proved that every finite split embedding problem over $K((t_1, t_2))$ has as many solutions as the cardinality of K. Moreover, this property is inherited by $K((t_1, t_2))_{\mathrm{ab}}$. Finally, the absolute Galois group of the latter field is projective. Together, this proves the following result:

Theorem 3.5. *Let K be a separably closed field and $E = K((t_1, t_2))$. Then* $\mathrm{Gal}(E_{\mathrm{ab}})$ *is isomorphic to the free profinite group $\hat{F}_m$ of cardinality $m = \mathrm{card}(E)$.*

Theorem 3.3 can be improved further.

Theorem 3.6 (Bary-Soroker, Haran, Harbater [1, Thm. 7.2]; Jarden [12, Thm. 11.7.1])**.** *Let E be a function field of one variable over an ample field K. Then* $\mathrm{Gal}(E)$ *is* **semi-free***. That is, every finite split embedding problem over E*

$$(\mathrm{res}\colon \mathrm{Gal}(E) \to \mathrm{Gal}(F/E),\ \alpha\colon G \to \mathrm{Gal}(F/E))$$

has $\mathrm{card}(E)$**-linearly disjoint solution fields** F_i *(i.e. the fields F_i are linearly disjoint extensions of F.)*

Combining this proposition with results of Bary-Soroker-Haran-Harbater, Efrat, and Pop, we were able to prove the following result.

Theorem 3.7 (Jarden [12, Thm. 11.7.6])**.** *Let K be a PAC field of cardinality c and x a variable. For each irreducible polynomial $p \in K[x]$ and every positive integer n satisfying* $\mathrm{char}(K) \nmid n$ *let $\sqrt[n]{p}$ be an nth root of p such that $(\sqrt[mn]{p})^m = \sqrt[n]{p}$ for all m, n. Let F be the compositum of all fields $K(\sqrt[n]{p})$. Then, F is Hilbertian and* $\mathrm{Gal}(F) \cong \hat{F}_c$*.*

Here is a special case:

Corollary 3.8 (Jarden [12, Example 11.7.8])**.** *Let K be an PAC field of cardinality m and x a variable. Suppose K contains all root of unity. Then* $\mathrm{Gal}(K(x)_{\mathrm{ab}}) \cong \hat{F}_m$*.*

And here is another example of a semi-free absolute Galois group:

Theorem 3.9 (Pop [12, Thm. 12.4.4])**.** *Each of the following fields K is Hilbertian and ample. Moreover,* $\mathrm{Gal}(K)$ *is semi-free of rank* $\mathrm{card}(K)$*.*

(a) $K = K_0((X_1, \dots, X_n))$*, where* K_0 *is an arbitrary field and* $n \geq 2$*.*

(b) $K = \mathrm{Quot}(R_0[[X_1, \dots, X_n]])$*, where* R_0 *is a Noetherian domain which is not a field and* $n \geq 1$*.*

More about Theorem 3.9 can be found in Chapter 12 of [12].

Problem 3.10. Give an example of a non-ample field K such that every finite split embedding over $K(x)$ is solvable.

Note that the existence of an example as in Problem 3.10 will give a negative answer to Problem C. Conversely, a positive answer to Problem C is a negative answer to Problem 3.10.

4 Axioms for Algebraic Patching

Let E be a field, G a finite group, and $(G_i)_{i \in I}$ a finite family of subgroups of G that generates G. Suppose that for each $i \in I$ we have a finite Galois extension F_i of E with Galois group G_i. We use these extensions to construct a Galois extension F of E (not necessarily containing F_i) with Galois group G. First we 'lift' each F_i/E to a Galois field extension Q_i/P_i, where P_i is an appropriate field extension of E (that we refer to as "analytic") such that $P_i \cap F_i = E$ and all of the Q_i's are contained in a common field Q. Then we define F to be the maximal subfield contained in $\bigcap_{i \in I} Q_i$ on which the Galois actions of $\mathrm{Gal}(Q_i/P_i)$ combine to an action of G.

$$\begin{tikzcd}
P_i \arrow[r, dash, "G_i"] \arrow[d, dash] & Q_i \arrow[r, dash] \arrow[d, dash] & Q \\
E \arrow[r, dash, "G_i"] & F_i &
\end{tikzcd}$$

The construction works if certain patching conditions on the initial data are satisfied.

Definition 4.1 (Patching data)**.** Let I be a finite set with $|I| \geq 2$. **Patching data**

$$\mathcal{E} = (E, F_i, P_i, Q; G_i, G)_{i \in I}$$

consists of fields $E \subseteq F_i, P_i \subseteq Q$ and finite groups $G_i \leq G$, $i \in I$, such that the following conditions hold.

(4.1a) F_i/E is a Galois extension with Galois group G_i, $i \in I$.

(4.1b) $F_i \subseteq P_i'$, where $P_i' = \bigcap_{j \neq i} P_j$, $i \in I$.

(4.1c) $\bigcap_{i\in I} P_i = E$.

(4.1d) $G = \langle G_i \mid i \in I\rangle$.

(4.1e) (Cartan's decomposition) Let $n = |G|$. Then for every $B \in \mathrm{GL}_n(Q)$ and each $i \in I$ there exist $B_1 \in \mathrm{GL}_n(P_i)$ and $B_2 \in \mathrm{GL}_n(P_i')$ such that $B = B_1B_2$.

We extend $\mathcal{E}$ by more fields. For each $i \in I$ let $Q_i = P_iF_i$ be the compositum of P_i and F_i in Q. Conditions (4.1b) and (4.1c) imply that $P_i \cap F_i = E$. Hence, Q_i/P_i is a Galois extension with Galois group isomorphic (via restriction of automorphisms) to $G_i = \mathrm{Gal}(F_i/E)$. We identify $\mathrm{Gal}(Q_i/P_i)$ with G_i via this isomorphism.

Definition 4.2 (Compound)**.** The **compound** of the patching data $\mathcal{E}$ is the set F of all $a \in \bigcap_{i\in I} Q_i$ for which there exists a function $f\colon G \to \bigcap_{i\in I} Q_i$ such that

(4.2a) $a = f(1)$ and

(4.2b) $f(\zeta\tau) = f(\zeta)^\tau$ for every $\zeta \in G$ and $\tau \in \bigcup_{i\in I} G_i$.

Note that f is already determined by $f(1)$. Indeed, by (4.1d), each $\tau \in G$ can be written as $\tau = \tau_1\tau_2\cdots\tau_r$ with $\tau_1,\dots,\tau_r \in \bigcup_{i\in I} G_i$. Hence, by (4.2b), $f(\tau) = f(1)^{\tau_1\cdots\tau_r}$.

We call f the **expansion** of a and denote it by f_a. Thus, $f_a(1) = a$ and $f_a(\zeta\tau) = f_a(\zeta)^\tau$ for all $\zeta \in G$ and $\tau \in \bigcup_{i\in I} G_i$.

We list some elementary properties of expansions:

Lemma 4.3. *Let F be the compound of $\mathcal{E}$. Then:*

(a) Every $a \in E$ has an expansion, namely the constant function $\zeta \mapsto a$.

(b) Let $a, b \in F$. Then $a+b, ab \in F$; in fact, $f_{a+b} = f_a + f_b$ and $f_{ab} = f_af_b$.

(c) If $a \in F^\times$, then $a^{-1} \in F$. More precisely: $f_a(\zeta) \neq 0$ for all $\zeta \in G$, and $\zeta \mapsto f_a(\zeta)^{-1}$ is the expansion of a^{-1}.

(d) Let $a \in F$ and $\sigma \in G$. Then $f_a(\sigma) \in F$; in fact, $f_{f_a(\sigma)}(\zeta) = f_a(\sigma\zeta)$.

Proof. Statement (a) holds, because $a^\tau = a$ for each $\tau \in \bigcup_{i\in I} G_i$. Next observe that the sum and the product of two expansions is again an expansion. Hence, Statement (b) follows from the uniqueness of expansions and from the observations $(f_{a+b})(1) = a + b = f_a(1) + f_b(1) = (f_a + f_b)(1)$ and $f_{ab}(1) = (f_af_b)(1)$.

Next we consider a nonzero $a \in F$ and let $\tau \in G$. Using the notation of Definition 4.2, we have $f_a(\tau) = \big((a^{\tau_1})^{\tau_2\cdots}\big)^{\tau_r} \neq 0$. Since taking the inverse in $\bigcap_{i\in I} Q_i$ commutes with the action of G, the map $\zeta \mapsto f_a(\zeta)^{-1}$ is the expansion of a^{-1}. This proves (c).

Finally, one can check that the map $\zeta \to f_a(\sigma\zeta)$ has the value $f_a(\sigma)$ at $\zeta = 1$ and it satisfies (4.2b). Hence, that map is an expansion of $f_a(\sigma)$, as claimed in (d). □

Definition 4.4 (G-action on F)**.** For $a \in F$ and $\sigma \in G$ put

$$a^\sigma = f_a(\sigma), \tag{4.3}$$

where f_a is the expansion of a.

Lemma 4.5. *The compound F of the patching data $\mathcal{E}$ is a field on which G acts by* (4.3) *such that $F^G = E$. Moreover, for each $i \in I$, the restriction of this action to G_i coincides with the action of $G_i = \mathrm{Gal}(Q_i/P_i)$ on F as a subset of Q_i.*

Proof. By Lemma 4.3(a),(b),(c), F is a field containing E. Furthermore, (4.3) defines an action of G on F. Indeed, $a^1 = f_a(1) = a$. Moreover, if ζ is another element of G, then by (4.3) and Lemma 4.3(d), $(a^\sigma)^\zeta = f_a(\sigma)^\zeta = f_{f_a(\sigma)}(\zeta) = f_a(\sigma\zeta) = a^{(\sigma\zeta)}$.

Claim: $F^G = E$. Indeed, by Lemma 4.3(a), elements of E have constant expansions, hence are fixed by G. Conversely, let $a \in F^G$. Then for each $i \in I$ we have $a \in Q_i^{G_i} = P_i$. Hence, by (4.1c), $a \in E$.

Finally, let $\tau \in G_i$ and $a \in F$. Then, $f_a(\tau) = f_a(1)^\tau = a^\tau$, where τ acts as an element of $G_i = \mathrm{Gal}(Q_i/P_i)$. Thus, that action coincides with the action given by (4.3). □

The next goal is to prove that F/E is a Galois extension with Galois group G. To achieve this goal we introduce more objects and invoke Cartan's decomposition. Let

$$N = \Big\{ \sum_{\zeta \in G} a_\zeta \zeta \mid a_\zeta \in Q \Big\} \tag{4.4}$$

be the vector space over Q with basis $(\zeta \mid \zeta \in G)$, where G is given some fixed ordering. Thus, $\dim_Q N = |G|$. For each $i \in I$ we consider the following subset of N:

$$N_i = \Big\{ \sum_{\zeta \in G} a_\zeta \zeta \in N \mid a_\zeta \in Q_i,\ a_\zeta^\eta = a_{\zeta\eta} \text{ for all } \zeta \in G,\ \eta \in G_i \Big\}. \tag{4.5}$$

It is a vector space over P_i.

$$\begin{tikzcd}
N_i \arrow[rr, dash] \arrow[d, dash] & & N \arrow[d, dash] \\
P_i \arrow[r, dash, "G_i"] \arrow[dd, dash] & Q_i \arrow[r, dash] \arrow[d, dash] \arrow[dd, dash, bend left] & Q \\
& F \arrow[dl, dash] & \\
E \arrow[r, dash, "G_i"] & F_i &
\end{tikzcd}$$

Lemma 4.6. *For each $i \in I$ the Q-vector space N has a basis which is contained in N_i.*

Proof. Let $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$ be a system of representatives of G/G_i and let $\eta_1, \ldots, \eta_r$ be a listing of the elements of G_i. Thus, $G = \{\lambda_k \eta_\nu \mid k = 1, \ldots, m;\ \nu = 1, \ldots, r\}$. Let z be a primitive element for Q_i/P_i. The following sequence of $|G|$ elements of N_i

$$\Big(\sum_{\nu=1}^{r} (z^{j-1})^{\eta_\nu}\ \lambda_k \eta_\nu \mid\ j = 1, \ldots, r;\ k = 1, \ldots, m\Big)$$

(in some order) is linearly independent over Q, hence it forms a basis of N over Q.

Indeed, let $a_{jk} \in Q$ such that $\sum_{j=1}^{r} \sum_{k=1}^{m} a_{jk} \big(\sum_{\nu=1}^{r} (z^{j-1})^{\eta_\nu}\ \lambda_k \eta_\nu \big) = 0$. Then

$$\sum_{k=1}^{m} \sum_{\nu=1}^{r} \Big(\sum_{j=1}^{r} a_{jk} (z^{j-1})^{\eta_\nu}\ \Big) \lambda_k \eta_\nu = 0.$$

This gives $\sum_{j=1}^{r} a_{jk} (z^{j-1})^{\eta_\nu} = 0$ for all k, ν. Thus, for each k, $(a_{1k}, \ldots, a_{rk})$ is a solution of the homogeneous system of equations with the Vandermonde matrix $\big((z^{j-1})^{\eta_\nu}\big)$. Since this matrix is invertible, $a_{jk} = 0$ for all j, k. □

Lemma 4.7 (Common basis lemma). *N has a Q-basis in $\bigcap_{i \in I} N_i$.*

Proof. Consider a nonempty subset J of I. Using induction on $|J|$, we find a Q-basis in $\bigcap_{j \in J} N_j$. For $J = I$ this gives the assertion of the lemma.

For each $i \in I$, Lemma 4.6 gives a Q-basis $\mathbf{v}_i$ of N in N_i, so the result follows when $|J| = 1$. Assume $|J| \geq 2$ and fix $i \in J$. By induction, N has a Q-basis $\mathbf{u}$ in $\bigcap_{j \in J \smallsetminus \{i\}} N_j$. The transition matrix $B \in \mathrm{GL}_n(Q)$ between $\mathbf{v}_i$ and $\mathbf{u}$ satisfies

$$\mathbf{u} = \mathbf{v}_i B. \tag{4.6}$$

By (4.1e), there exist $B_1 \in \mathrm{GL}_n(P_i)$ and $B_2 \in \mathrm{GL}_n(P_i') \subseteq \bigcap_{j \in J \smallsetminus \{i\}} \mathrm{GL}_n(P_j)$. such that $B = B_1 B_2$. Then $\mathbf{u} B_2^{-1} = \mathbf{v}_i B_1$ is a Q-basis of N in $\bigcap_{j \in J} N_j$. This finishes the induction. □

Lemma 4.8. *Let G be a finite group that acts on a field F and set $E = F^G$. If $[F : E] \geq |G|$, then F/E is a Galois extension whose Galois group is G.*

Proof. let $\bar{G}$ be the quotient of G by the kernel of the action of G on F. Then $\bar{G}$ is a finite group of automorphisms of F with fixed field E. By a lemma of Artin [14, Algebra, Lemma VI.1.8], F/E is a Galois extension with $\mathrm{Gal}(F/E) = \bar{G}$. By assumption, $|G| \geq |\bar{G}| = |\mathrm{Gal}(F/E)| = [F : E] \geq |G|$. Hence, $G = \bar{G} = \mathrm{Gal}(F/E)$. □

Now we are in a position to improve Lemma 4.5.

Proposition 4.9. *The compound F of the patching data $\mathcal{E}$ is a Galois extension of E with Galois group G acting by* (4.3). *Moreover, $Q_i = P_iF$ for each $i \in I$.*

Proof. We define a map $T\colon F \to N$ by

$$T(a) = \sum_{\zeta\in G} f_a(\zeta)\zeta.$$

By Lemma 4.3(a),(b), T is an E-linear map. By (4.2b), $f_a(\zeta)^\tau = f_a(\zeta\tau)$ for all $\zeta \in G$ and $\tau \in \bigcup_{i\in I} G_i$, so $\mathrm{Im}(T) \subseteq \bigcap_{i\in I} N_i$. Conversely, if $\xi = \sum_{\zeta\in G} a_\zeta\zeta \in \bigcap_{i\in I} N_i$, then $a_1 \in F$ and $T(a_1) = \xi$. Therefore, $\mathrm{Im}(T) = \bigcap_{i\in I} N_i$. By Lemma 4.7, $\mathrm{Im}(T)$ contains $|G|$ linearly independent elements over Q, hence over E. Therefore, $[F : E] = \dim_E F \geq \dim_E \mathrm{Im}(T) \geq |G|$. By Lemma 4.5, $F = E^G$. Hence, by Lemma 4.8, F/E is a Galois extension and $\mathrm{Gal}(F/E) = G$.

Finally, by what we have just proved and by Lemma 4.5, the restriction

$$\mathrm{Gal}(Q_i/P_i) \to \mathrm{Gal}(F/E)$$

is injective. Hence, $Q_i = P_iF$. □

5 Galois Action on Patching Data

Knowledge of the finite groups that can be realized over a field K does not determine $\mathrm{Gal}(K)$. The latter required control on the finite embedding problems that can be solved over K. Unfortunately, our methods can handle only “finite split embedding problems”. However, in some cases (like those that appear in our main results), being able to solve all finite split embedding problem suffices.

A **finite split embedding problem** over a field E_0 is an epimorphism

$$\mathrm{pr}\colon \Gamma \ltimes G \to \Gamma \tag{5.1}$$

of finite groups, where $\Gamma = \mathrm{Gal}(E/E_0)$ is the Galois group of a Galois extension E/E_0, G is a finite group on which Γ acts from the right, $\Gamma \ltimes G$ is the corresponding semidirect product, and pr is the projection on Γ. Each element of $\Gamma \ltimes G$ has a unique presentation as a product $\gamma\zeta$ with $\gamma \in \Gamma$ and $\zeta \in G$. The product and the inverse operation are given in $\Gamma \ltimes G$ by the formulas $\gamma\zeta \cdot \delta\eta = \gamma\delta \cdot \zeta^\delta\eta$ and $(\gamma\zeta)^{-1} = \gamma^{-1}(\zeta^{\gamma^{-1}})^{-1}$. A **solution** of (5.1) is a Galois extension F of E_0 that contains E and an isomorphism $\psi\colon \mathrm{Gal}(F/E_0) \to \Gamma \ltimes G$ such that $\mathrm{pr} \circ \psi = \mathrm{res}_E$. We call F a **solution field** of (5.1).

Suppose the compound F of a patching data $\mathcal{E}$ (§4) realizes G over E. A ‘proper’ action of Γ on $\mathcal{E}$ will then ensure that F is even a solution field for the embedding problem (5.1).

Definition 5.1. Let E/E_0 be a finite Galois extension with Galois group Γ. Let

$$\mathcal{E} = (E, F_i, P_i, Q; G_i, G)_{i \in I}$$

be patching data (Definition 4.1). A **proper action** of Γ on $\mathcal{E}$ is a triple that consists of an action of Γ on the group G, an action of Γ on the field Q, and an action of Γ on the set I such that the following conditions hold:

(5.2a) The action of Γ on Q extends the action of Γ on E.

(5.2b) $F_i^\gamma = F_{i^\gamma}$, $P_i^\gamma = P_{i^\gamma}$, and $G_i^\gamma = G_{i^\gamma}$, for all $i \in I$ and $\gamma \in \Gamma$.

(5.2c) $(a^\tau)^\gamma = (a^\gamma)^{\tau^\gamma}$ for all $i \in I$, $a \in F_i$, $\tau \in G_i$, and $\gamma \in \Gamma$.

The action of Γ on G defines a semidirect product $\Gamma \ltimes G$ such that $\tau^\gamma = \gamma^{-1}\tau\gamma$ for all $\tau \in G$ and $\gamma \in \Gamma$. Let $\mathrm{pr}\colon \Gamma \ltimes G \to \Gamma$ be the canonical projection.

Proposition 5.2. *In the notation of Definition 5.1 suppose that* $\Gamma = \mathrm{Gal}(E/E_0)$ *acts properly on the patching data $\mathcal{E}$ given in Definition 5.1. Let F be the compound of $\mathcal{E}$. Then Γ acts on F via the restriction from its action on Q and the actions of Γ and G on F combine to an action of $\Gamma \ltimes G$ on F with fixed field E_0. This gives an identification* $\mathrm{Gal}(F/E_0) = \Gamma \ltimes G$ *such that the following diagram of short exact sequences commutes:*

$$\begin{tikzcd}
1 \arrow[r] & G \arrow[r] \arrow[d, equal] & \Gamma \ltimes G \arrow[r, "\mathrm{pr}"] \arrow[d, equal] & \Gamma \arrow[r] \arrow[d, equal] & 1 \\
1 \arrow[r] & \mathrm{Gal}(F/E) \arrow[r] & \mathrm{Gal}(F/E_0) \arrow[r, "\mathrm{res}"] & \mathrm{Gal}(E/E_0) \arrow[r] & 1
\end{tikzcd}$$

Thus, F is a solution field of the embedding problem (5.1)*.*

Proof. We break the proof of the proposition into three parts.

- Part A: *The action of Γ on F.*

  Let $i \in I$ and $\gamma \in \Gamma$. Then $Q_i = P_iF_i$, so by (5.2b), $Q_i^\gamma = Q_{i^\gamma}$. Moreover, we have identified $\mathrm{Gal}(Q_i/P_i)$ with $G_i = \mathrm{Gal}(F_i/E)$ via restriction. Hence, by (5.2b), for all $a \in P_i$ and $\tau \in G_i$ we have $\tau^\gamma \in G_{i^\gamma}$ and $a^\gamma \in P_{i^\gamma}$, so $(a^\tau)^\gamma = a^\gamma = (a^\gamma)^{\tau^\gamma}$. Together with (5.2c), this gives

  $$(5.3) \qquad (a^\tau)^\gamma = (a^\gamma)^{\tau^\gamma} \quad \text{for all } a \in Q_i \text{ and } \tau \in G_i.$$

  Consider an $a \in F$ and let f_a be the expansion of a (Definition 4.1). Define $f_a^\gamma\colon G \to \bigcap_{i\in I} Q_i$ by $f_a^\gamma(\zeta) = f_a(\zeta^{\gamma^{-1}})^\gamma$. Then f_a^γ is the expansion f_{a^γ} of a^γ. Indeed, $f_a^\gamma(1) = f_a(1^{\gamma^{-1}})^\gamma = a^\gamma$ and if $\zeta \in G$ and $\tau \in G_i$, then $\tau^{\gamma^{-1}} \in G_{i^{\gamma^{-1}}}$.

Hence, by (5.3) with $i^{\gamma^{-1}}, f_a(\zeta^{\gamma^{-1}}), \tau^{\gamma^{-1}}$ replacing i, a, τ, respectively, we have

$$f_a^\gamma(\zeta\tau) = f_a(\zeta^{\gamma^{-1}}\tau^{\gamma^{-1}})^\gamma = \big(f_a(\zeta^{\gamma^{-1}})^{\tau^{\gamma^{-1}}}\big)^\gamma = \big(f_a(\zeta^{\gamma^{-1}})^\gamma\big)^{\tau^{\gamma^{-1}\gamma}} = f_a^\gamma(\zeta)^\tau.$$

Thus $a^\gamma \in F$. It follows that the action of Γ on Q restricts to an action of Γ on F.

- Part B: *The action of* $\Gamma \ltimes G$ *on* F. Let $a \in F$ and $\gamma \in \Gamma$. We claim that

$$(5.4) \qquad (a^\sigma)^\gamma = (a^\gamma)^{\sigma^\gamma} \qquad \text{for all } \sigma \in G,$$

where $a^\sigma = f_a(\sigma)$ (Definition 4.4). Indeed, write σ as a word in $\bigcup_{i\in I} G_i$. Then (5.4) follows from (5.3) by induction on the length of the word. If $\sigma = 1$, then (5.4) is an identity. Suppose (5.4) holds for some $\sigma \in G$ and let $\tau \in \bigcup_{i\in I} G_i$. Using the identification of the action of each $\tau \in G_i$ on F as an element of G_i with its action as an element of G (Lemma 4.5(a)) and (5.3) for a^σ rather than a, we have

$$(a^{\sigma\tau})^\gamma = \big((a^\sigma)^\tau\big)^\gamma = \big((a^\sigma)^\gamma\big)^{\tau^\gamma} = \big((a^\gamma)^{\sigma^\gamma}\big)^{\tau^\gamma} = (a^\gamma)^{\sigma^\gamma\tau^\gamma} = (a^\gamma)^{(\sigma\tau)^\gamma}.$$

Now we apply (5.4) to $a^{\gamma^{-1}}$ instead of a, to find that $\Big(\big(a^{\gamma^{-1}}\big)^\sigma\Big)^\gamma = a^{\sigma^\gamma}$. It follows that the actions of Γ and G on F combine to an action of $\Gamma \ltimes G$ on F.

(5.5)

$$\begin{tikzcd}
& P_i \arrow[r, dash] \arrow[dd, dash] & Q_i \arrow[r, dash] \arrow[d, dash] \arrow[dd, dash, bend left] & Q \arrow[dd, dash] \\
& & F \arrow[dl, dash] & \\
E_0 \arrow[r, dash] & E \arrow[r, dash] & F_i \arrow[r, dash] & P'_i
\end{tikzcd}$$

- Part C: *Conclusion of the proof.* Since $F^G = E$ (Lemma 4.5) and $E^\Gamma = E_0$, we have $F^{\Gamma\ltimes G} = E_0$. Furthermore, $[F : E_0] = [F : E] \cdot [E : E_0] = |G| \cdot |\Gamma| = |\Gamma \ltimes G|$. By Galois theory, $\mathrm{Gal}(F/E_0) = \Gamma \ltimes G$ and the map $\mathrm{res}\colon \mathrm{Gal}(F/E_0) \to \mathrm{Gal}(E/E_0)$ coincides with the canonical map $\mathrm{pr}\colon \Gamma \ltimes G \to \Gamma$.

□

6 Normed Rings

In Section 15 we construct patching data over fields $K(x)$, where K is a complete ultrametric valued field. The 'analytic' fields P_i will be the quotient fields of certain rings of convergent power series in several variables over K. At a certain point in a proof by induction we consider a ring of convergent power series in one variable over a complete ultrametric valued ring. So, we start by recalling the definition and properties of the latter rings.

Let A be a commutative ring with 1. An **ultrametric absolute value** of A is a function $|\ |\colon A \to \mathbb{R}$ satisfying the following conditions:

(6.1a) $|a| \geq 0$, and $|a| = 0$ if and only if $a = 0$.

(6.1b) There exists $a \in A$ such that $0 < |a| < 1$.

(6.1c) $|ab| = |a| \cdot |b|$.

(6.1d) $|a + b| \leq \max(|a|, |b|)$.

By (6.1a) and (6.1c), A is an integral domain. By (6.1c), the absolute value of A extends to an absolute value on the quotient field of A (by $|\frac{a}{b}| = \frac{|a|}{|b|}$). It follows also that $|1| = 1$, $|-a| = |a|$, and

(6.1d′) if $|a| < |b|$, then $|a + b| = |b|$.

Denote the ordered additive group of the real numbers by $\mathbb{R}^+$. The function $v\colon \mathrm{Quot}(A) \to \mathbb{R}^+ \cup \{\infty\}$ defined by $v(a) = -\log |a|$ satisfies the following conditions:

(6.2a) $v(a) = \infty$ if and only if $a = 0$.

(6.2b) There exists $a \in \mathrm{Quot}(A)$ such that $0 < v(a) < \infty$.

(6.2c) $v(ab) = v(a) + v(b)$.

(6.2d) $v(a + b) \geq \min\{v(a), v(b)\}$ (and $v(a + b) = v(b)$ if $v(b) < v(a)$).

In other words, v is a **real valuation** of $\mathrm{Quot}(A)$. Conversely, every real valuation $v\colon \mathrm{Quot}(A) \to \mathbb{R}^+ \cup \{\infty\}$ gives rise to a nontrivial ultrametric absolute value $|\cdot|$ of $\mathrm{Quot}(A)$: $|a| = \varepsilon^{v(a)}$, where ε is a fixed real number between 0 and 1.

An attempt to extend an absolute value from A to a larger ring A' may result in relaxing Condition (6.1c), replacing the equality by an inequality. This leads to the more general notion of a 'norm'.

Definition 6.1 (Normed rings)**.** Let R be an associative (and not necessarily commutative) ring with 1. An (ultrametric) **norm** on R is a function $\|\ \|\colon R \to \mathbb{R}$ that satisfies the following conditions for all $a, b \in R$:

(6.3a) $\|a\| \geq 0$, and $\|a\| = 0$ if and only if $a = 0$; further $\|1\| = \|-1\| = 1$.

(6.3b) There is an $x \in R$ with $0 < \|x\| < 1$.

(6.3c) $\|ab\| \leq \|a\| \cdot \|b\|$.

(6.3d) $\|a+b\| \leq \max(\|a\|, \|b\|)$.

The norm $\| \ \|$ naturally defines a topology on R whose basis is the collection of all sets $U(a_0, r) = \{a \in R \mid \ \|a - a_0\| < r\}$ with $a_0 \in R$ and $r > 0$. Both addition and multiplication are continuous under that topology. Thus, R is a **topological ring**.

In the sequel, whenever we speak about a "normed ring", we mean an "associative normed ring with 1".

Definition 6.2 (Complete rings)**.** Let R be a normed ring. A sequence $a_1, a_2, \ldots$ of elements of R is **Cauchy** if for each $\varepsilon > 0$ there exists m_0 such that $\|a_n - a_m\| < \varepsilon$ for all $m, n \geq m_0$. We say that R is **complete** if every Cauchy sequence converges.

Lemma 6.3. *Let R be a normed ring and let $a, b \in R$. Then:*

(a) $\|-a\| = \|a\|$.

(b) If $\|a\| < \|b\|$, then $\|a+b\| = \|b\|$.

(c) A sequence $a_1, a_2, a_3, \ldots$ of elements of R is Cauchy if for each $\varepsilon > 0$ there exists m_0 such that $\|a_{m+1} - a_m\| < \varepsilon$ for all $m \geq m_0$.

(d) The map $x \mapsto \|x\|$ from R to $\mathbb{R}$ is continuous.

(e) If R is complete, then a series $\sum_{n=0}^{\infty} a_n$ of elements of R converges if and only if $a_n \to 0$.

(f) If R is complete and $\|a\| < 1$, then $1 - a \in R^\times$. Moreover, $(1-a)^{-1} = 1 + b$ with $\|b\| < 1$.

Proof.

- *Proof of (a):* Observe that $\|-a\| \leq \|-1\| \cdot \|a\| = \|a\|$. Replacing a by $-a$, we get $\|a\| \leq \|-a\|$, hence the claimed equality.

- *Proof of (b):* Assume $\|a+b\| < \|b\|$. Then, by (a), $\|b\| = \|(-a) + (a+b)\| \leq \max(\|-a\|, \|a+b\|) < \|b\|$, which is a contradiction.

- *Proof of (c):* With m_0 as above let $n > m \geq m_0$. Then

$$\|a_n - a_m\| \leq \max(\|a_n - a_{n-1}\|, \ldots, \|a_{m+1} - a_m\|) < \varepsilon.$$

- *Proof of (d):* By (6.3d), $\|x\| = \|(x-y)+y\| \le \max(\|x-y\|, \|y\|) \le \|x-y\| + \|y\|$. Hence, $\|x\| - \|y\| \le \|x-y\|$. Symmetrically, $\|y\| - \|x\| \le \|y-x\| = \|x-y\|$. Therefore, $|\,\|x\| - \|y\|\,| \le \|x-y\|$. Consequently, the map $x \mapsto \|x\|$ is continuous.

- *Proof of (e):* Let $s_n = \sum_{i=0}^{n} a_i$. Then $s_{n+1} - s_n = a_{n+1}$. Thus, by (c), $s_1, s_2, s_3, \dots$ is a Cauchy sequence if and only if $a_n \to 0$. Hence, the series $\sum_{n=0}^{\infty} a_n$ converges if and only if $a_n \to 0$.

- *Proof of (f):* The elements a^i tend to 0 as i approaches ∞. Hence, by (e), $\sum_{i=0}^{\infty} a^i$ converges. The identities $(1-a)\sum_{i=0}^{n} a^i = 1-a^{n+1}$ and $\sum_{i=0}^{n} a^i(1-a) = 1-a^{n+1}$ imply that $\sum_{i=0}^{\infty} a^i$ is both the right and the left inverse of $1-a$. Moreover, $\sum_{i=0}^{\infty} a^i = 1+b$ with $b = \sum_{i=1}^{\infty} a^i$ and $\|b\| \le \max_{i \ge 1} \|a\|^i < 1$.

□

Example 6.4. (a) Every field K with an ultrametric absolute value is a normed ring. For example, for each prime number p, $\mathbb{Q}$ has a p-adic absolute value $|\cdot|_p$ which is defined by $|x|_p = p^{-m}$ if $x = \frac{a}{b}p^m$ with $a, b, m \in \mathbb{Z}$ and $p \nmid a, b$.

(b) The ring $\mathbb{Z}_p$ of p-adic integers and the field $\mathbb{Q}_p$ of p-adic numbers are complete with respect to the p-adic absolute value.

(c) Let K_0 be a field and let $0 < \varepsilon < 1$. The ring $K_0[[t]]$ (resp. field $K_0((t))$) of formal power series $\sum_{i=0}^{\infty} a_i t^i$ (resp. $\sum_{i=m}^{\infty} a_i t^i$ with $m \in \mathbb{Z}$) with coefficients in K_0 is complete with respect to the absolute value $|\sum_{i=m}^{\infty} a_i t^i| = \varepsilon^{\min(i \mid a_i \ne 0)}$.

(d) Let $\|\cdot\|$ be a norm of a commutative ring A. For each positive integer n we extend the norm to the associative (but usually not commutative) ring $\mathrm{M}_n(A)$ of all $n \times n$ matrices with entries in A by

$$\|(a_{ij})_{1 \le i,j \le n}\| = \max(\|a_{ij}\|_{1 \le i,j \le n}).$$

If $b = (b_{jk})_{1 \le j,k \le n}$ is another matrix and $c = ab$, then $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ and $\|c_{ik}\| \le \max(\|a_{ij}\| \cdot \|b_{jk}\|) \le \|a\| \cdot \|b\|$. Hence, $\|c\| \le \|a\|\|b\|$. This verifies Condition (6.3c). The verification of (6.3a), (6.3b), and (6.3d) is straightforward. Note that when $n \ge 2$, even if the initial norm of A is an absolute value, the extended norm satisfies only the weak condition (6.3c) and not the stronger condition (6.1c), so it is not an absolute value.

If A is complete, then so is $\mathrm{M}_n(A)$. Indeed, let $a_i = (a_{i,rs})_{1 \le r,s \le n}$ be a Cauchy sequence in $\mathrm{M}_n(A)$. Since $\|a_{i,rs} - a_{j,rs}\| \le \|a_i - a_j\|$, each of the sequences $a_{1,rs}, a_{2,rs}, a_{3,rs}, \dots$ is Cauchy, hence converges to an element b_{rs} of A. Set $b = (b_{rs})_{1 \le r,s \le n}$. Then $a_i \to b$. Consequently, $\mathrm{M}_n(A)$ is complete.

Like absolute valued rings, every normed ring has a completion:

Lemma 6.5. *Every normed ring $(R, \| \, \|)$ can be embedded into a complete normed ring $(\hat{R}, \| \, \|)$ such that R is dense in $\hat{R}$ and the following universal condition holds:*

(I) Each continuous homomorphism f of R into a complete ring S uniquely extends to a continuous homomorphism $\hat{f} \colon \hat{R} \to S$.

The normed ring $(\hat{R}, \| \, \|)$ is called the **completion** *of $(R, \| \, \|)$.*

Proof. We consider the set A of all Cauchy sequences $\mathbf{a} = (a_n)_{n=1}^\infty$ with $a_n \in R$. For each $\mathbf{a} \in A$, the values $\|a_n\|$ of its components are bounded. Hence, A is closed under componentwise addition and multiplication and contains all constant sequences. Thus, A is a ring. Let $\mathfrak{n}$ be the ideal of all sequences that converge to 0. We set $\hat{R} = A/\mathfrak{n}$ and identify each $x \in R$ with the coset $(x)_{n=1}^\infty + \mathfrak{n}$.

If $\mathbf{a} \in A \smallsetminus \mathfrak{n}$, then $\|a_n\|$ eventually becomes constant. Indeed, there exists $\beta > 0$ such that $\|a_n\| \geq \beta$ for all sufficiently large n. Choose n_0 large such that $\|a_n - a_m\| < \beta$ for all $n, m \geq n_0$. Then, $\|a_n - a_{n_0}\| < \beta \leq \|a_{n_0}\|$, so $\|a_n\| = \|(a_n - a_{n_0}) + a_{n_0}\| = \|a_{n_0}\|$. We define $\|\mathbf{a}\|$ to be the eventual absolute value of a_n and note that $\|\mathbf{a}\| \neq 0$. If $\mathbf{b} \in \mathfrak{n}$, we set $\|\mathbf{b}\| = 0$ and observe that $\|\mathbf{a} + \mathbf{b}\| = \|\mathbf{a}\|$. It follows that $\|\mathbf{a} + \mathfrak{n}\| = \|\mathbf{a}\|$ is a well defined function on $\hat{R}$ which extends the norm of R.

One checks that $\| \, \|$ is a norm on $\hat{R}$ and that R is dense in $\hat{R}$. Indeed, if $\mathbf{a} = (a_n)_{n=1}^\infty \in A$, then $a_n + \mathfrak{n} \to \mathbf{a} + \mathfrak{n}$. To prove that $\hat{R}$ is complete under $\| \, \|$ we consider a Cauchy sequence $(a_k)_{k=1}^\infty$ of elements of $\hat{R}$. For each k we choose an element $b_k \in R$ such that $\|b_k - a_k\| < \frac{1}{k}$. Then $(b_k)_{k=1}^\infty$ is a Cauchy sequence of R and the sequence $(\mathbf{a}_k)_{k=1}^\infty$ converges to the element $(b_k)_{k=1}^\infty + \mathfrak{n}$ of $\hat{R}$.

Finally, let S be a complete normed ring and $f \colon R \to S$ a continuous homomorphism. Then, for each $\mathbf{a} = (a_n)_{n=1}^\infty \in A$, the sequence $(f(a_n))_{n=1}^\infty$ of S is Cauchy, hence it converges to an element s. Define $\hat{f}(\mathbf{a} + \mathfrak{n}) = s$ and check that $\hat{f}$ has the desired properties. □

7 Rings of Convergent Power Series

Let A be a complete normed commutative ring and x a variable. Consider the following subset of $A[[x]]$:

$$A\{x\} = \Big\{ \sum_{n=0}^\infty a_n x^n \mid a_n \in A, \lim_{n \to \infty} \|a_n\| = 0 \Big\}.$$

For each $f = \sum_{n=0}^\infty a_n x^n \in A\{x\}$ we define $\|f\| = \max(\|a_n\|)_{n=0,1,2,\ldots}$. This definition makes sense because $a_n \to 0$, hence $\|a_n\|$ is bounded.

We prove the Weierstrass division theorem and the Weierstrass preparation theorem for $A\{x\}$ in analogy to the corresponding theorems for the ring of formal power series in one variable over a local ring.

Lemma 7.1. *(a) $A\{x\}$ is a subring of $A[[x]]$ containing A.*

(b) The function $\| \ \|\colon A\{x\} \to \mathbb{R}$ is a norm.

(c) The ring $A\{x\}$ is complete under that norm.

(d) Let B be a complete normed ring extension of A. Then each $b \in B$ with $\|b\| \le 1$ defines an **evaluation homomorphism** *$A\{x\} \to B$ given by*

$$f = \sum_{n=0}^{\infty} a_n x^n \mapsto f(b) = \sum_{n=0}^{\infty} a_n b^n.$$

Proof.

- *Proof of (a):* We prove only that $A\{x\}$ is closed under multiplication. To that end let $f = \sum_{i=0}^{\infty} a_i x^i$ and $g = \sum_{j=0}^{\infty} b_j x^j$ be elements of $A\{x\}$. Consider $\varepsilon > 0$ and let n_0 be a positive number such that $\|a_i\| < \varepsilon$ if $i \ge \frac{n_0}{2}$ and $\|b_j\| < \varepsilon$ if $j \ge \frac{n_0}{2}$. Now let $n \ge n_0$ and $i + j = n$. Then $i \ge \frac{n_0}{2}$ or $j \ge \frac{n_0}{2}$. It follows that $\|\sum_{i+j=n} a_i b_j\| \le \max(\|a_i\| \cdot \|b_j\|)_{i+j=n} \le \varepsilon \cdot \max(\|f\|, \|g\|)$. Thus, $fg = \sum_{n=0}^{\infty} \sum_{i+j=n} a_i b_j x^n$ belongs to $A\{x\}$, as claimed.

- *Proof of (b):* Standard checking.

- *Proof of (c):* Let $f_i = \sum_{n=0}^{\infty} a_{in} x^n$, $i = 1, 2, 3, \ldots$, be a Cauchy sequence in $A\{x\}$. For each $\varepsilon > 0$ there exists i_0 such that $\|a_{in} - a_{jn}\| \le \|f_i - f_j\| < \varepsilon$ for all $i, j \ge i_0$ and for all n. Thus, for each n, the sequence $a_{1n}, a_{2n}, a_{3n}, \ldots$ is Cauchy, hence converges to an element $a_n \in A$. If we let j tend to infinity in the latter inequality, we get that $\|a_{in} - a_n\| < \varepsilon$ for all $i \ge i_0$ and all n. Set $f = \sum_{i=0}^{\infty} a_n x^n$. Then $a_n \to 0$ and $\|f_i - f\| = \max(\|a_{in} - a_n\|)_{n=0,1,2,\ldots} < \varepsilon$ if $i \ge i_0$. Consequently, the f_i's converge in $A\{x\}$.

- *Proof of (d):* Note that $\|a_n b^n\| \le \|a_n\| \to 0$, so $\sum_{n=0}^{\infty} a_n b^n$ is an element of B.

□

Definition 7.2. Let $f = \sum_{n=0}^{\infty} a_n x^n$ be a nonzero element of $A\{x\}$. We define the **pseudo degree** of f to be the integer $d = \max\{n \ge 0 \,|\, \|a_n\| = \|f\|\}$ and set

$$\text{pseudo.deg}(f) = d.$$

The element a_d is the **pseudo leading coefficient** of f. Thus, $\|a_d\| = \|f\|$ and $\|a_n\| < \|f\|$ for each $n > d$. If $f \in A[x]$ is a polynomial, then pseudo.deg$(f) \le$ deg(f). If a_d is invertible in A and satisfies $\|ca_d\| = \|c\| \cdot \|a_d\|$ for all $c \in A$, we call f **regular**. In particular, if A is a field and $\| \ \|$ is an ultrametric absolute value, then each $0 \ne f \in A\{x\}$ is regular. The following lemma implies that in this case $\| \ \|$ is an absolute value of $A\{x\}$.

Lemma 7.3 (Gauss' Lemma). *Let $f, g \in A\{x\}$. Suppose f is regular of pseudo degree d and $f, g \neq 0$. Then $\|fg\| = \|f\| \cdot \|g\|$ and* $\text{pseudo.deg}(fg) = \text{pseudo.deg}(f) + \text{pseudo.deg}(g)$.

Proof. Let $f = \sum_{i=0}^{\infty} a_i x^i$ and $g = \sum_{j=0}^{\infty} b_j x^j$. Let a_d (resp. b_e) be the pseudo leading coefficient of f (resp. g). Then $fg = \sum_{n=0}^{\infty} c_n x^n$ with $c_n = \sum_{i+j=n} a_i b_j$.

If $i + j = d + e$ and $(i, j) \neq (d, e)$, then either $i > d$ or $j > e$. In each case, $\|a_i b_j\| \leq \|a_i\| \|b_j\| < \|f\| \cdot \|g\|$. By our assumption on a_d, we have $\|a_d b_e\| = \|a_d\| \cdot \|b_e\| = \|f\| \cdot \|g\|$. By Lemma 6.3(b), this implies $\|c_{d+e}\| = \|f\| \cdot \|g\|$.

If $i + j > d + e$, then either $i > d$ and $\|a_i\| < \|f\|$ or $j > e$ and $\|b_j\| < \|g\|$. In each case $\|a_i b_j\| \leq \|a_i\| \cdot \|b_j\| < \|f\| \cdot \|g\|$. Hence, $\|c_n\| < \|c_{d+e}\|$ for each $n > d + e$. Therefore, c_{d+e} is the pseudo leading coefficient of fg, and the lemma is proved. □

Proposition 7.4 (Weierstrass division theorem). *Let $f \in A\{x\}$ and let $g \in A\{x\}$ be regular of pseudo degree d. Then there are unique $q \in A\{x\}$ and $r \in A[x]$ such that $f = qg + r$ and* $\deg(r) < d$. *Moreover,*

$$\|qg\| = \|q\| \cdot \|g\| \leq \|f\| \qquad \textit{and} \qquad \|r\| \leq \|f\|. \tag{7.1}$$

Proof. We break the proof into several parts.

- Part A: *Proof of* (7.1). First we assume that there exist $q \in A\{x\}$ and $r \in A[x]$ such that $f = qg + r$ with $\deg(r) < d$. If $q = 0$, then (7.1) is clear. Otherwise, $q \neq 0$ and we let $e = \text{pseudo.deg}(q)$. By Lemma 7.3, $\|qg\| = \|q\| \cdot \|g\|$ and $\text{pseudo.deg}(qg) = e + d > \deg(r)$. Hence, the coefficient c_{d+e} of x^{d+e} in qg is also the coefficient of x^{d+e} in f. It follows that $\|qg\| = \|c_{d+e}\| \leq \|f\|$. Consequently, $\|r\| = \|f - qg\| \leq \|f\|$.

- Part B *Uniqueness.* Suppose $f = qg + r = q'g + r'$, where $q, q' \in A\{x\}$ and $r, r' \in A[x]$ are of degrees less than d. Then $0 = (q - q')g + (r - r')$. By Part A, applied to 0 rather than to f, $\|q - q'\| \cdot \|g\| = \|r - r'\| = 0$. Hence, $q = q'$ and $r = r'$.

- Part C: *Existence if g is a polynomial of degree d.* Write $f = \sum_{n=0}^{\infty} b_n x^n$ with $b_n \in A$ converging to 0. For each $m \geq 0$ let $f_m = \sum_{n=0}^{m} b_n x^n \in A[x]$. Then the $f_1, f_2, f_3, \ldots$ converge to f, in particular they form a Cauchy sequence. Since g is regular of pseudo degree d, its leading coefficient is invertible. Euclid's algorithm for polynomials over A produces $q_m, r_m \in A[x]$ with $f_m = q_m g + r_m$ and $\deg(r_m) < \deg(g)$. Thus, for all k, m we have $f_m - f_k = (q_m - q_k)g + (r_m - r_k)$. By Part A, $\|q_m - q_k\| \cdot \|g\|, \|r_m - r_k\| \leq \|f_m - f_k\|$. Thus, $\{q_m\}_{m=0}^{\infty}$ and $\{r_m\}_{m=0}^{\infty}$ are Cauchy sequences in $A\{x\}$. Since $A\{x\}$ is complete (Lemma 7.1), the q_m's converge to some $q \in A\{x\}$. Since A is complete, the r_m's converge to an $r \in A[x]$ of degree less than d. It follows that $f = qg + r$.

- Part D: *Existence for arbitrary g.* Let $g = \sum_{n=0}^{\infty} a_n x^n$ and set $g_0 = \sum_{n=0}^{d} a_n x^n \in A[x]$. Then $\|g - g_0\| < \|g\|$. By Part C, there are $q_0 \in A\{x\}$ and $r_0 \in A[x]$ such that $f = q_0 g_0 + r_0$ and $\deg(r_0) < d$. By Part A, $\|q_0\| \le \frac{\|f\|}{\|g\|}$ and $\|r_0\| \le \|f\|$. Thus, $f = q_0 g + r_0 + f_1$, where $f_1 = -q_0(g - g_0)$, and $\|f_1\| \le \frac{\|g-g_0\|}{\|g\|} \cdot \|f\|$.

  Set $f_0 = f$. By induction we get, for each $k \ge 0$, elements $f_k, q_k \in A\{x\}$ and $r_k \in A[x]$ such that $\deg(r_k) < d$ and

$$f_k = q_k g + r_k + f_{k+1}, \quad \|q_k\| \le \frac{\|f_k\|}{\|g\|}, \quad \|r_k\| \le \|f_k\|, \quad \text{and}$$
$$\|f_{k+1}\| \le \frac{\|g - g_0\|}{\|g\|} \|f_k\|.$$

  It follows that $\|f_k\| \le \left(\frac{\|g-g_0\|}{\|g\|}\right)^k \|f\|$, so $\|f_k\| \to 0$. Hence, also $\|q_k\|, \|r_k\| \to 0$. Therefore, $q = \sum_{k=0}^{\infty} q_k \in A\{x\}$ and $r = \sum_{k=0}^{\infty} r_k \in A[x]$. By construction, $f = \sum_{n=0}^{k} q_n g + \sum_{n=0}^{k} r_n + f_{k+1}$ for each k. Taking k to infinity, we get $f = qg + r$ and $\deg(r) < d$.

□

Corollary 7.5 (Weierstrass preparation theorem). *Let $f \in A\{x\}$ be regular of pseudo degree d. Then $f = qg$, where q is a unit of $A\{x\}$ and $g \in A[x]$ is a monic polynomial of degree d with $\|g\| = 1$. Moreover, q and g are uniquely determined by these conditions.*

Proof. By Proposition 7.4 there are $q' \in A\{x\}$ and $r' \in A[x]$ of degree $< d$ such that $x^d = q'f + r'$ and $\|r'\| \le \|x^d\| = 1$. Set $g = x^d - r'$. Then g is monic of degree d, $g = q'f$, and $\|g\| = 1$. It remains to show that $q' \in A\{x\}^\times$.

Note that g is regular of pseudo degree d. By Proposition 7.4, there are $q \in A\{x\}$ and $r \in A[x]$ such that $f = qg + r$ and $\deg(r) < d$. Thus, $f = qq'f + r$. Since $f = 1 \cdot f + 0$, the uniqueness part of Proposition 7.4 implies that $qq' = 1$. Hence, $q' \in A\{x\}^\times$.

Finally suppose $f = q_1 g_1$, where $q_1 \in A\{x\}^\times$ and $g_1 \in A[x]$ is monic of degree d with $\|g_1\| = 1$. Then $g_1 = (q_1^{-1} q) g + 0$ and $g_1 = 1 \cdot g + (g_1 - g)$, where $g_1 - g$ is a polynomial of degree at most $d - 1$. By the uniqueness part of Proposition 7.4, $q_1^{-1} q_2 = 1$, so $q_1 = q_2$ and $g_1 = g$. □

Corollary 7.6. *Let $f = \sum_{n=0}^{\infty} a_n x^n$ be a regular element of $A\{x\}$ such that $\|a_0 b\| = \|a_0\| \cdot \|b\|$ for each $b \in A$. Then $f \in A\{x\}^\times$ if and only if* $\text{pseudo.deg}(f) = 0$ *and $a_0 \in A^\times$.*

Proof. If there exists $g \in \sum_{n=0}^{\infty} b_n x^n$ in $A\{x\}$ such that $fg = 1$, then

$$\text{pseudo.deg}(f) + \text{pseudo.deg}(g) = 0,$$

so pseudo.deg$(f) = 0$. In addition, $a_0 b_0 = 1$, so $a_0 \in A^\times$.

Conversely, suppose pseudo.deg$(f) = 0$ and $a_0 \in A^\times$. Then f is regular. Hence, by Corollary 7.5, $f = q \cdot 1$ where $q \in A\{x\}^\times$. □

Corollary 7.7. *Let K be a complete field with respect to an absolute value $|\ |$ and let $O = \{a \in K \mid |a| \leq 1\}$ be its valuation ring. Then $K\{x\}$ is a principal ideal domain, hence a unique factorization domain. Moreover, every ideal of $K\{x\}$ is generated by an element of $O[x]$.*

Proof. By the Weierstass preparation theorem (Corollary 7.5), every nonzero ideal $\mathfrak{a}$ of $K\{x\}$ is generated by the ideal $\mathfrak{a} \cap K[x]$ of $K[x]$. Since $K[x]$ is a principal ideal domain, $\mathfrak{a} \cap K[x] = fK[x]$ for some $f \in K[x]$. Consequently, $\mathfrak{a} = K\{x\}$ is a principal ideal. Moreover, dividing f by one of its coefficients with highest absolute value, we may assume that $f \in O[x]$. □

8 Convergent Power Series

Let K be a complete field with respect to an ultrametric absolute value $|\ |$. We say that a formal power series $f = \sum_{n=m}^\infty a_n x^n$ in $K((x))$ **converges** at an element $c \in K$, if $f(c) = \sum_{n=m}^\infty a_n c^n$ converges, i.e. $a_n c^n \to 0$. In this case f converges at each $b \in K$ with $|b| \leq |c|$. For example, each $f \in K\{x\}$ converges at 1. We say that f **converges** if f converges at some $c \in K^\times$.

We denote the set of all convergent power series in $K((x))$ by $K((x))_0$ and prove that $K((x))_0$ is a field that contains $K\{x\}$ and is algebraically closed in $K((x))$.

Lemma 8.1. *A power series $f = \sum_{n=m}^\infty a_n x^n$ in $K((x))$ converges if and only if there exists a positive real number γ such that $|a_n| \leq \gamma^n$ for each $n \geq 0$.*

Proof. First suppose f converges at $c \in K^\times$. Then $a_n c^n \to 0$, so there exists $n_0 \geq 1$ such that $|a_n c^n| \leq 1$ for each $n \geq n_0$. Choose

$$\gamma = \max\{|c|^{-1}, |a_k|^{1/k} \mid k = 0, \ldots, n_0 - 1\}.$$

Then $|a_n| \leq \gamma^n$ for each $n \geq 0$.

Conversely, suppose $\gamma > 0$ and $|a_n| \leq \gamma^n$ for all $n \geq 0$. Increase γ, if necessary, to assume that $\gamma > 1$. Then choose $c \in K^\times$ such that $|c| \leq \gamma^{-1.5}$ and observe that $|a_n c^n| \leq \gamma^{-0.5n}$ for each $n \geq 0$. Therefore, $a_n c^n \to 0$, hence f converges at c. □

Lemma 8.2. *$K((x))_0$ is a field that contains* Quot$(K\{x\})$*, hence also $K(x)$.*

Proof. The only difficulty is to prove that if $f = 1 + \sum_{n=1}^\infty a_n x^n$ converges, then also $f^{-1} = 1 + \sum_{n=1}^\infty a'_n x^n$ converges.

Indeed, for $n \geq 1$, a'_n satisfies the recursive relation $a'_n = -a_n - \sum_{i=1}^{n-1} a_i a'_{n-i}$. By Lemma 8.1, there exists $\gamma > 1$ such that $|a_i| \leq \gamma^i$ for each $i \geq 1$. Set $a'_0 = 1$. Suppose, by induction, that $|a'_j| \leq \gamma^j$ for $j = 1, \ldots, n-1$. Then $|a'_n| \leq \max_i(|a_i| \cdot |a'_{n-i}|) \leq \gamma^n$. Hence, f^{-1} converges. □

Let v be the valuation of $K((x))$ defined by

$$v(\sum_{n=m}^{\infty} a_n x^n) = m \qquad \text{for } a_m, a_{m+1}, a_{m+2}, \ldots \in K \text{ with } a_m \neq 0.$$

It is discrete, complete, its valuation ring is $K[[x]]$, and $v(x) = 1$. The residue of an element $f = \sum_{n=0}^{\infty} a_n x^n$ of $K[[x]]$ at v is a_0, and we denote it by $\bar{f}$. We also consider the valuation ring $O = K[[x]] \cap K((x))_0$ of $K((x))_0$ and denote the restriction of v to $K((x))_0$ also by v. Since $K((x))_0$ contains $K(x)$, it is v-dense in $K((x))$. Finally, we also denote the unique extension of v to the algebraic closure of $K((x))$ by v.

Remark 8.3 ($K((x))_0$ is not complete)**.** Indeed, choose $a \in K$ such that $|a| > 1$. Then there exists no $\gamma > 0$ such that $|a^{n^2}| \leq \gamma^n$ for all $n \geq 1$. By Lemma 8.1, the power series $f = \sum_{n=0}^{\infty} a^{n^2} x^n$ does not belong to $K((x))_0$. Therefore, the valued field $(K((x))_0, v)$ is not complete.

Proposition 8.4. *The field $K((x))_0$ is separably algebraically closed in $K((x))$.*

Proof. Let $y = \sum_{n=m}^{\infty} a_n x^n$, with $a_n \in K$, be an element of $K((x))$ which is separably algebraic of degree d over $K((x))_0$. We have to prove that $y \in K((x))_0$.

- Part A: *A shift of y.* Assume that $d > 1$ and let $y_1, \ldots, y_d$, with $y = y_1$, be the (distinct) conjugates of y over $K((x))_0$. In particular $r = \max(v(y - y_i) \mid i = 2, \ldots, d)$ is an integer. Choose $s \geq r + 1$ and let

  $$y_i' = \frac{1}{x^s}\big(y_i - \sum_{n=m}^{s} a_n x^n\big), \qquad i = 1, \ldots, d.$$

  Then $y_1', \ldots, y_d'$ are the distinct conjugates of y_1' over $K((x))_0$. Also, $v(y_1') \geq 1$ and $y_i' = \frac{1}{x^s}(y_i - y) + y_1'$, so $v(y_i') \leq -1$, $i = 2, \ldots, d$. If y_1' belongs to $K((x))_0$, then so does y, and conversely. Therefore, we replace y_i by y_i', if necessary, to assume that

  $$v(y) \geq 1 \text{ and } v(y_i) \leq -1,\ i = 2, \ldots, d. \tag{8.1}$$

  In particular $y = \sum_{n=0}^{\infty} a_n x^n$ with $a_0 = 0$. The elements $y_1, \ldots, y_d$ are the roots of an irreducible separable polynomial

  $$h(Y) = p_d Y^d + p_{d-1} Y^{d-1} + \cdots + p_1 Y + p_0$$

  with coefficients $p_i \in O$. Let $e = \min(v(p_0), \ldots, v(p_d))$. Divide the p_i, if necessary, by x^e, to assume that $v(p_i) \geq 0$ for each i between 0 and d and that $v(p_j) = 0$ for at least one j between 0 and d.

- Part B: *We prove that* $v(p_0), v(p_d) > 0$, $v(p_k) > v(p_1)$ *if* $2 \le k \le d-1$ *and* $v(p_1) = 0$. Indeed, since $v(y) > 0$ and $h(y) = 0$, we have $v(p_0) > 0$. Since $v(y_2) < 0$ and $h(y_2) = 0$, we have $v(p_d) > 0$. Next observe that

$$\frac{p_1}{p_d} = \pm y_2 \cdots y_d \pm \sum_{i=2}^{d} \frac{y_1 \cdots y_d}{y_i}.$$

If $2 \le i \le d$, then $v(y_i) < v(y_1)$, so $v(y_2 \cdots y_d) < v(\frac{y_1}{y_i}) + v(y_2 \cdots y_d) = v(\frac{y_1 \cdots y_d}{y_i})$. Hence,

$$v\Big(\frac{p_1}{p_d}\Big) = v(y_2 \cdots y_d). \tag{8.2}$$

For k between 1 and $d-2$ we have

$$\frac{p_{d-k}}{p_d} = \pm \sum_{\sigma} \prod_{i=1}^{k} y_{\sigma(i)}, \tag{8.3}$$

where σ ranges over all monotonically increasing maps from $\{1, \ldots, k\}$ to $\{1, \ldots, d\}$. If $\sigma(1) \ne 1$, then $\{y_{\sigma(1)}, \ldots, y_{\sigma(k)}\}$ is properly contained in $\{y_2, \ldots, y_d\}$. Hence, $v(\prod_{i=1}^{k} y_{\sigma(i)}) > v(y_2 \cdots y_d)$. If $\sigma(1) = 1$, then

$$v\Big(\prod_{i=1}^{k} y_{\sigma(i)}\Big) > v\Big(\prod_{i=2}^{k} y_{\sigma(i)}\Big) > v(y_2 \cdots y_d).$$

Hence, by (8.2) and (8.3), $v(\frac{p_{d-k}}{p_d}) > v(\frac{p_1}{p_d})$, so $v(p_{d-k}) > v(p_1)$. Since $v(p_j) = 0$ for some j between 0 and d, since $v(p_i) \ge 0$ for every i between 0 and d, and since $v(p_0), v(p_d) > 0$, we conclude that $v(p_1) = 0$ and $v(p_i) > 0$ for all $i \ne 1$. Therefore,

$$p_k = \sum_{n=0}^{\infty} b_{kn} x^n, \qquad k = 0, \ldots, d \tag{8.4}$$

with $b_{kn} \in K$ such that $b_{1,0} \ne 0$ and $b_{k,0} = 0$ for each $k \ne 1$. In particular, $|b_{1,0}| \ne 0$. We shall need it to be larger that 1 in order to bound $|a_n|$.

- Part C: *Making* $|b_{1,0}|$ *large.* We choose $c \in K$ such that $|c^{d-1} b_{1,0}| \ge 1$ and let $z = cy$. Then z is a zero of the polynomial $g(Z) = p_d Z^d + c p_{d-1} Z^{d-1} + \cdots + c^{d-1} p_1 Z + c^d p_0$ with coefficients in O. Relation (8.4) remains valid except that the zero term of the coefficient of Z in g becomes $c^{d-1} b_{1,0}$. By the choice of c, its absolute value is at least 1. So, without loss, we may assume that

$$|b_{1,0}| \ge 1. \tag{8.5}$$

- PART D: *An estimate for* $|a_n|$. By Lemma 8.1, there exists $\gamma > 0$ such that $|b_{kn}| \leq \gamma^n$ for all $0 \leq k \leq d$ and $n \geq 1$. By induction we prove that $|a_n| \leq \gamma^n$ for each $n \geq 0$. This will prove that $y \in O$ and will conclude the proof of the lemma.

  Indeed, $|a_0| = 0 < 1 = \gamma^0$. Now assume that $|a_m| \leq \gamma^m$ for each $0 \leq m \leq n-1$. For each k between 0 and d we have that $p_k y^k = \sum_{n=0}^{\infty} c_{kn} x^n$, where

  $$c_{kn} = \sum_{\sigma \in S_{kn}} b_{k,\sigma(0)} \prod_{j=1}^{k} a_{\sigma(j)},$$

  and

  $$S_{kn} = \{\sigma \colon \{0, \ldots, k\} \to \{0, \ldots, n\} \mid \sum_{j=0}^{k} \sigma(j) = n\}.$$

  It follows that

  $$c_{0n} = b_{0n} \text{ and } c_{1n} = b_{1,0} a_n + b_{11} a_{n-1} + \cdots + b_{1,n-1} a_1. \tag{8.6}$$

  For $k \geq 2$ we have $b_{k,0} = 0$. Hence, if a term $b_{k,\sigma(0)} \prod_{j=1}^{k} a_{\sigma(j)}$ in c_{kn} contains a_n, then $\sigma(0) = 0$, so $b_{k,\sigma(0)} = 0$. Thus,

  $$\begin{aligned} c_{kn} = \text{sum of products of the form } b_{k,\sigma(0)} \prod_{j=1}^{k} a_{\sigma(j)}, \\ \text{with } \sigma(j) < n,\ j = 1, \ldots, k. \end{aligned} \tag{8.7}$$

  From the relation $\sum_{k=0}^{d} p_k y^k = h(y) = 0$ we conclude that $\sum_{k=0}^{d} c_{kn} = 0$ for all n. Hence, by (8.6),

  $$b_{1,0} a_n = -b_{0n} - b_{11} a_{n-1} - \cdots - b_{1,n-1} a_1 - c_{2n} - \cdots - c_{dn}.$$

  Therefore, by (8.7),

  $$\begin{aligned} b_{1,0} a_n = \text{sum of products of the form } -b_{k,\sigma(0)} \prod_{j=1}^{k} a_{\sigma(j)}, \\ \text{with } \sigma \in S_{kn},\ 0 \leq k \leq d, \text{ and } \sigma(j) < n,\ j = 1, \ldots, k. \end{aligned} \tag{8.8}$$

  Note that $b_{k,0} = 0$ for each $k \neq 1$ (by (8.4)), while $b_{1,0}$ does not occur on the right hand side of (8.8). Hence, for a summand in the right hand side of (8.8) indexed by σ we have

  $$|b_{k,\sigma(0)} \prod_{j=1}^{k} a_{\sigma(j)}| \leq \gamma^{\sum_{j=0}^{k} \sigma(j)} = \gamma^n.$$

We conclude from $|b_{1,0}| \geq 1$ that $|a_n| \leq \gamma^n$, as contended.

□

Proposition 8.5. *The field $K((x))_0$ is algebraically closed in $K((x))$. Thus, each $f \in K((x))$ which is algebraic over $K(x)$ converges at some $c \in K^\times$. Moreover, there exists a positive integer m such that f converges at each $b \in K^\times$ with $|b| \leq \frac{1}{m}$.*

Proof. In view of Proposition 8.4, we have to prove the proposition only for $\mathrm{char}(K) > 0$. Let $f = \sum_{n=m}^{\infty} a_n x^n \in K((x))$ be algebraic over $K((x))_0$. Then $K((x))_0(f)$ is a purely inseparable extension of a separable algebraic extension of $K((x))_0$. By Proposition 8.4, the latter coincides with $K((x))_0$. Hence, $K((x))_0(f)$ is a purely inseparable extension of $K((x))_0$.

Thus, there exists a power q of $\mathrm{char}(K)$ such that $\sum_{n=m}^{\infty} a_n^q x^{nq} = f^q \in K((x))_0$. By Lemma 8.1, there exists $\gamma > 0$ such that $|a_n^q| \leq \gamma^{nq}$ for all $n \geq 1$. It follows that $|a_n| \leq \gamma^n$ for all $n \geq 1$. By Lemma 8.1, $f \in K((x))_0$, so there exists $c \in K^\times$ such that f converges at c. If $\frac{1}{m} \leq |c|$, then f converges at each $b \in K^\times$ with $|b| \leq \frac{1}{m}$. □

9 Several Variables

Starting from a complete valued field $(K, |\ |)$, we choose an element $r \in K^\times$, a finite set I, and for each $i \in I$ an element $c_i \in K$ such that $|r| \leq |c_i - c_j|$ if $i \neq j$. Then we set $w_i = \frac{r}{x-c_i}$, with an indeterminate x, and consider the ring $R = K\{w_i \mid i \in I\}$ of all series

$$f = a_0 + \sum_{i\in I} \sum_{n=1}^{\infty} a_{in} w_i^n,$$

with $a_0, a_{in} \in K$ such that for each i the element a_{in} tends to 0 as $n \to \infty$. The ring R is complete under the norm defined by $\|f\| = \max_{i,n}(|a_0|, |a_{in}|)$ (Lemma 11.1). We prove that R is a principal ideal domain (Proposition 11.8) and denote its quotient field by Q. More generally for each subset J of I, we denote the quotient field of $K\{w_i \mid i \in J\}$ by P_J. We deduce (Proposition 12.1) that $P_J \cap P_{J'} = P_{J\cap J'}$ if $J, J' \subseteq I$ have a nonempty intersection and $P_J \cap P_{J'} = K(x)$ if $J \cap J' = \emptyset$. Thus, setting $P_i = P_{I \smallsetminus \{i\}}$ for $i \in I$, we conclude that $\bigcap_{i\in I} P_i = K(x)$. The fields $E = K(x)$ and P_i, $i \in I$, are the first objects of patching data (Definition 4.1) that we start to assemble.

10 A Normed Subring of $K(x)$

Let $E = K(x)$ be the field of rational functions in the variable x over a field K. Let I be a finite set and r an element of $K^\times$. For each $i \in I$ let c_i be an element of K. Suppose $c_i \neq c_j$ if $i \neq j$. For each $i \in I$ let $w_i = \frac{r}{x-c_i} \in K(x)$. We consider the subring $R_0 = K[w_i \mid i \in I]$ of $K(x)$, prove that each of its elements is a linear combination of the powers w_i^n with coefficients in K, and define a norm on R_0.

Lemma 10.1. *(a) For all $i \neq j$ in I and for each nonnegative integer m*

$$w_i w_j^m = \frac{r^m}{(c_i - c_j)^m} w_i - \sum_{k=1}^{m} \frac{r^{m+1-k}}{(c_i - c_j)^{m+1-k}} w_j^k. \tag{10.1}$$

(b) Given nonnegative integers m_i, $i \in I$, not all zero, there exist $a_{ik} \in K$ such that

$$\prod_{i \in I} w_i^{m_i} = \sum_{i \in I} \sum_{k=1}^{m_i} a_{ik} w_i^k. \tag{10.2}$$

(c) Every $f \in K[w_i \mid i \in I]$ can be uniquely written as

$$f = a_0 + \sum_{i \in I} \sum_{n=1}^{\infty} a_{in} w_i^n \tag{10.3}$$

where $a_0, a_{in} \in K$ and almost all of them are zero.

(d) Let $i \neq j$ be elements of I. Then $\frac{w_i}{w_j} = 1 + \frac{c_i - c_j}{r} w_i \in K[w_i]$ is invertible in $K[w_i, w_j]$.

Proof.

- *Proof of (a) and (b):* Starting from the identity

$$w_i w_j = \frac{r}{c_i - c_j} w_i - \frac{r}{c_i - c_j} w_j \tag{10.4}$$

  one proves (10.1) by induction on m. Then one proceeds by induction on $|I|$ and $\max_{i \in I} m_i$ to prove (10.2).

- *Proof of (c):* The existence of the presentation (10.3) follows from (b). To prove the uniqueness we assume that $f = 0$ in (10.3) but $a_{jk} \neq 0$ for some $j \in I$ and $k \in \mathbb{N}$. Then, $\sum_{n=1}^{\infty} a_{jn} w_j^n = -a_0 - \sum_{i \neq j} \sum_{n=1}^{\infty} a_{in} w_i^n$. The left hand side has a pole at c_j while the right hand side has not. This is a contradiction.

- *Proof of (d):* Multiplying $\frac{r}{w_j} - \frac{r}{w_i} = c_i - c_j$ by $\frac{w_i}{r}$ we get that

$$\frac{w_i}{w_j} = 1 + \frac{c_i - c_j}{r} w_i$$

is in $K[w_i]$. Similarly, $\frac{w_j}{w_i} \in K[w_j]$. Hence $\frac{w_i}{w_j}$ is invertible in $K[w_i, w_j]$.

□

Through the rest of the section and sections 11, 12, and 13, we make the following assumption:

Assumption 10.2. *The field K is complete with respect to a nontrivial ultrametric absolute value $|\ |$ and*

$$|r| \leq |c_i - c_j| \quad \textit{for all } i \neq j. \tag{10.5}$$

Geometrically, Condition (10.5) means that the open disks $\{a \in K \mid |a - c_i| < r\}$, $i \in I$, of K are disjoint.

Let $E = K(x)$ be the field of rational functions over K in the variable x. We define a function $\|\ \|$ on $R_0 = K[w_i \mid i \in I]$ using the unique presentation (10.3):

$$\|a_0 + \sum_{i \in I} \sum_{n \geq 1} a_{in} w_i^n\| = \max_{i,n}\{|a_0|, |a_{in}|\}.$$

Then $\|f\| \geq 0$ for each $f \in R_0$, $\|f\| = 0$ if and only if $f = 0$ (Lemma 10.1(c)), and $\|f + g\| \leq \max(\|f\|, \|g\|)$ for all $f, g \in R_0$. Moreover, $\|w_i\| = 1$ for each $i \in I$ but $\|w_i w_j\| = \frac{|r|}{|c_i - c_j|}$ (by (10.4)) is less than 1 if $|r| < |c_i - c_j|$. Thus, $\|\ \|$ is in general not an absolute value. However, by (10.1) and (10.5)

$$\|w_i w_j^m\| \leq \max_{1 \leq k \leq m}\Big(\Big|\frac{r}{c_i - c_j}\Big|^m, \Big|\frac{r}{c_i - c_j}\Big|^{m+1-k}\Big) \leq 1.$$

By induction, $\|w_i^k w_j^m\| \leq 1$ for each k, so $\|fg\| \leq \|f\| \cdot \|g\|$ for all $f, g \in R_0$. Moreover, if $a \in K$ and $f \in R_0$, then $\|af\| = \|a\| \cdot \|f\|$. Therefore, $\|\ \|$ is a norm on R_0 in the sense of Definition 6.1.

11 Mittag-Leffler Series

We keep the notation of Section 10 and Assumption 10.2 and proceed to define rings of convergent power series of several variables over K. In the language of rigid geometry, these are the rings of holomorphic functions on the complements of finitely many open discs of the projective line $\mathbb{P}^1(K)$.

Let $R = K\{w_i \mid i \in I\}$ be the completion of $R_0 = K[w_i \mid i \in I]$ with respect to $\|\ \|$ (Lemma 6.5). Our first result gives a Mittag-Leffler decomposition of each $f \in R$. It generalizes Lemma 10.1(c):

Lemma 11.1. *Each element f of R has a unique presentation as a* **Mittag-Leffler series**

$$f = a_0 + \sum_{i \in I} \sum_{n=1}^{\infty} a_{in} w_i^n, \tag{11.1}$$

where $a_0, a_{in} \in K$*, and* $|a_{in}| \to 0$ *as* $n \to \infty$*. Moreover,*

$$\|f\| = \max_{i,n}\{|a_0|, |a_{in}|\}.$$

Proof. Each f as in (11.1) is the limit of the sequence $(f_d)_{d\geq 1}$ of its partial sums $f_d = a_0 + \sum_{i\in I}\sum_{n=1}^{d} a_{in}w_i^n \in R_0$, so $f \in R$. Since $\|f_d\| = \max_{i,n}(|a_0|, |a_{in}|)$ for each sufficiently large d, we have $\|f\| = \max_{i,n}(|a_0|, |a_{in}|)$. If $f = 0$ in (11.1), then $0 = \max_{i,n}(|a_0|, |a_{in}|)$, so $a_0 = a_{in} = 0$ for all i and n. It follows that the presentation (11.1) is unique.

On the other hand, let $g \in R$. Then there exists a sequence of elements $g_k = a_{k,0} + \sum_{i\in I}\sum_{n=1}^{\infty} a_{k,in}w_i^n$, $k = 1, 2, 3, \ldots$, in R_0, that converges to g. In particular, for each pair (k, i) we have $a_{k,in} = 0$ if n is sufficiently large. Also, the sequence $(g_k)_{k=1}^{\infty}$ is Cauchy. Hence, each of the sequences $\{a_{k,0} \mid k = 1, 2, 3, \ldots\}$ and $\{a_{k,in} \mid k = 1, 2, 3, \ldots\}$ is Cauchy. Since K is complete, $a_{k,0} \to a_0$ and $a_{k,in} \to a_{in}$ for some $a_0, a_{in} \in K$. Fix $i \in I$ and let $\varepsilon > 0$ be a real number. There is an m such that for all $k \geq m$ and all n we have $|a_{k,in} - a_{m,in}| \leq \|g_k - g_m\| \leq \varepsilon$. If n is sufficiently large, then $a_{m,in} = 0$, and hence $|a_{k,in}| \leq \varepsilon$. Therefore, $|a_{in}| \leq \varepsilon$. It follows that $|a_{in}| \to 0$. Define f by (11.1). Then $f \in R$ and $g_k \to f$ in R. Consequently, $g = f$. □

If $I = \emptyset$, then $R = R_0 = K$.

We call the partial sum $\sum_{n=1}^{\infty} a_{in}w_i^n$ in (11.1) the i-**component** of f.

Remark 11.2. Let $i \in I$. Then $K\{w_i\} = \{\sum_{n=0}^{\infty} a_n w_i^n \mid a_n \to 0\}$ is a subring of R, the completion of $K[w_i]$ with respect to the norm. Consider the ring $K\{x\}$ of converging power series over K. By Lemma 7.1(d), there is a homomorphism $K\{x\} \to K\{w_i\}$ given by $\sum_{n=0}^{\infty} a_n x^n \mapsto \sum_{n=0}^{\infty} a_n w_i^n$. By Lemma 11.1, this is an isomorphism of normed rings.

Remark 11.3. Let $(L, |\;|)$ be a complete valued field extending $(K, |\;|)$. Each $c \in L$ with $|c - c_i| \geq |r|$, for all $i \in I$, defines a continuous **evaluation homomorphism** $R \to L$ given by $f = a_0 + \sum_{i\in I}\sum_n a_{in}w_i^n \mapsto f(c) = a_0 + \sum_{i\in I}\sum_n a_{in}(\frac{r}{c-c_i})^n$. Indeed, $x \mapsto c$ defines a K-homomorphism $\phi\colon K[x] \to L$. Let P be its kernel. Then ϕ extends to the localization $K[x]_P$. Since $\phi(x - c_i) = c - c_i \neq 0$, we have $w_i \in K[x]_P$, for each $i \in I$. Thus, ϕ restricts to a homomorphism $R_0 \to L$, given by the above formula. Since $\left|\frac{r}{c-c_i}\right| \leq 1$ for each i, we have $|f(c)| \leq \|f\|$ for each $f \in R_0$. Hence, ϕ uniquely extends to a continuous homomorphism $\phi\colon R \to L$.

Lemma 11.4 (Degree shifting)**.** *Let $f \in R$ be given by (11.1). Fix $i \neq j$ in I. Let $\sum_{n=1}^\infty a'_{in} w_i^n$ be the i-component of $\frac{w_j}{w_i} f \in R$. Then*

$$\begin{aligned} a'_{in} &= -\sum_{\nu=n+1}^{\infty} \frac{a_{i\nu} r^{\nu-n}}{(c_j - c_i)^{\nu-n}} \\ &= \frac{-r}{c_j - c_i} \sum_{\nu=n+1}^{\infty} a_{i\nu} \Big(\frac{r}{c_j - c_i}\Big)^{\nu-(n+1)}, \quad n = 1, 2, 3, \ldots. \end{aligned} \tag{11.2}$$

Furthermore, let $m \geq 1$ be an integer, and let $\sum_{n=1}^\infty b_{in} w_i^n$ be the i-component of $(\frac{w_j}{w_i})^m f$. Let $\varepsilon \geq 0$ be a real number and let d be a positive integer.

(a) If $|a_{in}| \leq \varepsilon$ for each $n \geq d+1$, then $|b_{in}| \leq |\frac{r}{c_j - c_i}|^m \varepsilon$ for each $n \geq d+1-m$.

(b) Suppose $d > m$. If $|a_{in}| < \varepsilon$ for each $n \geq d+1$ and $|a_{id}| = \varepsilon$, then $|b_{in}| < |\frac{r}{c_j - c_i}|^m \varepsilon$ for each $n \geq d+1-m$ and $|b_{i,d-m}| = |\frac{r}{c_j - c_i}|^m \varepsilon$.

(c) $\sum_{n=1}^\infty a_{in} w_i^n$ is a polynomial in w_i if and only if $\sum_{n=1}^\infty b_{in} w_i^n$ is.

Proof. By Lemma 10.1(d), $\frac{w_j}{w_i} \in R^\times$, so $(\frac{w_j}{w_i})^m f \in R$ for each m and the above statements make sense.

- Proof of (11.2): We may assume that $a_0 = a_{i1} = 0$ and $a_{k\nu} = 0$ for each $k \neq i$ and each ν. Indeed, $\frac{w_j}{w_i} = 1 + (c_j - c_i)\frac{w_j}{r} \in K\{w_j\}$. Hence, $\frac{w_j}{w_i} \cdot w_k^\nu \in K\{w_l \mid l \neq i\}$. Furthermore, $\frac{w_j}{w_i} \cdot w_i = w_j \in K\{w_l \mid l \neq i\}$. Hence, by (11.1), a_0, a_{i1}, and the $a_{k\nu}$ do not contribute to the i-component of $\frac{w_j}{w_i} f$. Thus, $f = \sum_{\nu=2}^\infty a_{i\nu} w_i^\nu$. Hence, by (10.1) of Section 10,

$$\begin{aligned} \frac{w_j}{w_i} f = \sum_{\nu=2}^{\infty} a_{i\nu} w_j w_i^{\nu-1} &= \sum_{\nu=2}^{\infty} a_{i\nu} \Big[\frac{r^{\nu-1}}{(c_j - c_i)^{\nu-1}} w_j - \sum_{n=1}^{\nu-1} \frac{r^{\nu-n}}{(c_j - c_i)^{\nu-n}} w_i^n\Big] \\ &= \sum_{\nu=2}^{\infty} \frac{a_{i\nu} r^{\nu-1}}{(c_j - c_i)^{\nu-1}} w_j - \sum_{n=1}^{\infty} \sum_{\nu=n+1}^{\infty} \frac{a_{i\nu} r^{\nu-n}}{(c_j - c_i)^{\nu-n}} w_i^n \end{aligned},$$

  from which (11.2) follows.

- Proof of (a) and (b): By induction on m it suffices to assume that $m = 1$. In this case we have to prove: (a) If $|a_{in}| \leq \varepsilon$ for each $n \geq d+1$, then $|a'_{in}| \leq |\frac{r}{c_j - c_i}| \varepsilon$ for each $n \geq d$; (b) assuming $d \geq 2$, if $|a_{in}| < \varepsilon$ for each $n \geq d+1$ and $|a_{id}| = \varepsilon$, then $|a'_{in}| < |\frac{r}{c_j - c_i}| \varepsilon$ for each $n \geq d$ and $|a'_{i,d-1}| = |\frac{r}{c_j - c_i}| \varepsilon$. By Condition (10.5) of Section 10, $|\frac{r}{c_i - c_j}| \leq 1$. Hence, (a) follows from (11.2) with $n = d, d+1, d+2, \ldots$ and (b) follows from (11.2) with $n = d-1, d, d+1, \ldots$.

- Proof of (c): Again, it suffices to prove that $\sum_{n=1}^\infty a_{in}w_i^n$ is a polynomial if and only if $\sum_{n=1}^\infty a'_{in}w_i^n$ is a polynomial.

  If $\sum_{n=1}^\infty a_{in}w_i^n$ is a polynomial, then $a_{i\nu}=0$ for all large ν. It follows from (11.2) that $a'_{i,n}=0$ for all large n. Hence, $\sum_{n=1}^\infty a'_{in}w_i^n$ is a polynomial.

  If $\sum_{n=1}^\infty a_{in}w_i^n$ is not a polynomial, then for each d_0 there exists $d>d_0$ such that $a_{id}\neq 0$. Since $|a_{in}|\to 0$ as $n\to\infty$, there are only finitely many $n\geq d$ with $|a_{in}|\geq |a_{id}|$. Replacing d with the largest of those n's, if necessary, we may assume that $|a_{in}|<|a_{id}|$ for each $n\geq d+1$. By (b), $a'_{i,d-1}\neq 0$. Consequently, $\sum_{n=1}^\infty a'_{in}w_i^n$ is not a polynomial.

□

We apply degree shifting (albeit not yet Lemma 11.4) to generalize the Weierstrass preparation theorem (Corollary 7.5) to Mittag-Leffler series.

Lemma 11.5. *Suppose $I\neq\emptyset$ and let $0\neq f\in R$. Then there is an $i\in I$ such that $f=pu$ with $p\in K[w_i]$ and $u\in R^\times$.*

Proof. Write f in the form (11.1). Then, there is a coefficient with absolute value $\|f\|$. Thus we are either in Case I or Case II below:

- Case I: $|a_0|=\|f\|>|a_{in}|$ *for all i and n.* Multiply f by a_0^{-1} to assume that $a_0=1$. Then $\|1-f\|<1$. By Lemma 6.3(f), $f\in R^\times$, so $p=1$ and $u=f$ satisfy the claim of the lemma for each $i\in I$.

- Case II: *There exist i and $d\geq 1$ such that $|a_{id}|=\|f\|$.* Increase d, if necessary, to assume that $|a_{in}|<|a_{id}|=\|f\|$ for all $n>d$.

  Let $A=K\{w_k\mid k\neq i\}$. This is a complete subring of R. We introduce a new variable z, and consider the ring $A\{z\}$ of convergent power series in z over A (Lemma 7.1(c)). Since $a_{id}\in K^\times\subseteq A^\times$, the element

  $$\hat f=(a_0+\sum_{k\neq i}\sum_{n=1}^\infty a_{kn}w_k^n)+\sum_{n=1}^\infty a_{in}z^n$$

  of $A\{z\}$ is regular of pseudo degree d. By Corollary 7.5, we have $\hat f=\hat p\hat u$, where $\hat u$ is a unit of $A\{z\}$ and $\hat p$ is a monic polynomial of degree d in $A[z]$.

  By definition, $\|w_i\|=1$. By Lemma 7.1(d), the evaluation homomorphism $\theta\colon A\{z\}\to R$ defined by $\sum c_nz^n\mapsto\sum c_nw_i^n$, with $c_n\in A$, maps $\hat f$ onto f, $\hat u$ onto a unit of R, and $\hat p$ onto a polynomial p of degree d in $A[w_i]$. Replacing f by p and using Lemma 10.1, we may assume that $f\in A[w_i]=A+K[w_i]$ is a polynomial of degree d in w_i, that is,

  $$f=(a_0+\sum_{k\neq i}\sum_{n=1}^\infty a_{kn}w_k^n)+\sum_{n=1}^d a_{in}w_i^n.$$

If $I = \{i\}$, then $A[w_i] = K[w_i]$, and we are done. If $|I| \ge 2$, we choose a $j \in I$ distinct from i. By Lemma 10.1(d), $\frac{w_j}{w_i} = 1 + \frac{c_j - c_i}{r} w_j$ is invertible in R_0, hence in R. Since $\frac{w_j}{w_i} \in A$, we have $\frac{w_j}{w_i}(\sum_{k \ne i} \sum_{n=1}^{\infty} a_{kn} w_k^n) \in A$. In addition, by Lemma 10.1,

$$\frac{w_j}{w_i} \sum_{n=1}^{d} a_{in} w_i^n = \sum_{n=1}^{d} a_{in} w_i^{n-1} w_j$$

is a polynomial in $A[w_i]$ of degree $\le d-1$. Using induction on d, we may assume that $f \in A$. Finally, we apply the induction hypothesis (on $|I|$) to conclude the proof.

□

Lemma 11.6. *Let $j \in I$. Then each $f \in R$ can be written as $f = pu$ with $p \in K[w_j]$, $\|p\| = 1$, and $u \in R^\times$.*

Proof. Lemma 11.5 gives a decomposition $f = p_1 u_1$ with $u_1 \in R^\times$ and $p_1 \in K[w_i]$ for some $i \in I$. If $i = j$, we set $p = p_1$ and $u = u_1$. If $i \ne j$, we may assume that $f \in K[w_i]$. Thus, $f = \sum_{n=0}^{d} a_n w_i^n$ with $a_d \ne 0$. By Lemma 10.1(d), $\frac{w_i}{w_j}$ is invertible in R_0, hence in R. Multiplying f by $\left(\frac{w_j}{w_i}\right)^d$ gives

$$\left(\frac{w_j}{w_i}\right)^d f = \sum_{n=0}^{d} a_n \left(\frac{w_j}{w_i}\right)^{d-n} w_j^n = \sum_{n=0}^{d} a_n \left(1 + \frac{c_j - c_i}{r} w_j\right)^{d-n} w_j^n \in K[w_j].$$

Thus, $f = pu$ with $p \in K[w_j]$ and $u \in R^\times$. Finally, we may divide p by a coefficient with the highest absolute value to get that $\|p\| = 1$. □

Corollary 11.7. *Let $0 \ne g \in R$. Then $R_0 + gR = R$.*

Proof. Since $R = \sum_{i \in I} K\{w_i\}$ and $R_0 = K[w_i \mid i \in I] = \sum_{i \in I} K[w_i]$ (Lemma 10.1), it suffices to prove for each $i \in I$ and for every $f \in K\{w_i\}$ that there is $h \in K[w_i]$ such that $f - h \in gR$. By Lemma 11.6, we may assume that $g \in K[w_i]$. By Remark 11.2, there is a K-isomorphism $K\{z\} \to K\{w_i\}$ that maps $K[z]$ onto $K[w_i]$. Therefore the assertion follows from the Weierstrass Division Theorem (Proposition 7.4) for the ring $K\{z\}$. □

The next result generalizes Proposition 7.7 to Mittag-Leffler series.

Proposition 11.8. *The ring $R = K\{w_i \mid i \in I\}$ is a principal ideal domain, hence a unique factorization domain. Moreover, for each $i \in I$, each ideal $\mathfrak{a}$ of R is generated by an element $p \in K[w_i]$ such that $\mathfrak{a} \cap K[w_i] = pK[w_i]$.*

Proof. Let $f_1, f_2 \in R$ with $f_1 f_2 = 0$. Choose an $i \in I$. By Lemma 11.6, $f_1 = p_1 u_1$ and $f_2 = p_2 u_2$ with $p_1, p_2 \in K[w_i]$ and $u_1, u_2 \in R^\times$. Then $p_1 p_2 = f_1 f_2 (u_1 u_2)^{-1} = 0$, and hence either $p_1 = 0$ or $p_2 = 0$. Therefore, either $f_1 = 0$ or $f_2 = 0$. Consequently, R is an integral domain.

By Lemma 11.6, each ideal $\mathfrak{a}$ of R is generated by the ideal $\mathfrak{a} \cap K[w_i]$ of $K[w_i]$. Since $K[w_i]$ is a principal ideal domain, $\mathfrak{a} \cap K[w_i] = pK[w_i]$ for some $p \in K[w_i]$. Consequently, $\mathfrak{a} = pR$ is a principal ideal. □

12 Fields of Mittag-Leffler Series

In the notation of Sections 10 and 11 we consider for each nonempty subset J of I the integral domain $R_J = K\{w_i \mid i \in J\}$ (Proposition 11.8) and let $P_J = \mathrm{Quot}(R_J)$. For $J = \emptyset$, we set $P_J = K(x)$. All of these fields are contained in the field $Q = P_I$. The fields $P_i = P_{I \smallsetminus \{i\}}$, $i \in I$, will be our 'analytic' fields in our patching data. As in Definition 4.1, $P_i' = \bigcap_{j \neq i} P_j$ will be useful auxiliary fields.

Proposition 12.1. *Let J and J' be subsets of I. Then, $P_J \cap P_{J'} = P_{J \cap J'}$.*

Proof. If either $J = \emptyset$ or $J' = \emptyset$, then $P_J \cap P_{J'} = K(x)$, by definition. We therefore assume that $J, J' \neq \emptyset$. Let $j \in J$. Then $K[w_j] \subseteq R_J$, hence $K(x) = K(w_j) \subseteq P_J$. Similarly $K(x) \subseteq P_{J'}$. Hence $K(x) \subseteq P_J \cap P_{J'}$. If $J \cap J' \neq \emptyset$, then, by the unique representation of elements in R appearing in (11.1) of Lemma 11.1, we have $R_{J \cap J'} = R_J \cap R_{J'}$, so $P_{J \cap J'} \subseteq P_J \cap P_{J'}$.

For the converse inclusion, let $0 \neq f \in P_J \cap P_{J'}$. Fix $j \in J$ and $j' \in J'$; if $J \cap J' \neq \emptyset$, take $j, j' \in J \cap J'$. Write f as f_1/g_1 with $f_1, g_1 \in R_J$. By Lemma 11.6, $g_1 = p_1 u_1$, where $0 \neq p_1 \in K[w_j]$ and $u_1 \in R_J^\times$. Replace f_1 by $f_1 u_1^{-1}$ to assume that $g_1 \in K[w_j]$. Similarly $f = f_2/g_2$ with $f_2 \in R_{J'}$ and $g_2 \in K[w_{j'}]$.

If $J \cap J' \neq \emptyset$, then $g_1, g_2 \in R_J \cap R_{J'} = R_{J \cap J'}$. Thus $g_2 f_1 = g_1 f_2 \in R_J \cap R_{J'} = R_{J \cap J'} \subseteq P_{J \cap J'}$, and hence $f = \frac{f_1 g_2}{g_1 g_2} \in P_{J \cap J'}$.

Now suppose $J \cap J' = \emptyset$. Let $g_1 = \sum_{n=0}^{d_1} b_n w_j^n$ with $b_n \in K$. Put $h_1 = (\frac{w_{j'}}{w_j})^{d_1} g_1$. Since $\frac{w_{j'}}{w_j} \in K[w_{j'}]$ (Lemma 10.1(d)), we have $h_1 = \sum_{n=0}^{d_1} b_n (\frac{w_{j'}}{w_j})^{d_1 - n} w_{j'}^n \in K[w_{j'}]$. Similarly there is an integer $d_2 \geq 0$ such that $h_2 = (\frac{w_j}{w_{j'}})^{d_2} g_2 \in K[w_j]$. Let $d = d_1 + d_2$. Then, for each $k \in J$

$$f_1 h_2 \cdot \left(\frac{w_{j'}}{w_k}\right)^d = f_2 h_1 \cdot \left(\frac{w_j}{w_k}\right)^d. \tag{12.1}$$

Note that $f_1 h_2 \in R_J$ while $f_2 h_1 \in R_{J'}$. In particular, the k-component of $f_2 h_1$ is zero. By Lemma 11.4(c), the k-component of $f_2 h_1 \cdot \left(\frac{w_j}{w_k}\right)^d$ is a polynomial in w_k. By (12.1), the k-component of $f_1 h_2 \cdot \left(\frac{w_{j'}}{w_k}\right)^d$ is a polynomial in w_k. Hence, again by Lemma 11.4(c), the k-component of $f_1 h_2$ is a polynomial in w_k.

We conclude that $f_1 h_2 \in K[w_k \mid k \in J]$, so $f = \frac{f_1 h_2}{g_1 h_2} \in K(x)$. □

Corollary 12.2. *For each $i \in I$ we have $P_i' = P_{\{i\}}$. Also, $\bigcap_{j\in I} P_j = K(x)$.*

Proof. We apply Proposition 12.1 several times:

$$P_i' = \bigcap_{j\neq i} P_j = \bigcap_{j\neq i} P_{I\smallsetminus\{j\}} = P_{\bigcap_{j\neq i} I\smallsetminus\{j\}} = P_{\{i\}}.$$

For the second equality we choose an $i \in I$. Then

$$\bigcap_{j\in I} P_j = P_{I\smallsetminus\{i\}} \cap \bigcap_{j\neq i} P_{I\smallsetminus\{j\}} = P_{I\smallsetminus\{i\}} \cap P_{\{i\}} = K(x),$$

as claimed. □

13 Factorization of Matrices over Complete Rings

We show in this section how to decompose a matrix over a complete ring into a product of matrices over certain complete subrings. This will establish the decomposition condition in the definition of patching data (Definition 4.1) in our setup.

Lemma 13.1. *Let $(M, \|\ \|)$ be a complete normed ring and let $0 < \varepsilon < 1$. Consider elements $a_1, a_2, a_3, \ldots \in M$ such that $\|a_i\| \leq \varepsilon$ for each i and $\|a_i\| \to 0$. Let*

$$p_i = (1-a_1)\cdots(1-a_i), \qquad i = 1,2,3,\ldots \quad .$$

Then, the sequence $(p_i)_{i=1}^\infty$ converges to an element of $M^\times$.

Proof. For each $i \geq 1$ we have $\|p_i\| \leq \|1-a_1\|\cdots\|1-a_i\| \leq 1$. Setting $p_0 = 1$, we also have $p_i = p_{i-1}(1-a_i)$. Hence,

$$\|p_i - p_{i-1}\| \leq \|p_{i-1}\| \cdot \|a_i\| \leq \|a_i\| \to 0.$$

Thus, $(p_i)_{i=1}^\infty$ is a Cauchy sequence, so it converges to some $p \in M$. Furthermore,

$$\|p_k - 1\| = \|\sum_{i=1}^k (p_i - p_{i-1})\| \leq \max\|a_i\| \leq \varepsilon.$$

Consequently, $\|p-1\| < 1$. By Lemma 6.3(f), $p \in M^\times$. □

Lemma 13.2 (Cartan's Lemma). *Let $(M, \|\ \|)$ be a complete normed ring. Let M_1 and M_2 be complete subrings of M. Suppose*

(II) for each $a \in M$ there are $a^+ \in M_1$ and $a^- \in M_2$ with $\|a^+\|, \|a^-\| \leq \|a\|$ such that $a = a^+ + a^-$.

Then for each $b \in M$ with $\|b-1\| < 1$ there exist $b_1 \in M_1^\times$ and $b_2 \in M_2^\times$ such that $b = b_1 b_2$.

Proof. Let $a_1 = b - 1$ and $\varepsilon = \|a_1\|$. Then $0 \le \varepsilon < 1$. The condition

$$(13.1) \qquad 1 + a_{j+1} = (1 - a_j^+)(1 + a_j)(1 - a_j^-),$$

with a_j^+, a_j^- associated to a_j by (II), recursively defines a sequence $(a_j)_{j=1}^\infty$ in M. Use the relation $a_j = a_j^+ + a_j^-$ to rewrite (13.1):

$$(13.2) \qquad a_{j+1} = a_j^+ a_j^- - a_j^+ a_j - a_j a_j^- + a_j^+ a_j a_j^-.$$

Inductively assume that $\|a_j\| \le \varepsilon^{2^{j-1}}$. Since $\|a_j^+\|, \|a_j^-\| \le \|a_j\|$, (13.2) implies that $\|a_{j+1}\| \le \max(\|a_j\|^2, \|a_j\|^3) = \|a_j\|^2 \le \varepsilon^{2^j}$. Therefore, $a_j \to 0$, $a_j^- \to 0$, and $a_j^+ \to 0$. Further, by (13.1),

$$(13.3) \qquad 1 + a_{j+1} = (1 - a_j^+) \cdots (1 - a_1^+)\ b\ (1 - a_1^-) \cdots (1 - a_j^-).$$

By Lemma 13.1, the partial products $(1 - a_1^-) \cdots (1 - a_j^-)$ converge to some $b_2' \in M_2^\times$. Similarly, the partial products $(1 - a_j^+) \cdots (1 - a_1^+)$ converge to some $b_1' \in M_1^\times$. Passing to the limit in (13.3), we get $1 = b_1' b b_2'$. Therefore, $b = (b_1')^{-1}(b_2')^{-1}$, as desired. □

Lemma 13.3. *Let A be a complete integral domain with respect to an absolute value $|\ |$, A_1, A_2 complete subrings of A, and A_0 a dense subring of A. Set $E_i = \mathrm{Quot}(A_i)$ for $i = 0, 1, 2$ and $E = \mathrm{Quot}(A)$. Suppose these objects satisfy the following conditions:*

(13.4a) For each $a \in A$ there are $a^+ \in A_1$ and $a^- \in A_2$ with $|a^+|, |a^-| \le |a|$ such that $a = a^+ + a^-$.

(13.4b) $A = A_0 + gA$ for each nonzero $g \in A_0$.

(13.4c) For every $f \in A$ there are $p \in A_0$ and $u \in A^\times$ such that $f = pu$.

(13.4d) $E_0 \subseteq E_2$.

Then, for every positive integer n and for each $b \in \mathrm{GL}_n(E)$ there are $b_1 \in \mathrm{GL}_n(E_1)$ and $b_2 \in \mathrm{GL}_n(E_2)$ such that $b = b_1 b_2$.

Proof. As in Example 6.4(d), we define the norm of a matrix $a = (a_{ij}) \in \mathrm{M}_n(A)$ by $\|a\| = \max_{ij} |a_{ij}|$ and note that $\mathrm{M}_n(A)$ is a complete normed ring, $\mathrm{M}_n(A_1), \mathrm{M}_n(A_2)$ are complete normed subrings of $\mathrm{M}_n(A)$, and $\mathrm{M}_n(A_0)$ is a dense subring of $\mathrm{M}_n(A)$. Moreover, by (13.4a), for each $a \in \mathrm{M}_n(A)$ there are $a^+ \in \mathrm{M}_n(A_1)$ and $a^- \in \mathrm{M}_n(A_2)$ with $\|a^+\|, \|a^-\| \le \|a\|$ such that $a = a^+ + a^-$.

By Condition (13.4c) each element of E is of the form $\frac{1}{h} f$, where $f \in A$ and $h \in A_0$, $h \ne 0$. Hence, there is $h \in A_0$ such that $hb \in \mathrm{M}_n(A)$ and $h \ne 0$. If

$hb = b_1b_2'$, where $b_1 \in \mathrm{GL}_n(E_1)$ and $b_2' \in \mathrm{GL}_n(E_2)$, then $b = b_1b_2$ with $b_2 = \frac{1}{h}b_2' \in \mathrm{GL}_n(E_2)$. Thus, we may assume that $b \in \mathrm{M}_n(A)$.

Let $d \in A$ be the determinant of b. By Condition (13.4c) there are $g \in A_0$ and $u \in A^\times$ such that $d = gu$. Let $b'' \in \mathrm{M}_n(A)$ be the adjoint matrix of b, so that $bb'' = d \cdot 1$, where 1 is here the unit of $M_n(A)$. Let $b' = u^{-1}b''$. Then $b' \in \mathrm{M}_n(A)$ and $bb' = g \cdot 1$.

We set

$$V = \{a' \in \mathrm{M}_n(A) \mid ba' \in g\mathrm{M}_n(A)\} \qquad \text{and} \qquad V_0 = V \cap \mathrm{M}_n(A_0).$$

Then V is an additive subgroup of $\mathrm{M}_n(A)$ and $g\mathrm{M}_n(A) \le V$. By (13.4b), $\mathrm{M}_n(A) = \mathrm{M}_n(A_0) + g\mathrm{M}_n(A)$. Hence $V = V_0 + g\mathrm{M}_n(A)$. Since $\mathrm{M}_n(A_0)$ is dense in $\mathrm{M}_n(A)$, the set $g\mathrm{M}_n(A_0)$ is dense in $g\mathrm{M}_n(A)$. Therefore, $V_0 = V_0 + g\mathrm{M}_n(A_0)$ is dense in $V = V_0 + g\mathrm{M}_n(A)$. Since $b' \in V$, there is $a_0 \in V_0$ such that $||b' - a_0|| < \frac{|g|}{||b||}$. In particular, $a_0 \in \mathrm{M}_n(A_0)$ and $ba_0 \in g\mathrm{M}_n(A)$.

Put $a = \frac{1}{g}a_0 \in \mathrm{M}_n(E_0)$. Then $ba \in \mathrm{M}_n(A)$ and $||1 - ba|| = ||\frac{1}{g}b(b' - a_0)|| \le \frac{1}{|g|}||b|| \cdot ||b' - a_0|| < 1$, where the first occurrence of 1 is the identity element of $M_n(A)$. It follows from Lemma 6.3(f) that $ba \in \mathrm{GL}_n(A)$. In particular $\det(a) \neq 0$ and therefore $a \in \mathrm{GL}_n(E_0) \le \mathrm{GL}(E_2)$. By Lemma 13.2, there are $b_1 \in \mathrm{GL}_n(A_1)$ and $b_2' \in \mathrm{GL}_n(A_2) \le \mathrm{GL}_n(E_2)$ such that $ba = b_1b_2'$. Thus $b = b_1b_2$, where $b_1 \in \mathrm{GL}_n(A_1) \le \mathrm{GL}_n(E_1)$ and $b_2 = b_2'a^{-1} \in \mathrm{GL}_n(E_2)$. □

We apply Corollary 13.3 to the rings and fields of Section 12.

Lemma 13.4. *Let $B \in \mathrm{GL}_n(Q)$.*

(a) For each partition $I = J \cup J'$ there exist $B_1 \in \mathrm{GL}_n(P_J)$ and $B_2 \in \mathrm{GL}_n(P_{J'})$ such that $B = B_1B_2$.

(b) For each $i \in I$ there exist $B_1 \in \mathrm{GL}_n(P_i)$ and $B_2 \in \mathrm{GL}_n(P_i')$ such that $B = B_1B_2$.

Proof. Assertion (b) of the lemma is a special case of Assertion (a). Thus, it suffices to prove Assertion (a).

We may assume without loss that both J and J' are nonempty and apply Lemma 13.3 to the rings $R, R_J, R_{J'}, R_0$ rather than A, A_1, A_2, A_0, where $R_0 = K[w_i \mid i \in I]$.

By definition, R, R_J, and $R_{J'}$ are complete rings (Second paragraph of Section 11). Given $f \in R$, say, $f = a_0 + \sum_{i\in I}\sum_{k=1}^{\infty} a_{ik}w_i^k$ (Lemma 11.1), we let $f_1 = a_0 + \sum_{i\in J}\sum_{k=1}^{\infty} a_{ik}w_i^k$ and $f_2 = \sum_{i\in J'}\sum_{k=1}^{\infty} a_{ik}w_i^k$. Then $|f_i| \le |f|$, $i = 1, 2$ and $f = f_1 + f_2$. This proves Condition (13.4a) in our context.

By definition, R is the completion of R_0, so R_0 is dense in R and $K(x) = \mathrm{Quot}(R_0)$ is contained in both $P_J = \mathrm{Quot}(R_j)$ and $P_{J'} = \mathrm{Quot}(R_{J'})$. Conditions (13.4b) and (13.4c) follow from Corollary 11.7 and Lemma 11.6, respectively. Our Corollary is therefore a special case of Lemma 13.3. □

We apply Corollary 12.2 and Lemma 13.4 to put together patching data whose analytic fields are the fields P_i introduced above.

Proposition 13.5. *Let K be a complete field with respect to an ultrametric absolute value $|\ |$. Let x be an indeterminate, G a finite group, r an element of $K^\times$, and I a finite set with $|I| \geq 2$. For each $i \in I$ let G_i be a subgroup of G, F_i a finite Galois extension of $E = K(x)$ with $\mathrm{Gal}(F_i/K) \cong G_i$, and $c_i \in K^\times$ such that $|r| \leq |c_i - c_j|$ if $i \neq j$. Set $w_i = \frac{r}{x-c_i}$, $P_i = \mathrm{Quot}(K\{w_j \mid j \in I \smallsetminus \{i\}\})$, $P_i' = \mathrm{Quot}(K\{w_i\})$, and $Q = \mathrm{Quot}(K\{w_i \mid i \in I\})$. Suppose $G = \langle G_i \mid i \in I\rangle$ and $F_i \subseteq P_i'$ for each $i \in I$. Then $\mathcal{E} = (E, F_i, P_i, Q, G_i, G)_{i\in I}$ is patching data.*

Proof. Our assumptions imply conditions (4.1a) and (4.1d) of Definition 4.1. By Corollary 12.2, $P_i' = P_{\{i\}} = \bigcap_{j\neq i} P_{I \smallsetminus \{j\}} = \bigcap_{j \neq i} P_j$ and $\bigcap_{i\in I} P_i = E$. Thus, Conditions (4.1b) and (4.1c) of Definition 4.1 hold. Finally, Condition (4.1e) of Definition 4.1 holds by Lemma 13.4. It follows that $\mathcal{E}$ is patching data. □

14 Cyclic Extensions

Every finite group is generated by cyclic groups whose orders are powers of prime numbers. Given a field K, a variable x, and a power q of a prime number, we construct a Galois extension F of $K(x)$ with $\mathrm{Gal}(F/K(x)) \cong \mathbb{Z}/q\mathbb{Z}$. If in addition K is complete with respect to a non-archimedean norm, we show how to embed F into $K\{x\}$.

Lemma 14.1. *Let K be a field, n a positive integer with $\mathrm{char}(K) \nmid n$, and x a variable. Then $K(x)$ has a cyclic extension F of degree n which is contained in $K((x))$.*

Proof. Choose a root of unity ζ_n of order n in K_s. Let $L = K(\zeta_n)$ and $G = \mathrm{Gal}(L/K)$. Consider the map $\chi\colon G \to \{1, \ldots, n-1\}$ such that $\sigma(\zeta_n) = \zeta_n^{\chi(\sigma)}$. Then $\gcd(\chi(\sigma), n) = 1$ and

$$\chi(\sigma\tau) \equiv \chi(\sigma)\chi(\tau) \mod n \tag{14.1}$$

for all $\sigma, \tau \in G$. By [9, Example 3.5.1], $K((x))$ is a regular extension of K and $L((x)) = K((x))(\zeta_n)$. Thus, we may identify G with $\mathrm{Gal}(L((x))/K((x)))$.

Consider the element

$$g(x) = \prod_{\sigma\in G} \left(1 + \sigma(\zeta_n)x\right)^{\chi(\sigma^{-1})}$$

of $L[x]$. Since $\mathrm{char}(K) \nmid n$, Hensel's lemma [9, Prop. 3.5.2] gives a $z \in L[[x]]$ with $z^n = 1 + \zeta_n x$. Then $y = \prod_{\sigma\in G} \sigma(z)^{\chi(\sigma^{-1})} \in L[[x]]$ and $y^n = \prod_{\sigma \in G} \sigma(z^n)^{\chi(\sigma^{-1})} = \prod_{\sigma\in G}(1 + \sigma(\zeta_n)x)^{\chi(\sigma^{-1})} = g(x)$. Since $\zeta_n \in L$, $F = L(x, y)$ is a cyclic extension of degree d of $L(x)$, where $d|n$ and $y^d \in L(x)$ [14, p. 289, Thm. 6.2(ii)]. The linear factors of $g(x)$ are not associate in $K[x]$. In particular, $(1+x)|g(x)$ and $(1+x)^2 \nmid g(x)$. Hence, by Eisensteins's criterion, $Y^n - g(x)$ is irreducible over $L(x)$. Therefore, $F/L(x)$ is a cyclic extension of degree n. Thus, the Galois group $\mathrm{Gal}(F/L(x))$ is generated by an element ω satisfying $\omega(y) = \zeta_n y$.

By (14.1), there exists for each $\tau, \rho \in G$ a positive integer $k(\tau, \rho)$ and a polynomial $f_\tau(x) \in L[x]$ such that

$$\begin{aligned}\tau(y) = \prod_{\sigma\in G} \tau\sigma(z)^{\chi(\sigma^{-1})} = \prod_{\rho\in G} \rho(z)^{\chi(\rho^{-1}\tau)} &= \prod_{\rho\in G} \rho(z)^{\chi(\rho^{-1})\chi(\tau)+k(\tau,\rho)n} \\ &= y^{\chi(\tau)} \prod_{\rho\in G} (1+\rho(\zeta_n)x)^{k(\tau,\rho)} = y^{\chi(\tau)} f_\tau(x).\end{aligned}$$

It follows that G leaves F invariant. Let E be the fixed field of G in F.

$$\begin{tikzcd}
K((x)) \arrow[r, dash] \arrow[d, dash] & L((x)) \arrow[d, dash] \\
E \arrow[r, dash] \arrow[d, dash] & F = L(x,y) \arrow[d, dash] \\
K(x) \arrow[r, dash] \arrow[d, dash] & L(x) \arrow[d, dash] \\
K \arrow[r, dash] & L = K(\zeta_n)
\end{tikzcd}$$

Denote the subgroup of $\mathrm{Aut}(F/K(x))$ generated by G and $\mathrm{Gal}(F/L(x))$ by H. Then the fixed field of H is $K(x)$, so $F/K(x)$ is a Galois extension with $\mathrm{Gal}(F/K(x)) = G \cdot \mathrm{Gal}(F/L(x))$. Moreover, given $\tau \in G$, put $m = \chi(\tau)$. Then $\tau\omega(y) = \tau(\zeta_n y) = \zeta_n^m y^m f_\tau(x) = \omega(y)^m f_\tau(x) = \omega(y^m f_\tau(x)) = \omega\tau(y)$. Thus, $\tau\omega = \omega\tau$, so G commutes with $\mathrm{Gal}(F/L(x))$. Therefore, $E/K(x)$ is a Galois extension with $\mathrm{Gal}(E/K(x)) \cong \mathrm{Gal}(F/L(x)) \cong \mathbb{Z}/n\mathbb{Z}$. □

Lemma 14.2. *Let A be a principal ideal domain, $E = \mathrm{Quot}(A)$, F a finite separable extension of E, and B the integral closure of A in F. Suppose B is unramified over A. Then there exists $b \in B$ such that $\mathrm{trace}_{F/E}(b) = 1$.*

Proof. We write trace for $\mathrm{trace}_{F/E}$ and notice that the set

$$B' = \{x \in F \mid \mathrm{trace}(xB) \subseteq A\}$$

is a fractional ideal of B that contains B [13, p. 58, Cor.]. Hence, $D = (B')^{-1}$ is an ideal of B (called the **different** of B over A). Moreover, a prime ideal P of B is ramified over A if and only if P divides D [13, p. 62, Prop. 8]. Since, by assumption, no P is ramified over A, we have $D = B$, so $B' = B$.

Next observe that since trace: $F \to E$ is an E-linear map, $\mathrm{trace}(B)$ is an ideal of A, hence there exists $a \in A$ with $\mathrm{trace}(B) = aA$. Since F/E is separable, the map trace: $F \to E$ is nonzero [14, p. 286, Thm. 5.2], so $a \neq 0$. Therefore, $\mathrm{trace}(a^{-1}B) = A$, hence $a^{-1} \in B' = B$. It follows that $a^{-1}B = B$, so $\mathrm{trace}(B) = A$. Consequently, there exists $b \in B$ with $\mathrm{trace}(b) = 1$. □

Lemma 14.3. *Suppose $p = \mathrm{char}(K) > 0$. Let F be a cyclic extension of $K(x)$ of degree p^n in $K((x))$, $n \geq 1$. Suppose the integral closure O of $K[x]$ in F is unramified over $K[x]$. Then $K(x)$ has a cyclic extension F' of degree p^{n+1} in $K((x))$ which contains F such that the integral closure O' of $K[x]$ in F' is unramified over $K[x]$.*

Proof. We define F' to be $F(z)$, where z is a zero of $Z^p - Z - a$ with a suitable element $a \in O$. The three parts of the proof produce a, and then show that F' has the desired properties.

- Part A: *Construction of a and z.* We apply Lemma 14.2 to choose $b \in O$ with $\mathrm{trace}_{F/K(x)}(b) = 1$. Then we set $c = b - b^p$ and notice that

  $$\mathrm{trace}_{F/K(x)}(c) = \mathrm{trace}_{F/K(x)}(b) - (\mathrm{trace}_{F/K(x)}(b))^p = 0.$$

  Let σ be a generator of $\mathrm{Gal}(F/K(x))$. Set $q = p^n$ and

  $$a_1 = \sum_{i=1}^{q-1} \sum_{j=0}^{i-1} b^{\sigma^i} c^{\sigma^j}.$$

  Then $a_1 \in O$ and

  $$a_1^\sigma = \sum_{i=1}^{q-1} \sum_{j=0}^{i-1} b^{\sigma^{i+1}} c^{\sigma^{j+1}} = \sum_{i=2}^{q} \sum_{j=1}^{i-1} b^{\sigma^i} c^{\sigma^j}.$$

  Hence,

  $$\begin{aligned} a_1 - a_1^\sigma &= b^\sigma c + b^{\sigma^2} c + \cdots + b^{\sigma^{q-1}} c - b^{\sigma^q} \sum_{j=1}^{q-1} c^{\sigma^j} \\ &= \sum_{i=0}^{q-1} b^{\sigma^i} c - b \sum_{j=0}^{q-1} c^{\sigma^j} = \mathrm{trace}_{F/K(x)}(b)c - b \cdot \mathrm{trace}_{F/K(x)}(c) = c. \end{aligned}$$

  Now note that O is integral over $K[x]$ and is contained in $K((x))$, so $O \subseteq K[[x]]$ (because $K[[x]]$ is integrally closed), in particular $a_1 \in K[[x]]$. Let v be the K-valuation of $K((x))$ with $v(x) = 1$. Since $K[x]$ is v-dense in $K[[x]]$, there is an $a_0 \in K[x]$ with $v(a_1 - a_0) > 0$. Set $a = a_1 - a_0$. Then $a \in O$, $v(a) > 0$, and

  $$a^\sigma - a = b^p - b. \tag{14.2}$$

  Thus, the polynomial $f(Z) = Z^p - Z - a$ satisfies $v(f(0)) = v(-a) > 0$ and $v(f'(0)) = v(-1) = 0$. By Hensel's lemma for $K((x))$, there exists $z \in K[[x]]$ such that $z^p - z - a = 0$.

- Part B: *Irreducibility of $Z^p - Z - a$.* Assume $Z^p - Z - a$ is reducible over F. Then $z \in F$ [14, p. 290, Thm. 6.4(b)]. By (14.2),

$$\begin{aligned}(z^\sigma - z)^p - (z^\sigma - z) - (b^p - b) &= (z^\sigma - z)^p - (z^\sigma - z) - (a^\sigma - a) \\ &= (z^p - z - a)^\sigma - (z^p - z - a) = 0.\end{aligned} \tag{14.3}$$

  Since b is a root of $Z^p - Z - (b^p - b)$, there is an integer i with $z^\sigma - z = b + i$ [14, p. 290, Thm. 6.4(b)]. Apply $\mathrm{trace}_{F/K(x)}$ to both sides to get 0 on the left and 1 on the right. This contradiction proves that $Z^p - Z - a$ is irreducible.

  It follows that $f = \mathrm{irr}(z, F) \in O[Z]$, $f(z) = 0$, and $f'(z) = -1$, in particular $\mathrm{discr}(f)$ is a unit of O [9, p. 109]. Hence, $O' = O[z]$ is the integral closure of O in $F' = F(z)$ and O'/O is a ring cover in the sense of [9, Def. 6.1.3]. In particular, no prime ideal of O is ramified in O' [9, Lemma 6.1.8(b)]. It follows from our assumption on O that O' is unramified over $K[x]$.

- Part C: *Extension of σ to σ' that maps z to $z + b$.* Equality (14.2) implies that $z + b$ is a zero of $Z^p - Z - a^\sigma$. Thus, by Part B, σ extends to an automorphism σ' of F' with $z^{\sigma'} = z + b$. We need only to prove that the order of σ' is p^{n+1}. Induction shows $z^{(\sigma')^j} = z + b + b^\sigma + \cdots + b^{\sigma^{j-1}}$ for each $j \geq 1$. In particular,

$$z^{(\sigma')^q} = z + \mathrm{trace}_{F/K(x)}(b) = z + 1.$$

  Hence, $z^{(\sigma')^{iq}} = z + i$ for $i = 1, \dots, p$, so the order of σ' is p^{n+1}. Consequently, F' is a cyclic extension of $K(x)$ of order p^{n+1}.

□

We can do even better, if K is a complete field under an absolute value $|\ |$.

Lemma 14.4. *Let K be a complete field under an absolute value $|\ |$, let x be a variable, and let q be a power of a prime number. Then $K(x)$ has a Galois extension in $K\{x\}$ in $K((x))$ with Galois group $\mathbb{Z}/q\mathbb{Z}$.*

Proof. By Lemmas 14.1 and 14.3, $K(x)$ has a Galois extension F in $K((x))$ with Galois group isomorphic to $\mathbb{Z}/q\mathbb{Z}$. We choose a primitive element y for $F/K(x)$ integral over $K[x]$. Then $y \in K[[x]]$, so $y = \sum_{n=0}^\infty a_n x^n$ with $a_n \in K$ for each $n \geq 0$. By Proposition 8.5, y converges at some $c \in K^\times$. Thus, the series $\sum_{n=0}^\infty a_n c^n$ converges in K, which means that $y' = \sum_{n=0}^\infty a_n c^n x^n \in K\{x\}$. Now, the map $x \to cx$ extends to an automorphism ϕ of $K((x))$ that leaves $K(x)$ invariant. It maps $K(x, y)$ onto the subfield $K(x, y')$ of $K\{x\}$. Since $K(x, y)/K$ is Galois with Galois group $\mathbb{Z}/q\mathbb{Z}$, so is the extension $K(x, y')/K$, as desired. □

15 Embedding Problems over Complete Fields

Let K/K_0 be a finite Galois extension of fields with Galois group Γ acting on a finite group G. Consider a variable x and set $E_0 = K_0(x)$ and $E = K(x)$. Then E/E_0 is a Galois extension and we identify $\mathrm{Gal}(E/E_0)$ with $\Gamma = \mathrm{Gal}(K/K_0)$ via restriction. We refer to

$$\mathrm{pr}\colon \Gamma \ltimes G \to \Gamma \tag{15.1}$$

as a **constant finite split embedding problem over** E_0. We prove that if K_0 is complete under an ultrametric absolute value, then (15.1) has a solution field (Section 5) equipped with a K-rational place.

Proposition 15.1. *Let K_0 be a complete field with respect to an ultrametric absolute value $|\ |$. Let K/K_0 be a finite Galois extension with Galois group Γ acting on a finite group G from the right. Then E has a Galois extension F such that*

(15.2a) F/E_0 is Galois;

(15.2b) there is an isomorphism $\psi\colon \mathrm{Gal}(F/E_0) \to \Gamma \ltimes G$ such that $\mathrm{pr}\circ\psi = \mathrm{res}_E$; and

(15.2c) F has a set of cardinality $|K_0|$ of K-rational places ϕ (so F/K is regular) such that $\phi(x) \in K_0$ and $\bar{F}_\phi = K$.

Proof. Our strategy is to attach patching data $\mathcal{E}$ to the embedding problem and to define a proper action of Γ on $\mathcal{E}$. Then we apply Proposition 5.2 to conclude that the compound F of $\mathcal{E}$ gives a solution to the embedding problem.

We fix a finite set I on which Γ acts from the right and a system of generators $\{\tau_i \mid i \in I\}$ of G such that for each $i \in I$

(15.3a) $\{\gamma \in \Gamma \mid i^\gamma = i\} = \{1\}$;

(15.3b) the order of the group $G_i = \langle \tau_i \rangle$ is a power of a prime number;

(15.3c) $\tau_i^\gamma = \tau_{i^\gamma}$, for every $\gamma \in \Gamma$; and

(15.3d) $|I| \geq 2$.

(E.g. assuming $G \neq 1$, let G_0 be the set of all elements of G whose order is a power of a prime number and note that Γ leaves G_0 invariant. Let $I = G_0 \times \Gamma$ and for each $(\sigma, \gamma) \in I$ and $\gamma' \in \Gamma$ let $(\sigma,\gamma)^{\gamma'} = (\sigma, \gamma\gamma')$ and $\tau_{(\sigma,\gamma)} = \sigma^\gamma$.)

Then $G_i^\gamma = G_{i^\gamma}$ for each $\gamma \in \Gamma$ and $G = \langle G_i \mid i \in I\rangle$. Choose a system of representatives J for the Γ-orbits of I. Then every $i \in I$ can be uniquely written as $i = j^\gamma$ with $j \in J$ and $\gamma \in \Gamma$.

- Claim A: *There exists a subset $\{c_i \mid i \in I\} \subseteq K$ such that $c_i^\gamma = c_{i^\gamma}$ and $c_i \neq c_j$ for all distinct $i, j \in I$ and $\gamma \in \Gamma$.*

  Indeed, it suffices to find $\{c_j \mid j \in J\} \subseteq K$ such that $c_j^\delta \neq c_j^\varepsilon$ for all $j \in J$ and all distinct $\delta, \varepsilon \in \Gamma$, and $c_j^\delta \neq c_k$ for all distinct $j, k \in J$ and all $\delta \in \Gamma$. Having done that, we can define c_i, for $i = j^\gamma \in I$, as c_j^γ.

  The first condition says that c_j is a primitive element for K/K_0; the second condition means that distinct c_j and c_k are not conjugate over K_0. Thus it suffices to show that there are infinitely many primitive elements for K/K_0. But if $c \in K^\times$ is primitive, then so is $c + a$, for each $a \in K_0$. Since K_0 is complete, hence infinite, the claim follows.

- Construction B: *Patching data.*

  We choose $r \in K_0^\times$ such that $|r| \leq |c_i - c_j|$ for all distinct $i, j \in I$. For each $i \in I$ we set $w_i = \frac{r}{x - c_i} \in K(x)$. As in Section 11, we consider the ring $R = K\{w_i \mid i \in I\}$ and let $Q = \mathrm{Quot}(R)$. For each $i \in I$ let

  $$P_i = P_{I \smallsetminus \{i\}} = \mathrm{Quot}(K\{w_j \mid j \neq i\}) \quad \text{and} \quad P_i' = P_{\{i\}} = \mathrm{Quot}(K\{w_i\})$$

  (we use the notation of Section 12).

  Let $\gamma \in \Gamma$. By our definition, $w_i^\gamma = \frac{r}{x - c_i^\gamma} = w_{i^\gamma}$, $i \in I$. Hence, γ leaves $R_0 = K[w_i \mid i \in I]$ invariant. Since $|\ |$ is complete on K_0, it has a unique extension to K, so $|a^\gamma| = |a|$ for each $a \in K$. Moreover, for each $f = a_0 + \sum_{i \in I} \sum_{n=1}^\infty a_{in} w_i^n \in R_0$, we have

  $$f^\gamma = a_0^\gamma + \sum_{i \in I} \sum_{n=1}^\infty a_{in}^\gamma (w_i^\gamma)^n \tag{15.4}$$

  and

  $$\begin{aligned} \|f^\gamma\| &= \|a_0^\gamma + \sum_{i \in I} \sum_{n=1}^\infty a_{in}^\gamma (w_i^\gamma)^n\| = \|a_0^\gamma + \sum_{i \in I} \sum_{n=1}^\infty a_{in}^\gamma w_{i^\gamma}^n\| \\ &= \max(|a_0^\gamma|, |a_{in}^\gamma|)_{i,n} = \max(|a_0|, |a_{in}|)_{i,n} = \|f\|. \end{aligned}$$

  By Lemma 6.5, γ uniquely extends to a continuous automorphism of the completion R of R_0, by formula (15.4) for $f \in R$. Hence, Γ lifts to a group of continuous automorphisms of R. Therefore, Γ extends to a group of automorphisms of $Q = \mathrm{Quot}(R)$. In addition, $P_i^\gamma = P_{i^\gamma}$ and $(P_i')^\gamma = P_{i^\gamma}'$.

  For each $j \in J$, Lemma 14.4 gives a cyclic extension F_j of E in $P_j' = K\{w_j\}$ with Galois group $G_j = \langle \tau_j \rangle$.

  For an arbitrary $i \in I$ there exist unique $j \in J$ and $\gamma \in \Gamma$ such that $i = j^\gamma$ (by (15.3a)). Since γ acts on Q and leaves E invariant, $F_i = F_j^\gamma$ is a Galois extension of E.

The isomorphism $\gamma\colon F_j \to F_i$ gives an isomorphism

$$\mathrm{Gal}(F_j/E) \cong \mathrm{Gal}(F_i/E)$$

that maps each $\tau \in \mathrm{Gal}(F_j/E)$ onto $\gamma^{-1} \circ \tau \circ \gamma \in \mathrm{Gal}(F_i/E)$ (notice that the elements of the Galois groups act from the right). In particular, it maps τ_j onto $\gamma^{-1} \circ \tau_j \circ \gamma$. We can therefore identify G_i with $\mathrm{Gal}(F_i/E)$ such that τ_i coincides with $\gamma^{-1} \circ \tau_j \circ \gamma$. This means that $(a^\tau)^\gamma = (a^\gamma)^{\tau^\gamma}$ for all $a \in F_j$ and $\tau \in G_j$. In particular, $F_i \subseteq P_i'$ for each $i \in I$. It follows from Proposition 13.5 that $\mathcal{E} = (E, F_i, P_i, Q; G_i, G)_{i \in I}$ is patching data. By construction, Γ acts properly on $\mathcal{E}$ (Definition 5.1). By Propositions 4.5 and 5.2, the compound F of $\mathcal{E}$ satisfies (15.2a) and (15.2b). Now we verify (15.2c).

- Claim C: *F/K has many prime divisors of degree* 1. Each $b \in K_0$ with

$$(15.5) \qquad |b| > \max_{i \in I}(|r|, |c_i|)$$

satisfies $\left|\frac{r}{b-c_i}\right| < 1$ for each $i \in I$, hence, the map $x \mapsto b$ extends to a homomorphism from R to K that maps w_i onto $\frac{r}{b-c_i}$. Since R is a principal ideal domain (Proposition 11.8), this homomorphism extends to a K-rational place $\phi_b\colon Q \to K \cup \{\infty\}$. Thus, $\phi_b|_F$ is a K-rational place of F with $\phi_b(x) = b \in K_0$, so it corresponds to a prime divisor of F/K of degree 1. If $b' \in K_0$ and $b' \neq b$, then $\phi_b \neq \phi_{b'}$, so also the prime divisors that ϕ_b and $\phi_{b'}$ define are distinct. Consequently, the cardinality of the prime divisors of F/K of degree 1 is that of K_0.

Finally, the regularity of F/K follows from the fact that $\phi_b(F) = K \cup \{\infty\}$ [9, Lemma 2.6.9].

□

16 Embedding Problems over Ample Fields

In this section K/K_0 is an arbitrary finite Galois extension with Galois group Γ and x is a variable. Suppose Γ acts on a finite group G. We look for a rational solution of the embedding problem

$$(16.1) \qquad \mathrm{pr}\colon \mathrm{Gal}(K(x)/K_0(x)) \ltimes G \to \mathrm{Gal}(K(x)/K_0(x))$$

over $K_0(x)$. We call (16.1) a **constant split embedding problem**. When K_0 is complete under an ultrametric absolute value, this problem reduces to the special case solved in Section 15.

Consider also a regular extension $\hat{K}_0$ of K_0 such that x is transcendental over $\hat{K}_0$ and let $\hat{K} = K\hat{K}_0$. Then $\hat{K}_0(x)$ is a regular extension of $K_0(x)$ [9, Lemma 2.6.8(a)], so $\hat{K}_0(x)$ is linearly disjoint from $K(x)$ over $K_0(x)$. Hence, res: $\mathrm{Gal}(\hat{K}(x)/\hat{K}_0(x)) \to \mathrm{Gal}(K(x)/K_0(x))$ is an isomorphism. This gives rise to a finite split embedding problem over $\hat{K}_0(x)$,

$$\mathrm{pr}\colon \mathrm{Gal}(\hat{K}(x)/\hat{K}_0(x)) \ltimes G \to \mathrm{Gal}(\hat{K}(x)/\hat{K}_0(x)) \tag{16.2}$$

such that $\mathrm{pr} \circ (\mathrm{res}_{K(x)} \times \mathrm{id}_G) = \mathrm{res}_{K(x)} \circ \mathrm{pr}$.

We identify each of the groups $\mathrm{Gal}(\hat{K}(x)/\hat{K}_0(x))$, $\mathrm{Gal}(K(x)/K_0(x))$, and $\mathrm{Gal}(\hat{K}/\hat{K}_0)$ with $\Gamma = \mathrm{Gal}(K/K_0)$ via restriction. Moreover, if F (resp. $\hat{F}$) is a solution field of embedding problem (16.1) (resp. (16.2)), then we identify $\mathrm{Gal}(F/K_0(x))$ (resp. $\mathrm{Gal}(\hat{F}/\hat{K}_0(x))$) with $\Gamma \ltimes G$ via an isomorphism θ (resp. $\hat{\theta}$) satisfying $\mathrm{pr}\circ\theta = \mathrm{res}$ (resp. $\mathrm{pr}\circ\hat{\theta} = \mathrm{res}$). We say that (F, θ) is a **split rational solution** of (16.1) if F has a K-rational place ϕ such that Γ is the **decomposition group**

$$D_\phi = \{\sigma \in \mathrm{Gal}(F/K_0(x)) \mid \sigma x = x \text{ for all } x \in F \text{ with } \phi(x) \neq \infty\}$$

of ϕ over $K_0(x)$. We say that (F, θ) is **unramified** if ϕ can be chosen to be unramified over $K_0(x)$.

Lemma 16.1. *In the above notation suppose K_0 is ample and existentially closed in $\hat{K}_0$. Let $\hat{F}$ be a solution field to embedding problem* (16.2) *with a $\hat{K}$-rational place $\hat{\phi}$, unramified over $\hat{K}_0(x)$, such that $\hat{\phi}(x) \in \hat{K}_0$. Then embedding problem* (16.1) *has a solution field F with a K-rational place ϕ unramified over $K_0(x)$ such that $\phi(x) \in K_0$.*

Proof. We break up the proof into several parts. First we solve embedding problem (16.1) over $\hat{K}_0(x)$, then we push the solution down to a solution over a function field $K_0(\mathbf{u}, x)$ which is regular over K_0, and finally we specialize the latter solution to a solution over $K_0(x)$ with a place satisfying all of the prescribed conditions.

- Part A: *A solution of* (16.1) *over $\hat{K}_0(x)$.* By assumption, there exists an isomorphism

  $$\hat{\theta}\colon \mathrm{Gal}(\hat{F}/\hat{K}_0(x)) \to \mathrm{Gal}(\hat{K}(x)/\hat{K}_0(x)) \ltimes G$$

  such that $\mathrm{pr} \circ \hat{\theta} = \mathrm{res}_{\hat{K}(x)}$. Let $\hat{F}_0$ be the fixed field in $\hat{F}$ of $D_{\hat{\phi}}$ $(= \Gamma)$. Then, $\hat{F}_0 \cap \hat{K}(x) = \hat{K}_0(x)$ and $\hat{F}_0 \cdot \hat{K}(x) = \hat{F}$, so $[\hat{F}_0 : \hat{K}_0(x)] = [\hat{F} : \hat{K}(x)]$. Then, $\hat{\phi}(\hat{F}_0) = \hat{K}_0 \cup \{\infty\}$. Hence, $\hat{F}_0/\hat{K}_0$ is regular [9, Lemma 2.6.9(b)].

  We choose a primitive element y for the extension $\hat{F}_0/\hat{K}_0(x)$ integral over $\hat{K}_0[x]$. By the preceding paragraph, $\hat{F} = \hat{K}(x, y)$.

  By [12, Lemma 5.1.2], there exists an absolutely irreducible polynomial $h \in \hat{K}_0[V, W]$ and elements $v, w \in \hat{F}_0$ such that $\hat{K}_0(v, w) = \hat{F}_0$, $h(v, w) = 0$, $h(0, 0) = 0$, and $\frac{\partial h}{\partial W}(0, 0) \neq 0$.

We also choose a primitive element c for K over K_0, a primitive element z for $\hat{F}$ over $\hat{K}_0(x)$ integral over $\hat{K}_0[x]$, and note that $\hat{F} = \hat{K}_0(c, x, y)$. Then there exist polynomials $f, p_0, p_1 \in \hat{K}_0[X, Z]$, $g, r_0, r_1, r_2 \in \hat{K}_0[X, Y]$, $q_0, q_1 \in \hat{K}_0[T, X, Y]$, and $s_0, s_1, s_2 \in \hat{K}_0[V, W]$ such that the following conditions hold:

(16.3a) $\hat{F} = \hat{K}_0(x, z)$ and $f(x, Z) = \mathrm{irr}(z, \hat{K}_0(x))$;
in particular $\mathrm{discr}(f(x, Z)) \in \hat{K}_0(x)^\times$.

(16.3b) $g(x, Y) = \mathrm{irr}(y, \hat{K}_0(x)) = \mathrm{irr}(y, \hat{K}(x))$; since $\hat{F}_0/\hat{K}_0$ is regular (by the first paragraph of Part A), $g(X, Y)$ is absolutely irreducible [9, Cor. 10.2.2(b)].

(16.3c) $y = \frac{p_1(x,z)}{p_0(x,z)}$, $z = \frac{q_1(c,x,y)}{q_0(c,x,y)}$, $p_0(x, z) \neq 0$, and $q_0(c, x, y) \neq 0$.

(16.3d) $v = \frac{r_1(x,y)}{r_0(x,y)}$, $w = \frac{r_2(x,y)}{r_0(x,y)}$, $x = \frac{s_1(v,w)}{s_0(v,w)}$, $y = \frac{s_2(v,w)}{s_0(v,w)}$, $r_0(x, y) \neq 0$, and $s_0(v, w) \neq 0$.

- Part B: *Pushing down.* The polynomials introduced in Part A depend on only finitely many parameters from $\hat{K}_0$. Thus, there are $u_1, \ldots, u_n \in \hat{K}_0$ with the following properties:

(16.4a) The coefficients of $f, g, h, p_0, p_1, q_0, q_1, r_0, r_1, r_2, s_0, s_1, s_2$ are in $K_0[\mathbf{u}]$.

(16.4b) $F_{\mathbf{u}} = K_0(\mathbf{u}, x, z)$ is a Galois extension of $K_0(\mathbf{u}, x)$,
$f(x, Z) = \mathrm{irr}(z, K_0(\mathbf{u}, x))$, and $\mathrm{discr}(f(x, Z)) \in K_0(\mathbf{u}, x)^\times$.

(16.4c) $g(x, Y) = \mathrm{irr}(y, K_0(\mathbf{u}, x)) = \mathrm{irr}(y, K(\mathbf{u}, x))$;
we set $F_{0,\mathbf{u}} = K_0(\mathbf{u}, x, y)$.

It follows that restriction maps the groups $\mathrm{Gal}(\hat{F}/\hat{K}_0(x))$, $\mathrm{Gal}(\hat{F}/\hat{F}_0)$, and $\mathrm{Gal}(\hat{F}/\hat{K}(x))$ isomorphically onto $\mathrm{Gal}(F_{\mathbf{u}}/K_0(\mathbf{u}, x))$, $\mathrm{Gal}(F_{\mathbf{u}}/F_{0,\mathbf{u}})$, and $\mathrm{Gal}(F_{\mathbf{u}}/K(\mathbf{u}, x))$, respectively. Therefore, restriction transfers $\hat{\theta}$ to an isomorphism

$$(16.5) \qquad \theta\colon \mathrm{Gal}(F_{\mathbf{u}}/K_0(\mathbf{u}, x)) \to \mathrm{Gal}(K(\mathbf{u}, x)/K_0(\mathbf{u}, x)) \ltimes G$$

satisfying $\mathrm{pr} \circ \theta = \mathrm{res}_{F_{\mathbf{u}}/K(\mathbf{u},x)}$.

Part C: *Specialization.* Since K_0 is existentially closed in $\hat{K}_0$, the field $\hat{K}_0$ and therefore also $K_0(\mathbf{u})$ are regular extensions of K_0 (Lemma 1.5). Thus, $\mathbf{u}$ generates an absolutely irreducible variety $U = \mathrm{Spec}(K_0[\mathbf{u}])$ over K_0 [9, Cor. 10.2.2]. The variety U has a nonempty Zariski-open subset U' that contains $\mathbf{u}$ such that for each $\mathbf{u}' \in U'$ the K_0-specialization $\mathbf{u} \to \mathbf{u}'$ extends to a $K(x)$-homomorphism $'\colon K(x)[\mathbf{u}, v, w, y, z] \to K(x)[\mathbf{u}', v', w', y', z']$ such that the following conditions, derived from (16.3a–d) and (16.4a–c), hold:

(16.6a) The coefficients of $f', g', h', p_0', p_1', q_0', q_1', r_0', r_1', r_2', s_0', s_1', s_2'$ belong to $K_0[\mathbf{u}']$.

(16.6b) $F = K_0(\mathbf{u}', x, z')$ is a Galois extension of $K_0(\mathbf{u}', x)$, $f'(x, z') = 0$, and $\mathrm{discr}(f'(x, Z)) \in K_0(\mathbf{u}', x)^\times$.

(16.6c) $y' = \frac{p_1'(x,z')}{p_0'(x,z')}$, $z' = \frac{q_1'(c,x,y')}{q_0'(c,x,y')}$, $p_0'(x, z') \neq 0$, and $q_0'(c, x, y') \neq 0$; we set $F_0 = K_0(\mathbf{u}', x, y')$ and find that $F = F_0K$.

(16.6d) $g'(X, Y)$ is absolutely irreducible, $\deg_Y(g'(x, Y)) = \deg_Y(g(x, Y))$, $g'(x, y') = 0$, and so $g'(x, Y) = \mathrm{irr}(y', K_0(\mathbf{u}', x)) = \mathrm{irr}(y', K(\mathbf{u}', x))$;

(16.6e) $h'(V, W)$ is absolutely irreducible, $h'(0, 0) = 0$, and $\frac{\partial h'}{\partial W}(0, 0) \neq 0$.

(16.6f) $v' = \frac{r_1'(x,y')}{r_0'(x,y')}$, $w' = \frac{r_2'(x,y')}{r_0'(x,y')}$, $x = \frac{s_1'(v',w')}{s_0'(v',w')}$, $y' = \frac{s_2'(v',w')}{s_0'(v',w')}$, $r_0'(x, y') \neq 0$, and $s_0'(v', w') \neq 0$; thus $F_0 = K_0(\mathbf{u}', v', w')$.

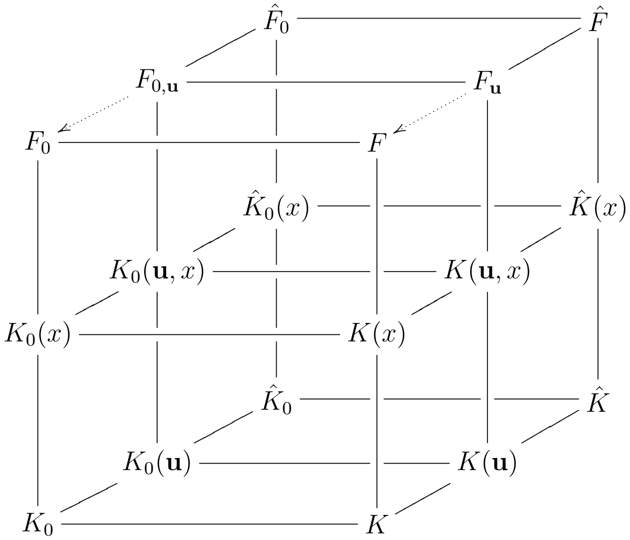
To achieve the absolute irreducibility of g' and h' we have used the Bertini-Noether theorem [9, Prop. 9.4.3].

- Part D: *Choosing* $\mathbf{u}' \in K_0^n$. Since K_0 is existentially closed in $\hat{K}_0$ and since $\mathbf{u} \in U'(\hat{K}_0)$, we can choose $\mathbf{u}' \in U'(K_0)$. Then $K_0[\mathbf{u}'] = K_0$, $K_0(\mathbf{u}', x) = K_0(x)$, $F_0 = K_0(x, y') = K_0(v', w')$, and $F = K_0(x, z')$.

  Since $\mathrm{discr}(f'(x, Z)) \neq 0$ (by (16.6b)) the homomorphism $'$ induces an embedding

$$\psi^* \colon \mathrm{Gal}(F/K_0(x)) \to \mathrm{Gal}(F_{\mathbf{u}}/K_0(\mathbf{u}, x)) \tag{16.7}$$

such that $(\psi^*(\sigma)(s))' = \sigma(s')$ for all $\sigma \in \mathrm{Gal}(F/K_0(x))$ and $s \in F_{\mathbf{u}}$ with $s' \in F$ [14, p. 344, Prop. 2.8]. Each $s \in K(x)$ is fixed by $'$, hence $\psi^*(\sigma)(s) = \sigma(s)$ for each $\sigma \in \mathrm{Gal}(F/K_0(x))$. It follows that ψ^* commutes with restriction to $K(x)$.



By (16.6c), $F = K(x,y') = F_0K$. By (16.6d) and [9, Cor. 10.2.2(b)], F_0/K_0 is a regular extension, so F_0 is linearly disjoint from K over K_0. Therefore, F_0 is linearly disjoint from $K(x)$ over $K_0(x)$, hence $F_0 \cap K(x) = K_0(x)$ and $[F_0 : K_0(x)] = [F : K(x)]$. It follows from (16.6d) that

$$\begin{aligned}
|\mathrm{Gal}(F/K_0(x))| &= [F : K_0(x)] \\
&= [F : K(x)][K(x) : K_0(x)] \\
&= \deg_Y g'(x,Y)[K : K_0] \\
&= \deg_Y g(x,Y)[K : K_0] \\
&= [F_{\mathbf{u}} : K(\mathbf{u},x)][K(\mathbf{u},x) : K_0(\mathbf{u},x)] \\
&= [F_{\mathbf{u}} : K_0(\mathbf{u},x)] = |\mathrm{Gal}(F_{\mathbf{u}}/K_0(\mathbf{u},x))|.
\end{aligned}$$

Therefore ψ^* is an isomorphism. Let

$$\rho\colon \mathrm{Gal}(K(\mathbf{u},x)/K_0(\mathbf{u},x)) \ltimes G \to \mathrm{Gal}(K(x)/K_0(x)) \ltimes G$$

be the isomorphism whose restriction to $\mathrm{Gal}(K(\mathbf{u},x)/K_0(\mathbf{u},x))$ is the restriction map and to G is the identity map. Then, $\theta' = \rho \circ \theta \circ \psi^*$ satisfies $\mathrm{pr} \circ \theta' = \mathrm{res}_{F/K(x)}$ (by (16.5)). This means that θ' is a solution of embedding problem (16.1).

- Part E: *Rational place.* Finally, by (16.6e) and (16.6f), the curve defined by $h'(X,Y) = 0$ is a model of F_0/K_0 and $(0,0)$ is a K_0-rational simple point of it. Therefore, by [12, Lemma 5.1.4(b)], F_0 has a K_0-rational place $\phi_0\colon F_0 \to K_0 \cup \{\infty\}$. Since K_0 is ample, F_0 has infinitely many K_0-places

(Lemma 1.1). Only finitely many of them are ramified over $K_0(x)$. Hence, we may choose ϕ_0 to be unramified over $K_0(x)$. Using the linear disjointness of F_0 and K over K_0, we extend ϕ_0 to a K-rational place $\phi\colon F \to K \cup \{\infty\}$ unramified over $K_0(x)$.

□

Theorem 16.2. *Let K_0 be an ample field. Then every constant finite split embedding problem over $K_0(x)$ has a split unramified rational solution.*

Proof. Consider a constant finite split embedding problem (16.1) over $K_0(x)$. Let $\hat{K}_0 = K_0((t))$. Then $\hat{K}_0$ is complete under a nontrivial discrete ultrametric absolute value with prime element t. Consequently, by Proposition 15.1, (16.2) has a split unramified rational solution. By Lemma 1.1(f), K_0 is existentially closed in $\hat{K}_0$. Hence, by Lemma 16.1, (16.1) has a split unramified rational solution. □

17 PAC Hilbertian Fields are ω-Free

The statement of the title was a major open problem of Field Arithmetic. Theorem 17.3 settles that problem.

Recall that the **rank** of a profinite group G is the least cardinality of a system of generators of G that converges to 1. If G is not finitely generated, then $\mathrm{rank}(G)$ is also the cardinality of the set of all open normal subgroups of G [9, Prop. 17.1.2]. We denote the free profinite group of rank m by $\hat{F}_m$.

An **embedding problem** for a profinite group G is a pair

$$(\phi\colon G \to A,\ \alpha\colon B \to A), \tag{17.1}$$

of homomorphisms of profinite groups with ϕ and α surjective. The embedding problem is said to be **finite** if B is finite. If there exists a homomorphism $\alpha'\colon A \to B$ such that $\alpha \circ \alpha' = \mathrm{id}_A$, we say that (17.1) **splits**. A **weak solution** to (17.1) is a homomorphism $\gamma\colon G \to B$ such that $\alpha \circ \gamma = \phi$. If γ is surjective, we say that γ is a **solution** to (17.1). We say that G is **projective** if every finite embedding problem for G has a weak solution.

An **embedding problem** over a field K is an embedding problem (17.1), where $G = \mathrm{Gal}(K)$. If L is the fixed field of $\mathrm{Ker}(\phi)$, we may identify A with $\mathrm{Gal}(L/K)$ and ϕ with $\mathrm{res}_{K_s/L}$ and then consider $\alpha\colon B \to \mathrm{Gal}(L/K)$ as the given embedding problem. This shows that our present definition generalizes the one given in Section 5. Note that if $\gamma\colon \mathrm{Gal}(K) \to B$ is a solution of (17.1) and F is the fixed field in K_s of $\mathrm{Ker}(\gamma)$, then F is a solution field of the embedding problem $\alpha\colon B \to \mathrm{Gal}(L/K)$ and γ induces an isomorphism $\bar{\gamma}\colon \mathrm{Gal}(F/K) \to B$ such that $\alpha \circ \bar{\gamma} = \mathrm{res}_{F/L}$.

The first statement of the following proposition is due to Gruenberg [9, Lemma 22.3.2], the second one is a result of Iwasawa [9, Cor. 24.8.2].

Proposition 17.1. *Let G be a projective group. If each finite split embedding problem for G is solvable, then every finite embedding problem for G is solvable. If in addition* $\mathrm{rank}(G) \le \aleph_0$, *then* $G \cong \hat{F}_\omega$.

We say that a field K is **ω-free** if every finite embedding problem over K (that is, finite embedding problem for $\mathrm{Gal}(K)$) is solvable.

Theorem 17.2. *Let K be an ample field.*

(a) If K is Hilbertian, then each finite split embedding problem over K is solvable.

(b) If in addition, $\mathrm{Gal}(K)$ *is projective, then K is ω-free.*

(c) If in addition, $\mathrm{Gal}(K)$ *has countably many generators, and in particular, if K is countable, then* $\mathrm{Gal}(K) \cong \hat{F}_\omega$.

Proof.

- *Proof of (a):* Every finite split embedding problem over K gives a finite split constant embedding problem over $K(x)$. The latter is solvable by Theorem 16.2. Now use the Hilbertianity and specialize to get a solution of the original embedding problem over K [9, Lemma 13.1.1].
- *Proof of (b):* By (a), every finite split embedding problem over K is solvable. Hence, by Proposition 17.1, every finite embedding problem over K is solvable.
- *Proof of (c)* Use (b) and Proposition 17.1.

□

The following special case of Theorem 17.2 is a solution of [8, Prob. 24.41].

Theorem 17.3. *Let K be a PAC field. Then K is ω-free if and only if K is Hilbertian.*

Proof. That 'K is ω-free' implies 'K is Hilbertian' is a result of Roquette [9, Cor. 27.3.3]. Conversely, if K is PAC, then $\mathrm{Gal}(K)$ is projective [9, Thm. 11.6.2]. By Example (a) of Section 2, K is ample. Hence, if K is Hilbertian, then by Theorem 17.2(b), K is ω-free. □

References

[1] L. Bary-Soroker, D. Haran, and D. Harbater, *Permanence criteria for semi-free profinite groups,* Mathematische Annalen **348** (2010), 539–583.

[2] J.-L. Colliot-Thélène, *Rational connectedness and Galois covers of the projective line,* Annals of Mathematics **151** (2000), 359–373.

[3] D. Haran and M. Jarden, *Regular split embedding problems over complete valued fields,* Forum Mathematicum **10** (1998), 329–351.

[4] D. Haran and M. Jarden, *The absolute Galois group of C(x),* Pacific Journal of Mathematics **196** (2000), 445–459.

[5] D. Haran and M. Jarden, *Regular split embedding problems over function fields of one variable over ample fields,* Journal of Algebra **208** (1998), 147–164.

[6] D. Harbater, *Fundamental groups and embedding problems in characteristic p,* Contemporary Mathematics **186** (1995), 353–369.

[7] A. Fehm, *Subfields of ample fields, Rational maps and definability.* Journal of Algebra **323** (2010) 1738–1744.

[8] M. D. Fried and M. Jarden, *Field Arithmetic,* Ergebnisse der Mathematik (3) **11**, Springer-Verlag, Heidelberg, 1986.

[9] M. D. Fried and M. Jarden, *Field Arithmetic, Third Edition, revised by Moshe Jarden,* Ergebnisse der Mathematik (3) **11**, Springer, Heidelberg, 2008.

[10] D. Harbater and K. Stevenson, *Local Galois theory in dimension two,* Advances in Mathematics **198** (2005), 623–653.

[11] M. Jarden, *Ample fields,* Archiv der Mathematik **80** (2003), 475–477.

[12] M. Jarden, *Algebraic Patching,* Springer Monographs in Mathematics, Springer 2011.

[13] S. Lang, *Algebraic Number Theory,* Addison-Wesley, Reading, 1970.

[14] S. Lang, *Algebra, Third Edition,* Addison-Wesley, Reading, 1993.

[15] F. Pop, *Étale Galois covers of affine smooth curves. The geometric case of a conjecure of Shafarevich. On Abhyankar's conjecture.* Inventiones mathematicae **120** (1995), 555–578.

[16] F. Pop, *Embedding problems over large fields,* Annals of Mathematics **144** (1996), 1–34.

[17] F. Pop, *Henselian implies large,* Annals of Mathematics **172** (2010), 101–113.

Moshe Jarden
School of Mathematics
Tel-Aviv University
Ramat Aviv, Tel Aviv 69978
Israel
jarden@post.tau.ac.il

Travaux mathématiques, Volume 23 (2013), 139–172, © by the author

Serre's Modularity Conjecture[1]

by Michael M. Schein

Abstract

A classical conjecture of Serre, which was recently proved by Khare, Wintenberger, and Kisin, specifies when a mod p Galois representation $\overline{\rho} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ arises from a modular form, and moreover states the weight and level of a suitable modular form. This expository article surveys the conjecture, its recent generalizations to mod p representations of the absolute Galois group of a totally real field, and the methods of proof both of the classical conjecture and of some results towards the generalizations.

MSC (2010): 11F80.

[1] This work was supported by Grant 2264/2010 from the German-Israeli Foundation for Scientific Research and Development and by a grant from the Pollack Family Foundation.

Contents

1 Introduction

These notes are based on lectures given by the author at the Winter School on Galois Theory held at the University of Luxembourg in February 2012. Their aim is to give an overview of Serre's modularity conjecture and of its proof by Khare, Wintenberger, and Kisin [36] [37] [39], as well as of the results of other mathematicians that played an important role in the proof. Along the way we will remark on some recent work concerning generalizations of the conjecture. We have tried as much as possible to concentrate on giving a broad picture of the structure of the arguments and have deliberately ignored technical details in places. The statements of some theorems omit the full list of technical hypotheses; we request the reader's forbearance.

Let F be a totally real number field. We will denote by G_F the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$. It was shown in Prof. Böckle's lectures in this volume how, under some hypotheses, a Hilbert modular eigenform f over F gives rise to a compatible system $\{\rho_{f,v}\}$ of p-adic Galois representations; see [34] for the most general theorem. These representations are extracted from the cohomology of a suitable algebraic variety, and this construction is more or less the only method we have for obtaining p-adic Galois representations. Therefore the following question is of acute interest: given a Galois representation $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$, when is ρ *modular*, i.e. when does there exist a Hilbert modular eigenform f and a place $v|p$ of F such that $\rho \simeq \rho_{f,v}$?

This question is a very difficult one. We will split it into two questions, which are still very difficult, by introducing the notion of the reduction of a Galois representation. The following result is classical; the proof given here is attributed to N. Katz and appears in print, for instance, at the beginning of section 2 of [55].

Proposition 1.1. *Let G be a compact Hausdorff group, and let $\rho : G \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ be a continuous representation. Then ρ is equivalent to a representation ρ' such that $\rho'(G) \subset \mathrm{GL}_n(\mathcal{O}_L)$, where $\mathcal{O}_L$ is the ring of integers of some finite extension $L/\mathbb{Q}_p$.*

Proof. Since G is compact and Hausdorff, it admits a Haar measure μ; without loss of generality, $\mu(G) = 1$. Now,

$$\mathrm{GL}_n(\overline{\mathbb{Q}}_p) = \bigcup_{[L:\mathbb{Q}_p]<\infty} \mathrm{GL}_n(L)$$

and hence

$$G = \bigcup_{[L:\mathbb{Q}_p]<\infty} \rho^{-1}(\mathrm{GL}_n(L)).$$

Since there are countably many finite extensions $L/\mathbb{Q}_p$, there must be some L such that $\mu(\rho^{-1}(\mathrm{GL}_n(L))) > 0$. Hence, $\rho^{-1}(\mathrm{GL}_n(L)) \subset G$ is a closed subgroup of finite index. Then $\rho^{-1}(\mathrm{GL}_n(\mathcal{O}_L))$ is an open subgroup of the compact group

$\rho^{-1}(\mathrm{GL}_n(L))$, so it has finite index in it and thus in G. Let $g_1, \ldots, g_m$ be a collection of coset representatives for $\rho^{-1}(\mathrm{GL}_n(\mathcal{O}_L))$. Let $\Lambda \subset L^n$ be the lattice generated by $\rho(g_1)\mathcal{O}_L^n, \ldots, \rho(g_m)\mathcal{O}_L^n$. This is a lattice of maximal rank, so $\Lambda \simeq \mathcal{O}_L^n$. Furthermore, Λ is stable under the action of G. Let $T \in \mathrm{GL}_n(L)$ be a linear transformation that takes $\mathcal{O}_L^n$ to Λ, and set $\rho'(g) = T^{-1}\rho(g)T$. □

The proposition above applies, in particular, to continuous representations $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ such as we have been considering. If ρ is equivalent to $\rho' : G_F \to \mathrm{GL}_2(\mathcal{O}_L)$, then we define the reduction $\overline{\rho}$ to be the semisimplification of the composition $\tilde{\rho}' : G_F \xrightarrow{\rho'} \mathrm{GL}_2(\mathcal{O}_L) \to \mathrm{GL}_2(k_L) \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$, where k_L is the residue field of L and the inclusion $k_L \hookrightarrow \overline{\mathbb{F}}_p$ is induced from $L \hookrightarrow \overline{\mathbb{Q}}_p$. In other words, $\overline{\rho}$ is the direct sum of the Jordan-Hölder constituents of $\tilde{\rho}'$. This definition is independent of all choices, and we call $\overline{\rho}$ the reduction modulo p of ρ.

Remark 1.2. Throughout these notes, except for Section 4, we will usually use $\overline{\rho}$ to denote a mod p Galois representation. The bar simply serves to emphasize that we are dealing with a mod p representation. It does not necessarily mean that we have any p-adic representation ρ in mind, of which $\overline{\rho}$ is to be the reduction.

We say that a mod p Galois representation $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is *modular* if it "arises from geometry" in a way that will be made precise in the next section. If $F = \mathbb{Q}$, then $\overline{\rho}$ is modular if and only if there exists a modular eigenform f such that $\overline{\rho} \simeq \overline{\rho_{f,p}}$. For larger totally real fields we will require a somewhat more subtle notion of modularity. If a p-adic Galois representation ρ is modular, then its reduction modulo p will be modular as well. We will consider two questions:

1. Let $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ be a p-adic Galois representation. Suppose that $\overline{\rho}$ is modular. Is ρ modular?
2. Let $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a mod p Galois representation. When is $\overline{\rho}$ modular?

It is clear that if we knew complete answers to both of these questions, their union would resolve the question of when a general p-adic representation is modular. Affirmative responses to the first question are known in a variety of different cases; results of this type are called modularity lifting theorems. A conjectural response to the second question is given by Serre's modularity conjecture and its generalizations. Serre's original conjecture, covering the case of $F = \mathbb{Q}$, is now a theorem of Khare, Wintenberger, and Kisin. However, as we shall see, even if we do not know whether a mod p Galois representation is modular, we can say a lot about the Hilbert modular forms f that it could come from if it were modular. The research towards resolving each of these two questions is tightly interconnected with work concerning the other, as shall become evident in these notes.

Acknowledgements The author is very grateful to Sara Arias de Reyna, Lior Bary-Soroker, and Gabor Wiese, the organizers of the Winter School on Galois Theory held in Luxembourg in February 2012, for inviting him to present these lectures. He is grateful to the audience for their stimulating questions, to Tommaso Centeleghe and Nicolas Billerey for allowing him to make use of the notes they took in his lectures, and to the anonymous referee for a thorough reading. The exposition given here has in places drawn on expositions of similar material given elsewhere, such as the lectures by Toby Gee, Richard Taylor, and Teruyoshi Yoshida at an MSRI workshop on modularity in 2006, by Richard Taylor at the summer school on Serre's conjecture at Luminy in 2007, and those by Fred Diamond at the Galois Trimester at the Institut Henri Poincaré in Paris in 2010. All errors and inaccuracies are, of course, entirely the responsibility of the author.

2 Statement of Serre's modularity conjecture

2.1 The classical conjecture

If we fix an embedding $\overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$, then complex conjugation is a well-defined element $c \in G_{\mathbb{Q}}$. Since it is an involution, any Galois representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ must send c to a matrix with determinant ± 1. We say that $\overline{\rho}$ is *odd* if $\det \overline{\rho}(c) = -1$. Similarly, if F is a totally real field with $[F : \mathbb{Q}] = d$, then the d embeddings $F \hookrightarrow \mathbb{R}$ induce d complex conjugation automorphisms $c_1, \ldots, c_d \in G_F$. We say that $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is *totally odd* if $\det \overline{\rho}(c_i) = -1$ for each $i = 1, \ldots, d$.

The original statement of Serre's conjecture, which essentially dates back to the 1960's, appeared in some cases in [52], and was properly published only in [54], is the following.

Conjecture 2.1. *Let $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a mod p Galois representation. If $\overline{\rho}$ is continuous, irreducible, and odd, then there exists a modular eigenform $f \in S_k(\Gamma_1(N))$, for some weight k and level N, such that $\overline{\rho} \simeq \overline{\rho_{f,p}}$.*

This statement has a natural generalization for totally real fields:

Conjecture 2.2 (Weak Serre conjecture)**.** *If F is a totally real field and $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is continuous, irreducible, and totally odd, then $\overline{\rho}$ is modular.*

While this conjecture is already very powerful, one generally wants to know in what weights and levels to look for a modular form giving rise to $\overline{\rho}$. In fact, Serre gave such a strengthened version of his own conjecture, in which he specified a minimal weight $k(\overline{\rho})$ and level $N(\overline{\rho})$ such that there should exist a modular eigenform $f \in S_{k(\overline{\rho})}(\Gamma_1(N(\overline{\rho})))$ with $\overline{\rho} \simeq \overline{\rho_{f,p}}$. We shall not give the explicit formulae for $k(\overline{\rho})$ and $N(\overline{\rho})$ here, but the reader will be able to extract them from our statement of a generalized conjecture later on.

2.2 Serre weights

One of the most basic and useful facts about mod p representation theory, and one which is responsible for much of the difference in flavor between it and representation theory in characteristic zero, is the following. A proof may be found in [19].

Proposition 2.3. *Let G be a profinite group, let $P \subset G$ be a normal pro-p-group, and let $\tau : P \to \mathrm{GL}(V)$ be a continuous finite-dimensional representation of P on an $\overline{\mathbb{F}}_p$-vector space V. Let $V^P = \{v \in V : \forall a \in P, \tau(a)v = v\}$. Then $V^P \neq \{0\}$. Moreover, if τ is irreducible, then $V^P = V$.*

Proof. Clearly we may restrict to the case where V is irreducible as a G-module. Since τ is continuous and hence has finite image, we in fact have $\tau : G \to \mathrm{GL}(W)$, where W is a finite-dimensional vector space over a finite extension of $\mathbb{F}_p$. Since P is normal in G, clearly $W^P \subset W$ is a sub-G-module. Since P is a pro-p-group, every non-trivial orbit of the P-action on W must have cardinality divisible by p. However, the cardinality of W is itself divisible by p. Hence, $|W^P| \neq 1$, and thus $W^P = W$. □

Definition 2.4. Let F be a number field, let v be a place of F, and let $k_v = \mathcal{O}_F/v$ be the residue field at v.

1. A *Serre weight* is an irreducible $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(\mathcal{O}_F/p)$.

2. A *local Serre weight at v* is an irreducible $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(k_v)$.

Since $\mathrm{GL}_2(\mathcal{O}_F/p)$ is a finite group, there are only finitely many Serre weights for any number field F. Moreover, suppose that the ideal $p\mathcal{O}_F$ decomposes into prime factors as $p\mathcal{O}_F = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\cdots\mathfrak{p}_r^{e_r}$. Then by the Chinese remainder theorem, $\mathrm{GL}_2(\mathcal{O}_F/p) = \mathrm{GL}_2(\mathcal{O}_F/\mathfrak{p}_1^{e_1}) \times \cdots \times \mathrm{GL}_2(\mathcal{O}_F/\mathfrak{p}_r^{e_r})$. For each $1 \leq i \leq r$, let k_i denote the residue field $\mathcal{O}_F/\mathfrak{p}_i$. Since the kernel of the natural projection

$$\mathrm{GL}_2(\mathcal{O}_F/p) \to \mathrm{GL}_2(k_1) \times \cdots \times \mathrm{GL}_2(k_r)$$

is a p-group, all Serre weights factor through this projection by Proposition 2.3. It follows that all Serre weights have the form

$$\sigma = \bigotimes_{v|p} \sigma_v,$$

where σ_v is a local Serre weight at v. The representation theory of GL_2 of a finite field is well known [27], and the distinct local Serre weights at v are precisely the following:

$$\sigma_v = \bigotimes_{\tau : k_v \hookrightarrow \overline{\mathbb{F}}_p} \det{}^{w_\tau} \otimes (\mathrm{Sym}^{r_\tau} k_v^2 \otimes_{k_v,\tau} \overline{\mathbb{F}}_p).$$

Here, k_v^2 is the standard action of $\mathrm{GL}_2(k_v)$ on a two-dimensional vector space over k_v, whereas $0 \leq r_\tau \leq p-1$ for all τ and $0 \leq w_\tau \leq p-1$ for all τ, with the stipulation that the w_τ are not all $p-1$.

What is the connection between these Serre weights and the modularity of Galois representations? Observe, first of all, that if $F = \mathbb{Q}$, then the Serre weights are just $\det^w \otimes \mathrm{Sym}^r \overline{\mathbb{F}}_p^2$, where $0 \leq r \leq p-1$ and $0 \leq w \leq p-2$. Let $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a mod p Galois representation, and let $\mathbf{T}$ be the Hecke algebra generated over $\mathbb{Z}$ by the Hecke operators T_l for $l \nmid pN(\overline{\rho})$. Define the maximal ideal $\mathfrak{m}_{\overline{\rho}} \subset \mathbf{T}$ to be the kernel of the map

$$\begin{array}{rcl} \mathbf{T} & \to & \overline{\mathbb{F}}_p \\ T_l & \mapsto & \mathrm{tr}\overline{\rho}(\mathrm{Frob}_l). \end{array}$$

The following was observed by Ash and Stevens [2].

Proposition 2.5. *Let $k \geq 2$. A Galois representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is modular of level N and weight k if and only if $H^1(\Gamma_1(N), \mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2)_{\mathfrak{m}_{\overline{\rho}}} \neq 0$.*

We remark that by Theorem 3.4 of [18], any collection of eigenvalues of $\mathbf{T}$ for which there exists an eigenform of some weight has an eigenform of weight at most $p+1$. Thus we do not lose generality by concentrating on forms associated to Serre weights. The twist by a power of the determinant that can occur in the Serre weight corresponds to Nebentypus.

2.3 Modularity

Inspired by the observations above, we will formulate a new definition of modularity. We use quaternionic Shimura curves in place of modular curves.

Recall that F is our fixed totally real field and that $d = [F : \mathbb{Q}]$. Let B/F be a quaternion algebra that splits at exactly one real place and at all places above p (note that we make no conditions at places away from p, so there are no parity issues). In other words, we are able to fix isomorphisms

$$\begin{array}{rcl} B \otimes \mathbb{R} & \simeq & M_2(\mathbb{R}) \times \mathbb{H}^{d-1} \\ B \otimes \mathbb{Q}_p & \simeq & M_2(F \otimes \mathbb{Q}_p). \end{array}$$

Denote by $\mathbb{A}$ the adeles of $\mathbb{Q}$, and if S is a finite set of places of $\mathbb{Q}$ then we set $\mathbb{A}^S$ to be the adeles away from S. Define the group $G = \mathrm{Res}_{F/\mathbb{Q}} B^*$, and let $U \subset G(\mathbb{A}^\infty)$ be an open compact subgroup of the form $U = U_p \times U^p$, where $U^p \subset G(\mathbb{A}^{\infty,p})$ and

$$U_p = \ker\left(\prod_{v|p} \mathrm{GL}_2(\mathcal{O}_{F_v}) \to \prod_{v|p} \mathrm{GL}_2(k_v)\right).$$

Then the Shimura curve

$$M_U(\mathbb{C}) = G(\mathbb{Q})\backslash(G(\mathbb{A}^\infty) \times (\mathbb{C} - \mathbb{R}))/U$$

has a model over F. Set $V \subset G(\mathbb{A}^\infty)$ to be the following open compact subgroup:

$$V = \left(\prod_{v|p} \mathrm{GL}_2(\mathcal{O}_{F_v}) \right) \times U^p.$$

It follows from a simple modification of an argument of Carayol [16] that if U^p is sufficiently small (see [46] for the precise definition of "sufficiently small") then the natural map $M_U \to M_V$ is a Galois cover of Shimura curves with Galois group $V/U \simeq \prod_{v|p} \mathrm{GL}_2(k_v)$.

Definition 2.6. Let F be a totally real field, let σ be a Serre weight, and let $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a mod p Galois representation. We say that $\overline{\rho}$ is *modular of weight* σ if there exist a quaternion algebra B/F and an open compact subgroup $U \subset G(\mathbb{A}^\infty)$ as above such that

$$\overline{\rho} \subset (\mathrm{Pic}^0(M_U)[p] \otimes \sigma)^{\prod_{v|p} \mathrm{GL}_2(k_v)}. \tag{2.1}$$

The reader is referred to the introduction of [14] and to Propositions 2.5 and 2.10 of the same paper for a discussion of why the naive notion of modularity (a Galois representation is modular if it arises from a Hilbert modular form of weight $(k_1, \dots, k_d)$) is insufficient and of the connection between this naive modularity and the notion of being modular of a certain Serre weight.

Remark 2.7. It is not hard to show, using the Eichler-Shimura relation, that the condition (2.1) is equivalent to the following:

$$\overline{\rho}^\vee = \mathrm{Hom}(\overline{\rho}, \overline{\mathbb{F}}_p) \subset H^1_{\acute{e}t}(M_V \otimes \overline{\mathbb{Q}}, \mathcal{L}_\sigma),$$

where $\mathcal{L}_\sigma$ is the mod p local system associated to σ and $\overline{\rho}^\vee$ is the dual of $\overline{\rho}$. While this fact is well-known, the author does not know of a clear proof in the literature. See Proposition 2.6 of [46].

Remark 2.8. The reader may wonder why we assume throughout that F is a totally real field. Why do we not work with arbitrary number fields? This question is strengthened by the fact that the proofs of many of the results about Galois representations $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ that are described below do not use global information about $\overline{\rho}$, but only local information about the restrictions of $\overline{\rho}$ to the decomposition subgroups at various places. It is very difficult to study modular Galois representations over arbitrary number fields F because of the lack of nice algebraic varieties over F that store automorphic data and in whose cohomology we could look for our Galois representations. The theory of Shimura varieties over totally real fields has no good analogue over general number fields. A number of mathematicians have studied the modularity for quadratic imaginary fields: see, for example, [28], [57], and [6]. In this case one can translate the problem to Siegel modular surfaces. See [7] and the end of [15] for modularity lifting theorems in the quadratic imaginary case.

2.4 Serre's weight conjecture

For every place v of F dividing p, denote by k_v the corresponding residue field and let its cardinality be $q_v = p^{f_v}$. Recall that G_{F_v} can be embedded non-canonically into G_F as follows. For every finite extension L/F, we choose a place v_L of L such that $v_L|v$ and such that if $L \subset L'$, then $v_{L'}|v_L$. Now let $G_v = \{\alpha \in G_F : \forall L/F, \alpha(v_L) = v_L\}$. This G_v is called a decomposition subgroup at v, and it is easy to show that $G_v \simeq G_{F_v}$. If we replace the system $\{v_L\}_L$ by a different compatible system of places, we will get a subgroup conjugate to G_v. Inside G_v, we have the inertia subgroup I_v consisting of all $\alpha \in G_v$ such that for each L/F, the induced automorphism of $\mathcal{O}_L/v_L$ is trivial. The wild inertia P_v is the pro-p-Sylow subgroup of I_v. Note that the isomorphism $G_v \simeq G_{F_v}$ induces isomorphisms $I_v \simeq \mathrm{Gal}(\overline{\mathbb{Q}}_p/F_v^{nr})$ and $P_v \simeq \mathrm{Gal}(\overline{\mathbb{Q}}_p/F_v^{tr})$, where F_v^{nr} and F_v^{tr} are the maximal unramified and maximal tamely ramified extensions of F_v, respectively. In particular, since F_v^{tr} is a Galois extension of F_v, we see that P_v is a normal subgroup of G_v. We have an exact sequence

$$1 \to I_v \to G_v \to \mathrm{Gal}(\overline{\mathbb{F}}_p/k_v) \to 1. \tag{2.2}$$

Now, let $r \geq 1$ and consider an embedding of fields $\tau : \mathbb{F}_{p^r} \hookrightarrow \overline{\mathbb{F}}_p$. Fix a uniformizer $\pi_v \in \mathcal{O}_{F_v}$ and define $\psi_\tau : I_v \to \overline{\mathbb{F}}_p^*$ to be the following composition of homomorphisms:

$$I_v \simeq \mathrm{Gal}(\overline{\mathbb{Q}}_p/F_v^{nr}) \to \mathrm{Gal}(F_v^{nr}(\sqrt[p^r-1]{\pi_v})/F_v^{nr}) \simeq \mathbb{F}_{p^r}^* \xrightarrow{\tau} \overline{\mathbb{F}}_p^*.$$

Let $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a mod p Galois representation. Then P_v acts trivially on $(\overline{\rho}_{|G_v})^{ss}$ by Proposition 2.3. Here $(\overline{\rho}^{ss}_{|G_v})$ is the semisimiplication of $\overline{\rho}_{|G_v}$, namely the direct sum of its irreducible Jordan-Hölder constituents. It follows that $((\overline{\rho}_{|G_v})^{ss})_{|I_v}$ factors through a representation of the abelian group I_v/P_v and thus is a sum of characters $\varphi \oplus \varphi'$. If Frob_v is a lift to G_v via (2.2) of the map $x \mapsto x^{q_v}$, which is a topological generator of $\mathrm{Gal}(\overline{\mathbb{F}}_p/k_v)$, then Frob_v acts on $((\overline{\rho}_{|G_v})^{ss})_{|I_v}$ by conjugation. Therefore, $\{\varphi, \varphi'\} = \{(\varphi)^{q_v}, (\varphi')^{q_v}\}$. Let $[k_v' : k_v] = 2$. We have two possibilities:

1. If $\overline{\rho}_{|G_v}$ is reducible, then $\overline{\rho}_{|I_v} \sim \begin{pmatrix} \varphi & * \\ 0 & \varphi' \end{pmatrix}$, where φ and φ' each factor through k_v.

2. If $\overline{\rho}_{|G_v}$ is irreducible, then $\overline{\rho}_{|I_v} \sim \begin{pmatrix} \varphi & 0 \\ 0 & \varphi' \end{pmatrix}$, where φ and φ' factor through $(k_v')^*$ and we have $\varphi' = \varphi^{q_v}$ and $\varphi = (\varphi')^{q_v}$.

We are finally ready to state Serre's conjecture. Given a mod p Galois representation $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$, our aim is to define a set $W(\overline{\rho})$ of Serre weights and then conjecture:

Conjecture 2.9 (Strong Serre conjecture)**.** *If $\overline{\rho}$ is continuous, irreducible, and totally odd, then it is modular. Moreover, if σ is a Serre weight, then $\overline{\rho}$ is modular of weight σ if and only if $\sigma \in W(\overline{\rho})$.*

A crucial property of the set $W(\overline{\rho})$ is that it is defined locally. This means that for each place $v|p$ we will specify a set $W_v(\overline{\rho})$ of local Serre weights at v (recall Definition 2.4) and then define

$$W(\overline{\rho}) = \left\{ \sigma = \bigotimes_{v|p} \sigma_v : \forall v|p, \sigma_v \in W_v(\overline{\rho}) \right\}. \tag{2.3}$$

Given a place $v|p$, let e_v be the ramification index of $v|p$, so that $[F_v : \mathbb{Q}_p] = e_v f_v$. Let S_v be the collection of all field embeddings $k_v \hookrightarrow \overline{\mathbb{F}}_p$. The definition of $W_v(\overline{\rho})$ involves several cases, corresponding to the cases above.

1. If $\overline{\rho}_{|G_v}$ is reducible and semisimple (i.e. the direct sum of two one-dimensional representations) then a local Serre weight at v

$$\sigma_v = \bigotimes_{\tau \in S_v} \left(\det{}^{w_\tau} \otimes \mathrm{Sym}^{r_\tau} k_v^2 \otimes_{k_v,\tau} \overline{\mathbb{F}}_p \right) \tag{2.4}$$

is contained in $W_v(\overline{\rho})$ if and only if there exists a subset $A \subset S_v$ and an integer $0 \leq \delta_\tau \leq e_v - 1$ for each $\tau \in S_v$, such that

(2.5)
$$\overline{\rho}_{|I_v} \sim \begin{pmatrix} \prod_{\tau \in A} \psi_\tau^{w_\tau + r_\tau + 1 + \delta_\tau} \prod_{\tau \notin A} \psi_\tau^{w_\tau + e_v - 1 - \delta_\tau} & 0 \\ 0 & \prod_{\tau \in A} \psi_\tau^{w_\tau + e_v - 1 - \delta_\tau} \prod_{\tau \notin A} \psi_\tau^{w_\tau + r_\tau + 1 + \delta_\tau} \end{pmatrix}.$$

2. If $\overline{\rho}_{|G_v}$ is irreducible, then σ_v as in (2.4) is contained in $W_v(\overline{\rho})$ if and only if for each $\tau \in S_v$ there exists an integer $0 \leq \delta_\tau \leq e_v - 1$ and a lift $\tilde{\tau} : k_v' \hookrightarrow \overline{\mathbb{F}}_p$ of τ (recall that k_v' is a quadratic extension of k_v) such that

$$\overline{\rho}_{|I_v} \sim \begin{pmatrix} \varphi & 0 \\ 0 & \varphi^{q_v} \end{pmatrix},$$

where

$$\varphi = \prod_{\tau \in S_v} \psi_{\tilde{\tau}}^{(q_v+1)w_\tau + r_\tau + 1 + \delta_\tau + q_v(e_v - 1 - \delta_\tau)}. \tag{2.6}$$

3. If $\overline{\rho}_{|G_v}$ is indecomposable, i.e. reducible but not semisimple, then
$$\overline{\rho}_{|G_v} \sim \begin{pmatrix} \varphi & * \\ 0 & \varphi' \end{pmatrix},$$
where $*$ corresponds to an element $c_{\overline{\rho}} \in \mathrm{Ext}^1(\varphi', \varphi) \simeq H^1(G_v, \varphi(\varphi')^{-1})$. For each local Serre weight σ_v and each subset $A \subset S_v$ one defines a subspace $L_{A,\sigma_v} \subset H^1(G_v, \varphi(\varphi')^{-1})$. The definition of L_{A,σ_v} is intricate, and we omit it here. Then a local Serre weight as in (2.4) is contained in $W_v(\overline{\rho})$ if and only if there exist a subset $A \subset S_v$ and an integer $0 \leq \delta_\tau \leq e_v - 1$ for each $\tau \in S_v$ such that $((\overline{\rho}_{|G_v})^{ss})_{|I_v}$ has the form of (2.5) and in addition $c_{\overline{\rho}} \in L_{A,\sigma_v}$.

As we mentioned above, this conjecture was first stated in the case of $F = \mathbb{Q}$ by Serre several decades ago, in a slightly different language. We will now indicate how to extract the minimal weight $k(\overline{\rho})$ from the conjecture formulated above. It follows from Proposition 2.5 and the definition of modularity that $\overline{\rho}$ arises from a modular form of weight $2 \leq k \leq p+1$ (and trivial Nebentypus) precisely when it is modular of the Serre weight $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$. Furthermore, it is easily shown by an argument with Eisenstein ideals that, for arbitrary $k \geq 2$, we have $H^1(\Gamma_1(N), \mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2)_{\mathfrak{m}_{\overline{\rho}}} \neq 0$ if and only if $H^1(\Gamma_1(N), W)_{\mathfrak{m}_{\overline{\rho}}} \neq 0$ for some Jordan-Hölder constituent W of $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p$. Thus, the minimal weight $k(\overline{\rho})$ conjectured by Serre is just the minimal $k \geq 2$ such that $\overline{\rho}$ is modular (in the sense of Conjecture 2.9) of some Jordan-Hölder constituent of $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$. This $k(\overline{\rho})$ is easily computed using the following decomposition.

Lemma 2.10. *Let $m > p-1$. Then the following holds, where "ss" denotes semisimplification:*
$$(\mathrm{Sym}^m\overline{\mathbb{F}}_p^2)^{ss} = (\det \otimes \mathrm{Sym}^{m-p-1}\overline{\mathbb{F}}_p^2)^{ss} \oplus \mathrm{Sym}^r\overline{\mathbb{F}}_p^2 \oplus \det{}^r \otimes \mathrm{Sym}^{p-1-r}\overline{\mathbb{F}}_p^2,$$
where r is the unique integer in the range $0 \leq r \leq p-2$ such that $r \equiv m$ modulo $p-1$.

The first generalization of Serre's conjecture beyond the original case of $F = \mathbb{Q}$, to the case of totally real fields F in which p is unramified, was by Buzzard, Diamond, and Jarvis [14] and circulated for nearly a decade before their paper appeared in print. A conjecture for arbitrary totally real fields F but semisimple $\overline{\rho}_{|G_v}$ (i.e. the first two of the three cases above) was made by the author; see [45], Theorems 2.4 and 2.5. At around the same time, Herzig [31] made a conjecture for n-dimensional representations $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ for arbitrary n, under the assumption that $\overline{\rho}_{|G_p}$ is semisimple. These conjectures were later restated by other authors [3] [21] to cover arbitrary $\overline{\rho}$ at the cost of becoming less explicit: if σ is a local Serre weight at $v|p$, then $\sigma \in W_v(\overline{\rho})$ is conjectured to be equivalent to the existence of a p-adic lift of $\overline{\rho}$ with some specified local properties. Such opacity appears already in the indecomposable case of the conjecture of [14], in the form of the spaces L_{A,σ_v}.

2.5 The level in Serre's conjecture

Before proceeding, we will say a few words about the level in Serre's conjecture. For this, the notion of the Artin conductor of a Galois representation is crucial. Roughly speaking, the Artin conductor $\mathfrak{n}(\overline{\rho})$ of $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is an ideal $\mathfrak{n}(\overline{\rho}) = \prod_v v^{a_v}$ of $\mathcal{O}_F$, where v runs over the finite places of F and a_v measures the ramification of $\overline{\rho}$ at v. If $\overline{\rho}$ is unramified at v, i.e. if $I_v \subset \ker \overline{\rho}$, then $a_v = 0$. Otherwise, the exponent a_v reflects how far down the upper ramification filtration of I_v one has to go to find a subgroup contained in $\ker \overline{\rho}$. More precisely, let L/F be a finite Galois extension such that $\overline{\rho}$ factors through $\mathrm{Gal}(L/F)$; this exists because $\overline{\rho}$ is continuous and G_F is compact, hence $\overline{\rho}$ has finite image. if $v \nmid p$, then let v_L be a place of L lying above v, and for each $i \geq 0$ let $G_i \subset G_{v_L} \simeq \mathrm{Gal}(\overline{L_{v_L}}/L_{v_L})$ be the i-th ramification subgroup; see, for instance, Chapter IV of [51] for the definition. If $V_{\overline{\rho}}$ is a two-dimensional $\overline{\mathbb{F}}_p$-vector space on which $\overline{\rho}$ acts, then we define

$$a_v = \sum_{i=0}^{\infty} \frac{1}{[I_{v_L} : G_i]} \dim_{\overline{\mathbb{F}}_p} (V_{\overline{\rho}}/V_{\overline{\rho}}^{G_i}),$$

where $V_{\overline{\rho}}^{G_i}$ is the subspace of G_i-invariants. More details about the Artin conductor may be found in Chapter VI of [51].

When $F = \mathbb{Q}$, any modular Galois representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ arises from a modular form of weight N that is prime to p, and Serre specified a minimal such weight $N(\overline{\rho})$: this is just the prime-to-p part of the Artin conductor of $\overline{\rho}$. If F is an arbitrary totally real field, we are no longer so lucky. We can no longer expect $\overline{\rho}$ to always arise in level prime to p (this would correspond to $U_p = \prod_{v|p} \mathrm{GL}_2(\mathcal{O}_{F_v})$ in Section 2.3); see the introduction of [14] for a discussion of why not. However, the prime-to-p part of the level may conjecturally always be taken to be the prime-to-p part of $\mathfrak{n}(\overline{\rho})$, which means that we may take $U^p = \prod_{v \nmid p} U_1(v^{a_v})$, where the group $U_1(v^{a_v}) \subset \mathrm{GL}_2(\mathcal{O}_{F_v})$ consists of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F_v})$ such that $a - 1 \in v^{a_v}$ and $c \in v^{a_v}$.

3 Weak Serre implies strong Serre

The "weak" version of Serre's modularity conjecture (Conjecture 2.2) is actually a very strong statement. It has been proved only in the case $F = \mathbb{Q}$. This was achieved by Khare, Wintenberger, and Kisin [36], [37], [39], and we will sketch some of their methods below. As we will see at the end of these notes, some of these methods fail crucially whenever F is a totally real field with $[F : \mathbb{Q}] > 1$, so that a major new idea is needed for any substantial further progress on the conjecture.

In the meantime, much research in the area has focused on proving, in various settings, that the weak version of Serre's conjecture (Conjecture 2.2) implies the strong version (Conjecture 2.9), in other words that if $\overline{\rho}$ is modular, then it is modular of precisely the predicted Serre weights. In the case $F = \mathbb{Q}$, this fact has essentially been known since the late 1970's except for a few cases where $p = 2$ and was an important input in the proof of Serre's conjecture (the stubborn cases with $p = 2$ were also settled by Khare, Wintenberger, and Kisin's work). It was proved by Deligne for $\overline{\rho}_{|G_p}$ reducible and by Fontaine for $\overline{\rho}_{|G_p}$ irreducible. Fontaine's work was never published, and a (somewhat different) proof of the theorem first appeared in print in [18].

We will now mention some of the "weak Serre implies strong Serre" theorems that have been proved in recent years. The conjecture of [14], for F in which p is unramified, was proved by Gee [23] for most Serre weights by deformation-theoretic methods; see also [46] for a more geometric proof of most cases of one direction of this conjecture for locally irreducible $\overline{\rho}$. We will say more about the methods of these papers in the remainder of this section. The remaining cases of the Buzzard-Diamond-Jarvis conjecture were attacked in a series of papers by Gee and coauthors, until it was finally proved completely in [24]. The results of [46] were extended in [45] to cases where p ramifies in F. Moreover, the conjecture of [45] was proved, for most cases where p is totally ramified in F, by Gee and Savitt [26]. More cases in the related, but not equivalent, unitary setting were resolved in [25]. Some non-totally ramified cases with $e = f = 2$ were addressed by R. Smith in his Ph.D. thesis at the University of Arizona, but it seems that new ideas are needed to make substantial further progress.

3.1 A sketch of Gee's argument

The claim that weak Serre implies strong Serre consists, of course, of two claims in opposite directions:

1. If $\overline{\rho}$ is modular of weight σ, then $\sigma \in W(\overline{\rho})$.
2. If $\overline{\rho}$ is modular of some weight and $\sigma \in W(\overline{\rho})$, then $\overline{\rho}$ is modular of weight σ.

The most successful method for proving "weak Serre implies strong Serre" has been that of relating the modularity of $\overline{\rho}$ to the existence of lifts of $\overline{\rho}$ with some specific local properties and then using p-adic Hodge theory to investigate the existence of such lifts.

An important breakthrough was Gee's paper [23], which proved the following result. Its statement involves the following definition: a local Serre weight at v is said to be regular if it is of the form (2.4) with $1 \leq r_\tau \leq p-3$ for all $\tau \in S_v$. A Serre weight $\sigma = \bigotimes_{v|p} \sigma_v$ is called regular if all the σ_v are regular.

Theorem 3.1. *Suppose that $p \geq 5$ is unramified in the totally real field F, that $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is modular of some Serre weight, and that $\overline{\rho}_{|G_{F(\zeta_p)}}$ is irreducible. Suppose that σ is a regular Serre weight. If $\overline{\rho}$ is modular of weight σ, then $\sigma \in W(\overline{\rho})$. Conversely, if $\sigma \in W(\overline{\rho})$ and some further technical conditions are satisfied at places $v|p$ where $\overline{\rho}_{|G_v}$ is reducible and σ_v arises from $A = S_v$ or $A = \varnothing$ in the recipe of Section 2.4, then $\overline{\rho}$ is modular of weight σ.*

We will give a very brief sketch of part of the argument of [23] to illustrate the method; the reader is referred to that paper (and the papers cited in it!) for further details. It should be noted that Gee works with a different notion of modularity than the one given above; he uses definite quaternion algebras, rather than indefinite ones. It is not possible to translate theorems directly from one setting to the other, but his local arguments can be translated. In this section, we will assume that p is unramified in the totally real field F. Let $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a continuous, irreducible, and totally odd mod p Galois representation. Recall that in Section 2.4 we defined a set $W_v(\overline{\rho})$ of local Serre weights at v for each $v|p$.

Let $v|p$; for the purposes of this section, we will say that a p-adic representation η_v of $\mathrm{GL}_2(k_v)$ is *good* if it is either an irreducible principal series or supercuspidal; in other words, η_v is any irreducible p-adic representation of $\mathrm{GL}_2(k_v)$ that is not one-dimensional or special. We regard η_v as a representation of the group $\mathrm{GL}_2(\mathcal{O}_{F_v})$ via the obvious inflation. Then η_v has an associated inertial type τ_v, namely a p-adic representation of I_v with the property that for any irreducible p-adic representation π of $\mathrm{GL}_2(F_v)$, we have $\eta_v \subset \pi_{|GL_2(\mathcal{O}_{F_v})}$ if and only if $\mathrm{LLC}(\pi)_{|I_v} \simeq \tau_v$. Here $\mathrm{LLC}(\pi)$ is the Weil-Deligne representation associated to π by the local Langlands correspondence. See Henniart's appendix to [11] for an exposition of the theory of types for GL_2.

Proposition 3.2 ([23], Lemma 2.1.4)**.** *Let $\overline{\rho}$ be as above, and for each $v|p$ let η_v be a good representation as above. Then $\overline{\rho}$ is modular of some Serre weight $\sigma \in \mathrm{JH}(\bigotimes_{v|p}(\eta_v \otimes \overline{\mathbb{F}}_p))$ if and only if $\overline{\rho}$ has a modular p-adic lift $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ such that for each $v|p$, the restriction $\rho_{|G_v}$ is potentially Barsotti-Tate (i.e. potentially crystalline with Hodge-Tate weights $(0,1)$) and $\mathrm{WD}(\rho_{|G_v})_{|I_v} = \tau_v$.*

Here, and subsequently, we write $JH(V)$ for the set of Jordan-Hölder constituents of a representation V, whereas $\mathrm{WD}(\rho_{|G_v})$ denotes the Weil-Deligne representation corresponding to the local Galois representation $\rho_{|G_v}$. The reader is referred to the classic article [56] for the correspondence between Galois and Weil-Deligne representations.

Suppose that we know how to prove the first of the two claims at the beginning of this section, namely that if $\overline{\rho}$ is modular of weight σ, then $\sigma \in W(\overline{\rho})$. Assuming that, here is a strategy for proving the second claim. Let $\sigma \in W(\overline{\rho})$ be a Serre weight. If it is regular, then for each $v|p$ there exists a good η_v as above such that $\mathrm{JH}(\otimes_{v|p}(\eta_v \otimes \overline{\mathbb{F}}_p)) \cap W(\overline{\rho}) = \{\sigma\}$. Then by Proposition 3.2 it suffices to find a modular lift ρ with the properties specified in the statement of that proposition.

The most daunting aspect of coming up with a lift ρ of $\overline{\rho}$ that satisfies the conditions of Proposition 3.2 is clearly that of showing that the ρ we have constructed is modular. Fortunately, Gee's adaption of a modularity lifting theorem of Kisin comes to the rescue. This is the first of many close connections that we will see in these lectures between Serre's modularity conjecture and modularity lifting theorems.

Proposition 3.3. *Suppose that the hypotheses of Theorem 3.1 hold and that $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ is a lift of $\overline{\rho}$ such that $\rho_{|G_v}$ is potentially Barsotti-Tate and $\mathrm{WD}(\rho_{|G_v})_{|I_v} = \tau_v$ for each $v|p$. Suppose that there exists a cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbb{A}_F)$ such that for every $v|p$, the local Galois representation $\rho_{\pi,v}$ is potentially ordinary if and only if $\rho_{|G_v}$ is potentially ordinary. Then ρ is modular.*

Note that the hypothesis on $\rho_{|G_{F(\zeta_p)}}$ in the statement of Theorem 3.1 is common in modularity lifting theorems à la Kisin, and this is the point in the proof where it is necessary.

Now we need to construct a lift ρ satisfying the conditions of Proposition 3.3. The theory of Breuil modules allows us to translate local conditions on Galois representations into linear-algebraic data.

Let k be a finite field of characteristic $p > 2$, let $W(k)$ be the associated ring of Witt vectors, and let $K_0 = W(k)[1/p]$ be its fraction field. Let K/K_0 be a totally tamely ramified Galois extension of degree e. Let $B \subset K_0$ be a subfield such that there exists a uniformizer $\pi \in \mathcal{O}_K$ satisfying $\pi^e \in B$. Choose such a π. Let $2 \leq k \leq p-1$ be an integer; this conflict of notation is standard and will produce no confusion. Let $E/\mathbb{F}_p$ be a finite extension. The category $\mathrm{BrMod}_{dd,B}^{k-1}$ of Breuil modules with descent data has as objects quintuples $(M, M_{k-1}, \varphi_{k-1}, N, \hat{g})$ such that:

1. M is a finitely generated $(k \otimes_{\mathbb{F}_p} E)[u]/u^{ep}$-module that is free over $k[u]/u^{ep}$.

2. M_{k-1} is a submodule such that $u^{e(k-1)}M \subset M_{k-1}$.

3. $\varphi_{k-1} : M_{k-1} \to M$ is an E-linear and Frobenius-semilinear homomorphism whose image generates M as a $(k \otimes_{\mathbb{F}_p} E)[u]/u^{ep}$-module. Frobenius-semilinear in this case means that if $a \in k[u]/u^{ep}$ and $m \in M_{k-1}$, then $\varphi_{k-1}(am) = a^p \varphi_{k-1}(m)$.

4. $N : M \to uM$ is a $(k \otimes_{\mathbb{F}_p} E)$-linear map satisfying

   (a) $N(um) = uN(m) - um$ for all $m \in M$.

   (b) $u^e N(M_{k-1}) \subset M_{k-1}$.

   (c) $\varphi_{k-1}(u^e N(m)) = -\frac{\pi^e}{p} N(\varphi_{k-1}(m))$ for all $m \in M_{k-1}$.

5. For each $g \in \mathrm{Gal}(K/B)$, there is an additive bijection $\hat{g} : M \to M$ such that

(a) Each $\hat{g}$ commutes with, φ_{k-1}, M, and the E-action.

(b) $\hat{1}$ is the identity map, where $1 \in \mathrm{Gal}(K/B)$ is the identity automorphism.

(c) $\hat{g} \circ \hat{h} = \widehat{g \circ h}$ for all $g, h \in \mathrm{Gal}(K/B)$.

(d) $\hat{g}(au^i m) = g(a)((g(\pi)/\pi)^i \otimes 1)u^i \hat{g}(m)$ for all $a \in k \otimes_{\mathbb{F}_p} E$, $m \in M$, and $i \geq 0$. To make sense of $g(a)$, note that k is the residue field of K_0, hence of K, and so is acted on by $\mathrm{Gal}(K/B)$. We let $\mathrm{Gal}(K/B)$ act trivially on the second component of $k \otimes_{\mathbb{F}_p} E$.

The connection between Breuil modules and potentially Barsotti-Tate Galois representations is evidenced, for instance, by the fact that the category $\mathrm{BrMod}^1_{dd,B}$ is equivalent to the category of finite flat group schemes over $\mathcal{O}_K$ with an action of E and descent data to B. Gee proves that the existence of a lift ρ which is potentially Barsotti-Tate at v of inertial type τ_v is equivalent to the existence of a Breuil module satisfying certain conditions. As we see from the definition above, Breuil modules with descent data are complicated objects but are very explicit, and one constructs the needed Breuil module by hand.

The proof for arbitrary (i.e. not necessarily regular) Serre weights follows the same lines, but the theory of Breuil modules, which itself is an extension of Fontaine-Laffaille theory, is not powerful enough. Here one uses Liu's theory of Kisin modules.

3.2 Modular weights are predicted ones: some algebraic geometry

In this section we will sketch how to prove that if $\overline{\rho}$ is modular of a Serre weight σ, then $\sigma \in W(\overline{\rho})$. In order to illustrate the variety of methods applicable to this problem, we will give an algebraic-geometry argument following [46] and [45]. This method was used to obtain the earliest results in this direction, but it has turned out to be less effective than the deformation-theoretic and p-adic Hodge-theoretic techniques of which a flavor was given in the previous section. The reader may, of course, find further details in [46].

In this section we will not impose such severe limitations on the ramification of p in the totally real field F, but we will suppose that $\overline{\rho}_{|G_v}$ is irreducible for all $v|p$. For each $v|p$, let e_v be the ramification index of $F_v/\mathbb{Q}_p$. Let $\sigma = \bigotimes_{v|p} \sigma_v$ be a Serre weight such that for each v and each $\tau \in S_v$ we have $0 \leq r_\tau \leq p - e_v - 1$ (so in particular we are assuming here that $e_v \leq p-1$).

Now we will recall some notions from Sections 2.2 and 2.3. Assume that $\overline{\rho} : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is modular of weight σ. By definition, this implies the existence of a quaternion algebra B/F, giving rise to an algebraic group G, and an open compact subgroup $V = \left(\prod_{v|p} \mathrm{GL}_2(\mathcal{O}_{F_v})\right) \times U^p \subset G(\mathbb{A}^\infty)$ such that we have

$H^1_{ét}(M_V \otimes \overline{\mathbb{Q}}, \mathcal{L}_\sigma)_{\mathfrak{m}_{\overline{\rho}}} \neq 0$. For each $v|p$, let $U_1^{bal}(v) \subset \mathrm{GL}_2(\mathcal{O}_{F_v})$ be the subgroup of matrices whose reductions modulo v are unipotent upper triangular, i.e. of matrices that are congruent to $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ modulo v. Consider the open compact subgroup

$$U_1^{bal}(p) = \left(\prod_{v|p} U_1^{bal}(v) \right) \times U^p \subset G(\mathbb{A}^\infty).$$

Fix a place $v|p$. Let $D = W(k_v)$ be a ring of Witt vectors, let $K = F_v^{nr}$ be the fraction field of D, let $K' = K(\sqrt[q_v-1]{\pi_v})$ be a totally tamely ramified extension with $\mathrm{Gal}(K'/K) \simeq k_v^*$, and let $D' = \mathcal{O}_{K'}$. Then $M_{U_1^{bal}(p)}$ has an integral model over D, which we shall denote $\mathbf{M}_{U_1^{bal}(p)}$. Moreover, $\mathbf{M}_{U_1^{bal}(p)} \times_D D'$ has a well-behaved special fiber consisting of two smooth curves intersecting transversally at finitely many points.

Let $j : \mathrm{Gal}(K'/K) \to \mathcal{O}_{F_v}^*/(1+v)$ be the isomorphism induced by the Artin reciprocity map of local class field theory (normalized so as to send arithmetic Frobenius to uniformizers). The special fiber of $\mathbf{M}_{U_1^{bal}(p)} \times_D D'$ is equipped with natural actions of $\mathrm{GL}_2(\mathcal{O}_{F_v})$ (coming from the p-component of $G(\mathbb{A}^\infty)$) and of $\mathrm{Gal}(K'/K)$. Carayol ([16], 10.3) shows that the action of $\gamma \in \mathrm{Gal}(K'/K)$ is equal to that of $\begin{pmatrix} j(\gamma)^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & j(\gamma)^{-1} \end{pmatrix}$, respectively, on the two components of the special fiber.

Recall that we are assuming that $\overline{\rho}$ is modular of a given Serre weight σ. Let $B(k_v) \subset \mathrm{GL}_2(k_v)$ be the Borel subgroup of upper triangular matrices, and let $\theta : B(k_v) \to \overline{\mathbb{F}}_p^*$ be a character such that $\sigma_v \in \mathrm{JH}(\mathrm{Ind}_{B(k_v)}^{\mathrm{GL}_2(k_v)}\theta)$. Let C be the Néron model over D' of the curve $\mathrm{Pic}^0(\mathbf{M}_{U_1^{bal}(p)}) \times K'$. Then $C[p^\infty]$ is a p-divisible group, and the reduction $C[p^\infty] \otimes \mathbf{T}/\mathfrak{m}_{\overline{\rho}}$ contains a finite piece G_θ on which the diagonal matrices in $\mathrm{GL}_2(\mathcal{O}_{F_v})$ act via the character θ. By the main result of [9], $G_\theta[\mathfrak{m}_{\overline{\rho}}]_K$ is a direct sum of a finite number of copies of $\overline{\rho}$. As in Section 2.4 above, $\overline{\rho}_{|I_v}$ is a direct sum of two characters, φ and φ', that satisfy $\varphi^{q_v} = \varphi'$ and $(\varphi')^{q_v} = \varphi$. We can pick out a subspace $H \subset G_\theta[\mathfrak{m}_{\overline{\rho}}]_K$ of rank q_v^2 on which $\mathrm{Gal}(\overline{K}/K) \simeq I_v$ acts by the character φ.

Let $\mathbb{F}$ be a finite field, sufficiently large so that $\mathrm{im}(\overline{\rho}_{|G_v}) \subset \mathrm{GL}_2(\mathbb{F})$ and $\mathbb{F}_{q_v^2} \subset \mathbb{F}$. We will apply Raynaud's theory of vector space schemes [44]. An $\mathbb{F}$-vector space scheme over D is a commutative group scheme W/D carrying an action of $\mathbb{F}$. Let $\mathcal{I} \subset \mathcal{O}_W$ be the augmentation ideal, so that $\mathcal{O}_W = \mathcal{I} \oplus \mathcal{O}_D$. Here $\mathcal{O}_W$ is the structure sheaf of W. It is easy to see that $\mathcal{I}$ decomposes as follows:

$$\mathcal{I} = \bigoplus_{\chi:\mathbb{F}^* \to D^*} \mathcal{I}_\chi,$$

where $\mathcal{I}_\chi$ is the piece of $\mathcal{I}$ on which $\mathbb{F}$ acts via the character χ. We see that H is an $\mathbb{F}_{q_v^2}$-vector space scheme, and it satisfies the additional crucial property that each $\mathcal{I}_\chi$ is a non-zero invertible sheaf. The vector space scheme H is endowed with two Galois actions:

1. As we noted before, $\mathrm{Gal}(\overline{K}/K) \simeq I_v$ acts on $H(\overline{K})$ by the character φ, which we are trying to determine.

2. $\mathrm{Gal}(K'/K) \simeq k_v^*$ acts on the cotangent space $\cot(H_{D'} \times_{D'} \overline{\mathbb{F}}_p)$. Thanks to Carayol's congruences mentioned above, we can express this action explicitly in terms of the character θ.

From Raynaud's work one deduces an explicit relation between these two different Galois actions. We will not perform the calculations here, but the reader can find them in Section 3 of [45]. At the end we obtain a collection $\Phi(\theta)$ of characters φ that are compatible with the known action of $\mathrm{Gal}(K'/K)$. It turns out that these are precisely the characters φ arising from mod p Galois representations $\overline{\rho}$ that are modular of some Serre weight $\sigma' \otimes \sigma^v$, where $\sigma' \in \mathrm{JH}(\mathrm{Ind}_{B(k_v)}^{\mathrm{GL}_2(k_v)} \theta)$ and $\sigma^v = \bigotimes_{w|p, w\neq v} \sigma_w$, where σ_w is an arbitrary local Serre weight at w. Observe that this is the best result that we can hope to obtain at this stage of the proof, since so far we have only used θ in our calculations and not σ_v itself.

To get a more precise result, we consider all the characters $\theta : B(k_v) \to \overline{\mathbb{F}}_p^*$ such that $\sigma_v \in \mathrm{JH}(\mathrm{Ind}_{B(k_v)}^{\mathrm{GL}_2(k_v)} \theta)$. Clearly all the φ associated to $\overline{\rho}$ that are modular of weight $\sigma_v \otimes \sigma^v$ lie in the intersection $\bigcap_{\sigma_v \in \mathrm{JH}(\mathrm{Ind}\theta)} \Phi_\theta$. We hope that this intersection will turn out to be exactly the collection of representations $\overline{\rho}$ such that $\sigma_v \in W_v(\overline{\rho})$. The hope comes true when σ_v is of the form (2.4) with $0 \leq r_\tau \leq p-1-e_v$ for all $\tau \in S_v$, which is the reason for the hypothesis to this effect that we made above. The combinatorial issues that prevent this method from giving us as good a theorem as we would like when r_τ does not satisfy the constraint $0 \leq r_\tau \leq p-1-e_v$ are essentially also what prevents the method of [23] from handling the non-regular Serre weights.

4 The mod p local Langlands correspondence

The Langlands philosophy postulates a deep connection between algebra and analysis and is one of the main motivations behind modern research on Serre's modularity conjecture and its generalizations. In this section we will show a very brief glimmer of the connection between them. Let $n \geq 1$, let $F/\mathbb{Q}_p$ be a p-adic field, and let E be a field. In very rough terms, we would like to have a correspondence between certain Galois representations $\rho : G_F \to \mathrm{GL}_n(E)$ and certain representations of $\mathrm{GL}_n(F)$ on vector spaces over E; one of the most difficult parts of this problem is finding the correct definition of "certain." Often one can attach L-functions to each of these types of objects, and the L-functions of the objects paired by the correspondence should match.

In the case of $E = \mathbb{C}$, the correspondence was proved by Harris and Taylor [29] and Henniart [30], working with Weil-Deligne representations instead of the closely related Galois representations. If $E = \overline{\mathbb{F}}_l$, with $l \neq p$, then considerable progress was made by Vignéras [60]. However, if $E = \overline{\mathbb{F}}_p$, then very little is known. In many respects the study of the mod p local Langlands correspondence is at the stage in its development where the complex local Langlands correspondence was in the 1970's: one tries to classify objects on both sides and pair them up explicitly in a natural way, but no deep underlying theory is yet available. Moreover, understanding the mod p representation theory of $\mathrm{GL}_n(F)$ has turned out to be remarkably difficult.

In this section, we will use the following notation. We let $G = \mathrm{GL}_n(F)$ and consider the maximal open compact subgroup $K = \mathrm{GL}_n(\mathcal{O}_F)$ and the center $Z = Z(G) \simeq F^*$. Let $\pi \in \mathcal{O}_F$ be a uniformizer, and let $k_F = \mathcal{O}_F/(\pi)$ be the residue field as usual. Let $q = p^f$ be the cardinality of k. Let $I \subset K$ be the Iwahori subgroup consisting of matrices that are upper triangular modulo π, and let $I(1)$ be the pro-p-Sylow subgroup of I. For instance, if $n = 2$ then

$$\begin{aligned} I &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_F) : c \in \pi\mathcal{O}_F \right\} \\ I(1) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_F) : c \in \pi\mathcal{O}_F; a, d \in 1 + \pi\mathcal{O}_F \right\}. \end{aligned}$$

Let σ be an irreducible $\overline{\mathbb{F}}_p$-representation of K. By Proposition 2.3, σ factors through the natural reduction map $K \to \mathrm{GL}_n(k)$, since the kernel of this map is a pro-p group. Therefore, σ arises from an irreducible $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_n(k)$ by inflation; these are exactly the objects that we called local Serre weights above in the case $n = 2$. Moreover, we can view σ as a representation of the larger group KZ by decreeing that $\begin{pmatrix} \pi & 0 \\ 0 & \pi \end{pmatrix}$ acts trivially.

If $H \subset G$ is any open subgroup, and τ is an $\overline{\mathbb{F}}_p$-representation of H, we can consider the compact induction $\mathrm{ind}_H^G \tau$. A model for this representation is given by the space of functions $f : G \to V_\tau$ that are locally constant, compactly supported modulo Z, and satisfy the condition $f(hg) = \tau(h) \cdot f(g)$ for every $h \in H$ and $g \in G$. Here V_τ is the underlying $\overline{\mathbb{F}}_p$-vector space of τ. The action of G is given by $(gf)(x) = f(xg)$ for all $g, x \in G$. Note that if H is a subgroup of finite index, then local constancy and compact support are automatic and this is just the usual induction. The endomorphisms of this compact induction were computed by Barthel and Livné [4] for $n = 2$. For $n \geq 2$, see [48] for an explicit computation and [33] for a more conceptual argument on the level of algebraic groups.

Proposition 4.1. *Let σ be an irreducible $\overline{\mathbb{F}}_p$-representation of K. The endomorphism algebra $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$ is equal to a polynomial ring $\overline{\mathbb{F}}_p[T_1, \ldots, T_{n-1}]$, where the T_i are explicitly defined endomorphisms.*

Let W be an irreducible $\overline{\mathbb{F}}_p$-representation of G with central character, i.e. such that the elements of Z act by scalars. Twisting by an unramified character, we may assume that $\begin{pmatrix} \pi & 0 \\ 0 & \pi \end{pmatrix}$ acts trivially. If $\sigma \subset W_{|K}$ is a K-submodule of G, then by Frobenius reciprocity we obtain a non-zero homomorphism $\mathrm{ind}_{KZ}^G \sigma \to W$ of G-modules, which must be a surjection by the irreducibility of W. We say that W is *admissible* if the space of invariants $W^U = \{w \in W : \forall u \in U, uw = w\}$ is finite-dimensional for any open subgroup $U \subset G$; since W is an $\overline{\mathbb{F}}_p$-representation this is in fact equivalent to $W^{I(1)}$ being finite-dimensional.

The endomorphism algebra $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$ is commutative by Proposition 4.1, and it acts on $\mathrm{Hom}_G(\mathrm{ind}_{KZ}^G \sigma, W)$ in the obvious way. If W is assumed to be admissible, then $\mathrm{Hom}_G(\mathrm{ind}_{KZ}^G \sigma, W) \simeq \mathrm{Hom}_{KZ}(\sigma, W_{|KZ})$ is finite-dimensional (because σ must contain a non-zero $I(1)$-invariant, which must map to an element of $W^{I(1)}$) and necessarily contains an eigenvector for the $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$-action. We obtain the following result.

Proposition 4.2. *Let W be a smooth irreducible $\overline{\mathbb{F}}_p[G]$-module with central character. Assume that W is admissible if $n \geq 3$. Let σ be an irreducible $\overline{\mathbb{F}}_p[K]$-module σ such that $\sigma \subset W_{|K}$. Then there exist an unramified character $\chi : F^* \to \overline{\mathbb{F}}_p^*$ and scalars $\lambda_1, \ldots, \lambda_{n-1} \in \overline{\mathbb{F}}_p$ such that there exists a surjection of G-modules*

$$(\chi \circ \det) \otimes \mathrm{ind}_{KZ}^G \sigma / (T_1 - \lambda_1, \ldots, T_{n-1} - \lambda_{n-1}) \mathrm{ind}_{KZ}^G \sigma \twoheadrightarrow W. \tag{4.1}$$

Proof. If W is admissible, then we have sketched out the proof. If $n = 2$, then Barthel and Livné (see Theorems 32 and 33 of [4]) obtain this result without assuming admissibility of W by using the fact that $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$ has Krull dimension 1. □

For the rest of this section, suppose that $n = 2$. In this case, the endomorphism algebra $\mathrm{End}_G(\mathrm{ind}_{KZ}^G \sigma)$ has a single generator T_1, which we will call T. Up to unramified twist, we know that every irreducible $\overline{\mathbb{F}}_p[G]$-module with central character is a quotient of $\mathrm{ind}_{KZ}^G \sigma / (T - \lambda)(\mathrm{ind}_{KZ}^G \sigma)$ for some σ and some $\lambda \in \overline{\mathbb{F}}_p$. We say that an irreducible W as above is *supersingular* if it is a quotient of some $\mathrm{ind}_{KZ}^G \sigma / (T - \lambda)(\mathrm{ind}_{KZ}^G \sigma)$. Barthel and Livné proved a partial classification of the irreducible $\overline{\mathbb{F}}_p[G]$-modules with central character as follows. Note that if $G = \mathrm{GL}_2(\mathbb{Q}_p)$, then Berger [5] recently showed that all irreducible $\overline{\mathbb{F}}_p[G]$-modules have central character, but this is not known even for $G = \mathrm{GL}_2(F)$ whenever $F \neq \mathbb{Q}_p$.

Theorem 4.3 (Barthel-Livné). *Let σ be an irreducible $\overline{\mathbb{F}}_p[K]$-module.*

1. *If σ has dimension other than 1 or p^f (the minimal and maximal dimensions possible) or if $\lambda \neq \pm 1$, then $\mathrm{ind}_{KZ}^G \sigma / (T - \lambda)(\mathrm{ind}_{KZ}^G \sigma)$ is irreducible and is isomorphic to the parabolic induction of a character from the upper triangular Borel subgroup $B \subset G$.*

2. *The induction* $\mathrm{ind}_B^G \mathbf{1}$*, where* $\mathbf{1}$ *is the trivial character of* B*, has length two. Its subquotients are a one-dimensional representation* det *and an infinite-dimensional analogue of the Steinberg representation, denoted* St*.*

3. *Up to unramified twist, every smooth irreducible* $\overline{\mathbb{F}}_p[G]$*-module* W *with central character satisfies exactly one of the following statements:*

   (a) $W \simeq \mathrm{ind}_{KZ}^G \sigma/(T-\lambda)(\mathrm{ind}_{KZ}^G \sigma)$*, where* σ *has dimension other than* 1 *or* p^f*, or* $\lambda \neq \pm 1$*.*

   (b) $W \simeq \chi \circ \det$ *for some smooth character* $\chi : F^* \to \overline{\mathbb{F}}_p^*$*.*

   (c) $W \simeq (\chi \circ \det) \otimes \mathrm{St}$ *for some smooth character* $\chi : F^* \to \overline{\mathbb{F}}_p^*$*.*

   (d) W *is supersingular.*

Remark 4.4. The previous theorem classifies all non-supersingular (smooth, with central character) $\overline{\mathbb{F}}_p$-representations of $\mathrm{GL}_2(F)$ for arbitrary finite extensions $F/\mathbb{Q}_p$. Herzig [32] proved a generalization of this theorem of $\mathrm{GL}_n(F)$ for $n > 2$, in which all smooth admissible representations of $\mathrm{GL}_n(F)$ with central character are classified in terms of the supersingular representations of $\mathrm{GL}_m(F)$ for $m \leq n$. A representation of $\mathrm{GL}_n(F)$ is called supersingular if it is a quotient of $\mathrm{ind}_{KZ}^G \sigma/(T_1, \ldots, T_{n-1})$. Abe [1] further generalized this result to a wider class of reductive groups.

Let L be a number field and v a place of L such that $L_v \simeq F$. If $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is an irreducible local Galois representation, let $\tilde{\rho} : G_L \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ be a global representation such that $\tilde{\rho}_{|G_v} \simeq \rho$. Recall that in Conjecture 2.9 we defined a set $W_v(\tilde{\rho})$ of local Serre weights at v, which in fact depends only on ρ. Thus we can speak of a set $W(\rho)$ of modular local Serre weights.

Now suppose that $G = \mathrm{GL}_2(\mathbb{Q}_p)$. In this case, the irreducible $\overline{\mathbb{F}}_p[K]$-modules have the form $\sigma = \det^w \otimes \mathrm{Sym}^r \overline{\mathbb{F}}_p^2$ with $0 \leq w \leq p-2$ and $0 \leq r \leq p-1$. We define an involution on the set of these local Serre weights as follows. For σ as above, define $\sigma' = \det^{w+r} \otimes \mathrm{Sym}^{p-1-r} \overline{\mathbb{F}}_p^2$. Note that $(\sigma')' = \sigma$. It is easy to compute from the statement of Conjecture 2.9 that if $\rho : G_{\mathbb{Q}_p} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is irreducible, then $W(\rho)$ is necessarily of the form $W(\rho) = \{\sigma, \sigma'\}$.

Breuil [10] completed the classification of the irreducible $\overline{\mathbb{F}}_p$-representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ with the following result.

Theorem 4.5 (Breuil)**.** *Let* $G = \mathrm{GL}_2(\mathbb{Q}_p)$*. Then for every local Serre weight* σ*, the* G*-module* $\mathrm{ind}_{KZ}^G \sigma/T(\mathrm{ind}_{KZ}^G \sigma)$ *is irreducible. Moreover, for every* σ *we have*

$$\mathrm{ind}_{KZ}^G \sigma/T(\mathrm{ind}_{KZ}^G \sigma) \simeq \mathrm{ind}_{KZ}^G \sigma'/T(\mathrm{ind}_{KZ}^G \sigma') \tag{4.2}$$

and these are the only isomorphisms among supersingular $\overline{\mathbb{F}}_p[G]$*-modules.*

Remark 4.6. Note that the two operators T appearing in (4.2) are different objects. The T on the left-hand side is the generator of the endomorphism algebra of $\mathrm{ind}_{KZ}^{G}\sigma$, while the one on the right-hand side generates the endomorphism algebra of $\mathrm{ind}_{KZ}^{G}\sigma'$.

Proof. Let $W = \mathrm{ind}_{KZ}^{G}\sigma / T(\mathrm{ind}_{KZ}^{G}\sigma)$ and let $U \subset W$ be an irreducible G-submodule. By explicit computation, one shows that $W^{I(1)}$ is two-dimensional and that every non-zero element of $W^{I(1)}$ generates W as a G-module. But $U^{I(1)} \neq 0$ by Proposition 2.3 and hence $U = W$. The isomorphisms of (4.2) are constructed explicitly, and one shows that

$$\mathrm{soc}_K(\mathrm{ind}_{KZ}^{G}\sigma / T(\mathrm{ind}_{KZ}^{G}\sigma)) \simeq \sigma \oplus \sigma', \tag{4.3}$$

implying that there are no other isomorphisms. Recall that for a G-module M, the socle $\mathrm{soc}_K(M)$ is the direct sum of all irreducible K-submodules of M. □

If $\rho : G_{\mathbb{Q}_p} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is an irreducible local Galois representation, define an $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(\mathbb{Q}_p)$ by $\pi(\rho) = \mathrm{ind}_{KZ}^{G}\sigma / T(\mathrm{ind}_{KZ}^{G}\sigma)$, where $\sigma \in W(\rho)$. It is immediate from the results just presented that $\pi(\rho)$ is well-defined and that this construction provides a bijection between irreducible Galois representations $\rho : G_{\mathbb{Q}_p} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ and supersingular representations of $\mathrm{GL}_2(\mathbb{Q}_p)$. Note that the following relation is satisfied:

$$\mathrm{soc}_K(\pi(\rho)) = \bigoplus_{\sigma \in W(\rho)} \sigma. \tag{4.4}$$

In the same paper [10], Breuil constructed $\pi(\rho)$ for semisimple reducible ρ, and eventually Colmez defined $\pi(\rho)$ for indecomposable ρ, thereby completing the mod p local Langlands correspondence for $\mathrm{GL}_2(\mathbb{Q}_p)$. These constructions are more complicated than the one presented above, and we will not give them here, nor shall we argue why these definitions of $\pi(\rho)$ are the "correct" ones. However, it is important to note that the property (4.4) remains true for all ρ.

If $F \neq \mathbb{Q}_p$, then almost nothing is known about the mod p local Langlands correspondence for $\mathrm{GL}_2(F)$, and the statements that are known are almost all negative. For instance, we know that there cannot be a bijection between irreducible Galois representations $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ and supersingular representations of $\mathrm{GL}_2(F)$ because there are far too many of the latter. The condition (4.4) does not isolate a $\pi(\rho)$ because, for unramified extensions $F/\mathbb{Q}_p$, Breuil and Paskunas [12] have proved the existence of infinite families of supersingular representations W satisfying $\mathrm{soc}_K(W) \simeq \bigoplus_{\sigma \in W(\rho)} \sigma$. Moreover, although it is immediate from Zorn's Lemma that supersingular representations of $\mathrm{GL}_2(F)$ exist, we do not have a single explicit construction of one; the proof of Breuil and Paskunas uses the theory of diagrams and involves the taking of injective envelopes, which makes their work very non-explicit. In fact, Schraen [49] has shown that if $F/\mathbb{Q}_p$ is quadratic, then no supersingular representation of $\mathrm{GL}_2(F)$ is finitely presented. This makes it difficult to contemplate generalizations of Colmez's construction.

Let e be the ramification index of $F/\mathbb{Q}_p$, and recall that the residue field k of F satisfies $[k : \mathbb{F}_p] = f$. Let F_0 be the maximal unramified subextension of $F/\mathbb{Q}_p$, and observe that, since F and F_0 have the same residue field, the Serre weights for F_0 are the same as those for F. If $\rho : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is irreducible, then (see [47]) one defines e^f irreducible representations $\rho_1, \dots, \rho_{e^f} : G_{F_0} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ such that one expects

$$\mathrm{soc}_K(\pi(\rho)) \simeq \bigoplus_{i=1}^{e^f} \bigoplus_{\sigma \in W(\rho_i)} \sigma.$$

The formula above explains the multiplicities of the different constituents of the K-socle of $\pi(\rho)$. Ongoing work of Breuil and Diamond aims to specify the K-socles of $\pi(\rho)$ for reducible ρ. This section has only scratched the surface of the mod p local Langlands correspondence and has said almost nothing about current research, but we hope that it has sufficiently piqued the reader's interest to consult the literature for more details about the field.

5 Potential modularity and compatible systems

After the digression about mod p local Langlands in the previous section, we return to our discussion of Serre's modularity conjecture. In particular, we return to the notation of Section 2, so that F is now again a totally real number field.

5.1 A wish list

Suppose that we have two mod p Galois representations $\overline{\rho}_1 : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_{p_1})$ and $\overline{\rho}_2 : G_F \to \mathrm{GL}_2(\overline{\mathbb{F}}_{p_2})$, where p_1 and p_2 are two primes, possibly distinct. It clearly would be useful to be able to prove statements of the form "if $\overline{\rho}_1$ is modular and certain conditions are satisfied, then $\overline{\rho}_2$ is modular as well." Such theorems would allow us to leverage knowledge of Serre's conjecture in some special cases to prove it for larger classes of Galois representations.

How can we relate the modularity of two different Galois representations? A crucial idea is to think about p-adic Galois representations, and to recall that when we first encountered them, in Böckle's lectures, they were constructed in families. Indeed, for a modular form f, we obtained a representation $\rho_{f,l} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_l)$ for each prime l. The $\rho_{f,l}$ for different l were very intimately related.

An important starting point for work on modularity is an axiomatization of this phenomenon: the notion of weakly and strongly compatible systems that we saw in Böckle's lectures. A strongly compatible system $\{\rho_l : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_l)\}$ of Galois representations behaves like a family of representations arising from a modular form. In particular, if one member of the system is modular, then they all are, and the same is true of their reductions. This gives us a general strategy for proving the "if $\overline{\rho}_1$ is modular, then $\overline{\rho}_2$ is too" theorems that we wished for at the beginning of this section. Suppose we could find a compatible system $\{\rho_l\}$ such that $\overline{\rho}_1 \simeq \overline{\rho_{p_1}}$ and $\overline{\rho}_2 \simeq \overline{\rho_{p_2}}$. We are assuming that $\overline{\rho}_1$ is modular. If we could somehow prove that ρ_{p_1} is modular, the compatible system would allow us to conclude that ρ_{p_2} is modular as well, and hence that $\overline{\rho}_2$ is modular.

Three major ingredients are involved in implementing this strategy. Starting with a representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$, we have the following **wish list**:

1. Find nice lifts $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ of $\overline{\rho}$. "Nice" will mean that ρ satisfies hypotheses that make the other parts of the wish list available.

2. Given such a lift ρ, embed it in a compatible system $\{\rho_l\}$ such that $\rho \simeq \rho_p$.

3. Modularity lifting theorems.

5.2 Potential modularity of mod p Galois representations

In this section we will sketch a proof of the following theorem of Taylor [58].

Proposition 5.1. *Let $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(k)$ be a continuous, irreducible, odd Galois representation, where k is a finite field of characteristic p. Then there exists a Galois totally real extension $F/\mathbb{Q}$ that is unramified at p and such that $\overline{\rho}_{|G_F}$ is modular.*

If $\overline{\rho}$ has solvable image, then this problem may be handled by the methods of Langlands and Tunnell, so we will assume that this is not the case. Passing to a suitable totally real extension $F/\mathbb{Q}$ that is unramified at p, we may assume that the determinant of $\overline{\rho}_{|G_v}$ is the mod p cyclotomic character for all places $v|p$ of F, and that $\overline{\rho}$ has the following form at all $v|p$:

$$\overline{\rho}_{|G_v} \sim \begin{pmatrix} \varepsilon\chi_v^{-1} & * \\ 0 & \chi_v \end{pmatrix},$$

where ε is the mod p cyclotomic character and $\chi_v : G_v \to k^*$ is a character. Now $\overline{\rho}_{G_F}$ looks like it could be the restriction of an ordinary p-adic representation of Hodge-Tate weights $\{0, 1\}$ coming from an abelian variety, and our task is to show that this is indeed the case.

Recall that if A/F is an abelian variety, then for every finite place v of F, the Galois group G_F acts on the torsion $A[v]$, and the reductions of these representations give us a strictly compatible system $\{\overline{\rho}_{A,v}\}$. We are looking for an abelian variety A/F such that $\overline{\rho}_{|G_F} \simeq \overline{\rho}_{A,v}$ for some v. Let $v' \neq v$ be another place of F lying over p. We will cleverly set up a moduli problem of abelian varieties in such a way that a handy theorem of Morel-Bailly [42], quoted below, will give us exactly the existence of the A that we need.

Let $M/\mathbb{Q}$ be an imaginary quadratic extension and $\psi : G_M \to \overline{\mathbb{Q}}_p^*$ a character. Consider the moduli problem of triples $(A, \varphi_v, \varphi_{v'})$ such that A is a Hilbert-Blumenthal abelian variety (this is an abelian variety carrying an action of the ring of integers $\mathcal{O}_E$ of a specified totally real field E and some additional structure such as a Rosati involution and a polarization; we will not give a precise definition here but only mention that there is a well-developed theory of moduli problems for these objects, founded by Rapoport in [43]), and the φ_v and $\varphi_{v'}$ are isomorphisms $\varphi_v : \overline{\rho}_{|G_F} \xrightarrow{\sim} \overline{\rho}_{A,v}$ and $\varphi_{v'} : \mathrm{Ind}_{G_M}^{G_{\mathbb{Q}}} \overline{\psi} \xrightarrow{\sim} \overline{\rho}_{A,v'}$. The general theory of Hilbert-Blumenthal abelian varieties tells us that this moduli problem is representable by some moduli space $X/\mathbb{Q}$, and if we knew that this space had a rational point, it would correspond to the abelian variety A that we are looking for.

Proposition 5.2 (Moret-Bailly)**.** *Let K be a number field and S a finite set of places of K. If X/K is a geometrically irreducible smooth quasi-projective scheme and $X(K_v) \neq \varnothing$ for all $v \in S$, then $X(K_S)$ is Zariski dense in X. Here K_S/K is the maximal extension of K in which all $v \in S$ split completely.*

By choosing E, v, v', ψ wisely, it can be arranged that the hypotheses of Moret-Bailly's theorem are satisfied for $K = \mathbb{Q}$ and $X/\mathbb{Q}$ the moduli space considered above. In fact, Moret-Bailly's result appears to be far stronger than what we need to prove the existence of a rational point. This gives us the freedom to strengthen Proposition 5.1 by imposing a number of additional properties on the totally real field F, such as requiring it to be linearly disjoint from any specified number field. These strengthenings turn out to be essential, as they allow $\overline{\rho}_{|G_F}$ to satisfy the hypotheses of the modularity lifting theorems that we will call upon later.

5.3 Deformation theory and modularity lifting results

In the previous section we laid out the ingredients of the proof of a potential modularity theorem for mod p representations. Now we want to build on that result to get a potential modularity theorem for p-adic representations, which will be used in Section 5.4. Suppose that we are given a continuous, odd, irreducible mod p Galois representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(k)$, as usual. First of all, we want to find a nice p-adic lift of $\overline{\rho}$ as in the first item of the wish list of Section 5.1.

Consider the following deformation problem. We want to study deformations $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(A)$, where A is a complete local noetherian algebra with residue field k, such that ρ lifts $\overline{\rho}$. In addition, for each prime l we fix an equivalence class τ_l of representations of the inertia group $I_l \simeq G_{\mathbb{Q}_l^{nr}}$, such that all but finitely many of the τ_l are trivial. Let $\chi_p : G_{\mathbb{Q}} \to \overline{\mathbb{Q}}_p^*$ be a character; we have $\chi_p = \omega_p^{k-1}$, where ω_p is the p-adic cyclotomic character. We require that $\det \rho = \chi_p$, that τ_l be the restriction to inertia of the Weil-Deligne representation associated to $\rho_{|G_l}$ for each l, and that $\rho_{|G_p}$ be crystalline of Hodge-Tate weights $\{0, k-1\}$.

By general deformation theory, this deformation problem is represented by a complete noetherian local ring $R^X_{\overline{\rho},\mathbb{Q}}$ with residue field k. It can be proved with very considerable effort, using Galois cohomology and the Euler characteristic formula (see [8]), that $\dim R^X_{\overline{\rho},\mathbb{Q}} \geq 1$ (by the dimension of a ring we mean the Krull dimension). To get our nice lift of $\overline{\rho}$, we need to show that $R^X_{\overline{\rho},\mathbb{Q}}$ has a point over an algebra of characteristic zero.

Consider the Hecke algebra $\mathbf{T} = \mathbb{Z}_p[T_l : l \neq p]$, which acts on the space of modular forms $S_k(\Gamma_1(N(\overline{\rho}))$. We get a natural surjection

$$R^X_{\overline{\rho},\mathbb{Q}} \twoheadrightarrow \mathbf{T}_{\mathfrak{m}_{\overline{\rho}}} \tag{5.1}$$

by the universal property of $R^X_{\overline{\rho},\mathbb{Q}}$. On the other hand, if every deformation classified by $R^X_{\overline{\rho},\mathbb{Q}}$ is modular, then the universal deformation must factor through $\mathbf{T}_{\mathfrak{m}_{\overline{\rho}}}$ and therefore $R^X_{\overline{\rho},\mathbb{Q}} \simeq \mathbf{T}_{\mathfrak{m}_{\overline{\rho}}}$. Proving a modularity lifting theorem, therefore, comes down to proving such an isomorphism, i.e. an "$R = T$" theorem. One of the breakthroughs of the Taylor-Wiles method [61] [59] was the understanding that modularity lifting results could often be reduced to statements about ring-theoretic properties of Hecke algebras. For example, if $R^X_{\overline{\rho},\mathbb{Q}}$ were an integral domain, then proving $\dim R^X_{\overline{\rho},\mathbb{Q}} = \dim \mathbf{T}_{\mathfrak{m}_{\overline{\rho}}}$ would suffice to establish that (5.1) is an isomorphism, since the quotient of an integral domain by a non-trivial ideal has strictly lower Krull dimension than the original ring. In general, Spec $R^X_{\overline{\rho},\mathbb{Q}}$ will have more than one irreducible component, and proving that (5.1) is an isomorphism often amounts to showing that each component overlaps the image of Spec $\mathbf{T}_{\mathfrak{m}_{\overline{\rho}}}$ as well as comparing Krull dimensions.

In fact, we do not know that (5.1) is an isomorphism. However, we know from Proposition 5.1 that $\overline{\rho}_{|G_F}$ is modular for some totally real fields F. We may consider an analogous deformation problem to the one studied above, but over F; it is represented by a deformation ring $R^X_{\overline{\rho},F}$. Moreover, if we choose F correctly, then a modularity lifting theorem is known by work of Diamond [17] and Fujiwara [22]; in that case one can prove that $R^X_{\overline{\rho},F}$ is isomorphic to a suitable localized Hecke algebra $\mathbf{T}_F$.

The map Spec $R^X_{\overline{\rho},\mathbb{Q}} \to$ Spec $R^X_{\overline{\rho},F}$ corresponding to restriction to the subgroup G_F of representations of $G_{\mathbb{Q}}$ is clearly quasi-finite, i.e. has finite fibers. Indeed, G_F has finite index in $G_{\mathbb{Q}}$ and it is not hard to see that there are only finitely many ways to extend a representation of G_F to the larger group $G_{\mathbb{Q}}$. Moreover, the Hecke algebra $\mathbf{T}_F$ is finitely generated as a $\mathbb{Z}_p$-module, since it embeds in the endomorphism algebra of a suitable abelian variety. This implies that $\mathbf{T}_F/(p)$ is a finite set, therefore that $R^X_{\overline{\rho},F}/(p)$ is finite, and therefore that $R^X_{\overline{\rho},\mathbb{Q}}/(p)$ is finite, hence has dimension zero. Hence, any prime ideal of $R^X_{\overline{\rho},\mathbb{Q}}$ containing (p) is necessarily maximal.

On the other hand, recall that $\dim R^X_{\overline{\rho},\mathbb{Q}} \geq 1$. This means that there exists a non-maximal prime ideal $P \subset R^X_{\overline{\rho},\mathbb{Q}}$. By the above, we know that P does not contain (p). Since $R^X_{\overline{\rho},\mathbb{Q}}$ is finitely generated over $\mathbb{Z}_p$, it follows that the quotient $R^X_{\overline{\rho},\mathbb{Q}}/P$ embeds into the ring of integers $\mathcal{O}_L$ of a suitable finite extension $L/\mathbb{Q}_p$. Now by the universal property of $R^X_{\overline{\rho},\mathbb{Q}}$, the embedding $R^X_{\overline{\rho},\mathbb{Q}}/P \hookrightarrow \mathcal{O}_L$ corresponds to a p-adic Galois representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathcal{O}_L)$ lifting $\overline{\rho}$.

We have now achieved the first part of the wish list in Section 5.1. In fact, by doing all of this more carefully we could ensure that the obtained lift ρ has a variety of good properties.

5.4 Constructing compatible systems

In the previous section, we started with a mod p representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ and found a finite extension $L/\mathbb{Q}_p$ and a p-adic representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathcal{O}_L)$ lifting $\overline{\rho}$, thereby fulfilling the first part of the wish list of section 5.1. In this section we will build a compatible system around ρ.

By Taylor's potential modularity theorem (Proposition 5.1), we know that there is a Galois totally real field $F/\mathbb{Q}$ such that $\overline{\rho}_{|G_F}$ is modular. By the modularity lifting theorems of Diamond and Fujiwara that were mentioned in the previous section, we know that $\rho_{|G_F}$ is modular as well. Let $G = \mathrm{Gal}(F/\mathbb{Q})$. By Brauer's theorem (see, for instance, chapter 10 of [53]) there exist solvable subgroups $H_i \subset G$, integers $n_i \in \mathbb{Z}$, and one-dimensional representations χ_i of H_i such that

$$\mathbf{1} = \sum_{i=1}^{t} n_i \mathrm{Ind}_{H_i}^{G} \chi_i, \tag{5.2}$$

in the Grothendieck group of G, where $\mathbf{1}$ is the trivial representation of G. Note that even though $\mathbf{1}$ is a true representation, some of the n_i might be negative. This will cause us problems later. Set F_i to be the fixed field of H_i. Tensoring with ρ, we obtain that

$$\rho = \sum_{i=1}^{t} n_i \mathrm{Ind}_{G_{F_i}}^{G_{\mathbb{Q}}} (\rho_{|G_{F_i}} \otimes \chi_i). \tag{5.3}$$

Since $\mathrm{Gal}(F/F_i) = H_i$ is solvable, we conclude by Langlands-Tunnell solvable base change that each $\rho_{|G_{F_i}}$ arises from an automorphic form π_i on F_i. Hence we can trivially rewrite (5.3) as

$$\rho = \rho_p = \sum_{i=1}^{t} n_i \mathrm{Ind}_{G_{F_i}}^{G_{\mathbb{Q}}} (\rho_{\pi_i,p} \otimes \chi_i). \tag{5.4}$$

Since each $\rho_{\pi_i,p}$ comes from an automorphic form and therefore sits in a compatible system of representations of G_{F_i}, it is very tempting to define

$$\rho_l = \sum_{i=1}^{t} n_i \mathrm{Ind}_{G_{F_i}}^{G_{\mathbb{Q}}} (\rho_{\pi_i,l} \otimes \chi_i) \tag{5.5}$$

for arbitrary primes l. In fact, this idea works. If we knew that the ρ_l were true representations and not just virtual ones, then the compatible system properties of the $\{\rho_{\pi_i,l}\}$ would easily imply that $\{\rho_l\}$ is a compatible system as well. In fact, it can indeed be checked that the $\{\rho_l\}$ are true representations. This fulfills the second part of the wish list.

6 Proof of Serre's conjecture

We are finally in a position to give an exceedingly impressionistic sketch of the strategy behind the proof of Serre's conjecture. For more detail, the reader is referred to Wintenberger's excellent expository article [62] and to Khare's exposition [35], which has somewhat fewer details but paints the big picture in bold strokes. For simplicity, we will only consider the level one case of Serre's conjecture. This means that we start with a Galois representation $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ that is continuous, irreducible, odd, and unramified outside p. Recall from Section 2.5 that the lack of ramification outside p means that the prime-to-p part of the Artin conductor $\mathfrak{n}(\overline{\rho})$ is trivial, and hence $N(\overline{\rho}) = 1$. We aim to prove that $\overline{\rho}$ is modular.

It is important to note that some special cases of Serre's conjecture were known well before Khare's idea of applying Taylor's potential modularity results and Kisin's modularity lifting techniques to this problem. In the 1970's Tate used discriminant bounds to prove that there are no continuous irreducible odd representations $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_2)$, and therefore that the level one case of Serre's modularity conjecture is vacuously true for $p = 2$. Serre extended his argument to $p = 3$ shortly afterwards, and these two results are essential to the work of Khare and Wintenberger, since they constitute the base cases of their induction argument. We note that the level one case of Serre's conjecture for $p = 5$ was proved by Brueggeman [13] contingent on the generalized Riemann hypothesis, and that, with some local hypotheses at small primes but no restriction on the level, the conjecture was proved for $\overline{\rho}$ with image lying in $\mathrm{GL}_2(\mathbb{F}_7)$ by Manoharmayum [40] and for $\overline{\rho}$ with image lying in $\mathrm{GL}_2(\mathbb{F}_9)$ by Ellenberg [20].

To give a taste of the inductive argument that proves Serre's conjecture, and to illustrate its crucial reliance on modularity lifting theorems, we will first flagrantly disregard the current reality and describe what the proof would look like if modularity lifting technology were more advanced than it actually is. Assume the following, for the moment:

Dream 6.1. *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ be continuous, irreducible, unramified outside p, and crystalline at p with Hodge-Tate weights $\{w, 0\}$ for some $w \leq 2p$. Suppose that its reduction $\overline{\rho}$ is modular. Then ρ is modular.*

This dream follows, of course, from the Fontaine-Mazur conjecture. It was considered totally out of reach when Khare and Wintenberger did their work, but such a modularity lifting result has since been proved in most cases by Kisin [38]. In fact, since Serre's conjecture is now known, his work implies most cases of Fontaine-Mazur for two-dimensional representations of $G_{\mathbb{Q}}$. In any case, let us assume the dream and then prove Serre's conjecture in level one.

Let p_n be the n-th prime. We will prove Serre's conjecture by induction on n. It is known for $p_1 = 2$ and $p_2 = 3$ by the theorems of Tate and Serre that were mentioned above. Suppose that it is also known for p_{n-1}. Let $\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_{p_n})$ be continuous, irreducible, odd, and unramified outside p_n. By the methods of Sections 5.3 and 5.4, we can find a lift ρ of $\overline{\rho}$ that sits in a compatible system $\{\rho_l\}$, so that $\rho_{p_n} \simeq \rho$. Consider $\overline{\rho}_{p_{n-1}}$; it is modular by induction. Moreover, by the properties of compatible systems, $\rho_{p_{n-1}}$ is unramified outside p_{n-1} and is crystalline of Hodge-Tate weight $(0, k(\overline{\rho})-1)$. As we saw at the beginning of these notes, up to a twist we can assume that $k(\overline{\rho}) \leq p_n \leq 2p_{n-1}$, where the second inequality is Bertrand's Postulate. By the Dream, $\rho_{p_{n-1}}$ is modular. Hence ρ_{p_n} is modular by the compatible system, and hence $\overline{\rho}$ is modular and we are done.

The powerful modularity lifting theorem of the Dream can be seen as a fulfillment of the third part of the wish list of Section 5.1. Even though the Dream is not yet known, the modularity lifting theorems available to Khare and Wintenberger in 2005 were enough to prove Serre's conjecture, albeit with lots of technical work. The modularity lifting theorems available now, and still more those available then, come with long lists of technical hypotheses, and one must be very careful to ensure that the liftings of $\overline{\rho}$ and the compatible systems obtained from the methods of Sections 5.3 and 5.4 satisfy these. In these notes we have entirely ignored these technical points, which complicate the work tremendously. However, at its core the basic idea is the simple one presented here.

We conclude with the unfortunate observation that it does not appear to be possible, at least not without a major new idea, to generalize the beautiful argument of Khare and Wintenberger to obtain a proof of the generalizations of Serre's conjecture to totally real fields that were incorporated into Conjecture 2.9 above. While all the ingredients of their proof – potential modularity, construction of lifts, compatible systems, modularity lifting theorems – are less developed for arbitrary totally real fields than for $\mathbb{Q}$, a more fundamental problem is that any inductive argument on the places of F would require that enough base cases be proved first, and it is not yet known how to prove them. Recall that Tate and Serre proved (the level one case of) Serre's conjecture for $p = 2$ and $p = 3$ by showing that it was vacuously true, i.e. that there were no $\overline{\rho}$ that were continuous, irreducible, odd, and unramified outside p. While analogous non-existence theorems have been proved for some small primes and a few specific quadratic real fields (see, for instance, [41] and [50]), we know that for general totally real fields, even quadratic ones, Serre's conjecture is never vacuously true. Indeed, for general totally real fields, for all p there exist continuous, irreducible, odd mod p Galois representations that are unramified outside p; see the introduction to [14] for an example in the case of $F = \mathbb{Q}(\sqrt{29})$. Thus, to get the base case for an induction argument, one would need to establish a sufficiently large number of non-vacuous cases of Serre's conjecture, and it is not clear at all at the present time how to attack this problem. Serre's modularity conjecture will likely continue to be an important motivation and source of research problems for some time to come.

References

[1] Noriyuki Abe. On a classification of irreducible admissible modulo p representations of a p-adic split reductive group. Preprint, available at `http://arxiv.org/pdf/1103.2525v3`.

[2] Avner Ash and Glenn Stevens. Cohomology of arithmetic groups and congruences between systems of Hecke eigenvalues. *J. Reine Angew. Math.*, 365:192–220, 1986.

[3] Thomas Barnet-Lamb, Toby Gee, and David Geraghty. Serre weights for rank two unitary groups. Preprint, 2012.

[4] L. Barthel and R. Livné. Irreducible modular representations of GL_2 of a local field. *Duke Math. J.*, 75(2):261–292, 1994.

[5] Laurent Berger. Central characters for smooth irreducible modular representations of $\mathrm{GL}_2(\mathbf{Q}_p)$. *Rendiconti del Seminario Matematico della Università di Padova*, 127, 2012.

[6] Tobias Berger and Gergely Harcos. l-adic representations associated to modular forms over imaginary quadratic fields. *Int. Math. Res. Not. IMRN*, (23):Art. ID rnm113, 16, 2007.

[7] Tobias Berger and Krzysztof Klosin. An $R = T$ theorem for imaginary quadratic fields. *Math. Ann.*, 349(3):675–703, 2011.

[8] Gebhard Böckle. A local-to-global principle for deformations of Galois representations. *J. Reine Angew. Math.*, 509:199–236, 1999.

[9] Nigel Boston, Hendrik W. Lenstra, Jr., and Kenneth A. Ribet. Quotients of group rings arising from two-dimensional representations. *C. R. Acad. Sci. Paris Sér. I Math.*, 312(4):323–328, 1991.

[10] Christophe Breuil. Sur quelques représentations modulaires et p-adiques de $\mathrm{GL}_2(\mathbf{Q}_p)$. I. *Compositio Math.*, 138(2):165–188, 2003.

[11] Christophe Breuil and Ariane Mézard. Multiplicités modulaires et représentations de $\mathrm{GL}_2(\mathbf{Z}_p)$ et de $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ en $l = p$. *Duke Math. J.*, 115(2):205–310, 2002. With an appendix by Guy Henniart.

[12] Christophe Breuil and Vytautas Paskunas. Towards a modulo p Langlands correspondence for GL_2. *Memoirs Amer. Math. Soc.*, 216, 2012.

[13] Sharon Brueggeman. The nonexistence of certain Galois extensions unramified outside 5. *J. Number Theory*, 75(1):47–52, 1999.

[14] Kevin Buzzard, Fred Diamond, and Frazer Jarvis. On Serre's conjecture for mod l Galois representations over totally real fields. *Duke Math. J.*, 55:105–161, 2010.

[15] Frank Calegari and David Geraghty. Modularity lifting beyond the Taylor-Wiles method. Preprint.

[16] Henri Carayol. Sur la mauvaise réduction des courbes de Shimura. *Compositio Math.*, 59(2):151–230, 1986.

[17] Fred Diamond. The Taylor-Wiles construction and multiplicity one. *Invent. Math.*, 128(2):379–391, 1997.

[18] Bas Edixhoven. The weight in Serre's conjectures on modular forms. *Invent. Math.*, 109(3):563–594, 1992.

[19] Bas Edixhoven. Serre's conjecture. In *Modular forms and Fermat's last theorem (Boston, MA, 1995)*, pages 209–242. Springer, New York, 1997.

[20] Jordan S. Ellenberg. Serre's conjecture over $\mathbb{F}_9$. *Ann. of Math. (2)*, 161(3):1111–1142, 2005.

[21] Matthew Emerton, Toby Gee, and Florian Herzig. Explicit Serre weight conjectures. In preparation.

[22] Kazuhiro Fujiwara. Galois deformations and arithmetic geometry of Shimura varieties. In *International Congress of Mathematicians. Vol. II*, pages 347–371. Eur. Math. Soc., Zürich, 2006.

[23] Toby Gee. On the weights of mod p Hilbert modular forms. *Invent. Math.*, 184:1–46, 2011.

[24] Toby Gee and Mark Kisin. The Breuil-Mézard conjecture for potentially Barsotti-Tate representations. Preprint, 2012.

[25] Toby Gee, Tong Liu, and David Savitt. Crystalline extensions and the weight part of Serre's conjecture. *Algebra and Number Theory*, To appear.

[26] Toby Gee and David Savitt. Serre weights for mod p Hilbert modular forms: the totally ramified case. *J. Reine Angew. Math.*, 660:1–26, 2011.

[27] J. A. Green. The characters of the finite general linear groups. *Trans. Amer. Math. Soc.*, 80:402–447, 1955.

[28] Michael Harris, David Soudry, and Richard Taylor. l-adic representations associated to modular forms over imaginary quadratic fields. I. Lifting to $\mathrm{GSp}_4(\mathbf{Q})$. *Invent. Math.*, 112(2):377–411, 1993.

[29] Michael Harris and Richard Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich.

[30] Guy Henniart. Une preuve simple des conjectures de Langlands pour $\mathrm{GL}(n)$ sur un corps p-adique. *Invent. Math.*, 139(2):439–455, 2000.

[31] Florian Herzig. The weight in a Serre-type conjecture for tame n-dimensional Galois representations. *Duke Math. J.*, 149:37–116, 2009.

[32] Florian Herzig. The classification of irreducible admissible mod p representations of a p-adic GL_n. *Invent. Math.*, 186:373–434, 2011.

[33] Florian Herzig. A Satake isomorphism in characteristic p. *Compos. Math.*, 147(1):263–283, 2011.

[34] Frazer Jarvis. On Galois representations associated to Hilbert modular forms of low weight. *J. Reine Angew. Math.*, 491:199–216, 1997.

[35] Chandrashekhar Khare. Serre's modularity conjecture: a survey of the level one case. In *L-functions and Galois representations*, volume 320 of *London Math. Soc. Lecture Note Ser.*, pages 270–299. Cambridge Univ. Press, Cambridge, 2007.

[36] Chandrashekhar Khare and Jean-Pierre Wintenberger. Serre's modularity conjecture. I. *Invent. Math.*, 178(3):485–504, 2009.

[37] Chandrashekhar Khare and Jean-Pierre Wintenberger. Serre's modularity conjecture. II. *Invent. Math.*, 178(3):505–586, 2009.

[38] Mark Kisin. The Fontaine-Mazur conjecture for GL_2. *J. Amer. Math. Soc.*, 22(3):641–690, 2009.

[39] Mark Kisin. Modularity of 2-adic Barsotti-Tate representations. *Invent. Math.*, 178(3):587–634, 2009.

[40] Jayanta Manoharmayum. Serre's conjecture for mod 7 Galois representations. In *Modular curves and abelian varieties*, volume 224 of *Progr. Math.*, pages 141–149. Birkhäuser, Basel, 2004.

[41] Hyunsuk Moon and Yuichiro Taguchi. The non-existence of certain mod 2 Galois representations of some small quadratic fields. *Proc. Japan Acad. Ser. A Math. Sci.*, 84(5):63–67, 2008.

[42] Laurent Moret-Bailly. Groupes de Picard et problèmes de Skolem. I, II. *Ann. Sci. École Norm. Sup. (4)*, 22(2):161–179, 181–194, 1989.

[43] Michael Rapoport. Compactifications de l'espace de modules de Hilbert-Blumenthal. *Compositio Math.*, 36(3):255–335, 1978.

[44] Michel Raynaud. Schémas en groupes de type $(p, \dots, p)$. *Bull. Soc. Math. France*, 102:241–280, 1974.

[45] Michael M. Schein. Weights in Serre's conjecture for Hilbert modular forms: the ramified case. *Israel J. Math.*, 166:369–391, 2008.

[46] Michael M. Schein. Weights of Galois representations associated to Hilbert modular forms. *J. Reine Angew. Math*, 622:57–94, 2008.

[47] Michael M. Schein. Reduction modulo p of cuspidal representations and weights in Serre's conjecture. *Bull. London Math. Soc.*, 41:147–154, 2009.

[48] Michael M. Schein. Weights in generalizations of Serre's conjecture and the mod p local Langlands correspondence. In *Symmetries in algebra and number theory*, pages 71–93. Universitätsverlag Göttingen, Göttingen, 2009.

[49] Benjamin Schraen. Sur la présentation des répresentations supersingulières de $\mathrm{GL}_2(F)$. Preprint, 2012.

[50] Mehmet Haluk Şengün. The nonexistence of certain representations of the absolute Galois group of quadratic fields. *Proc. Amer. Math. Soc.*, 137(1):27–35, 2009.

[51] Jean-Pierre Serre. *Corps locaux.* Hermann, Paris, 1968. Deuxième édition, Publications de l'Université de Nancago, No. VIII.

[52] Jean-Pierre Serre. Valeurs propres des opérateurs de Hecke modulo l. In *Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, 1974)*, pages 109–117. Astérisque, Nos. 24–25. Soc. Math. France, Paris, 1975.

[53] Jean-Pierre Serre. *Linear representations of finite groups.* Springer-Verlag, New York, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

[54] Jean-Pierre Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. J.*, 54(1):179–230, 1987.

[55] Christopher Skinner. A note on the p-adic Galois representations attached to Hilbert modular forms. *Doc. Math.*, 14:241–258, 2009.

[56] J. Tate. Number theoretic background. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 3–26. Amer. Math. Soc., Providence, R.I., 1979.

[57] Richard Taylor. l-adic representations associated to modular forms over imaginary quadratic fields. II. *Invent. Math.*, 116(1-3):619–643, 1994.

[58] Richard Taylor. Remarks on a conjecture of Fontaine and Mazur. *J. Inst. Math. Jussieu*, 1(1):125–143, 2002.

[59] Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995.

[60] Marie-France Vignéras. La conjecture de Langlands locale pour $\mathrm{GL}(n, F)$ modulo l quand $l \neq p,\ l > n$. *Ann. Sci. École Norm. Sup. (4)*, 34(6):789–816, 2001.

[61] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995.

[62] Jean-Pierre Wintenberger. La conjecture de modularité de Serre: le cas de conducteur (d'après C. Khare). *Astérisque*, (311):Exp. No. 956, viii, 99–121, 2007. Séminaire Bourbaki. Vol. 2005/2006.

Michael M. Schein
Department of Mathematics, Bar-Ilan University,
Ramat Gan 52900, Israel
mschein@math.biu.ac.il

Travaux mathématiques, Volume 23 (2013), 173–214, © by the authors

Model completeness of valued PAC fields

by Wulf-Dieter Geyer and Moshe Jarden[1]

Abstract

We present a theorem of Kollár on the density property of valued PAC fields and a theorem of Abraham Robinson on the model completeness of the theory of algebraically closed non-trivial valued fields. Then we prove that the theory T of non-trivial valued fields in an appropriate first order language has a model completion $\tilde{T}$. The models of $\tilde{T}$ are non-trivial valued fields (K, v) that are ω-imperfect, ω-free, and PAC.

MSC (2010): 12E30.

[1] Research supported by the Minkowski Center for Geometry at Tel Aviv University, established by the Minerva Foundation.

Contents

Introduction

A field K is said to be **PAC** (**pseudo algebraically closed**) if every absolutely irreducible variety V defined over K (i.e. a geometrically integral K-scheme) has a K-rational point. Here and throughout the paper we use $\tilde{K}$ to denote a fixed algebraic closure of K.

The notion of PAC fields has been introduced in [Ax68] (although not by this name) in connection with the decidability of the elementary theory of finite fields. Each countable Hilbertian field K has an abundance of separable algebraic extensions of K which are PAC. Indeed, for each positive integer e and for almost all $\boldsymbol{\sigma} \in \mathrm{Gal}(K)^e$, the fixed field $K_s(\boldsymbol{\sigma})$ is PAC [FrJ08, Thm. 18.6.1]. Here K_s is the separable algebraic closure of K, $\mathrm{Gal}(K) = \mathrm{Gal}(K_s/K)$ is the absolute Galois group of K, "almost all" is meant in the sense of the Haar measure of $\mathrm{Gal}(K)^e$ with respect to its Krull topology, and $K_s(\boldsymbol{\sigma})$ is the fixed field in K_s of the coordinates of $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_e)$.

Chapter 11 of [FrJ08] gives an extensive treatment of PAC fields. In particular, it points out that if K is PAC, then $V(K)$ is Zariski dense in $V(\tilde{K})$ for each absolutely irreducible variety V defined over K [FrJ08, p. 192, Prop. 11.1.1] and asks whether $V(K)$ is even v-dense in $V(\tilde{K})$ for each valuation v of $\tilde{K}$ [FrJ05, Problem 11.5.4]. If this happens, we say that K has the **density property**.

The latter problem goes back to [GeJ75, Problem 1], where the following theorem is proved: Let K be a countable Hilbertian field and e a positive integer. Then for every valuation v of $\tilde{K}$, for almost all $\boldsymbol{\sigma} \in \mathrm{Gal}(K)^e$, and for every absolutely irreducible variety V defined over K, the set $V(K_s(\boldsymbol{\sigma}))$ is v-dense in $V(\tilde{K})$ [GeJ75, Thm. 6.2]. Note that the order of the quantifiers "for every valuation v" and "for almost all $\boldsymbol{\sigma} \in \mathrm{Gal}(K)^e$" can not be exchanged without a substantial argument, because $\tilde{K}$ has in general uncountably many valuations. That argument is supplied in [FrJ76], where the "stability of PAC fields" is proved [FrJ76, Thm. 3.4]. As a result, it is proved that $K_s(\boldsymbol{\sigma})$ has the density property for almost all $\boldsymbol{\sigma} \in \mathrm{Gal}(K)^e$.

For a general PAC field K and an arbitrary valuation v of $\tilde{K}$, Prestel proved that K is v-dense in $\tilde{K}$ [FrJ08, p. 204, Prop. 11.5.3]. The proof is based on the observation that if $f \in K[X]$ is a nonconstant separable polynomial and $c \in K^\times$, then $f(X_1)f(X_2) - c^2$ is an absolutely irreducible polynomial. Thus, there exist $x_1, x_2 \in K$ with $f(x_1)f(x_2) = c^2$, so $v(f(x_1)) \geq v(c)$ or $v(f(x_2)) \geq v(c)$.

János Kollár refined Prestel's trick and proved that every PAC field has the density property [Kol07, Thm. 2]. Using the stability property of PAC fields with an extra condition (which Kollár proves), he reduces the theorem to proving that if K is a PAC field and $f \in K[X, Y]$ is an absolutely irreducible polynomial which is Galois in Y, then one can approximate each zero $(0, \tilde{b}) \in \tilde{K}^2$ of f by a zero $(a, b) \in K^2$. The main point is to find $c \in K^\times$ with $v(c)$ large such that the algebraic set defined by $f(X_1, Y_1) = 0$, $f(X_2, Y_2) = 0$, and $X_1X_2 = c^2$ is an absolutely irreducible variety defined over K. This is done by using a lemma of Enriques-Severi-Zariski followed by smoothness arguments.

The first goal of this note is to give a self contained presentation of Kollár's proof. Then we present Abraham Robinson's proof that the theory of non-trivial algebraically closed valued field is model complete. Finally, we apply the density property of PAC fields and Robinson's result to prove that the elementary theory of PAC valued fields, in an appropriate first order language, is itself model complete. Moreover, it admits elimination of quantifiers.

Acknowledgements: The authors are indebted to Dan Haran for a useful discussion on the subject and to Aharon Razon for a critical reading of the manuscript.

1 Convention

We follow [Wei62, Section I.1] and choose a **universal extension** $\mho$ of K that contains the algebraic closure $\tilde{K}$ of K. Thus, $\mho$ is an algebraically closed field containing $\tilde{K}$ with trans.deg$(\mho/K) = \infty$. For each non-negative integer n and every field $K \subseteq L \subseteq \mho$ we follow the classical algebraic geometry and consider $\mathbb{A}^n(L)$ as the set of all **points** $\mathbf{a} = (a_1, \ldots, a_n)$ with coordinates $a_1, \ldots, a_n \in L$. Likewise we consider $\mathbb{P}^n(L)$ as the set of all points $\mathbf{a} = (a_0{:}a_1{:}\cdots{:}a_n)$ which are, as usual, equivalence classes of $(n+1)$-tuples $(a_0, a_1, \ldots, a_n)$ of elements of $\mho$ modulo multiplication by a non-zero element of $\mho$ such that there exists $0 \leq i \leq n$ with $a_i \neq 0$ and $\frac{a_j}{a_i} \in L$ for $j = 0, \ldots, n$. In this case $K\big(\frac{a_0}{a_i}, \frac{a_1}{a_i}, \ldots, \frac{a_n}{a_i}\big)$ is the **residue field** of $\mathbf{a}$. The elements $a_0, a_1, \ldots, a_n$ are **homogeneous coordinates** of $\mathbf{a}$.

Next we consider the **affine** n**-dimensional space**

$$\mathbb{A}^n_K = \mathrm{Spec}(K[X_1, \ldots, X_n])$$

over K and the **projective** n**-dimensional space**

$$\mathbb{P}^n_K = \mathrm{Proj}(K[X_0, \ldots, X_n])$$

over K. We say that V is an **absolutely irreducible variety in** $\mathbb{A}^n_K$ **(resp.** $\mathbb{P}^n_K$**) defined over** K, if V is a Zariski-closed subscheme of $\mathbb{A}^n_K$ (resp. $\mathbb{P}^n_K$) such that the scheme $V \times_K \mho$ obtained from V by extending the field of scalars from K to $\mho$ is integral. Equivalently, $V \times_K \tilde{K}$ is an integral scheme.

If $V = \mathrm{Spec}(K[X_1, \ldots, X_n]/I)$ (resp. $V = \mathrm{Proj}(K[X_0, \ldots, X_n]/I)$) is an absolutely irreducible variety in $\mathbb{A}^n_K$ (resp. $\mathbb{P}^n_K$) defined over K, then for each field $K \subseteq L \subseteq \mho$, we consider $V(L)$ as the set of all zeros $\mathbf{a} \in \mathbb{A}^n(L)$ (resp. $\mathbf{a} \in \mathbb{P}^n(L)$) of I. Note that there is a natural bijective correspondence between $V(L)$ and $\mathrm{Mor}_K(\mathrm{Spec}(L), V)$. In particular, we may identify each point $\mathbf{a} \in V(K)$ with a unique scheme theoretic K-rational point of V (i.e. a point of V whose residue field is K).

A point $\mathbf{x} \in V(\mho)$ is a **generic point** of V over K if the field $F = K(\mathbf{x})$ is regular over K and trans.deg$(F/K) = \dim(V)$. In this case, F is **the function field of** V over K. Note that F is unique up to a K-isomorphism. Moreover, it is always possible to choose the homogeneous coordinates of $\mathbf{x}$ in F. Indeed, if in the projective case $\mathbf{x} = (x_0{:}\cdots{:}x_n)$ and $x_i \neq 0$, then $F = K(x_i^{-1}x_0, \ldots, x_i^{-1}x_n)$.

With this notation, a reduced closed subscheme V of $\mathbb{A}^n_K$ is an absolutely irreducible variety in $\mathbb{A}^n_K$ defined over K if and only if the scheme $V \times_K \mho$ (alternatively $V \times_K \tilde{K}$) is irreducible and the ideal of $\mho[X_1, \ldots, X_n]$ (alternatively $\tilde{K}[X_1, \ldots, X_n]$) of all polynomials that vanish on $V(\mho)$ (alternatively $V(\tilde{K})$) is generated by polynomials with coefficients in K.

Note that if $p = \mathrm{char}(K) > 0$, $a \in K$ has no pth root in K, and we set $V = \mathrm{Spec}(K[X]/K[X](X^p - a))$, then $V(\tilde{K})$ consists of one point, namely $\sqrt[p]{a}$. In particular $V_{\tilde{K}}$ is irreducible. However, the polynomial $X - \sqrt[p]{a}$ vanishes on $V(\tilde{K})$ but does not belong to $\tilde{K}[X](X^p - a)$, so V is not an absolutely irreducible variety in $\mathbb{A}^1_K$ defined over K.

A reduced closed subscheme V of $\mathbb{P}^n_K$ is **an absolutely irreducible variety in $\mathbb{P}^n_K$ which is defined over** K if and only if each of the standard affine open subsets of V is an absolutely irreducible variety in $\mathbb{A}^n_K$ defined over K.

By [FrJ08, p. 175, Cor. 10.2.2(a)] or [GoW10, p. 136, Prop. 5.51], a reduced irreducible closed subscheme V of $\mathbb{A}^n_K$ or $\mathbb{P}^n_K$ is an absolutely irreducible variety which is defined over K if and only if the function field F of V is a regular extension of K, i.e. K is algebraically closed in K and F/K is separable.

Thus, our definition of an "absolutely irreducible variety in $\mathbb{A}^n_K$ (resp. $\mathbb{P}^n_K$) defined over K" is equivalent to the definition of the same notion in the classical language of algebraic geometry (see also [FrJ08, Sec. 10.2]).

For each absolutely irreducible variety V defined over K we write $\tilde{V}$ for the variety $V \times_K \tilde{K}$ obtained from V by extending the base field from K to $\tilde{K}$. We also say that $\tilde{V}$ is **defined over** K. If $\varphi\colon V \to W$ is a morphism of absolutely irreducible varieties defined over K and L is a field extension of K in $\mho$, we abuse notation and write $\varphi\colon V(L) \to W(L)$ also for the set theoretic map induced from the morphism φ. Finally we write $\tilde{\varphi}\colon \tilde{V} \to \tilde{W}$ for the morphism obtained from φ by extending the base field from K to $\tilde{K}$.

2 Conics in $\mathbb{P}^1_K \times \mathbb{P}^1_K$

We consider the direct product $\mathbb{P}^1_K \times \mathbb{P}^1_K$ of two copies of the projective line over an arbitrary field K, define a pencil of conics in that space, and then blow it up at two points.

2.1 The Conics $H_{\mathbf{a}}$

We consider two copies of $\mathbb{P}^1_K$, one with homogeneous coordinates $(X_0{:}X_1)$ and the other one with homogeneous coordinates $(Y_0{:}Y_1)$. For each $\mathbf{a} = (a_0{:}a_1) \in \mathbb{P}^1(\tilde{K})$ we associate the conic $\tilde{H}_{\mathbf{a}}$ in $\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ defined by the bi-homogeneous equation

$$a_1 X_0 Y_0 = a_0 X_1 Y_1. \tag{2.1}$$

If $\mathbf{a} \in \mathbb{P}^1(K)$, we denote by $H_{\mathbf{a}}$ the conic in $\mathbb{P}^1_K \times \mathbb{P}^1_K$ defined by (2.1). In this case we also have $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$ and $\tilde{H}_{\mathbf{a}} = H_{\mathbf{a}} \times_K \tilde{K}$, in accordance with our convention.

The scheme $\mathbb{P}^1 \times \mathbb{P}^1$ is covered by the open subsets

$$U^{ij} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{P}^1 \times \mathbb{P}^1 \mid x_i y_j \neq 0\}, \qquad i, j = 0, 1$$

which are isomorphic to the affine plane $\mathbb{A}^2$ with the affine coordinates $X_{i'} = \frac{X_{i'}}{X_i}$ and $Y_{j'} = \frac{Y_{j'}}{Y_j}$, where $\{i, i'\} = \{j, j'\} = \{0, 1\}$. Therefore, the conic $\tilde{H}_{\mathbf{a}}$ is covered by the open affine subsets $\tilde{H}^{ij}_{\mathbf{a}} = U^{ij} \cap \tilde{H}_{\mathbf{a}}$ for $i, j = 0, 1$. As subsets of $\mathbb{A}^2$ the latter subsets are defined by the following equations:

$$\begin{aligned} &\tilde{H}^{00}_{\mathbf{a}}: \quad a_1 = a_0 X_1 Y_1, && \tilde{H}^{01}_{\mathbf{a}}: \quad a_1 Y_0 = a_0 X_1 \\ &\tilde{H}^{10}_{\mathbf{a}}: \quad a_1 X_0 = a_0 Y_1, && \tilde{H}^{11}_{\mathbf{a}}: \quad a_1 X_0 Y_0 = a_0. \end{aligned} \tag{2.2}$$

Thus, $\tilde{H}^{ij}_{\mathbf{a}}$ is a line or a hyperbola. Hence,

(3.3) if $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$ is not in the set $B = \{(1{:}0), (0{:}1)\}$ of the base points of $\mathbb{P}^1$, i.e. if $a_0 a_1 \neq 0$, then $\tilde{H}_{\mathbf{a}}$ is a smooth absolutely irreducible curve defined over $K(\mathbf{a})$.

The origins of the affine planes U^{01} and U^{10} are the points

$$\mathbf{q}_1 = ((1{:}0), (0{:}1)), \qquad \text{respectively} \qquad \mathbf{q}_2 = ((0{:}1), (1{:}0)),$$

which are the only points $(\mathbf{x}, \mathbf{y}) \in \mathbb{P}^1 \times \mathbb{P}^1$ with $x_0 y_0 = 0 = x_1 y_1$. They satisfy:

(3.4) For each $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$ we have $\mathbf{q}_1, \mathbf{q}_2 \in \tilde{H}_{\mathbf{a}}$. Moreover, if $\mathbf{a} \notin B$, then the slope of the tangent of $\tilde{H}_{\mathbf{a}}$ at $\mathbf{q}_1$ (resp. $\mathbf{q}_2$) is $\frac{a_0}{a_1}$ (resp. $\frac{a_1}{a_0}$) if $a_1 \neq 0$ (resp. $a_0 \neq 0$).

(3.5) The conic $\tilde{H}_{(1:0)}$ is defined by the equation $X_1 Y_1 = 0$. Let $L_1 = (1{:}0) \times \mathbb{P}^1_{\tilde{K}}$ be the line defined by $X_1 = 0$ and $L_2 = \mathbb{P}^1_{\tilde{K}} \times (1{:}0)$ be the line defined by $Y_1 = 0$. Then L_1 goes through $\mathbf{q}_1$, L_2 goes through $\mathbf{q}_2$, $\tilde{H}_{(1:0)} = L_1 \cup L_2$, and both lines go through $((1{:}0), (1{:}0))$ which is therefore a node of $\tilde{H}_{(1:0)}$ and actually its only singular point.

(3.6) The conic $\tilde{H}_{(0:1)}$ is defined by the equation $X_0Y_0 = 0$. Let $L'_1 = (0{:}1) \times \mathbb{P}^1_{\tilde{K}}$ be the line defined by $X_0 = 0$, $L'_2 = \mathbb{P}^1_{\tilde{K}} \times (0{:}1)$ be the line defined by $Y_0 = 0$. Then L'_1 goes through $\mathbf{q}_2$, L'_2 goes through $\mathbf{q}_1$, $\tilde{H}_{(0:1)} = L'_1 \cup L'_2$, and both lines go through $((0{:}1),(0{:}1))$ which is therefore a node of $\tilde{H}_{(0:1)}$ and actually its only singular point.

Summing up the information about the tangents of the $\tilde{H}_{\mathbf{a}}$'s at $\mathbf{q}_1$ and $\mathbf{q}_2$, we have:

(3.7) Let $i \in \{1,2\}$. If $\mathbf{a}$ and $\mathbf{a}'$ are distinct points of $\mathbb{P}^1(\tilde{K})$, then the tangents of $\tilde{H}_{\mathbf{a}}$ and $\tilde{H}_{\mathbf{a}'}$ at $\mathbf{q}_i$ are distinct. Moreover, as $\mathbf{a}$ ranges over all points of $\mathbb{P}^1(\tilde{K})$, the tangents of $\tilde{H}_{\mathbf{a}}$ at $\mathbf{q}_1$ (resp. $\mathbf{q}_2$) form the full pencil of lines through $\mathbf{q}_1$ (resp. $\mathbf{q}_2$) in U^{01} (resp. in U^{10}).

(3.8) If $\mathbf{a} \neq \mathbf{a}'$, then $\mathbf{q}_1$ and $\mathbf{q}_2$ are the only points of intersection of $\tilde{H}_{\mathbf{a}}$ and $\tilde{H}_{\mathbf{a}'}$.

Indeed, suppose $\mathbf{q} \in \tilde{H}_{\mathbf{a}} \cap \tilde{H}'_{\mathbf{a}}$ with $\mathbf{q} = ((x_0{:}x_1),(y_0{:}y_1)) \neq \mathbf{q}_1, \mathbf{q}_2$ and $\mathbf{a} \neq \mathbf{a}'$. Then,

(3.8a) $a_1x_0y_0 = a_0x_1y_1$ and $a'_1x_0y_0 = a'_0x_1y_1$,

and

(3.8b) $x_0y_0 \neq 0$ or $x_1y_1 \neq 0$ (otherwise $x_0 = y_1 = 0$ and $\mathbf{q} = \mathbf{q}_2$ or $y_0 = x_1 = 0$ and $\mathbf{q} = \mathbf{q}_1$).

If $\mathbf{a}, \mathbf{a}' \notin B$, then $\frac{a_1}{a_0}x_0y_0 = \frac{a'_1}{a'_0}x_0y_0$ and $\frac{a_0}{a_1}x_1y_1 = \frac{a'_0}{a'_1}x_1y_1$, hence $\mathbf{a} = \mathbf{a}'$. If $\mathbf{a} = (1{:}0)$, then $x_1y_1 = 0$ and $a'_1 \neq 0$, so $x_0y_0 = 0$ in contrast to (3.8b). If $\mathbf{a} = (0{:}1)$, then $x_0y_0 = 0$ and $a'_0 \neq 0$, so $x_1y_1 = 0$, in contrast to (3.8b). Similarly, the assumption $\mathbf{a}' \in B$ contradicts (3.8b).

(3.9) $\mathbb{P}^1(\tilde{K}) \times \mathbb{P}^1(\tilde{K}) = \bigcup_{\mathbf{a} \in \mathbb{P}^1(\tilde{K})} \tilde{H}_{\mathbf{a}}(\tilde{K})$.

Indeed, if $((x_0{:}x_1),(y_0{:}y_1)) \neq \mathbf{q}_1, \mathbf{q}_2$, then $x_0y_0 \neq 0$ or $x_1y_1 \neq 0$. Hence, the equality $(x_1y_1)x_0y_0 = (x_0y_0)x_1y_1$ implies that $((x_0{:}x_1),(y_0{:}y_1)) \in \tilde{H}_{(x_0y_0:x_1y_1)}$. This together with (3.8) proves (3.9).

Lemma 2.1. *Let $\varphi\colon H' \to H$ be a birational surjective morphism of a curve H' onto a normal curve H over $\tilde{K}$. Then φ is an isomorphism.*

Proof. Let F be the common field of functions of H and H' over $\tilde{K}$. Consider $\mathbf{p}' \in H'(\tilde{K})$ and let $\mathbf{p} = \varphi(\mathbf{p}')$. Then $\mathcal{O}_{H,\mathbf{p}} \subseteq \mathcal{O}_{H',\mathbf{p}'} \subset F$. By assumption, $\mathcal{O}_{H,\mathbf{p}}$ is a discrete valuation ring and $\mathcal{O}_{H',\mathbf{p}'}$ a proper local ring of F. Hence, $\mathcal{O}_{H,\mathbf{p}} = \mathcal{O}_{H',\mathbf{p}'}$. Therefore, φ is an isomorphism. □

2.2 Blowing up

We blow up $\mathbb{P}^1_K\times\mathbb{P}^1_K$ at the set $\{\mathbf{q}_1,\mathbf{q}_2\}$ to obtain a surface S in $(\mathbb{P}^1_K\times\mathbb{P}^1_K)\times\mathbb{P}^1_K\times\mathbb{P}^1_K$ such that the projection $\sigma\colon S\to\mathbb{P}^1_K\times\mathbb{P}^1_K$ on the first factor is a birational projective K-morphism, (the second factor $\mathbb{P}^1_K$ comes from blowing up at $\mathbf{q}_1$, and the third factor $\mathbb{P}^1_K$ comes from blowing up at $\mathbf{q}_2$). The morphism σ has the following properties [Mum88, pp. 219-225]:

(3.10a) S is an absolutely irreducible surface defined over K. We set $\tilde S = S\times_K\tilde K$ but use σ to denote also the map $\tilde S\to\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K}$ obtained by extending the base field from K to $\tilde K$. Note that $\tilde S(\tilde K)$ can be naturally identified with $S(\tilde K)$.

(3.10b) The restriction of σ to $S\smallsetminus\sigma^{-1}(\{\mathbf{q}_1,\mathbf{q}_2\})$ is an isomorphism onto $\mathbb{P}^1_K\times\mathbb{P}^1_K\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\}$. In particular, over each point $(\mathbf{a},\mathbf{a}')\in\mathbb{P}^1(\tilde K)\times\mathbb{P}^1(\tilde K)$ not in $\{\mathbf{q}_1,\mathbf{q}_2\}$ there lies a unique point of $S(\tilde K)$ which we also denote by $(\mathbf{a},\mathbf{a}')$.

(3.10c) For $i=1,2$, the fiber $\sigma^{-1}(\mathbf{q}_i)$ is of dimension 1, indeed the fiber is isomorphic to $\mathbb{P}^1_K$.

(3.10d) For each $\mathbf{a}\in\mathbb{P}^1(\tilde K)$ let $H'_\mathbf{a}$ be the Zariski-closure in $\tilde S$ of $\sigma^{-1}(\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\})$. Then σ maps $H'_\mathbf{a}$ isomorphically onto $\tilde H_\mathbf{a}$.

Proof of (3.10d). Since σ is a morphism, it is Zariski-continuous. Hence, using a bar to denote the Zariski-closure, we have

$$\sigma(H'_\mathbf{a})=\sigma\big(\overline{\sigma^{-1}(\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\})}\big)\subseteq\overline{\sigma(\sigma^{-1}(\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\}))}=\overline{\tilde H_\mathbf{a}-\{\mathbf{q}_1,\mathbf{q}_2\}}=\tilde H_\mathbf{a}.$$

Since σ is projective, it is closed [Liu06, p. 108]. Hence, $\sigma(H'_\mathbf{a})$ is a Zariski-closed subset of $\tilde H_\mathbf{a}$. Since $\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\}\subseteq\sigma(H'_\mathbf{a})$, we get that $\sigma(H'_\mathbf{a})=\tilde H_\mathbf{a}$. Further, since by (3.10b) σ maps $\sigma^{-1}(\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\})$ isomorphically onto $\tilde H_\mathbf{a}\smallsetminus\{\mathbf{q}_1,\mathbf{q}_2\}$, σ maps $H'_\mathbf{a}$ birationally onto $\tilde H_\mathbf{a}$.

If $\mathbf{a}\neq(0{:}1),(1{:}0)$, then, by (3.3), $\tilde H_\mathbf{a}$ is smooth, hence normal. By Lemma 2.1, σ maps $H'_\mathbf{a}$ isomorphically onto $\tilde H_\mathbf{a}$.

If $\mathbf{a}=(1{:}0)$, then by (3.5), $\tilde H_\mathbf{a}$ is defined in $\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K}$ by the equation $X_1Y_1=0$. Thus, $\tilde H_\mathbf{a}=L_1\cup L_2$, where L_1 is the line defined by $X_1=0$ and L_2 is the line defined by $Y_1=0$. For $i=1,2$ consider the closed set $L'_i=\sigma^{-1}(L_i)\cap H'_\mathbf{a}$. Since L_i is smooth, we have as in the previous case that σ maps L'_i isomorphically onto L_i. Next note that the point $L_1\cap L_2=((1{:}0),(1{:}0))$ (see (3.5)) is different from $\mathbf{q}_1$ and from $\mathbf{q}_2$. Since σ is an isomorphism beyond $\sigma^{-1}(\{\mathbf{q}_1,\mathbf{q}_2\})$, σ maps $H'_\mathbf{a}=L'_1\cup L'_2$ isomorphically onto $\tilde H_\mathbf{a}=L_1\cup L_2$.

The case where $\mathbf{a}=(0{:}1)$ is symmetric to the case where $\mathbf{a}=(1{:}0)$. □

(3.11) $S(\tilde K)=\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)}H'_\mathbf{a}(\tilde K)$.

Proof of (3.11). By the second part of (3.7), the lines $\sigma^{-1}(\mathbf{q}_1)$ and $\sigma^{-1}(\mathbf{q}_2)$ are contained in the right hand side of (3.11). Taking the inverse images of (3.9) under σ and using (3.10d), we have that $S(\tilde{K}) = \bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde{K})} H'_{\mathbf{a}}(\tilde{K})$. It remains to prove that the union is disjoint. To this end let $i \in \{1,2\}$ and note that since $\mathbf{q}_i$ is a simple point of $\mathbb{P}^1(K)\times\mathbb{P}^1(K)$, $\sigma^{-1}(\mathbf{q}_i)$ is isomorphic to the projectivised tangent cone of $\mathbb{P}^1_K \times \mathbb{P}^1_K$ at that point [Mum88, p. 225 (V.)]. Let $i \in \{1,2\}$ and let $\mathbf{a}, \mathbf{a}'$ be distinct points of $\mathbb{P}^1(\tilde{K})$. Then, $H'_{\mathbf{a}}(\tilde{K})\cap\sigma^{-1}(\mathbf{q}_i)(\tilde{K})$ and $H'_{\mathbf{a}'}(\tilde{K})\cap\sigma^{-1}(\mathbf{q}_i)(\tilde{K})$ are points of $S(\tilde{K})$ that correspond under that isomorphism to the tangents of $\tilde{H}_{\mathbf{a}}$ and $\tilde{H}_{\mathbf{a}'}$, respectively, at $\mathbf{q}_i$. By (3.7), those tangents are distinct. Hence, $H'_{\mathbf{a}}(\tilde{K}) \cap H'_{\mathbf{a}'}(\tilde{K}) \cap \sigma^{-1}(\{\mathbf{q}_1, \mathbf{q}_2\})(\tilde{K}) = \emptyset$. Since, by (3.8), $\tilde{H}_{\mathbf{a}}(\tilde{K}) \cap \tilde{H}_{\mathbf{a}'}(\tilde{K}) = \{\mathbf{q}_1, \mathbf{q}_2\}$, it follows from (3.10b) that $H'_{\mathbf{a}}(\tilde{K}) \cap H'_{\mathbf{a}'}(\tilde{K}) = \emptyset$, as claimed. □

(3.12) For each $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$, the projection π of $\tilde{S}$ on the second factor $\mathbb{P}^1_{\tilde{K}}$ of $(\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}) \times \mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ is an epimorphism that maps each $H'_{\mathbf{a}}$ onto $\mathbf{a}$.

Proof of (3.12). The affine (Y_0, X_1)-plane $A = U^{01}$ is a Zariski-open neighborhood of $\mathbf{q}_1$ in $\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$. By (3.2), the intersection $\tilde{H}^{01}_{\mathbf{a}} = \tilde{H}_{\mathbf{a}} \cap A$ is defined by $a_1Y_0 = a_0X_1$. The blow up of A at $\mathbf{q}_1$ is the subset A' of $A \times \mathbb{P}^1_{\tilde{K}}$ defined by the equation $Z_1Y_0 = Z_0X_1$, where $(Z_0{:}Z_1)$ are the homogeneous coordinates of $\mathbb{P}^1_{\tilde{K}}$. Let $\pi_1\colon A\times \mathbb{P}^1_{\tilde{K}} \to \mathbb{P}^1_{\tilde{K}}$ be the projection on the second factor. Then, $\pi_1^{-1}(\mathbf{a}) \cap A' = \tilde{H}^{01}_{\mathbf{a}}$. Since the blow up of $\mathbb{P}^1_K \times \mathbb{P}^1_K$ is done in two stages, first in $\mathbf{q}_1$ and then in the inverse image of $\mathbf{q}_2$ (which we identify with $\mathbf{q}_2$) and since $\mathbf{q}_2 \notin A(\tilde{K})$, we have $\pi(\sigma^{-1}(\tilde{H}^{01}_{\mathbf{a}})) = \mathbf{a}$. By the Zariski-continuity of π, we have $\pi(H'_{\mathbf{a}}) = \mathbf{a}$, as claimed. □

Since, by (3.11), the $H'_{\mathbf{a}}$ are disjoint,

(3.13) $\pi^{-1}(\mathbf{a}) = H'_{\mathbf{a}}$.

3 Irreducible Curves

Let $\tilde{K}$ be a fixed algebraic closure of a field K, $\tilde{D}$ a smooth projective irreducible curve over $\tilde{K}$, and $\tilde{\varphi}\colon \tilde{D} \to \mathbb{P}^1_{\tilde{K}}$ a surjective morphism. Given a morphism $\alpha\colon X \to Y$ of schemes and a point $y \in Y$, we say that α is **smooth over** y if α is smooth at each $x \in X$ with $\alpha(x) = y$. With this terminology, we assume

(4.1) $\tilde{\varphi}$ is smooth over (0:1).

As in Section 2 we consider two copies of $\mathbb{P}^1_{\tilde{K}}$ with respective homogeneous coordinates $(X_0{:}X_1)$ and $(Y_0{:}Y_1)$. For each $\mathbf{a} = (a_0{:}a_1) \in \mathbb{P}^1(\tilde{K})$ we consider the conic $\tilde{H}_{\mathbf{a}}$ defined in $\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ by the equation $a_1X_0Y_0 = a_0X_1Y_1$. Let $\tilde{\delta}\colon \tilde{D}\times\tilde{D} \to \mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ be the product $\tilde{\varphi}\times\tilde{\varphi}$ and consider the inverse image $\tilde{I}_{\mathbf{a}} = \tilde{\delta}^{-1}(\tilde{H}_{\mathbf{a}})$. By [AlK70, p. 129, Prop. 1.7(d)], (4.2a) $\tilde{\delta}\colon \tilde{D} \times \tilde{D} \to \mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ is smooth over $((0{:}1),(0{:}1))$.

Since $\tilde{D}$ is a smooth projective curve and $\tilde{\varphi}\colon \tilde{D} \to \mathbb{P}^1_{\tilde{K}}$ is a surjective morphism, $\tilde{\varphi}$ is finite (see [Sha77, p. 122, Thm. 11] or [Har77, p. 137, Prop. 6.8]). Since the property of being finite is stable under composition and base change [GoW, p. 325, Prop. 12.11(3)],

(4.2b) $\tilde{\delta}$ is a finite morphism, hence proper [GoW10, p. 325, Prop. 12.12].

Since $\dim(\tilde{H}_{\mathbf{a}}) = 1$, it follows from (4.2b) that $\dim(\tilde{I}_{\mathbf{a}}) = 1$. By (3.9) of Section 2, $\tilde{D}(\tilde{K}) \times \tilde{D}(\tilde{K}) = \bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde{K})} \tilde{I}_{\mathbf{a}}(\tilde{K})$.

Lemma 3.1. *For all* $\mathbf{a} \in \mathbb{P}^1(\tilde{K}) \smallsetminus \{(0{:}1),(1{:}0)\}$, $\tilde{I}_{\mathbf{a}}$ *is a connected scheme.*

Proof. First we note that

(4.3) each of the conics $\tilde{H}_{\mathbf{a}}$, with $\mathbf{a} = (a_0{:}a_1) \in \mathbb{P}^1(\tilde{K}) \smallsetminus \{(0{:}1),(1{:}0)\}$, considered as an irreducible divisor of $\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$, is very ample.

To this end we consider the Segre embedding $s\colon \mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}} \to \mathbb{P}^3_{\tilde{K}}$ given by $s((x_0{:}x_1),(y_0{:}y_1)) = (z_0{:}z_1{:}z_2{:}z_3)$, where $z_0 = x_0y_0$, $z_1 = x_0y_1$, $z_2 = x_1y_0$, and $z_3 = x_1y_1$. Then s is a closed immersion onto a closed subsurface P of $\mathbb{P}^3_{\tilde{K}}$ [GoW10, p. 112, Prop. 4.39]. Hence, s induces an isomorphism of $\mathcal{O}_P(1)$ represented by the divisor $P_{\mathbf{a}}$ of $\mathbb{P}^3_{\tilde{K}}$ defined by the linear equation $a_1Z_0 = a_0Z_3$ onto the invertible sheaf $\mathcal{L}_{\mathbf{a}}$ of $\mathbb{P}^1_{\tilde{K}} \times \mathbb{P}^1_{\tilde{K}}$ corresponding to $\tilde{H}_{\mathbf{a}}$. By definition, $\tilde{H}_{\mathbf{a}}$ is very ample [Har77, p. 120, Def. and p. 307].

By (4.3) and by definition [Har77, p. 307], $\tilde{H}_{\mathbf{a}}$ is an effective ample divisor. By (4.2b) and [Har77, p. 232, Exer. 5.7(d)], $\tilde{I}_{\mathbf{a}} = \tilde{\delta}^{-1}(\tilde{H}_{\mathbf{a}})$ is an effective ample divisor of $\tilde{D} \times \tilde{D}$. By assumption, $\tilde{D} \times \tilde{D}$ is an integral smooth (hence normal) projective variety. It follows from a Lemma of Enriques-Severi-Zariski [Har77, p. 244, Cor. 7.9] that $\tilde{I}_{\mathbf{a}}$ is connected. □

Remark 3.2 (Singular points)**.** Let $\kappa\colon X \to Y$ be a morphism of finite type between schemes of finite type over the algebraically closed field $\tilde{K}$. We set

$$\mathrm{Sing}(\kappa) = \{x \in X \mid \kappa \text{ is not smooth at } x\}.$$

By [Gro64, 6.8.7] or [Liu06, p. 224, Cor. 2.12], $\mathrm{Sing}(\kappa)$ is a closed subset of X. If $\lambda\colon Y \to Z$ is another morphism of finite type between schemes of finite type over $\tilde{K}$, then

$$\mathrm{Sing}(\lambda \circ \kappa) \subseteq \mathrm{Sing}(\kappa) \cup \kappa^{-1}(\mathrm{Sing}(\lambda)), \tag{3.4}$$

because composition of smooth morphisms is smooth [Liu06, p. 143, Prop. 3.38]. If $Y = \mathrm{Spec}(\tilde{K})$, then $\mathrm{Sing}(X) = \mathrm{Sing}(\kappa)$ is the set of singular points of X. Applying the equivalent definition of smoothness given by [Liu06, p. 142, Definition 3.35], we find that

$$\mathrm{Sing}(\kappa) = \bigcup_{y\in Y} \mathrm{Sing}(\kappa^{-1}(y)) \tag{3.5}$$

if κ is flat.

Lemma 3.3. *For almost all* $\mathbf{a}\in\mathbb{P}^1(\tilde K)$, $\tilde I_{\mathbf{a}}$ *is a smooth scheme.*

Proof. As in Section 2, we consider the points $\mathbf{q}_1 = ((1{:}0),(0{:}1))$ and $\mathbf{q}_2 = ((0{:}1),(1{:}0))$ of $\mathbb{P}^1(\tilde K)\times\mathbb{P}^1(\tilde K)$. As in Subsection 2.2, let S be the closed $\tilde K$-subsurface of $(\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K})\times\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K}$ obtained by blowing up $\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K}$ at the set $\{\mathbf{q}_1,\mathbf{q}_2\}$, let $\sigma\colon S\to(\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K})$ be the projection on the first factor, and $\pi\colon S\to\mathbb{P}^1_{\tilde K}$ the projection on the second factor. Both morphisms are projective, hence proper [Liu06, p. 108, Thm. 3.30]. Now we break up the rest of the proof into several parts.

Part A: *A commutative diagram.* Let $T = (\tilde D\times\tilde D)\times_{(\mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K})} S$ be the fibred product of $\tilde\delta$ and σ. Let $\tau\colon T\to S$ be the projection on the second factor and let $\pi_T = \pi\circ\tau$. Since $\tilde\delta$ is proper (by (4.2b)), so is τ [Liu06, p. 104, Cor. 3.16(c)]. Since also π is proper,

(4.6) $\pi_T = \pi\circ\tau$ is also proper [Liu06, p. 104, Cor. 3.16(b)].

By (3.11) of Section 2, $S(\tilde K) = \dot\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} H'_{\mathbf{a}}(\tilde K)$, where $H'_{\mathbf{a}}$ is a curve on S that σ maps isomorphically onto $\tilde H_{\mathbf{a}}$. For each $\mathbf{a}\in\mathbb{P}^1(\tilde K)$ let $I'_{\mathbf{a}} = \tau^{-1}(H'_{\mathbf{a}}) = \tilde I_{\mathbf{a}}\times_{\tilde H_{\mathbf{a}}} H'_{\mathbf{a}}$. Then, $T(\tilde K) = \dot\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} I'_{\mathbf{a}}(\tilde K)$. Moreover, by (3.13) of Section 2, $\pi^{-1}(\mathbf{a}) = H'_{\mathbf{a}}$, so with $\pi_T = \pi\circ\tau$, we have

(4.7) $\pi_T^{-1}(\mathbf{a}) = \tau^{-1}(\pi^{-1}(\mathbf{a})) = \tau^{-1}(H'_{\mathbf{a}}) = I'_{\mathbf{a}}$ for each $\mathbf{a}\in\mathbb{P}^1(\tilde K)$.

This gives a commutative diagram

$$\begin{tikzcd}
\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} \tilde I_{\mathbf{a}} = \tilde D\times\tilde D \arrow[d, "\tilde\delta"] & T = \dot\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} I'_{\mathbf{a}} \arrow[l, "\sigma_T"'] \arrow[r, "\pi_T"] \arrow[d, "\tau"] & \mathbb{P}^1_{\tilde K} \arrow[d, equal] \\
\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} \tilde H_{\mathbf{a}} = \mathbb{P}^1_{\tilde K}\times\mathbb{P}^1_{\tilde K} & S = \dot\bigcup_{\mathbf{a}\in\mathbb{P}^1(\tilde K)} H'_{\mathbf{a}} \arrow[l, "\sigma"'] \arrow[r, "\pi"] & \mathbb{P}^1_{\tilde K}\ ,
\end{tikzcd} \tag{3.8}$$

where the left square is cartesian. By (4.2b), $\tilde\delta$ is finite. Since finiteness of morphisms is preserved under base change [GoW10, p. 325, Prop. 12.11(2)],

(4.9) τ is a finite morphism.

Since $\dim(\mathbb{P}^1_{\tilde{K}}) = 1$ and $\mathbb{P}^1_{\tilde{K}}$ is smooth, $\mathbb{P}^1_{\tilde{K}}$ is a Dedekind scheme. In addition π is not a constant map. Since σ is birational, S is integral, so by [Liu06, p. 137, Cor. 3.10],

(4.10) π is flat.

Similarly, since σ_T is birational, T is integral, so by [Liu06, p. 137, Cor. 3.10],

(4.11) $\pi_T = \pi \circ \tau$ is flat.

Part B: *Finiteness of* $\tau^{-1}(\mathrm{Sing}(\pi))$. By (4.10) and (4.5), and by (3.13) of Section 2,

$$\mathrm{Sing}(\pi) = \bigcup_{\mathbf{a} \in \mathbb{P}^1(\tilde{K})} \mathrm{Sing}(\pi^{-1}(\mathbf{a})) = \bigcup_{\mathbf{a} \in \mathbb{P}^1(\tilde{K})} \mathrm{Sing}(H'_{\mathbf{a}}). \tag{3.12}$$

Let $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$. By (3.10d) of Section 2, $H'_{\mathbf{a}} \cong \tilde{H}_{\mathbf{a}}$. By (3.3) of Section 2, $\tilde{H}_{\mathbf{a}}$ is smooth if $\mathbf{a} \neq (1{:}0), (0{:}1)$, so $\mathrm{Sing}(H'_{\mathbf{a}})$ is empty. By (3.5) and (3.6) of Section 2, each of the conics $H_{(1:0)}$ and $H_{(0:1)}$ has a unique singular point. It follows from (4.12) that $\mathrm{Sing}(\pi)$ is finite. By (4.9), the set $\tau^{-1}(\mathrm{Sing}(\pi))$ is finite.

Part C: *Finiteness of* $\pi_T(\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau))$. Let $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$. By (3.10d) of Section 2, the morphism σ maps $H'_{\mathbf{a}}$ isomorphically onto $\tilde{H}_{\mathbf{a}}$. Since the diagram

$$\begin{tikzcd}
\tilde{\delta}^{-1}(\tilde{H}_{\mathbf{a}}) = \tilde{I}_{\mathbf{a}} \arrow[d, "\tilde{\delta}_{\mathbf{a}}"'] & I'_{\mathbf{a}} = \tau^{-1}(H'_{\mathbf{a}}) \arrow[l, "\sigma_{T,\mathbf{a}}"'] \arrow[d, "\tau_{\mathbf{a}}"] \\
\tilde{H}_{\mathbf{a}} & H'_{\mathbf{a}} \arrow[l, "\sigma_{\mathbf{a}}"'] ,
\end{tikzcd}$$

where the arrows are the corresponding restrictions of the arrows of the left square of Diagram (4.8), is cartesian,

(4.13) $\sigma_{T,\mathbf{a}}$ maps $I'_{\mathbf{a}}$ isomorphically onto $\tilde{I}_{\mathbf{a}}$.

By (3.6) of Section 2, $\tilde{H}_{(0:1)} = ((0{:}1) \times \mathbb{P}^1_{\tilde{K}}) \cup (\mathbb{P}^1_{\tilde{K}} \times (0{:}1))$. Therefore, by (4.13), $I'_{(0:1)}$ is isomorphic to $I_{(0:1)} = (\tilde{\varphi}^{-1}((0{:}1)) \times \tilde{D}) \cup (\tilde{D} \times \tilde{\varphi}^{-1}((0{:}1))$. Since $\tilde{D}$ is smooth, the latter scheme is smooth except for nodes lying over the intersection point of the two components of $\tilde{H}_{(0:1)}$, namely over $((0{:}1), (0{:}1))$. Therefore, using Convention (3.10b) of Section 2,

$$\tau(\mathrm{Sing}(I'_{(0:1)})) = \{((0{:}1), (0{:}1))\}. \tag{3.14}$$

By (3.13) of Section 2, $\pi^{-1}((0{:}1)) = H'_{(0:1)}$. Hence,

$$\pi_T^{-1}((0{:}1)) = \tau^{-1}(\pi^{-1}((0{:}1))) = \tau^{-1}(H'_{(0:1)}) = I'_{(0:1)}. \tag{3.15}$$

Therefore, by (4.11) and (4.5),

$$\mathrm{Sing}(I'_{(0:1)}) = \mathrm{Sing}(\pi_T^{-1}((0:1))) = \mathrm{Sing}(\pi_T) \cap I'_{(0:1)}.$$

It follows by (4.14) that

$$(3.16) \qquad \tau(\mathrm{Sing}(\pi_T) \cap I'_{(0:1)}) = \tau(\mathrm{Sing}(I'_{(0:1)})) = \{((0:1),(0:1))\}.$$

On the other hand, by (4.2a), $\tilde{\delta}$ is smooth over $((0:1),(0:1))$, so by [Liu06, p. 143, Prop. 3.38], τ is also smooth over the point $((0:1),(0:1))$ of S (Convention (3.10b) of Section 2). In other words,

$$(3.17) \qquad \tau^{-1}(((0:1),(0:1))) \cap \mathrm{Sing}(\tau) = \emptyset.$$

It follows from (4.16) and (4.17) that

$$(3.18) \qquad \mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau) \cap I'_{(0:1)} = \emptyset.$$

By (4.6), π_T is proper. By Remark 3.2, $\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau)$ is closed in T. Hence, the set $\pi_T(\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau))$ is closed in $\mathbb{P}^1_{\tilde{K}}$. Therefore, $\pi_T(\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau))$ is either $\mathbb{P}^1_{\tilde{K}}$ or a finite set. In the former case, each point in $I'_{(0:1)} = \pi_T^{-1}((0:1))$ (see (4.15)) lies in $\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau)$, which contradicts (4.18). Therefore,

(4.19) the set $\pi_T(\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau))$ is finite.

PART D: *Finiteness of* $\pi_T(\mathrm{Sing}(\pi_T))$. Now note by (4.4) that $\mathrm{Sing}(\pi_T) = \mathrm{Sing}(\pi \circ \tau) \subseteq \mathrm{Sing}(\tau) \cup \tau^{-1}(\mathrm{Sing}(\pi))$. Hence,

$$\pi_T(\mathrm{Sing}(\pi_T)) \subseteq \pi_T(\mathrm{Sing}(\pi_T) \cap \mathrm{Sing}(\tau)) \cup \pi_T(\tau^{-1}(\mathrm{Sing}(\pi))).$$

It follows from (4.19) and the finiteness of $\tau^{-1}(\mathrm{Sing}(\pi))$ (Part B) that $\pi_T(\mathrm{Sing}(\pi_T))$ is finite.

PART E: *End of proof.* We consider $\mathbf{a} \in \mathbb{P}^1(\tilde{K}) \smallsetminus \pi_T(\mathrm{Sing}(\pi_T))$. By (4.7), $\pi_T^{-1}(\mathbf{a}) = I'_{\mathbf{a}}$. Hence, $I'_{\mathbf{a}}$ is smooth. By (4.13), $I'_{\mathbf{a}}$ is isomorphic to $\tilde{I}_{\mathbf{a}}$. Hence, $\tilde{I}_{\mathbf{a}}$ is smooth. It follows from Part D that $\tilde{I}_{\mathbf{a}}$ is smooth for almost all $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$, as claimed. □

Since normal (in particular, smooth) connected $\tilde{K}$-schemes of finite type are irreducible (e.g. [GoW10, p. 168, Exer. 6.20]), a combination of Lemma 3.1 and Lemma 3.3 yields the following result:

Corollary 3.4. *For almost all* $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$, $\tilde{I}_{\mathbf{a}}$ *is an irreducible smooth curve.*

4 The Open Mapping Theorem

As in Section 2, we fix an algebraic closure $\tilde{K}$ of a field K. We also fix a valuation v of $\tilde{K}$ and prove an open map theorem for varieties over $\tilde{K}$ in the v-topology.

Our proof is based on a theorem about the continuity of roots for not necessarily separable polynomials. A convenient reference is [Jar91, Prop. 12.2].

Lemma 4.1. *Let $p(X) = \prod_{i=1}^n (X - a_i)$ be a monic polynomial with coefficients in $\tilde{K}$. Then, for each $c \in \tilde{K}^\times$ there exists $c' \in \tilde{K}^\times$ such that if $q \in \tilde{K}[X]$ is a monic polynomial of degree n and $v(q-p) > v(c')$, then $q(X)$ may be presented as a product $q(X) = \prod_{i=1}^n (X - b_i)$ such that $v(b_i - a_i) > v(c)$ for $i = 1, \ldots, n$.*

Here and throughout we write $v\big(\sum_{i=0}^n c_i X^i\big) > v(b)$ for $c_0, \ldots, c_n \in \tilde{K}$ and $b \in \tilde{K}^\times$ as an abbreviation for "$v(c_i) > v(b)$ for $i = 0, \ldots, n$."

Lemma 4.2. *Let V be a vector space of finite dimension d over an infinite field K_0. Then V has an infinite subset V_0 such that every subset of V_0 of cardinality d is a basis of V.*

Proof. By assumption, V has a subset of cardinality d which is a basis of V. Inductively suppose that U is a finite subset of V with $e \geq d$ elements such that every subset of U of cardinality d is a basis of V. We denote the collection of all subsets of U of cardinality $d-1$ by $\mathcal{U}$. By assumption, for each $U_0 \in \mathcal{U}$ the dimension of the subspace $\sum_{u \in U_0} K_0 u$ of V is $d-1$, so that subspace is properly contained in V. It follows that also $\bigcup_{U_0 \in \mathcal{U}} \sum_{u \in U_0} K_0 u$ is a proper subset of V. We choose an element $v \in V \smallsetminus \bigcup_{U_0 \in \mathcal{U}} \sum_{u \in U_0} K_0 u$. Then, for each $U_0 \in \mathcal{U}$ we have $\dim(K_0 v + \sum_{u \in U_0} K_0 u) = d$, so $\{v\} \cup U_0$ is a basis of V. Consequently, every subset of $U \cup \{v\}$ of cardinality d is a basis of V. This completes the induction and the proof of the lemma. □

Remark 4.3 (The v-topology on $\mathrm{Max}(A)$)**.** Let B be a finitely generated integral domain over $\tilde{K}$. We denote the set of all maximal ideals of B by $\mathrm{Max}(B)$. For each $\mathfrak{q} \in \mathrm{Max}(B)$ we identify $B/\mathfrak{q}$ with $\tilde{K}$ and let $x(\mathfrak{q})$ be the residue of $x \in B$ at $\mathfrak{q}$. The v-topology of $\tilde{K}$ induces a **v-topology** on $\mathrm{Max}(B)$. A basic open neighborhood of a point $\mathfrak{q}_0 \in \mathrm{Max}(B)$ is a set

$$\mathcal{V} = \bigcap_{i=1}^{r} \{\mathfrak{q} \in \mathrm{Max}(B) \mid v(y_i(\mathfrak{q}) - y_i(\mathfrak{q}_0)) > v(c_0)\}, \tag{4.1}$$

where $y_1, \ldots, y_r \in B$ and $c_0 \in \tilde{K}^\times$. Thus, each $x \in B$ may be viewed as a **v-continuous** (i.e. continuous in the v-topology) map from $\mathrm{Max}(B)$ to $\tilde{K}$.

Let A be an integral domain that contains $\tilde{K}$ with quotient field E. Suppose that $F = \mathrm{Quot}(B)$ is an extension of E and that B is an integral extension of A. If $\mathfrak{q} \in \mathrm{Max}(B)$, then $\mathfrak{p} = \mathfrak{q} \cap A \in \mathrm{Max}(A)$ and we identify $x(\mathfrak{p})$ with $x(\mathfrak{q})$ for each $x \in A$. Then, the canonical morphism $\varphi\colon \mathrm{Max}(B) \to \mathrm{Max}(A)$ defined by $\varphi(\mathfrak{q}) = \mathfrak{q} \cap A$ is v-continuous.

Indeed, consider a basic open neighborhood

$$\mathcal{U} = \bigcap_{i=1}^{m} \{\mathfrak{p} \in \mathrm{Max}(A) \mid \ v(x_i(\mathfrak{p}) - x_i(\mathfrak{p}_0)) > v(a_0)\}$$

of a point $\mathfrak{p}_0 \in \mathrm{Max}(A)$ with $x_1, \dots, x_m \in A$ and $a_0 \in \tilde{K}^\times$. By the going up theorem, $\varphi(\mathrm{Max}(B)) = \mathrm{Max}(A)$ [Lan93, p. 339, Prop. 1.10]. Hence, with $\mathfrak{q}_0 \in \varphi^{-1}(\mathfrak{p}_0)$,

$$\varphi^{-1}(\mathcal{U}) = \bigcap_{i=1}^{m} \{\mathfrak{q} \in \mathrm{Max}(B) \mid \ v(x_i(\mathfrak{q}) - x_i(\mathfrak{q}_0)) > v(a_0)\},$$

is v-open in $\mathrm{Max}(B)$.

We use Lemma 4.1 to prove that the map φ of Remark 4.3 is v-open if F/E is finite and separable and A is integrally closed.

Lemma 4.4. *Let A be an integrally closed domain which contains $\tilde{K}$, $E =$ Quot(A), F a finite separable extension of E, and B a subdomain of F which contains A and is integral over A. Then the canonical morphism φ:* $\mathrm{Max}(B) \to \mathrm{Max}(A)$ *defined by $\varphi(\mathfrak{q}) = \mathfrak{q} \cap A$ is a v-open map.*

Proof. Let $\hat{F}$ be the Galois closure of F/E, $\hat{B}$ the integral closure of B in $\hat{F}$, and $\psi\colon \mathrm{Max}(\hat{B}) \to \mathrm{Max}(B)$ the canonical map. Then, $\hat{B}$ is also the integral closure of A in $\hat{F}$ and $\hat{\varphi} = \varphi \circ \psi$ is the canonical map of $\mathrm{Max}(\hat{B})$ to $\mathrm{Max}(A)$. Let $\mathcal{V}$ be a v-open subset of $\mathrm{Max}(B)$. By Remark 4.3, ψ is v-continuous. Hence, $\hat{\mathcal{V}} = \psi^{-1}(\mathcal{V})$ is a v-open subset of $\mathrm{Max}(\hat{B})$. If $\hat{\varphi}(\hat{\mathcal{V}})$ is v-open in $\mathrm{Max}(A)$, then so is $\varphi(\mathcal{V}) = \hat{\varphi}(\psi^{-1}(\mathcal{V}))$. Therefore, replacing F by $\hat{F}$, we may assume that F/E is Galois of degree n with $G = \mathrm{Gal}(F/E)$ and B is the integral closure of A in F.

Consider a point $\mathfrak{q} \in \mathrm{Max}(B)$ and let $\mathfrak{p} = \varphi(\mathfrak{q})$. It suffices to prove that for every basic v-open neighborhood $\mathcal{V} = \bigcap_{i=1}^{r} \{\mathfrak{q}' \in \mathrm{Max}(B) \mid \ v(y_i(\mathfrak{q}') - y_i(\mathfrak{q})) > v(c)\}$ of $\mathfrak{q}$ in $\mathrm{Max}(B)$ with $y_1, \dots, y_r \in B$, and $c \in \tilde{K}^\times$, the point $\mathfrak{p}$ of $\mathrm{Max}(A)$ has a v-open neighborhood in $\varphi(\mathcal{V})$. We break the proof of this statement into several parts.

Part A: *Many bases of a vector space.* We consider the vector space $V = \sum_{i=1}^{r} \tilde{K} y_i$ spanned by $y_1, \dots, y_r$ over $\tilde{K}$ and let $d = \dim(V)$. By Lemma 4.2, there exists an infinite subset Z'' of V such that every subset of Z'' of cardinality d is a basis of V. We choose a finite subset Z' of Z'' of cardinality greater than $(d-1)n$. Since $\tilde{K} \subseteq A$, the vector space V is contained in B, hence $Z' \subseteq B$. In particular, every $z \in Z'$ is integral over A.

Let $\mathcal{Z}$ be the collection of all subsets of Z' of cardinality d. For every $1 \le i \le r$ and $Z \in \mathcal{Z}$ there exists a presentation

$$y_i = \sum_{z \in Z} a_{i,Z,z} z \text{ with } a_{i,Z,z} \in \tilde{K}. \tag{4.2}$$

We set

$$\alpha = \min(v(a_{i,Z,z})_{i=1,\ldots,r\,,Z\in\mathcal{Z},\,z\in Z}). \tag{4.3}$$

Part B: *Continuity of roots.* Since A is integrally closed, Part A implies that for each $z \in Z'$

(5.4) $f_z(X) = \prod_{\sigma\in G}(X - \sigma z)$ is a monic polynomial of degree n with coefficients in A.

Then,

$$f_z(\mathfrak{p})(X) = f_z(\mathfrak{q})(X) = \prod_{\sigma\in G}(X - (\sigma z)(\mathfrak{q})) = \prod_{\sigma\in G}(X - z(\sigma^{-1}\mathfrak{q})). \tag{4.5}$$

Let $\mathfrak{p}' \in \mathrm{Max}(A)$ and choose $\mathfrak{q}' \in \mathrm{Max}(B)$ over $\mathfrak{p}'$. As in (5.5), $f_z(\mathfrak{p}')(X) = \prod_{\sigma\in G}(X - z(\sigma^{-1}\mathfrak{q}'))$. By Lemma 4.1, there exists $c' \in \tilde{K}^\times$ and there exists $\sigma_z \in G$ such that

(5.6) if $v(f_z(\mathfrak{p}') - f_z(\mathfrak{p})) > v(c')$, then $v(z(\sigma_z^{-1}\mathfrak{q}') - z(\mathfrak{q})) > v(c) - \alpha$.

Let $s\colon Z' \to G$ be the map defined by $s(z) = \sigma_z$. Then $(d-1)n < |Z'| = \sum_{\sigma\in G}|s^{-1}(\sigma)|$. Hence, there exists $\sigma \in G$ such that $|s^{-1}(\sigma)| \geq d$. Choose a subset Z of $s^{-1}(\sigma)$ of cardinality d, in particular $Z \in \mathcal{Z}$. By (5.6), if $v(f_z(\mathfrak{p}') - f_z(\mathfrak{p})) > v(c')$ for each $z \in Z$, then

$$v(z(\sigma^{-1}\mathfrak{q}') - z(\mathfrak{q})) > v(c) - \alpha \text{ for all } z \in Z. \tag{4.7}$$

Part C: *Conclusion of the proof.* We prove that the open neighborhood

$$\mathcal{U} = \{\mathfrak{p}' \in \mathrm{Max}(A) \mid \bigwedge_{z\in Z} v(f_z(\mathfrak{p}') - f_z(\mathfrak{p})) > v(c')\}$$

of $\mathfrak{p}$ in $\mathrm{Max}(A)$ is contained in $\varphi(\mathcal{V})$.

Indeed, let $\mathfrak{p}' \in \mathcal{U}$ and choose $\mathfrak{q}' \in \mathrm{Max}(B)$ with $\varphi(\mathfrak{q}') = \mathfrak{p}'$. Then, $v(f_z(\mathfrak{p}') - f_z(\mathfrak{p})) > v(c')$ for each $z \in Z$, so (5.7) holds. We set $\mathfrak{q}'' = \sigma^{-1}\mathfrak{q}'$ and notice that $\varphi(\mathfrak{q}'') = \mathfrak{p}'$. By (5.7), $v(z(\mathfrak{q}'') - z(\mathfrak{q})) > v(c) - \alpha$ for each $z \in Z$. In addition, by (5.3), $v(a_{i,Z,z}) \geq \alpha$ for $i = 1, \ldots, r$ and for each $z \in Z$. It follows from (5.2) that

$$\begin{aligned} v(y_i(\mathfrak{q}'') - y_i(\mathfrak{q})) = v\big(\sum_{z\in Z} a_{i,Z,z}(z(\mathfrak{q}'') - z(\mathfrak{q}))\big) \\ \geq \min_{z\in Z}\big(v(a_{i,Z,z}) + v(z(\mathfrak{q}'') - z(\mathfrak{q}))\big) > v(c) \end{aligned}$$

for $i = 1, \ldots, r$. Consequently, $\mathfrak{q}'' \in \mathcal{V}$, as claimed. □

Proposition 4.5. *Let $(\tilde{K}, v)$ be an algebraically closed valued field and $\varphi\colon W \to V$ a finite morphism of absolutely irreducible varieties defined over $\tilde{K}$ with V normal. Then, the v-continuous map $\varphi\colon W(\tilde{K}) \to V(\tilde{K})$ is v-open.*

Proof. The morphism φ decomposes into a purely inseparable finite morphism followed by a separable finite morphism. Since inseparable finite morphisms induce v-homeomorphisms on the corresponding sets of K-rational points, we may assume that φ is separable.

By definition, V has a cover consisting of affine Zariski-open subsets V_i whose inverse images W_i under φ are also affine, such that for each i, $\Gamma(V_i, \mathcal{O}_V)$ is an integrally closed domain, $\Gamma(W_i, \mathcal{O}_W)$ is an integral domain which is finitely generated as a module over $\Gamma(V_i, \mathcal{O}_V)$. It follows that $\Gamma(W_i, \mathcal{O}_W)$ is integral over $\Gamma(V_i, \mathcal{O}_V)$. Since every Zariski-open subset of a variety is also v-open, our proposition is a consequence of Lemma 4.4. □

Remark 4.6. Proposition 4.5 is related to [GPR95, Thm. 9.4(1)]. The latter result says that if (K, v) is an arbitrary Henselian field and $\varphi\colon W \to V$ is a smooth surjective morphism of absolutely irreducible varieties V and W defined over K, then the map $\varphi\colon W(K) \to V(K)$ is v-open.

5 A Density Property of Smooth Curves over PAC Fields

The aim of this short section is to prove a density result for curves over PAC fields.

We start with an arbitrary field K. As in Section 2 we consider for each $\mathbf{a} = (a_0{:}a_1) \in \mathbb{P}^1(K)$ the conic $H_\mathbf{a}$ defined in $\mathbb{P}^1_K \times \mathbb{P}^1_K$ by the equation $a_1X_0Y_0 = a_0X_1Y_1$. We consider a smooth projective absolutely irreducible curve D defined over K. Using the Segre embedding [GoW10, p. 112, Prop. 4.39], we may consider $H_\mathbf{a}$ also as a closed subscheme of $\mathbb{P}^3_K$. Let $\varphi\colon D \to \mathbb{P}^1_K$ be a non-constant separable morphism. We assume that

(6.1) φ is smooth over $(0{:}1)$.

Let $\delta\colon D \times D \to \mathbb{P}^1_K \times \mathbb{P}^1_K$ be the product $\varphi \times \varphi$ and consider the inverse image $I_\mathbf{a} = \delta^{-1}(H_\mathbf{a})$.

Recall that we use a tilde to denote the constant extension from K to $\tilde{K}$ of algebro-geometrical objects.

Lemma 5.1. *Under Assumption (6.1), for almost all $\mathbf{a} \in \mathbb{P}^1(K)$, the scheme $I_\mathbf{a}$ is an absolutely irreducible smooth curve defined over K.*

Proof. Since D is absolutely irreducible and defined over K, the curve $\tilde{D}$ is $\tilde{K}$-irreducible, and the morphism $\tilde{\varphi}\colon \tilde{D} \to \mathbb{P}^1_{\tilde{K}}$ is surjective [Har77, p. 137, Prop. 6.8] and smooth over (0:1). Also, $\tilde{I}_{\mathbf{a}} = (\tilde{\delta})^{-1}(\tilde{H}_{\mathbf{a}})$. By Corollary 3.4, for almost all $\mathbf{a} \in \mathbb{P}^1(\tilde{K})$, $\tilde{I}_{\mathbf{a}}$ is irreducible. Moreover, by the latter corollary, $\tilde{I}_{\mathbf{a}}$ is also smooth. Hence, each point $\mathbf{p}$ of $\tilde{I}_{\mathbf{a}}$ is regular. Therefore, $\mathcal{O}_{\tilde{I}_{\mathbf{a}},\mathbf{p}}$ is an integral domain. It follows that $I_{\mathbf{a}}$ is an absolutely irreducible curve defined over K. □

Lemma 5.2. *Let $\psi\colon W \to V$ be a finite morphism of integral schemes over a field K such that V is normal. Consider the inclusion of the function field E of V into the function field F of W that ψ induces and assume that F/E is Galois. Suppose that the natural action of $G = \mathrm{Gal}(F/E)$ on the generic point of W extends to an action on W over V such that $\mathcal{O}_V$ is the fixed subsheaf of the induced action of G on $\mathcal{O}_W$. Then, for every $\mathbf{q}, \mathbf{q}' \in W$ with $\psi(\mathbf{q}) = \psi(\mathbf{q}')$ there exists $\sigma \in G$ such that $\sigma\mathbf{q} = \mathbf{q}'$.*

Proof. Let $\mathbf{p} \in V$ and $\mathbf{q}, \mathbf{q}' \in W$ such that $\psi(\mathbf{q}) = \mathbf{p}$ and $\psi(\mathbf{q}') = \mathbf{p}$. Let $V_0 = \mathrm{Spec}(A)$ be an affine Zariski-open neighborhood of $\mathbf{p}$ in V. Since V is integral and normal, A is an integrally closed domain with $\mathrm{Quot}(A) = E$. Since ψ is finite, $W_0 = \varphi^{-1}(V_0)$ is also affine, say $W_0 = \mathrm{Spec}(B)$, where B is an integral domain which is integral over A. By assumption, $F = \mathrm{Quot}(B)$ is a finite Galois extension of E. Also, G acts on B with A being the fixed ring of B under G. Finally, we may identify $\mathbf{q}$ and $\mathbf{q}'$ with prime ideals of B and $\mathbf{p}$ with the prime ideal of A lying under both $\mathbf{q}$ and $\mathbf{q}'$. By [Bou89, p. 331, Thm. 2(i)], there exists $\sigma \in G$ such that $\sigma\mathbf{q} = \mathbf{q}'$, as claimed. □

The proof of the following lemma uses a trick of Prestel [FrJ08, p. 204, proof of Prop. 11.5.3].

Lemma 5.3. *Let K be a PAC field and let v be a valuation of $\tilde{K}$. Let E be the function field of $\mathbb{P}^1_K$ and let $\hat{F}$ be a finite Galois extension of E which is regular over K. Suppose the morphism $\varphi\colon D \to \mathbb{P}^1_K$ introduced at the beginning of this section (in particular φ satisfies (6.1)) is the normalization of $\mathbb{P}^1_K$ in $\hat{F}$ [Liu06, p. 120, Def. 1.24].*

Let $\mathbf{p}$ be a point in $D(\tilde{K})$ such that $\tilde{\varphi}(\mathbf{p}) =$ (1:0) and let $\mathcal{V}$ be a v-open neighborhood of $\mathbf{p}$ in $D(\tilde{K})$. Then $\mathcal{V} \cap D(K) \neq \emptyset$.

Proof. Let $G = \mathrm{Gal}(\hat{F}/E)$. Since $\hat{F}$ is a regular extension of K, we may identify G with $\mathrm{Gal}(\hat{F}\tilde{K}/E\tilde{K})$. Note that $\hat{F}\tilde{K}/E\tilde{K}$ is the function field extension that corresponds to the morphism $\tilde{\varphi}\colon \tilde{D} \to \mathbb{P}^1_{\tilde{K}}$. Thus, the action of G on D extends to an action of G on $\tilde{D}$. Moreover, $\mathbb{P}^1_{\tilde{K}}$ is normal and $\mathcal{O}_{\mathbb{P}^1_{\tilde{K}}}$ is the fixed subsheaf under the action of G on $\mathcal{O}_{\tilde{D}}$. Also, by [Liu06, p. 121, Prop. 1.25], φ is finite, hence $\tilde{\varphi}$ is finite. It follows from Lemma 5.2 (with $\tilde{K}$ replacing K, $\tilde{\varphi}\colon \tilde{D} \to \mathbb{P}^1_{\tilde{K}}$ replacing $\psi\colon W \to V$, and $\hat{F}\tilde{K}/E\tilde{K}$ replacing F/E) that if $\mathbf{q}, \mathbf{q}' \in D(\tilde{K})$ and $\tilde{\varphi}(\mathbf{q}) = \tilde{\varphi}(\mathbf{q}')$, then there exists $\sigma \in G$ such that $\sigma\mathbf{q} = \mathbf{q}'$. Since σ fixes the elements of K, we have

(6.2) if $\mathbf{q} \in D(K)$, then $\mathbf{q}' = \sigma\mathbf{q} \in \sigma(D(K)) = D(K)$.

By Proposition 4.5, $\mathcal{U} = \tilde{\varphi}(\mathcal{V})$ is a v-open neighborhood of (1:0) in $\mathbb{P}^1(\tilde{K})$. Hence, there exists $a \in K^\times$ such that

(6.3) if $b \in \tilde{K}^\times$ and $v(b) \geq v(a)$, then $(1{:}b) \in \mathcal{U}$.

By definition, $I_{(1:a^2)} \subset D \times D$. Avoiding finitely many elements of $K^\times$, we may use Lemma 5.1 to choose a in K such that the curve $I_{(1:a^2)}$ is absolutely irreducible and is defined over K. Moreover, $H_{(1:a^2)} = \delta(I_{(1:a^2)})$ is defined by the equation $a^2X_0Y_0 = X_1Y_1$. Let U be the nonempty Zariski-open subset of $H_{(1:a^2)}$ defined by the inequalities $X_0 \neq 0$ and $Y_0 \neq 0$. Let $V = \delta^{-1}(U)$. Since K is PAC, there exists $(\mathbf{q}, \mathbf{r}) \in V(K)$. That is, $\mathbf{q}, \mathbf{r} \in D(K)$, $\varphi(\mathbf{q}) = (1{:}b)$, $\varphi(\mathbf{r}) = (1{:}c)$, and $bc = a^2$ for some $b, c \in K^\times$. Thus, $v(b) + v(c) = 2v(a)$. We may assume without loss that $v(b) \geq v(a)$. By (6.3), $\tilde{\varphi}(\mathbf{q}) = (1{:}b) \in \mathcal{U}$. Hence, there exists $\mathbf{q}' \in \mathcal{V}$ with $\tilde{\varphi}(\mathbf{q}') = \tilde{\varphi}(\mathbf{q})$. Since $\mathbf{q} \in D(K)$, it follows from (6.2) that $\mathbf{q}' \in D(K)$. Consequently, $\mathcal{V} \cap D(K) \neq \emptyset$. □

6 On the Density Property of PAC Fields

We prove Kollár's result saying that if V is an absolutely irreducible variety defined over a PAC field K and v is a valuation of $\tilde{K}$, then $V(K)$ is v-dense in $V(\tilde{K})$.

Remark 6.1. Let F/E be a Galois extension of degree n.

Claim: *If w, w' are valuations of F such that $O_w = O_{w'}$ and $w|_E = w'|_E$, then $w = w'$* Indeed, let $x \in F$. Since $(w(F^\times) : w(E^\times))|n$, there exists $a \in E$ such that $nw(x) = w(a)$. Hence, $w(x^na^{-1}) = 0$, so $x^na^{-1} \in O_w^\times$. Therefore, $x^na^{-1} \in O_{w'}^\times$, hence, $w'(x^na^{-1}) = 0$, so $nw'(x) = w'(a)$. Since $w(a) = w'(a)$, we get $nw(x) = nw'(x)$, consequently $w(x) = w'(x)$, as claimed.

In particular, if an element σ of $\mathrm{Gal}(F/E)$ belongs to the decomposition group D_w of w over E, then $O_{w\circ\sigma} = O_w$. In addition $w \circ \sigma|_E = w|_E$. Hence, by the claim, $w = w \circ \sigma$.

Now suppose that $w_1, \ldots, w_m$ are all of the extensions to F (up to equivalence) of a valuation v of E. Then $O_{w_1}, \ldots, O_{w_m}$ are all of the valuation rings of F whose intersections with E are O_v. Then for each $1 \leq i \leq n$ there exists $\sigma_i \in \mathrm{Gal}(F/E)$ such that $O_{w_1} = \sigma_i O_{w_i}$, that is $O_{w_1\circ\sigma_i} = O_{w_i}$. By the preceding paragraph, $w_i = w_1 \circ \sigma_i$.

Lemma 6.2. *Let (E, v) be a valued field, F a finite separable extension of E, and $\hat{F}$ the Galois closure of F/E. Suppose v totally splits in F. Then v totally splits in $\hat{F}$.*

Proof. Let w be a valuation of $\hat{F}$ lying over v. It suffices to prove that the decomposition group $D_{w/v}$ of w over v is trivial. Consider $\sigma \in D_{w/v}$.

Claim: $\sigma \in \mathrm{Gal}(\hat{F}/F)$. Let $d = [F : E]$. By assumption F has d distinct valuations $v_1, \ldots, v_d$ extending v. For each $1 \leq i \leq d$ we extend v_i to a valuation

w_i of $\hat{F}$ such that $w_1 = w$. By Remark 6.1, there exists $\sigma_i \in \mathrm{Gal}(\hat{F}/E)$ such that $w_i = w \circ \sigma_i$ and $\sigma_1 = 1$. If some $1 \le j \le d$ satisfies $\sigma_i \mathrm{Gal}(\hat{F}/F) = \sigma_j \mathrm{Gal}(\hat{F}/F)$, then for each $x \in F$ we have $v_i(x) = w_i(x) = w(\sigma_i x) = w(\sigma_j x) = w_j(x) = v_j(x)$, so $v_i = v_j$, hence $i = j$. Thus, $\sigma_1 \mathrm{Gal}(\hat{F}/F), \ldots, \sigma_d \mathrm{Gal}(\hat{F}/F)$ are distinct cosets of $\mathrm{Gal}(\hat{F}/F)$ in $\mathrm{Gal}(\hat{F}/E)$. Since $(\mathrm{Gal}(\hat{F}/E) : \mathrm{Gal}(\hat{F}/F)) = d$, we have $\mathrm{Gal}(\hat{F}/E) = \dot{\bigcup}_{i=1}^{d} \sigma_i \mathrm{Gal}(\hat{F}/F)$.

It follows that $\sigma = \sigma_i \eta$ for some $1 \le i \le d$ and $\eta \in \mathrm{Gal}(\hat{F}/F)$. If $2 \le i \le d$, then $v_1 = w|_F = w \circ \sigma|_F = w \circ \sigma_i \circ \eta|_F = w_i|_F = v_i$, which is a contradiction. It follows that $i = 1$, so $\sigma \in \mathrm{Gal}(\hat{F}/F)$, as claimed.

Now, since v totally splits in F, it totally splits in each of the conjugates F' of F over E. By the claim, σ belongs to $\mathrm{Gal}(\hat{F}/F')$. Since the compositum of all of the fields F' is $\hat{F}$, we conclude that $\sigma = 1$, as asserted. □

When we speak about a function field of one variable F/K, we always assume that F/K is a regular extension [FrJ08, Section 3.1]. If $D = \sum_{i=1}^{n} a_i P_i$ is a divisor of F/K with distinct prime divisors $P_1, \ldots, P_n$ and integral coefficients $a_1, \ldots, a_n$, we write $v_P(D) = a_i$ for a prime divisor P of F/K, if $P = P_i$ for some i between 1 and n, and $v_P(D) = 0$ otherwise. The **divisor** and the **pole divisor** of an element $f \in F^\times$ are $\mathrm{div}(f) = \sum_P v_P(f)P$ and $\mathrm{div}_\infty(f) = -\sum_{v_P(f)<0} v_P(f)P$, where P ranges over all prime divisors of F/K and v_P is here the normalized valuation of F associated with P. Note that $\deg(\mathrm{div}_\infty(f)) = [F : K(f)]$. We say that $\mathrm{div}_\infty(f)$ **totally splits** in F if $\mathrm{div}_\infty(f) = P_1 + \cdots + P_m$, where $m = [F : K(f)]$ and $P_1, \ldots, P_m$ are distinct prime divisors of F/K. This holds if and only if the valuation v_∞ of $K(f)/K$ defined by $v_\infty(f) = -1$ totally splits in F.

Lemma 6.3. *Let F be a function field of one variable over a PAC field K. Then F/K has a separating transcendental element f such that the Galois closure $\hat{F}$ of $F/K(f)$ is a regular extension of K. Moreover, given a prime divisor P of $F\tilde{K}/\tilde{K}$, we may choose $f \in F$ such that $v_P(f) > 0$ and $\mathrm{div}_\infty(f)$ totally splits in F and in $\hat{F}$.*

Proof. The prime divisor P is already defined over a finite extension M of K. Let $\sigma_1, \ldots, \sigma_d$ be the distinct K-embeddings of M into $\tilde{K}$. If $p = \mathrm{char}(K) > 0$, let $q = p^j$ be the inseparable degree of M/K. If $\mathrm{char}(K) = 0$, put $q = 1$. Then $D = \sum_{i=1}^{d} q\sigma_i P$ is a positive divisor of F/K and $v_P(D) > 0$.

Since K is PAC, F/K has for each positive integer m distinct prime divisors $Q_1, \ldots, Q_m$ of degree 1 with $Q_i \ne P$, $i = 1, \ldots, m$. Taking m sufficiently large, there exists by Riemann-Roch an element $f \in F^\times$ such that

(7.1a) $\mathrm{div}(f) + Q_1 + \cdots + Q_m \ge D$,

and

(7.1b) $\mathrm{div}_\infty(f) = Q_1 + \cdots + Q_m$.

By (7.1a), $v_P(f) > 0$. By (7.1b), $[F : K(f)] = \deg(\mathrm{div}_\infty(f)) = m$, so by (7.1b) again, $\mathrm{div}_\infty(f)$ totally splits in F. In particular, $F/K(f)$ is a finite separable extension. Let $\hat{F}$ be the Galois closure of $F/K(f)$. Then, by Lemma 6.2, $\mathrm{div}_\infty(f)$ totally splits in $\hat{F}$. In particular, $\hat{F}$ has a K-valuation with residue field K. Thus, by [FrJ08, p. 42, Lemma 2.6.9], $\hat{F}$ is regular over K. □

Remark 6.4 (Comparison of proofs)**.** We say that a field K is **stable** if each finitely generated regular extension F of K has a separating transcendence base $\mathbf{t}$ such that the Galois closure $\hat{F}$ of $F/K(\mathbf{t})$ is regular over K. The stability property of PAC fields is proved in [FrJ76]. The essential case in the proof is that where F is a function field of one variable over K. In that case [FrJ76, Thm. 2.3] constructs for each large prime number l a separating transcendental element t for F/K such that the pole divisor of t over $\tilde{K}(t)$ decomposes as $P_1 + \cdots + P_{l-2} + 2P_{l-1}$, where $P_1, \ldots, P_{l-2}, P_{l-1}$ are distinct prime divisors of $F\tilde{K}/\tilde{K}$. This leads to the conclusion that the Galois closure $\hat{F}$ of $F/K(t)$ satisfies $\mathrm{Gal}(\hat{F}/K(t)) \cong S_l$, from which the regularity of $\hat{F}/K$ easily follows [FrJ76, Lemma 2.1].

On the other hand, the proof of Lemma 6.3, due to Kollár, generates a separating transcendental element f of F/K for each prime divisor P of F/K of degree 1 such that P is a zero of f and the pole divisor of f over $K(t)$ totally splits in F and is of arbitrary large degree, not necessarily prime. This implies that each of the pole divisors of f in $\hat{F}/K$ is of degree 1, so $\hat{F}/K$ is regular. However, that proof gives no clue for the Galois group $\mathrm{Gal}(\hat{F}/K(f))$.

The next lemma is a standard result of algebraic geometry (see [Lan58, p. 152, Cor.] or [Har77, p. 43, Prop. 6.8]).

Lemma 6.5. *Let $f\colon C \to C'$ be a rational map of absolutely irreducible curves defined over a field K with C' projective and C normal. Then, f is a morphism.*

Proof. We have to prove that f is defined at each point $\mathbf{q}$ of C. Replacing C, if necessary by an affine open neighborhood of $\mathbf{q}$, we may assume that C is affine. Let $\mathbf{x}$ be a generic point of C and $\mathbf{y} = (y_0{:}y_1{:}\cdots{:}y_n)$ a homogeneous generic point of C' such that $f(\mathbf{x}) = \mathbf{y}$. Assume without loss that $y_0, y_1, \ldots, y_n$ belong to the function field F of C' over K. Since C is a normal curve, $\mathcal{O}_{C,\mathbf{q}}$ is a discrete valuation ring. Denote the corresponding valuation of F/K by $v_\mathbf{q}$. Now let u be an element of F with $v_\mathbf{q}(u) = \min(v_\mathbf{q}(y_0), v_\mathbf{q}(y_1), \ldots, v_\mathbf{q}(y_n))$. Then each of the elements $u^{-1}y_i$, $i = 0, \ldots, n$, belongs to $\mathcal{O}_{C,\mathbf{q}}$ and at least one of them is a unit. Hence, f is defined at $\mathbf{q}$ and its value is $((u^{-1}y_0)(\mathbf{q}){:}(u^{-1}y_1)(\mathbf{q}){:}\cdots{:}(u^{-1}y_n)(\mathbf{q}))$. □

Theorem 6.6 (Density theorem)**.** *Let K be a PAC field, v a valuation of $\tilde{K}$, and V an absolutely irreducible variety defined over K. Then, $V(K)$ is v-dense in $V(\tilde{K})$.*

Proof. We break up the proof into several parts.

Part A: *We prove that $C(K)$ is v-dense in $C(\tilde{K})$ for each absolutely irreducible projective normal curve C which is defined over K.* Let $\mathbf{p} \in C(\tilde{K})$ and let $\mathcal{U}$ be a v-open neighborhood of $\mathbf{p}$ in $C(\tilde{K})$. Denote the function field of C over K by F. Then F is an algebraic function field of one variable over K which is regular over K [FrJ08, p. 175, Cor. 10.2.2(a)]. Let P be a prime divisor of $F\tilde{K}/\tilde{K}$ whose center at $C_{\tilde{K}}$ is $\mathbf{p}$. Lemma 6.3 gives an $f \in F$ such that $F/K(f)$ is a finite separable extension and the Galois closure $\hat{F}$ of $F/K(f)$ is a regular extension of K. It follows that we may identify $G = \mathrm{Gal}(\hat{F}\tilde{K}/\tilde{K}(f))$ with $\mathrm{Gal}(\hat{F}/K(f))$ via the restriction map. Moreover,

(7.2) $v_P(f) > 0$ and $\mathrm{div}_\infty(f)$ totally splits in $\hat{F}$.

We consider f also as a rational map from C into $\mathbb{P}^1_K$. By Lemma 6.5, f is a morphism. Since C is projective, f is proper [GoW10, p. 386, Cor. 13.41]. By (7.2), f is not constant, hence each of the fibers of f is finite (Otherwise there exists a point $\mathbf{a} \in \mathbb{P}^1_K$ such that $f^{-1}(\mathbf{a})$ is infinite. Since the fiber is closed and C is an irreducible curve, $f^{-1}(\mathbf{a}) = C$, hence $f(C) = \mathbf{a}$, in contrast to the former conclusion.) i.e. f is quasi-finite. It follows that $f\colon C \to \mathbb{P}^1_K$ is a finite morphism [GoW10, p. 358, Cor. 89]. The corresponding function field extension is $F/K(f)$.

Now let $\pi\colon D \to C$ be the projective normalization of C in $\hat{F}$ [Lan58, p. 143, Thm. 5], in particular, D is normal and π is finite [Liu06, p. 121, Prop. 1.25]. Then $\varphi = f \circ \pi$ is a finite morphism of D onto $\mathbb{P}^1_K$. It follows that $\varphi\colon D \to \mathbb{P}^1_K$ is the normalization of $\mathbb{P}^1_K$ in $\hat{F}$ [Liu06, p. 120, Def. 1.24].

We may interpret (7.2) as

(7.3) $f(\mathbf{p}) = (1{:}0)$ and $|\varphi^{-1}((0{:}1))| = [\hat{F} : K(f)]$.

In particular, φ is unramified over (0:1). Since the local ring $\mathcal{O}_{\mathbb{P}^1_K,(0:1)}$ is a discrete valuation ring, it is a Dedekind domain. Therefore, each of the local rings of D lying over $\mathcal{O}_{\mathbb{P}^1_K,(0:1)}$ is flat over $\mathcal{O}_{\mathbb{P}^1_K,(0:1)}$ [Liu06, p. 11, Corollary 2.14]. It follows that φ is étale over (0:1), hence smooth over (0:1).

Since $\pi\colon D \to C$ is finite, so is $\tilde{\pi}\colon \tilde{D} \to \tilde{C}$ [GoW10, p. 325, Prop. 12.11(2)]. Hence, the map $\tilde{\pi}\colon D(\tilde{K}) \to C(\tilde{K})$ is surjective [GoW10, p. 339, Prop. 12.43(2)] and v-continuous. Let $\mathbf{q}$ be a point in $D(\tilde{K})$ lying over $\mathbf{p}$. By (7.3), $\tilde{\varphi}(\mathbf{q}) = (1{:}0)$ and $\mathcal{V} = \tilde{\pi}^{-1}(\mathcal{U})$ is an open neighborhood of $\mathbf{q}$ in $D(\tilde{K})$. By Lemma 5.3, there exists $\mathbf{q}' \in \mathcal{V} \cap D(K)$. Then, $\mathbf{p}' = \pi(\mathbf{q}') \in \mathcal{U} \cap C(K)$. Thus, $C(K)$ is v-dense in $C(\tilde{K})$.

Part B: *We prove that $C(K)$ is v-dense in $C(\tilde{K})$ for each absolutely irreducible curve C which is defined over K.* Again, let $\mathbf{p} \in C(\tilde{K})$ and let $\mathcal{U}$ be a v-open neighborhood of $\mathbf{p}$ in $C(\tilde{K})$. Let C_{simp} be the Zariski-open subset of C consisting of all simple points. Then $C_{\mathrm{simp}}(\tilde{K})$ is v-open and v-dense in $C(\tilde{K})$ [GeJ75, Lemma 2.2], in particular, $\mathcal{U}_0 = C_{\mathrm{simp}}(\tilde{K}) \cap \mathcal{U}$ is nonempty and v-open. Replacing $\mathcal{U}$ by $\mathcal{U}_0$ and C by C_{simp}, we may assume that C is smooth. Similarly, replacing C by a nonempty Zariski-open affine subset, we may assume that $C \subseteq \mathbb{A}_K^n$ for some positive integer n.

Let C^* be the Zariski-closure of C in $\mathbb{P}_K^n$. In particular, we may view C as a Zariski-open subset of C^*. Let $\pi\colon D \to C^*$ be the projective normalization of C^* [Lan58, p. 143, Thm. 5]. Since C is normal, the restriction of π to $\pi^{-1}(C)$ is an isomorphism. Since $\pi^{-1}(C)$ is Zariski-open in D, $\tilde{\pi}^{-1}(\mathcal{U})$ is a nonempty v-open subset of $D(\tilde{K})$. By Part A, $\tilde{\pi}^{-1}(\mathcal{U}) \cap D(K) \neq \emptyset$. Hence, by [GeJ75, Lemma 2.2], $\tilde{\pi}^{-1}(\mathcal{U}) \cap \tilde{\pi}^{-1}(C(K)) \neq \emptyset$. It follows that $\mathcal{U} \cap C(K) \neq \emptyset$, as desired.

Part C: *The general case.* Again, let $\mathcal{U}$ be a nonempty v-open subset of V. By [GeJ75, Lemma 2.4], $V(K_s)$ is v-dense in $V(\tilde{K})$. Hence, we may assume that $\mathcal{U} \cap V(K_s)$ contains a point $\mathbf{p}$. Let C be an absolutely irreducible subcurve of V defined over K with $\mathbf{p} \in C(K_s)$ [JaR98, Lemma 10.1]. Then, $\mathcal{U} \cap C(\tilde{K})$ is a v-open neighborhood of $\mathbf{p}$ in $C(\tilde{K})$. By Part B, $\mathcal{U} \cap C(K) \neq \emptyset$. Consequently, $\mathcal{U} \cap V(K) \neq \emptyset$, as desired. □

7 Embedding Lemma

The essential step in the proof of Abraham Robinson's result about the model completeness of the theory of algebraically closed valued fields is the embedding lemma 7.3 that we prove below. Our presentation follows that of Alexander Prestel in [Pre86, pp. 236–241]. In this section and the following ones we allow valuations of fields to be trivial.

Lemma 7.1. *Let K be an algebraically closed field, E a field extension of K, and v a valuation of E. We denote the valuation ring of v by O_v and use a bar to denote reduction modulo v.*

(a) $\bar{K}$ is algebraically closed, $v(K^\times)$ is a divisible group, and division in $v(K^\times)$ by each $n \in \mathbb{N}$ is unique. Moreover, if K is an algebraic closure of a subfield K_0, then $v(K^\times)$ is the divisible closure of $v(K_0^\times)$.

(b) The ordered group $\Gamma = v(K^\times)$ is dense in itself. That is, for all $\alpha, \beta \in \Gamma$ with $\alpha < \beta$ there exists $\gamma \in \Gamma$ such that $\alpha < \gamma < \beta$. Moreover, if v is non-trivial, then for each $\alpha \in \Gamma$ there exist $\delta, \delta' \in \Gamma$ such that $\delta < \alpha < \delta'$.

(c) Let x be an element of E such that $\bar{x}$ is transcendental over $\bar{K}$. Then for all $a_0, \ldots, a_n \in K$ we have $v(\sum_{i=0}^n a_i x^i) = \min(v(a_0), \ldots, v(a_n))$.

(d) Let x be an element of $E^\times$ such that $v(x) \notin v(K^\times)$. Then the order of $v(x)$ in $v(E^\times)$ is infinite and $v(K(x)^\times) = v(K^\times) \oplus \mathbb{Z}v(x)$.

Proof.

Proof of (a): Consider a polynomial $\bar{f}(X) = X^n + \bar{a}_{n-1}X^{n-1} + \cdots + \bar{a}_0$, with $n \ge 1$ and $a_0, \dots, a_{n-1} \in O_v \cap K$. Since K is algebraically closed, there exists $x \in K$ with $f(x) = 0$. If $v(x) < 0$, then $v(x^{-1}) > 0$ and $1 + a_{n-1}x^{-1} + \cdots + a_0x^{-n} = 0$. Taking residues on both sides, we get the contradiction $1 = 0$. Thus, $v(x) \ge 0$ and $\bar{f}(\bar{x}) = 0$, as desired.

Now let $a \in K^\times$ and let n a positive integer. Then $a^{1/n} \in K^\times$ and $v(a) = nv(a^{1/n})$. Hence, $v(K^\times)$ is divisible. Since $v(K^\times)$ is an ordered group, division by n is unique. Finally, $e = (v(K_0(a)^\times) : v(K_0^\times)) < \infty$. Hence, $ev(a) = v(a_0)$ for some $a_0 \in K_0$.

Proof of (b): Using (a), we may take $\gamma = \frac{\alpha+\beta}{2}$ to prove the first claim in (b). If v is non-trivial, there exists a positive $\varepsilon \in \Gamma$. Then $\alpha - \varepsilon < \alpha < \alpha + \varepsilon$, as desired.

Proof of (c): We choose a $0 \le j \le n$ with $v(a_j) = \min(v(a_0), \dots, v(a_n))$ and let $b_i = a_i a_j^{-1}$, $i = 1, \dots, n$. Then $\bar{b}_j = 1$ and $\bar{b}_i \in \bar{K}$ for $i = 0, \dots, n$. Since $\bar{x}$ is transcendental over $\bar{K}$, we have $\sum_{i=0}^n \bar{b}_i\bar{x}^i \ne 0$, so $v(\sum_{i=0}^n b_ix^i) = 0$. Therefore, $v(\sum_{i=0}^n a_ix^i) = v(a_j) + v(\sum_{i=0}^n b_ix^i) = \min(v(a_0), \dots, v(a_n))$, as claimed.

Proof of (d): First note that the order of $v(x)$ is not only infinite but even infinite modulo $v(K^\times)$. Indeed, if there exist $n \in \mathbb{N}$ and $a \in K^\times$ such that $nv(x) = v(a)$, then $v(x) = v(a^{1/n}) \in v(K^\times)$, in contrast to the assumption on x.

Next consider an element $\sum_{i=0}^n a_ix^i$ in $K[x]$. If $0 \le i < j \le n$, then $v(a_ix^i) \ne v(a_jx^j)$. Otherwise, $(j - i)v(x) = v(a_ia_j^{-1}) \in v(K)$, in contrast to the preceding paragraph. It follows that

$$v\big(\sum_{i=0}^n a_ix^i\big) = \min(v(a_0), v(a_1) + v(x), \dots, v(a_n) + nv(x)) \in v(K^\times) + \mathbb{Z}v(x).$$

By the preceding paragraph, the sum on the right hand side is direct. □

An **embedding** of a valued field (E, v) into a valued field (F, w) is a pair (φ, φ'), where $\varphi\colon E \to F$ is an embedding of fields, $\varphi'\colon v(E^\times) \to w(F^\times)$ is an embedding of ordered groups, and $w(\varphi(e)) = \varphi'(v(e))$ for each $e \in E^\times$. In the sequel we sometimes abuse our language and write φ also for φ' and also for the pair (φ, φ'). If K is a common subfield of E and F, φ is the identity on K, and φ' is the identity on $v(K^\times)$, we say that φ is a K-**embedding**.

Lemma 7.2. *Let $\varphi_0\colon E_0 \to F_0$ be an isomorphism of fields. Let (E, v) and (F, w) be valued fields such that E is an algebraic extension of E_0 and F is a field extension of F_0 which is algebraically closed. Let $\varphi_0'\colon v(E_0^\times) \to w(F_0^\times)$ be an isomorphism of valued groups. Suppose $w(\varphi_0(e)) = \varphi_0'(v(e))$ for each $e \in E_0^\times$. Then it is possible to extend φ_0 to an embedding $\varphi\colon E \to F$ and to extend φ_0' to an embedding $\varphi'\colon v(E^\times) \to w(F^\times)$ of ordered groups such that $w(\varphi(e)) = \varphi'(v(e))$ for each $e \in E^\times$.*

Proof. We choose an algebraic closure $\tilde{E}_0$ of E_0 that contains E and an algebraic closure $\tilde{F}_0$ of F_0 in F. Then we extend φ_0 to an isomorphism $\tilde{\varphi}\colon \tilde{E}_0 \to \tilde{F}_0$. By Chevalley's theorem, v extends to a valuation $\tilde{v}$ of $\tilde{E}_0$. Let $\tilde{\gamma} \in v(\tilde{E}_0^\times)$ (resp. $\tilde{\delta} \in w(\tilde{F}_0^\times)$). Then, there exists $n \in \mathbb{N}$ and there exists a unique $\gamma_0 \in v(E_0^\times)$ (resp. $\delta_0 \in w(F_0^\times)$) such that $\gamma_0 = n\tilde{\gamma}$ (resp. $\delta_0 = n\tilde{\delta}$) (Lemma 7.1(a)). Hence, φ_0' uniquely extends to an isomorphism $\widetilde{\varphi'}\colon \tilde{v}(\tilde{E}_0^\times) \to w(\tilde{F}_0^\times)$. Then, $\tilde{w} = w|_{\tilde{F}_0}$ and $\widetilde{\varphi'} \circ \tilde{v} \circ \tilde{\varphi}^{-1}$ are valuations of $\tilde{F}_0$ that coincide on F_0. Hence, there exists $\sigma \in \mathrm{Aut}(\tilde{F}_0/F_0)$ such that $\tilde{w} \circ \sigma = \widetilde{\varphi'} \circ \tilde{v} \circ \tilde{\varphi}^{-1}$ [Efr06, p. 131, Thm. 14.3.2], so $\tilde{w} = \widetilde{\varphi'} \circ \tilde{v} \circ (\sigma \circ \tilde{\varphi})^{-1}$. Then, $\varphi = \sigma \circ \tilde{\varphi}|_E$ is an embedding of E into F that extends φ_0 and $\varphi' = \widetilde{\varphi'}|_{E^\times}$ is an embedding of $v(E^\times)$ into $w(F^\times)$ that extends φ_0' such that $w(\varphi(e)) = \varphi'(v(e))$ for each $e \in E^\times$. □

Recall that a structure $\mathcal{A}$ of a language $\mathcal{L}$ with a domain A is **$\aleph_1$-saturated** if it satisfies the following condition: Let $\varphi_1, \varphi_2, \varphi_3, \dots$ be formulas of $\mathcal{L}$ in the free variables $X_1, X_2, X_3, \dots$ with parameters in A. Suppose for each n there exist $a_1, a_2, a_3, \dots \in A$ such that $\varphi_1(\mathbf{a}), \dots, \varphi_n(\mathbf{a})$ hold in $\mathcal{A}$. Then there exist $x_1, x_2, x_3, \dots \in A$ such that each $\varphi_n(\mathbf{x})$ holds in $\mathcal{A}$ [FrJ08, p. 143].

Lemma 7.3 (Embedding lemma)**.** *Let K be a countable algebraically closed field, (E, v) a valued field such that E is a function field of one variable over K, and (F, w) an $\aleph_1$-saturated algebraically closed non-trivial valued field such that $K \subseteq F$ and $v|_K = w|_K$. Then there exists a K-embedding $\varphi\colon (E, v) \to (F, w)$.*

Proof. Let $O = O_v \cap K = O_w \cap K$ and use a bar to denote reduction with respect to both v and w. In particular, $\bar{K}, \bar{E}, \bar{F}$ are the residue fields of K, E, F, respectively, $\bar{K} \subseteq \bar{E}, \bar{F}$, and both $\bar{K}$ and $\bar{F}$ are algebraically closed (Lemma 7.1(a)).

If $x \in E$ is transcendental over K, then by assumption, E is algebraic over $E_0 = K(x)$. Hence, in order to prove the lemma, it suffices by Lemma 7.2 to prove the following claim.

Claim: *There exist $x \in E$ and $y \in F$ transcendental over K, and there exists a K-isomorphism*

$$\varphi\colon (K(x), v(K(x)^\times)) \to (K(y), w(K(y)^\times))$$

such that $\varphi(x) = y$ and $w(\varphi(e)) = \varphi'(v(e))$ for each $e \in K(x)^\times$.

The proof of the Claim splits into three cases.

Case 1: $\bar{K} \neq \bar{E}$. We choose $x \in O_v$ such that $\bar{x} \notin \bar{K}$. Then $x \notin K$, so x is transcendental over K.

Since $\bar{K}$ is algebraically closed, $\bar{K}$ is infinite. Hence, for all $a_1, \dots, a_n \in O$ there exists $y \in O$ such that $\bar{y} \neq \bar{a}_i$, so $w(y - a_i) = v(y - a_i) = 0$ for $i = 1, \dots, n$. Since O is countable and (F, w) is $\aleph_1$-saturated, there exists $y \in O_w$ such that $w(y - a) = 0$ for each $a \in O$. This means that $\bar{y} \notin \bar{K}$. As in the preceding paragraph, y is transcendental over K. Therefore, there is a unique K-isomorphism $\varphi\colon K(x) \to$

$K(y)$ such that $\varphi(x) = y$. Moreover, since both $\bar{x}$ and $\bar{y}$ are transcendental over $\bar{K}$, Lemma 7.1(c) implies that $w(\sum_{i=0}^n a_i y^i) = v(\sum_{i=0}^n a_i x^i)$ for all $a_0, \dots, a_n \in K$. Thus, $w(\varphi(e)) = v(e) = \varphi'(v(e))$, where $\varphi' = \mathrm{id}_{v(K(x)^\times)}$, for all $e \in K(x)^\times$, as desired.

Case 2: $v(E^\times) \neq v(K^\times)$. We choose $x \in E^\times$ such that

(8.1) $v(x) \notin v(K^\times)$.

We consider $a_1, \dots, a_n \in K$ and assume without loss that

$$v(a_1) \le \cdots \le v(a_m) < v(x) < v(a_{m+1}) \le \cdots \le v(a_n) \tag{7.2}$$

for some m between 0 and n. By convention, if $m = 0$, then relation (8.2) becomes $v(x) < v(a_1) \le \cdots \le v(a_n)$ and if $m = n$, then relation (8.2) simplifies to $v(a_1) \le \cdots \le v(a_n) < v(x)$. By assumption, $w(a_i) = v(a_i)$ for $i = 1, \dots, n$. If $m = 0$, then Lemma 7.1(b) gives $y \in F$ such that $w(y) < w(a_1)$. If $m = n$, then Lemma 7.1(b) gives $y \in F$ such that $w(a_n) < w(y)$. Otherwise, Lemma 7.1(b) gives $y \in F$ such that $w(a_m) < w(y) < w(a_{m+1})$. Note that the first two cases use the assumption that w is non-trivial.

Since K is countable and (F, w) is $\aleph_1$-saturated, there exists $y \in F$ such that

(8.3) for all $a \in K$, $v(x) < v(a)$ implies that $w(y) < w(a)$, and $v(x) > v(a)$ implies that $w(y) > w(a)$.

In particular,

(8.4) $y \notin K$ and $w(y) \notin w(K^\times)$.

Since K is algebraically closed, both x and y are transcendental over K. Let $\varphi\colon K(x) \to K(y)$ be the unique K-isomorphism with $\varphi(x) = y$. By (8.1) (resp. (8.4)) and Lemma 7.1(d), the order of $v(x)$ (resp. $w(y)$) modulo $v(K^\times)$ is infinite and

(8.5) $v(K(x)^\times) = v(K^\times) \oplus \mathbb{Z}v(x)$ (resp. $w(K(y)^\times) = v(K^\times) \oplus \mathbb{Z}w(y)$).

Hence, there is an isomorphism $\varphi'\colon v(K(x)^\times) \to v(K(y)^\times)$ of ordered groups (by (8.3)) which is the identity map on $v(K^\times)$ such that $\varphi'(v(x)) = w(y)$. It follows from (8.3) that $\varphi'(v(x-a)) = w(y-a)$ for each $a \in K$. Hence,

$$\varphi'(v(a_0 \prod_{i=1}^m (x - a_i))) = w(a_0 \prod_{i=1}^m (y - a_i))$$

for all $a_0, a_1, \dots, a_m \in K$ with $a_0 \neq 0$. Since K is algebraically closed, this implies that $w(\varphi(e)) = \varphi'(v(e))$ for each $e \in K(x)^\times$, as claimed.

Case 3: *$\bar{E} = \bar{K}$ and $v(E^\times) = v(K^\times)$.* We choose $x \in E$, transcendental over K and prove:

(8.6) For all $a_1, \dots, a_n \in K$ there exists $y \in K$ such that $w(y - a_i) = v(x - a_i)$, $i = 1, \dots, n$.

Indeed, by assumption there exists for each $1 \le i \le n$ an element $b_i \in K^\times$ such that $v(x - a_i) = v(b_i)$. We choose $1 \le j \le n$ such that $v(b_j) = \max(v(b_1), \dots, v(b_n))$. Then, $v((x - a_j)b_j^{-1}) = 0$, so by assumption, there exists $c \in K$ with $\bar{c} = \overline{(x - a_j)b_j^{-1}}$. Hence,

$$v\Big(\frac{x - (a_j + cb_j)}{b_j}\Big) = v\Big(\frac{x - a_j}{b_j} - c\Big) > 0.$$

Set $y = a_j + cb_j$. Then, $v(x - y) > v(b_j) \ge v(b_i) = v(x - a_i)$, hence

$$w(y - a_i) = v(y - a_i) = v((x - a_i) - (x - y)) = v(x - a_i), \qquad i = 1, \dots, n,$$

as claimed.

Since K is countable and (F, w) is $\aleph_1$-saturated, there exists $y \in F$ such that

(8.7) $w(y - a) = v(x - a)$ for all $a \in K$.

Since $x \notin K$, (8.7) implies that $y \notin K$. Since K is algebraically closed, y is transcendental over K. Let $\varphi\colon K(x) \to K(y)$ be the unique K-isomorphism with $\varphi(x) = y$. Again, since K is algebraically closed, each non-constant $u \in K[x]$ can be written as $u = c_0 \prod_{i=1}^m (x - c_i)$ with $c_0, c_1, \dots, c_m \in K$ and $c_0 \neq 0$, so $\varphi(u) = c_0 \prod_{i=1}^m (y - c_i)$. It follows from (8.7) that $w(\varphi(u)) = v(u)$. Hence, the latter relation holds for each $u \in K(x)^\times$, as desired. □

8 Algebraically Closed Valued Fields

The goal of this section is to prove Abraham Robinson's model completeness theorem for the theory of algebraically closed valued fields and also the stronger result about the elimination of quantifiers in that theory. For both goals we have to fix the first order language with which we want to speak about the algebraically closed fields. If we restrict ourselves only to the model completeness, it suffices to extend the language $\mathcal{L}(\text{ring})$ of rings [FrJ08, p. 135, Example 7.3.1] with a unary predicate symbol O whose interpretation in a valued field (K, v) is the valuation ring O_v.

We start with the language of rings $\mathcal{L}(\text{ring})$ with the constant symbols $0, 1$, the binary function symbol $+$ (for addition), the binary function symbol $\cdot$ (for multiplication), and the unary function symbol $-$ (for taking the negative)[2]. The

[2] Note that the symbol $-$ does not appear among the function symbols of $\mathcal{L}(\text{ring})$ in [FrJ08, p. 135, Example 7.3.1]. We have been forced here to include this symbol in $\mathcal{L}(\text{ring})$ in order to be able to prove Lemma 8.1(a).

axioms for fields in that language are:

$$(8.1)\qquad \begin{aligned}
&(\forall X)(\forall Y)(\forall Z)[(X+Y)+Z = X+(Y+Z)];\\
&(\forall X)(\forall Y)[X+Y = Y+X];\\
&(\forall X)[X+0 = X];\\
&(\forall X)[X+(-X) = 0];\\
&(\forall X)(\forall Y)(\forall Z)[(XY)Z = X(YZ)];\\
&(\forall X)(\forall Y)[XY = YX];\\
&(\forall X)[1\cdot X = X];\\
&(\forall X)[X\neq 0 \to (\exists Y)[XY = 1]];\\
&1\neq 0;\ \text{and}\\
&(\forall X)(\forall Y)(\forall Z)[X(Y+Z) = XY+XZ].
\end{aligned}$$

Note that a substructure R of a field K in the language $\mathcal{L}(\text{ring})$ is a subset of K that contains $0, 1$ and is closed under addition, negation, and multiplication. Thus, R is an integral domain. If $K = \text{Quot}(R)$ is a valued field, then the main axiom for valued fields, "for all nonzero x we have $x \in O$ or $x^{-1} \in O$", does not make sense over R. This forces us to replace the monadic symbol O by the **division relation** associated with each valued field.

For each valued field (K, v) we define a binary relation $|_v$ on K by

$$(8.2)\qquad a|_v b \iff v(a) \le v(b).$$

It satisfies the following axioms (where we omit the index v):

(9.3a) $1|0$.

(9.3b) $(\forall X)[X|X]$.

(9.3c) $(\forall X)(\forall Y)(\forall Z)[X|Y \wedge Y|Z \to X|Z]$.

(9.3d) $(\forall X)(\forall Y)[X|Y \vee Y|X]$.

(9.3e) $(\forall X)(\forall Y)(\forall Z)[X|Y \to XZ|YZ]$.

(9.3f) $(\forall X)(\forall Y)(\forall Z)[X|Y \wedge X|Z \to X|(Y+Z)]$.

We call $|$ a **division relation** on K.

Conversely, if $|$ is a division relation on K, then $O = \{a \in K \mid 1|a\}$ is a valuation ring of K. Indeed, by (9.3a) and (9.3b), $0, 1 \in O$. By (9.3d), we have $1|-1$ or $-1|1$. By (9.3e), the latter case implies $1|-1$. Hence, in any case $-1 \in O$. If $a, b \in O$, then $1|a$ and $1|b$, so $a+b \in O$, by (9.3f). By (9.3e), $a|ab$, so by (9.3c), $1|ab$, which implies $ab \in O$. Finally, if $a \in K^\times$, then $1|a$ or $a|1$ (by (9.3d)). In the latter case $1|a^{-1}$ (by (9.3e)), so in each case either $a \in O$ or $a^{-1} \in O$. Thus,

O defines a valuation of K (which may be trivial) whose division relation is $|$. It follows that the correspondence between valuations and division relations on K is bijective. We denote the division relation on K that corresponds to a valuation v of K by $|_v$. The advantage of the division relation is that it allows to treat valuation rings as first order structures.

The language of valued fields will therefore be the extension of $\mathcal{L}(\text{ring})$ by the division symbol $|$. We denote the extended language by $\mathcal{L}_{\text{val}}(\text{ring})$. Let T_{val} be the theory of $\mathcal{L}_{\text{val}}(\text{ring})$ that consists of the axioms (9.1) of fields and the axioms (9.3) for $|$.

Lemma 8.1. *Let (E, v) and (F, w) be valued fields. Let $\varphi_0\colon (R, |_{v,0}) \to (S, |_{w,0})$ be an isomorphism of substructures of $(E, |_v)$ and $(F, |_w)$, respectively. Let* $K = \mathrm{Quot}(R)$ *and* $L = \mathrm{Quot}(S)$. *Then*

(a) φ_0 extends to an isomorphism $\varphi\colon (K, v|_K) \to (L, w|_L)$ of valued fields.

(b) If E and F are algebraically closed, then φ extends to an isomorphism $\tilde{\varphi}\colon (\tilde{K}, v|_{\tilde{K}}) \to (\tilde{L}, w|_{\tilde{L}})$.

Proof. As usual, φ_0 extends to an isomorphism $\varphi\colon K \to L$ of the quotient fields. If $x, y \in K$ satisfy $x|_v y$, then there exists $b \in R$, $b \neq 0$, such that $bx, by \in R$, and then $bx|_v by$. Hence, $\varphi(bx)|_w \varphi(by)$, so $\varphi(b)\varphi(x)|_w \varphi(b)\varphi(y)$. Since $\varphi(b) \neq 0$, we have $\varphi(x)|_w \varphi(y)$. It follows that $\varphi\colon (K, v|_K) \to (L, w|_L)$ is an isomorphism of valued fields, which proves (a).

In order to prove (b), we use the valuations rather than the division relations. Let $\psi\colon \tilde{K} \to \tilde{L}$ be an isomorphism of fields that extends φ. Then, there exists $\sigma \in \mathrm{Aut}(\tilde{L}/L)$ such that $\sigma \circ \psi\colon (\tilde{K}, v|_{\tilde{K}}) \to (\tilde{L}, w|_{\tilde{L}})$ extends $\varphi\colon (K, v|_K) \to (L, w|_L)$ [Efr06, p. 131, Thm. 14.3.21]. Thus, $\tilde{\varphi} = \sigma \circ \psi$ is the desired isomorphism. □

Whenever we speak about a "formula $\varphi(X_1, \dots, X_n)$" we mean that the free variables of that formula belong to the set $\{X_1, \dots, X_n\}$.

Definition 8.2. Let T be a theory in a first order language $\mathcal{L}$.

(a) We say that T is **model complete** if whenever a model $\mathcal{A}$ of T is a substructure of another model $\mathcal{B}$ of T, the model $\mathcal{A}$ is an elementary substructure of $\mathcal{B}$.

(b) We say that T has the **amalgamation property** if whenever two models $\mathcal{B}, \mathcal{C}$ of T contain a common $\mathcal{L}$-substructure $\mathcal{A}$, there exists a model $\mathcal{D}$ of T and embeddings $f\colon \mathcal{B} \to \mathcal{D}$ and $g\colon \mathcal{C} \to \mathcal{D}$ that coincide on $\mathcal{A}$.

(c) We say that T admits **elimination of quantifiers** if for every formula $\varphi(X_1, \dots, X_n)$ of the language $\mathcal{L}$ there exists a quantifier free formula $\psi(X_1, \dots, X_n)$ such that for every model $\mathcal{A}$ of T with a domain A and for all $a_1, \dots, a_n \in A$, the truth of $\varphi(\mathbf{a})$ in $\mathcal{A}$ is equivalent to the truth of $\psi(\mathbf{a})$ in $\mathcal{A}$.

We cite two theorems about the concepts just defined.

Proposition 8.3. *Let T be a theory in a first order language $\mathcal{L}$.*

(a) T admits elimination of quantifiers if and only if T is model complete and has the amalgamation property [Pre86, p. 193, Satz 3.22].

(b) T admits elimination of quantifiers if for every two models $\mathcal{B}, \mathcal{C}$ of T with domains B, C respectively, for every finitely generated common substructure $\mathcal{A}$ with domain A, and for every quantifier free formula $\eta(X)$ with parameters in A, the existence of $b \in B$ such that $\eta(b)$ holds in $\mathcal{B}$ implies the existence of $c \in C$ such that $\eta(c)$ holds in $\mathcal{C}$ [Pre86, p. 187, Satz 3.20].

A valued field (K, v) is non-trivial if and only if there exists $a \in K$ such that $v(a) > 0$, equivalently if

(9.4) there exists $x \in K$ such that $\neg(1|x)$.

Theorem 8.4 (Abraham Robinson)**.** *The theory of algebraically closed non-trivial valued fields in the language $\mathcal{L}_{\mathrm{val}}(\mathrm{ring})$ admits elimination of quantifiers, hence it is model complete and has the amalgamation property.*

Proof. By Proposition 8.3, it suffices to consider algebraically closed non-trivial valued fields (E, v) and (F, w), a common finitely generated substructure $(K, |)$ of $(E, |_v)$ and $(F, |_w)$, a quantifier free formula $\eta(X)$ with parameters in K such that there exists $x \in E$ with $E \models \eta(x)$, and to prove that there exists $y \in F$ such that $F \models \eta(y)$. In particular, K is a countable integral domain. By Lemma 8.1, we may replace K with the algebraic closure of $\mathrm{Quot}(K)$. We therefore assume that K is an algebraically closed field and use henceforth valuations rather than division relations.

Now recall that if $(F^*, w^*) = (F, w)^{\mathbb{N}}/\mathcal{D}$ is a nonprincipal ultrapower of (F, w), then (F^*, w^*) is $\aleph_1$-saturated [FrJ08, p. 143, Lemma 7.7.4] and is an elementary extension of (F, w) [FrJ08, p. 144, Prop. 7.7.5]. Replacing (F, w) by (F^*, w^*), if necessary, we may assume that (F, w) is $\aleph_1$-saturated.

Let η and x be as in the first paragraph of the proof. By the embedding lemma 7.3, there exists a K-embedding $\varphi\colon (K(x), v|_{K(x)}) \to (F, w)$ of valued fields. Set $y = \varphi(x)$. Since $\eta(X)$ is quantifier free, $\eta(x)$ holds in $(K(x), v|_{K(x)})$. Hence, $\eta(y)$ holds in $(K(y), w|_{K(y)})$ and therefore also in (F, w), as desired. □

9 Existential Closedness of PAC Fields

Using the density theorem 6.6, we prove that every PAC valued field (K, v) is existentially closed in each regular valued field extension (F, v) such that F is also PAC.

Theorem 9.1. *Let $(F,v)/(K,v)$ be an extension of non-trivial valued fields such that K is PAC and F/K is regular. Then (K,v) is existentially closed in (F,v) in the language $\mathcal{L}_{\mathrm{val}}(\mathrm{ring})$.*

Proof. We break up the proof into two parts.

Part A: *Simplified form for existential formulas.* Let $\mathbf{X} = (X_1,\dots,X_m)$ and $\mathbf{Y} = (Y_1,\dots,Y_n)$ be tuples of variables and let $f(\mathbf{X},\mathbf{Y})$ and $g(\mathbf{X},\mathbf{Y})$ be polynomials with coefficients in K. We may replace the formula $f(\mathbf{X},\mathbf{Y})|g(\mathbf{X},\mathbf{Y})$ by the equivalent formula $v(g(\mathbf{X},\mathbf{Y})) \ge v(f(\mathbf{X},\mathbf{Y}))$ and the formula $f(\mathbf{X},\mathbf{Y}) \nmid g(\mathbf{X},\mathbf{Y})$ by the equivalent formula $v(f(\mathbf{X},\mathbf{Y})) > v(g(\mathbf{X},\mathbf{Y}))$. Thus, every existential formula $\varphi(X_1,\dots,X_m)$ in $\mathcal{L}_{\mathrm{val}}(\mathrm{ring})$ with parameters in K is equivalent (in the theory of valued fields) to a formula of the form

$$\begin{aligned}(\exists Y_1)\cdots(\exists Y_n)\bigvee_{i\in I}\bigwedge_{j\in J}\big[f_{ij}(\mathbf{X},\mathbf{Y}) = 0 \wedge f'_{ij}(\mathbf{X},\mathbf{Y}) \ne 0 \\ \wedge\, v(g_{ij}(\mathbf{X},\mathbf{Y})) \ge v(g'_{ij}(\mathbf{X},\mathbf{Y})) \\ \wedge\, v(h_{ij}(\mathbf{X},\mathbf{Y})) > v(h'_{ij}(\mathbf{X},\mathbf{Y}))\big]\end{aligned} \tag{9.1}$$

where I and J are finite sets, and $f_{ij}, f'_{ij}, g_{ij}, g'_{ij}, h_{ij}, h'_{ij}$ are polynomials with coefficients in K for all $(i,j) \in I \times J$. We have to prove that

(10.2) if there exists $\mathbf{x} \in F^m$ such that $\varphi(\mathbf{x})$ holds in (F,v), then there exists $\mathbf{a} \in K^m$ such that $\varphi(\mathbf{a})$ holds in (K,v).

First note that the disjunction symbol commutes with the existential quantifiers. Moreover, if (10.2) holds for one of the disjuncts of φ, it also holds for φ. Thus, it suffices to consider φ of the form

$$\begin{aligned}(\exists Y_1)\cdots(\exists Y_n)\bigwedge_{j\in J}\big[f_j(\mathbf{X},\mathbf{Y}) = 0 \wedge f'_j(\mathbf{X},\mathbf{Y}) \ne 0 \\ \wedge\, v(g_j(\mathbf{X},\mathbf{Y})) \ge v(g'_j(\mathbf{X},\mathbf{Y})) \\ \wedge\, v(h_j(\mathbf{X},\mathbf{Y})) > v(h'_j(\mathbf{X},\mathbf{Y}))\big],\end{aligned} \tag{9.3}$$

where $f_j, f'_j, g_j, g'_j, h_j, h'_j \in K[\mathbf{X},\mathbf{Y}]$ for all $j \in J$.

The formula $f'_j(\mathbf{X},\mathbf{Y}) \ne 0$ is equivalent to $(\exists Z_{j1})[Z_{j1}f'_j(\mathbf{X},\mathbf{Y}) - 1 = 0]$. The formula $v(g_j(\mathbf{X},\mathbf{Y})) \ge v(g'_j(\mathbf{X},\mathbf{Y}))$ is equivalent to

$$\begin{aligned}&[g_j(\mathbf{X},\mathbf{Y}) = 0 \wedge g'_j(\mathbf{X},\mathbf{Y}) = 0] \\ &\qquad \vee \big[g'_j(\mathbf{X},\mathbf{Y}) \ne 0 \wedge (\exists Z_{j2})[g_j(\mathbf{X},\mathbf{Y}) = Z_{j2}g'_j(\mathbf{X},\mathbf{Y}) \wedge v(Z_{j2}) \ge 0]\big].\end{aligned}$$

Finally, the formula $v(h_j(\mathbf{X},\mathbf{Y})) > v(h'_j(\mathbf{X},\mathbf{Y}))$ is equivalent to

$$[h_j(\mathbf{X},\mathbf{Y}) = 0 \wedge h'_j(\mathbf{X},\mathbf{Y}) \neq 0] \\ \vee \big[h_j(\mathbf{X},\mathbf{Y}) \neq 0 \wedge (\exists Z_{j3})[h_j(\mathbf{X},\mathbf{Y}) = Z_{j3}h'_j(\mathbf{X},\mathbf{Y}) \wedge v(Z_{j3}) > 0]\big].$$

Since Z_{j1}, Z_{j2}, Z_{j3} do not occur among the coordinates of $\mathbf{Y}$, we may pull over the quantifiers $\exists Z_{j1}$, $\exists Z_{j2}$, and $\exists Z_{j3}$ to the left of (10.3). Then we rename each Z_{jk} as Y_r for some $r > n$ and finally enlarge n and repeat the first two simplification steps to conclude that φ has the form

$$(\exists Y_1)\cdots(\exists Y_n)\bigwedge_{j\in J}[f_j(\mathbf{X},\mathbf{Y}) = 0 \wedge v(Y_1) \succ_1 0 \wedge \cdots \wedge v(Y_n) \succ_n 0],$$

where for $i = 1,\dots,n$ the relation $\succ_i$ is either $\geq$, or $>$, or the trivial relation $0 = 0$.

Part B: *Existential closedness.* Let $\mathbf{x} \in F^m$ be such that $\varphi(\mathbf{x})$ holds in (F, v). Then there exists $\mathbf{y} \in F^n$ such that

$$\bigwedge_{j\in J}[f_j(\mathbf{x},\mathbf{y}) = 0 \wedge v(y_1) \succ_1 0 \wedge \cdots \wedge v(y_n) \succ_n 0].$$

Since F/K is a regular extension, so is $K(\mathbf{x},\mathbf{y})/K$. Thus, $(\mathbf{x},\mathbf{y})$ is a generic point of an affine absolutely irreducible variety W defined in $\mathbb{A}^{m+n}$ over K [FrJ08, p. 175, Cor. 10.2.2]. Now we extend v to a valuation v of $\tilde{F}$ and again denote the restriction of v to $\tilde{K}$ by v. By Theorem 8.4, the first order theory of algebraically closed non-trivial valued field is model complete. In particular, $(\tilde{K}, v)$ is an elementary substructure of $(\tilde{F}, v)$. Therefore, there exists $(\tilde{\mathbf{a}}, \tilde{\mathbf{b}}) \in W(\tilde{K})$ such that $v(\tilde{b}_i) \succ_i 0$ for each i. Since $W(K)$ is v-dense in $W(\tilde{K})$ (Theorem 6.6), we conclude that there exists $(\mathbf{a}, \mathbf{b}) \in W(K)$ such that $v(b_i) \succ_i 0$ for each i. Consequently, $\varphi(\mathbf{a})$ holds in (K, v). □

10 Model Companion

Model companions and model completions of first order theories generalize the relation that the theory of algebraically closed fields has relative to the theory of all fields. In this section we prove the existence of a model companion for the theory of non-trivial valued fields in a language that allows only extensions L/K of fields such that K is algebraically closed in L. In the next section we add more predicates to the language that force the field extensions we consider to be regular. This results in a model completion of the corresponding theory.

Definition 10.1. Let T and $\tilde{T}$ be theories in a first order language $\mathcal{L}$. We say that $\tilde{T}$ is a **model companion** of T if the following holds:

(11.1a) Each model of $\tilde{T}$ is a model of T.

(11.1b) Each model of T can be embedded into a model of $\tilde{T}$.

(11.1c) $\tilde{T}$ is model complete (Definition 8.2(a)).

We say that $\tilde{T}$ is a **model completion** of T if in addition

(11.1d) T has the amalgamation property (Definition 8.2(b)).

Remark 10.2. By (11.1a) and (11.1b) of Definition 10.1, (11.1d) is equivalent to the statement

(11.1.d') $\tilde{T}$ has the amalgamation property.

In this case $\tilde{T}$ admits, by Proposition 8.3(a), elimination of quantifiers.

Example 10.3. (a) The theory of algebraically closed fields is the model completion of the theory of fields in $\mathcal{L}(\text{ring})$.
(Essentially [FrJ08, p. 168, Cor. 9.3.2]).

(b) The theory RCF of real closed fields is the model companion of the theory OF of ordered fields in the language $\mathcal{L}(\text{ring}, <)$, where $<$ is the ordering symbol [Pre86, Kor. 4.8]. By [VdD78, p. 40], RCF is even the model completion of OF.

(c) The theory ACF_{val} of algebraically closed non-trivial valued fields is the model completion of the theory F_{val} of valued fields in the language $\mathcal{L}_{\text{val}}(\text{ring})$. This follows from Chevalley's extension theorem of valuations and from Theorem 8.4 of Abraham Robinson [Pre86, p. 241, Kor. 4.18].

Example 10.4. We augment the language $\mathcal{L}(\text{ring})$ to a language $\mathcal{L}_R(\text{ring})$ by an n-ary relation symbol R_n for each positive integer n. Let T_R be the theory of fields together with the axioms

$$R_n(X_1, \ldots, X_n) \leftrightarrow (\exists Z)[Z^n + X_1 Z^{n-1} + \cdots + X_n = 0]. \tag{10.2}$$

For each field K and every n we interpret R_n in K in the unique way such that (11.2) holds. Then consider K also as a model of T_R. If an $\mathcal{L}(\text{ring})$-structure L is a field, then an $\mathcal{L}_R(\text{ring})$-substructure K of the $\mathcal{L}_R(\text{ring})$-structure L is an integral domain contained in L such that every equation $X^n + a_{n-1}X^{n-1} + \cdots + a_0 = 0$ with coefficients in K that has a root in L has a root in K. In particular, if K is a field, then K is algebraically closed in L.

By [FrJ08, p. 663, Thm. 27.2.3], T_R has a model companion $\tilde{T}_R$. A field K is a model of $\tilde{T}_R$ if and only if K is **1-imperfect** (i.e. $\text{char}(K) = 0$ or $\text{char}(K) = p > 0$ and $[K : K^p] = p$), ω-free (i.e. every finite embedding problem over K is solvable), and PAC. However, T_R has no model completion, because T_R does not have the amalgamation property [FrJ08, p. 664, Example 27.2.4].

Let $T_{\rm val}$ be the theory of $\mathcal{L}_{\rm val}(\text{ring})$ that consists of the usual axioms of fields (9.1) of Section 8 and the axioms (9.3) of Section 8 for $|$.

Adding the division symbol to the language $\mathcal{L}_R(\text{ring})$, we get a first order language $\mathcal{L}_{R,\rm val}(\text{ring})$ for valued fields such that a valued field $(F, |)$ is a substructure of another valued field $(F', |')$ if and only if F is an algebraically closed subfield of F' and the restriction of $|'$ to F is $|$. Let $T_{R,\rm val}$ be the theory in $\mathcal{L}_{R,\rm val}(\text{ring})$ that consists of T_R and the axioms (9.3) of Section 8.

Remark 10.5. Every valued field (K, v) has an extension (K', v'), where K'/K is regular and v' is a non-trivial valuation of K'.

Indeed, if v is non-trivial, let $K' = K$ and $v' = v$. Otherwise, choose an indeterminate t, let $K' = K(t)$ and v' any of the valuations of K'/K (e.g. the one with $v'(t) = -1$).

Theorem 10.6. *The theory $T_{R,\rm val}$ of valued fields has a model companion $\tilde{T}_{R,\rm val}$ in the language $\mathcal{L}_{R,\rm val}(\text{ring})$. A non-trivial valued field (K, v) is a model of $\tilde{T}_{R,\rm val}$ if and only if K is 1-imperfect, ω-free, and PAC.*

Proof. Let $\tilde{T}_{R,\rm val}$ be the theory $\tilde{T}_R$ of Example 10.4 together with the axioms (9.3) and (9.4) of Section 8 for non-trivial valued fields.

If (K, v) is a model of $\tilde{T}_{R,\rm val}$, then K is a model of $\tilde{T}_R$. By Example 10.4, K is 1-imperfect, ω-free, and PAC. Conversely, if a field K is 1-imperfect, ω-free, and PAC, then by Example 10.4, K is a model of $\tilde{T}_R$. Hence, if v is a non-trivial valuation of K, then (K, v) is a model of $\tilde{T}_{R,\rm val}$.

If (K, v) is a valued field, we extend it to another model (K', v') with K'/K regular and v' non-trivial (Remark 10.5). By Example 10.4 and by Definition 10.1(b), there exists a field extension L of K' which is 1-imperfect, ω-free and PAC such that K' is algebraically closed in L. Then, K is algebraically closed in L. By Chevalley, v' extends to a valuation w of L [Lan58, p. 8, Thm. 1]. In particular w is non-trivial. Hence, (L, w) is a model of $\tilde{T}_{R,\rm val}$ that extends (K, v) in the language $\mathcal{L}_{R,\rm val}(\text{ring})$. Thus, $\tilde{T}_{R,\rm val}$ satisfies Condition (11.1a) and (11.1b) of Definition 10.1.

Claim: *If $(K, v) \subseteq (L, w)$ is an extension of models of $\tilde{T}_{R,\rm val}$, then the field L is a regular extension of K.* Indeed, by Example 10.4, K is algebraically closed in L. Moreover, by the second paragraph of the proof, K is 1-imperfect. Hence, by [FrJ08, p. 47, Lemma 2.7.5], L/K is a regular extension.

Next observe that if (K, v) and (L, w) are models of $\tilde{T}_{R,\rm val}$ and $(K, v) \subseteq (L, w)$, then K is PAC, and L/K is a regular extension (by the Claim). Moreover, both v and w are non-trivial valuations. Hence, by Theorem 9.1, (K, v) is existentially closed in (L, w) in the language $\mathcal{L}_{\rm val}(\text{ring})$. Since each of the axioms R_n is equivalent to an existential formula of $\mathcal{L}_{\rm val}(\text{ring})$ (Example 10.4), (K, v) is existentially closed in (L, w) in the language $\mathcal{L}_{R,\rm val}(\text{ring})$. By Abraham Robinson, $\tilde{T}_{R,\rm val}$

is model complete [FrJ08, p. 659, Lemma 27.1.11], that is it satisfies Condition (11.1c) of Definition 10.1. Consequently, $\tilde{T}_{R,\mathrm{val}}$ is a model companion of $T_{R,\mathrm{val}}$. □

11 The Non-Amalgamation Property of $T_{R,\mathrm{val}}$

Example 27.2.4 of [FrJ08] constructs fields K and L of positive characteristic p that are algebraically closed in no common field extension. However, K and L are isomorphic, so that example does not prove that the theory T_R does not have the amalgamation property, as the paragraph before that example claims.

In this section we correct that example by constructing three fields E, K, L of characteristic p such that E is algebraically closed in both K and L but K and L are algebraically closed in no common field extension. This implies that none of the theories T_R and $T_{R,\mathrm{val}}$ has the amalgamation property.

Our construction uses the notions of differentials of fields. For each field L of characteristic p we consider the vector space $\mathrm{Der}(L,L)$ over $\mathbb{F}_p$ of all derivations of L into L and the map $d_L\colon L \to \mathrm{Hom}(\mathrm{Der}(L,L),L)$ defined by $(d_La)(D) = Da$ for each $a \in L$. It satisfies the relations

$$p \cdot d_La = 0,\ d_L(a+b) = d_La + d_Lb,\ \text{and } d_L(ab) = ad_Lb + bd_La \tag{11.1}$$

for all $a, b \in L$. Each d_La with $a \in L$ is called a **differential** of L. Repeated application of (12.1) leads for $a_0, a_1, \ldots, a_n \in L$ to the following formula:

$$d_L(a_0^pa_1^{i_1}\cdots a_n^{i_n}) = \sum_{j=1}^n a_0^pa_1^{i_1}\cdots a_{j-1}^{i_{j-1}} \cdot i_ja_j^{i_j-1} \cdot a_{j+1}^{i_{j+1}}\cdots a_n^{i_n} \cdot d_La_j. \tag{11.2}$$

Lemma 11.1. *Let L be a field extension of $\mathbb{F}_p$ and let $a \in L$. Then $d_La = 0$ if and only if $a \in L^p$.*

Proof. If $a = a_0^p$ for some $a_0 \in L^p$, then $d_La = pa_0^{p-1} \cdot d_La_0 = 0$. Conversely, assume that $a \notin L^p$. Then a can be completed to a p-basis of L [FrJ08, p. 45, Lemma 2.7.1]. By [Lan58, Thm. 1] or [Gey13, Prop. (d) of Section 1.18], the trivial derivative of L^p extends to a derivative $D \in \mathrm{Der}(L,L)$ such that $Da = 1$. Hence, $(d_La)(D) = Da \neq 0$, so $d_La \neq 0$. □

Lemma 11.2. *Let E be a field extension of $\mathbb{F}_p$ with $[E : E^p] = p^n$, where $n \geq 2$ is an integer, let $a_1, \ldots, a_n$ be a p-basis for E over E^p, and let x be an indeterminate. For $i = 2, \ldots, n$ let*

$$y_i = a_{i-1}^{1/p}x + a_i^{1/p} \tag{11.3}$$

and let $L = E(x, y_2, \ldots, y_n)$. Then E is algebraically closed in L. Moreover, for $i = 2, \ldots, n$ we have

$$[E(x, y_2, \ldots, y_i) : E(x)] = p^{i-1}. \tag{11.4}$$

Proof. We break the proof into three parts.

PART A: *Proof of (12.4).* By assumption, $[E(a_i^{1/p}) : E] = p$ for $i = 1, \ldots, n$ and the fields $E(a_1^{1/p}), \ldots, E(a_n^{1/p})$ are linearly disjoint over E. Hence, $E' = E(a_1^{1/p}, \ldots, a_n^{1/p})$ satisfies $[E' : E] = p^n$, so $[E'(x) : E(x)] = p^n$. By (12.3), $y_i \in E'(x)$ for $i = 2, \ldots, n$, so $L \subseteq E'(x)$. Moreover, by (12.3), $y_i^p \in E(x)$, so $[E(x, y_2, \ldots, y_{i+1}) : E(x, y_2, \ldots, y_i)] \le p$ for $i = 2, \ldots, n-1$ and $[E(x, y_2) : E(x)] \le p$. Finally, by (12.3), $L(a_1^{1/p}) = E(x, y_2, \ldots, y_n, a_1^{1/p}) = E'(x)$ and $[E'(x) : L] \le p$. It follows that (12.4) holds for $i = 2, \ldots, n$ and also

$$[E'(x) : L] = p. \tag{11.5}$$

PART B: *For each $e \in E$ with $d_E e \neq 0$ we have $d_L e \neq 0$.* Otherwise, there exists $e \in E$ such that

$$d_E e \neq 0 \text{ but } d_L e = 0. \tag{11.6}$$

We denote the set of all n-tuples $\mathbf{j} = (j_1, \ldots, j_n)$ with $0 \le j_1, \ldots, j_n \le p-1$ by J. By our assumption on $a_1, \ldots, a_n$, there exist $e_\mathbf{j} \in E$, $\mathbf{j} \in J$, such that

$$e = \sum_{\mathbf{j} \in J} e_\mathbf{j}^p a_1^{j_1} a_2^{j_2} \cdots a_n^{j_n}. \tag{11.7}$$

Taking differentials of both sides of (12.7) and using formula (12.2), we get $\varepsilon_1, \ldots, \varepsilon_n \in E$ such that

$$d_E e = \sum_{i=1}^n \varepsilon_i d_E a_i \text{ and } 0 = d_L e = \sum_{i=1}^n \varepsilon_i d_L a_i. \tag{11.8}$$

Next we raise (12.3) to the pth power and get

$$a_n = y_n^p - a_{n-1} x^p, \quad a_{n-1} = y_{n-1}^p - a_{n-2} x^p, \quad \ldots, \quad a_2 = y_2^p - a_1 x^p. \tag{11.9}$$

Applying d_L on the latter equalities, we get

$$d_L a_n = -x^p d_L a_{n-1}, \quad d_L a_{n-1} = -x^p d_L a_{n-2}, \quad \ldots, \quad d_L a_2 = -x^p d_L a_1. \tag{11.10}$$

Now we successively substitute $d_L a_n, d_L a_{n-1}, \ldots, d_L a_2$ from (12.10) in the right equality of (12.8) to get a relation

$$0 = (\sum_{i=1}^n \pm \varepsilon_i x^{(i-1)p}) d_L a_1. \tag{11.11}$$

If $d_L a_1 = 0$, then by Lemma 11.1, $a_1 \in L^p$. Hence, by (12.9), $a_2, \ldots, a_n \in L^p$. It follows that $L = E'(x)$, in contrast to (12.5). It follows from this contradiction that $d_L a_1$ is a nonzero element of L. Hence, by (12.11), $\sum_{i=1}^{n} \pm\varepsilon_i x^{(i-1)p} = 0$. Therefore, $\varepsilon_1, \ldots, \varepsilon_n = 0$, so by (12.8), $d_E e = 0$, in contrast to (12.6).

Part C: *E is algebraically closed in L.* Let u be an element of L which is algebraic over E. Then u belongs to $E'(x)$ and is algebraic over E'. Hence, $u \in E'$. Since $(E')^p \subseteq E$, we have $u_0 = u^p \in E \cap L^p$. Therefore, $d_L u_0 = 0$. By Part B, $d_E u_0 = 0$. Hence, by Lemma 11.1, $u_0 \in E^p$. Consequently, $u \in E$, as claimed. □

Proposition 11.3. *None of the theories T_R and $T_{R,\mathrm{val}}$ have the amalgamation property.*

Proof. We choose a field E of positive characteristic p such that $[E : E^p] = p^3$ and let a, b, c be p-basis for E over E^p. For example, we may take a, b, c as algebraically independent elements over $\mathbb{F}_p$ and set $E = \mathbb{F}_p(a, b, c)$ [FrJ08, p. 45, proof of Lemma 2.7.2]. Then, let x be a transcendental element over E, set $y = a^{1/p}x + b^{1/p}$ and $z = b^{1/p}x + c^{1/p}$, and let $K = E(x, y)$ and $L = E(x, y, z)$. By Lemma 11.2, E is algebraically closed in both K and L.

We assume there exist a field M and embeddings $\varphi\colon K \to M$ and $\psi\colon L \to M$ that coincide on E such that $K' = \varphi(K)$ and $L' = \psi(L)$ are algebraically closed in M. We may assume without loss that $K' = K$ and φ is the identity map. Then $\varphi(E) = E$, so E is algebraically closed in M.

Raising y and z to the pth power we get

$$y^p = ax^p + b \text{ and } z^p = bx^p + c. \tag{11.12}$$

Hence, $x' = \psi(x)$, $y' = \psi(y)$, and $z' = \psi(z)$ satisfy

$$(y')^p = a(x')^p + b \text{ and } (z')^p = b(x')^p + c. \tag{11.13}$$

If $x' \neq x$, then by (12.12) and (12.13), $\left(\frac{y'-y}{x'-x}\right)^p = a$. Hence, $a^{1/p} \in M \cap \tilde{E} = E$, which is a contradiction.

If $x' = x$, then $y' = y$ and $z' = z$. Hence, $L' = L$, so $L \subseteq M$. However, by Lemma 11.2, L is a proper algebraic extension of K, so K is not algebraically closed in M, which is again a contradiction.

We conclude that an M as above does not exist, so T_R does not have the amalgamation property.

Now we choose a valuation v_E of E and extend v_E, by Chevalley, to valuations v_K and v_L of K and L, respectively. Then (E, v_E) is a $T_{R,\mathrm{val}}$-submodel of (K, v_K) and of (L, v_L). But (K, v_K) and (L, v_L) can not be embedded into a common $T_{R,\mathrm{val}}$-model (M, v_M) over (E, v_E), because K will then be algebraically closed in M, in contrast to the construction of K above. Thus, $T_{R,\mathrm{val}}$ does not have the amalgamation property. □

In the next section we rectify the deficiency of $T_{R,\mathrm{val}}$ expressed in Proposition 11.3 by adding more relation symbols to the language $\mathcal{L}_{R,\mathrm{val}}$. The new relations will ensure that if a model of the new language is contained in another model, then the underlying field of the latter model will be a regular extension of the underlying field of the former model.

12 Elimination of Quantifiers

Let (K,v) and (L,w) be valued fields and consider them as structures for the language $\mathcal{L}_{R,\mathrm{val}}(\mathrm{ring})$ (Example 10.4). If (K,v) is a substructure for (L,w), then by that example, K is algebraically closed in L. However, L is not necessarily a regular extension of K. We rectify this deficiency of $\mathcal{L}_{R,\mathrm{val}}(\mathrm{ring})$ by adding more relations to the language and prove that the theory of the non-trivial valued fields in the new language has a model completion admitting elimination of quantifiers.

Example 12.1. As in [FrJ08, p. 664, Sec. 27.3] we augment the language $\mathcal{L}_R(\mathrm{ring})$ (Example 10.4) to a language $\mathcal{L}_{R,Q}(\mathrm{ring})$ by adding n-ary relation symbols $Q_{p,n}$, one for each prime number p and each positive integer n. Let $T_{R,Q}$ be the theory of $\mathcal{L}_{R,Q}(\mathrm{ring})$ consisting of T_R (Example 10.4) together with the axioms

$$(12.1)\quad Q_{p,n}(X_1,\ldots,X_n) \leftrightarrow p=0 \wedge (\exists U_{\mathbf{i}})_{\mathbf{i}\in I}\Big(\sum_{\mathbf{i}\in I} U_{\mathbf{i}}^p X_1^{i_1}\cdots X_n^{i_n} = 0 \wedge \bigvee_{\mathbf{i}\in I} U_{\mathbf{i}} \neq 0\Big),$$

one for each pair (p,n), where I is the set of all n-tuples $(i_1,\ldots,i_n)$ of integers between 0 and $p-1$. Given a field K, we may uniquely regard K as a model of $T_{R,Q}$. Indeed, if $\mathrm{char}(K)=p>0$, we define $Q_{p,n}$ as the set of all n-tuples of elements of K satisfying the right hand side of (13.1), i.e. all n-tuples of p-dependent elements of K. In this case $Q_{p',n}$ is the empty relation for each prime number $p'\neq p$. Therefore, if $K\subseteq L$ is an extension of fields considered as models of $T_{R,Q}$, if $p=\mathrm{char}(K)>0$, and if $x_1,\ldots,x_n\in K$ are p-independent in K, then $\neg Q_{p,n}(x_1,\ldots,x_n)$ is true in K, hence in L, and therefore $x_1,\ldots,x_n$ are p-independent in L. It follows that L/K is a separable extension [FrJ08, p. 38, Lemma 2.6.1]. Considering L/K as an $\mathcal{L}_R(\mathrm{ring})$-extension, we find that K is algebraically closed in L (Example 10.4). Therefore, L/K is a regular extension [FrJ08, p. 39, Lemma 2.6.4].

Conversely, every regular extension L/K of fields is an extension of structures for the language $\mathcal{L}_{R,Q}(\mathrm{ring})$.

Definition 12.2. We say that a field K is **ω-imperfect** if either $\mathrm{char}(K)=0$ or $\mathrm{char}(K)=p>0$ and $[K^{1/p}:K]=\infty$. In the latter case, if L is a separable field extension of K, then L is linearly disjoint from $K^{1/p}$ over K, hence $[L^{1/p}:L]=\infty$, so L is also ω-imperfect.

Note that if $\mathrm{char}(K) = p > 0$, K is a model of $T_{R,Q}$, and $x_1, \ldots, x_n$ are elements of K such that $\neg Q_{p,n}(x_1, \ldots, x_n)$ holds in K, then, by (13.1), $[K(x_1^{1/p}, \ldots, x_n^{1/p}) : K] = p^n$. Hence, if each of the sentences

$$(12.2) \qquad (\exists X_1) \cdots (\exists X_n) \neg Q_{p,n}(X_1, \ldots, X_n)$$

holds in K, then K is ω-imperfect.

The following result is [FrJ08, p. 665, Thm. 27.3.1].

Proposition 12.3. *The theory $T_{R,Q}$ has a model completion $\tilde{T}_{R,Q}$ whose models are the ω-imperfect ω-free PAC fields.*

The axioms of $\tilde{T}_{R,Q}$ differ from the axioms of $\tilde{T}_R$ by the sentences (13.2) that replace the axioms for 1-imperfectness.

Adding the division symbol to the language $\mathcal{L}_{R,Q}(\mathrm{ring})$, we get a first order language $\mathcal{L}_{R,Q,\mathrm{val}}(\mathrm{ring})$ for valued fields. Then we augment the theory $T_{R,Q}$ to a theory $T_{R,Q,\mathrm{val}}$ in the language $\mathcal{L}_{R,Q,\mathrm{val}}(\mathrm{ring})$ by adding the axioms (9.3) of Section 8 to $T_{R,Q}$. We also augment $\tilde{T}_{R,Q}$ to a theory $\tilde{T}_{R,Q,\mathrm{val}}$ of the language $\mathcal{L}_{R,Q,\mathrm{val}}(\mathrm{ring})$ by adding the axioms (9.3) and (9.4) of Section 8 to $\tilde{T}_{R,Q}$. In particular, each of the models (K, v) of $\tilde{T}_{R,Q,\mathrm{val}}$ is a non-trivial valued field.

Lemma 12.4. *The theory $T_{R,Q,\mathrm{val}}$ has the amalgamation property.*

Proof. Let (L_1, v_1) and (L_2, v_2) be two models of $T_{R,Q,\mathrm{val}}$ that contain a common substructure for $T_{R,Q,\mathrm{val}}$. By Lemma 8.1(a), we may assume that this model is a common valued subfield (K, v). In particular, L_1 and L_2 are regular extensions of K. Replacing (L_2, v_2) by an isomorphic valued field extension (L_2', v_2') of (K, v), we may assume, in addition, that L_1 and L_2 are algebraically independent over K. Hence, L_1 and L_2 are linearly disjoint over K [FrJ08, p. 41, Lemma 2.6.7]. In particular, the compositum $L = L_1 L_2$ is a regular extension of K [FrJ08, p. 41, Cor. 2.6.8(b)]. By [FrJ08, p. 35, Lemma 2.5.5], L has a valuation w that extends both v_1 and v_2. Thus, (L, w) is a model of $T_{R,Q,\mathrm{val}}$ that extends both (L_1, v_1) and (L_2, v_2), as desired. □

We are now in a position to prove an analog of both Theorem 10.6 and Proposition 12.3.

Theorem 12.5. *The theory $T_{R,Q,\mathrm{val}}$ has a model completion $\tilde{T}_{R,Q,\mathrm{val}}$ whose models are the non-trivial valued fields (F, w) such that F is an ω-imperfect, ω-free PAC field. Moreover, $\tilde{T}_{R,Q,\mathrm{val}}$ admits elimination of quantifiers.*

Proof. If (K, v) is a model of $\tilde{T}_{R,Q,\mathrm{val}}$, then K is a model of $\tilde{T}_{R,Q}$. By Proposition 12.3, K is ω-imperfect, ω-free, and PAC. Conversely, if a field K is ω-imperfect,

ω-free, and PAC, then by Proposition 12.3, K is a model of $\tilde{T}_{R,Q}$. Hence, if v is a non-trivial valuation of K, then (K,v) is a model of $\tilde{T}_{R,Q,\mathrm{val}}$.

If (K,v) is a model of $T_{R,Q,\mathrm{val}}$, we replace it by a regular valued field extension, if necessary, to assume that v is non-trivial (Remark 10.5). By Proposition 12.3, K has a regular field extension L which is ω-imperfect, ω-free and PAC. Note that the regularity of L/K is forced by the language $T_{R,Q,\mathrm{val}}$, as noted in Example 12.1. By Chevalley, v extends to a valuation w of L [Lan58, p. 8, Thm. 1], so (L,w) is a model of $\tilde{T}_{R,Q,\mathrm{val}}$ that extends (K,v) in the language $\mathcal{L}_{R,Q,\mathrm{val}}(\mathrm{ring})$. Thus, $\tilde{T}_{R,Q,\mathrm{val}}$ satisfies Conditions (11.1a) and (11.1b) of Definition 10.1.

Next observe that if (E,v) and (F,w) are models of $\tilde{T}_{R,Q,\mathrm{val}}$ and $(E,v) \subseteq (F,w)$, then E is PAC and F/E is a regular extension. Moreover, both v and w are non-trivial valuations. Hence, by Theorem 9.1, (E,v) is existentially closed in (F,w) in the language $\mathcal{L}_{\mathrm{val}}(\mathrm{ring})$. Since each of the axioms R_n and $Q_{p,n}$ is equivalent to an existential formula of $\mathcal{L}_{\mathrm{val}}(\mathrm{ring})$ (Example 10.4 and Example 12.1), (E,v) is existentially closed in (F,w) in the language $\mathcal{L}_{R,Q,\mathrm{val}}(\mathrm{ring})$. Thus, $\tilde{T}_{R,Q,\mathrm{val}}$ is model complete, that is it satisfies Condition (11.1c) of Definition 10.1. Consequently, $\tilde{T}_{R,Q,\mathrm{val}}$ is a model companion of $T_{R,Q,\mathrm{val}}$. By Lemma 12.4, $\tilde{T}_{R,Q,\mathrm{val}}$ is a model completion of $T_{R,Q,\mathrm{val}}$. By Remark 10.2, $\tilde{T}_{R,Q,\mathrm{val}}$ admits elimination of quantifiers. □

References

[AlK70] A. Altman and S. Kleiman, *Introduction to Grothendieck Duality Theory,* Lecture Notes in Mathematics **146**, Springer-Verlag, Berlin, 1970.

[AtM69] M. F. Atiyah and I. G. MacDonald, *Introduction to Commutative Algebra,* Addison–Wesley, Reading, 1969.

[Ax68] J. Ax, *The elementary theory of finite fields,* Annals of Mathematics **88** (1968), 239–271.

[Bou89] N. Bourbaki, *Commutative Algebra, Chapters 1–7,* Springer, Berlin, 1989.

[Efr06] I. Efrat, *Valuations, Orderings, and Milnor K-Theory,* Mathematical surveys and monographs **124**, American Mathematical Society, Providence, 2006.

[FrJ76] M. Fried and M. Jarden, *Stable extensions with the global density property,* Canadian Journal of Mathematics **28** (1976), 774–787.

[FrJ05] M. D. Fried and M. Jarden, *Field Arithmetic, Second Edition, revised and enlarged by Moshe Jarden,* Ergebnisse der Mathematik (3) **11**, Springer, Heidelberg, 2005.

[FrJ08] M. D. Fried and M. Jarden, *Field Arithmetic, Third Edition, revised by Moshe Jarden,* Ergebnisse der Mathematik (3) **11**, Springer, Heidelberg, 2008.

[Gey13] W.-D. Geyer, *Field Theory,* manuscript, Erlangen, 2013.

[GeJ75] W.-D. Geyer and M. Jarden, *Fields with the density property,* Journal of Algebra **35** (1975), 178–189.

[GoW10] U. Görtz and T. Wedhorn, *Algebraic Geometry I,* Vieweg + Teubner Verlag, Wiesbaden 2010.

[GPR95] B. Green, F. Pop, and P. Roquette, *On Rumely's local-global principle,* Jahresbericht der Deutschen Mathematiker-Vereinigung **97** (1995), 43–74.

[Gro64] A. Grothendieck, *Éléments de Géométrie Algébrique IV, première partie,* Publications Mathématiques, IHES **20** (1964), 101–355.

[Har77] R. Hartshorne, *Algebraic Geometry,* Graduate Texts in Mathematics **52**, Springer, New York, 1977.

[JaR98] M. Jarden and A. Razon, *Rumely's local global principle for algebraic $\mathcal{PSC}$ fields over rings,* Transactions of AMS **350** (1998), 55-85.

[Jar91] M. Jarden, *Intersection of local algebraic extensions of a Hilbertian field,* in "Generators and Relations in Groups and Geometries", (A. Barlotti et al., eds), NATO ASI Series C **333** 343–405, Kluwer, Dordrecht, 1991.

[Kol07] J. Kollár, *Algebraic varieties over PAC fields,* Israel Journal of Mathematics **162** (2007), 89–101.

[Lan58] S. Lang, *Introduction to Algebraic Geometry,* Interscience Publishers, New York, 1958.

[Lan93] S. Lang, *Algebra, Third Edition,* Eddison-Wesley, Reading, 1993.

[Liu06] Q. Liu, *Algebraic Geometry and Arithmetic Curves,* Oxford Graduate Texts in Mathematics **6**, Oxford University Press, 2006.

[Mum88] D. Mumford, *The Red Book of Varieties and Schemes,* Lecture Notes in Mathematics **1358**, Springer, Berlin, 1988.

[Pre86] A. Prestel, *Einführung in die Mathematische Logik und Modelltheorie,* Vieweg, Braunschweig 1986.

[Sha77] I. R. Shafarevich, *Basic Algebraic Geometry,* Grundlehren der mathematischen Wissenschaften **213**, Springer Berlin, 1977.

[VdD78] L.P.D. v. d. Dries, *Model theory of fields*, Thesis, Utrecht, 1978.

[Wei62] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society, Providence, 1962.

Wulf-Dieter Geyer
Friedrich-Alexander-Universität Erlangen-Nürnberg
Department Mathematik
Cauerstraße 11
D-91058 Erlangen
GERMANY
geyer@mi.uni-erlangen.de

Moshe Jarden
School of Mathematics
Tel Aviv University
Ramat Aviv, Tel Aviv 69978
ISRAEL
jarden@post.tau.ac.il

Table of Content

Lecture notes

Research articles

Instructions to authors

Manuscripts should be written in English, French or German, and printed in **Latex** style. The final Latex source files, on which the publication will be based, should be prepared by the authors in a prescribed format using the macro packages available at the web page of the journal
http*://wwwen.uni.lu/recherche/fstc/mathematics_research_unit/journal.*

Each paper must include an **abstract** of not more than 200 words, which should contain a brief but informative summary of the contents of the paper. If possible, this abstract should be written in English.

Authors should include in their papers one or more **classification numbers**, following the AMS Mathematics Subject Classification. Details of this scheme can be found in each Annual Index of Mathematical Reviews or on the web at
http://www.ams.org/msc.

A few **key words** should also be indicated.

The manuscripts may be submitted either electronically or as hard copies (3 copies) to the following address:

University of Luxembourg
Campus Kirchberg
Mathematics Research Unit
6, rue Coudenhove-Kalergi
L-1359 Luxembourg
Grand-Duchy of Luxembourg
Email: carine.molitor@uni.lu or *norbert.poncin@uni.lu*

Subscription

In order to subscribe to the journal or to order previous volumes separately, please write to the address above.

University of Luxembourg
Mathematics Research Unit

Campus Kirchberg
6, rue Coudenhove-Kalergi
L-1359 Luxembourg

University of Luxembourg
Multilingual. Personalised. Connected.

ISBN 978-2-919940-09-7
ISSN 1024-1833

www.uni.lu

uni.lu
UNIVERSITÉ DU
LUXEMBOURG