Algebra & Number Theory

Volume 9

2015

No. 2

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2015 is US $255/year for the electronic version, and $440/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.267

Lifting harmonic morphisms II: Tropical curves and metrized complexes

Omid Amini, Matthew Baker, Erwan Brugallé and Joseph Rabinoff

We prove several lifting theorems for morphisms of tropical curves. We interpret the obstruction to lifting a finite harmonic morphism of augmented metric graphs to a morphism of algebraic curves as the nonvanishing of certain Hurwitz numbers, and we give various conditions under which this obstruction does vanish. In particular, we show that any finite harmonic morphism of (nonaugmented) metric graphs lifts. We also give various applications of these results. For example, we show that linear equivalence of divisors on a tropical curve C coincides with the equivalence relation generated by declaring that the fibers of every finite harmonic morphism from C to the tropical projective line are equivalent. We study liftability of metrized complexes equipped with a finite group action, and use this to classify all augmented metric graphs arising as the tropicalization of a hyperelliptic curve. We prove that there exists a d-gonal tropical curve that does not lift to a d-gonal algebraic curve.

This article is the second in a series of two.

Throughout this paper, unless explicitly stated otherwise, K denotes a complete algebraically closed nonarchimedean field with nontrivial valuation $\mathrm{val} : K \to \mathbb{R} \cup \{\infty\}$. Its valuation ring is denoted R, its maximal ideal is $\mathfrak{m}_R$, and the residue field is $k = R/\mathfrak{m}_R$. We denote the value group of K by $\Lambda = \mathrm{val}(K^\times) \subset \mathbb{R}$.

1. Introduction

This article is the second in a series of two. The first, entitled *Lifting harmonic morphisms I: metrized complexes and Berkovich skeleta*, will be cited as [ABBR1]; references of the form "Theorem I.1.1" will refer to Theorem 1.1 in [ABBR1].

We are grateful to Andrew Obus for a number of useful comments based on a careful reading of the first arXiv version of this manuscript. We thank Ye Luo for allowing us to include Example 5.13. M.B. was partially supported by NSF grant DMS-1201473. E.B. was partially supported by the ANR-09-BLAN-0039-01.

MSC2010: primary 14G22; secondary 14T05, 11G20.

Keywords: tropical lifting, skeleton, Berkovich space, analytic curve, harmonic morphism, Hurwitz number, metrized complex.

1.1. The basic motivation behind the investigations in this paper is to understand the relationship between tropical and algebraic curves. A fundamental problem along these lines is to understand which morphisms between tropical curves arise as tropicalizations[1] of morphisms of algebraic curves. More precisely, we are interested in the following question:

(Q) Given a curve X with tropicalization C, can we classify the branched covers of X in terms of (a suitable notion of) branched covers of C?

In addition to this lifting problem for morphisms of tropical curves, we also study questions such as "Which tropical curves arise as tropicalizations of hyperelliptic curves?". This naturally leads us to study group actions on tropical curves and how notions such as gonality change under tropicalization.

In this paper we will consider three different kinds of "tropical" objects which one can associate to a smooth, proper, connected algebraic curve X/K, each depending on the choice of a *triangulation* of X. Roughly speaking, a triangulation $(X, V \cup D)$ of X (with respect to a finite set of punctures $D \subset X(K)$) is a finite set V of points in the Berkovich analytification X^{an} of X whose removal partitions X^{an} into open balls and finitely many open annuli (with the punctures belonging to distinct open balls). Triangulations of (X, D) are naturally in one-to-one correspondence with semistable models $\mathfrak{X}$ of (X, D); see Section I.5. The skeleton of a triangulated curve is the dual graph of the special fiber $\mathfrak{X}_k$ of the corresponding semistable model, with infinite rays for the punctures, equipped with its canonical metric.

To any triangulated curve, one may associate the three following "tropical" objects, at each step adding some additional structure:

(1) a *metric graph* Γ: this is the skeleton of the triangulated curve $(X, V \cup D)$;

(2) an *augmented metric graph* (Γ, g), i.e., a metric graph Γ enhanced with a genus function $g : \Gamma \to \mathbb{Z}_{\geq 0}$ which is nonzero only at finitely many points: this is the above metric graph together with the function g satisfying $g(p) = 0$ for $p \notin V$ and $g(p) = \mathrm{genus}(C_p)$ for $p \in V$, where C_p is the (normalization of the) irreducible component of $\mathfrak{X}_k$ corresponding to p;

(3) a *metrized complex of curves* $\mathscr{C}$, i.e., an augmented metric graph Γ equipped with a vertex set V and a punctured algebraic curve over k of genus $g(p)$ for each point $p \in V$, with the punctures in bijection with the tangent directions

[1] In the present paper tropicalization is defined via Berkovich's theory of analytic spaces (see also [Payne 2009; Baker et al. 2012; Chambert-Loir and Ducros 2012]). Another framework for tropicalization has been proposed by Kontsevich and Soibelman [2001] and Mikhalkin (see for example [Mikhalkin 2006]), where the link between tropical geometry and complex algebraic geometry is provided by real one-parameter families of complex varieties. For some conjectural relations between the two approaches see [Kontsevich and Soibelman 2001; 2006].

to p in Γ: this is the above metric graph, together with the curves C_p for $p \in V$ and punctures given by the singular points of $\mathfrak{X}_k$.

An (augmented) metric graph or metrized complex of curves arising from a triangulated curve by the above procedure is said to be *liftable*. If $(X, V \cup D)$ and $(X, V' \cup D')$ are triangulations of the same curve X, with $D' \subset D$ and $V' \subset V$, then the corresponding metric graphs are related by a so-called *tropical modification*. Tropical modifications generate an equivalence relation on the set of (augmented) metric graphs, and an equivalence class for this relation is called an *(augmented) tropical curve*. The (augmented) *tropicalization* of a K-curve X is by definition the (augmented) tropical curve C corresponding to any triangulation of X. Tropical curves and augmented tropical curves can be thought of as "purely combinatorial" objects, whereas metrized complexes are a mixture of combinatorial objects (which one thinks of as living over the value group Λ of K) and algebrogeometric objects over the residue field k of K.

There is a natural notion of *finite harmonic morphism* between metric graphs which induces a natural notion of *tropical morphism* between tropical curves. There is a corresponding notion of tropical morphism for *augmented* tropical curves, where in addition to the harmonicity condition one imposes a "Riemann–Hurwitz condition" that the ramification divisor is effective. There is also a natural notion of finite harmonic morphism for metrized complexes of curves. Each kind of object (metric graphs, tropical curves, augmented tropical curves, metrized complexes) forms a category with respect to the corresponding notion of morphism. The construction of an (augmented) tropical curve C (resp. metrized complex $\mathscr{C}$) out of a K-curve X (resp. triangulated K-curve $(X, V \cup D)$) is functorial, in the sense that a finite morphism of curves induces in a natural way a tropical morphism $C' \to C$ (resp. a finite harmonic morphism $\mathscr{C}' \to \mathscr{C}$).

1.2. Our original question (Q) now breaks up into two separate questions:

(Q1) Which finite harmonic morphisms $\mathscr{C}' \to \mathscr{C}$ of metrized complexes can be lifted to finite morphisms of triangulated curves (with a prespecified lift X of $\mathscr{C}$)?

(Q2) Which tropical morphisms between augmented tropical curves can be lifted to finite harmonic morphisms of metrized complexes?

One can also forget the augmentation function $g : \Gamma \to \mathbb{Z}_{\geq 0}$ and ask the following variant of (Q2):

(Q2′) Which tropical morphisms between tropical curves can be lifted to finite harmonic morphisms of metrized complexes?

A consequence of the results of [ABBR1] is that the answer to question (Q1) is essentially "all", so the situation here is rather satisfactory; there is no obstruction to lifting a finite harmonic morphism $\mathscr{C}' \to \mathscr{C}$ to a branched cover of X, at least assuming everywhere tame ramification when k has characteristic $p > 0$. In particular, if $\mathrm{char}(k) = 0$ then there are no tameness issues, and we have the following result:

Theorem. *Assume* $\mathrm{char}(k) = 0$ *and let* $\varphi : \Sigma' \to \Sigma$ *be a finite harmonic morphism of* Λ*-metrized complexes of* k*-curves. Then there exists a finite morphism of triangulated punctured curves lifting* φ.

This follows immediately from Proposition I.7.15 and Theorem I.3.24. We stress that the genus and degree are automatically preserved by such lifts.

Essentially by definition, (Q2) reduces to an existence problem for ramified coverings $\varphi_{p'} : C'_{p'} \to C_p$ of a given degree with some prescribed ramification profiles. Hence the answer to (Q2) is intimately linked with the question of nonvanishing of Hurwitz numbers. See Proposition 3.3. In particular, one can easily construct tropical morphisms between augmented tropical curves which cannot be promoted to a finite harmonic morphism of metrized complexes (and hence cannot be lifted to a finite morphism of smooth proper curves over K). The simplest example of such a tropical morphism is depicted in Figure 1, and corresponds to the classical fact that, although it would not violate the Riemann–Hurwitz formula, there is no degree-four map of smooth proper connected curves over $\mathbb{C}$ having ramification profile $\{(2, 2), (2, 2), (3, 1)\}$; this is a consequence of the (easy part of the) Riemann existence theorem (see Example 3.4 below for more details).

Understanding when Hurwitz numbers vanish remains mysterious in general, so at present there is no satisfying "combinatorial" answer to question (Q2), in which we require that the genus of the objects in question be preserved by our lifts.

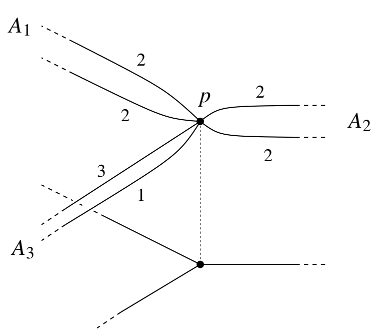


Figure 1. A tropical morphism of degree four which cannot be promoted to a degree-four morphism of metrized complexes of curves. The labels on the edges are the "expansion factors" of the harmonic morphism. See Definition I.2.4.

However, if we drop the latter condition, i.e., if we consider instead question (Q2′), we will see that the answer to (Q2′) is also "all" (see Theorem 3.11):

Theorem. *Any finite harmonic morphism $\overline{\varphi} : \Gamma' \to \Gamma$ of Λ-metric graphs is liftable if* $\mathrm{char}(k) = 0$.

1.3. Applications. We prove a number of additional results which supplement and provide applications of the above results. Some of these results are as follows.

1.3.1. Tame group actions. Let $\mathscr{C}$ be a metrized complex and let H be a finite subgroup of $\mathrm{Aut}(\mathscr{C})$. We say the action of H on $\mathscr{C}$ is *tame* if for any vertex p of Γ, the stabilizer group H_p acts freely on a dense open subset of C_p, and for any point x of C_p, the stabilizer subgroup H_x of H is cyclic of the form $\mathbb{Z}/d\mathbb{Z}$ for some integer d, with $(d, p) = 1$ if $\mathrm{char}(k) = p > 0$ (see Remark 4.6 for further explanation of this condition). It follows from Theorem I.7.4 (in its strong form, i.e., using the calculation of the automorphism group of a lift) that we can lift $\mathscr{C}$ together with a tame action of H if and only if the quotient $\mathscr{C}/H$ exists in the category of metrized complexes. We characterize when such a quotient exists in Theorem 4.9, of which the following result is a special case:

Theorem. *Suppose that the action of H is tame and has no isolated fixed points on the underlying metric graph of $\mathscr{C}$. Then there exists a smooth, proper, and geometrically connected algebraic K-curve X lifting $\mathscr{C}$ which is equipped with an action of H commuting with the tropicalization map.*

In the presence of isolated fixed points, there are additional hypotheses on the action of H to be liftable to a K-curve. As a concrete example, we prove the following characterization of all augmented tropical curves arising as the tropicalization of a hyperelliptic K-curve (see Corollary 4.15):

Theorem. *Let Γ be an augmented metric graph of genus $g \geq 2$ having no infinite vertices or degree one vertices of genus zero. Then there is a smooth proper hyperelliptic curve X over K of genus g having Γ as its minimal skeleton if and only if* (*a*) *there exists an involution s on Γ such that s fixes all the points $p \in \Gamma$ with $g(p) > 0$ and the quotient Γ/s is a metric tree, and* (*b*) *for every $p \in \Gamma$ the number of bridge edges adjacent to p is at most $2g(p) + 2$.*

1.3.2. Gonality of tropical curves. The *tropical projective line* is the augmented tropical curve $\mathbb{TP}^1$ represented by any tree with genus function identically zero. See Example 2.15. An augmented tropical curve C is called *d-gonal* if there exists a tropical morphism of degree d from C to $\mathbb{TP}^1$. By Corollary I.4.28, the gonality of an augmented tropical curve is always a lower bound for the gonality of any lift to a smooth proper curve over K. (See Remark 5.3 for a discussion of the various notions of gonality of tropical curves existing in the literature.) We prove in Section 5 that none of the lower bounds provided by tropical ranks and gonality are sharp. For example:

Theorem. (1) *There exists an augmented tropical curve C of gonality* 4 *such that the gonality of any lifting of C is at least* 5.

(2) *There exists an effective divisor D on a tropical curve C such that D has tropical rank equal to one*, *but any effective lifting of D has rank* 0.

The construction in (1) uses the vanishing of the degree-four Hurwitz number $H^4_{0,0}((2,2),(2,2),(3,1))$. In fact, we prove in Theorem 5.4 a much stronger statement: we exhibit an augmented (nonmetric) graph G such that *none* of the augmented tropical curves with G as underlying augmented graph can be lifted to a 4-gonal K-curve. This means that there is a finite graph with *stable gonality* 4 (in the sense of [Cornelissen et al. 2014]) which is not the (augmented) dual graph of any 4-gonal curve X/K.

The proof of (2) is based on our lifting results and an explicit example, due to Luo (see Example 5.13), of a degree three and rank one base-point free divisor D on a tropical curve C which does not appear as the fiber of any degree-three tropical morphism from C to $\mathbb{TP}^1$.

1.3.3. Linear equivalence of divisors. When the target curve has genus zero, we investigate in (3.16) a variant of question (Q2′) in which the genus of the source curve may be prescribed, at the cost of losing control over the degree of the morphism. As an application, we show in Theorem 4.3 that linear equivalence of divisors on a tropical curve C coincides with the equivalence relation generated by declaring that the fibers of every tropical morphism from C to the tropical projective line $\mathbb{TP}^1$ are equivalent:

Theorem. *Let Γ be a metric graph. Linear equivalence of divisors on Γ is the additive equivalence relation generated by* (*the retraction to Γ of*) *fibers of finite harmonic morphisms from a tropical modification of Γ to a metric graph of genus zero.*

1.4. Organization of the paper. The paper is organized as follows. Precise definitions of tropical modifications and tropical curves are given in Section 2, along with various kinds of morphisms between these objects. In that section we also use results from [ABBR1] to define tropicalizations of morphisms of curves, and provide a number of examples. Lifting results for (augmented) metric graphs and tropical curves are proved in Section 3. Section 4 contains applications of our lifting results. For example, lifting results for metrized complexes equipped with a finite group action are discussed in (4.5). In (4.5) we also give a complete classification of all hyperelliptic augmented tropical curves which can be realized as the minimal skeleton of a hyperelliptic curve. Finally, in Section 5 we study tropical rank and gonality and related lifting questions.

1.5. Related work. The definition of effective harmonic morphisms of augmented metric graphs that we use is the same as in [Bertrand et al. 2011]. The closely related, but slightly different, notion of an "indexed harmonic morphism" between weighted graphs was considered in [Caporaso 2014]. The indexed pseudoharmonic (resp. harmonic) morphisms in [ibid.] are closely related to harmonic (resp. effective harmonic) morphisms in our sense when the vertex sets are fixed (see Definition I.2.4), and nondegenerate morphisms in the sense of [ibid.] correspond to finite morphisms in our sense. One notable difference is that in [ibid.] only the combinatorial type of the metric graphs are fixed; the choice of positive indices in an indexed pseudoharmonic morphism determines the length of the edges in the source graph once the edge lengths in the target are fixed.

Tropical modifications and the "up-to-tropical-modification" point of view were introduced by Mikhalkin [2006].

In (5.1) we propose a definition for the stable gonality of a graph which coincides with the one used in the preprint [Cornelissen et al. 2014]. A slightly different notion of gonality for graphs was introduced in [Caporaso 2014]. We also define the gonality of an augmented tropical curve, which strikes us as a more natural and perhaps more useful notion than the stable gonality of a graph (where the lengths of the edges in the source and target metric graphs are not prespecified). We emphasize the importance of considering the dual graph of the special fiber of a semistable model of a smooth proper K-curve as an (augmented) *metric graph* and not just as a (vertex-weighted) graph. Keeping track of the natural edge lengths allows us to avoid pathological examples like Example 2.18 in [ibid.] of a 3-gonal graph which is not divisorally 3-gonal.

The question of lifting effective harmonic morphisms of metric graphs also occurs naturally (in a related but different archimedean framework) when one considers degenerating families of complex algebraic dynamical systems; see for example [DeMarco and McMullen 2008, Theorems 1.2 and 7.1], where the authors prove a lifting theorem for polynomial-like endomorphisms of (locally finite) simplicial trees which has applications to studying dynamical compactifications of the moduli space of degree-d polynomial maps. Our Theorem 3.15 was inspired by the results of DeMarco and McMullen.

2. Algebraic and tropical curves

In this section we introduce tropical curves and morphisms between them. We use the results of [ABBR1] to define functorial "intrinsic tropicalizations" of algebraic curves. We will freely use the definitions and notations in Section I.2. We reproduce some of them here for the convenience of the reader.

2.1. Metric graphs. A Λ-*metric graph* is a metric graph whose edge lengths belong to Λ. The length of an embedded segment e in a metric graph Γ is denoted $\ell(e)$. The set of tangent directions at a point p of Γ is denoted $T_p(\Gamma)$. To a harmonic morphism $\varphi : \Gamma' \to \Gamma$ of metric graphs are associated its degree $\deg\varphi$, its degree at a point $d_{p'}(\varphi)$, the degree of φ along an edge (also called the expansion factor) $d_{e'}(\varphi)$, the directional derivative of φ along a tangent direction at a vertex $d_{v'}(\varphi)$, and the induced map on tangent spaces $d\varphi(p')$ when $d_{p'}(\varphi) \neq 0$.

The group of divisors on a metric graph Γ is denoted $\mathrm{Div}(\Gamma)$. A harmonic morphism of metric graphs $\varphi : \Gamma' \to \Gamma$ gives rise to pushforward and pullback homomorphisms $\varphi_* : \mathrm{Div}(\Gamma') \to \mathrm{Div}(\Gamma)$ and $\varphi^* : \mathrm{Div}(\Gamma) \to \mathrm{Div}(\Gamma')$ defined by

$$\varphi^*(p) = \sum_{p' \mapsto p} d_{p'}(\varphi)\,(p') \quad \text{and} \quad \varphi_*(p') = (\varphi(p'))$$

and extended linearly. It is clear that for $D \in \mathrm{Div}(\Gamma)$ we have $\deg(\varphi^*(D)) = \deg\varphi \cdot \deg D$ and $\deg(\varphi_*(D)) = \deg D$.

2.2. Augmented metric graphs. An augmented metric graph Γ has a genus function $g : \Gamma \to \mathbb{Z}_{\geq 0}$. We say that Γ is *totally degenerate* provided that g is identically zero. The genus of Γ is

$$g(\Gamma) = h_1(\Gamma) + \sum_{p\in\Gamma} g(p),$$

where $h_1(\Gamma)$ is the first Betti number of Γ. If $g(\Gamma) = 0$ then we say that Γ is *rational*. The canonical divisor of an augmented metric graph Γ is

$$K_\Gamma = \sum_{p\in\Gamma} (\mathrm{val}(p) + 2g(p) - 2)\,(p).$$

The degree of K_Γ is $\deg K_\Gamma = 2g(\Gamma) - 2$.

Let $\varphi : \Gamma' \to \Gamma$ be a harmonic morphism of augmented metric graphs. The *ramification divisor* of φ is the divisor $R = \sum R_{p'}(p')$, where for $p' \in \Gamma'$,

$$R_{p'} = d_{p'}(\varphi) \cdot (2 - 2g(\varphi(p'))) - (2 - 2g(p')) - \sum_{v' \in T_{p'}(\Gamma')} (d_{v'}(\varphi) - 1).$$

We have the Riemann–Hurwitz formula

$$K_{\Gamma'} = \varphi^*(K_\Gamma) + R.$$

We say that φ is *generically étale* if R is supported on the set of infinite vertices of Γ and is *étale* if $R = 0$.

2.3. Effective harmonic morphisms. The following Riemann–Hurwitz condition will be used in formulating lifting problems for harmonic morphisms of augmented metric graphs. Given a vertex $p' \in V(\Gamma')$ with $d_{p'}(\varphi) \neq 0$, we define the *ramification degree* of φ at p' to be

$$r_{p'} = R_{p'} - \#\{v' \in T_{p'}(\Gamma') : d_{v'}(\varphi) = 0\}.$$

Clearly $r_{p'} \leq R_{p'}$, with $r_{p'} = R_{p'}$ if and only if $d_{v'}(\varphi) > 0$ for any $v' \in T_{p'}(\Gamma')$, i.e., the distinction between ramification divisors and ramification degrees only makes sense for nonfinite harmonic morphisms. Our motivation not to restrict ourselves to finite harmonic morphisms is that nonfinite harmonic morphisms show up naturally in many practical situations.

Definition 2.4. A harmonic morphism of augmented Λ-metric graphs $\varphi : \Gamma' \to \Gamma$ is said to be *effective* if $r_{p'} \geq 0$ for every finite vertex p' of Γ' with $d_{p'}(\varphi) \neq 0$.

The significance of the number $r_{p'}$ is given in Remark 2.7. In particular, only effective harmonic morphisms of augmented metric graphs have a chance to be liftable to a harmonic morphism of metrized complexes of curves, and possibly to a morphism of triangulated punctured K-curves. See Remark 2.10.

Note that a generically étale morphism of augmented metric graphs is effective.

Example 2.5. Consider the harmonic morphisms of graphs $\varphi : \Gamma' \to \Gamma$ represented in Figure 2. We use the following conventions in our pictures: black dots represent vertices of Γ' and Γ; we label an edge with its degree if and only if the degree is different from 0 and 1; we do not specify the lengths of edges of Γ' and Γ.

The morphisms in Figure 2(a,b,c) are effective provided that all the target graphs are totally degenerate. Suppose that all 1-valent vertices are infinite vertices in Figure 2(d,e), and that $g(p) = 0$ in Figure 2(e) and $g(p) = 1$ in Figure 2(e). Then $r_{p'} = 2g(p') - 1$ and $r_{p'_i} = 2g(p'_i) - 2$, so the morphism depicted in Figure 2(d) is effective if and only if $g(p') \geq 1$, and the morphism depicted in Figure 2(e) is effective if and only if both vertices p'_1 and p'_2 have genus at least one.

The morphism in Figure 1 is effective when both graphs are totally degenerate.

2.6. Metrized complexes of curves. Metrized complexes of curves and harmonic morphisms between them are defined in (I.2.16). We recall some of the definitions here. A *Λ-metrized complex of k-curves* $\mathscr{C}$ is the data of an underlying augmented Λ-metric graph Γ with a distinguished vertex set, and for each finite vertex $p \in \Gamma$ a smooth proper connected k-curve C_p of genus $g(p)$, called the *residue curve*, and an injective reduction map $\mathrm{red}_p : T_p(\Gamma) \hookrightarrow C_p(k)$. A *harmonic morphism* $\varphi : \mathscr{C}' \to \mathscr{C}$ is a harmonic morphism of underlying augmented metric graphs $\varphi : \Gamma' \to \Gamma$, taking finite vertices of Γ' to finite vertices of Γ, along with a finite morphism

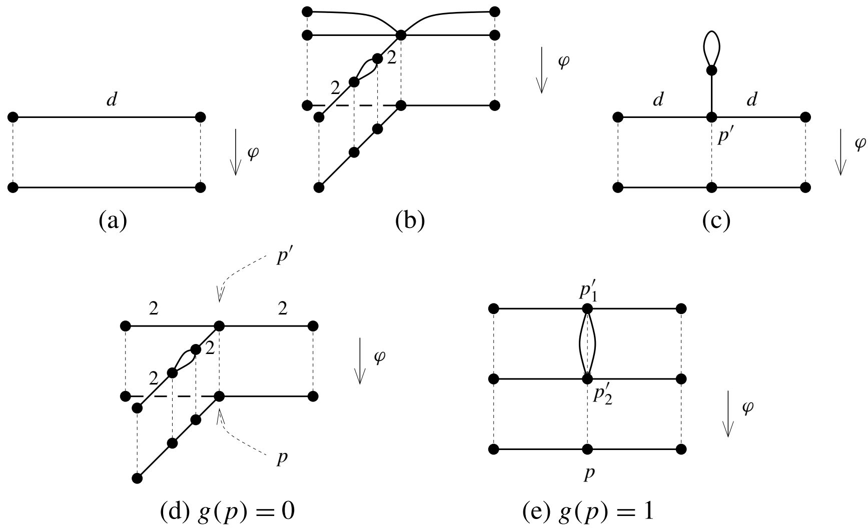


Figure 2. Examples of harmonic morphisms of augmented metric graphs.

$\varphi : C_{p'} \to C_{\varphi(p')}$ for every finite vertex p' of Γ' such that $d_{p'}(\varphi) \neq 0$, satisfying the following compatibility conditions:

(1) For every finite vertex $p' \in V(\Gamma')$ and every tangent direction $v' \in T_{p'}(\Gamma')$ with $d_{v'}(\varphi) > 0$, we have $\varphi_{p'}(\mathrm{red}_{p'}(v')) = \mathrm{red}_{\varphi(p')}(d\varphi(p')(v'))$, and the ramification degree of $\varphi_{p'}$ at $\mathrm{red}_{p'}(v')$ is equal to $d_{v'}(\varphi)$.

(2) For every finite vertex $p' \in V(\Gamma')$ with $d_{p'}(\varphi) > 0$, every tangent direction $v \in T_{\varphi(p')}(\Gamma)$, and every point $x' \in \varphi_{p'}^{-1}(\mathrm{red}_{\varphi(p')}(v)) \subset C'_{p'}(k)$, there exists $v' \in T_{p'}(\Gamma')$ such that $\mathrm{red}_{p'}(v') = x'$.

(3) For every finite vertex $p' \in V(\Gamma')$ with $d_{p'}(\varphi) > 0$ we have $d_{p'}(\varphi) = \deg \varphi_{p'}$.

Let $\varphi : \mathscr{C}' \to \mathscr{C}$ be a finite harmonic morphism of metrized complexes of curves. We say that φ is a *tame harmonic morphism* if $\varphi_{p'}$ is tamely ramified for all finite vertices $p' \in \Gamma'$. We call φ a *tame covering* if in addition it is a generically étale finite morphism of augmented metric graphs.

Remark 2.7. It follows from the Riemann–Hurwitz formula applied to the maps $\varphi_{p'} : C'_{p'} \to C_{\varphi(p')}$ that a harmonic morphism of metrized complexes of curves gives rise to an *effective* harmonic morphism of augmented metric graphs when each $\varphi_{p'}$ is a separable morphism of curves; the integer $r_{p'}$ is then the sum of ramification indices over all ramification points of $\varphi_{p'}$ not contained in $\mathrm{red}_{p'}(T_{p'}(\Gamma'))$. In particular, tame harmonic morphisms of metrized complexes of curves give rise to effective harmonic morphisms of augmented metric graphs.

2.8. Triangulated punctured curves and skeleta. Let X be a smooth, connected, proper algebraic K-curve and let $D \subset X(K)$ be a finite set of *punctures*. Recall from Definitions I.3.8 and I.3.9 that a *semistable vertex set of* (X, D) is a finite set V of type-2 points of X^{an} such that $X^{an} \setminus (V \cup D)$ is a disjoint union of open balls and finitely many once-punctured open balls and open annuli. If V is a semistable vertex set of (X, D), then $(X, V \cup D)$ is called a *triangulated punctured curve*. The semistable vertex sets of (X, D) are in bijective correspondence with the semistable R-models of (X, D).

To a triangulated punctured curve $(X, V \cup D)$ one associates a canonical Λ-metrized complex of curves $\Sigma(X, V \cup D)$ called its *skeleton*. The genus of the underlying augmented metric graph Γ is equal to the genus $g(X)$ of X. There is a canonical closed embedding $\Gamma \hookrightarrow X^{an}$ and a retraction map $\tau : X^{an} \to \Gamma$.

A *finite morphism* of triangulated punctured K-curves $\varphi : (X', V' \cup D') \to (X, V \cup D)$ consists of a finite morphism $\varphi : X' \to X$ such that $\varphi^{-1}(V) = V'$, $\varphi^{-1}(D) = D'$ and $\varphi^{-1}(\Sigma(X, V \cup D)) = \Sigma(X', V' \cup D')$ as sets. Here we restate Corollary I.4.28:

Proposition. *Let $\varphi : (X', V' \cup D') \to (X, V \cup D)$ be a finite morphism of triangulated punctured curves. Then φ naturally induces a finite harmonic morphism of Λ-metrized complexes of curves*

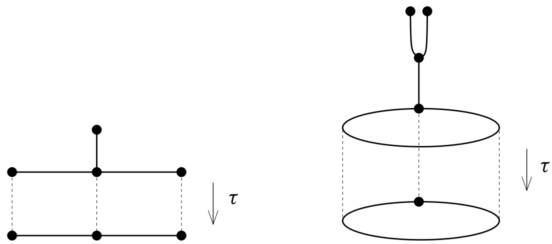
$$\Sigma(X',\ V' \cup D') \longrightarrow \Sigma(X,\ V \cup D).$$

Definition 2.9. A finite harmonic morphism $\bar{\varphi} : \Gamma' \to \Gamma$ of metrized complexes of curves (resp. augmented metric graphs, metric graphs) is said to be *liftable* provided that there exists a finite morphism of triangulated punctured K-curves $\varphi : (X', V' \cup D') \to (X, V \cup D)$ and an isomorphism of $\bar{\varphi}$ with the induced finite harmonic morphism of skeleta $\Sigma(X', V' \cup D') \to \Sigma(X, V \cup D)$ (resp. of augmented metric graphs underlying the skeleta, of metric graphs underlying the skeleta).

Remark 2.10. Among all finite harmonic morphisms of augmented metric graphs, only the effective ones have a chance to be liftable to a tame finite morphism of triangulated punctured K-curves. Since the induced morphism of skeleta is a finite harmonic morphism of metrized complexes of curves, this follows from Remark 2.7.

2.11. Tropical modifications and tropical curves. Here we introduce an equivalence relation among metric graphs; an equivalence class for this relation will be called a *tropical curve*.

Definition 2.12. An *elementary tropical modification* of a Λ-metric graph Γ_0 is a Λ-metric graph $\Gamma = [0, +\infty] \cup \Gamma_0$ obtained from Γ_0 by attaching the segment $[0, +\infty]$ to Γ_0 in such a way that $0 \in [0, +\infty]$ gets identified with a finite Λ-point $p \in \Gamma_0$. If Γ_0 is augmented, then Γ naturally inherits a genus function from Γ_0 by declaring that every point of $(0, +\infty]$ has genus zero.



(a) (b)

Figure 3. Two tropical modifications.

An (augmented) Λ-metric graph Γ obtained from an (augmented) Λ-metric graph Γ_0 by a finite sequence of elementary tropical modifications is called a *tropical modification* of Γ_0.

If Γ is a tropical modification of Γ_0, then there is a natural retraction map $\tau : \Gamma \to \Gamma_0$ which is the identity on Γ_0 and contracts each connected component of $\Gamma \setminus \Gamma_0$ to the unique point in Γ_0 lying in the topological closure of that component. The map τ is a (nonfinite) harmonic morphism of (augmented) metric graphs.

Example 2.13. We depict an elementary tropical modification in Figure 3(a), and a tropical modification which is a sequence of two elementary tropical modifications in Figure 3(b).

Tropical modifications generate an equivalence relation $\sim$ on the set of (augmented) Λ-metric graphs.

Definition 2.14. A Λ-*tropical curve* (resp. an *augmented* Λ-*tropical curve*) is an equivalence class of Λ-metric graphs (resp. augmented Λ-metric graphs) with respect to $\sim$.

In other words, a Λ-tropical curve is a Λ-metric graph considered up to tropical modifications and their inverses (and similarly for augmented tropical curves). By abuse of terminology, we will often refer to a tropical curve in terms of one of its metric graph representatives.

Example 2.15. There exists a unique rational (augmented) tropical curve, which we denote by $\mathbb{TP}^1$. Any rational (augmented) metric graph whose 1-valent vertices are all infinite is obtained by a sequence of tropical modifications from the metric graph consisting of a unique finite vertex (of genus zero).

Example 2.16. Let Γ_0 be a Λ-metric graph, $p \in \Gamma_0$ a finite Λ-point, and $l \in \Lambda \setminus \{0\}$. We can construct a new Λ-metric graph Γ by attaching the segment $[0, l]$ to Γ_0 via the identification of $0 \in [0, l]$ with p. Then Γ_0 and Γ represent the same tropical curve, since the elementary tropical modification of Γ_0 at p and the elementary tropical modification of Γ at the right-hand endpoint of $[0, l]$ are the same metric graph.

Definition 2.17. Let Γ and Γ' be representatives of Λ-tropical curves C and C', respectively, and assume we are given a harmonic morphism of Λ-metric graphs $\varphi : \Gamma' \to \Gamma$.

An *elementary tropical modification of* φ is a harmonic morphism $\varphi_1 : \Gamma'_1 \to \Gamma_1$ of Λ-metric graphs, where $\tau : \Gamma_1 \to \Gamma$ is an elementary tropical modification, $\tau' : \Gamma'_1 \to \Gamma'$ is a tropical modification, and such that $\varphi \circ \tau' = \tau \circ \varphi_1$.

A *tropical modification of* φ is a finite sequence of elementary tropical modifications of φ.

Two harmonic morphisms φ_1 and φ_2 of Λ-metric graphs are said to be *tropically equivalent* if there exists a harmonic morphism which is a tropical modification of both φ_1 and φ_2.

A *tropical morphism of tropical curves* $\varphi : C' \to C$ is a harmonic morphism of Λ-metric graphs between some representatives of C' and C, considered up to (the equivalence relation generated by) tropical equivalence, and which has a finite representative.

One makes similar definitions for morphisms of augmented tropical curves, with the additional condition that all harmonic morphisms should be effective.

Note that it might happen that two nonequivalent morphisms of augmented metric graphs represent the same tropical morphisms of nonaugmented tropical curves.

Remark 2.18. The collection of Λ-metric graphs (resp. augmented Λ-metric graphs), together with harmonic morphisms (resp. effective harmonic morphisms) between them, forms a category. Except for the condition of having a finite representative, one could try to think of tropical curves, together with tropical morphisms between them, as the localization of this category with respect to tropical modifications. However, there are some technical problems which arise when one tries to make this rigorous (at least if we demand that the localized category admit a calculus of fractions): as we will see in Example 2.19, tropical equivalence is *not* a transitive relation between morphisms of Λ-metric graphs. On the other hand, the restriction of tropical equivalence of morphisms (resp. of augmented morphisms) to the collection of finite morphisms (resp. of generically étale morphisms) *is* transitive (and hence an equivalence relation). This is one reason why we include the condition that φ has a finite representative in our definition of a morphism of tropical curves; another reason is that all morphisms of tropical curves which arise from algebraic geometry automatically satisfy this condition. See (2.21).

Example 2.19. The morphism of (totally degenerate augmented) metric graphs depicted in Figure 2(b) (resp. 4(b)) is an elementary tropical modification of the one depicted in 4(a) (resp. 2(b)).

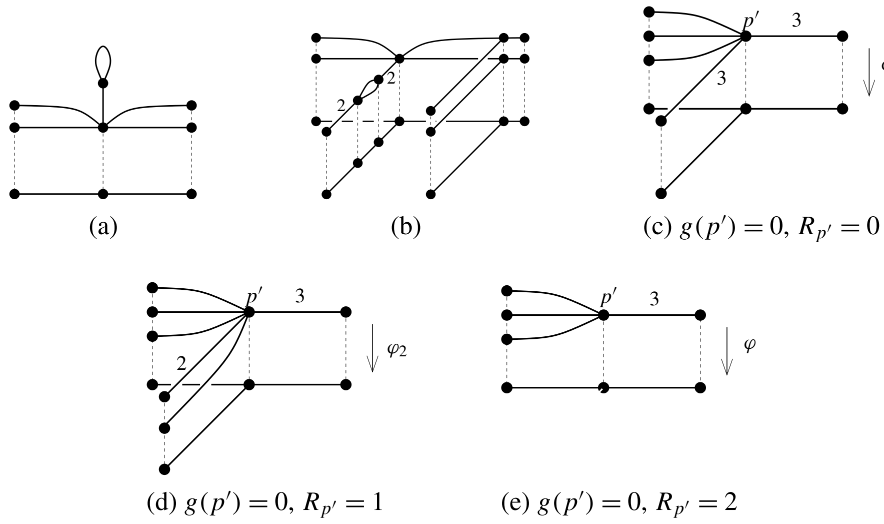


Figure 4. Figure 4(b) is an elementary tropical modification of Figure 4(a), and Figures 4(c) and (d) are elementary tropical modifications of Figure 4(e).

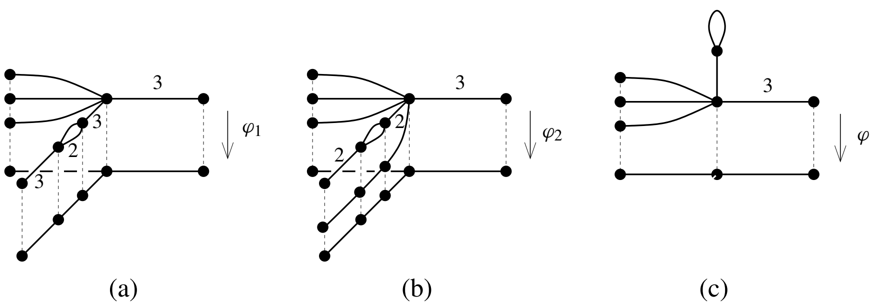


Figure 5. Figures 5(a) and (b) are elementary tropical modifications of Figure 5(c).

The tropical morphisms φ_1 and φ_2 of totally degenerate augmented tropical curves depicted in Figure 4(c) and (d) are both elementary tropical modifications of the morphism φ depicted in Figure 4(e).

The tropical morphisms φ_1 and φ_2 depicted in Figure 5(a) and (b) are both elementary tropical modifications of the morphism φ depicted in Figure 5(c).

On the other hand, the harmonic morphism $\varphi : \Gamma' \to \Gamma$ depicted in Figure 2(c) with $d = 1$ is not tropically equivalent to any finite morphism: since φ has degree one, the cycle of the source graph will be contracted to a point by any harmonic morphism of metric graphs tropically equivalent to φ. In particular, φ does not give rise to a tropical morphism.

As mentioned above, tropical equivalence is not transitive among morphisms of metric graphs (resp. of augmented metric graphs). For example, the two morphisms φ_1 and φ_2 depicted in Figure 4(c) and (d) are not tropically equivalent as augmented morphisms: since $R_{p'} = 0$ in Figure 4(c), any edge appearing in a tropical modification of φ_1 will have degree one.

Note that the preceding harmonic morphisms φ_1 and φ_2 *are* tropically equivalent as morphisms of metric graphs (i.e., forgetting the genus function). However, tropical equivalence is not transitive for tropical morphisms either, for essentially the same reason: the two tropical morphisms φ_1 and φ_2 depicted in Figure 5(a) and (b) are not tropically equivalent.

Nevertheless, the restriction of tropical equivalence of morphisms to the set of finite morphisms (resp. generically étale morphisms) is an equivalence relation. Hence a tropical morphism (resp. an augmented tropical morphism) can also be thought of as an equivalence class of finite harmonic morphisms (resp. generically étale morphisms). In particular, there exists a natural composition rule for tropical morphisms (resp. augmented tropical morphisms), turning tropical curves (resp. augmented tropical curves) equipped with tropical morphisms into a category.

Remark 2.20. In the definition of a tropical morphism of augmented tropical curves, in addition to the condition of being a harmonic morphism and the "up to tropical modifications" considerations, we imposed two rather strong conditions, namely being effective and having a finite representative. We already saw in Example 2.19 that the finiteness condition is nontrivial. The effectiveness condition is also nontrivial: for example, the harmonic morphism $\varphi : \Gamma' \to \Gamma$ of totally degenerate augmented metric graphs depicted in Figure 2(c) with $d = 2$ is not tropically equivalent to any finite effective morphism of totally degenerate augmented metric graphs. Indeed, for any tropical modification of φ which is effective, at most two edges adjacent to p' can have degree two; since Γ' already has two such edges for φ, any tropical modification of φ which is finite and effective will contract the cycle of Γ' to a point.

We refer to [Brugallé and Mikhalkin $\geq$ 2015] for a general definition of a tropical morphism $\varphi : C \to X$ from an augmented tropical curve to a nonsingular tropical variety, including Definition 2.17 as a particular case.

2.21. Algebraic and tropical curves. Restating Lemma I.3.15 and Remark I.3.16, we have:

Proposition. *Let $(X, V \cup D)$ be a triangulated punctured K-curve. Let $D' \subset X(K)$ be a finite set and let V' be a semistable vertex set of (X, D'), so $(X, V' \cup D')$ is another triangulated punctured K-curve with underlying curve X. Then the augmented metric graphs underlying $\Sigma(X, V' \cup D')$ and $\Sigma(X, V \cup D)$ represent the same tropical curve.*

This proposition implies that one can associate a canonical (augmented) tropical curve to any smooth proper connected K-curve X. This association is *functorial* by Corollary I.4.26:

Proposition. *Let $\varphi : X' \to X$ be a finite morphism of smooth proper connected K-curves, let $D \subset X(K)$ be a finite set, and let $D' = \varphi^{-1}(D)$. Then there exist semistable vertex sets V, V' of (X, D) and (X', D'), respectively, such that φ induces a finite morphism of triangulated punctured curves $\varphi : (X', V' \cup D') \to (X, V \cup D)$. In particular, φ induces a finite harmonic morphism on suitable choices of skeleta.*

Again we emphasize that a tropical morphism of tropical curves functorially induced by a finite morphism of algebraic curves is *effective* and has a *finite representative*.

Definition 2.22. We say that a tropical morphism of tropical curves $\bar{\varphi} : C' \to C$ is *liftable* provided that there exists a finite morphism of smooth proper connected K-curves $\varphi : X' \to X$ functorially inducing $\bar{\varphi}$ on skeleta in the above sense.

We will also make use in the text of the notion of tropical modifications of metrized complexes of curves.

Definition 2.23. Let $\mathscr{C}_0$ be a Λ-metrized complex of k-curves.

- A *refinement* of $\mathscr{C}_0$ is any Λ-metrized complex of k-curves $\mathscr{C}$ obtained from $\mathscr{C}_0$ by adding a finite set of Λ-points S of $\mathscr{C}_0 \setminus V(\mathscr{C}_0)$ to the set $V(\mathscr{C}_0)$ of vertices of $\mathscr{C}_0$ (see Definition I.2.17), setting $C_p = \mathbb{P}^1_k$ for all $p \in S$, and defining the map red_p by choosing any two distinct closed points of C_p.
- An *elementary tropical modification* of $\mathscr{C}_0$ is a Λ-metrized complex of k-curves $\mathscr{C}$ obtained from $\mathscr{C}_0$ by an elementary tropical modification of the underlying metric graph at a vertex p of $\mathscr{C}$, with the map red_p extended to e by choosing any closed point of C_p not in the image of the reduction map for $\mathscr{C}_0$.
- Any metrized complex of curves $\mathscr{C}$ obtained from a metrized complex of curves $\mathscr{C}_0$ by a finite sequence of refinements and elementary tropical modifications is called a *tropical modification* of $\mathscr{C}_0$.

3. Lifting harmonic morphisms of metric graphs to morphisms of metrized complexes

There is an obvious forgetful functor which sends metrized complexes of curves to (augmented) metric graphs, and harmonic morphisms of metrized complexes to harmonic morphisms of (augmented) metric graphs. A harmonic morphism of (augmented) metric graphs is said to be *liftable to a harmonic morphism of metrized complexes of k-curves* if it lies in the image of the forgetful functor.

We proved in Theorem I.7.7 that every tame covering of metrized complexes of curves can be lifted to a tame covering of algebraic curves. In this section we study the problem of lifting harmonic morphisms of (augmented) metric graphs to finite morphisms of metrized complexes (and thus to tame coverings of proper smooth curves, thanks to Proposition I.7.15).

3.1. Lifting finite augmented morphisms. Recall that k is an algebraically closed field of characteristic $p \geq 0$. A finite harmonic morphism φ of (augmented) metric graphs is called a *tame harmonic morphism* if either $p = 0$ or all the local degrees of φ along edges are prime to p. Lifting of tame harmonic morphisms of augmented metric graphs to tame harmonic morphisms of metrized complexes of k-curves is equivalent to the existence of tamely ramified covers of k-curves of given genus with some given prescribed ramification profile.

3.1.1. A partition μ of an integer d is a multiset of natural numbers $d_1, \dots, d_l \geq 1$ with $\sum_i d_i = d$. The integer l, called the length of μ, will be denoted by $l(\mu)$.

Let $g', g \geq 0$ and $d > 0$ be integers, and let $M = \{\mu_1, \dots, \mu_s\}$ be a collection of s partitions of d. Assume that the integer R defined by

$$R := d(2-2g) + 2g' - 2 - sd + \sum_{i=1}^{s} l(\mu_i) \tag{3.1.2}$$

is nonnegative. Denote by $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ the set of all tame coverings $\varphi : C' \to C$ of smooth proper curves over k, with the following properties:

(i) The curves C and C' are irreducible of genus g and g', respectively.

(ii) The degree of φ is equal to d.

(iii) The branch locus of φ contains (at least) s distinct points $x_1, \dots, x_s \in C$, and the ramification profile of φ at the points $\varphi^{-1}(x_i)$ is given by μ_i, for $1 \leq i \leq s$.

As we will explain now, the lifting problem for morphisms of augmented metric graphs to morphisms of metrized complexes over a field k reduces to the emptiness or nonemptiness of certain sets $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$. This latter problem is quite subtle, and no complete satisfactory answer is yet known (see also (3.3.1)). In some simple cases, however, one can ensure that $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ is nonempty. For example, if all the partitions μ_i are trivial (i.e., they each consist of d 1's), then $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ is nonempty. Here is another simple example.

Example 3.2. For an integer d that is prime to the characteristic p of k, the set $\mathscr{A}^d_{0,0}((d), (d))$ is nonempty since it contains the map $z \mapsto z^d$. This is in fact the only map in $\mathscr{A}^d_{0,0}((d), (d))$ up to the action of the group $\mathrm{PGL}(2, k)$ on the target curve and $\mathbb{P}^1$-isomorphisms of coverings.

3.2.1. Let $\varphi : \Gamma' \to \Gamma$ be a finite harmonic morphism of augmented metric graphs. Using the definition of a harmonic morphism, one can associate to any point p' of Γ' a collection $\mu_1(p'), \dots, \mu_s(p')$ of s partitions of the integer $d_{p'}(\varphi)$, where $s = \mathrm{val}(\varphi(p'))$, as follows: if $T_{\varphi(p)}(\Gamma) = \{v_1, \dots, v_s\}$ denotes all the tangent directions to Γ at $\varphi(p')$, then $\mu_i(p')$ is the partition of $d_{p'}(\varphi)$ which consists of the various local degrees of φ in all tangent directions $v' \in T_{p'}(\Gamma')$ mapping to v_i.

The next proposition is an immediate consequence of the various definitions involved once we note that, by Example 3.2, there are only finitely points $p' \in \Gamma'$ for which the question of nonemptiness of the sets $\mathscr{A}^{d_{p'}(\varphi)}_{g(p'),g(\varphi(p'))}$ arises. It provides a "numerical criterion" for a tame harmonic morphism of augmented metric graphs to be liftable to a tame harmonic morphism of metrized complexes of curves.

Proposition 3.3. *Let $\varphi : \Gamma' \to \Gamma$ be a tame harmonic morphism of augmented metric graphs. Then φ can be lifted to a tame harmonic morphism of metrized complexes over k if and only if for every point p' in Γ', the set*

$$\mathscr{A}^{d_{p'}(\varphi)}_{g(p'),g(\varphi(p'))}(\mu_1(p'), \dots, \mu_{\mathrm{val}(\varphi(p'))}(p'))$$

is nonempty.

3.3.1. In characteristic zero, the lifting problem for finite augmented morphisms of metric graphs can be further reduced to a vanishing question for certain Hurwitz numbers.

Fix an irreducible smooth proper curve C of genus g over k, and let $x_1, \dots, x_s$, $y_1, \dots, y_R$ be a set of distinct points on C. The Hurwitz set $\mathscr{H}^d_{g',g}(\mu_1, \dots, \mu_s)$ is the set of C-isomorphism classes of all coverings in $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ satisfying (i), (ii) and (iii) in (3.1.1) for the curve C and the points $x_1, \dots, x_s$ that we have fixed, and which in addition satisfy:

(iv) The integer R is given by (3.1.2), and for each $1 \le i \le R$, φ has a unique simple ramification point y'_i lying above y_i.

Note that, by this condition, the branch locus of φ consists precisely of the points x_i, y_j. The Hurwitz number $H^d_{g',g}(\mu_1, \dots, \mu_s)$ is defined as

$$H^d_{g',g}(\mu_1, \dots, \mu_s) := \sum_{\varphi \in \mathscr{H}^d_{g',g}(\mu_1, \dots, \mu_s)} \frac{1}{|\mathrm{Aut}_C(\varphi)|},$$

and does not depend on the choice of C and the closed points $x_1, \dots, x_s, y_1, \dots, y_R$ in C.

Example 3.4. It is known (see for example [Edmonds et al. 1984]) that

$$H^2_{g,0} = \tfrac{1}{2}, \quad H^3_{g,0}((3), \dots (3)) > 0, \quad H^4_{0,0}((2,2),(2,2),(3,1)) = 0.$$

For the reader's convenience, and since we will use it several times in the sequel, we sketch a proof of the fact that $H^4_{0,0}((2,2),(2,2),(3,1)) = 0$. By the Riemann–Hurwitz formula and the Riemann existence theorem, $H^4_{0,0}((2,2),(2,2),(3,1)) \neq 0$ if and only if there exist elements $\sigma_1, \sigma_2, \sigma_3$ in the symmetric group $\mathfrak{S}_4$ having cycle decompositions of type $(2,2)$, $(2,2)$, $(3,1)$, respectively, such that $\sigma_1\sigma_2\sigma_3 = 1$ and such that the σ_i generate a transitive subgroup of $\mathfrak{S}_4$. However, elementary group theory shows that the product $\sigma_1\sigma_2$ cannot be of type $(3,1)$ (the transitivity condition does not intervene here). For a proof which works in any characteristic, see Lemma 5.10 below.

All Hurwitz numbers can be theoretically computed, for example using the Frobenius formula (see [Lando and Zvonkin 2004, Theorem A.1.9]). Nevertheless, the problem of understanding their vanishing is wide open. The above example shows that Hurwitz numbers in degree at most three are all positive, which is not the case in degree four. Some families of (non)vanishing Hurwitz numbers are known (see Example 3.5). However, in general one has to explicitly compute a given Hurwitz number to decide if this latter vanishes or not. We refer the reader to [Edmonds et al. 1984; Pervova and Petronio 2006; 2008], along with the references therein, for an account of what is known about this subject. We will use the vanishing of $H^4_{0,0}((2,2),(2,2),(3,1))$ in Section 5 to construct a 4-gonal augmented graph (see Section 5 for the definition) which cannot be lifted to any 4-gonal proper smooth algebraic curve over K.

Example 3.5. Some partial results are known concerning the (non)vanishing of Hurwitz numbers. For example, it is known that double Hurwitz numbers (i.e., when $s = 2$) are all positive (this can be seen for example from the presentation of the cut-join equation given in [Cavalieri et al. 2010]), as well as all the Hurwitz numbers $H^d_{g',g}(\mu_1, \ldots, \mu_s)$ when $g \geq 1$ and $R \geq 0$ [Husemoller 1962; Edmonds et al. 1984]. On the other hand, it is proved in [Pervova and Petronio 2008] that

$$H^d_{0,0}\big((d-2,2),(2,\ldots,2),(\tfrac{1}{2}d+1,1,\ldots,1)\big) = 0 \quad \text{for all } d \geq 4 \text{ even.}$$

Example 3.6. As another example of nonvanishing Hurwitz numbers, one has $H^{d'}_{0,0}(\mu_1, \ldots, \mu_s, (d')) > 0$ for all integers $d' \geq 1$ when the integer R defined in (3.1.2) is zero (i.e., if the combinatorial Riemann–Hurwitz formula holds); see [Edmonds et al. 1984, Proposition 5.2] or [DeMarco and McMullen 2008, Proposition 7.2].

The nonemptiness of $\mathscr{A}^d_{g',g}(\mu_1, \ldots, \mu_s)$ can be reduced to the nonemptiness of the Hurwitz set $\mathscr{H}^d_{g',g}(\mu_1, \ldots, \mu_s)$.

Lemma 3.7. *Suppose that k has characteristic zero. Then $\mathscr{A}^d_{g',g}(\mu_1, \ldots, \mu_s)$ is nonempty if and only if $H^d_{g',g}(\mu_1, \ldots, \mu_s) \neq 0$.*

Proof. Since $\mathcal{H}^d_{g',g}(\mu_1,\dots,\mu_s)$ is a subset of $\mathcal{A}^d_{g',g}(\mu_1,\dots,\mu_s)$, obviously we only need to prove that if $\mathcal{A}^d_{g',g}(\mu_1,\dots,\mu_s)\neq\varnothing$, then the Hurwitz set is also nonempty. Let $\varphi: C'\to C$ be an element of $\mathcal{A}^d_{g',g}(\mu_1,\dots,\mu_s)$, branched over $x_i\in C$ with ramification profile μ_i for $i=1,\dots,s$, and let $z_1,\dots,z_t$ be all the other points in the branch locus of φ. Denote by ν_i the ramification profile of φ above the point z_i. Fix a closed point $\star$ of $C\setminus\{x_1,\dots,x_s,z_1,\dots,z_t\}$. The étale fundamental group $\pi_1(C\setminus\{x_1,\dots,x_s,z_1,\dots,z_t\},\star)$ is the profinite completion of the group generated by a system of generators $a_1,b_1,\dots,a_g,b_g,c_1,\dots,c_{s+t}$ satisfying the relation

$$[a_1,b_1]\cdots[a_g,b_g]c_1\cdots c_{s+t}=1,$$

where $[a,b]=aba^{-1}b^{-1}$ (see [SGA 1 1971]). In addition, the data of φ is equivalent to the data of a surjective morphism ρ from $\pi_1(C\setminus\{x_1,\dots,x_s,z_1,\dots,z_t\},\star)$ to a transitive subgroup of the symmetric group $\mathfrak{S}_d$ of degree d such that the partition μ_i (resp. ν_i) of d corresponds to the lengths of the cyclic permutations in the decomposition of $\rho(c_i)$ (resp. $\rho(c_{s+i})$) in $\mathfrak{S}_d$ into products of cycles, for $1\le i\le s$ (resp. $1\le i\le t$). By the Riemann–Hurwitz formula, we have $R=\sum_{i=1}^{t}(d-l(\nu_i))$.

Now note that each $\rho(c_{s+i})$ can be written as a product of $d-l(\nu_i)$ transpositions $\tau_i^1,\dots,\tau_i^{d-l(\nu_i)}$ in $\mathfrak{S}_d$, i.e., $\rho(c_{s+i})=\tau_i^1\dots\tau_i^{d-l(\nu_i)}$. Rename the set of R distinct points $y_1,\dots,y_R$ of $C\setminus\{x_1,\dots,x_s,\star\}$ as $z_i^1,\dots,z_i^{d-l(\nu_i)}$ for $1\le i\le t$.

The étale fundamental group $\pi_1(C\setminus\{x_1,\dots,x_s,z_1^1,\dots,z_1^{d-l(\nu_1)},\dots,z_t^{d-l(\nu_t)}\},\star)$ has, as a profinite group, a system of generators $a_1,b_1,\dots,a_g,b_g,c_1,\dots,c_s$, $c_{s+1}^1,\dots,c_{s+1}^{d-l(\nu_1)},\dots,c_{s+t}^{d-l(\nu_t)}$ verifying the relation

$$[a_1,b_1]\cdots[a_g,b_g]c_1\cdots c_s c_{s+1}^1\cdots c_{s+1}^{d-l(\nu_1)}\cdots c_{s+t}^1\cdots c_{s+t}^{d-l(\nu_t)}=1,$$

and admits a surjective morphism to $\mathfrak{S}_d$ which coincides with ρ on $a_1,b_1,\dots,a_g,b_g$, and which sends c_{s+i}^j to τ_i^j for each $1\le i\le t$ and $1\le j\le d-l(\nu_i)$. The corresponding cover $C''\to C$ obviously belongs to $\mathcal{A}^d_{g',g}(\mu_1,\dots,\mu_s)$ and in addition has simple ramification profile (2) above each y_i, i.e., it verifies condition (iv) above. This shows that $\mathcal{H}^d_{g',g}(\mu_1,\dots,\mu_s)$ is nonempty. □

Corollary 3.8. *Suppose that k has characteristic zero. Let $\varphi:\Gamma'\to\Gamma$ be a finite morphism of augmented metric graphs, and let $\mathfrak{C}$ be a metrized complex over k lifting Γ. There exists a lifting of φ to a finite harmonic morphism of metrized complexes $\mathfrak{C}'\to\mathfrak{C}$ over k (and thus to a morphism of smooth proper curves over K) if and only if*

$$\prod_{p'\in V(\Gamma')} H^{d_{p'}(\varphi)}_{g(p'),g(\varphi(p'))}(\mu_1(p'),\dots,\mu_{\mathrm{val}(\varphi(p'))})\neq 0.$$

In particular, if φ is effective and $g(p)\ge 1$ for all the points of valency at least three in Γ, then φ lifts to a finite harmonic morphism of metrized complexes over k.

Remark 3.9. If k has positive characteristic $p > d$, then $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ has the same cardinality as in characteristic zero. (This follows from [SGA 1 1971], which provides an isomorphism between the tame fundamental group in positive characteristic p and the prime-to-p part of the étale fundamental group in characteristic zero.) In particular, Lemma 3.7 also holds under the assumption that $p > d$.

3.10. Lifting finite harmonic morphisms. Now we turn to the lifting problem for finite morphisms of nonaugmented metric graphs to morphisms of metrized complexes of k-curves. In this case there are no obstructions to the existence of such a lift.

Theorem 3.11. *Let $\varphi : \Gamma' \to \Gamma$ be a tame harmonic morphism of metric graphs, and suppose that Γ is augmented. There exists an enrichment of Γ' to an augmented metric graph (Γ', g') such that $\varphi : (\Gamma', g') \to (\Gamma, g)$ lifts to a tame harmonic morphism of metrized complexes of curves over k (and thus to a morphism of smooth proper curves over K).*

Theorem 3.11 is an immediate consequence of Proposition 3.3 and the following theorem. (For the statement, we say that a partition μ of d is *tame* if either $\mathrm{char}(k) = 0$ or all the integers appearing in μ are prime to p.)

Theorem 3.12. *Let $g \geq 0$, $d \geq 2$, $s \geq 1$ be integers. Let $\mu_1, \dots, \mu_s$ be a collection of s tame partitions of d. Then there exists a sufficiently large nonnegative integer g' such that $\mathscr{A}^d_{g',g}(\mu_1, \dots, \mu_s)$ is nonempty.*

Proof. We first give a simple proof which works in characteristic zero, and more generally, in the case of a tame monodromy group. The proof in characteristic $p > 0$ is based on our lifting theorem and a deformation argument.

Suppose first that the characteristic of k is zero. By Lemma 3.7, we need to show that for large enough g' the set $\mathscr{H}^d_{g',g}(\mu_1, \dots, \mu_s)$ is nonempty.

If $g \geq 1$, for any large enough g' giving $R \geq 0$, we have $\mathscr{H}^d_{g',g}(\mu_1, \dots, \mu_s) \neq \varnothing$ [Husemoller 1962]. So suppose $g = 0$. Consider $s + R + 1$ distinct points $x_1, \dots, x_s, z_1, \dots, z_R, \star$ in C. The étale fundamental group

$$\pi_1(R) := \pi_1(C \setminus \{x_1, \dots, x_s, z_1, \dots, z_R\}, \star)$$

has, as a profinite group, a system of generators $c_1, \dots, c_s, c_{s+1}, \dots, c_{s+R}$ verifying the relation

$$c_1 \cdots c_r c_{s+1} \cdots c_{s+R} = 1.$$

It will be enough to show that for a large enough R, there exists a surjective morphism ρ from $\pi_1(R)$ to $\mathfrak{S}_d$ so that $\rho(c_{s+i})$ is a transposition for any $i = 1, \dots, R$, and that for any $i = 1, \dots, s$, the partition of d given by the lengths of the cyclic permutations in the decomposition of $\rho(c_i)$ is equal to μ_i. In this case, the genus g' of the corresponding cover C' of C in $\mathscr{H}^d_{g',0}(\mu_1, \dots, \mu_s)$ will be given by

$$g' = 1 - d + \tfrac{1}{2}\Big[sd + R - \sum_{i=1}^{s} l(\mu_i)\Big].$$

Consider an arbitrary map ρ from $\{c_1, \dots, c_s\}$ to $\mathfrak{S}_d$ verifying the ramification profile condition for $\rho(c_1), \dots, \rho(c_r)$. Choose a system of d transpositions $\tau_1, \dots, \tau_d$ generating $\mathfrak{S}_d$, and consider a set of transpositions $\tau_{d+1}, \dots, \tau_R$ such that

$$\rho(c_1) \cdots \rho(c_s)\, \tau_1 \cdots \tau_d = \tau_R \cdots \tau_{d+1}\,.$$

This proves Theorem 3.12 when k has characteristic zero.

Consider now the case of a base field k of positive characteristic $p > 0$. Note that since the prime-to-p part of the tame fundamental group has the same representation as in the case of characteristic zero, the group theoretic method we used in the previous case can be applied if the monodromy group is tame, i.e., has size prime to p. However, in general it is impossible to impose such a condition on the monodromy group. For example in the case when p divides d, the size of the monodromy group is always divisible by p.

We first describe how to reduce the proof of Theorem 3.12 to the case $s = 1$ and $g = 0$. Suppose that for each μ_i, $1 \le i \le s$, there exists a large enough g_i such that $\mathscr{A}^d_{g_i,0}(\mu_i)$ is nonempty, and consider a tame cover $\varphi_i : C_i \to \mathbb{P}^1_k$ in $\mathscr{A}_{g_i,0}(\mu_i)$ such that the ramification profile over $0 \in \mathbb{P}^1$ is given by μ_i, and choose two regular points $x_i, y_i \in \mathbb{P}^1$ (i.e., x_i, y_i are outside the branch locus of φ_i). Choose also a smooth proper curve C_0 of genus g which admits a tame cover $\varphi_0 : C'_0 \to C_0$ of degree d from a smooth proper curve C'_0 of large enough genus g'_0. (The existence of such a cover can be deduced by a similar trick as that discussed at the end of the proof below and depicted in Figure 7.) Let $y_0 \in C_0$ be a regular point of φ_0.

Let $\mathscr{C}_0$ be the metrized complex over k whose underlying metric graph is $[0, +\infty]$, with one finite vertex v_0 and one infinite vertex v_∞, equipped with the metric induced by $\mathbb{R}$, and with $C_{v_0} = C_0$ and $\mathrm{red}_{v_0}(\{v_0, v_\infty\}) = y_0$. Denote by $\mathscr{C}$ the modification of $\mathscr{C}_0$ obtained by taking a refinement at r distinct points $0 < v_1 < \cdots < v_s < \infty$, as depicted in Figure 6, and by setting $C_{v_i} = \mathbb{P}^1$ and $\mathrm{red}_{v_i}(\{v_i, v_{i-1}\}) = x_i$ and $\mathrm{red}_{v_i}(\{v_i, v_{i+1}\}) = y_i$ (here $v_{s+1} = v_\infty$), and by adding an infinite edge e_i to each v_i, and defining $\mathrm{red}_{v_i}(e_i) = 0 \in \mathbb{P}^1$. Denote by Γ the underlying metric graph of $\mathscr{C}$. See Figure 6.

Define now the metric graph $B_{s,d}$ as the chain of s banana graphs of size d: $B_{s,d}$ has $s+1$ finite vertices $u_0, \dots, u_s$ and $u'_1, \dots, u'_d$ infinite vertices adjacent to u_s such that u_i is connected to u_{i+1} with d edges of length $\ell_\Gamma(\{v_{i+1} - v_i\})$. We denote by $\widetilde{B}_{s,d}$ the tropical modification of $B_{r,d}$ at $u_1, \dots, u_s$, obtained by adding $l(\mu_i)$

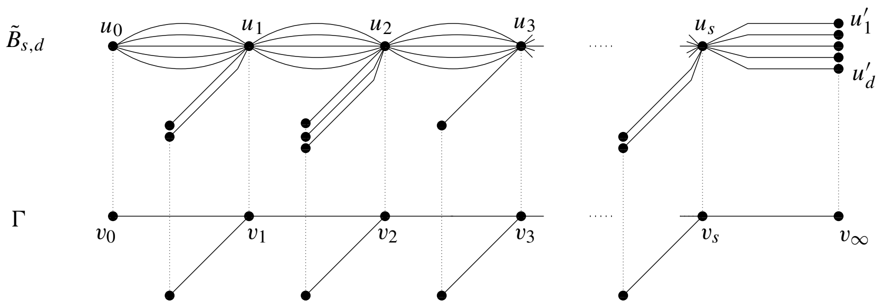


Figure 6. Construction of the graph $\widetilde{B}_{s,d}$ used in the proof of Theorem 3.12.

infinite edges to u_i. Eventually we turn $\widetilde{B}_{s,d}$ into a metrized complex $\mathscr{C}_{s,d}$ over k by setting $C_{u_i} = C_i$, and defining red_{u_i} on the d edges between u_i and u_{i+1} by a bijection to the d points in $\varphi_i^{-1}(y_i)$, red_{u_i} on the edges between u_i and u_{i-1} by a bijection to the d points in $\varphi_i^{-1}(x_i)$, and red_{u_i} on the $l(\mu_i)$ infinite edges adjacent to u_i by a bijection to the $l(\mu_i)$ points in $\varphi_i^{-1}(0)$.

Obviously, there exists a degree-d tame morphism $\varphi : \mathscr{C}_{s,d} \to \mathscr{C}$ of curve complexes over k which sends u_i to v_i, and has degrees given by integers in μ_i above the infinite edge of Γ adjacent to v_i, for $i = 1, \ldots, s$, and $\varphi_{u_i} = \varphi_i$ (see Figure 6). According to Proposition I.7.15, the map φ lifts to a tame morphism of smooth proper curves $\varphi_K : X \to X'$ over K the completion of the algebraic closure of $k[\![t]\!]$. The map φ_K has partial ramification profile $\mu_1, \ldots, \mu_s$. To deduce now the nonemptiness of $\mathscr{A}^d_{g',g}(\mu_1, \ldots, \mu_s)$, we note that there exists a subring R of K, finitely presented over k, such that the map φ_K descends to a finite morphism $\varphi_R : \mathfrak{X} \to \mathfrak{X}'$ between smooth proper curves over $\mathrm{Spec}(R)$. In addition, over a nonempty open subset U of $\mathrm{Spec}(R)$, φ_R specializes to a tame cover with the same ramification profile as φ_K. Since U contains a k-rational point, we infer the existence of a large enough g' such that $\mathscr{A}^d_{g',g}(\mu_1, \ldots, \mu_s) \neq \varnothing$.

We are thus led to consider the case where $s = 1$, $g = 0$, $\mu = (d_1, d_2, \ldots, d_l)$ with $\sum_i d_i = d$, $d_1, \ldots, d_t > 1$ and $d_{t+1} = \cdots = d_l = 1$. Figure 7 shows that, just as in the previous reduction, one can reduce to the case where $s = 1$ and $\mu_1 = \{d\}$ with $(d, p) = 1$. (Note that in Figure 7(a) the degree of the morphism at some of the middle vertices is two; Figure 7(b) is arranged so that the degrees are all odd.) But this is just nonemptiness of $\mathscr{A}_{0,0}((d))$ (see Example 3.2). $\square$

Remark 3.13. As the above proof shows, when k has characteristic zero one can get an explicit upper bound on the least positive integer g' with $\mathscr{H}^d_{g',0}(\mu_1, \ldots, \mu_s) \neq \varnothing$. Indeed, the permutation $\rho(c_1) \cdots \rho(c_s)\tau_1 \cdots \tau_d$ can be written as the product of $d + \sum_{i=1}^s (d - l(\mu_i))$ transpositions. So without loss of generality we have $R - d = d + \sum_{i=1}^s (d - l(\mu_i))$, which means that one can take g' to be $1 + \sum_{i=1}^r (d - l(\mu_i))$.

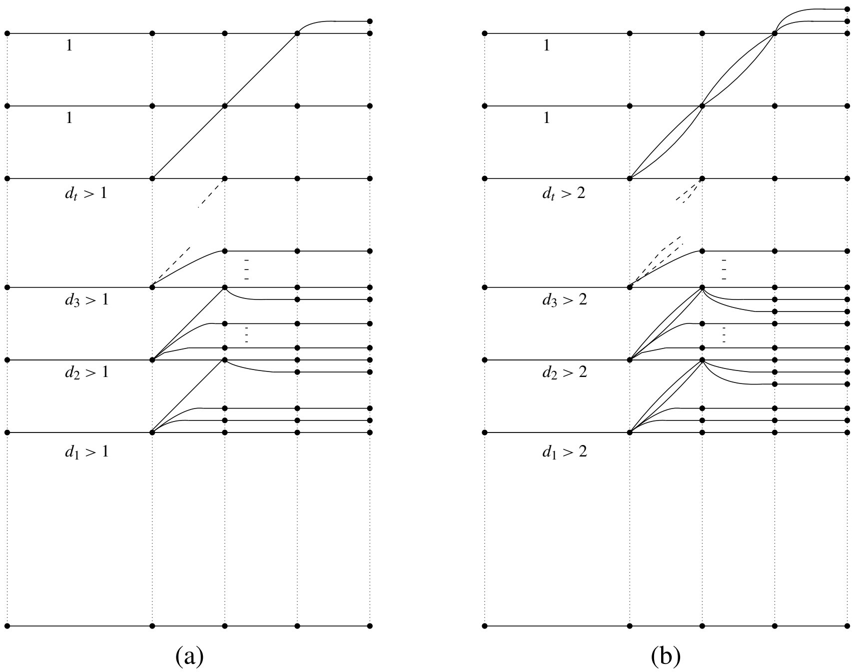


Figure 7. (a) Reduction in the case $p \neq 2$ (in this example, $d_1 = 4$, $d_2 = 4$, $d_3 = 3$ and $d_t = 2$). (b) Reduction in the case $p = 2$.

For $g \geq 1$, $\mathcal{H}_{g',g}$ is nonempty as soon as R is nonnegative, which means in this case that one can take g' to be $1 + (g-1)d + \frac{1}{2}\sum_i (d - l(\mu_i))$.

3.14. Lifting polynomial-like harmonic morphisms of trees. There is a special case of Theorem 3.12 in which one does not need to increase the genus of the source curve. To state the result, we say (following [DeMarco and McMullen 2008]) that a degree-d finite harmonic morphism $\bar{\varphi} : T' \to T$ of metric trees is *polynomial-like* if there exists an infinite vertex of T' with local degree equal to d.

Theorem 3.15. *Assume that the residue characteristic of K is zero or bigger than d. Let $\bar{\varphi} : T' \to T$ be a generically étale polynomial-like harmonic morphism of metric trees. Then there exists a degree-d polynomial map $\varphi : \mathbb{P}^1 \to \mathbb{P}^1$ over K lifting $\bar{\varphi}$.*

Proof. It suffices to prove that $\bar{\varphi}$ can be extended to a degree-d harmonic morphism of genus-zero metrized complexes of curves. By Theorem I.7.7, Proposition 3.3, and Remark 3.9, this reduces to showing that the Hurwitz numbers given by the ramification profiles around each finite vertex of T' are all nonzero. Fix an infinite vertex ∞ of T' with local degree d. Then it is easy to see that, for any such vertex v', the local degree of $\bar{\varphi}$ at v' is equal to the local degree of $\bar{\varphi}$ in the tangent direction corresponding to the unique path from v' to ∞. (This is analogous to [DeMarco and McMullen 2008, Lemma 2.3].) The result now follows from Example 3.6. $\square$

3.16. Lifting of harmonic morphisms when the base has genus zero. We now consider the special case where Γ has genus zero and present more refined lifting results in this case. As explained in (2.11), a given harmonic morphism of (augmented) metric graphs does not necessarily have a tropical modification which is finite. We present below a weakened notion of finiteness of a harmonic morphism, and prove that any harmonic morphism from an (augmented) metric graph to an (augmented) rational metric graph satisfies this weak finiteness property. We discuss in Section 4 some consequences concerning linear equivalence of divisors on metric graphs.

Definition 3.17. A harmonic morphism $\varphi : \Gamma \to T$ from an augmented metric graph Γ to a metric tree T is said to admit a *weak resolution* if there exists a tropical modification $\tau : \widetilde{\Gamma} \to \Gamma$ and an augmented harmonic morphism $\widetilde{\varphi} : \widetilde{\Gamma} \to T$ such that the restriction $\widetilde{\varphi}|_\Gamma$ is equal to φ, and some tropical modification of $\widetilde{\varphi}$ is finite.

In other words, the morphism φ has a weak resolution if it can be extended, up to increasing the degree of φ using the modification τ, to a tropical morphism $\widetilde{\varphi} : \widetilde{\Gamma} \to T$.

Example 3.18. The harmonic morphism depicted in Figure 2(c) with $d = 1$ can be weakly resolved by the harmonic morphisms depicted in Figures 4(b) and 2(b). Another example of a weak resolution is depicted in Figure 8.

Definition 3.19. Let $\varphi : \Gamma \to T$ be a harmonic morphism from a metric graph Γ to a metric tree T. A point $p \in \Gamma$ is *regular* if φ is nonconstant on all neighborhoods of p.

The *contracted set* of φ, denoted by $\mathscr{E}(\varphi)$, is the set of all nonregular points of φ. A *contracted component* of φ is a connected component of $\mathscr{E}(\varphi)$.

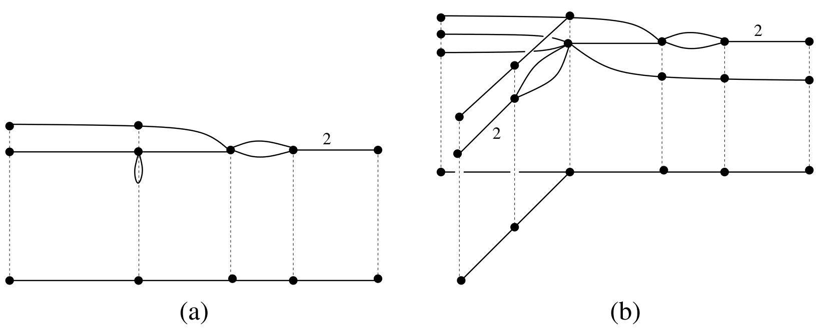


Figure 8. (a) A harmonic morphism not tropically equivalent to any finite harmonic morphism. (b) A weak resolution of the morphism in Figure 8(a).

The next proposition, together with Proposition I.7.15, allows us to conclude that any harmonic morphism from an augmented metric graph to a metric tree can be realized, up to weak resolutions, as the induced morphism on skeleta of a finite morphism of triangulated punctured curves. Recall that $\Lambda = \mathrm{val}(K^\times)$ is divisible since K is algebraically closed.

Proposition 3.20 (weak resolution of contractions). *Let $\varphi : \Gamma \to T$ be a harmonic morphism of degree d from a metric graph Γ to a metric tree T.*

(1) *There exist tropical modifications $\tau : \widetilde{\Gamma} \to \Gamma$ and $\tau' : \widetilde{T} \to T$, and a harmonic morphism of metric graphs (of degree $\widetilde{d} \geq d$) $\widetilde{\varphi} : \widetilde{\Gamma} \to \widetilde{T}$, such that $\widetilde{\varphi}|_{\Gamma \setminus \mathscr{E}(\varphi)} = \varphi$, where $\mathscr{E}(\varphi)$ is the contracted part of Γ.*

(2) *Suppose in addition that Γ is augmented, and if $p > 0$ that all the nonzero degrees of φ along tangent directions at Γ are prime to p. Then there exist tropical modifications of Γ, T, and φ as above such that $\widetilde{\varphi}$ is tame and, in addition, there exists a tame harmonic morphism of metrized complexes of k-curves with $\widetilde{\varphi}$ as the underlying finite harmonic morphism of augmented metric graphs.*

Proof. Up to tropical modifications, we may assume that all 1-valent vertices of T are infinite vertices.

The proof of (1) goes by giving an algorithm to exhibit a weak resolution of φ. Note that this algorithm does not produce the weak resolutions presented in Example 3.18, since in these cases we could find simpler ones.

Let $V(\Gamma)$ be any vertex set of Γ with no loop edge. We denote by d the degree of φ, and by α the number of nonregular vertices of φ. Given v a finite nonregular vertex of Γ, we consider the tropical modification $\tau_v : \widetilde{\Gamma}_v \to \Gamma$ such that $(\widetilde{\Gamma}_v \setminus \Gamma) \cup \{v\}$ is isomorphic to T as a metric graph. Considering all those modifications for all nonregular vertices of φ, we obtain a modification $\tau : \widetilde{\Gamma} \to \Gamma$. We can naturally extend φ to a harmonic morphism $\widetilde{\varphi} : \widetilde{\Gamma} \to T$ of degree $d + \alpha$ such that $\widetilde{\varphi}|_\Gamma = \varphi$ and all degrees of $\widetilde{\varphi}$ on edges not in Γ are equal to 1 (see Figure 9(a) in the case of the harmonic morphism depicted in Figure 2(c) with $d = 1$).

By construction, any contracted component of $\widetilde{\varphi}$ is an open edge of Γ, and this can be easily resolved. Indeed, if e is a finite contracted edge of $\widetilde{\varphi}$, we do the following (see Figure 9(b)):

- consider the tropical modification $\tau_T : \widetilde{T} \to T$ of T at $\widetilde{\varphi}(e)$; denote by e_1 the new end of $\widetilde{T}$;
- consider the composition $\tau_e : \widetilde{\Gamma}_e \to \widetilde{\Gamma}$ of two elementary tropical modifications of $\widetilde{\Gamma}$ at the middle of the edge e; denote by e_2 and e_3 the two new infinite edges of $\widetilde{\Gamma}_e$, and by e_4 and e_5 the two new finite edges of $\widetilde{\Gamma}_e$;

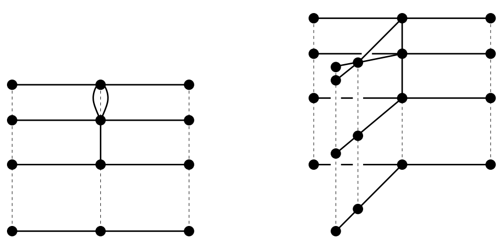
- subdivide e_1 into a finite edge e_1^0 of length equal to the lengths of e_4 and e_5, and an infinite edge e_1^∞;
- consider the morphism of metric graphs $\widetilde{\varphi}_e : \widetilde{\Gamma}_e \to \widetilde{T}$ defined by
  - $\widetilde{\varphi}_e|_{\widetilde{\Gamma}\setminus\{e_2,e_3,e_4,e_5\}} = \widetilde{\varphi}$,
  - $\widetilde{\varphi}_e(e_2) = \widetilde{\varphi}_e(e_3) = e_1^\infty$, and $\widetilde{\varphi}_e(e_4) = \widetilde{\varphi}_e(e_5) = e_1^0$,
  - $d_{e_i}(\widetilde{\varphi}_e) = 1$ for $i = 2, 3, 4, 5$;
- extend $\widetilde{\varphi}_e$ to a harmonic morphism of metric graphs $\psi_e : \Gamma' \to \widetilde{T}$, where Γ' is a modification of $\widetilde{\Gamma}_e$ at regular vertices in $\widetilde{\varphi}_e^{-1}(\widetilde{\varphi}(e))$, with all degrees of $\widetilde{\varphi}$ on edges not in $\widetilde{\Gamma}_e$ equal to 1.

We resolve in the same way a contracted infinite end of $\widetilde{\Gamma}$. By applying this process to all contracted edges of $\widetilde{\varphi}$, we end up with a finite harmonic morphism of metric graphs which is a tropical modification of $\widetilde{\varphi}$.

Note that in the proof of (1) we increased some of the local degrees by one, but we could have increased these local degrees by any amount by inserting an arbitrary number of copies of T in the construction of $\widetilde{\Gamma}$. Based on this remark, the proof of (2) now follows the same steps as the proof of (1), using in addition the following claim:

Claim. Let $g' \geq 0$ and $d, s > 0$ be integers. Let $\mu_1, \ldots, \mu_s$ be a collection of s tame partitions of d. Then there exist arbitrarily large nonnegative integers d' such that $\mathscr{A}^{d'}_{g',0}(\mu'_1, \ldots, \mu'_s)$ is nonempty, where μ'_i is the partition of d' obtained by adding a sequence of $d' - d$ numbers 1 to each partition μ_i.

Figure 10, Figure 7(a), our resolution procedure, and the argument used for the positive characteristic case of the proof of Theorem 3.12 reduce the proof of the claim to the case $s = 1$ and $\mu_1 = \{d\}$ with $(d, p) = 1$. But in this case, for any $g' \geq 0$, by the group theoretic method we used in the proof of Theorem 3.12, there exists a (tame) covering of $\mathbb{P}^1$ by a curve of genus g' having (tame) monodromy group the cyclic group $\mathbb{Z}/d\mathbb{Z}$, and with the property that the ramification profile



(a) (b)

Figure 9. The harmonic morphisms (a) $\widetilde{\varphi}$, and (b) ψ_e in the case of Figure 2(c) with $d = 1$.

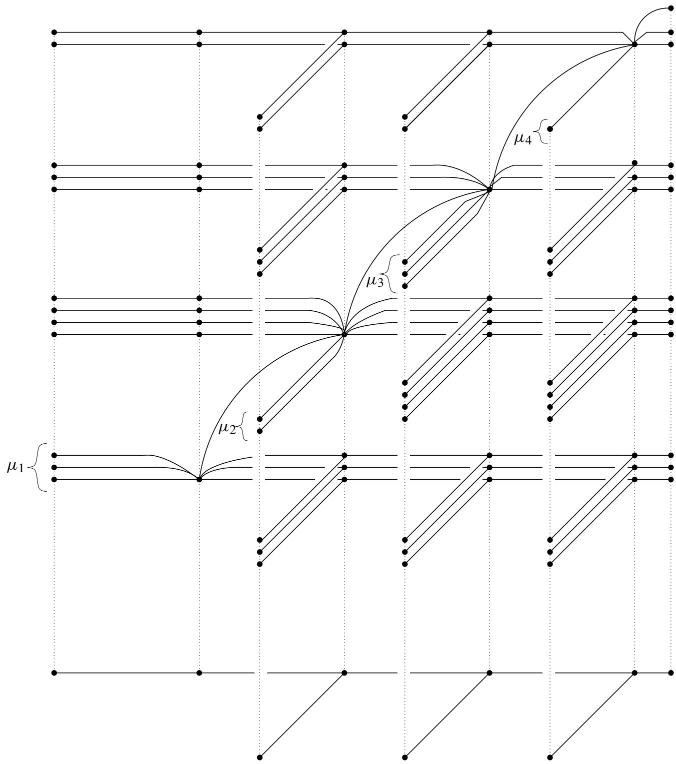


Figure 10. Reduction to the case $s = 1$ in the proof of (2) in Proposition 3.20. Degrees on (infinite) edges related to μ_i are exactly the integers appearing in μ_i. All the other degrees are one. Degrees over each infinite edge consist of a μ_i and precisely $(s-1)d$ numbers 1.

above the point 1 of $\mathbb{P}^1$ is given by $\mu = \{d\}$. This finishes the proof of the claim, and the proposition follows. □

4. Applications

4.1. Linear equivalence of divisors. A *(tropical) rational function* on a metric graph Γ is a continuous piecewise affine function $F : \Gamma \to \mathbb{R}$ with integer slopes. If F is a rational function on Γ, $\mathrm{div}(F)$ is the divisor on Γ whose coefficient at a point x of Γ is given by $\sum_{v \in T_x} d_v F$, where the sum is over all tangent directions to Γ at x and $d_v F$ is the outgoing slope of F at x in the direction v. Two divisors D and D' on a metric graph Γ are called *linearly equivalent* if there exists a rational function F on Γ such that $D - D' = \mathrm{div}(F)$, in which case we write $D \sim D'$. For a divisor D on Γ, the *complete linear system of D*, denoted $|D|$, is the set of all effective divisors E linearly equivalent to D. The *rank* of a divisor $D \in \mathrm{Div}(\Gamma)$ is defined to be

$$r_\Gamma(D) := \min_{\substack{E\,:\,E\geq 0\\ |D-E|=\varnothing}} \deg E - 1.$$

Let $\varphi : \Gamma \to T$ be a finite harmonic morphism from Γ to a metric tree T of degree d. For any point $x \in T$, the (local degree of φ at the points of the) fiber $\varphi^{-1}(x)$ defines a divisor of degree d in $\mathrm{Div}(\Gamma)$ that we denote by $D_x(\varphi)$. We have

$$D_x(\varphi) := \sum_{y\in\varphi^{-1}(x)} d_y(\varphi)(y),$$

where $d_y(\varphi)$ denotes the local degree of φ at y.

Proposition 4.2. *Let $\varphi : \Gamma \to T$ be a finite harmonic morphism of degree d from Γ to a metric tree. Then for any two points x_1 and x_2 in T, we have $D_{x_1}(\varphi) \sim D_{x_2}(\varphi)$ in Γ. Moreover, for every $x \in T$ the rank of the divisor $D_x(\varphi)$ is at least one.*

Proof. Since T is connected, we may assume that x_1 and x_2 are sufficiently close; more precisely, we can suppose that x_2 lies on the same edge as x_1 with respect to some model G for Γ. Removing the open segment (x_1, x_2) from T leaves two connected components T_{x_1} and T_{x_2} which contain x_1 and x_2, respectively. Identifying the segment $[x_1, x_2]$ with the interval $[0, \ell]$ by a linear map (where $\ell = \ell([x_1, x_2])$ denotes the length in T of the segment $[x_1, x_2]$) gives a rational function $F : \Gamma \to [0, \ell]$ by sending $\varphi^{-1}(T_{x_1})$ and $\varphi^{-1}(T_{x_2})$ to 0 and ℓ, respectively. It is easy to verify that $D_{x_1}(\varphi) - D_{x_2}(\varphi) = \mathrm{div}(F)$, which establishes the first part.

The second part follows from the first, since y belongs to the support of the divisor $D_{\varphi(y)}(\varphi) \sim D_x(\varphi)$ for all $y \in \Gamma$, which shows that $r_\Gamma(D_x(\varphi)) \geq 1$. □

By Theorem 3.11, any finite morphism $\varphi : \Gamma \to T$ can be lifted to a morphism $\varphi : X \to \mathbb{P}^1$ of smooth proper curves, possibly with $g(X) > g(\Gamma)$. This shows that any effective divisor on Γ which appears as a fiber of a finite morphism to a metric tree can be lifted to a divisor of rank at least one on a smooth proper curve *of possibly higher genus.*

We are now going to show that the (additive) equivalence relation generated by fibers of "tropicalization" of finite morphisms $X \to \mathbb{P}^1$ coincides with tropical linear equivalence of divisors. To give a more precise statement, let Γ be a metric graph with first Betti number $h_1(\Gamma)$, and consider the family of all smooth proper curves of genus $h_1(\Gamma)$ over K which admit a semistable vertex set V and a finite set of K-points D such that the metric graph $\Sigma(X, V \cup D)$ is a modification of Γ. Given such a curve X and a finite morphism $\varphi : X \to \mathbb{P}^1$, there is a corresponding finite harmonic morphism $\varphi : \Sigma(X, V \cup D) \to T$ from a modification of Γ to a metric tree T. Two effective divisors D_0 and D_1 on Γ are called *strongly effectively linearly equivalent* if there exists a morphism $\varphi : \Sigma(X, V \cup D) \to T$ as above such that $D_0 = \tau_*(D_{x_0}(\varphi))$ and $D_1 = \tau_*(D_{x_1}(\varphi))$ for two points x_0 and x_1 in T. Here

$\tau_* : \mathrm{Div}(\Sigma(X, V \cup D)) \to \mathrm{Div}(\Gamma)$ is the extension by linearity of the retraction map $\tau : \Sigma(X, V \cup D) \to \Gamma$. The equivalence relation on the abelian group $\mathrm{Div}(\Gamma)$ generated by this relation is called *effective linear equivalence* of divisors. In other words, two divisors D_0 and D_1 on Γ are effectively linearly equivalent if and only if there exists an effective divisor E on Γ such that $D_0 + E$ and $D_1 + E$ are strongly effectively linearly equivalent. This can be summarized as follows: D_0 and D_1 on Γ are effectively linearly equivalent if and only if there exists a lifting of Γ to a smooth proper curve X/K of genus $h_1(\Gamma)$, and a finite morphism $\varphi : X \to \mathbb{P}^1$ such that $\tau_*(\varphi^{-1}(0)) = D_0 + E$ and $\tau_*(\varphi^{-1}(\infty)) = D_1 + E$ for some effective divisor E, where τ_* is the natural retraction map from $\mathrm{Div}(X)$ to $\mathrm{Div}(\Gamma)$.

Theorem 4.3. *The two notions of linear equivalence and effective linear equivalence of divisors on a metric graph Γ coincide. As a consequence, linear equivalence of divisors is the additive equivalence relation generated by (the retraction to Γ of) fibers of finite harmonic morphisms from a tropical modification of Γ to a metric graph of genus zero.*

Proof. Consider two divisors D_0 and D_1 which are effectively linearly equivalent. There exists an effective divisor E and a finite harmonic morphism $\varphi : \widetilde{\Gamma} \to T$, from a tropical modification of Γ to a metric tree, such that $D_0 + E = D_{x_0}(\varphi)$ and $D_1 + E = D_{x_1}(\varphi)$ for two points $x_0, x_1 \in T$. By Proposition 4.2 we have $D_0 + E \sim D_1 + E$, which implies that D_0 and D_1 are linearly equivalent in $\widetilde{\Gamma}$, and hence in Γ.

To prove the other direction, it will be enough to show that if D is linearly equivalent to zero, then there exists an effective divisor E such that $D + E$ and E are fibers of a finite harmonic morphism φ from a modification of Γ to a metric tree T, and such that φ can be lifted to a morphism $X \to \mathbb{P}^1$.

By assumption, there exists a rational function $f : \Gamma \to \mathbb{R} \cup \{\pm\infty\}$ such that $D + \mathrm{div}(f) = 0$. We claim that there is a tropical modification $\widetilde{\Gamma}$ of Γ together with an extension of f to a (not necessarily finite) harmonic morphism $\varphi_0 : \widetilde{\Gamma} \to \mathbb{R} \cup \{\pm\infty\}$. The tropical modification $\widetilde{\Gamma}$ is obtained from Γ by choosing a vertex set which contains all the points in the support of D, adding an infinite edge to any finite vertex in Γ with $\mathrm{ord}_v(f) \neq 0$, and extending f as an affine linear function of slope $-\mathrm{ord}_v(f)$ along this infinite edge. It is clear that the resulting map φ_0 is harmonic.

Consider now the retraction map $\tau : \widetilde{\Gamma} \to \Gamma$, and note that for the two divisors $D_{\pm\infty}(\varphi_0)$, we have $\tau_*(D_{\pm\infty}(\varphi_0)) = D_\pm$, where D_+ and D_- denote the positive and negative parts of D, respectively. By Proposition 3.20, there exist tropical modifications $\overline{\Gamma}$ of $\widetilde{\Gamma}$ and T of $\mathbb{R} \cup \{\pm\infty\}$ such that φ_0 extends to a finite harmonic morphism $\varphi : \overline{\Gamma} \to T$ which can be lifted to a finite morphism $X \to \mathbb{P}^1$. If we denote (again) the retraction map $\overline{\Gamma} \to \Gamma$ by τ, then $\tau_*\big(D_{\pm\infty}(\varphi)\big) = D_\pm + E_0$ for

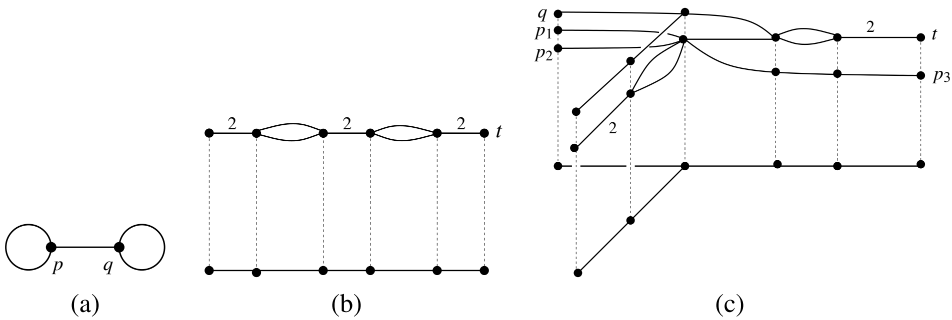


Figure 11. Illustration of the distinction between effective linear equivalence and strongly effective linear equivalence of divisors in Example 4.4. (a) $K_\Gamma = (p) + (q)$. (b) An effective lift of $2(t)$. (c) A noneffective lift of K_Γ.

some effective divisor E_0 in Γ. Setting $E = D_- + E_0$, the divisors $D + E$ and E are strongly effectively linearly equivalent, and the theorem follows. $\square$

Example 4.4. Here is an example which illustrates the distinction between the notions of (effective) linear equivalence and strongly effective linear equivalence of divisors, as introduced above.

Let Γ be the metric graph depicted in Figure 11(a), with arbitrary lengths, and $K_\Gamma = (p) + (q)$ the canonical divisor on Γ.

We claim that K_Γ is not the specialization of any effective divisor of degree two representing the canonical class of a smooth proper curve of genus two over K. More precisely, we claim that for any triangulated punctured curve $(X, V \cup D)$ over K such that $\Sigma(X, V \cup D)$ is a tropical modification of Γ, and for any effective divisor $\mathscr{D}$ in $\mathrm{Div}(X)$ with $K_\Gamma = \tau_*(\mathscr{D})$, we must have $r_X(\mathscr{D}) = 0$. (Here τ_* denotes the specialization map from $\mathrm{Div}(X)$ to $\mathrm{Div}(\Gamma)$ and $r_X(D) = \dim_K(H^0(X, \mathscr{O}(\mathscr{D}))) - 1$.) Indeed, otherwise there would exist a degree-two finite harmonic morphism $\pi : \widetilde{\Gamma} \to T$ from some tropical modification of Γ to a metric tree with the property that $\pi(p) = \pi(q)$. Restricting such a harmonic morphism to the preimage in $\widetilde{\Gamma}$ of the loop containing p would imply, by Proposition 4.2, that the divisor (p) has rank one in a genus-one metric graph, which is impossible. On the other hand, Figure 11(b) shows that the divisor $2(t) \sim (p) + (q)$ can be lifted to an effective representative of the canonical class K_X, where t is the middle point of the loop edge with vertex q. This shows that the two linearly equivalent divisors $D_0 = (p) + (q)$ and $D_1 = 2(t)$ are not strongly effectively linearly equivalent.

However, D_0 and D_1 are effectively linearly equivalent. Indeed, adding $E = (p)$ to D_0 and D_1, respectively, gives the two divisors $2(p) + (q)$ and $2(t) + (p)$ which are retractions of fibers of a degree 3 finite harmonic morphism from a tropical modification of Γ to a tree, as shown in Figure 11(c). Consequently, $D_0 + (p)$ and $D_1 + (p)$ can be lifted to linearly equivalent effective divisors on a smooth proper curve X.

Note also that Figure 11(c) shows that since $(p_1) + (p_2) + (q) - (p_3)$ can be lifted to a noneffective representative of the canonical class K_X, there exists a noneffective divisor $\mathscr{D}$ in the canonical class K_X of X such that $\tau_*(\mathscr{D}) = (p) + (q)$.

4.5. Tame actions and quotients. Let $\mathscr{C}$ be a metrized complex of k-curves, and denote by Γ the underlying metric graph of $\mathscr{C}$. An *automorphism* of $\mathscr{C}$ is a (degree-one) finite harmonic morphism of metrized complexes $h : \mathscr{C} \to \mathscr{C}$ which has an inverse. The group of automorphisms of $\mathscr{C}$ is denoted by $\mathrm{Aut}(\mathscr{C})$.

Let H be a finite subgroup of $\mathrm{Aut}(\mathscr{C})$. The action of H on $\mathscr{C}$ is *generically free* if for any vertex v of Γ, the inertia (stabilizer) group H_v acts freely on an open subset of C_v. A finite subgroup H of $\mathrm{Aut}(\mathscr{C})$ is called *tame* if the action of H on $\mathscr{C}$ is generically free and all the inertia subgroups H_x for x belonging to some C_v are cyclic of the form $\mathbb{Z}/d\mathbb{Z}$ for some positive integer d, with $(d, p) = 1$ if $p > 0$. In this case we say that the action of H on $\mathscr{C}$ is tame.

Remark 4.6. The stabilizer condition in the definition of tame actions is equivalent to requiring the cover $C_v \to C_v/H_v$ be tame, where H_v is the stabilizer of the vertex v. To see that this latter condition implies that all the stabilizers of points on C_v are cyclic, consider a uniformizer π at a point x, and consider the map $H_x \to k^\times$ which sends an element $h \in H_x$ to $h(\pi)/\pi$. This is independent of the choice of the uniformizer, and embeds H_x in the subgroup of roots of unity in $k^\times$, from which the assertion follows. The other direction is clear from the definition. Note that, more generally, one has a filtration of H_v with higher ramification groups $H_v \supseteq H_0 = H_x \supseteq H_1 \supseteq H_2 \supseteq \cdots$, the quotient H_0/H_1 is a finite cyclic group of order prime to the characteristic p, and H_i/H_{i+1} are all p-groups. In the case of tame actions, H_1 is trivial.

In this section, we characterize tame group actions H on $\mathscr{C}$ which lift to an action of H on some smooth proper curve X/K lifting $\mathscr{C}$. The main problem to consider is whether there exists a refinement $\widetilde{\mathscr{C}}$ of $\mathscr{C}$ and an extension of the action of H to $\widetilde{\mathscr{C}}$ such that the quotient $\widetilde{\mathscr{C}}/H$ can be defined, and such that the projection map $\pi : \widetilde{\mathscr{C}} \to \widetilde{\mathscr{C}}/H$ is a tame harmonic morphism. The lifting of the action of H to a smooth proper curve X as above will then be a consequence of our lifting theorem.

4.7. Let H be a tame group of automorphisms of a metrized complex $\mathscr{C}$. Let $W_H = W_H(\mathscr{C})$ be the set of all $w \in \Gamma$ lying in the middle of an edge e such that there is an element $h \in H$ having w as an isolated fixed point. Denote by H_w the stabilizer of $w \in W_H$. It is easy to see that H_w consists of all elements h of H which restrict on e either to the identity or to the symmetry with center w. In particular, if $h|_e \neq \mathrm{id}$, then h permutes the two vertices p and q adjacent to e. For $w \in W_H$, the inertia group $H_{\mathrm{red}_p(e)} = H_{\mathrm{red}_q(e)} \cong \mathbb{Z}/d_e\mathbb{Z}$ (for some integer d_e) is a normal subgroup of index two in H_w:

$$0 \longrightarrow H_{\mathrm{red}_p(e)} \longrightarrow H_w \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow 0.$$

We make the following assumption on the groups H_w:

Definition 4.8. A tame group of automorphisms H of a metrized complex $\mathscr{C}$ satisfies the *dihedral condition* provided that, for all $w \in W_H$, the stabilizer group H_w is isomorphic to the dihedral group generated by two elements σ and ζ with the relations

$$\sigma^2 = 1, \quad \zeta^d = 1 \quad \text{and} \quad \sigma\zeta\sigma = \zeta^{-1}$$

for some integer d, such that $H_{\mathrm{red}_p(e)} = \langle \zeta \rangle$.

The dihedral condition means that the above short exact sequence splits, and the action of $\mathbb{Z}/2\mathbb{Z} \cong \{\pm 1\}$ on $H_{\mathrm{red}_p(e)}$ is given by $h \to h^{\pm 1}$ for $h \in H_{\mathrm{red}_p(e)}$.

We can now formulate our main theorem on lifting tame group actions:

Theorem 4.9. *Let H be a finite group with a tame action on a metrized complex $\mathscr{C}$.*

(1) *If $W_H \neq \varnothing$, then the dihedral condition and* $\mathrm{char}(k) \neq 2$ *are the necessary and sufficient conditions for the existence of a refinement $\widetilde{\mathscr{C}}$ of $\mathscr{C}$ such that the action of H on $\mathscr{C}$ extends to a tame action on $\widetilde{\mathscr{C}}$ such that $W_H(\widetilde{\mathscr{C}}) = \varnothing$.*[2]

(2) *If $W_H = \varnothing$, then the quotient $\mathscr{C}/H$ exists in the category of metrized complexes. In addition, the action of H on $\mathscr{C}$ can be lifted to an action of H on a triangulated punctured K-curve $(X, V \cup D)$ such that $\Sigma(X, V \cup D_0) \cong \mathscr{C}$ with $D_0 \subset D$, the action of H on $X \setminus D$ is étale, and the inertia group H_x for $x \in D$ coincides with the inertia group $H_{\tau(x)}$ of the point $\tau(x) \in \Sigma(X, V \cup D_0) = \mathscr{C}$.*

Proof. Suppose that $W_H \neq \varnothing$, that the dihedral condition holds, and that $\mathrm{char}(k) \neq 2$. Fix an orientation of the edges of Γ, and for an oriented edge e, denote by p_0 and p_∞ the two vertices of Γ which form the tail and the head of e, respectively. Let w be a point lying in the middle of an oriented edge $e = (p_0, p_\infty)$ of Γ which is an isolated fixed point of some elements of H. Take the refinement $\widetilde{\mathscr{C}}$ of $\mathscr{C}$ obtained by adding all such points w to the vertex set of Γ and by setting $C_w = \mathbb{P}^1_k$, $\mathrm{red}_e(\{w, p_0\}) = 0$ and $\mathrm{red}(\{w, p_\infty\}) = \infty$. To see that the action of H on $\mathscr{C}$ extends to $\widetilde{\mathscr{C}}$, first note that one can define a generically free action of H_w on $\mathbb{P}^1_k$ (equivalently, one can embed H_w in $\mathrm{Aut}(\mathbb{P}^1_k)$) in a way compatible with the action of H_w on Γ, i.e., such that all the elements of $H_{\mathrm{red}_{p_0}(e)} = H_{\mathrm{red}_{p_\infty}(e)}$ fix the two points 0 and ∞ of $\mathbb{P}^1_k$, and such that the other elements of H_w permute the two points $0, \infty \in \mathbb{P}^1_k$. Indeed, the

[2]See [Raynaud 1999, §2.3] for a related discussion, including remarks on the situation in characteristic 2.

dihedral condition is the necessary and sufficient condition for the existence of such an action. Under this condition and upon a choice of a $d_e = |H_{\mathrm{red}_{p_0}(e)}|$-th root of unity $\zeta_{d_e} \in k$, and upon the choice of the point $1 \in \mathbb{P}^1_k$ as a fixed point of σ, the actions of the two generators σ and ζ of H_w on $\mathbb{P}^1$ are given by $\sigma(z) = 1/z$ and $\zeta(z) = \zeta_{d_e} z$, respectively.

Fix once and for all a d-th root of unity $\zeta_d \in k$ for each positive integer d (with $(d, p) = 1$ in the case $p > 0$). Given $h \in H$, we extend the action of h on $\mathscr{C}$ to an action on $\widetilde{\mathscr{C}}$ in the following way. Let $w \in W_H(\mathscr{C})$ and let e be the edge containing w, with the orientation chosen above. If $h(w) \neq w$, we define $h_w : C_w \to C_{h(w)}$ by $h_w = \mathrm{id}_{\mathbb{P}^1}$ if h is compatible with the orientations of e and $h(e)$, and we set $h_w(z) = z^{-1}$ otherwise. If $h \in H_w$, we define the action of h on C_w as above. This defines a generically free action of H on $\widetilde{\mathscr{C}}$. The inertia groups of the points 0, ∞ and ± 1 in C_w are $\mathbb{Z}/d_e\mathbb{Z}$, $\mathbb{Z}/d_e\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$, respectively. Since $p \neq 2$, this shows that the action of H on $\widetilde{\mathscr{C}}$ is tame. By construction we have $W_H(\widetilde{\mathscr{C}}) = \varnothing$.

Working backward, one recovers the necessity of the dihedral condition and $\mathrm{char}(k) \neq 2$. Indeed, any $\widetilde{\mathscr{C}}$ satisfying the conditions of the theorem must contain each $w \in W_H(\mathscr{C})$ as a vertex. Since H_w acts on $\mathbb{P}^1_k$ in the manner described above, it must be a dihedral group; since its action on C_w has stabilizers of order ± 2, we must have $\mathrm{char}(k) \neq 2$.

Now we assume that the action of H on $\mathscr{C}$ is tame and that no element of H has an isolated fixed point in the middle of an edge. We will define the quotient metrized complex $\mathscr{C}/H$. The metric graph underlying $\mathscr{C}/H$ is the quotient graph Γ/H equipped with the following metric: given an edge e of Γ of length ℓ and stabilizer H_e, we define the length of its projection in Γ/H to be $\ell \cdot |H_e|$. The projection map $\Gamma \to \Gamma/H$ is a tame finite harmonic morphism.

For any vertex p of Γ, the k-curve associated to its image in $\mathscr{C}/H$ is C_p/H_p. The marked points of C_p/H_p are the different orbits of the marked points of C_p, and are naturally in bijection with the edges of Γ/H adjacent to the projection of p. The projection map $\mathscr{C} \to \mathscr{C}/H$ is a tame harmonic morphism of metrized complexes.

To see the second part, let $\mathscr{C}'$ be the (tropical) modification of $\mathscr{C}$ obtained as follows: for any closed point $x \in C_p$ with a nontrivial inertia group and which is not the reduction $\mathrm{red}_p(e)$ of any edge e adjacent to p, consider the elementary tropical modification of $\mathscr{C}$ at x. Extend the action of H to a tame action on $\mathscr{C}'$ by defining $h_x : e_x \to e_{h(x)}$ to be affine with slope one for any such point. Let $\pi : \mathscr{C}' \to \mathscr{C}'/H$ be the projection map. Let $(X', V' \cup D')$ be a triangulated punctured K-curve such that $\mathscr{C}(X', V' \cup D') \cong \mathscr{C}'/H$. By Theorem I.7.4, the tame harmonic morphism π lifts to a morphism of triangulated punctured K-curves $(X, V \cup D) \to (X', V' \cup D')$. By Remark I.7.5, we have an injection $\iota : \mathrm{Aut}_{X'}(X) \hookrightarrow \mathrm{Aut}_{\mathscr{C}'/H}(\mathscr{C}')$. By the construction given in the proof of Theorem I.7.4, it is easy to see that every $h \in H$ lies in the image of ι, and thus $H \subset \mathrm{Aut}_{X'}(X)$. The last part follows formally from the definition of the modification $\mathscr{C}'$ and the choice of X as the lifting of $\pi : \mathscr{C}' \to \mathscr{C}'/H$. □

Remark 4.10. (Compare with Remark 4.6.) If the characteristic of k is positive, the lifting of the action of a finite group on a metrized complexes cannot be guaranteed in general without further assumptions. Indeed, even in the smooth case, i.e., where the metrized complex consists of a single vertex v and a single curve C_v, there are obstructions to the lifting [Oort 1987; Sekiguchi et al. 1989; Green and Matignon 1998; Bertin and Mézard 2000], e.g., due to the fact that the automorphism group of a smooth proper curve in positive characteristic does not respect the Hurwitz upper bound $84(g-1)$. However, Pop's proof [2014] of the Oort conjecture, based on the results of Obus and Wewer [2014], shows that in the smooth case the action can be lifted under the assumption that the stabilizers of points are all cyclic. A natural question is then to see whether our theorem can be extended by only requiring all the stabilizers of points to be cyclic (without the tameness assumption).

4.11. Characterization of liftable hyperelliptic augmented metric graphs. Let Γ be an augmented metric graph and denote by $r^{\#}$ the weighted rank function on divisors introduced in [Amini and Caporaso 2013]. Recall that this is the rank function on the nonaugmented metric graph $\Gamma^{\#}$ obtained from Γ by attaching $g(p)$ cycles, called *virtual cycles*, of (arbitrary) positive lengths to each $p \in \Gamma$ with $g(p) > 0$. We say that an augmented metric graph Γ is *hyperelliptic* if $g(\Gamma) \geq 2$ and there exists a divisor D in Γ of degree two such that $r^{\#}_{\Gamma}(D) = 1$. An augmented metric graph is said to be *minimal* if it contains neither infinite vertices nor 1-valent vertices of genus zero. Every augmented metric graph Γ is tropically equivalent to a minimal augmented metric graph Γ', which is furthermore unique if $g(\Gamma) \geq 2$. Since the tropical rank and weighted rank functions are invariant under tropical modifications, an augmented metric graph Γ is hyperelliptic if and only if Γ' is. Hence we restrict in this section to the case of minimal augmented metric graphs.

The following proposition is a refinement of a result from [Chan 2013] on vertex-weighted metric graphs (itself a strengthening of results from [Baker and Norine 2009]):

Proposition 4.12. *For a minimal augmented metric graph Γ of genus at least two, the following assertions are equivalent*:

(1) Γ *is hyperelliptic*;

(2) *There exists an involution s on Γ such that*:

   (a) *s fixes all the points $p \in \Gamma$ with $g(p) > 0$*;

   (b) *the quotient Γ/s is a metric tree*;

(3) *There exists an effective finite harmonic morphism of degree two* $\varphi : \Gamma \to T$ *from* Γ *to a metric tree* T *such that the local degree at any point* $p \in \Gamma$ *with* $g(p) > 0$ *is two.*

Furthermore, if the involution s *exists, then it is unique.*

Proof. The implication (2) $\Rightarrow$ (3) is obtained by taking $T = \Gamma/s$ and letting φ be the natural quotient map.

To prove (3) $\Rightarrow$ (1), we observe that a finite harmonic morphism of degree two $\varphi : \Gamma \to T$ with local degree two at each vertex p with $g(p) > 0$ naturally extends to an effective finite harmonic morphism of degree two from a tropical modification Γ' of $\Gamma^{\#}$ to a tropical modification T' of T as follows: Γ' is obtained by modifying $\Gamma^{\#}$ once at the midpoint of each of its virtual cycles, and T' is obtained by modifying T precisely $g(p)$ times at each point $\varphi(p)$ with $g(p) > 0$. The map φ extends uniquely to an effective finite degree-two harmonic morphism $\varphi' : \Gamma' \to T'$, since φ has local degree two at p whenever $g(p) > 0$. By Proposition 4.2, the linearly equivalent degree-two divisors $D_x(\varphi')$ have rank one in Γ' as x varies over all points of T', which shows that Γ is hyperelliptic.

It remains to prove (1) $\Rightarrow$ (2). A *bridge edge* of Γ is an edge e such that $\Gamma \setminus e$ is not connected. Let Γ' be the augmented metric graph obtained by removing all bridge edges from Γ. Since Γ is minimal, any connected component of Γ' has positive genus. In particular, the involution s, if it exists, has to restrict to an involution on each such connected component. This implies that s has to fix pointwise any bridge edge. Hence we may now assume without loss of generality that Γ has no bridge edge. In this case s has the following simple definition: for any point $p \in \Gamma$, since $r_{\Gamma^{\#}}(D) = 1$ and Γ is two-edge connected, there exists a unique point $q = s(p)$ such that $D \sim (p) + (q)$. This also proves that the involution is unique. $\square$

From now until the end of the section we assume that $\mathrm{char}(k) \neq 2$. An *involution* on a metrized complex $\mathscr{C}$ is a finite harmonic morphism $s : \mathscr{C} \to \mathscr{C}$ with $s^2 = \mathrm{id}_{\mathscr{C}}$. An involution is called *tame* if the action of the group generated by $\langle s \rangle \cong \mathbb{Z}/2\mathbb{Z}$ on $\mathscr{C}$ is tame.

If X/K is a (smooth proper) hyperelliptic curve, then the augmented metric graph Γ associated to stable model of X is hyperelliptic. Indeed, if s_X is an involution on X, then the quotient map $X \to X/s$ tropicalizes to an effective tropical morphism $\varphi : \Gamma \to T$ of degree two. The condition that φ has local degree 2 at each point p with $g(p) > 0$ comes from the fact that any nonconstant algebraic map from a positive genus curve to $\mathbb{P}^1$ has degree at least two. The next theorem, combined with Proposition 4.12, provides a complete characterization of hyperelliptic augmented metric graphs which can be realized as the skeleton of a hyperelliptic curve over K.

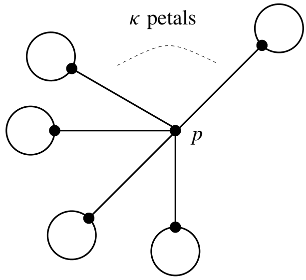


Figure 12. This graph can be lifted to a hyperelliptic curve of genus g if and only if $2g(p) \geq \kappa - 2$. See Example 4.14.

Theorem 4.13. *Let Γ be a minimal hyperelliptic augmented metric graph, and let $s : \Gamma \to \Gamma$ be the involution given by Proposition 4.12*(2). *Then the following assertions are equivalent*:

(1) *There exists a hyperelliptic smooth proper curve X over K and an involution $s_X : X \to X$ such that Γ is the minimal skeleton of X, and s coincides with the reduction of s_X to Γ.*

(2) *For every $p \in \Gamma$ we have*

$$2g(p) \geq \kappa(p) - 2,$$

where $\kappa(p)$ denotes the number of tangent directions at p which are fixed by s.

Proof. Consider the finite harmonic morphism $\pi : \Gamma \to \Gamma/s$. We note that the tangent directions at p which are fixed by s are exactly those along which π has local degree two. Thus the condition $2g(p) \geq \kappa(p) - 2$ is equivalent to the ramification index R_p being nonnegative: see Section 2. This proves (1) $\Rightarrow$ (2).

To prove (2) $\Rightarrow$ (1), we use Proposition I.7.15 and Theorem 4.9. According to these results, it suffices to prove that the involution $s : \Gamma \to \Gamma$ lifts to an involution $\bar{s} : \mathscr{C} \to \mathscr{C}$ for some metrized complex $\mathscr{C}$ with underlying augmented metric graph Γ such that $\mathscr{C}/\bar{s}$ has genus zero. The existence of such a lift follows from the observation that Hurwitz numbers of degree two are all positive (see Example 3.4). $\square$

Example 4.14. Let Γ be the augmented metric graph of genus g depicted in Figure 12 with arbitrary positive lengths. It is clearly hyperelliptic, and since the involution s restricts to the identity on each bridge edge, it fixes all tangent directions at p. Then one can lift Γ as a hyperelliptic curve of genus g if and only if $2g(p) \geq \kappa - 2$. In particular, if $g(p) = 0$ then this metric graph cannot be realized as the skeleton of a hyperelliptic curve as soon as $\kappa \geq 3$.

Since the hyperelliptic involution is unique for both curves and minimal augmented metric graphs, and since the tangent directions fixed by the hyperelliptic involution on an augmented metric graph correspond to bridge edges, we can reformulate Theorem 4.13 as follows, obtaining a metric strengthening of [Caporaso 2014, Theorem 4.8]:

Corollary 4.15. *Let Γ be a minimal augmented metric graph of genus $g \geq 2$. Then there is a smooth proper hyperelliptic curve X over K of genus g having Γ as its minimal skeleton if and only if Γ is hyperelliptic and for every $p \in \Gamma$ the number of bridge edges adjacent to p is at most $2g(p)+2$.*

5. Gonality and rank

A fundamental (if vaguely formulated) question in tropical geometry is the following: if X is an algebraic variety and $\mathbb{T}X$ is a tropicalization of X (whatever it means), which properties of X can be read off from $\mathbb{T}X$? In this section, we discuss more precisely (for curves) the relation between the classical and tropical notions of gonality, and of the rank of a divisor. It is not difficult to prove that the gonality of a tropical curve (resp. the rank of a tropical divisor) provides a lower bound for the gonality (resp. an upper bound for the rank) of any lift (this is a consequence, for example, of Corollary I.4.28). Here we address the question of sharpness for these inequalities:

(1) Can a d-gonal (augmented or nonaugmented) tropical curve C always be lifted to a d-gonal algebraic curve?

(2) Can a divisor D on an (augmented or nonaugmented) tropical curve C always be lifted to divisor of the same rank on an algebraic curve lifting C?

It follows immediately from Theorem 3.11 that the answer to Question (1) is *yes* if C is not augmented, i.e., if we are allowed to arbitrarily increase the genus of finitely many points in C. On the other hand, we prove in this section that the answer to Question (1) in the case when C is augmented, and the answer to Question (2) in both cases, is *no*.

We refer to [Baker and Norine 2007; Mikhalkin and Zharkov 2008; Amini and Caporaso 2013; Amini and Baker 2014] for the basic definitions concerning ranks of divisors on metric graphs, augmented metric graphs, and metrized complexes of curves.

5.1. Gonality of augmented graphs versus gonality of algebraic curves. An augmented tropical curve C is said to have an augmented (nonmetric) graph G as its *combinatorial type* if C admits a representative whose underlying augmented graph is G. Given an augmented graph G, we denote by $\mathcal{M}(G)$ the set of all augmented metric graphs which define a tropical curve C with combinatorial type G. In other words, $\mathcal{M}(G)$ consists of all augmented metric graphs which can be obtained by a finite sequence of tropical modifications (and their inverses) from an augmented metric graph Γ with underlying augmented graph G. When no confusion is possible, we identify an (augmented) tropical curve with any of its representatives as an (augmented) metric graph: in what follows, we deliberately write $C \in \mathcal{M}(G)$ for a tropical curve C with combinatorial type G. Note that the spaces $\mathcal{M}(G)$ appear naturally in the stratification of the moduli space of tropical curves of genus $g(G)$; see for example [Caporaso 2014].

Definition 5.2. An augmented tropical curve C is called *d-gonal* if there exists a tropical morphism $C \to \mathbb{TP}^1$ of degree d.

An augmented graph G is called *stably d-gonal* if there exists a d-gonal augmented tropical curve C whose combinatorial type is G.

In other words, an augmented graph G is stably d-gonal if and only if there is an augmented metric graph $\Gamma \in \mathcal{M}(G)$ which admits an effective finite harmonic morphism of degree d to a metric tree.

Remark 5.3. Our definition of the stable gonality of a graph is equivalent to the one given in [Cornelissen et al. 2014]. See Appendix A of that reference for a detailed discussion of the relationship between stable gonality and other tropical or graph-theoretic notions of gonality in the literature, e.g., Caporaso's definition [2014].

In this section we prove the following theorem, which is an immediate consequence of Corollary I.4.28 and Propositions 5.5 and 5.6 below.

Theorem 5.4. *There exists an augmented stably d-gonal graph G such that for any augmented metric graph $\Gamma \in \mathcal{M}(G)$ and any smooth proper connected K-curve X lifting Γ, the gonality of X is strictly larger than d.*

Let G_{27} be the graph depicted in Figure 13, which we promote to a totally degenerate augmented graph by taking the genus function to be identically zero. Note that $g(G_{27}) = 27$, and that $G_{27} \setminus \{p\}$ has three connected components, which we denote by A_1, A_2 and A_3 according to Figure 13.

Given an element $\Gamma \in \mathcal{M}(G_{27})$ and a tropical morphism $\varphi : C \to \mathbb{TP}^1$ from the tropical curve represented by Γ to $\mathbb{TP}^1$, we denote by φ_i the restriction of φ to (the metric subgraph in Γ which corresponds to) A_i, and by φ_p the restriction of φ to a small neighborhood of the point p.

Proposition 5.5. *The graph G_{27} depicted in Figure 13 is stably 4-gonal.*

Proof. We need to show the existence of a suitable tropical curve C with combinatorial type G_{27} which admits a tropical morphism of degree four to $\mathbb{TP}^1$. For a suitable choice of edge lengths on G_{27}, we get an element $\Gamma \in \mathcal{M}(G_{27})$ such that there exists a harmonic morphism from Γ to a star-shaped genus-zero augmented metric graph with three infinite edges, which has restrictions $\varphi_1, \varphi_2, \varphi_3, \varphi_v$ to A_1, A_2, A_3 and a small neighborhood of p, respectively, given as in Figure 14. We claim that φ induces a tropical morphism, i.e., that there exists a tropical modification of φ which is finite and effective.

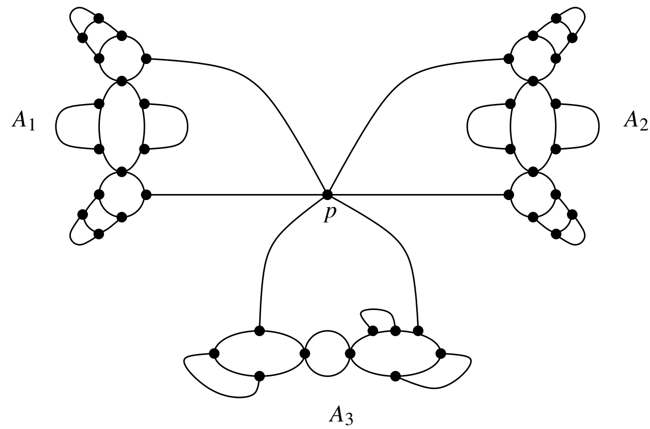


Figure 13. The graph G_{27}.

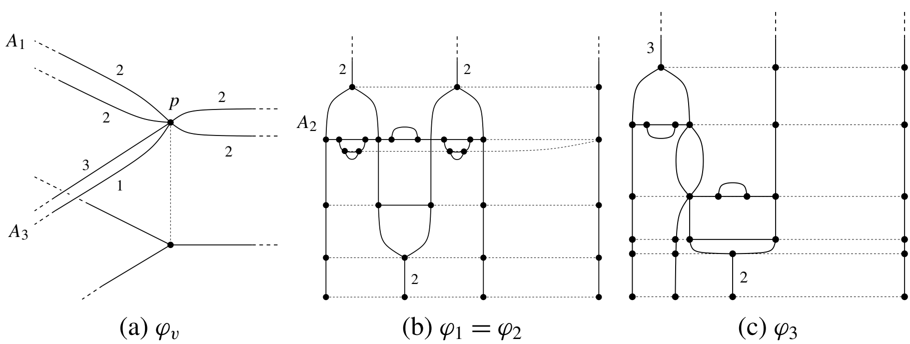


Figure 14. A tropical morphism of degree four.

Note that each of the morphisms φ_1 and φ_2 contains a fiber of genus five, while the morphism φ_3 has two different fibers of genus one. All the other fibers of φ_1, φ_2 and φ_3 are either finite or connected of genus zero. We depict in Figure 15 a few patterns which show how to resolve contractions of φ, turning φ into an augmented tropical morphism. Figure 15(a) shows how to resolve a contracted segment (resolving contracted fibers of genus zero). Figure 15(b) shows how to resolve a contracted cycle (resolving the contracted cycles in φ_3 and the middle contracted cycle in φ_1 and φ_2): the idea is to reduce to the case of a contracted segment, in which case one can use the resolution given in Figure 15(a) to finish. And finally, Figure 15(c) shows how to resolve the two contracted double-cycles in φ_1 and φ_2 by reducing to the case already treated in Figure 15(b). Note that performing these tropical modifications imposes conditions on the length of the contracted edges in Γ, e.g., in Figure 15(b), the two edges adjacent to the contracted cycle should have the same length. Nevertheless, by appropriately choosing the

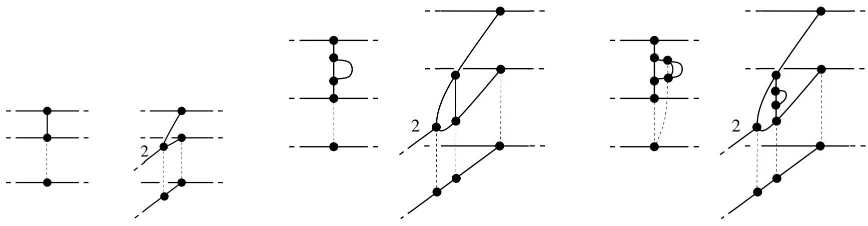


Figure 15. Patterns to resolve contractions in the harmonic morphisms φ_1, φ_2 and φ_3. (a) Resolution in one step. (b) Resolution in two steps (combined with case (a)). (c) Resolution in three steps (combined with case (b)).

edge lengths, we get the existence of a metric graph $\Gamma \in \mathcal{M}(G_{27})$ which admits a finite morphism of degree four to a metric tree. It is easily seen that this morphism is effective; thus we get a tropical curve C with combinatorial type G_{27} and a tropical morphism of degree four to $\mathbb{TP}^1$, finishing the proof of the proposition. $\square$

To conclude the proof of Theorem 5.4, we now show the following:

Proposition 5.6. *There is no metrized complex of k-curves with underlying augmented metric graph in $\mathcal{M}(G_{27})$ and admitting a finite morphism of degree four to a metrized complex of k-curves of genus zero.*

We emphasize that the statement holds for any (algebraically closed) field k. The proof of Proposition 5.6 relies on some technical lemmas that we state now.

We first recall a formula given in [Amini and Baker 2014] for the rank of divisors on a metric graph $\Gamma = \Gamma_1 \vee \Gamma_2$ which is obtained as a wedge sum of two metric graphs Γ_1 and Γ_2. Recall that, given two metric graphs Γ_1 and Γ_2 and distinguished points $t_1 \in \Gamma_1$ and $t_2 \in \Gamma_2$, the *wedge sum* or *direct sum* of (Γ_i, t_i), denoted $\Gamma = \Gamma_1 \vee \Gamma_2$, is the metric graph obtained by identifying the points t_1 and t_2 in the disjoint union of Γ_1 and Γ_2. Denoting by $t \in \Gamma$ the image of t_1 and t_2 in Γ, one refers to $t \in \Gamma$ as a *cut-vertex* and to $\Gamma = \Gamma_1 \vee \Gamma_2$ as the decomposition corresponding to the cut-vertex t. (By abuse of notation, we will use t to denote both t_1 in Γ_1 and t_2 in Γ_2.) There is an addition map $\mathrm{Div}(\Gamma_1) \oplus \mathrm{Div}(\Gamma_2) \to \mathrm{Div}(\Gamma)$ which sends a pair of divisors D_1 and D_2 in $\mathrm{Div}(\Gamma_1)$ and $\mathrm{Div}(\Gamma_2)$ to the divisor $D_1 + D_2$ on Γ defined by pointwise addition of the coefficients in D_1 and D_2.

Let $D_1 \in \mathrm{Div}(\Gamma_1)$ and $D_2 \in \mathrm{Div}(\Gamma_2)$. For any nonnegative m, define $\eta_{\Gamma_1,D_1}(m)$ as the minimum integer h such that $r_{\Gamma_1}(D_1 + h(t_1)) = m$. Then

$$r_\Gamma(D) = \min_{m \geq 0}\{m + r_{\Gamma_2}(D_2 - \eta_{\Gamma_1,D_1}(m)(t_2))\} \tag{5.6.1}$$

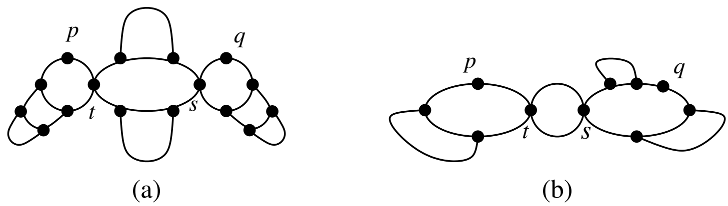


Figure 16. (a) A metric graph Γ in $\mathcal{M}(A_1) = \mathcal{M}(A_2)$. (b) A metric graph Γ in $\mathcal{M}(A_3)$.

(see [ibid.] for details).

In what follows, (5.6.1) will be applied to a metric graph $\Gamma \in \mathcal{M}(A_1) = \mathcal{M}(A_2)$ (see Figure 16(a) and Lemma 5.7), to a metric graph $\Gamma \in \mathcal{M}(A_3)$ (see Figure 16(b) and Lemma 5.9), and to $\Gamma_{27} \in \mathcal{M}(G_{27})$ with cut-vertex p in the proof of Proposition 5.6.

Lemma 5.7. *Let Γ be a metric graph in $\mathcal{M}(A_1) = \mathcal{M}(A_2)$ as depicted in Figure 16(a). For any nonnegative integers $a \leq 3$ and $b \leq 1$, the divisors $a(p) + b(q)$ and $b(p) + a(q)$ have rank zero in Γ.*

Proof. By symmetry, it is enough to prove the lemma for the divisor $D = 3(p) + (q)$. Consider the decomposition $\Gamma = \Gamma_p \vee \Gamma_q$ associated to the cut-vertex t in Γ, where Γ_p and Γ_q denote the closure in Γ of the two connected components of $\Gamma \setminus \{t\}$ which contain the points p and q, respectively.

We claim that $\eta_{\Gamma_q,(q)}(1) = 3$. Assume for the moment that this is true. Then by (5.6.1), we have

$$0 \leq r_\Gamma(3(p) + (q)) \leq 1 + r_{\Gamma_p}(3(p) - 3(t)).$$

By Lemma 5.8 below, in Γ_p we have $r_{\Gamma_p}(3(p) - 3(t)) = -1$. We thus infer that $r_\Gamma(3(p) + (q)) = 0$.

It remains to prove that $\eta_{\Gamma_q,(q)}(1) = 3$. In other words, we need to show that in Γ_q we have $r_{\Gamma_q}(2(t) + (q)) = 0$. For this, consider the decomposition $\Gamma_q = \Gamma_q^t \vee \Gamma_q^q$ corresponding to the cut-vertex s in Γ_q, where Γ_q^t and Γ_q^q denote the components which contain t and q, respectively. We claim that $\eta_{\Gamma_q^t,2(t)}(1) = 1$. Assuming the claim, we have $0 \leq r_{\Gamma_q}(2(t) + (q)) \leq 1 + r_{\Gamma_q^q}((q) - (s)) = 0$ (since q and s are not linearly equivalent in Γ_q^q; see Lemma 5.8). So it remains to prove that $\eta_{\Gamma_q^t,2(t)}(1) = 1$. This is equivalent to $r_{\Gamma_q^t}(2(t)) = 0$, which is obviously the case. $\square$

Lemma 5.8. *Let Γ be any metric graph in $\mathcal{M}(G_3)$, where G_3 is the totally degenerate graph depicted in Figure 17(a). Then the two divisors $3(p)$ and $3(t)$ are not linearly equivalent in Γ.*

Proof. By symmetry we can assume that the length of the edge $\{u, p\}$ is less than or equal to the length of the edge $\{t, w\}$. Then there exists a point t' in the segment

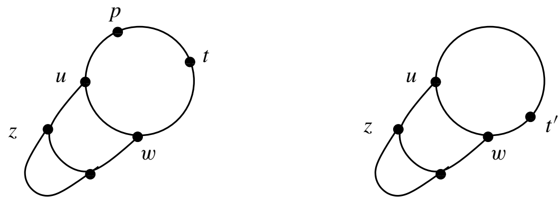


Figure 17. (a) The divisor $3(p)-3(t)$ is not rationally equivalent to zero. (b) $3(p)-3(t)\sim 3(u)-3(t')$.

$[t, w]$ so that $3(p)-3(t)\sim 3(u)-3(t')$ — see Figure 17(b) — and we are led to prove that $D=3(u)-3(t')$ is not linearly equivalent to zero. Consider the unique t'-reduced divisor $D_{t'}$ linearly equivalent to D in Γ (see, e.g., [Amini 2013; Baker and Norine 2007] for the definition and basic properties of reduced divisors). It will be enough to show that $D_{t'}\neq 0$. Three cases can occur, depending on the lengths ℓ_z, ℓ_w and $\ell_{t'}$ in Γ of the edges $\{u, z\}$, $\{u, w\}$ and the segment $\{u, t'\}$, respectively:

- If $\ell_z=\min\{\ell_z, \ell_u, \ell_{t'}\}$, then there are two points w' and t'' on the segments $\{u, w\}$ and $\{u, t'\}$, respectively, such that $D_{t'}=(z)+(w')+(t'')-3(t')$.
- If $\ell_u=\min\{\ell_z, \ell_u, \ell_{t'}\}$, then there are two points z' and t'' on the segments $\{u, z\}$ and $\{u, t'\}$, respectively, such that $D_{t'}=(z')+(w)+(t'')-3(t')$.
- If $\ell_{t'}=\min\{\ell_z, \ell_u, \ell_{t'}\}$, then there are two points z' and w' on the segments $\{u, z\}$ and $\{u, w\}$, respectively, such that $D_{t'}=(z')+(w')-2(t')$.

In all the three cases, we have $D_{t'}\neq 0$, which shows that D cannot be equivalent to zero in Γ. □

Lemma 5.9. *Let $\Gamma\in\mathcal{M}(A_3)$ be a metric graph as depicted in Figure 16(b). For any $a, b\le 2$, the divisor $a(p)+b(q)$ has rank zero on Γ.*

Proof. The arguments are similar to the ones used in the proof of Lemma 5.7. Consider the cut-vertex t in Γ and denote by Γ_p and Γ_q the corresponding components containing p and q, respectively. We claim that $\eta_{\Gamma_q,2(q)}(1)=2$. This obviously implies the lemma. Indeed, $r_{\Gamma_p}(2(p)-2(t))=-1$ (which can be verified by an analogue of Lemma 5.8 in Γ_p), and thus (5.6.1) implies that $r_\Gamma(2(p)+2(q))\le 1+r_{\Gamma_p}(2(p)-2(t))=0$.

To show that $\eta_{\Gamma_q,2(q)}(1)=2$, it will be enough to show that $r_{\Gamma_q}(2(q)+(t))=0$. This can be done in exactly the same way by considering the other cut-vertex s adjacent to t in Γ_q. □

Lemma 5.10. *Let x_1, x_2 and x_3 be distinct points in $\mathbb{P}^1(k)$. Then there does not exist a morphism $f : \mathbb{P}^1 \to \mathbb{P}^1$ of degree four branched over x_1, x_2 and x_3 and having ramification profile* $(2, 2)$, $(2, 2)$ *and* $(3, 1)$ *at these three points.*

Proof. Suppose that such a rational map $f : \mathbb{P}^1 \to \mathbb{P}^1$ exists. The monodromy group of f is a subgroup of $\mathfrak{S}_4$, so its cardinality is of the form $2^a 3^b$. In particular, if the characteristic of k is neither 2 nor 3, then f has a tame monodromy group and the nonexistence of f then comes from the fact that $H^4_{0,0}((2, 2), (2, 2), (3, 1)) = 0$ (see Example 3.4).

Hence it remains to check the lemma for $\mathrm{char}(k) = 2, 3$. Note that the same technique we use in this case works in any characteristic, but the computations are a bit more tedious in characteristic different from 2 and 3.

Up to the action of $\mathrm{GL}(2, k)$ on $\mathbb{P}^1$ via automorphisms, we may assume that $x_1 = 0$, $x_2 = \infty$ and $x_3 = 1$, and that

$$f(X) = a\frac{X^2(X+1)^2}{(X+b)^2}$$

with $a \neq 0$ and $b \neq 0, -1$. Hence the condition on the ramification profile of x_3 translates as

$$aX^2(X+1)^2 - (X+b)^2 = c(X-d)^3(X-e)$$

with $c \neq 0$, $d \neq 0, -1, b$ and $e \neq 0, -1, b, d$. Looking at the coefficients of the two polynomials, we obtain the five equations

$$(E_1):\ a = c, \quad (E_2):\ ec = -2a - 3cd, \quad (E_3):\ a - 1 = 3cd(d+e),$$
$$(E_4):\ 2b = cd^2(d+3e), \quad (E_5):\ -b^2 = cd^3e.$$

If k has characteristic 2, then (E_2) becomes $ec = cd$, which contradicts the fact that $e \neq d$.

If k has characteristic 3, then these five equations become

$$(E_1): a = c, \ (E_2): ec = a, \ (E_3): a = 1, \ (E_4): -b = cd^3, \ (E_5): -b^2 = cd^3e.$$

Equations (E_1), (E_2), (E_3) imply $a = c = e = 1$. Then (E_4) and (E_5) become $-b = d^3 = -b^2$; hence $b = 1 = e$, which contradicts our assumptions. $\square$

We can now give the promised proof of Proposition 5.6.

Proof of Proposition 5.6. Suppose that there exists a metrized complex of k-curves $\mathscr{C}_{27}$ of genus 27 with underlying augmented metric graph Γ_{27} in $\mathcal{M}(G_{27})$, and admitting a finite harmonic morphism of metrized complexes of degree four $\varphi : \mathscr{C}_{27} \to \mathscr{T}$, for $\mathscr{T}$ of genus zero with underlying metric tree denoted by T. Without loss of generality, we may assume that T has no infinite vertex $q \in V_\infty(T)$ such that any infinite edge e' adjacent to an infinite vertex $q' \in \varphi^{-1}(q)$ has $d_{e'}(\varphi) = 1$.

We are going to prove below that the local degree at p is 4. Assuming that this is the case, we show how the proposition follows. Denote by Γ_1, Γ_2 and Γ_3 the three components of $\Gamma_{27} \setminus \{p\}$ which contain A_1, A_2 and A_3, respectively. Since the degree of φ at p is four, we have $\varphi^{-1}(\varphi(p)) = \{p\}$. Therefore, by the connectivity of Γ_i, the images of Γ_i under φ are pairwise disjoint in T. This shows that for x sufficiently close to $\varphi(p)$ in T, the support of the divisor $D_x(\varphi)$ lives entirely in one of the Γ_i for $i \in \{1, 2, 3\}$. Choose x_i sufficiently close to $\varphi(p)$ such that the support of $D_{x_i}(\varphi)$ is contained in Γ_i. Applying Proposition 4.2, we see that each divisor $D_{x_i}(\varphi)$ has rank one in Γ_i. Now, according to Lemma 5.7, the degree-four divisor $D_{x_1}(\varphi)$ (resp. $D_{x_2}(\varphi)$) must be of the form $2(a)+2(b)$ for two points a and b sufficiently close to p and lying on the two different branches of Γ_1 (resp. Γ_2) adjacent to p. Similarly, by Lemma 5.9, the divisor $D_{x_3}(\varphi)$ has to be of the form $3(a) + (b)$ for two points a and b sufficiently close to p and lying on the two different branches of Γ_3 adjacent to p. This shows that the map φ_p, the restriction of φ to a sufficiently small neighborhood of p in Γ_{27}, coincides with the map depicted in Figure14(a). The proposition now follows from Lemma 5.10.

It remains to prove that $d_p(\varphi) = 4$. We first claim that φ maps one of the components Γ_i, for $i = 1, 2, 3$, onto a connected component of $T \setminus \{\varphi(p)\}$. Otherwise, for the sake of contradiction, suppose that $\varphi^{-1}(\varphi(p))$ consists of p and one point p_i in each of the components Γ_i for $i = 1, 2, 3$. Then φ has local degree one at each of the points p_i. By Proposition 4.2, $D_{\varphi(p)}(\varphi) = (p) + (p_1) + (p_2) + (p_3)$ has rank one in Γ. By (5.6.1) applied to the cut-vertex p in Γ_{27}, we infer that the divisor $(p) + (p_i)$ has rank one in the metric graph $\overline{\Gamma}_i$, the closure of Γ_i in Γ_{27}. In other words, the metric graphs $\overline{\Gamma}_i$ are hyperelliptic, which is clearly not the case. This gives a contradiction and the claim follows.

Summarizing, there must exist at least one Γ_i such that φ maps Γ_i onto one of the connected components of $T \setminus \{\varphi(p)\}$. Reasoning again as in the first part of the proof, it follows from Proposition 4.2 and Lemmas 5.7 and 5.9 that the restriction of φ to Γ_i has degree four, which implies that $d_p(\varphi) = 4$. □

5.11. Lifting divisors of given rank. First, recall that to a smooth proper curve X over K together with a semistable vertex set V and a subset D_0 of $X(K)$ compatible with V, we can naturally associate a metrized complex of curves $\mathscr{C} = \Sigma(X, V \cup D_0)$ with underlying augmented metric graph Γ. As in [Amini and Baker 2014], there are natural specialization maps on divisors, which we denote for simplicity by the same letter τ_*:

$$\tau_* : \mathrm{Div}(X) \longrightarrow \mathrm{Div}(\mathscr{C}) \quad \text{and} \quad \tau_* : \mathrm{Div}(\mathscr{C}) \longrightarrow \mathrm{Div}(\Gamma).$$

Since this discussion is pointless in the case of rational curves, we may assume that X (equivalently, $\mathscr{C}$ or the augmented metric graph Γ) has positive genus. We will also assume that Γ does not have any infinite vertices, i.e., that D_0 is empty, which does not lead to any real loss of generality and which makes various statements easier to write down and understand. We may also assume without loss of generality that V is a strongly semistable vertex set of X.

According to the specialization inequality [Baker 2008; Amini and Caporaso 2013; Amini and Baker 2014]), for any divisor D in $\mathrm{Div}(X)$ one has

$$r_X(D) \le r_{\mathscr{C}}(\tau_*(D)) \le r^{\#}_{\Gamma}(\tau_*(D)) \le r_{\Gamma}(\tau_*(D)), \tag{5.11.1}$$

where r_X, $r_{\mathscr{C}}$ and r_Γ denote rank of divisors on X, $\mathscr{C}$ and (unaugmented) Γ, respectively, and $r^{\#}_{\Gamma}$ denotes the weighted rank in the augmented metric graph (Γ, g) (see (4.11)).

We spend the rest of this section discussing the sharpness of the inequalities appearing in (5.11.1).

Definition 5.12. Let $\mathscr{C}$ be a metrized complex of curves whose underlying metric graph Γ has no infinite leaves, and let $\mathscr{D}$ be a Λ-rational divisor in $\mathrm{Div}_\Lambda(\mathscr{C})$. A *lifting* of the pair $(\mathscr{C}, \mathscr{D})$ consists of a triple $(X, V; D_X)$ where X is a smooth proper curve over K, V is a strongly semistable vertex set for which $\mathscr{C} = \Sigma(X, V)$, and D_X is a divisor in $\mathrm{Div}(X)$ with $\mathscr{D} \sim \tau_*(D_X)$. We say that the inequality $r_X \le r_{\mathscr{C}}$ is *sharp* if for any metrized complex of curves $\mathscr{C}$ and any divisor $\mathscr{D} \in \mathrm{Div}(\mathscr{C})$, there exists a lifting $(X, V; D_X)$ of $(\mathscr{C}, \mathscr{D})$ such that $r_X(D_X) = r_{\mathscr{C}}(\mathscr{D})$.

We can define in a similar way what it means to lift a divisor on an (augmented) metric graph to a divisor on a metrized complex of curves or to a smooth proper curve over K, and what it means for the corresponding specialization inequalities to be sharp.

It is easy to see that the inequality $r^{\#}_{\Gamma} \le r_\Gamma$ is not sharp (see [Amini and Baker 2014] for a precise formula relating the two rank functions).

The following example is due to Ye Luo (unpublished); we thank him for his permission to include it here. Together with Corollary I.4.28, it implies that the inequality $r_X \le r_\Gamma$ is not sharp.

Example 5.13 (Luo). Let Γ be a metric graph in $\mathscr{M}(G_7)$, where G_7 is the graph of genus seven depicted in Figure 18(a), such that all edge lengths in Γ are equal, and let $D = (p) + (q) + (s) \in \mathrm{Div}(\Gamma)$. Then $r_\Gamma(D) = 1$, but there does not exist any finite harmonic morphism of metric graphs $\varphi : \Gamma' \to T$ of degree three to a metric tree for any $\Gamma' \in \mathscr{M}(G_7)$. In particular, this shows that the stable gonality of an augmented graph can be greater than its divisorial gonality.

We briefly sketch a proof. Suppose that such a finite harmonic morphism $\varphi :$ $\Gamma' \to T$ exists. Since Γ' is not hyperelliptic, one easily verifies that $D_{\varphi(p)}(\varphi) = 3(p)$, $D_{\varphi(q)}(\varphi) = 3(q)$, and $D_{\varphi(s)}(\varphi) = 3(s)$. This shows the existence of a finite morphism $\varphi' : \Gamma'_1 \to T'$ of degree three to a metric tree T', where Γ'_1 is depicted in Figure 18(b),

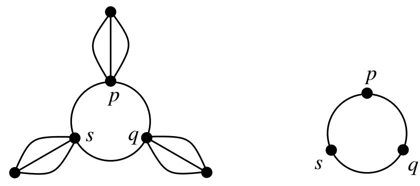


Figure 18. (a) The graph G_7. (b) The metric graph $\Gamma'_1 \subset \Gamma'$.

so that $D_{\varphi'(p)}(\varphi') = 3(p)$, $D_{\varphi'(q)}(\varphi') = 3(q)$ and $D_{\varphi(s)}(\varphi') = 3(s)$. But it is easy to verify by hand that such a morphism φ' does not exist.

Proposition 5.14. *Neither of the inequalities $r_X \le r_{\mathscr{C}}$ and $r_{\mathscr{C}} \le r_\Gamma^\#$ is sharp.*

Proof. To show the nonsharpness of the inequality $r_X \le r_{\mathscr{C}}$, let $\mathscr{C}$ be a metrized complex of curves whose underlying metric graph Γ belongs to the family depicted in Figure 12, with first Betti number κ, and whose genus function is positive at each vertex. Consider the divisor $\mathscr{D}_d = d(p) \oplus d(x)$ in $\mathscr{C}$ for a closed point x in C_p and d a positive integer. If d is sufficiently large compared to the genera of the vertices, then $r_{\mathscr{C}}(\mathscr{D}_d) \ge 1$. If the pair $(\mathscr{C}, \mathscr{D}_d)$ lifted to a triple $(X, V; D_X)$ with $\tau_*(D_X) \sim \mathscr{D}_d$, then there would exist a finite harmonic morphism $\varphi : \widetilde{\mathscr{C}} \to \mathscr{T}$ from a modification of $\mathscr{C}$ to a metrized complex of curves of genus zero. But this would imply the existence of a degree-d morphism $\varphi_p : C_p \to \mathbb{P}^1$ such that the image of red_p (on edges adjacent to p in Γ) is contained in the set of critical values of φ_p. By the Riemann–Hurwitz formula, this is impossible for κ large enough compared to d.

To show the nonsharpness of the inequality $r_{\mathscr{C}} \le r^\#$, let again (Γ, g) be an augmented metric graph with underlying graph depicted in Figure 12 with $\kappa \ge 3$ and $2 \le 2g(p) < \kappa - 2$, and let $D = 2(p)$. One easily computes that $r_\Gamma^\#(D) = 1$. An algebraic curve of genus $g(p) \ge 1$ contains at most $2g(p) + 2$ distinct points p such that $2(p)$ is in a given linear system of degree two, which implies that (Γ, g) cannot be lifted to a hyperelliptic metrized complex of curves. This shows that the inequality $r_{\mathscr{C}} \le r_\Gamma^\#$ is not sharp. □

References

[ABBR1] O. Amini, M. Baker, E. E. Brugallé, and J. Rabinoff, "Lifting harmonic morphisms I: Metrized complexes and Berkovich skeleta", preprint, 2013. To appear in *Res. Math. Sci.* arXiv 1303.4812

[Amini 2013] O. Amini, "Reduced divisors and embeddings of tropical curves", *Trans. Amer. Math. Soc.* **365**:9 (2013), 4851–4880. MR 3066772 Zbl 06204978

[Amini and Baker 2014] O. Amini and M. Baker, "Linear series on metrized complexes of algebraic curves", *Math. Ann.* (online publication October 2014).

[Amini and Caporaso 2013] O. Amini and L. Caporaso, "Riemann–Roch theory for weighted graphs and tropical curves", *Adv. Math.* **240** (2013), 1–23. MR 3046301 Zbl 1284.14087

[Baker 2008] M. Baker, "Specialization of linear systems from curves to graphs", *Algebra Number Theory* **2**:6 (2008), 613–653. With an appendix by Brian Conrad. MR 2010a:14012

[Baker and Norine 2007] M. Baker and S. Norine, "Riemann–Roch and Abel–Jacobi theory on a finite graph", *Adv. Math.* **215**:2 (2007), 766–788. MR 2008m:05167 Zbl 1124.05049

[Baker and Norine 2009] M. Baker and S. Norine, "Harmonic morphisms and hyperelliptic graphs", *Int. Math. Res. Not.* **2009**:15 (2009), 2914–2955. MR 2010e:14031 Zbl 1178.05031

[Baker et al. 2012] M. Baker, S. Payne, and J. Rabinoff, "Nonarchimedean geometry, tropicalization, and metrics on curves", preprint, 2012. arXiv 1104.0320

[Bertin and Mézard 2000] J. Bertin and A. Mézard, "Déformations formelles des revêtements sauvagement ramifiés de courbes algébriques", *Invent. Math.* **141**:1 (2000), 195–238. MR 2001f:14023 Zbl 0993.14014

[Bertrand et al. 2011] B. Bertrand, E. Brugallé, and G. Mikhalkin, "Tropical open Hurwitz numbers", *Rend. Semin. Mat. Univ. Padova* **125** (2011), 157–171. MR 2866125 Zbl 1226.14066

[Brugallé and Mikhalkin $\geq$ 2015] E. Brugallé and G. Mikhalkin, "Realizability of superabundant curves", in preparation.

[Caporaso 2014] L. Caporaso, "Gonality of algebraic curves and graphs", in *Algebraic and complex geometry: in honour of Klaus Hulek's 60th birthday*, Proceedings in Mathematics and Statistics **71**, Springer, Cham, 2014. arXiv 1201.6246

[Cavalieri et al. 2010] R. Cavalieri, P. Johnson, and H. Markwig, "Tropical Hurwitz numbers", *J. Algebraic Combin.* **32**:2 (2010), 241–265. MR 2011m:14089 Zbl 1218.14058

[Chambert-Loir and Ducros 2012] A. Chambert-Loir and A. Ducros, "Formes différentielles réelles et courants sur les espaces de Berkovich", preprint, 2012. arXiv 1204.6277

[Chan 2013] M. Chan, "Tropical hyperelliptic curves", *J. Algebraic Combin.* **37**:2 (2013), 331–359. MR 3011346 Zbl 1266.14050

[Cornelissen et al. 2014] G. Cornelissen, F. Kato, and J. Kool, "A combinatorial Li–Yau inequality and rational points on curves", *Math. Ann.* (online publication July 2014).

[DeMarco and McMullen 2008] L. G. DeMarco and C. T. McMullen, "Trees and the dynamics of polynomials", *Ann. Sci. Éc. Norm. Supér.* (4) **41**:3 (2008), 337–382. MR 2010d:37087 Zbl 1202.37067

[Edmonds et al. 1984] A. L. Edmonds, R. S. Kulkarni, and R. E. Stong, "Realizability of branched coverings of surfaces", *Trans. Amer. Math. Soc.* **282**:2 (1984), 773–790. MR 85k:57005 Zbl 0603.57001

[Green and Matignon 1998] B. Green and M. Matignon, "Liftings of Galois covers of smooth curves", *Compositio Math.* **113**:3 (1998), 237–272. MR 99k:14045a Zbl 0923.14006

[Husemoller 1962] D. H. Husemoller, "Ramified coverings of Riemann surfaces", *Duke Math. J.* **29** (1962), 167–174. MR 25 #188 Zbl 0196.34001

[Kontsevich and Soibelman 2001] M. Kontsevich and Y. Soibelman, "Homological mirror symmetry and torus fibrations", pp. 203–263 in *Symplectic geometry and mirror symmetry* (Seoul, 2000), edited by K. Fukaya et al., World Sci. Publ., River Edge, NJ, 2001. MR 2003c:32025 Zbl 1072.14046

[Kontsevich and Soibelman 2006] M. Kontsevich and Y. Soibelman, "Affine structures and non-Archimedean analytic spaces", pp. 321–385 in *The unity of mathematics* (Cambridge, MA, 2003), edited by P. Etingof et al., Progr. Math. **244**, Birkhäuser, Boston, MA, 2006. MR 2006j:14054 Zbl 1114.14027

[Lando and Zvonkin 2004] S. K. Lando and A. K. Zvonkin, *Graphs on surfaces and their applications*, Encyclopaedia of Mathematical Sciences **141**, Springer, Berlin, 2004. With an appendix by Don B. Zagier, Low-Dimensional Topology, II. MR 2005b:14068 Zbl 1040.05001

[Mikhalkin 2006] G. Mikhalkin, "Tropical geometry and its applications", pp. 827–852 in *International Congress of Mathematicians* (Madrid, 2006), vol. II, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR 2008c:14077 Zbl 1103.14034

[Mikhalkin and Zharkov 2008] G. Mikhalkin and I. Zharkov, "Tropical curves, their Jacobians and theta functions", pp. 203–230 in *Curves and abelian varieties* (Athens, GA, 2007), edited by V. Alexeev et al., Contemp. Math. **465**, Amer. Math. Soc., Providence, RI, 2008. MR 2011c:14163 Zbl 1152.14028

[Obus and Wewers 2014] A. Obus and S. Wewers, "Cyclic extensions and the local lifting problem", *Ann. of Math.* (2) **180**:1 (2014), 233–284. MR 3194815 Zbl 06316070

[Oort 1987] F. Oort, "Lifting algebraic curves, abelian varieties, and their endomorphisms to characteristic zero", pp. 165–195 in *Algebraic geometry* (Bowdoin College, Brunswick, ME, 1985), edited by S. J. Bloch, Proc. Sympos. Pure Math. **46**, Amer. Math. Soc., Providence, RI, 1987. MR 89c:14069 Zbl 0645.14017

[Payne 2009] S. Payne, "Analytification is the limit of all tropicalizations", *Math. Res. Lett.* **16**:3 (2009), 543–556. MR 2010j:14104 Zbl 1193.14077

[Pervova and Petronio 2006] E. Pervova and C. Petronio, "On the existence of branched coverings between surfaces with prescribed branch data. I", *Algebr. Geom. Topol.* **6** (2006), 1957–1985. MR 2008h:57005 Zbl 1132.57002

[Pervova and Petronio 2008] E. Pervova and C. Petronio, "On the existence of branched coverings between surfaces with prescribed branch data. II", *J. Knot Theory Ramifications* **17**:7 (2008), 787–816. MR 2009c:57004 Zbl 1160.57001

[Pop 2014] F. Pop, "The Oort conjecture on lifting covers of curves", *Ann. of Math.* (2) **180**:1 (2014), 285–322. MR 3194816 Zbl 06316071

[Raynaud 1999] M. Raynaud, "Spécialisation des revêtements en caractéristique $p > 0$", *Ann. Sci. École Norm. Sup.* (4) **32**:1 (1999), 87–126. MR 2000e:14016 Zbl 0999.14004

[Sekiguchi et al. 1989] T. Sekiguchi, F. Oort, and N. Suwa, "On the deformation of Artin–Schreier to Kummer", *Ann. Sci. École Norm. Sup.* (4) **22**:3 (1989), 345–375. MR 91g:14041 Zbl 0714.14024

[SGA 1 1971] A. Grothendieck, *Revêtements étales et groupe fondamental* (Séminaire de Géométrie Algébrique du Bois Marie 1960–1961 = SGA 1), Lecture Notes in Math. **224**, Springer, Berlin, 1971. Updated and annotated reprint, Soc. Math. de France, Paris, 2003. MR 50 #7129 Zbl 0572.55016

Communicated by Ravi Vakil
Received 2013-06-25 Revised 2014-07-09 Accepted 2014-12-06

oamini@math.ens.fr — *CNRS, Département de mathématiques et applications, École Normale Supérieure, 45 Rue d'Ulm, 75005 Paris, France*

mbaker@math.gatech.edu — *School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, United States*

erwan.brugalle@imj-prg.fr — *Université Pierre et Marie Curie (Paris 6), 4 Place Jussieu, 75005 Paris, France*

rabinoff@math.gatech.edu — *School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, United States*

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.317

Noncommutative Hilbert modular symbols

Ivan Horozov

The main goal of this paper is to construct noncommutative Hilbert modular symbols. However, we also construct commutative Hilbert modular symbols. Both the commutative and the noncommutative Hilbert modular symbols are generalizations of Manin's classical and noncommutative modular symbols. We prove that many cases of (non)commutative Hilbert modular symbols are periods in the Kontsevich–Zagier sense. Hecke operators act naturally on them.

Manin defined the noncommutative modular symbol in terms of iterated path integrals. In order to define noncommutative Hilbert modular symbols, we use a generalization of iterated path integrals to higher dimensions, which we call iterated integrals on membranes. Manin examined similarities between noncommutative modular symbol and multiple zeta values in terms of both infinite series and of iterated path integrals. Here we examine similarities in the formulas for noncommutative Hilbert modular symbol and multiple Dedekind zeta values, recently defined by the current author, in terms of both infinite series and iterated integrals on membranes.

1. Introduction

Classical elliptic modular symbols were introduced by Birch [1971] and Manin [1972] in connection with the Birch–Swinnerton-Dyer conjecture for certain congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$. We recall that a modular symbol $\{p, q\}$ is associated to a pair of cusp points $p, q \in \mathbb{P}^1(\mathbb{Q})$ on the completed upper half-plane $\mathbb{H}^1 \cup \mathbb{P}^1(\mathbb{Q})$. One can think of the modular symbol $\{p, q\}$ as a homology class of the geodesic

MSC2010: primary 11F41; secondary 11F67, 11M32.
Keywords: modular symbols, Hilbert modular groups, Hilbert modular surfaces, iterated integrals.

connecting p and q, in $H_1(X_\Gamma, \{\text{cusps}\})$, where X_Γ is the modular curve associated to a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. One can pair $\{p, q\}$ with a cusp form f by

$$\{p, q\} \times f \mapsto \int_p^q f\, dz$$

If f is a cusp form of weight 2 then $f\, dz$ can be viewed as a cohomology class in $H^1(X_\Gamma)$. This gives a pairing between homology (Betti) and cohomology (de Rham) that leads to periods. Modular symbols are a useful tool for studying L-functions and in the computation of cohomology groups. For a review of such topics, one can consult [Manin 2009].

Their theory was developed in [Manin 1972; Drinfeld 1973; Shokurov 1980; Mazur 1973]. Later the theory was extended to higher ranks in [Ash and Rudolph 1979; Ash and Borel 1990; Gunnells 2000b].

Elliptic modular symbols are important tool in the study of modular forms. They are particularly useful in computations with modular forms. J. Cremona [1997] designed algorithms for computations with elliptic curves, based on modular symbols ("modular symbol algorithms"). Some of the applications include the computation of homology and cohomology. Also, the study of special values of L-functions became a vast area of applications of classical modular symbols; see [Mazur and Swinnerton-Dyer 1974; Kazhdan et al. 2000].

Later, W. Stein also contributed to the difficult area of computations with modular forms. See his excellent book [Stein 2007], which contains both theory and computational methods. For higher-rank groups, one can consult the appendix of this book, by P. Gunnells.

Manin's noncommutative modular symbol [2006] is a generalization of both the classical modular symbol and of multiple zeta values in terms of Chen's iterated integral theory in the holomorphic setting. Manin showed that the noncommutative modular symbol is a noncommutative 1-cocycle. He also showed that each of the iterated integrals on Hecke eigenforms that enter in the noncommutative modular symbol are periods.

The main goal of this paper is to construct noncommutative Hilbert modular symbols. However, we also construct an analog of the classical modular symbol for Hilbert modular varieties. Both symbols are generalizations of the corresponding constructions by Manin.

We compute explicit integrals in terms of the noncommutative Hilbert modular symbol of type $\boldsymbol{b}$, and present similar formulas for the recently defined multiple Dedekind zeta values (see [Horozov 2014b]). We prove that the iterated integrals on membranes that enter in the noncommutative modular symbol of type $\boldsymbol{c}$ are periods. We also give some explicit and some categorical arguments in support of a conjecture that a certain type of noncommutative Hilbert modular symbol satisfies a noncommutative 2-cocycle condition.

Before describing our results, let us recall the noncommutative modular symbol of Manin [2006]. Let $\nabla = d - \sum_{i=1}^{m} X_i f_i \, dz$ be a connection on the upper half-plane, where $f_1, \ldots, f_m$ are cusp forms and $X_1, \ldots, X_m$ are formal variables. One can think of $X_1, \ldots, X_m$ as constant square matrices of the same size.

Let J_a^b be the parallel transport to the point b of the identity matrix 1 at the point a. Alternatively, J_b^a can be written as a generating series of iterated path integrals of the forms $f_1 \, dz, \ldots, f_m \, dz$, (see [Chen 1977] and [Manin 2006]), namely,

$$J_a^b = 1 + \sum_{i=1}^{m} X_i \int_a^b f_i \, dz + \sum_{i,j=1}^{m} X_i X_j \int_a^b f_i \, dz \cdot f_j \, dz + \cdots$$

Then $J_a^b J_b^c = J_a^c$. This property leads to the 1-cocycle $c_a^1(\gamma) = J_{\gamma a}^a$ for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, which is the noncommutative modular symbol (see [Manin 2006] and Section 2 of this paper). If $f_1, \ldots, f_m$ are normalized cusp Hecke eigenforms, then each iterated integral appearing in the generating series J_a^b is a period. In this paper we introduce both commutative and noncommutative modular symbols for Hilbert modular surfaces. As it turned out we need some new tools, in comparison to the classical modular symbols. In particular, for the noncommutative Hilbert modular symbol we need iterated integrals in dimension higher than one. We introduce them and study their properties in the special case of Hilbert modular surfaces for the Hilbert modular group $\mathrm{SL}_2(\mathbb{O}_K)$. For the Hilbert modular group, one may consult [Bruinier et al. 2008] and [Freitag 1990]. In the case of Hilbert modular surface, it is not possible to repeat Manin's constructions for the noncommutative modular symbols, since the integration domain is two-dimensional over the complex numbers. Instead, we develop a new approach (Section 3), which we call iterated integrals on membranes. This is a higher-dimensional analogue of iterated path integrals. In Section 4G, we explore similar relations between noncommutative Hilbert modular symbols and multiple Dedekind zeta values (see [Horozov 2014b]).

In Section 4, we associate modular symbols for $\mathrm{SL}_2(\mathbb{O}_K)$ to geodesic triangles and geodesic diangles (2-cells whose boundaries have two vertices and two edges, which are geodesics). We are going to explain how the geodesic triangles and the geodesic diangles are constructed. Consider four cusp points in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$. We can map any three of them to 0, 1 and ∞ with a linear fractional transformation $\gamma \in \mathrm{GL}_2(K)$. There is a diagonal map $\mathbb{H}^1 \to \mathbb{H}^2$, whose image Δ contains 0, 1 and ∞. We can take a pullback of Δ with respect to the map γ in order to obtain a holomorphic (or antiholomorphic) curve that passes through the given three points. If $\det \gamma$ is totally positive or totally negative then $\gamma^* \Delta$ is a holomorphic curve in $\mathbb{H}^2$. If $\det \gamma$ is not totally positive or totally negative (that is, in one of the real embeddings it is positive and in the other it is negative) then $\gamma^*\Delta$ is antiholomorphic. This means that it is holomorphic in $\mathbb{H}^2$ if we conjugate the complex structure in one of the copies of $\mathbb{H}^1$. The same type of change of the complex structure is considered in [Freitag 1990].

On each holomorphic (or antiholomorphic) curve $\gamma^*\Delta$, there is a unique geodesic triangle connecting the three given points. However, if we take two of the points, we see that they belong to two geodesic triangles. Thus they belong to two holomorphic, (antiholomorphic) curves. Therefore, there are two geodesic connecting the two points — each lying on different holomorphic (antiholomorphic) curves, as faces of the corresponding geodesic triangles defining the curves. There are two pairings that we consider: the first is an integral of a cusp form over a geodesic triangle and the second is an integral of a cusp form over a geodesic diangle. If we integrate a holomorphic 2-form coming from a cusp form over a geodesic triangle, we obtain 0, if the triangle lies on an holomorphic curve. Thus the only nonzero pairings come from integration of a cusp form over a diangle or over a triangle lying on an antiholomorphic curve.

Now let us look again at the four cusp points together with the geodesics that we have just described. We obtain four geodesic triangles, corresponding to each triple of points among the four points, and six diangles, corresponding to the six "edges" of a tetrahedron with vertices the four given points. Thus, we obtain a "tetrahedron" with thickened edges. We will use tetrahedrons with thickened edges as an intuition for a noncommutative 2-cocycle relation (see Conjecture 4.15) for the noncommutative Hilbert modular symbol, which is an analogue of Manin's noncommutative 1-cocycle relation for the noncommutative modular symbol.

Usually, the four vertices are treated as a tetrahedron and a 2-cocycle is functional on the faces, considered as 2-chains. The boundary is defined as a sum of the 2-cocycles on each of the faces (which are triangles). The boundary of the tetrahedron gives a boundary relation of a 2-cocycle.

In our case the analogue of a 2-cocycle is a functional on diangles and on triangles. And the boundary map is a sum over the faces of the thickened tetrahedron. Thus, the faces of the thickened tetrahedron are four triangles and six diangles, corresponding to the six edges of a tetrahedron.

We show that the geodesics on the boundary of a diangle or of a geodesic triangle lie on a holomorphic curves $\gamma^*\Delta$ for various elements γ with totally positive or totally negative determinant. This implies that when we take the quotient by a Hilbert modular group the holomorphic curve $\gamma^*\Delta$ becomes a Hirzebruch–Zagier divisor [Hirzebruch and Zagier 1976]. Then we prove that the commutative Hilbert modular symbols paired with a cusp forms of weight (2, 2) gives periods in the sense of [Kontsevich and Zagier 2001].

In order to construct a noncommutative Hilbert modular symbol, first we define a suitable generalization of iterated path integrals, which we call iterated integrals on membranes (see Section 3). We choose the word "membrane" since such integrals are invariant under suitable variation of the domain of integration.

There is a topological reason for considering a noncommutative Hilbert modular symbol as opposed to only a commutative one. Let us first make such a comparison for the case of $\mathrm{SL}_2(\mathbb{Z})$. The commutative modular symbol captures $H_1(X_\Gamma)$, while the noncommutative symbol captures the rational homotopy type of the modular curve X_Γ. Now, let $\widetilde{X}$ be a smooth Hilbert modular surface, by which we mean the minimal desingularization of the Borel–Baily compactification due to Hirzebruch. Then the rational fundamental group of a Hilbert modular surface vanishes: $\pi_1(\widetilde{X})_{\mathbb{Q}} = 0$ (see [Bruinier et al. 2008]). The noncommutative Hilbert modular symbol is an attempt to capture more from the rational homotopy type than $H_2(\widetilde{X})$ captures.

For the convenience of the reader, we first define type $\boldsymbol{a}$ iterated integrals on membranes (Definition 3.3). They are simpler to define and more intuitive. However, they do not have enough properties. (For example, they do not have an integral shuffle relation.) Then we define type $\boldsymbol{b}$ iterated integrals on membranes (Definition 3.4), which involves two permutations. Type $\boldsymbol{b}$ has integral shuffle relation (Theorem 3.21(i)), and type $\boldsymbol{a}$ is a particular case of type $\boldsymbol{b}$.

We are mostly interested in iterated integrals of type $\boldsymbol{b}$. If there is no index specifying the type of iterated integral over membranes, we assume that it is of type $\boldsymbol{b}$.

Similarly to Manin's approach, we define a generating series of iterated integrals over membrane of type $\boldsymbol{b}$ over U, which we denote by $J(U)$. We also define a shuffle product of generating series of iterated integrals over membranes of type $\boldsymbol{b}$ (see Theorem 3.21(iii)),

$$\phi(J(U_1) \times_{\mathrm{Sh}} J(U_2)) = J(U_1 \cup U_2),$$

for U_1, U_2 disjoint 2-dimensional manifolds with corners contained in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$ (see [Borel and Serre 1973]). This shuffle product generalizes the composition of generating series of iterated path integrals, namely, $J_a^b J_b^c = J_a^c$, to dimension 2. Note that a similar definition is also possible in higher dimensions. Also, $J(U)$ is invariant under homotopy. This allows us to consider cocycles and coboundaries, where the relations use homotopy invariance and values at different cells can be composed via the shuffle product.

We define noncommutative Hilbert modular symbols, which we call c^1 and c^2; c^1 is the functional J on certain geodesic diangles and c^2 is the functional J on geodesic triangles. Conjecturally, c^1 is a 1-cocycle such that if we change the base point of c^1 then c^1 is modified by a coboundary. Also, conjecturally, c^2 is a 2-cocycle up to finitely many multiples of different values of c^1. Also, if we change the base point of c^2 then c^2 is modified by a coboundary up to a finitely many multiples of different values of c^1. In Section 4E we give explicit formulas in support of the interpretation of the noncommutative symbols as cocycles.

In Section 4F, we give a categorical construction, which might help to prove that the noncommutative symbols are cocycles.

In Section 4G, we define multiple L-values associated to cusp forms, and we compare them to multiple Dedekind zeta values (see [Horozov 2014b]).

We also briefly recall the Riemann zeta values and multiple zeta values (MZVs). The Riemann zeta values are defined as

$$\zeta(k) = \sum_{n>0} \frac{1}{n^k},$$

where n is an integer. MZVs are defined as

$$\zeta(k_1, \ldots, k_m) = \sum_{0<n_1<\cdots<n_m} \frac{1}{n_1^{k_1} \cdots n_m^{k_m}},$$

where $n_1, \ldots, n_m$ are integers. The above MZV is of depth m. Riemann zeta values $\zeta(k)$ and MZV $\zeta(k_1, \ldots, k_m)$ were defined by Euler [1748] for $m = 1, 2$.

The common feature of MZVs and the noncommutative modular symbol is that they both can be written as iterated path integrals (see [Goncharov 2001a; 2001b]). Moreover, Manin's noncommutative modular symbol resembles the generating series of MZV, which is the Drinfeld associator. Let us recall that the Drinfeld associator is a generating series of iterated integrals of the type J_a^b associated to the connection

$$\nabla = d - A\frac{dx}{x} - B\frac{dx}{1-x}$$

on $Y_{\Gamma(2)} = \mathbb{P}^1 - \{0, 1, \infty\}$. One can think of $Y_{\Gamma(2)}$ as the modular curve associated to the congruence subgroup $\Gamma(2)$ of $\mathrm{SL}_2(\mathbb{Z})$. Then the differential forms dx/x and $dx/(1-x)$ are Eisenstein series of weight 2 on the modular curve $Y_{\Gamma(2)}$.

Relations between MZV and modular forms have been examined by many authors. For example, Goncharov [2001b; 2001c] considered a mysterious relation between MZV (multiple zeta values) of given weight and depth 3 and the cohomology of $\mathrm{GL}_3(\mathbb{Z})$, which is closely related to the cohomology of $\mathrm{SL}_3(\mathbb{Z})$. In the pursuit of such a relation in depth 4, Goncharov suggested and the current author computed the group cohomology of $\mathrm{GL}_4(\mathbb{Z})$ with coefficients in a family of representations [Horozov 2014a]. Another relation between modular forms and MZV is presented in [Gangl et al. 2006].

Similarly to Manin's approach, we explore relations between the noncommutative Hilbert modular symbols and multiple Dedekind zeta values (see [Horozov 2014b]). Let us recall multiple Dedekind zeta values. Let each of $C_1, \ldots, C_m$ be a suitable subset of the ring of integers $\mathbb{O}_K$ of a number field K. We call each of $C_1, \ldots, C_m$ a cone. Then *multiple Dedekind zeta values* are defined as

$$\zeta_{K;C_1,\ldots,C_m}(k_1,\ldots,k_m) = \sum_{\alpha_i \in C_i} \frac{1}{N(\alpha_1)^{k_1} N(\alpha_1+\alpha_2)^{k_2} \cdots N(\alpha_1+\cdots+\alpha_m)^{k_m}}.$$

The connection between noncommutative Hilbert modular symbols and multiple Dedekind zeta values lies both in the similarities in the infinite sum formulas and in the definition in terms of iterated integrals on membranes (see [Horozov 2014b]).

We consider a noncommutative Hilbert modular symbol of type $\boldsymbol{b}$ over one diangle and compare it with (multiple) Dedekind zeta values with summation over one discrete cone [Horozov 2014b]. However, in this case the two series look very different. We obtain that the multiple L-values are noncommutative modular symbols defined as J evaluated at an infinite union of diangles. We obtain that such L-values are very similar to the sum of multiple Dedekind zeta values, in the same way that the integrals in Manin's noncommutative modular symbol are similar to the multiple zeta values (MZV). Then the sum of the multiple Dedekind zeta values is over an infinite union of cones. The idea of considering cones originated in [Zagier 1976] and more generally in [Shintani 1976].

Classical or commutative modular symbols for $\mathrm{SL}_3(\mathbb{Z})$ and $\mathrm{SL}_4(\mathbb{Z})$ were constructed in [Ash and Borel 1990] and [Gunnells 2000a]. For $\mathrm{GL}_2(\mathbb{O}_K)$, where K is a real quadratic field, Gunnells and Yasaki [2008] defined a modular symbol based on Voronoi decomposition of a fundamental domain, in order to compute the third cohomology group of $\mathrm{GL}_2(\mathbb{O}_K)$. (For the Hilbert modular group $\mathrm{SL}_2(\mathbb{O}_K)$ one may consult [Bruinier et al. 2008; Freitag 1990].) In contrast, here we use a geodesic triangulation of $\mathbb{H}^2/\,\mathrm{SL}_2(\mathbb{O}_K)$. We are interested mostly in 2-cells, whose boundaries are geodesics. One of the (commutative) symbols that we define here resembles combinatorially the symplectic modular symbol of [Gunnells 2000b]. However, the meanings of the two types of symbols and their approaches are different.

There are several different directions for further work on Hilbert modular symbols. First of all, the commutative Hilbert modular symbols behave well with respect to Hecke operators. It will be interesting to extend the Hecke operators to cases of higher equal weights (k, k). To apply Hecke operators to Hilbert modular groups one either assumes a trivial narrow class group or one has to work with adeles. Another possible continuation of the current work is to extend commutative Hilbert modular symbols to the adelic setting. Then, one may try to extend these properties — higher equal weight cusp forms and Hecke operators in the adelic setting — to the noncommutative Hilbert modular symbols. Hopefully, the abelian Hilbert modular symbol would lead to computational tools for cohomology of some Hilbert modular groups with coefficients in various representations.

For the noncommutative Hilbert modular symbols we expect that some of the continuations would be establishing the 2-categorical framework that define nonabelian 2-cohomology sets. This work would also have applications to noncommutative reciprocity laws on algebraic surfaces. In dimension 1, we have a noncommutative reciprocity law as a reciprocity law for a generating series of iterated path integrals on a complex curve [Horozov 2011]. In dimension 2 we have proven both the Parshin reciprocity and a new reciprocity for a 4-function local symbols [Horozov 2014c] defined by the author, which are particular cases in the generating series. A 2-categorical second cohomology set would capture algebraically the generating series of iterated integrals on membranes needed for the general reciprocity on algebraic surfaces.

Finally, we expect that the (non)commutative Hilbert modular symbols would be useful for studying L-functions and multiple L-functions together with their special values.

2. Manin's noncommutative modular symbol

In this section we recall the definition and main properties of Manin's [2006] noncommutative modular symbol. In this paper, Manin uses iterated path integrals on a modular curve and on its universal cover — the upper half-plane. Our main constructions are parallel to some extent to Manin's approach, and for that reason we recall it below. However, instead of iterated path integrals we introduce a new tool — iterated integrals on membranes (see Section 3). Only this notion is adequate for studying noncommutative Hilbert modular symbols, by generalizing the iteration process to higher dimensions.

2A. ***Iterated path integrals.*** Here we recall iterated path integrals (see also [Parshin 1966; Chen 1977; Goncharov 2001a; Manin 2006]). In Section 3, we generalize them to iterated integrals over membranes.

Definition 2.1. Let $\omega_1, \dots, \omega_m$ be m holomorphic 1-forms on $\mathbb{H}^1 \cup \mathbb{P}^1(\mathbb{Q})$, the upper half-plane together with the cusps. Let

$$g : [0, 1] \to \mathbb{H}^1 \cup \mathbb{P}^1(\mathbb{Q}),$$

be a piecewise smooth path. We define an iterated integral

$$\int_g \omega_1 \cdots \omega_m = \int \cdots \int_{0<t_1<t_2<\cdots<t_m<1} g^*\omega_1(t_1) \wedge \cdots \wedge g^*\omega_m(t_m).$$

Let $X_1, \dots, X_m$ be formal variables. Consider the differential equation

$$dF(\Omega) = F(\Omega)(X_1\omega_1 + \cdots + X_n\omega_m) \tag{1}$$

with values in the associative but noncommutative ring of formal power series in the noncommuting variables $X_1, \ldots, X_m$ over the ring of holomorphic functions on the upper half-plane. There is a unique solution with initial condition $F(\Omega)(g(0)) = 1$; that is, equal to 1 at the starting point $g(0)$. Then at the end of the path, that is, at the point $g(1)$, $F(\Omega)$ has the value

$$F_g(\Omega)=1+\sum_{i=1}^{m} X_i \int_g \omega_i+\sum_{i,j=1}^{m} X_i X_j \int_g \omega_i\omega_j+\sum_{i,j,k=1}^{m} X_i X_j X_k \int_g \omega_i\omega_j\omega_k+\cdots. \quad (2)$$

Using the solution (2) to (1), we prove the following theorem:

Theorem 2.2. *Let g_1 and g_2 be two paths such that the end of g_1 (i.e., $g_1(1)$) is equal to the beginning of g_2 (i.e., $g_2(0)$). Let g_1g_2 denote the concatenation of g_1 and g_2. Then*

$$F_{g_1g_2}(\Omega) = F_{g_1}(\Omega)F_{g_2}(\Omega).$$

Proof. The left-hand side is the value of the solution of the linear first-order ordinary differential equation at the point $g_2(1)$. From the uniqueness of the solution, we have that the solution along g_2 gives the same result, when the initial condition at $g_2(0)$ is $F_{g_1}(\Omega)$. That result is $F_{g_1}(\Omega)F_{g_2}(\Omega)$. □

The same result can be proven via product formula for iterated integrals. We need this alternative proof in order to generalize to higher dimensions.

Lemma 2.3 (product formula). *Let $\omega_1, \ldots, \omega_m$ be holomorphic 1-forms on $\mathbb{C}$ and g_1, g_2 two paths such that the end of g_1 is the beginning of g_2, that is, $g_1(1) = g_2(0)$. As before we denote by g_1g_2 the concatenation of the paths g_1 and g_2. Then*

$$\int_{g_1g_2} \omega_1 \cdots \omega_m = \sum_{i=0}^{m} \int_{g_1} \omega_1 \cdots \omega_i \int_{g_2} \omega_{i+1} \cdots \omega_m.$$

Proof. Let $g_1 : [0, 1] \to \mathbb{C}$ and let $g_2 : [1, 2] \to \mathbb{C}$. We consider the concatenation g_1g_2 to be a map $g_1g_2 : [0, 2] \to \mathbb{C}$ whose restriction to the interval $[0, 1]$ gives the path g_1 and whose restriction to the interval $[1, 2]$ gives g_2. From Definition 2.1, we have that

$$\int_{g_1g_2} \omega_1 \cdots \omega_m = \int \cdots \int_{0<t_1<\cdots<t_m<2} (g_1g_2)^*\omega_1(t_1) \wedge \cdots \wedge (g_1g_2)^*\omega_m(t_m).$$

In the domain of integration $0 < t_1 < \cdots < t_m < 2$ insert the number 1. Geometrically, we cut the simplex $0 < t_1 < \cdots < t_m < 2$ into a disjoint union of products of pairs of simplices such that $t_k \in [0, 1]$ for $k \le i$ and $t_k \in [1, 2]$ for $k > i$. Thus, the union is over distinct values of i for $i = 0, \ldots, m$. And for each fixed i the two simplices are $0 < t_1 < \cdots < t_i < 1$ and $1 < t_{i+1} < \cdots < t_m < 2$. Then we have

$$\begin{aligned}
\int_{g_1g_2} \omega_1 \cdots \omega_m &= \sum_{i=0}^{n} \int \cdots \int_{\substack{0<t_1<\cdots<t_i<1 \\ 1<t_{i+1}<\cdots<t_m<2}} (g_1g_2)^*\omega_1(t_1) \wedge \cdots \wedge (g_1g_2)^*\omega_m(t_m) \\
&= \sum_{i=0}^{n} \Bigl(\int \cdots \int_{0<t_1<\cdots<t_i<1} g_1^*\omega_1(t_1) \wedge \cdots \wedge g_1^*\omega_i(t_i) \Bigr) \\
&\qquad \times \Bigl(\int \cdots \int_{1<t_{i+1}<\cdots<t_m<2} g_2^*\omega_{i+1}(t_{i+1}) \wedge \cdots \wedge g_2^*\omega_m(t_m) \Bigr) \\
&= \sum_{i=0}^{m} \int_{g_1} \omega_1 \cdots \omega_i \int_{g_2} \omega_{i+1} \cdots \omega_m. \qquad \square
\end{aligned}$$

Definition 2.4. The set of all shuffles $\mathrm{sh}(i, j)$ is a subset of all permutations σ of the set $\{1, 2, \ldots, i + j\}$ such that

$$\sigma(1) < \cdots < \sigma(i)$$

and

$$\sigma(i+1) < \cdots < \sigma(i+j).$$

Such a permutation σ is called a shuffle.

Lemma 2.5 (shuffle relation). *Let $\omega_1, \ldots, \omega_m$ be holomorphic 1-forms on $\mathbb{C}$ and let g be a path. Then*

$$\int_g \omega_1 \cdots \omega_i \int_g \omega_{i+1} \cdots \omega_m = \sum_{\sigma \in \mathrm{sh}(i,m-i)} \int_g \omega_{\rho(1)} \cdots \omega_{\rho(m)},$$

where $\mathrm{sh}(i, j)$ *is the set of shuffles from* Definition 2.4.

2B. ***Manin's noncommutative modular symbol.*** Now let g be a geodesic connecting two cusps a and b in the completed upper half-plane $\mathbb{H}^1 \cup \mathbb{P}^1(\mathbb{Q})$. Let $\Omega = \{f_1\,dz, \ldots, f_m\,dz\}$ be a finite set of holomorphic forms with respect to a congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ such that $f_1, \ldots, f_m$ are cusp forms of weight 2. Let

$$J_a^b = F_g(\Omega).$$

As a reformulation of Theorem 2.2, we obtain:

Lemma 2.6. $$J_a^b J_b^c = J_a^c.$$

The following is a direct consequence:

Corollary 2.7. $$J_b^a = (J_a^b)^{-1}.$$

Now we are ready to define Manin's noncommutative modular symbol. Note that there is a natural action of Γ on J_a^b. If $\gamma \in \Gamma$ then γJ_a^b is defined as $J_{\gamma a}^{\gamma b}$. If $f_1, \ldots, f_m$ are cusp forms of weight 2, then $\omega_1 = f_1\, dz, \ldots, \omega_m = f_m\, dz$ are forms of weight 0, that is, they are invariant forms with respect to the group Γ. Then

$$\gamma J_a^b = F_{\gamma g}(\omega_1, \ldots, \omega_m) = F_g(g^*\omega_1, \ldots, g^*\omega_m) = F_g(\omega_1, \ldots, \omega_m) = J_a^b.$$

Let Π be the subgroup of invertible elements of $\mathbb{C}\langle\langle X_1, \ldots, X_m\rangle\rangle$ with constant term 1. We extend the action of Γ on J_a^b to a trivial action of Γ on Π.

Following Manin, we present the key theorem and definition for the noncommutative modular symbol:

Theorem 2.8. *Let*

$$c_a^1(\gamma) = J_{\gamma a}^a.$$

Then c_a^1 *represent a cohomology class in* $H^1(\Gamma, \Pi)$, *independently of the base point* a.

Proof. First, c_a^1 is a cocycle:

$$dc_a^1(\beta, \gamma) = J_{\beta a}^a(\beta \cdot J_{\gamma a}^a)(J_{\beta\gamma a}^a)^{-1} = J_{\beta a}^a J_{\beta\gamma a}^{\beta a} J_a^{\beta\gamma a} = 1.$$

Second, c_a^1 and c_b^1 are homologous:

$$c_a^1(\gamma) = J_{\gamma a}^a = J_b^a J_{\gamma b}^b J_{\gamma a}^{\gamma b} = J_b^a c_b^1(\gamma)(\gamma \cdot J_b^a)^{-1}. \qquad \square$$

Definition 2.9. A noncommutative modular symbol is a nonabelian cohomology class in $H^1(\Gamma, \Pi)$, with representative

$$c_a^1(\gamma) = J_{\gamma a}^a,$$

3. Iterated integrals on membranes

Iterated integrals on membranes are a higher-dimensional analogue of iterated path integrals. This technical tool was used in [Horozov 2014b] for constructing multiple Dedekind zeta values and in [Horozov 2014c] for proving new and classical reciprocity laws on algebraic surfaces. It appeared first in the preprint [Horozov 2006] for the purpose of noncommutative Hilbert modular symbols.

3A. ***Definition and properties.*** Let $\mathbb{H}^1$ be the upper half-plane. Let $\mathbb{H}^2$ be a product of two upper half-planes. We are interested in the action of $\mathrm{GL}_2(K)$, where K is a real quadratic field. This group acts on $\mathbb{H}^2$ by linear fractional transformations. It is convenient to introduce cusp points $\mathbb{P}^1(K)$ as boundary points of $\mathbb{H}^2$.

Let $\omega_1, \dots, \omega_m$ be holomorphic 2-forms on $\mathbb{H}^2$ which are continuous at the cusps $\mathbb{P}^1(K)$. Let

$$g : [0, 1]^2 \to \mathbb{H}^2 \cup \mathbb{P}^1(K)$$

be a continuous map which is smooth almost everywhere. Denote by F^1 and F^2 the following coordinatewise foliations: for any $a \in [0, 1]$, define the leaves

$$F^1_a = \{(t_1, t_2) \in [0, 1]^2 \mid t_1 = a\} \quad \text{and} \quad F^2_a = \{(t_1, t_2) \in [0, 1]^2 \mid t_2 = a\}.$$

Definition 3.1. We call the above map $g : [0, 1]^2 \to \mathbb{H}^2 \cup \mathbb{P}^1(K)$ a membrane on $\mathbb{H}^2$ if it is a continuous and piecewise differentiable map such that $g(F^1_a)$ and $g(F^2_a)$ belong to a finite union of holomorphic curves in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$ for all constants a.

Similarly, we define a membrane of a Hilbert modular variety. Let $\omega_1, \dots, \omega_m$ be holomorphic 2-forms on $Y_\Gamma = \mathbb{H}^2/\Gamma$ which are continuous at the cusps $\mathbb{P}^1(K)/\Gamma$. Let

$$g : [0, 1]^2 \to X_\Gamma$$

be a continuous map which is smooth almost everywhere, where $X_\Gamma = \mathbb{H}^2 \cup \mathbb{P}^1(K)$. Let $f_i : X_\Gamma \to \mathbb{P}^1(\mathbb{C})$ for $i = 1, 2$ be two algebraically independent rational functions on the Hilbert modular surface X_Γ. Denote by F^1 and F^2 the following coordinatewise foliations: for any $a \in [0, 1]$, define the leaves

$$F^1_a = \{(t_1, t_2) \in [0, 1]^2 \mid t_1 = a\} \quad \text{and} \quad F^2_a = \{(t_1, t_2) \in [0, 1]^2 \mid t_2 = a\}.$$

Let also

$$P^1_x = \{P \in X_\Gamma \mid f_1(P) = x\} \quad \text{and} \quad P^2_x = \{P \in X_\Gamma \mid f_2(P) = x\}.$$

Definition 3.2. We call the above map $g : [0, 1]^2 \to X_\Gamma$ a membrane on X_Γ if it is a continuous and piecewise differentiable map such that for each a there are x_1 and x_2 such that $g(F^1_a) \subset P^1_{x_1}$ and $g(F^2_a) \subset P^2_{x_2}$.

We define three types of iterated integrals over membranes — type $\boldsymbol{a}$, type $\boldsymbol{b}$ and type $\boldsymbol{c}$. Type $\boldsymbol{a}$ consists of linear iterations, while type $\boldsymbol{b}$ is more general and involves permutations. Type $\boldsymbol{a}$ is less general, but more intuitive. The advantage of type $\boldsymbol{b}$ is that it satisfies integral shuffle relation (Theorem 3.21). In other words a product of two integrals of type $\boldsymbol{b}$ can be expresses as a finite sum of iterated integrals over membranes of type $\boldsymbol{b}$. However, one might not be able to express a product of two integrals of type $\boldsymbol{a}$ as a sum of finitely many integrals of type $\boldsymbol{a}$. Both type $\boldsymbol{a}$ and type $\boldsymbol{b}$ are defined on a product of two upper half-planes. Type $\boldsymbol{c}$

is defined on a Hilbert modular surface; that is, on a quotient of a product of upper half-planes by an arithmetic group which is commensurable to $\mathrm{SL}_2(\mathbb{O}_K)$. Type $\boldsymbol{c}$ also satisfies a shuffle product, that is, a product of two integrals of this type can be expresses a finite sum of such integrals.

Definition 3.3 (type $\boldsymbol{a}$, ordered iteration over membranes). Let

$$g : [0, 1]^2 \to \mathbb{H}^2 \cup \mathbb{P}^1(K)$$

be a membrane on $\mathbb{H}^2 \cup \mathbb{P}^1(K)$. Then define

$$\int_g \omega_1 \cdots \omega_m = \int_D \textstyle\bigwedge_{j=1}^m g^* \omega_i(t_{1,j}, t_{2,j}),$$

where

$$D = \left\{(t_{1,1}, \ldots, t_{2,m}) \in [0, 1]^{2m} \mid 0 \le t_{1,1} \le \cdots \le t_{1,m} \le 1,\ 0 \le t_{2,1} \le \cdots \le t_{2,m} \le 1\right\}.$$

Definition 3.4 (type $\boldsymbol{b}$, two permutations). Let

$$g : [0, 1]^2 \to \mathbb{H}^2 \cup \mathbb{P}^1(K)$$

be a membrane on $\mathbb{H}^2 \cup \mathbb{P}^1(K)$, and let ρ_1, ρ_2 be two permutations of the set $\{1, 2, \ldots, m\}$. Then define

$$\int_g^{\rho_1,\rho_2} \omega_1 \cdots \omega_m = \int_D \textstyle\bigwedge_{j=1}^m g^* \omega_j(t_{1,\rho_1(j)}, t_{2,\rho_2(j)}),$$

where

$$D = \left\{(t_{1,1}, \ldots, t_{2,m}) \in [0, 1]^{2m} \mid \ \le t_{1,1} \le \cdots \le t_{1,m} \le 1,\ 0 \le t_{2,1} \le \cdots \le t_{2,m} \le 1\right\}.$$

Definition 3.5 (type $\boldsymbol{c}$, two permutations). Let

$$g : [0, 1]^2 \to X_\Gamma$$

be a membrane on the Hilbert modular surface $X_\Gamma = (\mathbb{H}^2 \cup \mathbb{P}^1(K))/\Gamma$, and let ρ_1, ρ_2 be two permutations of the set $\{1, 2, \ldots, m\}$. Then define

$$\int_g^{\rho_1,\rho_2} \omega_1 \cdots \omega_m = \int_D \textstyle\bigwedge_{j=1}^m g^* \omega_j(t_{1,\rho_1(j)}, t_{2,\rho_2(j)}),$$

where

$$D = \left\{(t_{1,1}, \ldots, t_{2,m}) \in [0, 1]^{2m} \mid 0 \le t_{1,1} \le \cdots \le t_{1,m} \le 1,\ 0 \le t_{2,1} \le \cdots \le t_{2,m} \le 1\right\}.$$

Examples (iterated integrals of type b). Let $\alpha_i(t_1, t_2) = g^*\omega_i(t_1, t_2)$. Denote by (1) the trivial permutation and by (1 2) the permutation exchanging 1 and 2.

(1) The four diagrams

$$\begin{array}{c|c|c|}
\cline{2-3}
t_{2,2} & & \alpha_2(t_{1,2}, t_{2,2}) \\
\cline{2-3}
t_{2,1} & \alpha_1(t_{1,1}, t_{2,1}) & \\
\cline{2-3}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{t_{1,1}} & \multicolumn{1}{c}{t_{1,2}}
\end{array}
\qquad
\begin{array}{c|c|c|}
\cline{2-3}
t_{2,2} & \alpha_2(t_{1,1}, t_{2,2}) & \\
\cline{2-3}
t_{2,1} & & \alpha_1(t_{1,2}, t_{2,1}) \\
\cline{2-3}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{t_{1,1}} & \multicolumn{1}{c}{t_{1,2}}
\end{array}$$

$$\begin{array}{c|c|c|}
\cline{2-3}
t_{2,2} & & \alpha_1(t_{1,2}, t_{2,2}) \\
\cline{2-3}
t_{2,1} & \alpha_2(t_{1,1}, t_{2,1}) & \\
\cline{2-3}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{t_{1,1}} & \multicolumn{1}{c}{t_{1,2}}
\end{array}
\qquad
\begin{array}{c|c|c|}
\cline{2-3}
t_{2,2} & \alpha_1(t_{1,1}, t_{2,2}) & \\
\cline{2-3}
t_{2,1} & & \alpha_2(t_{1,2}, t_{2,1}) \\
\cline{2-3}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{t_{1,1}} & \multicolumn{1}{c}{t_{1,2}}
\end{array}$$

correspond, respectively, to the integrals

$$\int_g^{(2),(1)} \omega_1 \cdot \omega_2, \qquad \int_g^{(12),(1)} \omega_1 \cdot \omega_2,$$
$$\int_g^{(12),(12)} \omega_1 \cdot \omega_2, \qquad \int_g^{(1),(12)} \omega_1 \cdot \omega_2.$$

(2) The diagram

$$\begin{array}{c|c|c|c|}
\cline{2-4}
t_{2,3} & & & \alpha_2(t_{1,3}, t_{2,3}) \\
\cline{2-4}
t_{2,2} & \alpha_2(t_{1,1}, t_{2,2}) & & \\
\cline{2-4}
t_{2,1} & & \alpha_1(t_{1,2}, t_{2,1}) & \\
\cline{2-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{t_{1,1}} & \multicolumn{1}{c}{t_{1,2}} & \multicolumn{1}{c}{t_{1,3}}
\end{array}$$

corresponds to the integral

$$\int_g^{(12),(1)} \omega_1 \cdot \omega_2 \cdot \omega_3.$$

Remark 3.6. Let us give more intuition for Definition 3.4. Each of the differential forms $g^*\omega_1, \dots, g^*\omega_m$ has two arguments. Consider the set of first arguments for each of the differential forms $g^*\omega_1, \dots, g^*\omega_m$. They are ordered as

$$0 < t_{1,1} < t_{1,2} < \cdots < t_{1,m} < 1 \tag{3}$$

(they are the coordinates of the domain D). Since $g^*\omega_j$ depends on $t_{1,\rho_1(j)}$, we have that $t_{1,k}$ is an argument of $g^*\omega_{\rho_1^{-1}(k)}$, where $k = \rho_1(j)$. Then we can order the differential forms $g^*\omega_1, \dots, g^*\omega_m$ according to the order of their first arguments given by the inequalities (3), which is

$$g^*\omega_{\rho_1^{-1}(1)}, g^*\omega_{\rho_1^{-1}(2)}, \dots, g^*\omega_{\rho_1^{-1}(m)}.$$

Similarly, we can order the differential forms $g^*\omega_1, \dots, g^*\omega_m$ with respect to the order of their second arguments:

$$g^*\omega_{\rho_2^{-1}(1)}, g^*\omega_{\rho_2^{-1}(2)}, \dots, g^*\omega_{\rho_2^{-1}(m)}.$$

We call the first ordering *horizontal* and the second ordering *vertical*.

Now we are going to examine the homotopy of a domain of integration and what that reflects about the integral. Let $g_s : [0, 1]^2 \to \mathbb{H}^2 \cup \mathbb{P}^1(K)$ be a family of membranes such that $g_s(0, 0) = \infty$ and $g_s(1, 1) = 0$. Assume that the parameter s is in the interval $[0, 1]$.

Let $h(s, t_1, t_2) = g_s(t_1, t_2)$ be a homotopy between g_0 and g_1. Let

$$G_s : [0, 1]^{2m} \to (\mathbb{H}^2 \cup \mathbb{P}^1(K))^m$$

be the map

$$G_s(t_{1,1}, \dots, t_{2,m}) = \big(g_s(t_{1,\sigma_1(1)}, t_{2,\sigma_2(1)}), g_s(t_{1,\sigma_1(2)}, t_{2,\sigma_2(2)}), \dots, g_s(t_{1,\sigma_1(m)}, t_{2,\sigma(m)})\big).$$

Let H be the induced homotopy between G_0 and G_1, defined by

$$H(s, t_{1,1}, \dots, t_{2,m}) = G_s(t_{1,1}, \dots, t_{2,m}).$$

We define diagonals in the domain $D \subset (0, 1)^{2m}$, where

$$D = \big\{(t_{1,1}, t_{2,1}, \dots, t_{1,m}, t_{2,m}) \in (0, 1)^{2m} \mid 0 \le t_{1,1} \le t_{1,2} \le \cdots \le t_{1,m} \le 1 \\ \text{and}\, 0 \le t_{2,1} \le t_{2,2} \le \cdots \le t_{2,m} \le 1\big\}.$$

We define $D_{1,k}$ for $k = 0, \dots, m$ by $D_{1,0} = D|_{t_{1,1}=0}$, $D_{1,k} = D|_{t_{1,k}=t_{1,k+1}}$ for $k = 1, \dots, m-1$ and $D_{1,m} = D|_{t_{1,m}=1}$. Similarly, we define $D_{2,k}$ for $k = 0, \dots, m$ by $D_{2,0} = D|_{t_{2,1}=0}$, $D_{2,k} = D|_{t_{2,k}=t_{2,k+1}}$ for $k = 1, \dots, m-1$ and $D_{2,m} = D|_{t_{2,m}=1}$.

For iterated integrals of types $\boldsymbol{a}$ and $\boldsymbol{b}$, we define diagonals in $V = (\mathbb{H}^2 \cup \mathbb{P}^1(K))^m$. We denote a generic coordinate of $V = (\mathbb{H}^2 \cup \mathbb{P}^1(K))^m$ by $(z_{1,1}, z_{2,1}, \dots, z_{1,m}, z_{2,m})$ For $k = 1, \dots, m-1$, let $V_{1,k} = V|_{z_{1,k}=z_{1,k+1}}$. Let also $V_{1,0} = V|_{z_{1,1}=0}$ and $V_{1,m} = V|_{z_{1,m}=1}$. Similarly, for $k = 1, \dots, m-1$, let $V_{2,k} = V|_{z_{2,k}=z_{2,k+1}}$. Let also $V_{2,0} = V|_{z_{2,1}=0}$ and $V_{2,m} = V|_{z_{2,m}=1}$.

For iterated integrals of type $\boldsymbol{c}$, we define "diagonals" as fibers product of schemes corresponding to certain varieties (for fiber products of schemes one may look at the book [Hartshorne 1977]). Occasionally, it will be more natural to realize multiple fiber products as finite limits in the category of schemes of finite type over $\mathbb{C}$. Let $X_{i,j} = X_\Gamma$ for $i, j = 1, \dots, n$. Let V be the universal scheme (finite limit) that maps to X_{ij} for each i and j as a part of a commutative diagram. The commutative diagram is defined as follows: $X_{i,j}$ and $X_{i+1,j}$ both map to $\mathbb{P}^1(\mathbb{C})$ via the morphism f_1 for $1 \le i \le n-1$ and all j, and $X_{i,j}$ and $X_{i,j+1}$ both map to $\mathbb{P}^1(\mathbb{C})$ via the morphism f_2 for $1 \le j \le n-1$ and all i. Define the following:

- Let $V_{1,0}$ be the subscheme of V defined by putting $\mathbb{P}^1(\mathbb{C})$ in the place of $X_{1,j}$, so that $f_1 : X_{1,j} \to \mathbb{P}^1(\mathbb{C})$ is replaced by the identity map and the corresponding $f_2 : X_{1,j} \to \mathbb{P}^1(\mathbb{C})$ is deleted.
- Let $V_{2,0}$ be the subscheme of V defined by putting $\mathbb{P}^1(\mathbb{C})$ in the place of $X_{i,1}$, so that $f_2 : X_{1,j} \to \mathbb{P}^1(\mathbb{C})$ is replaced by the identity map and the corresponding $f_1 : X_{1,j} \to \mathbb{P}^1(\mathbb{C})$ is deleted.
- Let $V_{1,n}$ be the subscheme of V defined by putting $\mathbb{P}^1(\mathbb{C})$ in the place of $X_{n,j}$, so that $f_1 : X_{n,j} \to \mathbb{P}^1(\mathbb{C})$ is replaced by the identity map and the corresponding $f_2 : X_{n,j} \to \mathbb{P}^1(\mathbb{C})$ is deleted.
- Let $V_{2,n}$ be the subscheme of V defined by putting $\mathbb{P}^1(\mathbb{C})$ in the place of $X_{i,n}$, so that $f_2 : X_{n,j} \to \mathbb{P}^1(\mathbb{C})$ is replaced by the identity map and the corresponding $f_1 : X_{n,j} \to \mathbb{P}^1(\mathbb{C})$ is deleted.
- Additionally, let $V_{1,i}$ be the subscheme of V obtained by replacing each factor $X_{i,j} \times_{\mathbb{P}^1(\mathbb{C})} X_{i+1,j}$ by the corresponding diagonal for fixed i and for all j.
- Finally, let $V_{2,j}$ be the subscheme of V obtained by replacing each factor $X_{i,j} \times_{\mathbb{P}^1(\mathbb{C})} X_{i,j+1}$ by the corresponding diagonal for fixed j and all i.

Theorem 3.7 (homotopy invariance theorem I). *The iterated integrals on membranes from Definition 3.4 (of type* $\boldsymbol{b}$*) are homotopy-invariant with respect to homotopies that preserve the boundary of the membrane.*

Proof. Let

$$\Omega = \textstyle\bigwedge_{j=1}^m \omega_j(z_{1,\sigma_1(j)}, z_{2,\sigma_2(j)}).$$

Note that Ω is a closed form, since ω_i is a form of top dimension. By Stokes' theorem, we have

$$0 = \int_{s=0}^{s=1} \int_D H^* d\Omega \tag{4}$$

$$= \int_D G_1^*\Omega - \int_D G_0^*\Omega \tag{5}$$

$$\pm \int_{s=0}^{s=1} \sum_{k=1}^{m-1} \left(\int_{D_{1,k}} \pm \int_{D_{2,k}} \right) H^*\Omega \tag{6}$$

$$\pm \int_{s=0}^{s=1} \left(\int_{D_{1,0}} \pm \int_{D_{2,0}} \right) H^*\Omega \tag{7}$$

$$\pm \int_{s=0}^{s=1} \left(\int_{D_{1,m}} \pm \int_{D_{2,m}} \right) H^*\Omega. \tag{8}$$

We want to show that the difference in the terms in (5) is zero. It is enough to show that each of the terms (6), (7) and (8) are zero. If $z_{1,k} = z_{1,k+1}$ for types ***a*** and ***b*** (or on $V_{1,k}$ for type ***c***), then the wedge of the corresponding differential forms will vanish. Thus the terms in (6) are zero. If $z_1 = 0$ then $dt_1 = 0$, defined via the pullback H^*. Then the terms (7) are equal to zero. Similarly, we obtain that the last integral (8) vanishes. □

Let A be a 2-dimensional manifold with corners in $[0, 1]^2$. We recall the domain of integration

$$D = \left\{ (t_{1,1}, \ldots, t_{2,m}) \in [0, 1]^{2m} \mid 0 \leq t_{1,1} \leq \cdots \leq t_{1,m} \leq 1,\ 0 \leq t_{2,1} \leq \cdots \leq t_{2,m} \leq 1 \right\}.$$

Let us define

$$A^D = \{ (t_{1,1}, \ldots, t_{2,m}) \in D \mid (t_{1,i}, t_{2,j}) \in A \text{ for } i, j = 1, \ldots, m \}$$

Let ρ_1 and ρ_2 be two permutations of m elements. We define a function on A^D

$$G(t_{1,1}, \ldots, t_{2,m}) = \left(g(t_{1,\rho_1(1)}, t_{2,\rho_2(1)}), g(t_{1,\rho_1(2)}, t_{2,\rho_2(2)}), \ldots, g(t_{1,\rho_1(m)}, t_{2,\rho(m)}) \right).$$

Recall that

$$\Omega = \bigwedge_{j=1}^m \omega_j (z_{1,\rho_1(j)}, z_{2,\rho_2(j)}).$$

Definition 3.8. With the above notation, we define an iterated integral over a membrane of type ***b*** restricted to a domain $U = g(A)$ by

$$^b\!\!\int_{g,U}^{\rho_1,\rho_2} \omega_1 \cdots \omega_m = \int_{A^D} G^*\Omega.$$

Now we are going to define iterated integrals of type ***c***:

Definition 3.9. Let $\Omega_0 = \bigwedge_{i,j=1}^m \Omega_{i,j}$, where $\Omega_{i,j} = \omega_i \delta_{i,j}$ on $X_{i,i} \equiv X_\Gamma$ and where $\Omega_{i,j} = 1$ for $i \neq j$. Let $\mathrm{in} : X \to \prod_{i,j=1}^n X_{i,j}$ be the inclusion of the finite limit into the product of the schemes $X_{i,j}$. Let $\Omega = \mathrm{in}^* \Omega_0$.

With this definition of Ω, we define iterated integrals of type $\boldsymbol{c}$ restricted to a domain $U = g(A)$ by

$$^{c}\!\int_{g,U}^{\rho_1,\rho_2} \omega_1 \cdots \omega_m = \int_{A^D} G^* \Omega.$$

Let A_1 and A_2 be two manifolds with corners which are subsets of $[0, 1]^2$, with a common component of the boundary. Let $A = A_1 \cup A_2$. Let s be a map of sets with values 1 or 2:

$$s : \{1, \dots, m\} \to \{1, 2\}.$$

We define a certain subset A_s^D of A^D as follows: consider the image of the map G. It has m coordinates. The first coordinate, $g(t_{1,\rho_1(1)}, t_{2,\rho_2(1)})$, will be restricted to the set $A_{s(1)}$. The second coordinate, $g(t_{1,\rho_1(2)}, t_{2,\rho_2(2)})$, will be restricted to $A_{s(2)}$, and so on, and the last m coordinate $g(t_{1,\rho_1(m)}, t_{2,\rho_2(m)})$ will be restricted to $A_{s(m)}$. Formally, this can be written as

$$A_s^D = \{(t_{1,1}, \dots, t_{2,m}) \in A^D \mid (t_{1,\rho_1(i)}, t_{2,\rho_2(i)}) \in A_{s(i)} \text{ for } i = 1, \dots, m\}.$$

Note that the image of the map s is 1 or 2.

Definition 3.10. With the above notation, we define an iterated integral of type $\boldsymbol{b}$ or $\boldsymbol{c}$ over two domains U_1 and U_2, where $U_i = g(A_i)$ and $U = U_1 \cup U_2$, by

$$\int_{g,U,s}^{\rho_1\rho_2} \omega_1 \cdots \omega_m = \int_{A_s^D} G^* \Omega. \tag{9}$$

For type $\boldsymbol{b}$ we have that U is in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$ and for type $\boldsymbol{c}$ we have that U is in $X_\Gamma = (\mathbb{H}^2 \cup \mathbb{P}^1(K))/\Gamma$.

Again we examine the homotopy of iterated integrals on membranes. Now we restrict the domain of integration to a manifold with corners A that is a subset of $[0, 1]^2$. Assume that for the boundary of a domain A, denoted by ∂A, we have that $g(\partial A)$ belongs to a finite union of complex analytic curves in $\mathbb{H}^2$ for type $\boldsymbol{b}$ and in X_Γ for type $\boldsymbol{c}$. We call the minimal union of complex analytic (holomorphic) curves such that $g(\partial A)$ belongs to a finite union of complex analytic curves in $\mathbb{H}^2$ for type $\boldsymbol{b}$ and in X_Γ for type $\boldsymbol{c}$ the *complex boundary* of $g(\partial A)$.

Theorem 3.11 (homotopy invariance theorem II). *Iterated integrals over membranes are homotopy invariant with respect to homotopies that change the boundary ∂U of the domain of integration U so that the boundary varies on a finite union of complex analytic curves.*

Proof. Assume that $g_0(\partial A)$ and $g_1(\partial A)$ have the same complex boundary. Let h be a homotopy between g_0 and g_1, such that for each value of s we have that $h(s, \partial A)$ has the same complex boundary as $h(0, \partial A) = g_0(\partial A)$. Let $A \subset B$ be a strict inclusion of disks. Identify $B - A^\circ$ with $A \times [0, 1]$. Let $i : B - A^\circ \to [0, 1] \times \partial A$. Here A° is the interior of A and ∂A is the boundary of A. Let $\tilde{g}_0$ be a map from B to $\mathbb{H}^2$ such that $\tilde{g}_0(a) = g_0(a)$ for $a \in A$ and $\tilde{g}_0(b) \in h(i(b))$. Since the restriction of the pullback $(\tilde{g}_0^*\omega_i)|_{B-A} = 0$ is mapped to a finite union of complex curves, it vanishes. Therefore

$$\int_A g_0^*\Omega = \int_B \tilde{g}_0^*\Omega. \tag{10}$$

Let $\tilde{g}_1$ be a membrane from B defined by $\tilde{g}_1(a) = g_1(a)$ for $a \in A$ and $\tilde{g}_1(b) = \tilde{g}_1(a)$ for $i(b) = (s, a)$. (Note that $i(b) \in [0, 1] \times \partial A$.) Again,

$$\int_A g_1^*\Omega = \int_B \tilde{g}_1^*\Omega. \tag{11}$$

However, the boundary of B is mapped to the same set (pointwise) by both $\tilde{g}_0$ and $\tilde{g}_1$. Moreover, the homotopy between g_0 and g_1 extends to a homotopy between $\tilde{g}_0$ and $\tilde{g}_1$ that respects the inclusion into the complex boundary. Thus by Theorem 3.7, we have that

$$\int_B \tilde{g}_0^*\Omega = \int_B \tilde{g}_1^*\Omega.$$

Using (10) and (11), we complete the proof of this theorem. □

3B. ***Generating series.*** We are going to define two types of generating series — type ***a*** and type ***b***, corresponding to the iterated integrals on membranes of type ***a*** and type ***b***.

Definition 3.12 (type ***a***). Let A be a domain in $\mathbb{R}^2$. Let g be a membrane. Let $U = g(A) \subset \mathbb{H}^2$. And let $\omega_1, \dots, \omega_m$ be holomorphic 2-forms on $\mathbb{H}^2$. We define a generating series of type ***a*** by

$$J^a(U) = 1 + \sum_{k=1}^{\infty} \sum_{c:\{1,\dots,k\}\to\{1,\dots,m\}} X_{c(1)} \otimes \cdots \otimes X_{c(k)} \int_{g,U} \omega_{c(1)} \cdots \omega_{c(k)},$$

where $c : \{1, \dots, k\} \to \{1, \dots, m\}$ is a map of sets.

Consider a map of sets $c : \{1, \dots, k\} \to \{1, \dots, m\}$ and two permutations ρ_1, ρ_2 of $\{1, 2, \dots, k\}$. We call two triples (c', ρ_1', ρ_2') and $(c'', \rho_1'', \rho_2'')$ equivalent if they are in the same orbit of the permutation group S_k. That is, $(c'', \rho_1'', \rho_2'') \sim (c', \rho_1', \rho_2')$ if for some $\tau \in S_k$ we have $c'' = c'\tau^{-1}$, $\rho_1'' = \rho_1'\tau^{-1}$ and $\rho_2'' = \rho_2'\tau^{-1}$. Then for the equivalence class of a triple (c, ρ_1, ρ_2), we can associate a unique pair $(c \circ \rho_1, c \circ \rho_2)$ (which are precisely the indices of the X- and Y-variables in (12) and (13), respectively.) The reason for using such an equivalence is that the integral in (13) is invariant by the above action of $\tau \in S_k$ on the triple (c, ρ_1, ρ_2).

Definition 3.13 (ring R, values of the generating series). The values of the generation series of iterated integrals on membranes will be in a ring R, which we define as follows. Let

$$R_0 = \mathbb{C}\langle\langle X_1, Y_1, \dots, X_m, Y_m\rangle\rangle / I$$

be the quotient of the ring of formal power series modulo the two-sided ideal I generated by $X_iY_j - Y_jX_i$ for $i, j = 1, \dots, m$. Let $R \subset R_0$ be the subring of formal power series whose monomials have the following property: in every monomial of R, X_i occurs as many times as Y_i.

Definition 3.14 (type b). We define the generating series of type $\boldsymbol{b}$ on U by

$$J^b(U) = 1 + \sum_{k=1}^{\infty} \sum_{(c,\rho_1,\rho_2)/\sim} X_{c(\rho_1^{-1}(1))} \otimes \cdots \otimes X_{c(\rho_1^{-1}(k))} \tag{12}$$

$$\otimes Y_{c(\rho_2^{-1}(1))} \otimes \cdots \otimes Y_{c(\rho_2^{-1}(k))} \int_{g,U}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(k)}, \tag{13}$$

where the second summation is over all maps of sets $c : \{1, \dots, k\} \to \{1, \dots, m\}$ and all permutations ρ_1, ρ_2 of k elements, up to the above equivalence.

Let Y_Γ be a Hilbert modular surface. Let α and β be two rational functions on Y_Γ. We denote by D the union of the divisors $(\alpha)_\infty$ and $(\beta)_\infty$ at infinity. Let $F : Y_\Gamma - D \to \mathbb{C}^2$ be defined as $F(y) = (\alpha(y), \beta(y))$. Let $g : (0, 1)^2 \to Y_\Gamma - D$ be a membrane, so that the composition $F \circ g$ respects the coordinatewise foliations. Consider the differential forms ω_i from the definition of type $\boldsymbol{b}$. They are invariant under the action of the arithmetic group Γ. Thus, we can treat them as differential forms on the Hilbert modular variety Y_Γ.

Definition 3.15 (type c). With the new definition of a membrane g and a domain $U \subset Y_\Gamma$, we define the generating series of type $\boldsymbol{c}$ by

$$J^c(U) = 1 + \sum_{k=1}^{\infty} \sum_{(c,\rho_1,\rho_2)/\sim} X_{c(\rho_1^{-1}(1))} \otimes \cdots \otimes X_{c(\rho_1^{-1}(k))} \tag{14}$$

$$\otimes Y_{c(\rho_2^{-1}(1))} \otimes \cdots \otimes Y_{c(\rho_2^{-1}(k))} \int_{g,U}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(k)}, \tag{15}$$

where the second summation is over all maps of sets $c : \{1, \dots, k\} \to \{1, \dots, m\}$ and all permutations ρ_1, ρ_2 of k elements, up to the above equivalence.

Definition 3.16 (ring R', generating series $J(U_1, U_2)$). We define a generating series of iterated integrals on two disjoint domain U_1 and U_2 (see Definition 3.10). Let $U_i = g(A_i)$. Define

$$\begin{aligned} J(U_1, U_2) \\ = 1 + \sum_{k=1}^{\infty} \sum_{s:\{1,\dots,k\}\to\{1,2\}} \sum_{(c,\rho_1,\rho_2)/\sim} X_{c(\rho_1^{-1}(1)),s(1)} \otimes \cdots \otimes X_{c(\rho_1^{-1}(k)),s(k)} \quad (16) \\ \otimes Y_{c(\rho_2^{-1}(1)),s(1)} \otimes \cdots \otimes Y_{c(\rho_2^{-1}(k)),s(k)} \int_{g,U,s}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(k)}, \quad (17) \end{aligned}$$

The generating series takes values in a ring R' defined as follows. Let

$$R'_0 = \mathbb{C}\langle\langle X_{1,1}, X_{1,2}, Y_{1,1}, Y_{1,2}, \dots, X_{m,1}, X_{m,2}, Y_{m,1}, Y_{m,2}\rangle\rangle / I'$$

be a quotient of the ring of formal power series, where I' is the two-sided ideal generated by the Lie commutators of all the $X_{i,j}$ and $Y_{k,l}$. Let R' be the subring of R'_0 with the property that in every monomial of R', $X_{i,j}$ occurs as many times as $Y_{i,j}$.

Lemma 3.17. *Let $\phi : R' \to R$ be the homomorphism of rings defined by $\phi(X_{i,1}) = \phi(X_{i,2}) = X_i$ and $\phi(Y_{i,1}) = \phi(Y_{i,2}) = Y_i$. If $U = U_1 \cup U_2$ is in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$, then*

$$\phi(J(U_1, U_2)) = J^b(U).$$

If $U = U_1 \cup U_2$ is in X_Γ, then

$$\phi(J(U_1, U_2)) = J^c(U).$$

Proof. After applying the homomorphism ϕ the formal variables on the left-hand side become independent of the map s. Therefore, we have to examine what happens when we sum over all possible maps s. The value $s(i)$ is 1 or 2. This has the following significance: if $s(i) = 1$, then we restrict the form $g^*\omega_{c(i)}$ to A_1 (instead of to A). Similarly, if $s(i) = 2$, we restrict $g^*\omega_{c(i)}$ to A_2. If we add both choices (restriction to A_1 and restriction to A_2) then we obtain the restriction of $g^*\omega_{c(i)}$ to $A = A_1 \cup A_2$. Thus, we obtain the formula

$$\sum_{s:\{1,\dots,k\}\to\{1,2\}} \int_{g,U,s}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(k)} = \int_{g,U}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(k)}.$$

We do the same for every monomial in R. That proves the above lemma for the generating series. $\square$

3C. ***Shuffle product of generating series.*** The regions of integration that we are mostly interested in will be ideal diangles, that is, 2-cells whose boundaries have two vertices and two edges, and ideal triangles. All other regions that we will deal with are going to be a finite union of ideal diangles and ideal triangles. The first type of decomposition is based on a union of two diangles with a common vertex. The second type of decomposition will be based on two of the cells (diangles or triangles) with a common edge.

Let g_1 and g_2 be two membranes. Let $P=(0,0)$ and $Q=(1,1)$ be the vertices of a diangle $A\subset\mathbb{R}^2$ and $Q=(1,1)$ and $R=(2,2)$ be the two points of a diangle $B\subset\mathbb{R}^2$. Assume that A lies within the rectangle with vertices $(0,0)$, $(0,1)$, $(1,1)$, $(1,0)$. Similarly, assume that B lies within the rectangle $(1,1)$, $(1,2)$, $(2,2)$, $(2,1)$. Let $U=g(A)$ and $V=g(B)$.

Theorem 3.18. (i) $$\int_{g,U\cup V}\omega_1\cdots\omega_m=\sum_{j=0}^{m}\int_{g,U}\omega_1\cdots\omega_j\int_{g,V}\omega_{j+1}\cdots\omega_m.$$

(ii) *The generating series of type* $\boldsymbol{a}$ *from* Definition 3.12 *satisfies*

$$J^a(g;\,A\cup B;\,\Omega)=J^a(g;\,A;\,\Omega)J^a(g;\,B;\,\Omega).$$

The proof of the first statement is essentially the same as the combinatorial proof for composition of paths, when one considers iterated path integrals (see Lemma 2.3). The proof of the second statement combines all compositions into generating series (see Definition 3.12), resembling Manin's approach for the noncommutative modular symbol.

For generating series of type $\boldsymbol{b}$, we have a similar statement:

Definition 3.19. Let ρ' and ρ'' be two permutations of the sets $\{1,\dots,i\}$ and $\{i+1,\dots,i+j\}$, respectively. We define the permutation $\rho'^{-1}\cup\rho''^{-1}$ of $\{1,\dots,i+j\}$, which acts on $\{1,\dots,i\}$ as ρ'^{-1} and on $\{i+1,\dots,i+j\}$ as ρ''^{-1}. We define the set of shuffles of two given permutations, denoted by $\mathrm{sh}(\rho',\rho'')$, as the set of all permutations ρ of the set $\{1,2,\dots,i+j\}$ such that ρ^{-1} is the composition of a shuffle of sets $\tau\in\mathrm{sh}(i,j)$ (see Definition 2.4) with $\rho'^{-1}\cup\rho''^{-1}$. That is,

$$\rho^{-1}=\tau\circ(\rho'^{-1}\cup\rho''^{-1}).$$

Definition 3.20. We define a shuffle of two monomials

$$M'=X_{c'(\rho_1'^{-1}(1))}\otimes\cdots\otimes X_{c'(\rho_1'^{-1}(i))}\otimes Y_{c'(\rho_2'^{-1}(1))}\otimes\cdots\otimes Y_{c'(\rho_2'^{-1}(i))}\int_{g,U'}^{\rho_1',\rho_2'}\omega_{c'(1)}\cdots\omega_{c'(i)}$$

and

$$M''=X_{c''(\rho_1''^{-1}(1))}\otimes\cdots\otimes X_{c''(\rho_1''^{-1}(j))}\otimes Y_{c''(\rho_2''^{-1}(1))}\otimes\cdots\otimes Y_{c''(\rho_2''^{-1}(j))}\\ \times\int_{g,U''}^{\rho_1'',\rho_2''}\omega_{c''(i+1)}\cdots\omega_{c''(i+j)},$$

where ρ_1' and ρ_2' are permutations of $\{1,\dots,i\}$, c' is a map of sets $\{1,\dots,i\}\to\{1,\dots,m\}$, ρ_1'' and ρ_2'' are permutations of $\{i+1,\dots,i+j\}$, and c'' is a map of sets $\{i+1,\ldots,i+j\} \to \{1,\ldots,m\}$. By a shuffle product of the monomials M' and M'', we mean the sum

$$M' \times_{\mathrm{Sh}} M'' = \sum_{\substack{\rho_1 \in \mathrm{sh}(\rho_1',\rho_1'') \\ \rho_2 \in \mathrm{sh}(\rho_2',\rho_2'')}} X_{c(\rho_1^{-1}(1)),s(1)} \otimes \cdots \otimes X_{c(\rho_1^{-1}(i+j)),s(i+j)}$$

$$\otimes Y_{c(\rho_2^{-1}(1)),s(1)} \otimes \cdots \otimes Y_{c(\rho_2^{-1}(i+j)),s(i+j)} \int_{g,U'\cup U'',s}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(i+j)},$$

where $c : \{1,\ldots,i+j\} \to \{1,\ldots,m\}$ is such that its restriction to the first i elements is c' and its restriction to the last j elements is c''. Here the maps s takes the value 1 on the set $c^{-1}\{1,\ldots,i\} = c'^{-1}\{1,\ldots,i\}$ and the value 2 on the set $c^{-1}\{i+1,\ldots,i+j\} = c''^{-1}\{i+1,\ldots,i+j\}$.

Theorem 3.21 (shuffle product). *For iterated integrals of type* $\boldsymbol{b}$ *and the corresponding generating series*, *we have the following shuffle relations*:

(i) $$\int_{g,U}^{\rho_1',\rho_2'} \omega_1 \cdots \omega_j \int_{g,U}^{\rho_1'',\rho_2''} \omega_{j+1} \cdots \omega_m = \sum_{\substack{\rho_1 \in \mathrm{sh}(\rho_1',\rho_1'') \\ \rho_2 \in \mathrm{sh}(\rho_2',\rho_2'')}} \int_{g,U}^{\rho_1,\rho_2} \omega_1 \cdots \omega_m. \tag{18}$$

(ii) $$\int_{g,U'}^{\rho_1',\rho_2'} \omega_1 \cdots \omega_j \int_{g,U''}^{\rho_1'',\rho_2''} \omega_{j+1} \cdots \omega_m = \sum_{\substack{\rho_1 \in \mathrm{sh}(\rho_1',\rho_1'') \\ \rho_2 \in \mathrm{sh}(\rho_2',\rho_2'')}} \int_{g,U,s}^{\rho_1,\rho_2} \omega_1 \cdots \omega_m, \tag{19}$$

where s is a map from $\{1,\ldots,m\}$ *to* $\{1,2\}$ *such that* $\{1,\ldots,j\}$ *are mapped to* 1 *and the remaining elements are mapped to* 2.

(iii) $$\phi(J^b(U') \times_{\mathrm{Sh}} J^b(U'')) = J^b(U' \cup U''). \tag{20}$$

(iv) $$\phi(J^c(U') \times_{\mathrm{Sh}} J^c(U'')) = J^c(U' \cup U''). \tag{21}$$

Proof. For part (i), it is useful to consider the two orderings of differential forms, given in Remark 3.6. Note that we need to order the forms both horizontally and vertically in the terminology of that remark. Let us consider first the horizontal order. That is the order with respect to the first variables of the differential forms $g^*\omega_{\rho_1'^{-1}(1)},\ldots,g^*\omega_{\rho_1'^{-1}(j)}$ and $g^*\omega_{\rho_1''^{-1}(j+1)},\ldots,g^*\omega_{\rho_1''^{-1}(m)}$, corresponding to the two integrals on the left-hand side of (18). In order to arrange both of the above orderings in one sequence of increasing first arguments, we need to shuffle them (similarly to shuffling a deck of cards). That leads to $\rho_1 \in \mathrm{sh}(\rho_1', \rho_1'')$ (see Definition 3.19). We proceed similarly with the second arguments and the permutations ρ_2', ρ_2'' and ρ_2.

For (ii), apply the equality from part (i) with the differential forms $g^*\omega_1,\ldots,g^*\omega_j$ multiplied by the function $\mathbf{1}_{A'}$, defined by

$$\mathbf{1}_{A'}(x) = \begin{cases} 1 & \text{for } x \in A', \\ 0 & \text{for } x \notin A', \end{cases}$$

and the differential forms $g^*\omega_{j+1}, \dots, g^*\omega_m$ multiplied by $\mathbf{1}_{\mathbf{A}''}$.

For part (iii), we are going to establish similar relation among generating series as elements of R'. Applying the homomorphism $\phi : R' \to R$ from Lemma 3.17, we obtain the desired equality. Every monomial from $J(U_1)$ is of the form

$$M' = X_{c'(\rho_1'^{-1}(1))} \otimes \cdots \otimes X_{c'(\rho_1'^{-1}(i))} \otimes Y_{c'(\rho_2'^{-1}(1))} \otimes \cdots \otimes Y_{c'(\rho_2'^{-1}(i))} \int_{g,U'}^{\rho_1',\rho_2'} \omega_{c'(1)} \cdots \omega_{c'(i)}$$

and similarly every monomial from $J(U_2)$ is of the form

$$M'' = X_{c''(\rho_1''^{-1}(1))} \otimes \cdots \otimes X_{c''(\rho_1''^{-1}(j))} \otimes Y_{c''(\rho_2''^{-1}(1))} \otimes \cdots \otimes Y_{c''(\rho_2''^{-1}(j))} \times \int_{g,U''}^{\rho_1'',\rho_2''} \omega_{c''(i+1)} \cdots \omega_{c''(i+j)},$$

where ρ_1' and ρ_2' are permutations of $\{1, \dots, i\}$, c' is a map of sets $\{1, \dots, i\} \to \{1, \dots, m\}$, ρ_1'' and ρ_2'' are permutations of $\{i+1, \dots, i+j\}$, and c'' is a map of sets $\{i+1, \dots, i+j\} \to \{1, \dots, m\}$. We take the shuffle product of the monomials M' and M'' (see Definition 3.20):

$$M' \times_{\mathrm{Sh}} M'' = \sum_{\substack{\rho_1 \in \mathrm{sh}(\rho_1', \rho_1'') \\ \rho_2 \in \mathrm{sh}(\rho_2', \rho_2'')}} X_{c(\rho_1^{-1}(1)),s(1)} \otimes \cdots \otimes X_{c(\rho_1^{-1}(i+j)),s(i+j)} \otimes Y_{c(\rho_2^{-1}(1)),s(1)} \otimes \cdots \otimes Y_{c(\rho_2^{-1}(i+j)),s(i+j)} \int_{g,U,s}^{\rho_1,\rho_2} \omega_{c(1)} \cdots \omega_{c(i+j)},$$

where the map s takes the value 1 on the set $c^{-1}\{1, \dots, i\}$ and the value 2 on the set $c^{-1}\{i+1, \dots, i+j\}$. This determines the map s uniquely.

In order to complete the proof, we have to show that every monomial in $J(U_1, U_2)$ can be obtained in exactly one way as a result (on the right-hand side) of a shuffle product of a pair of monomials (M_1, M_2) from $J(U_1)$ and $J(U_2)$. Every monomial from $J(U_1, U_2)$ is characterized by two permutation ρ_1, ρ_2 and two maps of sets $c : \{1, \dots, k\} \to \{1, \dots, m\}$ and $s : \{1, \dots, k\} \to \{1, 2\}$. Let i be the number of elements in $s^{-1}(1)$ and j the number of elements in $s^{-1}(2)$. Then $i + j = k$. Then i is the number of differential forms among $g^*\omega_{c(1)}, \cdots, g^*\omega_{c(k)}$ which are restricted to the set A_1. The remaining j differential forms are restricted to A_2. Also, every permutation ρ_1 can be written in an unique way as a composition of a shuffle $\tau_1 \in \mathrm{sh}(i, j)$ and two disjoint permutations ρ_1' and ρ_1'' of i and of j elements, respectively (see Definition 3.19). Similarly, ρ_2 can be written in a unique way as a product of a shuffle $\tau_2 \in \mathrm{sh}(i, j)$ and two disjoint permutation ρ_2' and ρ_2''. The map of sets c_1 is defined as a restriction of the map c to the image of ρ_1'. Similarly, the map c_2 is defined as a restriction of the map c to the image of ρ_1''. Now we can define the monomials M' and M'' in $J(U_1)$ and $J(U_2)$ based on the triples ρ_1', ρ_2', c'

and ρ_1'', ρ_2'', c'', respectively. Such monomials are unique. One can show that the shuffle product of M' and M'' contains the monomial in $J(U_1, U_2)$ that we started with exactly once. The proof of part (iii) is complete after applying Lemma 3.17. $\square$

4. Hilbert modular symbols

In this section, we recall the Hilbert modular group and its action on the product of two upper half-planes. Then we define commutative Hilbert modular symbols (Section 4A) and its pairing with the cohomology of the Hilbert modular surface (Section 4B). In Sections 4C and 4D, we define noncommutative Hilbert modular symbols (Definition 4.13) as generating series of iterated integrals over membranes of type $\boldsymbol{b}$. We also examine relations among the noncommutative Hilbert modular symbols (Theorem 4.12), which we interpret as cocycle conditions or as a difference by a coboundary (Conjecture 4.14). In Section 4E, we consider a two-category C with a sheaf J on C. Then the noncommutative Hilbert modular symbol is a sheaf on a two-category. This is done in order to give a plausible approach to defining a suitable noncommutative cohomology set. In Section 4F, we make explicit computations and compare them to computations for multiple Dedekind zeta values.

4A. ***Commutative Hilbert modular symbols.*** In this subsection, we define commutative Hilbert modular symbols, using geodesics, geodesic triangles and geodesic diangles. Then, we prove certain relations among the commutative Hilbert modular symbols, which are generalized to relations among noncommutative Hilbert modular symbols (Section 4D).

Let $K = \mathbb{Q}(\sqrt{d})$ be a real quadratic extension of $\mathbb{Q}$. Then the ring of integers in K is

$$\mathbb{O}_K = \begin{cases} \mathbb{Z}[(1+\sqrt{d})/2] & \text{for } d = 1 \bmod 4, \\ \mathbb{Z}[\sqrt{d}] & \text{for } d = 2, 3 \bmod 4. \end{cases}$$

Then $\Gamma = \mathrm{SL}_2(\mathbb{O}_K)$ is called a Hilbert modular group. Let $\gamma \in \Gamma$. We recall the action of γ on a product of two upper half-planes $\mathbb{H}^2$. Let

$$\gamma = \gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}.$$

Let a_2, b_2, c_2, d_2 be the Galois conjugates of a_1, b_1, c_1, d_1, respectively. Let us define γ_2 by

$$\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}.$$

Let $z = (z_1, z_2)$ be any point of the product of two upper half-planes $\mathbb{H}^2$.

For an element $\gamma \in \mathrm{GL}_2(K)$, we define the following action: If $\det\gamma$ is totally positive, that is $\det\gamma_1 > 0$ and $\det\gamma_2 > 0$, then the action of γ on $z = (z_1, z_2) \in \mathbb{H}^2$ is essentially the same as for $\gamma \in \mathrm{SL}_2(K)$, namely,

$$\gamma z = (\gamma_1 z_1, \gamma_2 z_2),$$

where

$$\gamma_1 z_1 = \frac{a_1 z_1 + b_1}{c_1 z_1 + d_1} \quad \text{and} \quad \gamma_2 z_2 = \frac{a_2 z_2 + b_2}{c_2 z_2 + d_2}$$

are linear fractional transforms. If $\det\gamma$ is totally negative, that is, $\det\gamma_1 < 0$ and $\det\gamma_2 < 0$, then we define

$$\gamma z = \left(-\frac{a_1 \bar{z}_1 + b_1}{c_1 \bar{z}_1 + d_1}, -\frac{a_2 \bar{z}_2 + b_2}{c_2 \bar{z}_2 + d_2}\right).$$

Similarly if one of the embeddings of $\det\gamma$ is positive and the other is negative, for example, $\det\gamma_1 > 0$ and $\det\gamma_2 < 0$, e.g., for $\det\gamma = \sqrt{d}$, then

$$\gamma z = \left(\frac{a_1 z_1 + b_1}{c_1 z_1 + d_1}, -\frac{a_2 \bar{z}_2 + b_2}{c_2 \bar{z}_2 + d_2}\right).$$

We add cusp points $\mathbb{P}^1(K)$ to $\mathbb{H}^2$. Then the quotient $\mathrm{SL}_2(\mathbb{O}_K)\backslash(\mathbb{P}^1(K) \cup \mathbb{H}^2)$ is compact.

We are going to carefully examine geodesics joining the cusps 0, 1 and ∞.

Let z_0, z_1, z_∞ be three distinct cusp points. There is a unique $\gamma \in \mathrm{PGL}_2(K)$ that sends z_0, z_1 and z_∞ to 0, 1 and ∞, respectively.

Let

$$i : \mathbb{H} \to \mathbb{H}^2, \quad i(x) = (x, x)$$

be the diagonal map and Δ be its image. Consider the Hirzebruch–Zagier divisor $X = \gamma^*\Delta$. It is an analytic curve that passes through the points z_0, z_1 and z_∞. Then X is a holomorphic curve in $\mathbb{H}^2$ if $\det\gamma$ is totally positive or totally negative. If $\det\gamma$ is not totally positive or totally negative, then X is a holomorphic curve in $\mathbb{H}^1 \times \overline{\mathbb{H}}^1 \cup \mathbb{P}^1(K)$; in other words, it is an antiholomorphic curve in $\mathbb{H}^2$, such as $z_1 = -\bar{z}_2$. Let $\Delta_X = \gamma^*\Delta$ be the pullback of the geodesic triangle Δ formed by the points $0, 1, \infty$ in the analytic curve X.

Given four points on the boundary in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$, we are tempted to consider them as vertices of a geodesic tetrahedron in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$, whose faces are triangles of the type Δ_X. However, there is one problem that we encounter: Two distinct cusps could be connected by two different geodesics in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$. In particular, two triangles from the faces of the "tetrahedron" might not have a common edge, but only two common vertices. Thus, we are led to consider a *thickened tetrahedron* with two types of faces on the boundary: the first type is an ideal triangle that we have just defined and the other type is an ideal diangle — a union of geodesics connecting two fixed points, which has the homotopy type of a disc with two vertices and two edges. The two edges of an ideal diangle in the boundary of a thickened tetrahedron correspond to the two geodesics connecting the same two cusps, where two geodesics belong to the geodesic triangles that have the two cusps in common.

Let us describe a diangle $D_{0,\infty;1,\alpha}$ whose two vertices are 0 and ∞ and whose two sides are geodesics that belong to each of the ideal triangles $0, 1, \infty$ and $0, \alpha, \infty$. The geodesic l_0 between the points 0 and ∞ that lie on the geodesic triangle $0, 1, \infty$ can be parametrized in the following way: $\{(it, it) \mid t \in \mathbb{R}, t \geq 0\} \subset \mathrm{Im}(\mathbb{H}) \times \mathrm{Im}(\mathbb{H})$. Here by $\mathrm{Im}(\mathbb{H})$ we mean the imaginary part of the upper half-plane. The element $\gamma \in \Gamma$ that sends $0, \alpha, \infty$ to $0, 1, \infty$ is $\gamma = \left(\begin{smallmatrix} \alpha^{-1} & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Then $(\gamma^{-1})^*(it, it) = (|\alpha_1|it, |\alpha_2|it)$. Therefore, the geodesic l_α between the points 0 and ∞ that lie on the geodesic triangle $0, \alpha, \infty$ can be parametrized as $\{(|\alpha_1|it, |\alpha_2|it) \mid t \in \mathbb{R},\ t \geq 0\} \subset \mathrm{Im}(\mathbb{H}) \times \mathrm{Im}(\mathbb{H})$. Then, we define the diangle $D_{0,\infty;1,\alpha}$ as the two-dimensional region in $\mathrm{Im}(\mathbb{H}) \times \mathrm{Im}(\mathbb{H})$ between the lines l_0 and l_α. We also consider the diangle with orientation. If $|\alpha_1| > |\alpha_2|$ then it is positively oriented. If the inequality is reversed then the diangle is negatively oriented; if $|\alpha_1| = |\alpha_2|$ then it is a degenerate diangle, which consists of a single geodesic. All other diangles that we will consider are translates of $D_{0,\infty;1,\alpha}$ via the action of any element $\gamma \in \mathrm{PGL}_2(K)$.

Lemma 4.1. (i) *Each geodesic triangle Δ_X lies either on a holomorphic curve or on an antiholomorphic curve.*

(ii) *Each geodesic in a geodesic triangle Δ_X belongs both to a holomorphic curve and to an antiholomorphic curve.*

Proof. Part (i) follows from the construction of a geodesic triangle before the lemma. For part (ii), consider the following: Let $\Delta(0, 1, \infty)$ be the geodesic triangle in the diagonal of $\mathbb{H}^2$ connecting the points 0, 1 and ∞. It is a holomorphic curve. Thus, a geodesic $\{(it, it) \in \mathbb{H}^2 \mid t > 0\}$, connecting the points 0 and ∞ as a face of the geodesic triangle $\Delta(0, 1, \infty)$, lies on a holomorphic curve. Now consider the geodesic triangle $D(0, \sqrt{d}, \infty)$. It lies on an antiholomorphic curve in $\mathbb{H}^2$, by which we mean a complex curve in $\mathbb{H}^2$ (where we have taken the complex conjugate complex structure in one of the upper half-planes), since the linear fractional transformation that sends $D(0, \sqrt{d}, \infty)$ to $D(0, 1, \infty)$ does not have totally positive (or totally negative) determinant. Explicitly, the linear fractional transformation that sends $(0, \sqrt{d}, \infty)$ to $(0, 1, \infty)$ is

$$\gamma = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{d} \end{pmatrix}.$$

Then

$$(\gamma_1, \gamma_2) = \left(\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{d} \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -\sqrt{d} \end{pmatrix} \right).$$

We have $\gamma_1(it) = (1/\sqrt{d})it$ and $\gamma_2(it) = -(1/\sqrt{d})\overline{it} = \gamma_1(it)$. Then the same geodesic (it, it) belongs to the antiholomorphic curve given by the pullback of the diagonal with respect to the linear fractional map γ. Thus, we obtain that the geodesic (it, it) connecting 0 and ∞ belongs to both a holomorphic curve and an antiholomorphic curve. Similarly, any translate of the geodesic (it, it) via a linear fractional map from $\mathrm{GL}_2(K)$ would belong to both a holomorphic curve and an antiholomorphic curve. That proves part (ii). $\square$

Definition 4.2. Let p_1, p_2, p_3, p_4 be cusp points in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$. To each triple of points p_1, p_2, p_3, we associate the geodesic triangle $\{p_1, p_2, p_3\}$ with coefficient 1 as an element of the singular chain complex in $C_2(\mathbb{H}^2 \cup \mathbb{P}^1(K), \mathbb{Q})$. Also, to each quadruple of points p_1, p_2, p_3, p_4, we associate the geodesic diangle between the two geodesic connecting p_1 and p_2 so that the first geodesic is a face of the geodesic triangle $\{p_1, p_2, p_3\}$ and the second geodesic is a face of the geodesic triangle $\{p_1, p_2, p_4\}$. We denote such a diangle by $\{p_1, p_2; p_3, p_4\}$. We call the geodesic triangle $\{p_1, p_2, p_3\}$ and the geodesic diangle $\{p_1, p_2; p_3, p_4\}$, considered as elements of $C_2(\mathbb{H}^2 \cup \mathbb{P}^1(K), \mathbb{Q})$, *commutative Hilbert modular symbols*.

Theorem 4.3. *The commutative Hilbert modular symbols, modulo the boundary of singular* 3*-chains* $\partial C_3(\mathbb{H}^2 \cup \mathbb{P}^1(K), \mathbb{Q})$, *satisfy the following properties*:

(1) *If* σ *is a permutation of the set* $\{1, 2, 3\}$ *then*

$$\{p_{\sigma(1)}, p_{\sigma(2)}, p_{\sigma(3)}\} = \mathrm{sign}(\sigma)\{p_1, p_2, p_3\}.$$

(2) *If* p_1, p_2, p_3, p_4 *are four points on the same holomorphic* (*or antiholomorphic*) *curve of the type* $\gamma^*\Delta$, *then*

$$\{p_1, p_2, p_3\} + \{p_2, p_3, p_4\} = \{p_1, p_2, p_4\} + \{p_1, p_3, p_4\}.$$

For every four points p_1, p_2, p_3, p_4, *we associate a diangle with vertices* p_1 *and* p_2. *Let* $\{p_1, p_2; p_3, p_4\}$ *be the corresponding symbol.*

(3) *If* p_1, p_2, p_3, p_4 *are four points on the same holomorphic* (*or antiholomorphic*) *curve of type* $\gamma^*\Delta$, *then*

$$0 = \{p_1, p_2; p_3, p_4\}.$$

(4) *For every distinct four points* p_1, p_2, p_3, p_4, *we have the following relations, based on the orientation of the domain*:

$$\{p_2, p_1; p_3, p_4\} = \{p_1, p_2; p_4, p_3\} = -\{p_2, p_1; p_4, p_3\} = -\{p_1, p_2; p_3, p_4\}.$$

(5) *For every five points* p_1, p_2, p_3, p_4, p_5, *we have*

$$\{p_1, p_2; p_3, p_4\} + \{p_1, p_2; p_4, p_5\} = \{p_1, p_2; p_3, p_5\}.$$

(6) *We also have a relation between the two types of commutative Hilbert modular symbols. For every four distinct points* p_1, p_2, p_3, p_4, *we have*

$$\begin{aligned} 0 = &\{p_1, p_2, p_3\} + \{p_2, p_3, p_4\} - \{p_1, p_2, p_4\} - \{p_1, p_3, p_4\} \\ &+ \{p_1, p_2; p_3, p_4\} + \{p_2, p_3; p_1, p_4\} + \{p_3, p_1; p_2, p_4\} \\ &+ \{p_3, p_4; p_1, p_2\} + \{p_1, p_4; p_2, p_3\} + \{p_2, p_4; p_3, p_1\}. \end{aligned}$$

Proof. Part (1) follows from the orientation of the simplex in singular homology. Part (2) is an equality induced by two different triangulations on a holomorphic (or antiholomorphic) curve with four vertices. In that setting the diangles are trivial, which proves part (3). Part (4) follows from the orientation of the diangle. Part (5) corresponds to a union of two geodesic diangles with a common face, given by a third geodesic diangle. Part (5) will be used for a noncommutative 1-cocycle relation for the noncommutative Hilbert modular symbol (see Conjecture 4.14). Part (6) is a boundary relation for the boundary of a thickened tetrahedron. By a thickened tetrahedron, we mean a union of four geodesic triangles corresponding to each triple of points among the four points p_1, p_2, p_3, p_4 together with six geodesic diangles that correspond to the area between the faces of the geodesic triangles. They correspond exactly to the thickening of the six edges of a tetrahedron. □

Part 6 will be used to derive explicit formulas for the noncommutative Hilbert modular symbol of type $\boldsymbol{c}'$ resembling a noncommutative 2-cocycle relation (see Conjecture 4.15).

4B. ***Pairing of the modular symbols with cohomology.*** In this subsection, we consider pairings between commutative Hilbert modular symbols and cusp forms. In some cases, we prove that such pairings give periods in the sense of [Kontsevich and Zagier 2001].

We are interested in holomorphic cusp forms with respect to Γ. Equivalently, we can consider the holomorphic 2-forms on $\widetilde{X}$, the minimal smooth algebraic compactification of X [Hirzebruch 1973]. At this point we should distinguish between geodesic triangles that lie on a holomorphic curve and those that lie on an antiholomorphic curve. The reason for this distinction is that a holomorphic 2-form restricted to a holomorphic curve vanishes. The way to distinguish the two types of geodesic triangles is the following: Let γ be a linear fractional transform that sends the points p_1, p_2, p_3 to $0, 1, \infty$. If $\det\gamma$ is totally positive or totally negative, then the geodesic triangle p_1, p_2, p_3 lies on a holomorphic curve. If $\det\gamma$ is neither totally positive nor totally negative, then the geodesic triangle p_1, p_2, p_3 lies on an antiholomorphic curve.

Definition 4.4. Define $M_2(\mathbb{H}^2 \cup \mathbb{P}^1(K), \mathbb{Q})$ to be the span of the Hilbert modular symbols $\{p_1, p_2, p_3\}$ and $\{p_1, p_2; p_3, p_4\}$ as a subspace of the singular chain $C_2(\mathbb{H}^2 \cup \mathbb{P}^1(K), \mathbb{Q})$. We define the pairing

$$\langle\ ,\ \rangle : M_2(\mathbb{H}^2 \cup \mathbb{P}^1(K)) \times S_{2,2}(\Gamma) \to \mathbb{C}$$

by setting

$$\langle \{p_1, p_2, p_3\}, f\, dz_1 \wedge dz_2 \rangle = \int_{\{p_1, p_2, p_3\}} f\, dz_1 \wedge dz_2$$

for geodesic triangles and

$$\langle \{p_1, p_2; p_3, p_4\}, f\, dz_1 \wedge dz_2 \rangle = \int_{\{p_1, p_2; p_3, p_4\}} f\, dz_1 \wedge dz_2$$

for geodesic diangles.

We are going to use that a Hilbert modular surface $X(\mathbb{C})$ can be realized as the complex points of an arithmetic surface defined over a number field F.

Theorem 4.5. *The image of the above pairing is a period over a number field F when we integrate a normalized cusp Hecke eigenform f of weight* $(2, 2)$. (For Hecke eigenforms, see [Shimura 1978; Berger et al. 2013].)

Proof. From Lemma 3.1(ii), the boundary of the geodesic triangles of the diangles are geodesics that lie on holomorphic curves in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$. Therefore, in the quotient by the congruence group Γ, the geodesic lies in a Hirzebruch–Zagier divisor on the Hilbert modular surface. Thus, we integrate a closed algebraic differential 2-form (that is, a global differential 2-form with algebraic coefficients) on the Hilbert modular surface, with boundaries Hirzebruch–Zagier divisors. □

Conjecture 4.6. Let $f \in S_{k,k}(\Gamma)$ be a normalized cusp Hecke eigenform of weight (k, k). Then

$$\int_{\{p_1, p_2, p_3\}} f\, dz_1 \wedge dz_2$$

for geodesic triangles and

$$\int_{\{p_1, p_2; p_3, p_4\}} f\, dz_1 \wedge dz_2$$

for geodesic diangles are periods.

Theorem 4.5 is a proof of Conjecture 4.6 for the case of cusp form of weight (2, 2).

4C. ***Iteration — revisited.*** We defined iterated integrals on diangles in Definitions 3.14 and 3.15. However, these definitions have to be extended to other domains of integration in order to consider iterated integrals on geodesic triangles.

A consequence of the results from this subsection is the following:

Theorem 4.7. *Iterated integrals of type* $\boldsymbol{c}$ *on a geodesic diangle and on a geodesic triangle of algebraic differential* 2*-forms on a Hilbert modular surface are periods in the sense of Kontsevich–Zagier.*

Before giving the proof, we need definitions of several objects, as well as their properties. In the process, we will be able to extend the definition of iterated integrals on membranes when the domain of integration is a geodesic triangle.

For type $\boldsymbol{b}$, in Definition 3.14, we have a map $g : U \to \mathbb{H}^2$ that sends the two $\mathbb{R}$-foliations on U into two coordinatewise $\mathbb{C}$-foliations of $\mathbb{H}^2$. The same definition does not work when the domain U is a geodesic triangle. The reason is that a geodesic triangle is either a holomorphic curve or an antiholomorphic curve. In both cases, a pullback of one leaf to the geodesic triangle is a point not a line (which is the case for the diangles).

In order to extend Definitions 3.14 and 3.15 to the case when the domain U is a geodesic triangle, we are going to construct a new space using the fiber products multiple times.

Now, we are going to define a space Y_n associated to an iterated integral on n 2-forms on $\mathbb{H}^2$. We are going to use fiber products (see [Hartshorne 1977]). Let p_1 and p_2 be the projections of $\mathbb{H}^2$ on the first and the second component, respectively. Define $X_{ij} = \mathbb{H}^2$ for $1 \le i \le n$ and $1 \le j \le n$. (One should think of the component X_{ij} as the complexification of the real coordinate (s_i, t_j).) Let $C_i = \mathbb{H}$ for $1 \le i \le n$ and $C'_j = \mathbb{H}$ for $1 \le j \le n$. Let

$$X_j = X_{1j} \times_{C'_j} X_{2j} \times_{C'_j} \cdots \times_{C'_j} X_{nj}.$$

(X_j corresponds to the variable t_j.) Then

$$X_j \subset X_{1j} \times X_{2j} \times \cdots \times X_{nj}.$$

Let also

$$P_j = (p_1, \dots, p_1) : X_{1j} \times X_{2j} \times \cdots \times X_{nj} \to C_1 \times \cdots \times C_n.$$

Let $P_j^\circ = P_j|_{X_j}$ be the restriction of P_j to the subset X_j. We define Y_n as the fiber product of $X_1, \dots, X_n$ with respect to the morphisms $P_1^\circ, \dots, P_n^\circ$ over the base $C_1 \times \cdots \times C_n$, namely

$$Y_n = X_1 \times_C \cdots \times_C X_n, \tag{22}$$

where $C = C_1 \times \cdots \times C_n$. Note that X_j is isomorphic to X_{j+1}. Let Z_j be the subspace of Y_n defined by setting the j- and the $(j+1)$-components of $Y_n = X_1 \times_C \cdots \times_C X_n$

to be equal. (The space Z_j corresponds to a boundary components obtained by letting $t_j = t_{j+1}$.) Similarly, we could have defined Y_n by defining first

$$X_i' = X_{i1} \times_{C_i} X_{i2} \times_{C_i} \cdots \times_{C_i} X_{in},$$

(X_i' corresponds to s_i) so that

$$X_i' \subset X_{i1} \times X_{i2} \times \cdots \times X_{in}.$$

Let

$$P_i' = (p_2, \ldots, p_2) : X_{i1} \times X_{i2} \times \cdots \times X_{in} \to C_1' \times \cdots \times C_n'.$$

Define $P_i'^{\circ} = P_i'|_{X_i'}$ to be the restriction of P_i' to X_i'. We define Y_n as the fiber product of $X_1', \ldots, X_n'$ with respect to the morphisms $P_1'^{\circ}, \ldots, P_n'^{\circ}$ over the base $C_1' \times \cdots \times C_n'$, namely,

$$Y_n = X_1' \times_{C'} \cdots \times_{C'} X_n', \tag{23}$$

where $C' = C_1' \times \cdots \times C_n'$. Similarly we define Z_i' to be the subspace of Y_n defined by setting the i- and the $(i+1)$-components of $Y_n = X_1' \times_{C'} \cdots \times_{C'} X_n'$ to be equal. (The space Z_i' corresponds to a boundary components obtained by letting $s_i = s_{i+1}$.)

We have given two definitions (22) and (23) of the space Y_n. In the two definitions we have only exchanged the role of p_1 and p_2. We will prove that both definitions lead to the same object in the case $n = 2$. The general case is left to the reader.

Lemma 4.8. *For $n = 2$, the two definitions (22) and (23) define isomorphic objects Y_2.*

Proof. The space Y_2 can be defined as a finite limit (in a categorical sense) of a diagram in the following way. Consider the commutative diagram

$$\begin{tikzcd}
 & & C_1' & & \\
 & X_{11} \arrow[ur] \arrow[dl] & & X_{12} \arrow[ul] \arrow[dr] & \\
C_1 & & & & C_2 \\
 & X_{21} \arrow[ul] \arrow[dr] & & X_{22} \arrow[ur] \arrow[dl] & \\
 & & C_2' & &
\end{tikzcd}$$

For any space W such that

$$\begin{tikzcd}[column sep=small]
 & & C_1' & & \\
 & X_{11} \arrow[ur] \arrow[dl] & & X_{12} \arrow[ul] \arrow[dr] & \\
C_1 & & W \arrow[ul] \arrow[ur] \arrow[dl] \arrow[dr] & & C_2 \\
 & X_{21} \arrow[ul] \arrow[dr] & & X_{22} \arrow[ur] \arrow[dl] & \\
 & & C_2' & &
\end{tikzcd} \tag{24}$$

commutes, we have that the maps $f_{ij} : W \to X_{ij}$ factor through $g_{ij} : Y_2 \to X_{ij}$, so that $f_{ij} = g_{ij} \circ h$ for some $h : W \to Y_2$, and Y_2 is part of the commutative diagram

$$\begin{tikzcd}[column sep=small]
 & & C_1' & & \\
 & X_{11} \arrow[ur] \arrow[dl] & & X_{12} \arrow[ul] \arrow[dr] & \\
C_1 & & Y_2 \arrow[ul] \arrow[ur] \arrow[dl] \arrow[dr] & & C_2 \\
 & X_{21} \arrow[ul] \arrow[dr] & & X_{22} \arrow[ur] \arrow[dl] & \\
 & & C_2' & &
\end{tikzcd}$$

In order to prove this universal property of Y_2 we follow the first definition of Y_2. This leads to the commutative diagram

$$\begin{tikzcd}
 & & C_1' & & \\
 & X_{11} \arrow[ur] \arrow[dl] & X_{11} \times_{C_1'} X_{12} \arrow[l] \arrow[r] & X_{12} \arrow[ul] \arrow[dr] & \\
C_1 & & W \arrow[u] \arrow[d] & & C_2 \\
 & X_{21} \arrow[ul] \arrow[dr] & X_{21} \times_{C_2'} X_{22} \arrow[l] \arrow[r] & X_{22} \arrow[ur] \arrow[dl] & \\
 & & C_2' & &
\end{tikzcd}$$

Then we have that $X_1 = X_{11} \times_{C_1'} X_{12}$ maps to $C = C_1 \times C_2$ and also $X_2 = X_{21} \times_{C_2'} X_{22}$ maps to $C = C_1 \times C_2$. Thus the maps from W to any element of the diagram factors through $Y_2 = X_1 \times_C X_2$. Similarly, W factors through $X_1' \times_{C'} X_2'$, where $X_1' = X_{11} \times_{C_1} X_{21}$, $X_2' = X_{12} \times_{C_2} X_{22}$ and $C' = C_1' \times C_2'$. Since both $X_1 \times_C X_2$ and $X_1' \times_{C'} X_2'$ are universal objects with respect to the diagram (24), we have that they are isomorphic. □

Now, we return to the initial question of this subsection, namely, how to iterate over a geodesic triangle so that it is consistent with the current definition of iteration over a diangle.

For an n-fold iteration of 2-forms of types $\boldsymbol{b}$ or $\boldsymbol{c}$, we have to specify a domain $U \subset \mathbb{H}^2$, $\dim_{\mathbb{R}} U = 2$, and a pair of permutations ρ_1 and ρ_2 of n elements. We make an essential assumption that the boundary of $U \subset \mathbb{H}^2$, denoted by ∂U, projected onto the Hilbert modular surface Y_Γ lies on a finite union of Hirzebruch–Zagier divisors. We will denote the finite union of such Hirzebruch–Zagier divisors by HZ.

Let

$$P_{\rho_1,\rho_2} : X_{11} \times \cdots \times X_{nn} \to X_{\rho_1(1)\rho_2(1)} \times \cdots \times X_{\rho_1(n)\rho_2(n)}$$

be a projection to n of the factors. Let $U_{ij} \cong U$ for $1 \le i \le n$ and $1 \le j \le n$. Let

$$I : U_{\rho_1(1)\rho_2(1)} \times \cdots \times U_{\rho_1(n)\rho_2(n)} \to X_{\rho_1(1)\rho_2(1)} \times \cdots \times X_{\rho_1(n)\rho_2(n)}$$

be induced from the product of inclusion of the domains $U \to X$. We will use the notation

$$\underline{U}^\rho = U_{\rho_1(1)\rho_2(1)} \times \cdots \times U_{\rho_1(n)\rho_2(n)}$$

and

$$\underline{X}^\rho = X_{\rho_1(1)\rho_2(1)} \times \cdots \times X_{\rho_1(n)\rho_2(n)}.$$

Then the map I becomes

$$I : \underline{U}^\rho \to \underline{X}^\rho.$$

Let

$$J : Y_n \to \underline{X}^\rho$$

be the composition of the natural inclusion $Y_n \to X_{11} \times \cdots \times X_{nn}$ and the projection P_{ρ_1,ρ_2}. Then we define the domain of integration to be

$$\underline{U}^\rho_{Y_n} = \underline{U}^\rho \times_{\underline{X}^\rho} Y_n,$$

which is the fiber product of the maps I and J. Since $I : \underline{U}^\rho \to \underline{X}^\rho$ is an inclusion, we have that the induced map

$$\underline{U}^\rho_{Y_n} \to Y_n$$

is an inclusion.

In the above constructions, we have used a parallel between type $\boldsymbol{b}$ and type $\boldsymbol{c}$ iterated integrals on membranes. The following definition allows us to extend in some sense the two types when the domain of integration is an ideal triangle:

Definition 4.9 (iterated integrals on membranes of types $\boldsymbol{b}'$ or $\boldsymbol{c}'$). For any manifold with corners of dimension 2 on a Hilbert modular variety, we define an iterated integral

$$\int_U^{\Sigma_n(\rho_1,\rho_2)} (f_1\,dz_1 \wedge dz_2)\cdots(f_n\,dz_1 \wedge dz_2) = \int_{\underline{U}^{\rho}_{Y_n}} J^*(f_1\,dz_1 \wedge dz_2, \dots, f_n\,dz_1 \wedge dz_2), \quad (25)$$

where $f_k\,dz_1 \wedge dz_2$ is a form defined on $X_{\rho_1(k)\rho_2(k)}$ for $1 \le k \le n$. If Y_n and $\underline{U}^{\rho}_{Y_n}$ are constructed in the setting of type $\boldsymbol{b}$ iterated integrals on membranes, then the above definition is of iterated integrals on membranes of type $\boldsymbol{b}'$. Similarly, if Y_n and $\underline{U}^{\rho}_{Y_n}$ are constructed in the setting of type $\boldsymbol{c}$ iterated integrals on membranes, then the above definition is of iterated integrals on membranes of type $\boldsymbol{c}'$.

If U is a diangle, then the relation of the above integral to the ones defined by iterated integrals over membranes is the following: The integral

$$\int_U^{\Sigma_n(\rho_1,\rho_2)} (f_1\,dz_1 \wedge dz_2)\cdots(f_n\,dz_1 \wedge dz_2)$$

is the sum of the integrals from Definitions 3.14 or 3.15, namely, the sum

$$\sum_{\rho\in\Sigma_n} \int_U^{(\rho\rho_1,\rho\rho_2)} (f_1\,dz_1 \wedge dz_2)\cdots(f_n\,dz_1 \wedge dz_2)$$

over the orbit of the diagonal action of the permutation group Σ_n on any chosen pair of permutations (ρ_1, ρ_2).

Proposition 4.10 (properties of the iterated integral (25)). (1) *The iterated integral (25) is well-defined when U is an ideal triangle both for types $\boldsymbol{b}$ and $\boldsymbol{c}$.*

(2) *The iterated integral (25) for type $\boldsymbol{c}$ is a period if U is an ideal triangle or an ideal diangle, when $f_1, \dots, f_n$ are normalized Hecke eigenforms of weight* $(2, 2)$.

(3) *The iterated integral (25), both for types $\boldsymbol{b}$ and $\boldsymbol{c}$, is homotopy invariant with respect to homotopies that vary within the divisors*

$$J^{-1}\big(X_{\rho_1(1)\rho_2(1)} \times \cdots \times p_1^{-1}(q_i) \times \cdots \times X_{\rho_1(n)\rho_2(n)}\big),$$

where q_i is a point of $X_{\rho_1(i)\rho_2(i)}$ for fixed i and $p_1 : X_{\rho_1(i)\rho_2(i)} \to C$; or homotopies that vary within the divisors

$$J^{-1}\big(X_{\rho_1(1)\rho_2(1)} \times \cdots \times p_2^{-1}(q_i) \times \cdots \times X_{\rho_1(n)\rho_2(n)}\big),$$

where q_i is a point of $X_{\rho_1(i)\rho_2(i)}$ for fixed i and $p_2 : X_{\rho_1(i)\rho_2(i)} \to C'$.

Proof. (a) The integral (25) is well-defined for any two-dimensional submanifold with corners of the Hilbert modular variety [Borel and Serre 1973].

(b) The iterated integral (25) is a period since:

(1) A Hilbert modular variety can be defined over a number field.

(2) The normalized Hecke eigenforms $f_1, \dots, f_n$ of weight $(2, 2)$ can be realized as algebraic differential forms on the Hilbert modular variety.

(3) The boundary of the region of integration $\overline{U}^{\rho}_{Y_n}$ is a divisor on Y_n, namely,

$$\bigcup_{i=1}^{n} \mathrm{HZ}_i,$$

where

$$\mathrm{HZ}_i = J^{-1}(X_{\rho_1(1)\rho_2(1)} \times \cdots \times \mathrm{HZ} \times \cdots \times X_{\rho_1(n)\rho_2(n)})$$

is a divisor of Y_n obtained as a pullback of a divisor whose i-th component is a Hirzebruch–Zagier divisor HZ, and the rest of the factors are $X_{\rho_1(k)\rho_2(k)}$ for $k \neq i$.

(c) The proof is essentially the same as that of Theorem 3.7. □

The domain $\underline{U}_{Y_n}$ might be cut into disconnected components by the Z_i and Z'_j. In order to choose a connected component we need to define another region of integration. Recall that for the case of iterated integrals on membranes of type $\boldsymbol{b}$, $p_1 : \mathbb{H}^2 \to C$ and $p_2 : \mathbb{H}^2 \to C'$ are projections onto the first and the second component, with $C \cong \mathbb{H}$ and $C' \cong \mathbb{H}$.

For the case of iterated integrals on membranes of type $\boldsymbol{c}$, $p_1 = \alpha_1 \circ \pi$ and $p_2 = \alpha_2 \circ \pi$ are compositions of

$$\pi : \mathbb{H}^2 \to X_\Gamma,$$

the map from the universal cover to the Hilbert modular surface, with

$$\alpha_1, \alpha_2 : X_\Gamma \to \mathbb{P}^1,$$

two algebraically independent rational functions on the Hilbert modular surface, and $C_i \cong \mathbb{P}^1$ and $C'_j \cong \mathbb{P}^1$ for $1 \le i \le n$ and $1 \le j \le n$.

Let $q_0, q_1, r_0, r_1 \in \mathbb{P}^1$ be points. Let Q_0, Q_1, R_0 and R_1 be connected components of $p_1^{-1}(q_0)$, $p_1^{-1}(q_1)$, $p_2^{-1}(r_0)$, and $p_2^{-1}(r_1)$, respectively.

Let $V \to \mathbb{H}^2$ be a domain in $\mathbb{H}^2$ with boundary on the union

$$Q_0 \cup Q_1 \cup R_0 \cup R_1,$$

but with interior disjoint from this union. We define the divisors Z_0, Z_n, Z'_0, Z'_n of Y_n as follows: Z_0 will be the beginning of the integration of the t_1 variable ($t_1 = 0$),

Z_n will be the end of the integration of the t_n variable ($t_n = 1$), Z'_0 will be the beginning of the integration of the s_1 variable ($s_1 = 0$), and Z'_n will be the end of the integration of the s_n variable ($s_n = 1$). We define them as the fiber product

$$\begin{aligned} Z_0 &= Q_0 \times_C X_2 \times_C \cdots \times_C X_n, & Z_n &= X_1 \times_C \cdots \times_C X_{n-1} \times_C Q_1, \\ Z'_0 &= R_0 \times_{C'} X'_2 \times_{C'} \cdots \times_{C'} X'_n, & Z'_n &= X'_1 \times_{C'} \cdots \times_{C'} X'_{n-1} \times_{C'} R_1. \end{aligned}$$

We will use the notation

$$\underline{V}^\rho = V_{\rho_1(1)\rho_2(1)} \times \cdots \times V_{\rho_1(n)\rho_2(n)} \quad \text{and} \quad \underline{X}^\rho = X_{\rho_1(1)\rho_2(1)} \times \cdots \times X_{\rho_1(n)\rho_2(n)}.$$

Then the map I' becomes

$$I' : \underline{V}^\rho \to \underline{X}^\rho.$$

Let

$$J' : Y_n \to \underline{X}^\rho$$

be the composition of the natural inclusion $Y_n \to X_{11} \times \cdots \times X_{nn}$ and the projection P_{ρ_1,ρ_2}. Then we define the domain of integration to be

$$\underline{V}^\rho_{Y_n} = \underline{V}^\rho \times_{\underline{X}^\rho} Y_n,$$

which is the fiber product of the maps I' and J'. Since $I' : \underline{U}^\rho \to \underline{X}^\rho$ is an inclusion, we have that the induced map

$$\underline{V}^\rho_{Y_n} \to Y_n$$

is an inclusion.

Then the divisors $Z_0, Z_1, \ldots, Z_{n-1}, Z_n$ and $Z'_0, Z'_1, \ldots, Z'_{n-1}, Z'_n$ cut out from $\underline{V}_{Y_n}$ a product of two n-simplices, which corresponds to the region where the product $\{0 \le s_1 \le \cdots \le s_n \le 1\} \times \{0 \le t_1 \le \cdots \le t_n \le 1\}$ is embedded. Denote by $\overline{V}^\rho_{Y_n}$ the connected components of $\overline{V}^\rho_{Y_n}$ that contains the image of the product $\{0 \le s_1 \le \cdots \le s_n \le 1\} \times \{0 \le t_1 \le \cdots \le t_n \le 1\}$ under the map g from Definition 3.5.

Proof of Theorem 4.7. We consider the type of iterated integrals defined in Definition 3.15. Using the above notation, the domain of integration is $U \subset V$. We define

$$\overline{U}^\rho_{Y_n} = \underline{U}^\rho_{Y_n} \cap \overline{V}^\rho_{Y_n}.$$

Then the boundary of $\overline{U}^\rho_{Y_n}$ lies on the union of divisors

$$\partial \overline{U}^\rho_{Y_n} \subset \left(\bigcup_{i=1}^n Z_i \right) \cup \left(\bigcup_{j=1}^n Z'_j \right).$$

The normalized Hecke eigenforms of weight $(2, 2)$ can be realized as algebraic differential forms on the Hilbert modular variety. Then the iterated integrals on a membrane of type $\boldsymbol{c}$ over the domain U are periods, since:

(1) A Hilbert modular variety can be defined over a number field.

(2) The normalized Hecke eigenforms $f_1, \dots, f_n$ of weight $(2, 2)$ can be realized as algebraic differential forms on the Hilbert modular variety.

(3) The boundary of the region of integration $\overline{U}^{\rho}_{Y_n}$ is a divisor on Y_n, namely,

$$\left(\bigcup_{i=1}^{n} Z_i\right) \cup \left(\bigcup_{j=1}^{n} Z'_j\right). \qquad \square$$

4D. ***Generating series and relations.*** In this subsection, we examine the generating series of iterated integrals on membranes (of types $\boldsymbol{b}'$ or $\boldsymbol{c}'$), evaluated at geodesic triangles and geodesic diangles. We prove relations among them. Most importantly, the generating series J will be used in Section 4E to define noncommutative Hilbert modular symbols. Moreover, the relations that we prove in this subsection will be interpreted as cocycles or as coboundaries of the noncommutative Hilbert modular symbols satisfy in Section 4E.

Definition 4.11. Let $f_1, \dots, f_m$ be m cusp forms with respect to a Hilbert modular group Γ. Let $f_1\, dz_1 \wedge dz_2, \dots, f_m\, dz_1 \wedge dz_2$ be the corresponding differential forms defining the generating series. Let $J(p_1, p_2, p_3)$ be the generating series J evaluated at the geodesic triangle with vertices p_1, p_2, p_3. Let $J(p_1, p_2; p_3, p_4)$ be the generating series J evaluated at the geodesic diangle $\{p_1, p_2; p_3, p_4\}$.

Both $J(p_1, p_2, p_3)$ and $J(p_1, p_2; p_3, p_4)$ will be called noncommutative Hilbert modular symbols after the action of the arithmetic group is included (see Definition 4.13).

Theorem 4.12. *The generating series J is one of the types $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{b}'$ or $\boldsymbol{c}'$. Note that $J(p_1, p_2; p_3, p_4)$ is defined for all types, while $J(p_1, p_2, p_3)$ is defined only for types $\boldsymbol{b}'$ or $\boldsymbol{c}'$. Then the generating series $J(p_1, p_2, p_3)$ and $J(p_1, p_2; p_3, p_4)$ satisfy the following relations*:

(1) *If σ is a permutation of the set $\{1, 2, 3\}$, then*

$$J(p_{\sigma(1)}, p_{\sigma(2)}, p_{\sigma(3)}) = J^{\operatorname{sign}(\sigma)}(p_1, p_2, p_3).$$

(2) *If p_1, p_2, p_3, p_4 are four points on the same holomorphic (or antiholomorphic) curve of type $\gamma^*\Delta$, then*

$$\begin{aligned} 1 = {} & J(p_1, p_2, p_3) J(p_2, p_3, p_4) \\ & \times J(p_2, p_1, p_4) J(p_1, p_4, p_3)\,. \end{aligned}$$

(3) *If p_1, p_2, p_3, p_4 are four points on the same holomorphic (or antiholomorphic) curve of type $\gamma^*\Delta$, then*

$$1 = J(p_1, p_2; p_3, p_4).$$

(4) *For every four points* p_1, p_2, p_3, p_4, *we have the following relation based on the orientation of the domain*:

$$\begin{aligned} J(p_2, p_1; p_3, p_4) &= J(p_1, p_2; p_4, p_3) \\ &= J^{-1}(p_2, p_1; p_4, p_3) \\ &= J^{-1}(p_1, p_2; p_3, p_4). \end{aligned}$$

(5) *For every five points* p_1, p_2, p_3, p_4, p_5, *we have*

$$J(p_1, p_2; p_3, p_4)J(p_1, p_2; p_4, p_5) = J(p_1, p_2; p_3, p_5).$$

(6) *For every four points* p_1, p_2, p_3, p_4, *we have the following relation*, *based on the boundary of a thickened tetrahedron*:

$$\begin{aligned} 1 = J&(p_1, p_2, p_3)J(p_2, p_3, p_4) \\ &\times J(p_2, p_1, p_4)J(p_1, p_4, p_3) \\ &\times J(p_1, p_2; p_3, p_4)J(p_2, p_3; p_1, p_4)J(p_3, p_1; p_2, p_4) \\ &\times J(p_3, p_4; p_1, p_2)J(p_1, p_4; p_2, p_3)J(p_2, p_4; p_3, p_1). \end{aligned}$$

Proof. For part (1), let σ be an odd permutation. Let U be the union of two triangles along one of their edges. Let the first triangle have vertices p_1, p_2, p_3 and the second triangle have vertices p_3, p_2, p_1 with the opposite orientation. We can glue the two triangles along the edge p_1p_2. (Gluing along any other edge would lead to the same result for the corresponding generating series.) From the shuffle product formula (Theorem 3.21(iii)), it follows that $J(U) = J(p_1, p_2, p_3)J(p_3, p_2, p_1)$. (Note that the product is not the product in the ring R. It is induced by a shuffle product of iterated integrals on membranes.) From the second homotopy invariance theorem (Theorem 3.11) it follows that the generating series $J(U)$ depends on U up to homotopy, which keeps the boundary components p_2p_3, p_3p_2, p_1p_3 and p_3p_1 on fixed unions of holomorphic curves. We can contract U to its boundaries ∂U so that the contracting homotopy keeps the boundary components on a fixed union of holomorphic curves. Therefore, $J(U) = J(\partial U) = 1$.

Parts (2), (4) and (5) can be proven similarly.

For part (3), if p_1, p_2, p_3, p_4 belong to the same holomorphic (or antiholomorphic) curve, then the corresponding diangle has no interior, since the two edges will coincide. Recall that the edges of the diangle are defined via unique geodesic triangles lying on a holomorphic (or antiholomorphic) curve.

The proof of part (6) is essentially the same as the one for part (1); however, we will prove it independently, since it is a key property of the noncommutative Hilbert modular symbol. Consider a thickened tetrahedron with vertices p_1, p_2, p_3, p_4.

The faces of the thickened tetrahedron are precisely the ones listed in the product of part (6). The whole product is equal to $J(V)$, where V is the union of all faces of the thickened tetrahedron. From the second homotopy invariance theorem it follows that the generating series $J(V)$ depends on V up to homotopy, which keeps the boundary components on a fixed union holomorphic curves. Since V bounds a contractible 3-dimensional region (a thickened tetrahedron), from Theorem 3.11, it follows that $J(V) = J(\text{point}) = 1$. □

4E. ***Definition of noncommutative Hilbert modular symbols.*** In this subsection, we define noncommutative Hilbert modular symbols. They are analogues of Manin's [2006] noncommutative modular symbol, applicable to the Hilbert modular group. Instead of the iterated path integrals that Manin uses, we use a higher-dimensional analogue, defined in Section 3.

Usually, a modular symbol represents a cohomology class. Manin's noncommutative modular symbol represents a noncommutative first cohomology class. We would like to say that the noncommutative Hilbert modular symbols represent noncommutative cohomology classes; this is formulated in Conjectures 4.14 and 4.15.

After defining the noncommutative Hilbert modular symbols, we prove some of their properties. These properties will be interpreted intuitively as cocycle or coboundary conditions. The approach in this subsection is more geometric. The purpose of presenting them here is to give many examples of relations and to help establish a suitable cohomology theory that truly captures these relations in a more structured way.

The cocycle interpretation is only for intuition; it is not precise. The formula holds for geometric reasons. Note that the composition is not the multiplication in the ring R; it is given by the shuffle product (see Theorem 3.21), which works for the generating series on iterated integrals on membranes. The multiplication is written linearly as we would multiply several elements in a group or in a ring; however, the multiplication is two-dimensional among regions with common boundaries.

In the next subsection will give some intuition about higher categories, for the purpose of giving more structure to the noncommutative Hilbert modular symbols and for a possible approach to defining a first and second noncommutative cohomology class.

For definitions of iterated integrals on membranes, see Definitions 3.4 and 3.5 for types $\boldsymbol{b}$ and $\boldsymbol{c}$ and Definition 4.9 for types $\boldsymbol{b}'$ and $\boldsymbol{c}'$.

Definition 4.13. We define noncommutative Hilbert modular symbols as generating series of iterated integrals on membranes of types $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{b}'$, or $\boldsymbol{c}'$ over a geodesic diangle by

$$c^1_{p_1,p_2;p_3}(\gamma) = J(p_1, p_2; p_3, \gamma p_3).$$

We also define noncommutative Hilbert modular symbols as generating series of iterated integrals on membranes of types $\boldsymbol{b}'$ or $\boldsymbol{c}'$ over a geodesic triangle by

$$c_p^2(\gamma, \delta) = J(p, \gamma p, \gamma\delta p),$$

where p, p_1, p_2, p_3 are cusp points in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$ and $\beta, \gamma, \delta \in \mathrm{SL}_2(\mathbb{O}_K)$.

We are going to define an action of $\mathrm{Mat}_2(\mathbb{O}_K)^+$ on the generating series J^c, where $\mathrm{Mat}_2(\mathbb{O}_K)^+$ is the semigroup of 2×2 matrices with totally positive determinant.

In order to interpret $c^1(\gamma)$ and $c^2(\gamma, \delta)$ as cocycles, we are going to define an action of the semigroup $\mathrm{Mat}_2(\mathbb{O}_K)^+$ on the whole ring R where the generating series take values. Such an action can be given via Hecke operators.

For simplicity, we shall assume that $\mathbb{O}_K$ has narrow class number 1. We consider all Hecke eigenforms of weight $(2, 2)$ with respect to $\mathrm{Mat}_2(\mathbb{O}_K)^+$. Now, let u be a unit such that $u_1 > 0$ and $u_2 < 0$, where u_1 and u_2 are the images of u under the two real embeddings of K into $\mathbb{R}$. It exists, since the narrow class group is trivial. (For example, $K = \mathbb{Q}(\sqrt{2})$ is such a field.) We define an action of $\gamma \in \mathrm{Mat}_2(\mathbb{O}_K)$ on the ring R (Definition 3.13) where the generating series takes values. We define

$$\gamma \bullet f \mapsto T_\gamma(f)$$

for $\gamma \in \mathrm{Mat}_2(\mathbb{O}_K)^+$. Let $f_1, \dots, f_m$ be a basis of Hecke eigenforms of the space of cusp form of weight $(2, 2)$. Let $X_1, Y_1, \dots, X_m, Y_m$ be generators of R, and to each f_i associate X_i and Y_i. Then the action of $\gamma \in \mathrm{Mat}_2(\mathbb{O}_K)^+$ is given by $T_\gamma(X_i) = c(\gamma, f_i)X_i$ and $T_\gamma(Y_i) = Y_i$, where $c(\gamma, f_i)$ is the eigenvalue of the Hecke operator.

In this setting the group action, namely, the action of the Hilbert modular group, is trivial. This trivial action extend to the action of $T_1 = \mathrm{id}$ on the whole ring R. In fact, for an element $\beta \in \mathrm{SL}_2(\mathbb{O}_K)$, the trivial action on $c^1_{p_1,p_2;p_3}$ and c^2_p can be realized as

$$(\beta c^1_{p_1,p_2;p_3})(\gamma) = c^1_{\beta p_1,\beta p_2;\beta p_3}(\beta\gamma)$$

and

$$(\beta c^2_p)(\gamma, \delta) = c^2_{\beta p}(\beta p, \beta\gamma p, \beta\gamma\delta p).$$

The last two relations hold because for a cusp form of weight $(2, 2)$ the differential form $f\,dz_1 \wedge dz_2$ is invariant under the action of the Hilbert modular group Γ. Algebraically, for any geodesic diangle, we have

$$\beta J(p_1, p_2; p_3, p_4) = J(p_1, p_2; p_3, p_4) = J(\beta p_1, \beta p_2; \beta p_3, \beta p_4).$$

Similarly, for a geodesic triangle,

$$\beta J(p_1, p_2, p_3) = J(p_1, p_2, p_3) = J(\beta p_1, \beta p_2, \beta p_3).$$

The relations among the symbols are based on two properties: composition via shuffle product Theorem 3.21(iii) and homotopy invariance (Theorems 3.7 and 3.11).

Conjecture 4.14. The noncommutative Hilbert modular symbol $c^1_{p_1,p_2;p_3}$ is a 1-cocycle. Moreover, if we change the point p_3 to q_3, then the cocycle changes by a coboundary.

Property (5) of Theorem 4.12 can be interpreted as a 1-cocycle relation. Consider the analogy with a noncommutative 1-cocycle of a group acting on a noncommutative ring; we define the boundary of $c^1_{p_1,p_2;p_3}$ by

$$dc^1_{p_1,p_2;p_3}(\beta,\gamma)=c^1_{p_1,p_2;p_3}(\beta)(\beta c^1_{p_1,p_2;p_3})(\gamma)(c^1_{p_1,p_2;p_3}(\beta\gamma))^{-1}.$$

The action of β on the cocycle is given in Definition 4.13. In contrast to a first noncommutative cocycle (see for example [Brown 1982]), here we have two-dimensional composition of symbols, that is, one can compose the symbols as two-morphisms in a two-category.

Then

$$\begin{aligned} dc^1_{p_1,p_2;p_3}(\beta,\gamma) &= J(p_1,p_2;p_3,\beta p_3)(\beta J(p_1,p_2;p_3,\gamma p_3))J^{-1}(p_1,p_2;p_3,\beta\gamma p_3)\\ &= J(p_1,p_2;p_3,\beta p_3)J(p_1,p_2;\beta p_3,\beta\gamma p_3)J^{-1}(p_1,p_2;p_3,\beta\gamma p_3)\\ &= 1. \end{aligned} \tag{26}$$

If we change p_3 to q_3 then the cocycle changes by a coboundary. Let $b^0 = J(p_1,p_2;p_3,q_3)$ be a 0-cochain. Then

$$\begin{aligned} c^1_{p_1,p_2;q_3}(\gamma) &= J(p_1,p_2;p_3,\gamma p_3)\\ &= J(p_1,p_2;p_3,q_3)J(p_1,p_2;q_3,\gamma q_3)J(p_1,p_2;\gamma q_3,\gamma p_3)\\ &= J(p_1,p_2;p_3,q_3)J(p_1,p_2;q_3,\gamma q_3)(\gamma J(p_1,p_2;p_3,q_3))^{-1}\\ &= b^0 c^1_{p_1,p_2;q_3}(\gamma)(\gamma b^0)^{-1} \end{aligned} \tag{27}$$

Conjecture 4.15. The noncommutative Hilbert modular symbol $c^2_p(\beta,\gamma)$ satisfies a 2-cocycle relation. Moreover, if we change the point p to q, then the cocycle changes by a coboundary up to terms involving c^1.

Recall

$$c^2_p(\beta,\gamma)=J(p,\beta p,\beta\gamma p).$$

Then c_p^2 satisfies a 2-cocycle condition up to a multiple of the 1-cocycle $c^1_{q_1,q_2;q_3}$ for various points q_1, q_2, q_3. For the 2-cocycle relation, we compute $dc_p^2(\beta,\gamma,\delta)$:

$$\begin{aligned} dc_p^2(\beta,\gamma,\delta) &= c_p^2(\beta,\gamma)c^2(\beta,\gamma\delta)(c^2(\beta\gamma,\delta))^{-1}(\beta\cdot c^2(\gamma,\delta))^{-1} \\ &= J(p,\beta p,\beta\gamma p)J(p,\beta p,\beta\gamma\delta p) \\ &\quad \times J(p,\beta\gamma p,\beta\gamma\delta p)^{-1}J(\beta p,\beta\gamma p,\beta\gamma\delta p)^{-1}. \end{aligned} \tag{28}$$

In order to have $dc_p^2(\beta,\gamma,\delta)=1$, we must multiply by suitable values of c^1, corresponding to edges of a certain thickened tetrahedron. Then

$$\begin{aligned} & dc_p^2(\beta,\gamma,\delta)\times\big[c^1_{p,\beta p;\beta\gamma p}((\beta\gamma)\delta(\beta\gamma)^{-1})c^1_{\beta p,\beta\gamma p;p}(\beta\gamma\delta)c^1_{\beta\gamma p,p;\beta p}((\beta)\gamma\delta\beta^{-1}) \\ &\qquad\qquad \times c^1_{\beta\gamma p,\beta\gamma\delta p;p}(\beta)c^1_{p,\beta\gamma\delta p;\beta p}(\beta\gamma\beta^{-1})c^1_{\beta p,\beta\gamma\delta p;\beta\gamma p}((\beta\gamma)^{-1})\big] \\ &= [c_p^2(\beta,\gamma)c^2(\beta,\gamma\delta)(c^2(\beta\gamma,\delta))^{-1}(\beta\cdot c^2(\gamma,\delta))^{-1}] \\ &\quad \times\big[c^1_{p,\beta p;\beta\gamma p}((\beta\gamma)\delta(\beta\gamma)^{-1})c^1_{\beta p,\beta\gamma p;p}(\beta\gamma\delta)c^1_{\beta\gamma p,p;\beta p}((\beta)\gamma\delta\beta^{-1}) \\ &\qquad\qquad \times c^1_{\beta\gamma p,\beta\gamma\delta p;p}(\beta)c^1_{p,\beta\gamma\delta p;\beta p}(\beta\gamma\beta^{-1})c^1_{\beta p,\beta\gamma\delta p;\beta\gamma p}((\beta\gamma)^{-1})\big] \\ &= [J(p,\beta p,\beta\gamma p)J(p,\beta p,\beta\gamma\delta p)J(p,\beta\gamma p,\beta\gamma\delta p)^{-1}J(\beta p,\beta\gamma p,\beta\gamma\delta p)^{-1}] \\ &\quad \times\big[J(p,\beta p;\beta\gamma p,\beta\gamma\delta p)J(\beta p,\beta\gamma p;p,\beta\gamma\delta p)J(\beta\gamma p,p;\beta p,\beta\gamma\delta p) \\ &\qquad \times J(\beta\gamma p,\beta\gamma\delta p;p,\beta p)J(p,\beta\gamma\delta p;\beta p,\beta\gamma p)J(\beta p,\beta\gamma\delta p;\beta\gamma p,p)\big] \\ &= 1. \end{aligned}$$

The first equality follows from (28). The second equality follows from the definition of the symbols. And the last equality follows from property (6) of Theorem 4.12 with $(p_1,p_2,p_3,p_4)=(p,\beta p,\beta\gamma p,\beta,\gamma\delta p)$. Therefore, we obtain that $dc_p^2(\beta,\gamma,\delta)$ is 1 up to values of the 1-cocycle c^1.

Conjecture 4.16. The conjectural cocycles c_p^2 and c_q^2 are homologous:

$$c_p^2(\beta,\gamma)=c_q^2(\beta,\gamma)[db^1_{pq}(\beta,\gamma)]\prod_i J(D_i),$$

up to a product of $J(D_i)$, where the D_i are geodesic diangles.

Before we proceed, we would like to make an analogy between 1-dimensional and 2-dimensional cocycles. For the 1-dimensional cocycle, the property that it is a cocycle uses the geometry of a triangle, where the faces of the triangle are essentially the 1-cocycle. We want commutativity of the triangular diagram. We think of the commutativity of the diagram as follows: consider the interior of the triangle as a homotopy of paths and think of the 1-cocycle as a homotopy-invariant function. The 2-cocycle relation is represented by the faces of a tetrahedron. By "commutativity" of the diagram, we mean a homotopy invariant 2-cocycle and a homotopy from one of the faces to the union of the other three faces.

The comparison that $c^1_{p_1,p_2;p_3}$ and $c^1_{p_1,p_2;q_3}$ are homologous is given by a square-shaped diagram. The analogy with dimension 2 is that the cocycles c^2_p and c^2_q are two faces of an octahedron. The vertices associated to $c^2_p(\beta,\gamma)$ are $(p,\beta p,\beta\gamma p)$ and the vertices associated to c^2_q are $(q,\beta q,\beta\gamma q)$. The two faces will be opposite to each other on the octahedron Oct so that the three pairs of opposite vertices are $(p,\beta\gamma q)$, $(\beta p,q)$ and $(\beta\gamma p,\beta q)$. The remaining six faces are combined into two triples. Each of them corresponds to a coboundary of a 1-chain.

Let

$$b^1_{p,q}(\beta)=[J(p,q,\beta p)J(q,\beta q,\beta p)][J(q,\beta p;\, p,\beta q)].$$

Consider the action of $\gamma\in\Gamma$ on b^1 by acting on each point in the argument of J, denoted as before by $\gamma\cdot b^1$. Then, we define

$$db^1_{p,q}(\beta,\gamma)=b^1_{p,q}(\beta)[\beta\cdot b^1_{p,q}(\gamma)][b^1_{p,q}(\beta\gamma)]^{-1},$$

where $\beta\cdot b^1_{p,q}(\gamma)=[J(\beta p,\beta q,\beta\gamma p)J(\beta q,\beta\gamma q,\beta\gamma p)][J(\beta q,\beta\gamma p;\, p;\,\beta\gamma q)]$.

Consider the above octahedron Oct. Remove from it the tetrahedron T with vertices $(p,q,\beta\gamma q,\beta\gamma p)$. Then the triangles of the remaining geometric figure are precisely the triangles in the definitions of $c^2_p(\beta,\gamma)$, $c^2_q(\beta,\gamma)$ and $db^1_{p,q}(\beta,\gamma)$. Now, consider thickenings of the edges, which are common for two triangles. It can be done in the following way. Instead of any triangle, we can take a geodesic triangle. The two triangles that had a common edge might have only two common vertices. Then the region between the two geodesic, one for each of the geodesic triangles, forms the induced diangle. Take J of the induces diangles from the octahedron Oct and J^{-1} of the induced diangles from the tetrahedron T. Their product gives $\prod_i J(D_i)$. The equality holds because we apply J to the union of the faces of the thickened Oct $-T$, which gives 1.

4F. ***A two-category.*** Why do we need a two-category? Is there an example of a sheaf on this category/topology? How does the noncommutative Hilbert modular symbols represents a sheaf?

The ideas presented in this subsection will be developed in a follow-up paper. Here we present the basic constructions that give justification for the conjectures that the noncommutative Hilbert modular symbols c^1 and c^2 are cocycles in some categorical and sheaf-theoretic setting. For sheaves on 2-categories one may consult [Street 1982]. Since our 2-morphisms are invertible one may also use Lurie's constructions [2009] of sheaves on higher categories.

We are going to construct a 2-category C and a sheaf J on C. We define p to be an object of the 2-category C if p is a cusp point, that is $p\in\mathbb{P}^1(K)$. We define 1-morphisms in the following way. Let σ be the geodesic connecting 0 and ∞ that lies on the diagonal $\Delta=i(\mathbb{H})\subset\mathbb{H}\times\mathbb{H}$. There is unique such geodesic. All geodesics $\gamma^*\sigma$ together with a choice of orientation are defined to be 1-morphisms, where $\gamma \in \mathrm{PGL}_2(K)$. We define the 1-morphisms of C to be finite concatenations of geodesics of type $\gamma^*\sigma$ or the trivial path whose image coincides with a cusp point. Consider ideal triangles and ideal diangles as cells from which we build manifolds with corners. A 2-morphism is a finite union of manifolds with corners, made from finitely many ideal diangles and ideal triangles, which is path-connected and has orientation.

The boundary of a 1-morphism is a union of two objects — the starting point and the ending point of the directed path. The boundary of a 2-morphism (a 2-manifold with corners) is a finite union of 1-morphisms (oriented loops), where the orientation of the loops on the boundary is induced by the orientation of the 2-manifold with corners.

Now we are going to define a 2-sheaf J, whose values on a 2-morphism will be in a subset of the ring R and whose values on an object and on a 1-morphism will be a subset of a countable product of the ring R with itself.

As always, $S_{2,2}(\Gamma)$ denotes the space of cusp forms of weight (2, 2) with respect to the group Γ. Here we will consider this space as the space of holomorphic 2-forms on $\mathbb{H} \times \mathbb{H}$ which vanish on the cusps and which can descend to the Hilbert modular surface $X_\Gamma = \Gamma\backslash(\mathbb{H}^2 \cup \mathbb{P}^1(K))$. Every n-tuple of such holomorphic forms $\Omega \in (S_{2,2}(\Gamma))^n$ defines a value of a 2-morphism f in C. Let this value be the generating series $J_f(\Omega)$. Let J_f be the collection of all values $J_f(\Omega)$ for $\Omega \in (S_{2,2}(\Gamma))^n$. Let e be a 1-morphism. We say that e is in the boundary of a 2-morphism f, denoted $e \subset \partial f$, if the image of the loop e is in the boundary of the image of the membrane f together with the induced orientation on e from f. We say that an object p is in the boundary of a 1-morphism e, denoted by $p \in \partial e$, if p is a source or a target of e. We define the values of J on a 1-morphism e to be the product

$$\prod_{e \subset \partial f} J_f \subset \prod_{e \subset \partial f} R.$$

We define the values of J on objects p to be

$$\prod_{\substack{p \in \partial e \\ e \subset \partial f}} J_f \subset \prod_{\substack{p \in \partial e \\ e \subset \partial f}} R.$$

The sheaf conditions for 1- and 2-morphisms resemble the classical conditions for a presheaf to be a sheaf.

Let $f_i : A_i \to [0, 1]^2$ be a finite collection of disjoint 2-morphisms, whose union is a morphism $f : A \to [0, 1]^2$. We define a finite collection f_{ij}^k of 1-morphisms and 0-morphisms (objects) such that the union $\bigcup_k \mathrm{im}(f_{ij}^k) = \mathrm{im}(f_i) \cap \mathrm{im}(f_j)$ is a disjoint union of the intersection.

Then the equalizer

$$J_f \to \prod_i J_{f_i} \rightrightarrows \prod_{ijk} J_{f_{ij}^k}$$

is exact (for a definition of equalizer one may consult [Borceux 1994]).

Similarly, let e be a 1-morphism and let $\{e_i\}$ be a finite set of disjoint 1-morphisms such that the union $\bigcup_i \mathrm{im}(e_i)$ is equal to $\mathrm{im}(e)$. We can write the intersection $\mathrm{im}(e_i) \cap \mathrm{im}(e_j)$ as a finite union of 0-morphisms $\bigcup_k \mathrm{im}(e_{ij}^k)$, for some 1-morphisms e_{ij}^k.

Then the equalizer

$$J_e \to \prod_i J_{e_i} \rightrightarrows \prod_{ijk} J_{e_{ij}^k}$$

is exact.

The cochain is defined as

$$\prod_{p\colon \text{0-morph}} J_p \to \prod_{e\colon \text{1-morph}} J_e \to \prod_{f\colon \text{2-morph}} J_f \to \prod_{\substack{g\colon \text{2-morph}\\ \partial g=\varnothing}} J_g$$

The maps $J_e \to J_p$ and $J_p \to J_f$ are surjective when they are defined, resembling flabby sheaves. Thus, we should have trivial zeroth and first cohomology set. The only nontrivial cohomology will be the second cohomology set. The cocycle conditions for both noncommutative Hilbert modular symbols c^1 and c^2 can be interpreted as a particular case of maps

$$\prod_{f\colon \text{2-morph}} J_f \to \prod_{\substack{g\colon \text{2-morph}\\ \partial g=\varnothing}} J_g.$$

For c^1, the boundary condition is that a union of two diangles with a common edge is a third diangle. One can think of the these three diangles as the boundary of a degenerate 3-dimensional region. One can realize this cocycle condition as a sheaf-theoretic one by modifying the above definition so that the 2-morphisms consists of a finite union of ideal diangles (without using the ideal triangles). Then the sheaf-theoretic second cocycle condition is the one for noncommutative Hilbert modular symbol c^1.

If we are able to quotient the 2-category described in the beginning of this subsection by the 2-morphisms generated by diangles, then we have only two morphisms generated by ideal triangles. The noncommutative Hilbert modular symbol c^2 is exactly the one that considers ideal triangles. Note that its cocycle relation for c^2 is satisfied up to 2-morphisms generated by diangles.

4G. ***Explicit computations. Multiple Dedekind zeta values.*** In this subsection, we make explicit computations of some ingredients in the noncommutative Hilbert modular symbol. Manin [2006] compared explicit formulas of integrals in the noncommutative modular symbol to multiple zeta values. The similarities are in terms of both infinite series formulas and formulas via iterated path integrals. Here we compare certain integrals in the noncommutative Hilbert modular symbol to multiple Dedekind zeta values (for multiple Dedekind zeta values, see [Horozov 2014b]). Again the similarities are in terms of both infinite series formulas and formulas via iterated integrals over membranes.

We are going to consider the Fourier expansion of two Hilbert cusp forms f and g. Let $\omega_f = f\,dz_1 \wedge dz_2$, $\omega_g = g\,dz_1 \wedge dz_2$ and $\omega_0 = dz_1 \wedge dz_2$. We are going to associate L-values to iterated integrals of the forms ω_f and ω_g. The L-values will be iterated integrals over an union of diangles. One can think of a diangle connecting 0 and ∞ as a segment or as a real cone. The union will be a disjoint union of all such real cones connecting 0 and ∞ or simply $\mathrm{Im}(\mathbb{H}) \times \mathrm{Im}(\mathbb{H})$. We also recall the definition of a multiple Dedekind zeta values via (discrete) cones. Finally, we prove analogous formulas for iterated L-values associated to Hilbert cusp forms and for multiple Dedekind zeta values.

We will be mostly interested in the modular symbol associated to a diangle. Let us recall what we mean by a diangle.

Let p_1, p_2, p_3, p_4 be four cusp points. Let $\gamma_1 \in \mathrm{GL}_2(K)$ be a linear fractional transformation that sends $\gamma_1(p_1)$ to 0, $\gamma_1(p_2)$ to ∞, and $\gamma_1(p_3)$ to 1. Let Δ be the image of the diagonal embedding of $\mathbb{H}^1$ into $\mathbb{H}^2$. Then 0, 1 and ∞ are boundary points of Δ. Let $\lambda(0, \infty)$ be the unique geodesic in Δ that connects 0 and ∞. And let

$$\lambda_1(p_1, p_2) = \gamma_1^{-1}\lambda(0, \infty)$$

be the pullback of the geodesic λ to a geodesic connecting p_1 and p_2.

Now consider the triple p_1, p_2 and p_4. Let $\gamma_2 \in \mathrm{GL}_2(K)$ be a linear fractional transformation that sends $\gamma_2(p_1)$ to 0, $\gamma_2(p_2)$ to ∞ and $\gamma_2(p_4)$ to 1. Let Δ be the image of the diagonal embedding of $\mathbb{H}^1$ into $\mathbb{H}^2$. Then 0, 1 and ∞ are boundary points of Δ. Let $\lambda(0, \infty)$ be the unique geodesic in Δ that connects 0 and ∞. And let

$$\lambda_2(p_1, p_2) = \gamma_2^{-1}\lambda(0, \infty)$$

be the pullback of the geodesic λ to a geodesic connecting p_1 and p_2.

By a diangle, we mean a region in $\mathbb{H}^2 \cup \mathbb{P}^1(K)$ with the homotopy type of a disc, bounded by the geodesics $\lambda_1(0, \infty)$ and $\lambda_2(0, \infty)$.

We are going to present a computation for the diangle D_u defined by the points $(0, \infty, u^1, u^{-1})$, where u is a generator for the group of units modulo ± 1 in K. Let (1) be the trivial permutation.

Lemma 4.17. *Let u be a totally positive unit. Then*

$$\iint_{D_u}^{(1)(1)} e^{2\pi i(\alpha_1 z_1+\alpha_2 z_2)}\, dz_1\wedge dz_2 = \frac{1}{(2\pi i)^2}\frac{u_2^2-u_1^2}{(\alpha_1u_1+\alpha_2u_2)(\alpha_1u_2+\alpha_2u_1)}.$$

Proof. Let u_1 and u_2 be the two embeddings of u into $\mathbb{R}$. Then $(0,\infty,u)$ can be mapped to $(0,\infty,1)$ by $\gamma_1=\left(\begin{smallmatrix}u^{-1} & 0\\ 0 & 1\end{smallmatrix}\right)$. The geodesic $\lambda(0,\infty)$ can be parametrized by (it, it) for $t\in\mathbb{R}$. Then the geodesic $\lambda_1(0,\infty)$ on the geodesic triangle $(0,\infty,u)$ can be parametrized by $\{(iu_1t, iu_2t)\mid t>0\}$. Similarly, the geodesic $\lambda_2(0,\infty)$ on the geodesic triangle $(0,\infty,u^{-1})$ can be parametrized by $\{(iu_2t, iu_1t)\mid t>0\}$. Then the diangle D_u can be parametrized by

$$\left\{(z_1,z_2)\in\mathbb{H}^2 \mid \mathrm{Re}(z_1)=\mathrm{Re}(z_2)=0,\ \mathrm{Im}(z_1)\in\left(\frac{u_1}{u_2}t,\frac{u_2}{u_1}t\right),\ \mathrm{Im}(z_2)=t\in(0,\infty)\right\}.$$

Then we have

$$\begin{aligned}\iint_{D_u}^{(1)(1)} e^{2\pi i(\alpha_1 z_1+\alpha_2 z_2)}\, dz_1\wedge dz_2 &= \int_\infty^0\left(\int_{\frac{u_2}{u_1}t}^{\frac{u_1}{u_2}t} e^{2\pi i(\alpha_1 z_1+\alpha_2 t)}\,dz_1\right)dt\\ &=\frac{1}{2\pi i\alpha_1}\int_\infty^0\left(e^{\alpha_1\frac{u_1}{u_2}t+\alpha_2 t}-e^{\alpha_1\frac{u_2}{u_1}t+\alpha_2 t}\right)dt\\ &=\frac{1}{(2\pi i)^2}\frac{1}{\alpha_1}\left(\frac{1}{\alpha_1\frac{u_1}{u_2}+\alpha_2}-\frac{1}{\alpha_1\frac{u_2}{u_1}+\alpha_2}\right)\\ &=\frac{1}{(2\pi i)^2}\frac{u_2^2-u_1^2}{(\alpha_1u_1+\alpha_2u_2)(\alpha_1u_2+\alpha_2u_1)}\end{aligned}$$ □

Therefore, one term of the Fourier expansion of a Hilbert cusp form paired with a symbol given by one diangle does not resemble a norm of an algebraic integer. However, if we integrate over an infinite union of diangles, then a similarity with Dedekind zeta and multiple Dedekind zeta values occurs.

Consider the limit of D_{u^n} when $n\to\infty$. It is the product of the two imaginary axes of the two upper half-planes. Set

$$\mathrm{Im}(\mathbb{H}^2)=\mathrm{Im}(\mathbb{H})\times\mathrm{Im}(\mathbb{H}).$$

One can think of this region as an infinite union of diangles.

Denote by αz the sum of products $\alpha_1z_1+\alpha_2z_2$. Using the methods of [Horozov 2014b, Section 1], we obtain

$$\frac{(2\pi i)^{-2}}{N(\alpha)N(\alpha+\beta)}=\int_{\mathrm{Im}(\mathbb{H}^2)}^{(1)(1)} e^{2\pi i\alpha z}\,dz_1\wedge dz_2\cdot e^{2\pi i\beta z}\,dz_1\wedge dz_2$$

and

$$\frac{1}{(2\pi i)^2}\frac{1}{N(\alpha)^3 N(\alpha+\beta)^2}$$
$$= \int_{\mathrm{Im}(\mathbb{H}^2)}^{(1)(1)} e^{2\pi i\alpha z}\, dz_1\wedge dz_2\cdot(dz_1\wedge dz_2)\cdot(dz_1\wedge dz_2)\cdot e^{2\pi i\beta z}\, dz_1\wedge dz_2\cdot(dz_1\wedge dz_2).$$

Let f and g be two cusp forms of weights $(2k, 2k)$ and $(2l, 2l)$, respectively. Consider the Fourier expansion of both of the cusp forms. Let

$$f = \sum_{\alpha\gg 0} a_\alpha e^{2\pi i\alpha z} \quad \text{and} \quad g = \sum_{\beta\gg 0} b_\beta e^{2\pi i\beta z}.$$

Since f is of weight $(2k, 2k)$, we have that $a_{u\alpha} = a_\alpha$, where u is a unit. For such a modular form the modular factor with respect to the transformation $z \to uz$ is 1. The L-values of f are

$$L_f(n) = \int_{\mathrm{Im}(\mathbb{H}^2)}^{(1)(1)} \sum_{\alpha\in\mathbb{O}_K^+/U^+} a_\alpha e^{2\pi i\alpha z}\, dz_1\wedge dz_2\cdot(dz_1\wedge dz_2)^{\cdot(n-1)}$$
$$= \frac{1}{(2\pi i)^{2n}} \sum_{\alpha\in\mathbb{O}_K^+/U^+} \frac{a_\alpha}{N(\alpha)^n}.$$

Here $\mathbb{O}_K^+$ denotes the totally positive algebraic integers in K and U^+ denotes the totally positive units.

We recall some of the definitions from [Horozov 2014b]. We fix a positive cone C in $\mathbb{O}_K$, by which we mean

$$C = \mathbb{N}\cup\{\alpha\in\mathbb{O}_K \mid a+b\epsilon,\ a, b\in\mathbb{N}\},$$

where ϵ is a generator of the group of totally positive units. By $\epsilon^k C$, we mean the collection of products $\epsilon^k\alpha$, where α varies in the cone C.

The following infinite sum is an example of a multiple Dedekind zeta value:

$$\zeta_{K;C,\epsilon^k C}(m, n) = \sum_{\alpha\in C}\sum_{\beta\in\epsilon^k C} \frac{1}{N(\alpha)^m N(\alpha+\beta)^n}.$$

Let $Z(m, n) = \sum_{k\in\mathbb{Z}} \zeta_{K;C,\epsilon^k C}(m, n)$, where C is any set representing the totally positive algebraic integers $\mathbb{O}_K^+$ modulo totally positive units U^+.

Lemma 4.18. *The values $Z(m, n)$ are finite for $m > n > 1$.*

Proof. Let ϵ be a generators of the group of totally positive units U^+ in K. For the two real embeddings ϵ_1 and ϵ_2 of ϵ, we can assume that $\epsilon_1 > 1 > \epsilon_2$. Otherwise we can take its reciprocal.

$$Z(m,n) = \sum_{k\in Z}\sum_{\alpha,\beta\in C} \frac{1}{N(\alpha)^m N(\alpha+\epsilon^k\beta)^n} \tag{29}$$

$$< \sum_{\alpha,\beta\in C} \frac{1}{N(\alpha)^m}\left(\frac{1}{N(\alpha+\beta)^n} + \sum_{k=1}^{\infty}\frac{2^n}{\epsilon_1^k}\left(\frac{1}{\alpha_1^n\beta_2^n}+\frac{1}{\alpha_2^n\beta_1^n}\right)\right) \tag{30}$$

$$< \sum_{\alpha,\beta\in C} \frac{1}{N(\alpha)^m}\left(\frac{1}{N(\alpha+\beta)^n} + \sum_{k=1}^{\infty}\frac{2}{\epsilon_1^k}\left(\frac{N(\alpha+\beta)^n - N(\alpha)^n}{N(\alpha+\beta)^n}\right)\right) \tag{31}$$

$$= \sum_{\alpha,\beta\in C} \frac{1}{N(\alpha)^m}\left(\frac{1}{N(\alpha+\beta)^n} + \frac{2}{\epsilon_1-1}\left(\frac{N(\alpha+\beta)^n - N(\alpha)^n}{N(\alpha+\beta)^n}\right)\right) \tag{32}$$

$$= \sum_{\alpha,\beta\in C} \frac{1}{N(\alpha)^m N(\alpha+\beta)^n} + \frac{2}{\epsilon_1-1}\frac{1}{N(\alpha)^n} - \frac{2}{N(\alpha)^{m-n}N(\alpha+\beta)^n} \tag{33}$$

$$= \zeta_K(C;m,n) - \frac{2}{\epsilon_1-1}(\zeta_K(C;n)+\zeta_K(C;m-n,n)). \tag{34}$$

Equation (29) is the definition. Inequality (30) is based on the following: $\epsilon_2 < 1$ is replaced with 1 when $k > 0$. For $k < 0$ we use $\epsilon_2^k = \epsilon_1^{-k}$. We put 1 for ϵ_1^k for $k < 0$. The case $k = 0$ is treated separately. Finally we group the terms with equal powers of ϵ_1. In inequality (31) we estimate the mixed terms in the brackets. In (32) we take the sum of the geometric series in ϵ_1^{-1}. Then in (33) we open the brackets. And finally, in (34), we express the sums as a finite linear combinations of a Dedekind zeta value and multiple Dedekind zeta values. □

The following definition of an iterated L-value is a coefficient of one monomial from the noncommutative Hilbert modular symbol of type $\boldsymbol{b}$:

Definition 4.19. For a pair of Hilbert cusp forms f and g with Fourier expansions

$$f = \sum_{\alpha\gg 0} a_\alpha e^{2\pi i\alpha z} \quad \text{and} \quad g = \sum_{\beta\gg 0} b_\beta e^{2\pi i\beta z},$$

we define iterated L-values

$$L_{f,g}(m,n) = \int_{\mathrm{Im}(\mathbb{H}^2)}^{(1)(1)} \sum_{(\alpha,\beta)\in(\mathbb{O}_K^+,\mathbb{O}_K^+)/U} (a_\alpha e^{2\pi i\alpha z}\, dz_1\wedge dz_2)\cdot(dz_1\wedge dz_2)^{\cdot(m-1)} \cdot (b_\beta e^{2\pi i\beta z}\, dz_1\wedge dz_2)\cdot(dz_1\wedge dz_2)^{\cdot(n-1)}.$$

Theorem 4.20. *Using the above definition, we have*

$$L_{f,g}(m,n)=\sum_{k\in\mathbb{Z}}\sum_{\alpha\in C,\beta\in\epsilon^k C}\frac{a_\alpha b_\beta}{N(\alpha)^m N(\alpha+\beta)^n}.$$

Proof.

$$\begin{aligned}L_{f,g}(m,n)&=\int_{\mathrm{Im}(\mathbb{H}^2)}^{(1)(1)}\sum_{(\alpha,\beta)\in(\mathbb{O}_K^+,\mathbb{O}_K^+)/U}(a_\alpha e^{2\pi i\alpha z}\,dz_1\wedge dz_2)\cdot(dz_1\wedge dz_2)^{\cdot(m-1)}\\&\qquad\qquad\cdot(b_\beta e^{2\pi i\beta z}\,dz_1\wedge dz_2)\cdot(dz_1\wedge dz_2)^{\cdot(n-1)}\\&=\sum_{(\alpha,\beta)\in(\mathbb{O}_K^+,\mathbb{O}_K^+)/U}\frac{a_\alpha b_\beta}{N(\alpha)^m N(\alpha+\beta)^n}\\&=\sum_{k\in\mathbb{Z};\alpha,\beta\in C}\frac{a_\alpha b_\beta}{N(\alpha)^m N(\alpha+\epsilon^k\beta)^n}\\&=\sum_{k\in\mathbb{Z},}\sum_{\alpha\in C}\sum_{\beta\in\epsilon^k C}\frac{a_\alpha b_\beta}{N(\alpha)^m N(\alpha+\beta)^n}.\end{aligned}$$

□

We would like to bring to the attention of the reader Definition 4.19, the definition of the multiple L- values. More specifically, we would like to point out that the region of integration is an infinite union of diangles (or equivalently an infinite union of real cones; see the beginning of this section). Note also that in Theorem 4.20 the values of the multiple L-functions are expressed as an infinite sums over different discrete cones, namely, over $\epsilon^k C$ for $k\in\mathbb{Z}$. However, a single real cone D_u, as in Lemma 4.18, does not correspond to a single discrete cone. Only a good union of real cones $\mathrm{Im}(\mathbb{H})\times\mathrm{Im}(\mathbb{H})$ corresponds to a good union of discrete cones $\bigcup_{k\in\mathbb{Z}}(C,\epsilon^k C)$ as a fundamental domain of $(\mathbb{O}_K^+,\mathbb{O}_K^+)/U^+$.

Acknowledgements

I would like to thank Ronnie Brown for his talk on higher categories at the University of Durham, which helped me to define iterated integrals over membranes. I would like to thank Mladen Dimitrov for his interest in my approach to Hilbert modular symbols. Most of all I would like to thank to Yuri Manin for his inspiring talk at the Max Planck Institute for Mathematics in Bonn on noncommutative modular symbols and for his encouragement and interest in the work presented here. Finally, I would like to thank the referees for the useful suggestions and commentaries.

I would like to thank the Max Planck Institute for Mathematics in Bonn and the Marie Curie Research Training Network, Arithmetic Algebraic Geometry Network at Durham University for financial support in 2004–2005 and 2005–2006, respectively.

References

[Ash and Borel 1990] A. Ash and A. Borel, "Generalized modular symbols", pp. 57–75 in *Cohomology of arithmetic groups and automorphic forms* (Luminy-Marseille, 1989), edited by J.-P. Labesse and J. Schwermer, Lecture Notes in Math. **1447**, Springer, Berlin, 1990. MR 92e:11058 Zbl 0719.11033

[Ash and Rudolph 1979] A. Ash and L. Rudolph, "The modular symbol and continued fractions in higher dimensions", *Invent. Math.* **55**:3 (1979), 241–250. MR 82g:12011 Zbl 0426.10023

[Berger et al. 2013] L. Berger, G. Böckle, L. Dembélé, M. Dimitrov, T. Dokchitser, and J. Voight, *Elliptic curves, Hilbert modular forms and Galois deformations*, Birkhäuser/Springer, Basel, 2013. MR 3155300 Zbl 1276.11004

[Birch 1971] B. J. Birch, "Elliptic curves over Q: a progress report", pp. 396–400 in *1969 Number Theory Institute* (Stony Brook, 1969), edited by D. J. Lewis, Proc. Symp. Pure Math. **20**, Amer. Math. Soc., Providence, R.I., 1971. MR 47 #3395 Zbl 0214.19801

[Borceux 1994] F. Borceux, *Handbook of categorical algebra I: Basic category theory*, Encyclopedia of Mathematics and its Applications **50**, Cambridge University Press, 1994. MR 96g:18001a Zbl 0803.18001

[Borel and Serre 1973] A. Borel and J.-P. Serre, "Corners and arithmetic groups", *Comment. Math. Helv.* **48** (1973), 436–491. MR 52 #8337 Zbl 0274.22011

[Brown 1982] K. S. Brown, *Cohomology of groups*, Graduate Texts in Mathematics **87**, Springer, Berlin, 1982. MR 83k:20002 Zbl 0584.20036

[Bruinier et al. 2008] J. H. Bruinier, G. van der Geer, G. Harder, and D. Zagier, *The 1-2-3 of modular forms: lectures at a summer school in Nordfjordeid, Norway*, Springer, Berlin, 2008. MR 2009d:11002 Zbl 1197.11047

[Chen 1977] K. T. Chen, "Iterated path integrals", *Bull. Amer. Math. Soc.* **83**:5 (1977), 831–879. MR 56 #13210 Zbl 0389.58001

[Cremona 1997] J. E. Cremona, *Algorithms for modular elliptic curves*, 2nd ed., Cambridge University Press, 1997. MR 99e:11068 Zbl 0872.14041

[Drinfeld 1973] V. G. Drinfeld, "Two theorems on modular curves", *Funkcional. Anal. i Prilozhen.* **7**:2 (1973), 83–84. In Russian; translated in *Functional Anal. Appl.* **7**:2 (1973), 155–156. MR 47 #6705 Zbl 0285.14006

[Euler 1748] L. Euler, *Introductio in analysin infinitorum, I*, 1748. Translated as *Introduction to analysis of the infinite, Book I*, Springer, New York, 1988. MR 89g:01067 Zbl 0657.01013

[Freitag 1990] E. Freitag, *Hilbert modular forms*, Springer, Berlin, 1990. MR 91c:11025 Zbl 0702.11029

[Gangl et al. 2006] H. Gangl, M. Kaneko, and D. Zagier, "Double zeta values and modular forms", pp. 71–106 in *Automorphic forms and zeta functions*, edited by S. Böcherer et al., World Sci. Publ., Hackensack, NJ, 2006. MR 2006m:11138 Zbl 1122.11057

[Goncharov 2001a] A. B. Goncharov, "Multiple polylogarithms and mixed Tate motives", preprint, 2001. arXiv math/0103059

[Goncharov 2001b] A. B. Goncharov, "Multiple ζ-values, Galois groups, and geometry of modular varieties", pp. 361–392 in *European Congress of Mathematics* (Barcelona, 2000), vol. I, edited by C. Casacuberta et al., Progr. Math. **201**, Birkhäuser, Basel, 2001. MR 2003g:11073 Zbl 1042.11042

[Goncharov 2001c] A. B. Goncharov, "The dihedral Lie algebras and Galois symmetries of $\pi_1^{(l)}(\mathbb{P}^1 - (\{0, \infty\} \cup \mu_N))$", *Duke Math. J.* **110**:3 (2001), 397–487. MR 2002m:14020 Zbl 1113.14020

[Gunnells 2000a] P. E. Gunnells, "Finiteness of minimal modular symbols for SL_n", *J. Number Theory* **82**:1 (2000), 134–139. MR 2001e:11051 Zbl 0987.11040

[Gunnells 2000b] P. E. Gunnells, "Symplectic modular symbols", *Duke Math. J.* **102**:2 (2000), 329–350. MR 2001i:11062 Zbl 0988.11023

[Gunnells and Yasaki 2008] P. E. Gunnells and D. Yasaki, "Hecke operators and Hilbert modular forms", pp. 387–401 in *Algorithmic number theory*, edited by A. J. van der Poorten and A. Stein, Lecture Notes in Comput. Sci. **5011**, Springer, Berlin, 2008. MR 2010a:11081 Zbl 1205.11056

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Hirzebruch 1973] F. E. P. Hirzebruch, "Hilbert modular surfaces", *Enseign. Math.* (2) **19** (1973), 183–281. MR 52 #10750 Zbl 0285.14007

[Hirzebruch and Zagier 1976] F. Hirzebruch and D. Zagier, "Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus", *Invent. Math.* **36** (1976), 57–113. MR 56 #11909 Zbl 0332.14009

[Horozov 2006] I. Horozov, "Non-commutative two dimensional modular symbol", preprint, 2006. arXiv math/0611955

[Horozov 2011] I. Horozov, "Non-abelian reciprocity laws on a Riemann surface", *Int. Math. Res. Not.* **2011**:11 (2011), 2469–2495. MR 2012d:19006 Zbl 1218.30127

[Horozov 2014a] I. Horozov, "Cohomology of $GL(4, \mathbb{Z})$ with nontrivial coefficients", *Math. Res. Lett.* **21**:5 (2014), 1111–1136.

[Horozov 2014b] I. Horozov, "Multiple Dedekind zeta functions", *J. Reine Angew. Math.* (online publication July 2014).

[Horozov 2014c] I. Horozov, "Reciprocity laws on algebraic surfaces via iterated integrals", *J. K-Theory* **14**:2 (2014), 273–312. MR 3264264 Zbl 06369821

[Kazhdan et al. 2000] D. Kazhdan, B. Mazur, and C.-G. Schmidt, "Relative modular symbols and Rankin–Selberg convolutions", *J. Reine Angew. Math.* **519** (2000), 97–141. MR 2001j:11026 Zbl 0941.11020

[Kontsevich and Zagier 2001] M. Kontsevich and D. Zagier, "Periods", pp. 771–808 in *Mathematics unlimited—2001 and beyond*, edited by B. Engquist and W. Schmid, Springer, Berlin, 2001. MR 2002i:11002 Zbl 1039.11002

[Lurie 2009] J. Lurie, *Higher topos theory*, Annals of Mathematics Studies **170**, Princeton University Press, NJ, 2009. MR 2010j:18001 Zbl 1175.18001

[Manin 1972] J. I. Manin, "Parabolic points and zeta functions of modular curves", *Izv. Akad. Nauk SSSR Ser. Mat.* **36** (1972), 19–66. In Russian; translated in *Math. USSR Izv.* **6**:1 (1972), 19–64. MR 47 #3396 Zbl 0243.14008

[Manin 2006] Y. I. Manin, "Iterated integrals of modular forms and noncommutative modular symbols", pp. 565–597 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progr. Math. **253**, Birkhäuser, Boston, 2006. MR 2008a:11062 Zbl 1184.11019

[Manin 2009] Y. I. Manin, "Lectures on modular symbols", pp. 137–152 in *Arithmetic geometry*, edited by H. Darmon et al., Clay Math. Proc. **8**, Amer. Math. Soc., Providence, RI, 2009. MR 2010i:11068 Zbl 1222.11066

[Mazur 1973] B. Mazur, "Courbes elliptiques et symboles modulaires", exposeé no. 414, 277–294 in *Séminaire Bourbaki,* 1971/1972, Lecture Notes in Math. **317**, Springer, Berlin, 1973. MR 55 #2930 Zbl 0276.14012

[Mazur and Swinnerton-Dyer 1974] B. Mazur and P. Swinnerton-Dyer, "Arithmetic of Weil curves", *Invent. Math.* **25** (1974), 1–61. MR 50 #7152 Zbl 0281.14016

[Parshin 1966] A. N. Parshin, "On a certain generalization of Jacobian manifold", *Izv. Akad. Nauk SSSR Ser. Mat.* **30** (1966), 175–182. In Russian. MR 33 #4956 Zbl 0163.15303

[Shimura 1978] G. Shimura, "The special values of the zeta functions associated with Hilbert modular forms", *Duke Math. J.* **45**:3 (1978), 637–679. MR 80a:10043 Zbl 0394.10015

[Shintani 1976] T. Shintani, "On evaluation of zeta functions of totally real algebraic number fields at non-positive integers", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **23**:2 (1976), 393–417. MR 55 #266 Zbl 0349.12007

[Shokurov 1980] V. V. Shokurov, "Shimura integrals of cusp forms", *Izv. Akad. Nauk SSSR Ser. Mat.* **44**:3 (1980), 670–718. In Russian; translated in *Math. USSR Izv.* **16**:3 (1981), 603–646. MR 82b:10029 Zbl 0444.14030

[Stein 2007] W. Stein, *Modular forms, a computational approach*, Graduate Studies in Mathematics **79**, Amer. Math. Soc., Providence, RI, 2007. MR 2008d:11037 Zbl 1110.11015

[Street 1982] R. Street, "Two-dimensional sheaf theory", *J. Pure Appl. Algebra* **23**:3 (1982), 251–270. MR 83d:18014 Zbl 0478.18004

[Zagier 1976] D. Zagier, "On the values at negative integers of the zeta-function of a real quadratic field", *Enseignement Math.* (2) **22**:1-2 (1976), 55–95. MR 53 #10742 Zbl 0334.12021

Communicated by Yuri Manin
Received 2013-08-22 Revised 2014-09-17 Accepted 2014-11-26

horozov@math.wustl.edu *Department of Mathematics, Washington University in St. Louis, One Brookings Drive, Campus Box 1146, Saint Louis, MO 63130, United States*

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.371

p-adic Hodge theory in rigid analytic families

Rebecca Bellovin

We study the functors $\mathbf{D}_{\mathbf{B}_*}(V)$, where $\mathbf{B}_*$ is one of Fontaine's period rings and V is a family of Galois representations with coefficients in an affinoid algebra A. We first relate them to (φ,Γ)-modules, showing that $\mathbf{D}_{\mathrm{HT}}(V)=\bigoplus_{i\in\mathbf{Z}}(\mathbf{D}_{\mathrm{Sen}}(V)\cdot t^i)^{\Gamma_K}$, $\mathbf{D}_{\mathrm{dR}}(V)=\mathbf{D}_{\mathrm{dif}}(V)^{\Gamma_K}$, and $\mathbf{D}_{\mathrm{cris}}(V)=\mathbf{D}_{\mathrm{rig}}(V)[1/t]^{\Gamma_K}$; this generalizes results of Sen, Fontaine, and Berger. We then deduce that the modules $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ are coherent sheaves on $\mathrm{Sp}(A)$, and $\mathrm{Sp}(A)$ is stratified by the ranks of submodules $\mathbf{D}_{\mathrm{HT}}^{[a,b]}(V)$ and $\mathbf{D}_{\mathrm{dR}}^{[a,b]}(V)$ of "periods with Hodge–Tate weights in the interval $[a,b]$". Finally, we construct functorial $\mathbf{B}_*$-admissible loci in $\mathrm{Sp}(A)$, generalizing a result of Berger and Colmez to the case where A is not necessarily reduced.

1. Introduction

1.1. ***Background.*** In this article, we study rigid analytic families of representations of Gal_K, where K is a finite extension of $\mathbf{Q}_p$ and $\mathrm{Gal}_K := \mathrm{Gal}(\overline{K}/K)$ is its absolute Galois group. More precisely, we consider vector bundles $\mathscr{V}$ over a rigid analytic space X over $\mathbf{Q}_p$ equipped with a continuous $\mathscr{O}_X$-linear action of Gal_K. Thus, if we specialize $\mathscr{V}$ at any closed point of X, we get a representation of Gal_K on a finite-dimensional $\mathbf{Q}_p$-vector space. Families of Galois representations arise, for example, on the generic fibers of Galois deformation rings, as in [Kisin 2008]. Such families of Galois representations also arise from families of p-adic modular forms.

MSC2010: primary 11S20; secondary 14G22.
Keywords: p-adic Hodge theory, rigid analytic geometry.

The study of p-adic representations of p-adic Galois groups is quite technical, so we put off precise definitions to the body of this paper and give an overview here. Given a finite-dimensional $\mathbf{Q}_p$-vector space V equipped with a continuous $\mathbf{Q}_p$-linear action of Gal_K, one can capture the information of V in terms of a semilinear Frobenius φ and a semilinear action of a 1-dimensional p-adic Lie group, at the expense of making the coefficients more complicated. More precisely, work of Fontaine and many others defines equivalences of categories between the category $\mathrm{Rep}_E(\mathrm{Gal}_K)$ of finite-dimensional E-linear representations of Gal_K, where E is some finite-dimensional $\mathbf{Q}_p$-algebra, and various kinds of étale (φ,Γ)-modules (see, e.g., [Wintenberger 1983; Cherbonnier and Colmez 1998]).

The same theory lets us sort p-adic Galois representations based on how "nice" or arithmetically significant they are. One accomplishes this by defining certain "period rings" $\mathbf{B}_*$, such as $\mathbf{B}_{\mathrm{HT}}$, $\mathbf{B}_{\mathrm{dR}}$, $\mathbf{B}_{\mathrm{st}}$, and $\mathbf{B}_{\mathrm{cris}}$, which are equipped with Galois actions and "linear algebra structures", and defining $\mathbf{D}_{\mathbf{B}_*}(V) := (\mathbf{B}_* \otimes_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K}$. We say that V is $\mathbf{B}_*$-admissible (or, for the specific examples of $\mathbf{B}_*$ listed above, "Hodge–Tate", "de Rham", "semistable", or "crystalline") if the $\mathbf{Q}_p$-dimension of V is the same as the $\mathbf{B}_*^{\mathrm{Gal}_K}$-dimension of $\mathbf{D}_{\mathbf{B}_*}(V)$ (as part of the definition of a period ring, $\mathbf{B}_*^{\mathrm{Gal}_K}$ is required to be a field).

Berger and Colmez [2008] associate to a rank-d Galois representation V with coefficients in a Banach algebra A a family of (φ,Γ)-modules $\mathbf{D}^\dagger(V)$, under the supplementary hypothesis that V admits a Galois-stable integral lattice. As an application, they show that if A is an affinoid algebra, then the locus of closed points $x \in \mathrm{Sp}(A)$ where the specialization V_x is $\mathbf{B}_*$-admissible with Hodge–Tate weights in a fixed interval $[a,b]$ is a closed analytic set, and if A is reduced and V_x is $\mathbf{B}_*$-admissible for every $x \in \mathrm{Sp}(A)$, then $\mathbf{D}_{\mathbf{B}_*}(V) := ((A \mathbin{\widehat{\otimes}} \mathbf{B}_*) \otimes_A V)^{\mathrm{Gal}_K}$ is a locally free $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module of rank d.

We make a closer study of the functors $\mathbf{D}_{\mathbf{B}_*}(V)$ for $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$, where V is a finite projective A-module of rank d equipped with a continuous A-linear action of Gal_K, for some affinoid algebra A. We actually treat vector bundles over rigid analytic spaces in Sections 4 and 5, but we state our results with affinoid coefficients here.

The first theorem we prove relates $\mathbf{D}_{\mathbf{B}_*}(V)$ to families of (φ,Γ)-modules. The modules $\mathbf{D}_{\mathrm{Sen}}(V)$, $\mathbf{D}_{\mathrm{dif}}(V)$, $\mathbf{D}^\dagger_{\mathrm{rig}}(V)$, and $\mathbf{D}^\dagger_{\log}(V)$ are defined in Section 2.2.

Theorem 1.1.1. *Let A and V be as above.*

(1) $\mathbf{D}^K_{\mathrm{HT}}(V) = \bigoplus_{i\in\mathbf{Z}} (\mathbf{D}^K_{\mathrm{Sen}}(V)\cdot t^i)^{\Gamma_K}$ *as submodules of* $(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{HT}}) \otimes_A V$.

(2) $\mathbf{D}^K_{\mathrm{dR}}(V) = (\mathbf{D}^K_{\mathrm{dif}}(V))^{\Gamma_K}$ *as submodules of* $(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_A V$.

(3) $\mathbf{D}^K_{\mathrm{cris}}(V) = (\mathbf{D}^\dagger_{\mathrm{rig,K}}(V)[1/t])^{\Gamma_K}$ *and* $\mathbf{D}^K_{\mathrm{st}}(V) = (\mathbf{D}^\dagger_{\log,K}(V)[1/t])^{\Gamma_K}$. *The first equality is as submodules of* $(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}) \otimes_A V$ *and the second is as submodules of* $(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^\dagger_{\log}) \otimes_A V$.

Remark 1.1.2. The first two parts of Theorem 1.1.1 are used in the proofs of Théorèmes 5.1.4 and 5.3.2 of [Berger and Colmez 2008], respectively. We are not aware of proofs of these facts in the literature when A is not $\mathbf{Q}_p$-finite, so for the convenience of the reader we provide proofs for general $\mathbf{Q}_p$-affinoid algebras.

We can then deduce that $\mathbf{D}_{\mathbf{B}_*}(V)$ is a finite $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module, and that the formation of $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ commutes with flat base change on A. In particular, $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ are coherent sheaves on the rigid analytic space $\mathrm{Sp}(A)$. Together with Theorem 1.1.1, this is used in [Diao and Liu 2014] to prove properness of the eigencurve over weight space. We further conjecture that the formation of $\mathbf{D}_{\mathrm{st}}(V)$ and $\mathbf{D}_{\mathrm{cris}}(V)$ also commutes with flat base change.

The key to our base change theorems is that we can express $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ as cohomology groups of a complex which has finite cohomology. We do not know how to do the same for $\mathbf{D}_{\mathrm{st}}(V)$ and $\mathbf{D}_{\mathrm{cris}}(V)$. However, the cohomological finiteness theorem of [Kedlaya et al. 2014] implies that if $K/\mathbf{Q}_p$ is finite then, for any $\alpha \in A^\times$, the formation of $\mathbf{D}_{\mathrm{cris}}(V)^{\varphi=\alpha}$ commutes with flat base change on A [Kedlaya et al. 2014, Theorem 4.4.3(2)]. Cohomological finiteness similarly underlies the results of [Liu 2015] on interpolation of semistable periods.

We then have a pair of theorems about the $\mathbf{B}_*$-admissible loci in $\mathrm{Sp}(A)$, generalizing the results of [Berger and Colmez 2008] to a base that is not necessarily reduced:

Theorem 1.1.3. *Let A and V be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$. Then there is a quotient $A \twoheadrightarrow A_{\mathbf{B}_*}^{[a,b]}$ such that for any $\mathbf{Q}_p$-finite algebra B, a map $A \to B$ factors through $A_{\mathbf{B}_*}^{[a,b]}$ if and only if the induced $\mathbf{Q}_p$-finite B-linear Galois representation $V_B := V \otimes_A B$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a, b]$.*

Theorem 1.1.4. *Let A and V be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$. Suppose that V_B is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a, b]$ for every homomorphism $A \to B$, where B is an $\mathbf{Q}_p$-finite algebra.*

(1) *The module $\mathbf{D}_{\mathbf{B}_*}^K(V)$ is a projective $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module of rank d.*

(2) *The natural homomorphism*

$$(A \otimes_{\mathbf{Q}_p} \mathbf{B}_*) \otimes_{A \otimes \mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}_{\mathbf{B}_*}^K(V) \to (A \otimes_{\mathbf{Q}_p} \mathbf{B}_*) \otimes_A V$$

is an isomorphism.

(3) *The formation of $\mathbf{D}_{\mathbf{B}_*}^K(V)$ commutes with arbitrary $\mathbf{Q}_p$-affinoid base change on A.*

In fact, assuming part (1), parts (2) and (3) are equivalent. We do not know whether part (1) implies (2) and (3).

For $* \in \{\mathrm{HT}, \mathrm{dR}\}$, we actually prove a more general pair of theorems. Namely, we let $\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V)$ be the module of "$\mathbf{B}_*$-admissible periods with Hodge–Tate weights in the interval $[a, b]$" (this is precisely defined in Section 5), and we show that $\mathrm{Sp}(A)$ is stratified by the rank of the fibral modules $\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V_x)$.

Theorem 1.1.5. *Let A and V be as above, let $X = \mathrm{Sp}(A)$, and let $* \in \{\mathrm{HT}, \mathrm{dR}\}$. There is a Zariski-open subspace $X_{\mathbf{B}_*,\le d'}^{[a,b]} \subset X$ and a Zariski-closed subspace $X_{\mathbf{B}_*,\ge d'}^{[a,b]} \hookrightarrow X$ such that $x : \mathrm{Sp}(B) \to X$ factors through*

$$X_{\mathbf{B}_*,d'}^{[a,b]} := X_{\mathbf{B}_*,\le d'}^{[a,b]} \cap X_{\mathbf{B}_*,\ge d'}^{[a,b]}$$

if and only if $\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V_x)$ is a free $B \otimes_{\mathbf{Q}_p} K$-module of rank d', where B is a $\mathbf{Q}_p$-finite Artin local ring and $V_B := V \otimes_A B$.

The subspaces $X_{\mathbf{B}_*,d'}^{[a,b]}$ give a stratification of X, in the sense that $X_{\mathbf{B}_*,\le d'-1}^{[a,b]} = X_{\mathbf{B}_*,\le d'}^{[a,b]} \setminus X_{\mathbf{B}_*,d'}^{[a,b]}$ and $X = X_{\mathbf{B}_*,\le d}^{[a,b]}$.

Theorem 1.1.6. *Let X and V be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}\}$. Suppose that for every $\mathbf{Q}_p$-finite artinian point $x : A \to B$, the $B \otimes_{\mathbf{Q}_p} K$-module $\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V_x)$ is free of rank d', where $0 \le d' \le d$.*

(1) *$\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V)$ is a rank-d' locally free $A \otimes_{\mathbf{Q}_p} K$-module, and the $(t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{K_n}(V))^{\Gamma_K}$ are locally free $A \otimes_{\mathbf{Q}_p} K$-modules.*

(2) *The formation of $\mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V)$ commutes with arbitrary $\mathbf{Q}_p$-affinoid base change $A \to A'$.*

If $d' = d$, then:

(3) $\mathbf{D}_{\mathbf{B}_*}(V) = \mathbf{D}_{\mathbf{B}_*}^{[a,b]}(V)$.

(4) *The natural morphism*

$$\alpha_V : (A \mathbin{\widehat{\otimes}} \mathbf{B}_*) \otimes_{A \otimes_{\mathbf{Q}_p} K} \mathbf{D}_{\mathbf{B}_*}^K(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_*) \otimes_A V$$

is an isomorphism.

To prove Theorems 1.1.5 and 1.1.6, we use Theorem 1.1.1 and Pottharst's [2013] theory of Galois cohomology with affinoid coefficients. This approach makes transparent the role of boundedness of Hodge–Tate weights in the results of [Berger and Colmez 2008]: boundedness of Hodge–Tate weights is equivalent to finiteness of a certain Galois cohomology group, and finiteness of cohomology groups is the essential ingredient in cohomology and base change results.

Remark 1.1.7. Shah [2013] has obtained similar results on the behavior of $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ when A is reduced, by studying the Galois cohomology of $\mathbf{B}_{\mathrm{dR}}^{+}$ directly.

The proofs of Theorems 1.1.3 and 1.1.4 when $* \in \{\mathrm{st}, \mathrm{cris}\}$ are quite different. In the latter case, we largely follow the strategy of [Berger and Colmez 2008]; the new ingredient that permits us to handle nonreduced coefficients is Lemma 2.1.2, which gives a generalization of the Shilov boundary points of a reduced Berkovich space. This permits us to prove that "de Rham implies uniformly potentially semistable" when the coefficient ring is nonreduced; when the coefficients are reduced, this is [Berger and Colmez 2008, Théorème 6.3.2].

Our results about the behavior of $\mathbf{D}_{\mathrm{cris}}(V)$ and $\mathbf{D}_{\mathrm{st}}(V)$ under base change are quite limited, except in the case where V is crystalline or semistable, when the formation of $\mathbf{D}_{\mathrm{cris}}(V)$ and $\mathbf{D}_{\mathrm{st}}(V)$ commutes with arbitrary base change. This is because the continuous Γ-cohomology of a (φ, Γ)-module is not finite, nor is there an apparent subquotient which does have finite Galois cohomology. In addition, individual de Rham periods are not necessarily potentially semistable, so we are unable to extend our present technique.

We remark further that all of our results are limited to the case where $K/\mathbf{Q}_p$ is a finite extension. This is primarily because overconvergence of families of Galois representations is only known in this case. However, our use of cohomology and base change in the study of $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ means we would be restricted to the case where K is discretely valued in any case. We similarly have no access to information about the behavior of $\mathbf{D}_{\mathbf{B}_*}(V)$ under general analytic field extension on the coefficients.

1.2. *Structure of this paper.*

Throughout this paper, we consider representations of Gal_K, where $K/\mathbf{Q}_p$ is finite, on vector bundles over $\mathbf{Q}_p$-rigid analytic spaces, where $E/\mathbf{Q}_p$ is finite.

In Section 2.1, we review some of rigid analytic geometry that we will need. The rigid analytic geometry is primarily standard. However, we prove that an affinoid algebra over a discretely valued field can be embedded in a finite product of Artin rings which are module-finite over a complete discretely valued field with perfect residue field. This generalizes the fact that a reduced affinoid algebra can be embedded in the product of the residue fields at the points of its Shilov boundary, and we expect it to be of independent interest. We then briefly recall the theory of families of (φ, Γ)-modules attached to families of Galois representations, and the subsequent construction of $\mathbf{D}_{\mathrm{Sen}}(V)$ and $\mathbf{D}_{\mathrm{dif}}(V)$. We give a criterion for a coherent sheaf over a quasi-Stein space to have finitely generated global sections, and we deduce that families of (φ, Γ)-modules over the Robba ring have finitely generated global sections.

In Section 3, we review Pottharst's results on Galois cohomology with affinoid coefficients. We generalize some of his results to modules M which are finite flat over $A[\![t]\!]$ and equipped with a continuous $A[\![t]\!]$-semilinear action of a profinite group G, where the action of G on $A[\![t]\!]$ is trivial on A and preserves the t-adic filtration. We then show that the inverse system $\{\mathrm{H}^0(G, M/t^k)\}_{k\geq 0}$ is eventually constant, and thus satisfies Mittag–Leffler.

In Section 4, we prove Theorem 1.1.1. This generalizes results of Sen, Fontaine, and Berger. We deduce that each $\mathbf{D}_{\mathbf{B}_*}(V)$ is A-finite, and that the formation of $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ commutes with flat base change on $\mathrm{Sp}(A)$. We conjecture that the formation of $\mathbf{D}_{\mathrm{st}}(V)$ and $\mathbf{D}_{\mathrm{cris}}(V)$ also commutes with flat base change on $\mathrm{Sp}(A)$.

In Section 5, we prove Theorems 1.1.3 and 1.1.4. We first study the behavior of $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$ under nonflat base changes. We use Pottharst's [2013] theory of cohomology and base change for Galois cohomology with affinoid coefficients to prove Theorems 1.1.5 and 1.1.6, giving us functorial Hodge–Tate and de Rham loci.

We can then use the existence of a functorial de Rham locus to construct functorial semistable and crystalline loci, following the argument of [Berger and Colmez 2008]. More precisely, we prove that "de Rham implies uniformly potentially semistable", using Lemma 2.1.2 to generalize a similar result of that paper when the base is reduced. The locus of points where the Galois representation is semistable with Hodge–Tate weights in the interval $[a, b]$ is a union of connected components of the locus "de Rham with Hodge–Tate weights in the interval $[a, b]$". We can then cut out the crystalline locus as the subspace where the monodromy operator N vanishes.

The appendix contains results on sheaves of period rings. We need to work with sheaves of various rings of p-adic Hodge theory, which requires us to be careful about the topologies on these rings. In the Appendix, we describe some of these rings and indicate how to sheafify them.

Notation and conventions. Throughout this paper, K is a finite extension of $\mathbf{Q}_p$. The rings of p-adic Hodge theory are as defined in [Berger 2002]. We let χ denote the p-adic cyclotomic character. Our Hodge–Tate weights are normalized so that χ has Hodge–Tate weight -1. All of our p-adic Hodge-theoretic rings are as in [Berger 2002]; in particular, we use $\mathbf{B}_{\max}$ instead of $\mathbf{B}_{\mathrm{cris}}$ to define the functor $\mathbf{D}_{\mathrm{cris}}$, and we let $\mathbf{B}_{\mathrm{st}}^+$ and $\mathbf{B}_{\mathrm{st}}$ denote $\mathbf{B}_{\max}^+[\log([\bar{\pi}])]$ and $\mathbf{B}_{\max}[\log([\bar{\pi}])]$, respectively.

2. Preliminaries

We will review some of the rigid geometry we will need, before recalling the theory of families of (φ, Γ)-modules attached to families of Galois representations.

2.1. ***Rigid geometry.*** We will use the language of classical rigid spaces. The standard reference for such spaces, and the rings of restricted power series which underlie them, is [Bosch et al. 1984]. The goal of the theory is to provide a robust theory of analytic spaces over nonarchimedean fields, mirroring the theory of complex analytic spaces over **C**. We will also assume that the reader is familiar with Raynaud's theory of formal models of quasicompact quasiseparated rigid analytic spaces, as treated in [Bosch and Lütkebohmert 1993].

2.1.1. *Affinoid algebras.* Let k be a field complete with respect to a nonarchimedean valuation $|\cdot|$, let R denote its valuation ring $\{x \in k \mid |x| \leq 1\}$, and let $\mathfrak{m}$ denote the maximal ideal of R, which consists of elements with absolute value strictly less than 1. Let $T_n(k)$ (or T_n, if the ground field is clear) denote the n-variable Tate algebra over k.

If k is discretely valued, the value group of the norm on a k-affinoid algebra is discrete, and it is often possible to reduce questions about affinoid algebras to questions about discretely valued fields. For example, if A is reduced, we have the following result, due to Berkovich [1990, Corollary 2.4.5; Berger and Colmez 2008, Corollary 2.1.4]:

Proposition 2.1.1. *If A is a reduced k-affinoid algebra with k discretely valued, there exist complete discretely valued fields $B_1, \dots, B_m$ such that there is a closed embedding $A \hookrightarrow \prod_{i=1}^m B_i$.*

We will need the following strengthening of this result, in which we drop the reducedness hypothesis:

Lemma 2.1.2. *Let A be a k-affinoid algebra, where k is discretely valued. Then there is a closed embedding $A \to \prod_i R_i$ into a product of finitely many artinian k-Banach algebras R_i, each module-finite over a complete discretely valued field B_i over k (as valued fields) with perfect residue field (and each Artin ring R_i is topologized as a finite-dimensional vector space over B_i).*

Remark 2.1.3. Kedlaya and Liu [2010, Lemma 6.4] claim a similar result, but we do not understand their argument.

Proof. First of all, we note that this is true if we take A to be a Tate algebra T_n, for then we may embed T_n into $Q(T_n)^\wedge$, the completion of its quotient field for the multiplicative Gauss norm. This field will not have perfect residue field (unless $n = 0$), but by [Matsumura 1989, Theorem 29.1] applied to $R\langle X_1, \dots, X_n\rangle^\wedge_{(\pi)}$, where π is a uniformizer of R, we can find a complete discretely valued extension $B/Q(T_n)^\wedge$ with perfect residue field and $e(B|Q(T_n)^\wedge) = 1$.

Now let A be a general k-affinoid algebra. Since A is noetherian, we can find a minimal primary decomposition $(0) = \mathfrak{q}_1 \cap \dots \cap \mathfrak{q}_r$. This yields a module-finite injective (hence closed) map $A \to \prod_j A/\mathfrak{q}_j$, so if we can embed each $A/\mathfrak{q}_i$, we can embed A.

Since each $\mathfrak{q}_j$ is a primary ideal, every zero-divisor in $A/\mathfrak{q}_j$ is nilpotent. Thus, we may replace A with $A/\mathfrak{q}_j$ and assume that all zero-divisors are nilpotent.

By Noether normalization, we can find an integral (and finite) monomorphism $T_n \to A$ for some n. Since $A \otimes_{T_n} Q(T_n)^\wedge$ is a module-finite algebra over $Q(T_n)^\wedge$, it decomposes as the product of finitely many Artin local rings finite over $Q(T_n)^\wedge$.

We claim that the natural map $A \to A \otimes_{T_n} Q(T_n)^\wedge$ is an injection. Since the natural map

$$A \otimes_{T_n} Q(T_n) \to A \otimes_{T_n} Q(T_n)^\wedge$$

is injective, it is enough to show that $A \to A \otimes_{T_n} Q(T_n)$ is injective. If a nonzero $r \in A$ dies in $A \otimes_{T_n} Q(T_n)$, there is some nonzero $\omega \in T_n$ such that $r\omega = 0$ in A. Then since all zero-divisors are nilpotent, ω lands in the nilradical of A, so some power of ω is zero in A. But $T_n \to A$ is injective and T_n is reduced, so we have a contradiction.

Moreover, we claim that the natural Banach topology on A agrees with its subspace topology from the finite-dimensional $Q(T_n)^\wedge$-vector space $A \otimes Q(T_n)^\wedge$. To see this, we first note that the natural topology on A as an affinoid algebra is the same as the topology on A as a T_n-module. When $A = T_n[x]/(f(x))$, for f a monic polynomial over T_n, A is free over T_n, and so clearly $A \to A \otimes_{T_n} Q(T_n)^\wedge$ is a closed embedding.

Now consider the general case. For any $r \in A$, when viewing r in $A \otimes_{T_n} Q(T_n)$, its minimal polynomial over $Q(T_n)$ is a monic polynomial $f(x)$ with coefficients in T_n, because T_n is integrally closed in $Q(T_n)$. Moreover, $T_n[r] \cong T_n[x]/\operatorname{ann}_{T_n[x]}(r)$, where $\operatorname{ann}_{T_n[x]}(r)$ is the annihilator of r. Because A is torsion-free as a T_n-module (as we saw above),

$$\operatorname{ann}_{T_n[x]}(r) = T_n[x] \cap f \cdot Q(T_n)[x] = f \cdot T_n[x],$$

since f is monic. Thus, $T_n[r] = T_n[x]/(f(x))$. Therefore, the subring $T_n[r] \subset A$ is finite free as a T_n-module, so it is a closed T_n-submodule of

$$(Q(T_n)^\wedge)[r] \subset A \otimes_{T_n} Q(T_n)^\wedge.$$

We will show that $A \to A \otimes_{T_n} Q(T_n)^\wedge$ is a closed embedding by considering the collection of T_n-submodules $A' \subset A$ which are closed in $A \otimes_{T_n} Q(T_n)^\wedge$. If $A' \subset A$ is a T_n-submodule, then $A' \otimes_{T_n} Q(T_n)^\wedge \to A \otimes_{T_n} Q(T_n)^\wedge$ is an injection, because $T_n \to Q(T_n)^\wedge$ is flat. We begin with $A' = T_n$. If $A' = A$, we are done. If not, choose $s \in A - A'$, so $T_n[s] \subset A$ is a finite free T_n-submodule. We claim that $A' + T_n[s]$ is closed in $A \otimes_{T_n} Q(T_n)^\wedge$. Then we can replace A' with $A' + T_n[s]$ and repeat the process. Since A is finite over T_n, this process terminates eventually at A, so we will be done.

Thus, it suffices to show that if we have a finite-dimensional $Q(T_n)^\wedge$-vector space V (equipped with its natural topology) and two closed finite T_n-submodules F and F', with F' free and $(F+F')\otimes_{T_n} Q(T_n)^\wedge \to V$ an injection, then $F+F'$ is also closed in V. We may assume by induction on the rank of F' that F' is free of rank one. We may also replace V with $(F+F')\otimes_{T_n} Q(T_n)^\wedge$, since all subspaces of a finite-dimensional vector space over a nonarchimedean field (such as $Q(T_n)^\wedge$) are closed. If $F\cap F'=\{0\}$, it is clear that $F+F' \xleftarrow{\sim} F\oplus F'$ is closed, because $V \xleftarrow{\sim} (F\otimes_{T_n} Q(T_n)^\wedge)\oplus(F'\otimes_{T_n} Q(T_n)^\wedge)$ as $Q(T_n)^\wedge$-vector spaces, and the topology on the right side is the product topology (and F is closed in $F\otimes_{T_n} Q(T_n)^\wedge$ due to the closedness of F in V). If the intersection is nonzero, there is some $\omega\in T_n-\{0\}$ such that $F'\subset(1/\omega)F$, so $F+F'\subset(1/\omega)F$. Since $(1/\omega)F$ is closed in V, it only remains to show that $F+F'$ is closed in $(1/\omega)F$. But submodules of finite modules over affinoid algebras (equipped with their natural topologies) are always closed, by [Bosch et al. 1984, Prop. 3.7.3/1], so we are done.

The upshot of this is that we have a closed embedding $A\to\prod_i R_i$, where the R_i are a finite collection of $Q(T_{n_i})^\wedge$-finite Artin rings (equipped with the topologies of finite-dimensional $Q(T_{n_i})^\wedge$-vector spaces). Finally, we replace R_i with $R_i\otimes_{Q(T_n)^\wedge} B_i$. □

2.1.2. *Quasi-Stein spaces.*

Definition 2.1.4. A rigid analytic space Y over k is said to be *quasi-Stein* if it admits an admissible covering by a rising union of affinoid subdomains $Y_0\subset Y_1\subset\cdots$ such that the transition maps $\Gamma(Y_{n+1},\mathscr{O}_{Y_{n+1}})\to\Gamma(Y_n,\mathscr{O}_{Y_n})$ are flat with dense image.

In particular, $A_\infty:=\Gamma(Y,\mathscr{O}_Y)=\varprojlim_n\Gamma(Y_n,\mathscr{O}_{Y_n})$ is a Fréchet–Stein algebra.

By Kiehl's results on coherent sheaves on rigid analytic spaces, a coherent sheaf $\mathscr{F}$ on Y is simply a compatible system of coherent sheaves $\{\mathscr{F}_n\}$ on $\{Y_n\}$. Then $F_\infty:=\Gamma(Y,\mathscr{F})=\varprojlim_n\Gamma(Y_n,\mathscr{F}_n)$ is a coadmissible module over A_∞, in the sense of [Schneider and Teitelbaum 2003].

Example 2.1.5. Fix $s>0$, and let Y be the coordinate on the closed unit disk. Then the half-open annulus $0<v_p(X)\le 1/s$ is a quasi-Stein space, as it is the rising union of the closed annuli $1/s'\le v_p(X)\le 1/s$ as $s'\to\infty$.

Quasi-Stein spaces behave much as affine schemes do in algebraic geometry. In particular, Kiehl [1967, Satz 2.4] proved the following theorem on the cohomology of coherent sheaves on quasi-Stein spaces (which also follows from [Schneider and Teitelbaum 2003, Theorem 3]):

Theorem 2.1.6. *Let Y be a quasi-Stein space, and let $\mathscr{F}$ be a coherent sheaf on Y.*

(1) $\mathrm{H}^i(Y,\mathscr{F})=0$ *for* $i>0$.

(2) *The image of $\mathscr{F}(Y)$ in $\mathscr{F}(Y_n)$ is dense for all n.*

There is no a priori reason for F_∞ to be a finite module over A_∞. For example, let $Y_n = \mathrm{Sp}\big(\prod_{i=0}^n \mathbf{Q}_p(\zeta_{p^i})\big)$, and let $\mathscr{F}_n$ be the sheaf on Y_n associated to $\prod_{i=0}^n \mathbf{Q}_p(\zeta_{p^i})^{\oplus i}$. Then F_∞ is not A_∞-finite, because the fiber of F_∞ at $\mathrm{Sp}(\mathbf{Q}_p(\zeta_{p^n}))$ is a $\mathbf{Q}_p(\zeta_{p^n})$-vector space of dimension n. Happily, this is the only thing that can go wrong:

Lemma 2.1.7. *Let $\mathscr{F}$ be a coherent sheaf over a finite-dimensional quasi-Stein space Y over $\mathbf{Q}_p$. Then $\mathrm{H}^0(Y, \mathscr{F})$ is finitely generated as an $\mathrm{H}^0(Y, \mathscr{O}_Y)$-module if and only if there is some integer d such that $\dim_{\kappa(y)} \mathscr{F}(y) \leq d$ for all $y \in Y$.*

Proof. Necessity is clear. To prove sufficiency, we proceed by induction on the dimension of Y. If Y is a zero-dimensional Stein space, the result is clear.

Now suppose we have the desired result when $\dim Y < n$, which is to say when every irreducible component of Y has dimension at most $n-1$ for some $n \geq 1$, and suppose $\dim Y = n$. Choose $i : Y' \hookrightarrow Y$ consisting of a point on every irreducible component of Y. By the settled zero-dimensional case, $i^*\mathscr{F}$ is finitely generated over $\mathscr{O}_{Y'}$, so by Theorem 2.1.6 there is some coherent $\mathscr{O}_Y$-submodule $\mathscr{F}' \subset \mathscr{F}$ on Y such that $i^*\mathscr{F}' \twoheadrightarrow i^*\mathscr{F}$. Thus, we have an exact sequence of coherent $\mathscr{O}_Y$-modules

$$0 \longrightarrow \mathscr{F}' \longrightarrow \mathscr{F} \longrightarrow \mathscr{G} \longrightarrow 0$$

such that $i^*\mathscr{G} = 0$. Then $\mathscr{G}$ vanishes on a Zariski-open subspace of Y containing Y'; its complement is a quasi-Stein space Z with structure sheaf $\mathscr{O}_Y / \mathrm{ann}_{\mathscr{O}_X} \mathscr{G}$, all of whose irreducible components have dimension at most $n-1$, since Y' intersects each irreducible component of Y. Then we may apply our inductive hypothesis to $\mathscr{G}$ ($\mathscr{F}(y) \twoheadrightarrow \mathscr{G}(y)$, so the fibral ranks of $\mathscr{G}$ are bounded). Therefore, $\mathscr{G}(Y)$ is $\mathscr{O}_Y(Y)$-finite. Since $\mathscr{F}'$ is $\mathscr{O}_Y$-finite by construction, $\mathscr{F}$ is as well, since $\mathrm{H}^1(Y, \mathscr{F}') = 0$. $\square$

Corollary 2.1.8. *Suppose that $\mathscr{F}$ is a flat coherent sheaf over $\mathscr{O}_Y$, where Y is a finite-dimensional quasi-Stein space. Then $\mathrm{H}^0(Y, \mathscr{F})$ is projective of rank d over $\mathrm{H}^0(Y, \mathscr{O}_Y)$ if and only if $\dim_{\kappa(y)} \mathscr{F}(y) = d$ for all $y \in Y$.*

Proof. As flat finitely presented modules are finite locally free, and hence projective, it is enough to prove that $\mathscr{F}$ is finitely presented over $\mathscr{O}_Y$. Lemma 2.1.7 implies that there is a surjection $\mathscr{O}_Y^{\oplus m} \twoheadrightarrow \mathscr{F}$, and we will apply Lemma 2.1.7 to the kernel $\mathscr{G}$. To do this, we need to know that $\dim_{\kappa(y)} \mathscr{G}(y)$ is bounded over all $y \in Y$. But if we specialize the short exact sequence

$$0 \longrightarrow \mathscr{G} \longrightarrow \mathscr{O}_X^{\oplus m} \longrightarrow \mathscr{F} \longrightarrow 0$$

at $y \in Y$, we get a short exact sequence

$$0 \longrightarrow \mathscr{G}(y) \longrightarrow \kappa(y)^{\oplus m} \longrightarrow \mathscr{F}(y) \longrightarrow 0,$$

because $\mathscr{F}$ was assumed flat. Therefore, $\dim_{\kappa(y)} \mathscr{G}(y) \leq m$ and we are done. $\square$

2.2. *Families of* (φ, Γ)*-modules.* We briefly recall the construction of families of (φ, Γ)-modules associated to families of Galois representations. Let A be an E-affinoid algebra, and let $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}}$ denote one of the rings $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger}_{K}$, $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{K}$, $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{\mathrm{rig},\mathrm{K}}$, or $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger}_{\mathrm{rig},\mathrm{K}}$. Similarly, let $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,(s)}$ denote one of the rings $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}$ or $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s}$.

Throughout this subsection, let $s_0 = (p-1)/p$ and $s_n = p^n s_0 = p^{n-1}(p-1)$. For L a p-adic field, let $L_n = L(\zeta_{p^n})$.

Definition 2.2.1. A φ-module over $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}}$ is a finitely presented projective module $\mathbf{D}^{(s)}$ over $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}}$, together with a map $\varphi : \mathbf{D}^{(s)} \to \mathbf{D}^{(ps)}$ which is semilinear over $\varphi : \mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}} \to \mathbf{B}^{\dagger,(ps)}_{(\mathrm{rig}),\mathrm{K}}$, such that the linearization

$$\varphi' : \mathbf{B}^{\dagger,(ps)}_{(\mathrm{rig}),\mathrm{K}} \;{}_{\varphi}\!\otimes_{\mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}}} \mathbf{D}^{(s)} \to \mathbf{D}^{(ps)}$$

is an isomorphism. A (φ, Γ)-module over A is a φ-module over A together with a continuous A-linear action of Γ_K which is semilinear over the action of Γ_K on $\mathbf{B}^{\dagger,(s)}_{(\mathrm{rig}),\mathrm{K}}$ and commutes with φ.

Remark 2.2.2. A φ-module $\mathbf{D}$ over $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{(\mathrm{rig}),\mathrm{K}}$ is in particular a finite $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{(\mathrm{rig}),\mathrm{K}}$-module. It is therefore a finite module over either a Banach algebra or a Fréchet–Stein algebra. It follows that $\mathbf{D}$ has a unique structure as a Fréchet $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{(\mathrm{rig}),\mathrm{K}}$-module. Thus, we may speak unambiguously of the continuity of any action of Γ_K.

Remark 2.2.3. Kedlaya and Lie [2010] define a family of (φ, Γ)-modules over $A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{\mathrm{rig},\mathrm{K}}$, for $s \gg 0$, to be a coherent locally free sheaf over the product of the half-open annulus $0 < v_p(X) \leq 1/s$ with $\mathrm{Sp}(A)$ in the category of rigid analytic spaces. By Lemma 2.1.7 and Corollary 2.1.8, this is equivalent to the definition we have given. This equivalence is also proven in [Kedlaya et al. 2014, Proposition 2.2.7], where the authors use the φ-module structure on a family of (φ, Γ)-modules to prove finite generation of its global sections.

The main source of (φ, Γ)-modules is Galois representations; to any family of p-adic Galois representations, we can functorially associate a family of (φ, Γ)-modules, and this functor is fully faithful.

Definition 2.2.4. Let X be a rigid analytic space over E. A family of Galois representations over X is a locally free $\mathscr{O}_X$-module $\mathscr{V}$ of rank d together with an $\mathscr{O}_X$-linear action of Gal_K which acts continuously on $\Gamma(U, \mathscr{V})$ for every admissible affinoid open $U \subset X$.

Remark 2.2.5. It is enough to check continuity on a single admissible affinoid cover $\{U_i\}$ of X. For if $U_i = \mathrm{Sp}(A_i)$ is affinoid and Gal_K acts continuously on $\mathscr{V}(U_i)$, then Gal_K certainly acts continuously on $\mathscr{V}(W) = \mathscr{V}(U_i) \otimes_{A_i} \mathscr{O}_X(W)$ for any affinoid subdomain $W \subset U_i$.

On the other hand, suppose that $\{U_i = \mathrm{Sp}(A_i)\}$ is an admissible affinoid covering of $U = \mathrm{Sp}(A)$, and suppose that Gal_K acts continuously on $\mathscr{V}(U_i)$. Since

$$0 \longrightarrow A \longrightarrow \prod_i A_i \longrightarrow \prod_{i,j} A_i \mathbin{\widehat{\otimes}}_A A_j$$

is exact, $\mathscr{V}(U)$ inherits its topology from its embedding in $\prod_i \mathscr{V}(U_i)$ and $\mathrm{GL}(\mathscr{V}(U))$ inherits its topology from its embedding in $\prod_i \mathrm{GL}(\mathscr{V}(U_i))$. Therefore, Gal_K acts continuously on $\mathscr{V}(U)$.

Then we have the following theorem:

Theorem 2.2.6 [Berger and Colmez 2008]. *Let $\mathscr{A}$ be a formal $\mathscr{O}_E$-model for A, and let V be a free A-module of rank d equipped with a continuous A-linear action of Gal_K. Suppose that V contains a free Gal_K-stable $\mathscr{A}$-submodule V_0 of rank d. Then for $s \gg 0$, there is a φ- and Gal_K-stable $A \mathbin{\widehat{\otimes}} \mathbf{B}_K^{\dagger,s}$-submodule (compatible with change in s)*

$$\mathbf{D}_K^{\dagger,s}(V) \subset ((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s}) \otimes_{\mathbf{Q}_p} V)^{H_K}$$

which is a locally free $A \mathbin{\widehat{\otimes}} \mathbf{B}_K^{\dagger,s}$-module of constant rank d such that the natural map

$$(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s}) \otimes_{A \widehat{\otimes} \mathbf{B}_K^{\dagger,s}} \mathbf{D}_K^{\dagger,s}(V) \to (A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s}) \otimes_A V$$

is an isomorphism (by "φ-stable", we mean that $\varphi(\mathbf{D}_K^{\dagger,s}(V)) \subset \mathbf{D}_K^{\dagger,ps}(V)$). If Gal_K acts trivially on $V_0/12pV_0$, then $\mathbf{D}_K^{\dagger,s}(V)$ is $A \mathbin{\widehat{\otimes}} \mathbf{B}_K^{\dagger,s}$-free of rank d.

The formation of $\mathbf{D}_K^{\dagger,s}(V)$ is compatible with base change in $\mathscr{A}$.

The base change property is not stated in [Berger and Colmez 2008], but follows easily from the construction.

Remark 2.2.7. The construction of families of (φ, Γ)-modules given in Proposition 4.2.8 and Théorème 4.2.9 of [Berger and Colmez 2008] in fact only requires the coefficients to be a Banach algebra, not an affinoid algebra.

If V admits a Gal_K-stable locally free $\mathscr{A}$-submodule V_0 of rank d, we may construct $\mathbf{D}_K^{\dagger,s}(V)$ by working on a cover $\{\mathrm{Spf}\,\mathscr{A}_i\}$ of $\mathrm{Spf}\,\mathscr{A}$ trivializing V_0. Since we know that the formation of $\mathbf{D}_K^{\dagger,s}(V)$ is functorial in maps $\mathscr{A}_i \to \mathscr{A}_i \mathbin{\widehat{\otimes}}_{\mathscr{O}_E} \mathscr{A}_j$, we can glue the $\mathbf{D}_K^{\dagger,s}(V|_{\mathscr{A}_i[1/p]})$ to get a sheaf of $A \mathbin{\widehat{\otimes}} \mathbf{B}_K^{\dagger,s}$-modules on $\mathrm{Sp}(A)$. By [Kedlaya and Liu 2010, Proposition 3.10], there is a finite locally free $A \mathbin{\widehat{\otimes}} \mathbf{B}_K^{\dagger,s}$-module $\mathbf{D}_K^{\dagger,s}(V)$ which induces this sheaf.

By [Chenevier 2009, Lemme 3.18], for any family of Galois representations $\mathscr{V}$ over a quasicompact quasiseparated rigid analytic space X, there is a formal model $\mathscr{X}$ of X such that $\mathscr{V}$ admits a Galois-stable $\mathscr{O}_{\mathscr{X}}$-lattice. In fact, $\mathbf{D}_K^{\dagger,s}(V)$ is independent of the formal model $\mathscr{A}$:

Proposition 2.2.8. *Let A and V be as above. Then $\mathbf{D}_K^{\dagger,s}(V)$ is independent of $\mathscr{A}$.*

Proof. It suffices to check independence of the integral model for an admissible formal blowing-up $\mathscr{X}' \to \operatorname{Spf} \mathscr{A}$ with center $\mathscr{I} = (f_0, \dots f_m)$. More precisely, if V admits both a Galois-stable $\mathscr{A}$-lattice and a Galois-stable $\mathscr{A}'$-lattice, then $\operatorname{Spf} \mathscr{A}$ and $\operatorname{Spf} \mathscr{A}'$ have a common admissible blow-up $\mathscr{X}$, so it suffices to check that $\mathbf{D}_K^{\dagger,s}(V)$ yields the same result on the generic fibers of $\mathscr{X}$ and $\operatorname{Spf} \mathscr{A}$.

Temporarily let $\mathbf{D}_{K,\mathscr{X}}^{\dagger,s}(V)$ denote the construction using the integral structure $\mathscr{X}$ and $\mathbf{D}_{K,\mathscr{A}}^{\dagger,s}(V)$ denote the construction using the integral structure $\mathscr{A}$. Now $\mathscr{X}$ admits a covering by the formal schemes

$$\mathscr{X}_i := \operatorname{Spf} \mathscr{A}\left\langle \frac{f_0}{f_i}, \dots, \frac{f_m}{f_i} \right\rangle$$

and the morphism $\mathscr{X}_i \to \operatorname{Spf} \mathscr{A}$ is induced by $\mathscr{A} \to \mathscr{A}\langle f_0/f_i, \dots, f_m/f_i \rangle$. In other words,

$$\mathbf{D}_{K,\mathscr{X}}^{\dagger,s}(V)|_{\operatorname{Sp}(A\langle \frac{f_0}{f_i},\dots,\frac{f_m}{f_i}\rangle)} = A\left\langle \frac{f_0}{f_i}, \dots, \frac{f_m}{f_i} \right\rangle \widehat{\otimes}_A \mathbf{D}_{K,\mathscr{A}}^{\dagger,s}(V).$$

It follows that $\mathbf{D}_{K,\mathscr{X}}^{\dagger,s}(V) = \mathbf{D}_{K,\mathscr{A}}^{\dagger,s}(V)$. □

Corollary 2.2.9. *The formation of* $\mathbf{D}_K^{\dagger,s}(V)$ *commutes with arbitrary base change on A.*

Proof. Let $A \to A'$ be a homomorphism of E-affinoid algebras, and let $X = \operatorname{Sp}(A)$ and $X' = \operatorname{Sp}(A')$. We may choose an admissible formal $\mathscr{O}_E$-model $\mathscr{X}_1$ of X such that the family of Galois representations on X extends to a family of Galois representations V_0 over $\mathscr{X}_1$. By [Bosch and Lütkebohmert 1993, Theorem 4.1], we can find a formal model $\mathscr{X}_2$ of X' together with an admissible formal blow-up $\psi : \mathscr{X}_2 \to \mathscr{X}'$ and a morphism $\psi : \mathscr{X}_2 \to \mathscr{X}_1$ which induces f on the generic fiber. Thus, functoriality of $\mathbf{D}_K^{\dagger,s}(V)$ follows from functoriality in the integral model. □

Furthermore, it is straightforward to check the following functorial properties of the assignment $V \rightsquigarrow \mathbf{D}_K^{\dagger,s}(V)$:

Proposition 2.2.10. *Let A be an E-affinoid algebra, and let V and V' be families of* Gal_K*-representations over A as above. Then, for $s \gg 0$:*

(1) $\mathbf{D}_K^{\dagger,s}(V \oplus V') = \mathbf{D}_K^{\dagger,s}(V) \oplus \mathbf{D}_{X,K}^{\dagger,s}(V')$.

(2) $\mathbf{D}_K^{\dagger,s}(V \otimes_A V') = \mathbf{D}_K^{\dagger,s}(V) \otimes_{A\widehat{\otimes}\mathbf{B}_K^{\dagger,s}} \mathbf{D}_K^{\dagger,s}(V')$.

(3) $\mathbf{D}_K^{\dagger,s}(\operatorname{Hom}_A(V, V')) = \operatorname{Hom}_{A\widehat{\otimes}\mathbf{B}_K^{\dagger,s}}(\mathbf{D}_K^{\dagger,s}(V), \mathbf{D}_K^{\dagger,s}(V))$.

In particular, the third part implies that the assignment $V \rightsquigarrow \mathbf{D}_K^{\dagger,s}(V)$ is a fully faithful functor. We omit the details; they are written out in [Bellovin 2013, §4.3].

Combined with various refinements of [Kedlaya and Liu 2010] and [Liu 2014], we have the following corollary:

Corollary 2.2.11. *Let X be a quasicompact quasiseparated rigid analytic space over E, and let $\mathscr{V}$ be a rank-d family of* Gal_K*-representations over X. Then, for $s \gg 0$, there is a family of (φ, Γ)-modules $\mathscr{D}^{\dagger,(s)}_{K,(\mathrm{rig})}(\mathscr{V})$ which has rank d over $\mathscr{O}_X \hat{\otimes} \mathbf{B}^{\dagger,(s)}_{K,(\mathrm{rig})}$ such that the natural map*

$$(\mathscr{O}_X \hat{\otimes} \widetilde{\mathbf{B}}^{\dagger,(s)}_{(\mathrm{rig})}) \otimes_{\mathscr{O}_X \hat{\otimes} \mathbf{B}^{\dagger,(s)}_{K,(\mathrm{rig})}} \mathscr{D}^{\dagger,(s)}_{K,(\mathrm{rig})}(\mathscr{V}) \to (\mathscr{O}_X \hat{\otimes} \widetilde{\mathbf{B}}^{\dagger,(s)}_{(\mathrm{rig})}) \otimes_A \mathscr{V}$$

is an isomorphism.

The formation of $\mathscr{D}^{\dagger,(s)}_{K,(\mathrm{rig})}(\mathscr{V})$ is compatible with base change in X and the assignment $\mathscr{V} \rightsquigarrow \mathscr{D}^{\dagger,(s)}_{K,(\mathrm{rig})}(\mathscr{V})$ is a fully faithful functor compatible with direct sums, duals, and tensor products.

Remark 2.2.12. We do not know whether there is an intrinsic characterization of $\mathbf{D}^{\dagger,s}_K(V)$ as a submodule of $(A \hat{\otimes} \widetilde{\mathbf{B}}^{\dagger,s}) \otimes_A V$, or an intrinsic characterization of $\mathbf{D}^{\dagger,s}_{K,\mathrm{rig}}(V)$ as a subsheaf of $\widetilde{\mathbf{B}}^{\dagger,s}_{\mathrm{rig}} \otimes_A V$.

We pause to briefly discuss the objects we have constructed. For simplicity, we temporarily assume that $A = \mathbf{Q}_p$. Given a Galois representation V of dimension d, we have constructed a module over $\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$ of rank d, equipped with a semilinear Frobenius and a semilinear action of Γ_K. There is some s so that these structures descend to $\mathbf{B}^{\dagger,s}_{\mathrm{rig,K}}$, which is (noncanonically) the ring of analytic functions on the half-open annulus $0 < v_p(X) \leq 1/e_K s$; we think of $p^{-1/e_K s(V)}$ as the minimal inner radius of an annulus to which everything descends.

Consider the analytic function $\log(1+X) \in \mathbf{B}^{\dagger,s}_{\mathrm{rig,K}}$. It has infinitely many zeroes, at the points $X = 1 - \zeta_{p^n}$, which accumulate towards the boundary of the unit disk. For a given s, we think of $n(s)$ as the minimal n so that $X = 1 - \zeta_{p^n}$ lies in the annulus $0 < v_p(X) \leq 1/e_K s$.

Returning to our general setup, we use (φ, Γ)-modules to construct modules $\mathbf{D}_{\mathrm{Sen}}(V)$ and $\mathbf{D}_{\mathrm{dif}}(V)$, which we will use to study Hodge–Tate and de Rham representations.

Recall that there is a family of injections $i_n : \mathbf{B}^{\dagger,s}_K \to K_n[\![t]\!]$ for every $n \geq n(s)$, which extend to injections $i_n : \mathbf{B}^{\dagger,s}_{\mathrm{rig,K}} \to K_n[\![t]\!]$. It is defined as the composition

$$\mathbf{B}^{\dagger,s_n}_K \subset \widetilde{\mathbf{B}}^{\dagger,s_n} \xrightarrow{\varphi^{-n}} \widetilde{\mathbf{B}}^{\dagger,s_0} \longrightarrow \mathbf{B}^+_{\mathrm{dR}},$$

where the last map sends $\sum p^k [x_k]$ (viewed as an element of $\widetilde{\mathbf{B}}^+$) to its image in $\mathbf{B}^+_{\mathrm{dR}}$, and factors through $K_n[\![t]\!]$.

Definition 2.2.13. Let X be a quasicompact quasiseparated rigid analytic space and let $\mathscr{V}$ be a locally free $\mathscr{O}_X$-module of rank d equipped with a continuous $\mathscr{O}_X$-linear action of Gal_K. Then by the preceding discussion, there is a finite extension L/K such that $\mathscr{D}^{\dagger,s}_{\mathrm{rig,L}}(\mathscr{V})$ is X-locally free.

(1) For any $n \geq n(s)$, we put

$$\mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V}) := \mathscr{D}^{\dagger,s}_{L}(\mathscr{V}) \otimes^{i_n}_{\mathscr{B}^{\dagger,s}_{L}} (\mathscr{O}_X \otimes_{\mathbf{Q}_p} L_n).$$

Then $\mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V})$ is an X-locally free $\mathscr{O}_X \otimes L_n$-module of rank d with a linear action of Γ_{L_n}.

(2) For any $n \geq n(s)$, we put

$$\mathscr{D}^{L_n,+}_{\mathrm{dif}}(\mathscr{V}) := \mathscr{D}^{\dagger,s}_{L}(\mathscr{V}) \otimes^{i_n}_{\mathscr{B}^{\dagger,s}_{L}} (\mathscr{O}_X \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} L_n[\![t]\!]),$$

and we define $\mathscr{D}^{L_n}_{\mathrm{dif}}(\mathscr{V}) := \mathscr{D}^{L_n,+}_{\mathrm{dif}}(\mathscr{V})[1/t]$. Then $\mathscr{D}^{L_n,+}_{\mathrm{dif}}(\mathscr{V})$ is an X-locally free $\mathscr{O}_X \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} L_n[\![t]\!]$-module of rank d with a continuous semilinear action of Γ_{L_n}, where $L_n[\![t]\!]$ is equipped with its natural Fréchet topology (i.e., as the inverse limit $\varprojlim_k L_n[t]/t^k$ of finite-dimensional $\mathbf{Q}_p$-vector spaces). Here Γ_{L_n} acts trivially on L_n, but acts on t via $\gamma \cdot t = \chi(\gamma)t$.

Remark 2.2.14. Both $\mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V})$ and $\mathscr{D}^{L_n,+}_{\mathrm{dif}}(\mathscr{V})$ actually have semilinear actions of all of Gal_K, ultimately by Gal_K-stability of $\mathbf{D}^{\dagger,s_0}_{L,n}(V)$ inside $(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s_0}) \otimes_A V$. We define $\mathscr{D}^{K_n}_{\mathrm{Sen}}(\mathscr{V}) := \mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V})^{H_K}$ and $\mathscr{D}^{K_n,+}_{\mathrm{dif}}(\mathscr{V}) := \mathscr{D}^{L_n,+}_{\mathrm{dif}}(\mathscr{V})^{H_K}$.

Remark 2.2.15. If A is a general $\mathbf{Q}_p$-Banach algebra with valuation ring $\mathscr{A}$, V_0 is a free $\mathscr{A}$-module of rank d equipped with a continuous $\mathscr{A}$-linear action of Gal_K, and $V := V_0[1/p]$, then we may similarly define

$$\mathbf{D}^{L_n}_{\mathrm{Sen}}(V) := \mathbf{D}^{\dagger,s}_{L}(V) \otimes^{i_n}_{A \widehat{\otimes} \mathbf{B}^{\dagger,s}_{L}} (A \otimes_{\mathbf{Q}_p} L_n),$$
$$\mathbf{D}^{L_n}_{\mathrm{dif}}(V) := \mathbf{D}^{\dagger,s}_{L}(V) \otimes^{i_n}_{A \widehat{\otimes} \mathbf{B}^{\dagger,s}_{L}} (A \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} L_n[\![t]\!]).$$

Remark 2.2.16. It is also possible to construct $\mathbf{D}^{L_n}_{\mathrm{Sen}}(V)$ directly by means of Tate–Sen theory applied to semilinear representations of Gal_K on finite X-locally free $\mathscr{O}_X \mathbin{\widehat{\otimes}} \mathbf{C}_K$-modules. In particular, there is a constant c_3 (fixed at the outset such that $1/(p-1) < c_3 < \frac{1}{2}\,\mathrm{ord}_p(12p)$) such that $\mathbf{D}^{L_n}_{\mathrm{Sen}}(V)$ admits a c_3-*fixed basis*; i.e., there is a basis $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$ such that if U_γ is the matrix of the action of a topological generator γ of Γ_n, then every entry of $U_\gamma - \mathrm{Id}$ has p-adic valuation greater than c_3. We exploit this point of view in the proof of Theorem 4.2.5.

Proposition 2.2.17. (1) *$\mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V})$ is an X-locally free $\mathscr{O}_X \otimes L_n$-module of rank d, and we have a Galois-equivariant isomorphism*

$$\mathbf{C}_K \mathbin{\widehat{\otimes}}_{L_n} \mathscr{D}^{L_n}_{\mathrm{Sen}}(\mathscr{V}) \to \mathbf{C}_K \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} V.$$

(2) $\mathscr{D}_{\mathrm{dif}}^{L_n,+}(\mathscr{V})$ *is an* X*-locally free* $\mathscr{O}_X \mathbin{\widehat{\otimes}} L_n[\![t]\!]$*-module of rank* d*, and we have a Galois-equivariant isomorphism*

$$(\mathscr{O}_X \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}^+) \otimes_{\mathscr{O}_X \widehat{\otimes} L_n[\![t]\!]} \mathscr{D}_{\mathrm{dif}}^{L_n,+}(\mathscr{V}) \to (\mathscr{O}_X \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}^+) \otimes_{\mathscr{O}_X} \mathscr{V}$$

which respects the filtrations on each side.

Proof. For both of these, the starting point is the isomorphism

$$\widetilde{\mathscr{B}}^{\dagger,s} \otimes_{\mathscr{B}_L^{\dagger,s}} \mathscr{D}_L^{\dagger,s}(\mathscr{V}) \to \widetilde{\mathscr{B}}^{\dagger,s} \otimes_{\mathscr{O}_X} \mathscr{V}.$$

The composition

$$\mathbf{B}_L^{\dagger,s} \xrightarrow{i_n} L_n[\![t]\!] \to \mathbf{B}_{\mathrm{dR}}^+$$

is the same as the composition

$$\mathbf{B}_L^{\dagger,s} \subset \widetilde{\mathbf{B}}^{\dagger,s} \xrightarrow{\varphi^{-n}} \widetilde{\mathbf{B}}^{\dagger,p^{-n}s} \to \mathbf{B}_{\mathrm{dR}}^+$$

by definition, so extending scalars on each side from $\widetilde{\mathscr{B}}^{\dagger,s}$ to $\mathscr{B}_{X,\mathrm{dR}}^+$ or $\mathscr{O}_X \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathbf{C}_K$ gives the desired result. □

3. Cohomology of procyclic groups

3.1. ***Overview.*** Let G be a profinite group with finite p-cohomological dimension e such that $\mathrm{H}^i(G, T)$ has finite cardinality for all finite p-torsion discrete G-modules T. Let M be a topological abelian group equipped with a continuous action of G. We consider the continuous cochain complex $C^\bullet(G, M)$ and its cohomology groups $\mathrm{H}^\bullet(G, M)$. Specifically, we define the n-cochains $C^n(G, M)$ to be the set of continuous functions

$$f : G^n \to M,$$

and we define the differential $d^n : C^n(G, M) \to C^{n+1}(G, M)$ by

$$\begin{aligned} d^n(f)(g_1, \dots, g_{n+1}) := g_1 \cdot f(g_2, \dots, g_{n+1}) \\ + \sum_{i=1}^{n} (-1)^i f(g_1, \dots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \dots, g_{n+1}) \\ + (-1)^{n+1} f(g_1, \dots, g_n). \end{aligned}$$

Thus, we get a complex

$$C^\bullet(G, M) : 0 \longrightarrow M = C^0(G, M) \longrightarrow C^1(G, M) \longrightarrow \cdots,$$

and we define $\mathrm{H}^n(G, M) := \ker d_n / \operatorname{im} d_{n-1}$. If $M = \varinjlim_{i \in I} M_i$ is the filtered colimit of topological abelian groups equipped with continuous actions of G (compatible with the transition maps), we define $\mathrm{H}^n(G, M) := \varinjlim_{i \in I} \mathrm{H}^n(G, M_i)$.

Suppose now that M is actually a $\mathbf{Q}_p$-Banach space and the action of G on M is $\mathbf{Q}_p$-linear. Exact sequences of $\mathbf{Q}_p$-Banach spaces of are $\mathbf{Q}_p$-linearly split, so a G-equivariant exact sequence of $\mathbf{Q}_p$-Banach spaces

$$0 \longrightarrow M' \longrightarrow M \longrightarrow M'' \longrightarrow 0$$

admits a continuous $\mathbf{Q}_p$-linear section $M'' \to M$. Therefore, there is a long exact sequence in cohomology

$$0 \longrightarrow \mathrm{H}^0(G, M') \longrightarrow \mathrm{H}^0(G, M) \longrightarrow \mathrm{H}^0(G, M'') \longrightarrow \mathrm{H}^1(G, M') \longrightarrow \cdots .$$

If A is an E-affinoid algebra, and M is a finite flat A-module and the action of G is A-linear, Pottharst has shown that the cohomology groups $\mathrm{H}^i(G, M)$ satisfy a number of good properties. In particular, $\mathrm{H}^i(G, M)$ is a finite A-module for all $i \geq 0$, by [Pottharst 2013, Theorem 1.2], and $\mathrm{H}^i(G, M) = 0$ whenever $i > e$, by [Pottharst 2013, Proposition 1.1].

Crucially, the finiteness of the cohomology groups $\mathrm{H}^i(G, M)$ makes it possible to deduce the following "cohomology and base change" result:

Theorem 3.1.1 [Pottharst 2013, Theorem 1.4]. *Let M be a finite flat A-module. Then if A' is an A-affinoid algebra, there is a base change spectral sequence of A'-modules*

$$\mathrm{E}_2^{ij} = \mathrm{Tor}_{-i}^A(\mathrm{H}^j(G, M), A') \Rightarrow \mathrm{H}^{i+j}(G, M \otimes_A A'),$$

in which the edge map $\mathrm{E}_2^{0,j} = \mathrm{H}^j(G, M) \otimes_A A' \to \mathrm{H}^j(G, M \otimes_A A')$ is the natural map.

In particular, if A' is flat over A, then the formation of continuous group cohomology commutes with base change.

Remark 3.1.2. The base change spectral sequence follows from the natural isomorphism

$$C^\bullet(G, M) \otimes_A^{\mathbf{L}} A' \to C^\bullet(G, M \otimes_A A')$$

in the bounded derived category $\mathbf{D}_{\mathrm{coh}}^b(A')$ of finite A'-modules.

We will be primarily concerned with the Galois cohomology of groups G of p-cohomological dimension 1. In that case, the base change theorem takes a particularly nice form:

Corollary 3.1.3. *Suppose G has p-cohomological dimension* 1.

(1) *The formation of* $\mathrm{H}^1(\Gamma, M)$ *commutes with affinoid base change on A.*

(2) *The spectral sequence degenerates at the* E_3 *page.*

(3) *There is a low-degree exact sequence*

$$0 \longrightarrow \mathrm{H}^0(\Gamma, M) \otimes_A A' / \operatorname{Tor}_2^A(\mathrm{H}^1(\Gamma, M), A') \\ \longrightarrow \mathrm{H}^0(\Gamma, M \otimes_A A') \longrightarrow \operatorname{Tor}_1^A(\mathrm{H}^1(\Gamma, M), A') \longrightarrow 0.$$

3.2. *Cohomology of semilinear G-modules.* We will need to extend some of Pottharst's results. Throughout this subsection, let M be a finite flat $A[\![t]\!]$-module, equipped with its natural Fréchet topology (i.e., as the inverse limit $\varprojlim_k M/t^k$ of finite Banach A-modules), and suppose M is equipped with a continuous $A[\![t]\!]$-semilinear G-action, where G acts on $A[\![t]\!]$ so that the action on A is trivial and the action preserves the t-adic filtration.

Proposition 3.2.1. *Let M be as above.*

(1) $\mathrm{H}^i(G, M) = 0$ *for $i > e$.*

(2) *If $A \to A'$ is a quotient of affinoid algebras, the formation of* $\mathrm{H}^e(G, M)$ *commutes with base change to A', i.e., the natural map* $\mathrm{H}^e(G, M) \otimes_A A' \to \mathrm{H}^e(G, M \widehat{\otimes}_A A')$ *is an isomorphism.*

Proof. (1) For each quotient M/t^k, we have the continuous cochain complex $C^\bullet(G, M/t^k)$, and the transition maps $C^\bullet(G, M/t^{k+n}) \to C^\bullet(G, M/t^k)$ are surjective. Therefore, by [Weibel 1994, Theorem 3.5.8], for each i we have an exact sequence

$$0 \longrightarrow \varprojlim_k{}^1 \mathrm{H}^{i-1}(G, M/t^k) \longrightarrow \mathrm{H}^i(G, M) \longrightarrow \varprojlim_k \mathrm{H}^i(G, M/t^k) \longrightarrow 0.$$

Then for $i > e+1$, we have $\mathrm{H}^i(G, M/t^k) = 0$ and $\mathrm{H}^{i-1}(G, M/t^k) = 0$, by [Pottharst 2013, Theorem 1.1(4)]. Therefore, $\mathrm{H}^i(G, M) = 0$. If $i = e+1$, $\mathrm{H}^i(G, M/t^k) = 0$, by the same theorem. Then we use the long exact sequence associated to

$$0 \longrightarrow C^\bullet(G, t^n M/t^{k+n}) \longrightarrow C^\bullet(G, M/t^{k+n}) \longrightarrow C^\bullet(G, M/t^k) \longrightarrow 0$$

and the vanishing of $\mathrm{H}^i(G, t^n M/t^{k+n})$ to see that $\{\mathrm{H}^i(G, M/t^k)\}_k$ has surjective transition maps. Therefore, $\varprojlim_k^1 \mathrm{H}^{i-1}(G, M/t^k)$ and $\mathrm{H}^i(G, M)$ vanish as well.

(2) Let $A' = A/J$ be a quotient of A. Since A' is a finitely presented A-module, $- \otimes_A A'$ commutes with taking inverse limits with surjective transition maps. It follows that

$$C^\bullet(G, M) \otimes_A A' \xrightarrow{\sim} \varprojlim_k (C^\bullet(G, M/t^k) \otimes_A A').$$

But the natural map $C^\bullet(G, M/t^k) \otimes_A A' \to C^\bullet(G, (M/t^k) \otimes_A A')$ is a quasiisomorphism by [Pottharst 2013, Lemma 1.5], so we obtain a quasiisomorphism

$$C^\bullet(G, M) \otimes_A A' \to \varprojlim_k (C^\bullet(G, (M/t^k) \otimes_A A')) = C^\bullet(G, M \widehat{\otimes}_A A').$$

As A' is finitely presented as an A-module, we may find a projective resolution $D^\bullet \to A'$. Because the terms of $C^\bullet(G, M)$ are A-flat, the induced map $C^\bullet(G, M) \otimes_A D^\bullet \to C^\bullet(G, M) \otimes_A A'$ is a quasiisomorphism, and we obtain a second-quadrant spectral sequence abutting to the homology of $C^\bullet(G, M \widehat{\otimes}_A A')$. Consideration of the E_2-page yields the desired result. □

Proposition 3.2.2. *Let the notation be as above, and suppose in addition that G is procyclic with topological generator γ and that $\mathrm{H}^0(G, \mathrm{gr}^\bullet M)$ is A-finite, where $\mathrm{gr}^\bullet M$ is the associated graded module to M. Then there is some N such that $\mathrm{H}^0(G, M) \xrightarrow{\sim} \mathrm{H}^0(G, M/t^k M)$ for any $k \geq N$.*

Before we prove Proposition 3.2.2, we record several useful consequences about $\mathrm{H}^1(G, M)$:

Corollary 3.2.3. *Let the notation be as in Proposition 3.2.2.*

(1) $\mathrm{H}^1(G, M) = \varprojlim_k \mathrm{H}^1(G, M/t^k)$.

(2) *For any $k \geq N$, the sequence*

$$0 \longrightarrow \mathrm{H}^1(G, t^k M) \longrightarrow \mathrm{H}^1(G, M) \longrightarrow \mathrm{H}^1(G, M/t^k) \longrightarrow 0$$

is exact.

(3) *For any $k \in \mathbf{Z}$ and $k' \in \mathbf{N}$, the kernel of the natural map $\mathrm{H}^1(G, t^{k+k'} M) \to \mathrm{H}^1(G, t^k M)$ is a quotient of the A-finite module $\mathrm{H}^0(G, t^k M/t^{k+k'})$, the cokernel is the A-finite module $\mathrm{H}^1(G, t^k M/t^{k+k'})$, and for all but finitely many k, k', it is an injection.*

(4) $\varprojlim_k \mathrm{H}^1(G, t^k M) = 0$.

Proof. (1) Proposition 3.2.2 implies that the projective system $\{\mathrm{H}^0(G, M/t^k)\}_{k \geq 0}$ satisfies the Mittag–Leffler condition, so $\varprojlim_k^1 \mathrm{H}^0(G, M/t^k) = 0$. Together with the exact sequence

$$0 \longrightarrow \varprojlim_k{}^1 \mathrm{H}^0(G, M/t^k) \longrightarrow \mathrm{H}^1(G, M) \longrightarrow \varprojlim_k \mathrm{H}^1(G, M/t^k) \longrightarrow 0$$

this yields the desired result.

(2) For each $n \geq 0$, the exact sequence of Banach A-modules

$$0 \longrightarrow t^k M/t^{k+n} \longrightarrow M/t^{k+n} \longrightarrow M/t^k \longrightarrow 0$$

induces a long exact sequence in cohomology. If $k \geq N$, then Proposition 3.2.2 implies that $\mathrm{H}^0(G, M/t^{k+n}) \to \mathrm{H}^0(G, M/t^k)$ is a surjection. Therefore, the connecting homomorphism $\delta : \mathrm{H}^0(G, M/t^k) \to \mathrm{H}^1(G, t^k M/t^{k+n})$ is zero and

$$0 \longrightarrow \mathrm{H}^1(G, t^k M/t^{k+n}) \longrightarrow \mathrm{H}^1(G, M/t^{k+n}) \longrightarrow \mathrm{H}^1(G, M/t^k) \longrightarrow 0$$

is exact. Part (1) above (applied to $\mathrm{H}^1(G, t^k M)$) implies that $\mathrm{H}^1(G, t^k M) = \varprojlim_n \mathrm{H}^1(G, t^k M/t^{k+n})$, and, since the inverse system $\{\mathrm{H}^1(G, t^k M/t^{k+n})\}_n$ has surjective transition maps, we obtain an exact sequence

$$0 \longrightarrow \mathrm{H}^1(G, t^k M) \longrightarrow \mathrm{H}^1(G, M) \longrightarrow \mathrm{H}^1(G, M/t^k) \longrightarrow 0,$$

as desired.

(3) For every $n \geq 0$, the exact sequence

$$0 \longrightarrow t^{k+k'} M/t^{k+k'+n} \longrightarrow t^k M/t^{k+k'+n} \longrightarrow t^k M/t^{k+k'} \longrightarrow 0$$

induces an exact sequence

$$\begin{aligned} \mathrm{H}^0(G, t^k M/t^{k+k'}) &\longrightarrow \mathrm{H}^1(G, t^{k+k'} M/t^{k+k'+n}) \\ &\longrightarrow \mathrm{H}^1(G, t^k M/t^{k+k'+n}) \longrightarrow \mathrm{H}^1(G, t^k M/t^{k+k'}) \longrightarrow 0. \end{aligned}$$

If $k \gg 0$ or $k \ll 0$, and $k' \gg 0$, $\mathrm{H}^0(G, t^k M/t^{k+k'}) = 0$ and we may take the projective limit as $n \to \infty$ to obtain an injection $\mathrm{H}^1(G, t^{k+1} M) \to \mathrm{H}^1(G, t^k M)$.

Otherwise, let

$$\begin{aligned} K_{k,k',n} &:= \ker\bigl(\mathrm{H}^1(G, t^{k+k'} M/t^{k+k'+n}) \to \mathrm{H}^1(G, t^k M/t^{k+k'+n})\bigr) \\ &= \mathrm{H}^0(G, t^k M/t^{k+k'})/\mathrm{H}^0(G, t^k M/t^{k+k'+n}). \end{aligned}$$

To show that $\varprojlim_n K_{k,k',n}$ is A-finite, it suffices to show that the natural map

$$\mathrm{H}^0(G, t^k M/t^{k+k'}) \to \varprojlim_n K_{k,k',n}$$

is a surjection. But, by Proposition 3.2.2 (applied to $t^k M$), $\{\mathrm{H}^0(G, t^k M/t^{k+k'+n})\}_n$ is stationary for $n \gg 0$, so $\{\mathrm{H}^0(G, t^k M/t^{k+k'+n})/\mathrm{H}^0(G, t^{k+k'} M/t^{k+k'+n})\}_n$ satisfies the Mittag–Leffler condition. This implies that

$$\begin{aligned} 0 \longrightarrow \varprojlim_n &\bigl(\mathrm{H}^0(G, t^k M/t^{k+k'+n})/\mathrm{H}^0(G, t^{k+k'} M/t^{k+k'+n})\bigr) \\ &\longrightarrow \mathrm{H}^0(G, t^k M/t^{k+k'}) \longrightarrow \varprojlim_n K_{k,k',n} \longrightarrow 0 \end{aligned}$$

is exact.

To identify the cokernel of $\mathrm{H}^1(G, t^{k+k'} M) \to \mathrm{H}^1(G, t^k M)$, we again consider the exact sequences

$$\begin{aligned} 0 \longrightarrow K_{k,k',n} &\longrightarrow \mathrm{H}^1(G, t^{k+k'} M/t^{k+k'+n}) \\ &\longrightarrow \mathrm{H}^1(G, t^k M/t^{k+k'+n}) \longrightarrow \mathrm{H}^1(G, t^k M/t^{k+k'}) \longrightarrow 0 \end{aligned}$$

for $n \geq 0$. As n varies, the natural transition maps are surjections, so we see that the cokernel $\operatorname{coker}\bigl(\mathrm{H}^1(G, t^{k+k'} M) \to \mathrm{H}^1(G, t^k M)\bigr)$ is identified with the A-finite module $\mathrm{H}^1(G, t^k M/t^{k+k'})$.

(4) By Proposition 3.2.2, there is some N such that, for $k \geq N$ and any $n \geq 0$, the sequence

$$0 \longrightarrow \mathrm{H}^1(G, t^k M/t^{k+n}) \longrightarrow \mathrm{H}^1(G, M/t^{k+n}) \longrightarrow \mathrm{H}^1(G, M/t^k) \longrightarrow 0$$

is exact. Taking the projective limit as $n \to \infty$ and applying part (1) to M and $t^k M$, we obtain an exact sequence

$$0 \longrightarrow \mathrm{H}^1(G, t^k M) \longrightarrow \mathrm{H}^1(G, M) \longrightarrow \mathrm{H}^1(G, M/t^k).$$

Taking the projective limit as $k \to \infty$, we obtain an exact sequence

$$0 \longrightarrow \varprojlim_k \mathrm{H}^1(G, t^k M) \longrightarrow \mathrm{H}^1(G, M) \longrightarrow \varprojlim_k \mathrm{H}^1(G, M/t^k).$$

But the map $\mathrm{H}^1(G, M) \to \varprojlim_k \mathrm{H}^1(G, M/t^k)$ is an isomorphism by part (1), so $\varprojlim_k \mathrm{H}^1(G, t^k M) = 0$. □

Remark 3.2.4. If we knew that surjections of $\mathbf{Q}_p$-Fréchet spaces admit continuous sections (so that short exact sequences of G-representations would yield long exact sequences in cohomology), the proof of Corollary 3.2.3 could be simplified in a number of places. As we are not aware of a result to that effect, we are instead forced to use the Mittag–Leffler property proved in Proposition 3.2.2.

We now turn to the proof of Proposition 3.2.2; we will require a number of preliminaries. We observe at the outset that the cohomology group $\mathrm{H}^0(G, M)$ is computed by H^0 of the complex

$$C^\bullet_{\mathrm{alg}} : 0 \longrightarrow M \xrightarrow{\gamma - 1} M \longrightarrow 0.$$

For the remainder of this section, we therefore take $\mathrm{H}^\bullet(G, M)$ to be the homology of $C^\bullet_{\mathrm{alg}}$. It is purely algebraic, with no input from the topology of G or M.

Lemma 3.2.5. *Let M and G be as above.*

(1) *There is some N_0 such that $t^k M/t^{k+1} M$ has no nonzero G-invariants for any $k \geq N_0$.*

(2) *For any $k \geq N_0$, $(t^k M)^{G=1} = \{0\}$.*

(3) *For any $k \geq N_0$, the natural maps*

$$(M/t^{k+1} M)^{G=1} \to (M/t^k M)^{G=1} \quad \text{and} \quad M^{G=1} \to (M/t^k M)^{G=1}$$

are injections.

(4) *$\mathrm{H}^0(G, M)$ is finite.*

Proof. (1) This follows from the finiteness of $\mathrm{H}^0(G, \mathrm{gr}^\bullet M)$.

(2) Since $t^kM = \varprojlim_h t^kM/t^{k+h}M$ and taking G-invariants is left-exact, it is enough to show that $(t^kM/t^{k+h}M)^{G=1} = 0$ for all $h \geq 0$. But this follows from repeated applications of the exact sequence

$$0 \to (t^{k'+1}M/t^{k+h}M)^{G=1} \to (t^{k'}M/t^{k+h}M)^{G=1} \to (t^{k'}M/t^{k'+1}M)^{G=1} = 0$$

for $k \leq k' \leq k+h$.

(3) We have an exact sequence

$$0 \longrightarrow (t^kM/t^{k+1}M)^{G=1} \longrightarrow (M/t^{k+1}M)^{G=1} \longrightarrow (M/t^kM)^{G=1}.$$

By the choice of k, $(t^kM/t^{k+1}M)^{G=1} = 0$. Similarly, we have an exact sequence

$$0 \longrightarrow (t^kM)^{G=1} \longrightarrow (M)^{G=1} \longrightarrow (M/t^kM)^{G=1}.$$

By the choice of k, $(t^kM)^{G=1} = \{0\}$.

(4) We have seen that $\mathrm{H}^0(G, M)$ injects into $(M/t^kM)^{G=1}$ for sufficiently large k. But M/t^kM is a finite A-module, so $\mathrm{H}^0(G, M)$ is A-finite as well. $\square$

Lemma 3.2.6. *Let M and G be as above, and let*

$$0 \longrightarrow I \longrightarrow B \longrightarrow B' \longrightarrow 0$$

be a small extension of Artin local A-algebras, and let $\mathfrak{m}_B$ be the maximal ideal of B. Let M_B, $M_{B/\mathfrak{m}_B}$, and $M_{B'}$ denote $M \otimes_A B$, $M \otimes_A B/\mathfrak{m}_B$, and $M \otimes_A B'$, respectively. Suppose that the natural maps

$$\mathrm{H}^0(G, M_{B/\mathfrak{m}_B}/t^{k+1}) \to \mathrm{H}^0(G, M_{B/\mathfrak{m}_B}/t^k),$$
$$\mathrm{H}^0(G, M_{B'}/t^{k+1}) \to \mathrm{H}^0(G, M_{B'}/t^k)$$

are isomorphisms. Then $\mathrm{H}^0(G, M_B/t^{k+1}) \to \mathrm{H}^0(G, M_B/t^k)$ is an isomorphism as well.

Remark 3.2.7. It is crucial for our application of Lemma 3.2.6 in the proof of Proposition 3.2.2 that B and B' are not assumed to be $\mathbf{Q}_p$-finite. This is why we are working with the "algebraic" cohomology groups computed by the complex $0 \longrightarrow M \xrightarrow[\gamma-1]{} M \longrightarrow 0$.

Proof. Since $0 \longrightarrow I \longrightarrow B \longrightarrow B' \longrightarrow 0$ is a small extension, I is a principal ideal killed by $\mathfrak{m}_B$. It follows that the complex

$$0 \longrightarrow IM_B/t^{k+1} \xrightarrow{\gamma-1} IM_B/t^{k+1} \longrightarrow 0$$

(resp. $0 \longrightarrow IM_B/t^k \xrightarrow{\gamma-1} IM_B/t^k \longrightarrow 0$) is isomorphic as a complex of $B/\mathfrak{m}_B$-vector spaces to the complex

$$0 \longrightarrow (M_{B/\mathfrak{m}_B})/t^{k+1} \xrightarrow{\gamma-1} (M_{B/\mathfrak{m}_B})/t^{k+1} \longrightarrow 0$$

(resp. $0 \longrightarrow M_{B/\mathfrak{m}_B}/t^k \xrightarrow{\gamma-1} (M_{B/\mathfrak{m}_B})/t^k \longrightarrow 0$). Then the hypothesis that

$$\mathrm{H}^0(G, M_{B/\mathfrak{m}_B}/t^{k+1}) \to \mathrm{H}^0(G, M_{B/\mathfrak{m}_B}/t^k)$$

is an isomorphism implies that $\mathrm{H}^0(G, IM_B/t^{k+1}) \to \mathrm{H}^0(G, IM_B/t^k)$ is an isomorphism as well.

Since M is A-flat, we have a commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & IM_B/t^{k+1} & \longrightarrow & M_B/t^{k+1} & \longrightarrow & M_{B'}/t^{k+1} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & IM_B/t^k & \longrightarrow & M_B/t^k & \longrightarrow & M_{B'}/t^k & \longrightarrow & 0
\end{array}$$

where the rows are exact.

Taking G-invariants, we get a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & (IM_B/t^{k+1})^G & \to & (M_B/t^{k+1})^G & \to & (M_{B'}/t^{k+1})^G & \to & \mathrm{H}^1(G, IM_B/t^{k+1}) \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
0 & \to & (IM_B/t^k)^G & \to & (M_B/t^k)^G & \to & (M_{B'}/t^k)^G & \to & \mathrm{H}^1(G, IM_B/t^k)
\end{array}$$

To show that $\mathrm{H}^0(G, M_B/t^{k+1}) \to \mathrm{H}^0(G, M_B/t^k)$ is an isomorphism, it suffices by the five lemma to show that $\mathrm{H}^1(G, IM_B/t^{k+1}) \to \mathrm{H}^1(G, IM_B/t^k)$ is an isomorphism. But IM_B/t^{k+1} and IM_B/t^k are finite $B/\mathfrak{m}_B$-vector spaces, so to show this, it is enough to show that they have the same dimension as $B/\mathfrak{m}_B$-vector spaces (since the map is a priori a surjection). But

$$\begin{aligned}
\dim \mathrm{H}^1&(G, IM_B/t^{k+1}) \\
&= \dim IM_B/t^{k+1} - \dim IM_B/t^{k+1} + \dim \mathrm{H}^0(G, IM_B/t^{k+1}) \\
&= \dim \mathrm{H}^0(G, IM_B/t^{k+1}) = \dim \mathrm{H}^0(G, IM_B/t^k) \\
&= \dim IM_B/t^k - \dim IM_B/t^k + \dim \mathrm{H}^1(G, IM_B/t^k) \\
&= \dim \mathrm{H}^1(G, IM_B/t^k). \qquad \square
\end{aligned}$$

Now we are in a position to prove Proposition 3.2.2:

Proof of Proposition 3.2.2. We proceed by noetherian induction on $\mathrm{Spec}(A)$. By Lemma 3.2.5, we may first choose N_0 such that, for $k \geq N_0$,

$$M^{G=1} \hookrightarrow \cdots \hookrightarrow (M/t^{k+1})^{G=1} \hookrightarrow (M/t^k)^{G=1} \hookrightarrow \cdots \hookrightarrow (M/t^{N_0})^{G=1}.$$

It follows that, for any $k \geq N_0$, the cokernel of $(M/t^{k+1})^{G=1} \to (M/t^k)^{G=1}$ is supported on a Zariski-closed subspace of $\mathrm{Spec}(A)$, namely, the support of the cokernel of $M^{G=1} \hookrightarrow (M/t^{N_0})^{G=1}$. Let $\{\mathfrak{q}_j\}$ be the (finitely many) primes corresponding to the irreducible components of this subspace. We will find some $N_1 \gg 0$ such that for $k \geq N_1$, the natural map $(M/t^{k+1})^{G=1} \otimes_A \left(\prod_j A_{\mathfrak{q}_j}^\wedge\right) \to (M/t^k)^{G=1} \otimes_A \left(\prod_j A_{\mathfrak{q}_j}^\wedge\right)$ is an isomorphism.

Since $(M/t^{k+1})^{G=1} \otimes_A A_{\mathfrak{q}_j}^\wedge = ((M/t^{k+1}) \otimes_A A_{\mathfrak{q}_j}^\wedge)^{G=1}$ by flatness of $A \to A_{\mathfrak{q}_i}^\wedge$, it is enough to produce some $N_{1,j}$ such that

$$((M/t^{k+1}) \otimes_A A_{\mathfrak{q}_j}/\mathfrak{q}_j^m)^{G=1} \to ((M/t^k) \otimes_A A_{\mathfrak{q}_j}/\mathfrak{q}_j^m)^{G=1}$$

is an isomorphism for all $k \geq N_{1,j}$ and for all m. But this follows from Lemma 3.2.6 and the fact that any surjection of Artin local rings can be factored into a sequence of small extensions.

Let $N_1 = \max\{N_{1,j}\}$. Then the natural map $M^{G=1} \to (M/t^{N_1}M)^{G=1}$ is an injection with cokernel supported on a strictly smaller Zariski-closed subspace of $\mathrm{Spec}(A)$. If it is actually an isomorphism, we are done; otherwise, we repeat the argument with primes of A corresponding to irreducible components of the support of the cokernel of $M^{G=1} \hookrightarrow (M/t^{N_1}M)^{G=1}$.

This process terminates in finitely many steps, so we find that, for k large enough, the natural maps $(M/t^{k+1}M)^{G=1} \to (M/t^k M)^{G=1}$ are isomorphisms of A-modules. It follows that $\mathrm{H}^0(G, M) = \mathrm{H}^0(G, M/t^k M)$ for sufficiently large k. □

4. The functors $\mathbf{D}_{\mathbf{B}_*}(V)$

4.1. ***Overview.*** In this section, we discuss the functors $\mathbf{D}_{\mathrm{HT}}(V)$ and $\mathbf{D}_{\mathrm{dR}}(V)$; we relate them to (φ, Γ)-modules, and we prove they are coherent sheaves on $\mathrm{Sp}(A)$. We also relate $\mathbf{D}_{\mathrm{st}}(V)$ and $\mathbf{D}_{\mathrm{cris}}(V)$ to (φ, Γ)-modules, and conjecture that they are coherent sheaves on $\mathrm{Sp}(A)$.

Throughout this section, we let E and K be finite extensions of $\mathbf{Q}_p$, and we let X be a quasiseparated rigid analytic space over E.

Definition 4.1.1. A *family of Galois representations* over X is a locally free $\mathscr{O}_X$-module $\mathscr{V}$ of rank d together with an $\mathscr{O}_X$-linear action of Gal_K which acts continuously on $\Gamma(U, V)$ for every affinoid subdomain $U \subset X$.

Definition 4.1.2. Let $\mathbf{B}_*$ be one of the period rings $\mathbf{B}_{\mathrm{HT}}$, $\mathbf{B}_{\mathrm{dR}}$, $\mathbf{B}_{\max}$, or $\mathbf{B}_{\mathrm{st}}$. Then for any family of Galois representations, we define the presheaf

$$\mathscr{D}_*^K(\mathscr{V})(U) := (\mathscr{B}_{X,*}(U) \otimes_{\mathscr{O}_X(U)} \mathscr{V}(U))^{\mathrm{Gal}_K},$$

where $\mathscr{B}_{X,*}$ is one of the sheaves of period rings defined in Section A.2. We say that $\mathscr{V}$ is $\mathbf{B}_*$-*admissible* (or simply Hodge–Tate, de Rham, semistable, or crystalline) if

$\mathscr{D}_*^K(\mathscr{V})$ is a projective $\mathscr{O}_X \otimes \mathbf{B}_*^{\mathrm{Gal}_K}$-module of rank d, and the natural morphism

$$\alpha_{\mathscr{V}} : \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X \otimes \mathbf{B}_*^{\mathrm{Gal}_K}} \mathscr{D}_*^K(\mathscr{V}) \to \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}$$

is an isomorphism.

Let $\{U_i\}_{i\in I}$ be an admissible covering of X. Then, because $\mathscr{V}$ and $\mathscr{B}_{X,*}$ are both sheaves on X, we have an exact sequence

$$\begin{aligned}0 &\longrightarrow \Gamma(X, \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}) \\ &\longrightarrow \prod_{i\in I} \Gamma(U_i, \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}) \longrightarrow \prod_{i,j\in I} \Gamma(U_i \cap U_j, \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}).\end{aligned}$$

Each of these terms has a continuous action of Gal_K by assumption, and, since the formation of Gal_K-invariants is left-exact, we have an exact sequence

$$0 \longrightarrow \mathscr{D}_*^K(\mathscr{V})(X) \longrightarrow \prod_{i\in I} \mathscr{D}_*^K(\mathscr{V})(U_i) \longrightarrow \prod_{i,j\in I} \mathscr{D}_*^K(\mathscr{V})(U_i \cap U_j).$$

It follows that $\mathscr{D}_*^K(\mathscr{V})$ is actually a sheaf of $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-modules. However, we do not know at this stage that $\mathscr{D}_*^K(\mathscr{V})(U)$ is finite, let alone that $\mathscr{D}_*^K(\mathscr{V})$ is a coherent sheaf of $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-modules.

Suppose that $X = \mathrm{Sp}(A)$, where A is a $\mathbf{Q}_p$-finite Artin ring, and V is a finite projective A-module equipped with a continuous A-linear action of Gal_K. Then $((A \otimes_{\mathbf{Q}_p} \mathbf{B}_*) \otimes_A V)^{\mathrm{Gal}_K} = (\mathbf{B}_* \otimes_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K}$ as $\mathbf{B}_*^{\mathrm{Gal}_K}$-vector spaces. In fact, V is $\mathbf{B}_*$-admissible as an A-linear representation in the sense above if and only if the underlying $\mathbf{Q}_p$-linear representation is $\mathbf{B}_*$-admissible:

Proposition 4.1.3. *Let A be a $\mathbf{Q}_p$-finite Artin local ring with maximal ideal $\mathfrak{m}$, and let V be a finite free A-module of rank d equipped with a continuous A-linear Gal_K-action. Then V is $\mathbf{B}_*$-admissible as an A-representation if and only if its underlying $\mathbf{Q}_p$-representation is $\mathbf{B}_*$-admissible.*

Proof. Let $n := \dim_{\mathbf{Q}_p} A$. It is clear that $\mathbf{B}_*$-admissibility over A implies $\mathbf{B}_*$-admissibility over $\mathbf{Q}_p$. For the converse, assume V is $\mathbf{B}_*$-admissible when viewed as a $\mathbf{Q}_p$-representation, so $((A \otimes_{\mathbf{Q}_p} \mathbf{B}_*) \otimes_A V)^{\mathrm{Gal}_K} = (\mathbf{B}_* \otimes_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K}$ is an nd-dimensional $\mathbf{B}_*^{\mathrm{Gal}_K}$-vector space and the natural map $\mathbf{B}_* \otimes_{\mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}_{\mathbf{B}_*}(V) \to \mathbf{B}_* \otimes_{\mathbf{Q}_p} V$ is an isomorphism.

We first assume that $A = E$ is a field. In that case, $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$ is a product of fields, so $\mathbf{D}_{\mathbf{B}_*}(V)$ is certainly locally free and therefore projective. In addition, the isomorphism $\alpha_V : \mathbf{B}_* \otimes_{\mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}_{\mathbf{B}_*}(V) \to \mathbf{B}_* \otimes_{\mathbf{Q}_p} V$ tells us that the natural map

$$(A \otimes_{\mathbf{Q}_p} \mathbf{B}_*) \otimes_{A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}_{\mathbf{B}_*}(V) \to (A \otimes_{\mathbf{Q}_p} \mathbf{B}) \otimes_A V$$

is an isomorphism. In particular, $\mathbf{D}_{\mathbf{B}_*}(V)$ is locally free of rank d over $A \otimes \mathbf{B}_*^{\mathrm{Gal}_K}$. Since $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$ is semilocal, $\mathbf{D}_{\mathbf{B}_*}(V)$ is free over $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$.

Now consider the general case. We will factor the extension of $\mathbf{Q}_p$-finite Artin rings $A \twoheadrightarrow A/\mathfrak{m}V$ as a sequence of small extensions and proceed by induction. So suppose we have a small extension $f : A \twoheadrightarrow A'$, so that $\mathfrak{m}\ker(f) = 0$ and $\ker(f) = (t) \cong A/\mathfrak{m}$, and suppose the result holds for A'-representations.

We have a surjection of $\mathbf{Q}_p$-representations $V \twoheadrightarrow V \otimes_A A'$, with kernel tV. By the formalism of admissible representations, $V \otimes_A A'$ is $\mathbf{B}_*$-admissible and we have a surjection $\mathbf{D}_{\mathbf{B}_*}(V) \twoheadrightarrow \mathbf{D}_{\mathbf{B}_*}(V \otimes_A A')$ with kernel $\mathbf{D}_{\mathbf{B}_*}(tV)$.

We claim that the kernel of this surjection is $t\mathbf{D}_{\mathbf{B}_*}(V)$. Clearly, $t\mathbf{D}_{\mathbf{B}_*}(V) \subset \mathbf{D}_{\mathbf{B}_*}(tV)$, since there is no Galois action on the coefficients. On the other hand, suppose that $mv \in \mathbf{D}_{\mathbf{B}_*}(tV)$ for some $m \in \ker(f)$, $v \in \mathbf{B}_* \otimes V$. Then $mv = mg(v)$ for any $g \in \mathrm{Gal}_K$, so $v = g(v)$ in $\mathbf{B}_* \otimes (V \otimes_A A/I)$ for all $g \in \mathrm{Gal}_K$, where I is the ideal of elements of A killed by m. But again by the formalism of admissible representations we have a surjection $\mathbf{D}_{\mathbf{B}_*}(V) \twoheadrightarrow \mathbf{D}_{\mathbf{B}_*}(V/IV)$, so there is some $\tilde{v} \in \mathbf{D}_{\mathbf{B}_*}(V)$ such that $\tilde{v} \cong v \mod I$. Since v and $\tilde{v}$ differ by an element of $\mathbf{D}_{\mathbf{B}_*}(IV)$ and m kills $\mathbf{D}_{\mathbf{B}_*}(IV)$, $mv = m\tilde{v} \in t\mathbf{D}_{\mathbf{B}_*}(V)$, as desired.

By the assumption on A'-representations, $\mathbf{D}_{\mathbf{B}_*}(V \otimes_A A')$ is a free module of rank d over $A' \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$. Furthermore, $\mathbf{D}_{\mathbf{B}_*}(V) \otimes_A A' = \mathbf{D}_{\mathbf{B}_*}(V \otimes_A A')$. If A_i is a local factor of the semilocal ring $A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$, then $A_i' := A_i/t$ is a local factor of $A' \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$. Therefore, Nakayama's lemma implies that $\mathbf{D}_{\mathbf{B}_*}(V) \otimes_A A_i$ is generated by d elements for all i, so we have a surjection $(A \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K})^{\otimes d} \twoheadrightarrow \mathbf{D}_{\mathbf{B}_*}(V)$. But we also have an isomorphism of $\mathbf{B}_*$-modules $\alpha_V : \mathbf{B}_* \otimes_{\mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}_{\mathbf{B}_*}(V) \to \mathbf{B}_* \otimes_{\mathbf{Q}_p} V$, so, by comparing the $\mathbf{B}_*^{\mathrm{Gal}_K}$-dimensions of $\mathbf{D}_{\mathbf{B}_*}(V)$ and $(A \otimes \mathbf{B}_*^{\mathrm{Gal}_K})^{\otimes d}$, we see that $\mathbf{D}_{\mathbf{B}_*}(V)$ is a free $A \otimes \mathbf{B}_*^{\mathrm{Gal}_K}$-module of rank d. $\square$

4.2. (φ, Γ)-*modules and* $\mathbf{D}_{\mathbf{B}_*}(V)$. When $X = \mathrm{Sp}(\mathbf{Q}_p)$, the overconvergence of Galois representations is important in part because it allows us to recover the p-adic Hodge theoretic invariants $D_{\mathbf{B}_*}(V)$ from the (φ, Γ)-module. This allows us to convert questions about Galois groups with cohomological dimension 2 into questions about profinite groups with cohomological dimension 1, at the cost of making the coefficients more complicated.

Specifically, when $X = \mathrm{Sp}(\mathbf{Q}_p)$, we have the following results:

Theorem 4.2.1 [Sen 1973]. *Let V be a finite-dimensional $\mathbf{Q}_p$-linear representation of* Gal_K. *Then* $\mathbf{D}_{\mathrm{HT}}^K(V) = \bigoplus_{i \in \mathbf{Z}} (\mathbf{D}_{\mathrm{Sen}}^K(V) \cdot t^i)^{\Gamma_K}$.

Theorem 4.2.2 [Fontaine 2004]. *Let V be a finite-dimensional $\mathbf{Q}_p$-linear representation of* Gal_K. *Then* $\mathbf{D}_{\mathrm{dR}}^K(V) = (\mathbf{D}_{\mathrm{dif}}^K(V))^{\Gamma_K}$.

Theorem 4.2.3 [Berger 2002]. *Let V be a finite-dimensional $\mathbf{Q}_p$-linear representation of* Gal_K. *Then* $\mathbf{D}_{\mathrm{st}}^K(V) = (\mathbf{D}_{\log}^{\dagger}(V)[1/t])^{\Gamma_K}$ *and* $\mathbf{D}_{\mathrm{cris}}^K(V) = (\mathbf{D}_{\mathrm{rig}}^{\dagger}(V)[1/t])^{\Gamma_K}$.

Remark 4.2.4. Sen and Fontaine used different constructions of $\mathbf{D}_{\mathrm{Sen}}^{K}(V)$ and $\mathbf{D}_{\mathrm{dif}}^{K}(V)$ than the one we have given. The equivalence of the two constructions is shown in [Berger 2002].

We will prove analogues of these results for families of Galois representations over an E-rigid analytic space X.

Theorem 4.2.5. *Let $\mathscr{V}$ be a family of representations of* Gal_K *of rank d. Then*

$$\mathscr{D}_{\mathrm{HT}}^{K}(\mathscr{V}) = \bigoplus_{i\in\mathbf{Z}} (\mathscr{D}_{\mathrm{Sen}}^{K}(\mathscr{V})\cdot t^i)^{\Gamma_K}$$

as subsheaves of $\mathscr{B}_{\mathrm{HT}}\otimes_{\mathscr{O}_X}\mathscr{V}$.

Proof. Since both $\mathscr{D}_{\mathrm{HT}}^{K}(\mathscr{V})$ and $\bigoplus_{i\in\mathbf{Z}}(\mathscr{D}_{\mathrm{Sen}}^{K}(\mathscr{V})\cdot t^i)^{\Gamma_K}$ are subsheaves of $\mathscr{B}_{\mathrm{HT}}\otimes_{\mathscr{O}_X}\mathscr{V}$, we may work locally on X. Therefore, we may assume that $X=\mathrm{Sp}(A)$ for some E-affinoid algebra A and $V := \Gamma(X,\mathscr{V})$ admits a free Gal_K-stable $\mathscr{A}$-lattice V_0 of rank d, where $\mathscr{A}$ is some formal $\mathscr{O}_E$-model for A.

Since $\mathbf{D}_{\mathrm{HT}}^{K'}(V) = K'\otimes_K \mathbf{D}_{\mathrm{HT}}^{K}(V)$ for any finite extension K'/K, we may replace K with any finite extension. Let L/K be a finite extension such that Gal_L acts trivially on $V_0/12pV_0$. Then $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ is a finite free $A\otimes_{\mathbf{Q}_p} L_n$-module of rank d, and we have a natural Galois-equivariant isomorphism

$$(A\mathbin{\widehat{\otimes}}\mathbf{C}_K)\otimes_{A\otimes_{\mathbf{Q}_p}L_n}\mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \xrightarrow{\sim} (A\mathbin{\widehat{\otimes}}\mathbf{C}_K)\otimes_A V.$$

Taking H_L-invariants, we get

$$(A\mathbin{\widehat{\otimes}}\mathbf{C}_K^{H_L})\otimes_{A\otimes L_n}\mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \xrightarrow{\sim} ((A\mathbin{\widehat{\otimes}}\mathbf{C}_K)\otimes_A V)^{H_L},$$

since H_L acts trivially on $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ by construction. We need to take Γ_{L_n}-invariants of both sides.

It suffices to show that

$$\big((A\mathbin{\widehat{\otimes}}\mathbf{C}_K)^{H_L}\otimes_{A\otimes L_n}\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)\big)^{\Gamma_n} = \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)^{\Gamma_n}.$$

To see this, we fix an $(A\otimes_{\mathbf{Q}_p}L_n)$-basis $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_d)$ of $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ which is c_3-fixed by Γ_{L_n} and choose some $\boldsymbol{x}\in((A\mathbin{\widehat{\otimes}}\mathbf{C}_K)^{H_L}\otimes_{A\otimes_{\mathbf{Q}_p}L_n}\mathbf{D}_{\mathrm{Sen}}^{L_n}(V))^{\Gamma_n}$. Then $\boldsymbol{x}=\sum x_i\boldsymbol{e}_i$ for some $x_i\in(A\mathbin{\widehat{\otimes}}\mathbf{C}_K)^{H_L}$. By the semilinearity of the Galois action, this means that $U_\gamma\cdot\gamma(\boldsymbol{x})=\boldsymbol{x}$ for any $\gamma\in\Gamma_n$, where $\boldsymbol{x}$ is the column vector of the x_i. But then we may invoke [Berger and Colmez 2008, Lemme 3.2.5] with $V_1=U_\gamma^{-1}$ and $V_2=1$ to get that $\boldsymbol{x}\in A\otimes_{\mathbf{Q}_p}L_n$.

Since $\mathbf{B}_{\mathrm{HT}}=\mathbf{C}_K[t,t^{-1}]$, it follows that $\mathbf{D}_{\mathrm{HT}}^{K}(V)=\bigoplus_{i\in\mathbf{Z}}(\mathbf{D}_{\mathrm{Sen}}^{K}(V)\cdot t^i)^{\Gamma_K}$, as desired. $\square$

Lemma 4.2.6. *Let M be a finite A-module, where A is a Banach algebra whose value group is discrete, and let $m_1, \dots, m_r$ generate M over A. Equip M with the norm $|\cdot|_M$ induced by the natural quotient $A^{\oplus r} \twoheadrightarrow M$, where $A^{\oplus r}$ has the norm $|(a_1, \dots, a_r)| = \max_i\{|a_i|\}$. Let $T : M \to M$ be an A-linear map such that $|T(m_i)| \leq C|m_i|$ for all m_i. Then the operator norm of T on M is at most C.*

Proof. Let $m \in M$. We wish to show that $|T(m)|_M \leq C|m|_M$. Because the value group of A is discrete, we can write $m = a_1 m_1 + \dots + a_r m_r$ such that $|m|_M = \max_i\{|a_i|\}$. Then

$$|T(m)|_M \leq \max_i\{|a_i|\cdot|T(m_i)|_M\} \leq C \max_i\{|a_i|\cdot|m_i|_M\} \leq C \max_i\{|a_i|\} = C|m|. \quad \square$$

Lemma 4.2.7. *Let V be a finite free A-module of rank d, equipped with a continuous A-linear action of* Gal_K. *Then the module generated by the Γ_{L_n}-orbit of $x \in \widehat{L}_\infty \widehat{\otimes}_{L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ is A-finite if and only if $x \in \bigcup_{n' \geq n} L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$.*

Proof. The Γ_{L_n}-orbit of any element of $L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ certainly generates an A-finite module. Conversely, suppose that the Γ_{L_n}-orbit of $x \in \widehat{L}_\infty \widehat{\otimes}_{L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ generates a finite A-module M. Let $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_d\}$ be an $A \otimes_{\mathbf{Q}_p} L_n$-basis of $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$, so that the action of $\gamma \in \Gamma_{L_n}$ with respect to $\boldsymbol{e}_1, \dots, \boldsymbol{e}_d$ is given by a matrix (a_{ij}) with $a_{ij} \in A \otimes_{\mathbf{Q}_p} L_n$. Write $x = \sum_i c_i \boldsymbol{e}_i$.

By assumption, M is finite over $A \otimes_{\mathbf{Q}_p} L_n$, so it is generated by a finite collection $f_1, \dots, f_r$ of elements of $(A \widehat{\otimes} \widehat{L}_\infty) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$. Then the coefficients (with respect to $\{\boldsymbol{e}_i\}$) of elements of M are contained in the $A \otimes_{\mathbf{Q}_p} L_n$-submodule of $A \widehat{\otimes} \widehat{L}_\infty$ generated by the coefficients of $f_1, \dots, f_r$, which is finite. But

$$\gamma(x) = \sum \gamma(c_i)\gamma(\boldsymbol{e}_i) = \sum_j \Big(\sum_i a_{ji} \cdot \gamma(c_i)\Big)\boldsymbol{e}_j.$$

Since γ is invertible, this shows that the Γ_{L_n}-orbit of the c_i is in the $A \otimes_{\mathbf{Q}_p} L_n$-span of the coefficients with respect to $\{\boldsymbol{e}_i\}$ of $\Gamma_{L_n} \cdot x$.

Thus, we are reduced to the rank-1 case. That is, we need to show that if the Γ_{L_n}-orbit of $c \in A \widehat{\otimes}_{\mathbf{Q}_p} \widehat{L}_\infty$ generates an $A \otimes_{\mathbf{Q}_p} L_n$-finite module $M \subset A \widehat{\otimes}_{\mathbf{Q}_p} \widehat{L}_\infty$, then $c \in \bigcup_{n' \geq n} L_{n'} \otimes_{L_n} (A \otimes_{\mathbf{Q}_p} L_n)$.

Choose a finite set $x_1, \dots, x_r \in M$ which generates M, and give M the quotient norm $|\cdot|_M$ coming from the natural surjection $A^{\oplus r} \twoheadrightarrow M$. Since $A \widehat{\otimes}_{\mathbf{Q}_p} \widehat{L}_\infty$ is a potentially orthonormalizable A-module, M is closed in $A \widehat{\otimes}_{\mathbf{Q}_p} \widehat{L}_\infty$ by [Buzzard 2007, Lemma 2.3], and therefore also acquires a p-adic norm $|\cdot|_p$. All norms on a finite Banach module are equivalent by [Bosch et al. 1984, Proposition 3.7.3/3], so $|\cdot|_M$ and $|\cdot|_p$ are equivalent, meaning that there are positive constants C_1, C_2 such that $C_1|x|_p \leq |x|_M \leq C_2|x|_p$ for all $x \in M$.

Then for any $\varepsilon > 0$ there is some m_ε such that $|(\gamma^{p^m}-1)(x_i)|_p < \varepsilon |x_i|_p$ for all i and any $m \geq m_\varepsilon$. We choose $\varepsilon = \frac{1}{2}C_1^2/(C_2^2 p^{c_3})$. This implies that

$$|(\gamma^{p^m}-1)(x_i)|_M \leq C_2|(\gamma^{p^m}-1)(x_i)|_p < C_2\varepsilon|x_i|_p \leq \frac{C_2\varepsilon}{C_1}\cdot|x_i|_M.$$

By Lemma 4.2.6, $\gamma^{p^m}-1$ has operator norm at most $C_2\varepsilon/C_1$ with respect to $|\cdot|_M$. But then

$$|(\gamma^{p^m}-1)(x)|_p \leq \frac{1}{C_1}|(\gamma^{p^m}-1)(x)|_M < \frac{C_2\varepsilon}{C_1^2}|x|_M \leq \frac{C_2^2\varepsilon}{C_1^2}|x|_p,$$

so $\gamma^{p^m}-1$ has operator norm at most $C_2^2\varepsilon/C_1^2$ with respect to $|\cdot|_p$.

Next, we observe that for any integer $m \geq 1$ the kernel of $\gamma^{p^m}-1$ on $A\widehat{\otimes}_{\mathbf{Q}_p}\widehat{L}_\infty$ is $A\otimes_{\mathbf{Q}_p}L_{m+n}$. Therefore, if $(\gamma^{p^m}-1)(M)=0$ for some $m \gg 0$, we are done. Now recall that by the third Tate–Sen axiom, for any $n' \geq n$, there is a $\Gamma_{L_{n'}}$-equivariant topological splitting

$$A\,\widehat{\otimes}_{\mathbf{Q}_p}\,\widehat{L}_\infty = (A\otimes_{\mathbf{Q}_p}L_{n'})\oplus X_{H,n'},$$

and, for $n' \gg_{m_\varepsilon} n$, $\gamma^{p^{m_\varepsilon}}-1$ acts invertibly on $X_{H,n'}$, with the norm of $(\gamma^{p^{m_\varepsilon}}-1)^{-1}$ bounded above by the constant p^{c_3}. Since $\gamma^{p^{n'-n}}-1$ kills $A\otimes_{\mathbf{Q}_p}L_{n'}$, it follows that $(\gamma^{p^{n'-n}}-1)(M)\subset X_{H,n'}$. But $(\gamma^{p^{n'-n}}-1)(M)\subset M$, so $\gamma^{p^{m_\varepsilon}}-1$ has p-adic operator norm at most $C_2^2\varepsilon/C_1^2$ on $(\gamma^{p^{n'-n}}-1)(M)$. Then for any $x\in M$,

$$\begin{aligned}\left|(\gamma^{p^{n'-n}}-1)(x)\right|_p &= \left|(\gamma^{p^{m_\varepsilon}}-1)^{-1}(\gamma^{p^{m_\varepsilon}}-1)(\gamma^{p^{n'-n}}-1)(x)\right|_p\\ &\leq p^{c_3}\left|(\gamma^{p^{m_\varepsilon}}-1)(\gamma^{p^{n'-n}}-1)(x)(x)\right|_p\\ &\leq p^{c_3}\frac{C_2^2\varepsilon}{C_1^2}\left|(\gamma^{p^{n'-n}}-1)(x)\right|_p = \frac{1}{2}\left|(\gamma^{p^{n'-n}}-1)(x)\right|_p.\end{aligned}$$

This forces $\left|(\gamma^{p^{n'-n}}-1)(x)\right|_p$ to be 0, so $(\gamma^{p^{n'-n}}-1)(x)=0$. Therefore $(\gamma^{p^{n'-n}}-1)(M)=0$, and we are done. $\square$

We can bootstrap this result to relate $\mathscr{D}_{\mathrm{dR}}(\mathscr{V})$ and $\mathscr{D}_{\mathrm{dif}}(\mathscr{V})$, just as in the case when $X=\mathrm{Sp}(\mathbf{Q}_p)$:

Theorem 4.2.8. *Let $\mathscr{V}$ be a family of representations of* Gal_K *of rank d. Then $\mathscr{D}_{\mathrm{dR}}^K(\mathscr{V}) = (\mathscr{D}_{\mathrm{dif}}^K(\mathscr{V}))^{\Gamma_K}$ as subsheaves of $\mathscr{B}_{\mathrm{dR}}\otimes_{\mathscr{O}_X}\mathscr{V}$.*

Proof. As before, we reduce to the case when $X=\mathrm{Sp}(A)$ for some E-affinoid algebra A and $V:=\Gamma(X,\mathscr{V})$ admits a free Gal_K-stable $\mathscr{A}$-lattice V_0 of rank d, where $\mathscr{A}$ is some formal $\mathscr{O}_E$-model for A.

Since $\mathbf{D}_{\mathrm{dR}}^{K'}(V) = K' \otimes_K \mathbf{D}_{\mathrm{dR}}^{K}(V)$ for any finite extension K'/K, we may again replace K with any finite extension; we choose L/K such that $\mathrm{Gal}_{L/K}$ acts trivially on $V_0/12pV_0$. Then $\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$ is a free $A \hat{\otimes} L_n[\![t]\!]$-module of rank d, and we have a Galois-equivariant isomorphism

$$(A \hat{\otimes} \mathbf{B}_{\mathrm{dR}}^{+}) \otimes_{A \hat{\otimes} L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+} \xrightarrow{\sim} (A \hat{\otimes} \mathbf{B}_{\mathrm{dR}}^{+}) \hat{\otimes}_A V$$

which respects the t-adic filtration on both sides.

After twisting V by some power of the cyclotomic character, it therefore suffices to show that

$$\big((A \hat{\otimes} \mathbf{L}_{\mathrm{dR}}^{+}) \otimes_{A \hat{\otimes} L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)\big)^{\Gamma_{L_n}} = \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)^{\Gamma_{L_n}},$$

where $\mathbf{L}_{\mathrm{dR}}^{+} := (\mathbf{B}_{\mathrm{dR}}^{+})^{H_K}$. In fact, it suffices to show that

$$\big((A \hat{\otimes} \mathbf{L}_{\mathrm{dR}}^{+}/t^m) \otimes_{A \otimes_{\mathbf{Q}_p} L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)\big)^{\Gamma_{L_n}} = (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m)^{\Gamma_{L_n}}$$

for all m, because taking inverse limits commutes with taking Γ_{L_n}-invariants.

We will do this by showing that if $x \in (A \hat{\otimes} \mathbf{L}_{\mathrm{dR}}^{+}/t^m) \otimes_{A \otimes_{\mathbf{Q}_p} L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$ and the Γ_{L_n}-orbit of x generates a finite $A \otimes_{\mathbf{Q}_p} L_n[\![t]\!]/t^m$-module, then x actually lives in $\bigcup_{n' \geq n} L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m$. For then if x is Γ_{L_n}-fixed, it lives in $L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m$ for some $n' \geq n$. Since $(L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m)^{\Gamma_{L_n}} = \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m$, we conclude that $x \in \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m$.

We proceed by induction on m. We first consider $m = 1$. Then we considering elements of $(A \hat{\otimes} \hat{L}_\infty) \otimes \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ whose Γ_{L_n}-orbits generate finite $A \otimes_{\mathbf{Q}_p} L_n$-modules. But such elements actually live in $\bigcup_{n' \geq n} L_{n'} \otimes_{L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$, by Lemma 4.2.7.

Now we assume the result for m, and we consider the exact sequence

$$\begin{aligned} 0 \longrightarrow t^m (\mathbf{L}_{\mathrm{dR}}^{+}/t^{m+1}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V) \longrightarrow (\mathbf{L}_{\mathrm{dR}}^{+}/t^{m+1}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V) \\ \longrightarrow (\mathbf{L}_{\mathrm{dR}}^{+}/t^{m}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V) \longrightarrow 0. \end{aligned}$$

If the Γ_{L_n}-orbit of $c \in (\mathbf{L}_{\mathrm{dR}}^{+}/t^{m+1}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$ generates a finite $A \otimes_{\mathbf{Q}_p} L_n$-module, then its image $\bar{c}$ in $(\mathbf{L}_{\mathrm{dR}}^{+}/t^{m}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$ does as well. By the inductive hypothesis, $\bar{c} \in \bigcup_{n' \geq n} L_{n'} \otimes_{L_n} (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^m)$. We may choose $\hat{c} \in \bigcup_{n' \geq n} L_{n'} \otimes_{L_n} (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^{m+1})$ lifting $\bar{c}$, so that the Γ_{L_n}-orbit of $c - \hat{c}$ still generates a finite $A \otimes_{\mathbf{Q}_p} L_n$-module. Then $c - \hat{c}$ is an element of

$$t^m (\mathbf{L}_{\mathrm{dR}}^{+}/t^{m+1}) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V),$$

which is isomorphic to $t^m \cdot (\mathbf{L}_{\mathrm{dR}}^{+}/t) \hat{\otimes}_{L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$ as a Γ_{L_n}-representation. But the $m = 1$ case applies to this latter space, so we are done. $\square$

We can similarly relate $\mathscr{D}_{\mathrm{cris}}^{K}(\mathscr{V})$ and $\mathscr{D}_{\mathrm{st}}^{K}(\mathscr{V})$ to the family of (φ, Γ)-modules $\mathscr{D}_{\mathrm{rig,K}}^{\dagger}(\mathscr{V})$, following [Berger 2002]:

Theorem 4.2.9. *Let $\mathscr{V}$ be a family of representations of* Gal_K. *Then* $\mathscr{D}^K_{\mathrm{cris}}(\mathscr{V}) = (\mathscr{D}^\dagger_{\mathrm{rig,K}}(\mathscr{V})[1/t])^{\Gamma_K}$ *and* $\mathscr{D}^K_{\mathrm{st}}(\mathscr{V}) = (\mathscr{D}^\dagger_{\log,K}(\mathscr{V})[1/t])^{\Gamma_K}$. *The first equality is as subsheaves of $\widetilde{\mathscr{B}}^\dagger_{\mathrm{rig}} \otimes_{\mathscr{O}_X} \mathscr{V}$ and the second is as subsheaves of $\widetilde{\mathscr{B}}^\dagger_{\log} \otimes_{\mathscr{O}_X} \mathscr{V}$.*

We will need a number of preparatory results. Throughout the proofs of these results, we will use freely the fact that if A is a $\mathbf{Q}_p$-Banach algebra, then A is potentially orthonormalizable in the sense of [Buzzard 2007]. This follows from [Schneider 2002, Proposition 10.1] since $\mathbf{Q}_p$ is discretely valued. This has the consequence that injections of Fréchet spaces are preserved under completed tensor products with A over $\mathbf{Q}_p$.

Lemma 4.2.10. *Let A be an orthonormalizable $\mathbf{Q}_p$-Banach algebra, and let $\mathscr{A}$ be its unit ball. Let h be a positive integer. Then*

$$\bigcap_{k=0}^{\infty} p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}) = \mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{+} \quad \textit{and} \quad \bigcap_{k=0}^{\infty} p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}_{\mathrm{rig}}) \subset A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}.$$

Proof. This is an A-linear analogue of [Berger 2002, Lemme 3.1]. We prove the first assertion here; with this in place, the proof of the second carries over verbatim from [Berger 2002]. Note that the first assertion is an equality of topological $\mathbf{Z}_p$-modules inside $A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}$, not algebras, because we do not know that there is an algebra norm on A making it into an orthonormalizable $\mathbf{Q}_p$-Banach space.

Choose an orthonormal basis $\{e_i\}_{i\in I}$ of A. Then we may compute the intersection $\bigcap_{k=0}^{\infty} p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s})$ inside $A\,\widehat{\otimes}\,\widetilde{\mathbf{B}} \xrightarrow{\sim} c_I(\mathbf{B})$. But if

$$x = \sum_{i\in I} a_i e_i \in p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s})$$

for all k, then $a_i \in p^{-hk}\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}$ for all k, implying that $x \in \mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{+}$. □

Remark 4.2.11. The completed tensor product $\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{+}$ appearing in the first assertion of Corollary 4.2.12 is with respect to the weak topology on $\widetilde{\mathbf{A}}^{+}$, not with respect to the p-adic topology.

Corollary 4.2.12. *Let A be a $\mathbf{Q}_p$-Banach algebra, equipped with an algebra norm $|\cdot|$, and let $\mathscr{A}$ be its valuation ring. Then* $\bigcap_{k=0}^{\infty} p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}_{\mathrm{rig}}) \subset A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}$.

Proof. By [Schneider 2002, Proposition 10.1], there is an equivalent norm $|\cdot|'$ on A with respect to which A is orthonormalizable; let $\mathscr{A}'$ be the unit ball with respect to $|\cdot|'$. Then there exists a constant $c \geq 0$ such that $p^c\mathscr{A} \subset \mathscr{A}'$, so that $\bigcap_{k=0}^{\infty} p^c p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}_{\mathrm{rig}}) \subset A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}$. But p is invertible in $\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}$, so $\bigcap_{k=0}^{\infty} p^{-hk}(\mathscr{A}\,\widehat{\otimes}\,\widetilde{\mathbf{A}}^{\dagger,p^{-k}s}_{\mathrm{rig}}) \subset A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}$. □

As in [Berger 2002], "Frobenius regularization" follows immediately from Corollary 4.2.12.

Proposition 4.2.13 [Berger 2002, Proposition 3.2]. *Let d_1, d_2, and h be three positive integers, and let $M \in \mathrm{Mat}_{d_2\times d_1}(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})$ be a matrix. Suppose there exists $P \in \mathrm{GL}_{d_1}(A\otimes_{\mathbf{Q}_p} F)$ such that $M = \varphi^{-h}(M)P$. Then $M \in \mathrm{Mat}_{d_2\times d_1}(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\log})$.*

Corollary 4.2.14. *Let d be a positive integer, and let $M \in \mathrm{GL}_d(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log}[t])$ be an invertible matrix. Suppose there exists $P \in \mathrm{GL}_d(A\otimes_{\mathbf{Q}_p} F)$ such that $M = \varphi(M)P$. Then $M \in \mathrm{GL}_d(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{+}_{\log}[1/t])$.*

Remark 4.2.15. Proposition 4.2.13 is stated and proved in [Berger 2002] for $h = 1$. However, the proof carries over verbatim for $h > 1$.

Proposition 4.2.16. *Let A be a discretely valued $\mathbf{Q}_p$-Banach field with perfect residue field, and let V be an A-vector space of dimension d equipped with a continuous A-linear action of Gal_K. Then the natural map*

$$((A\,\widehat{\otimes}\,\mathbf{B}^{+}_{\mathrm{st}})\otimes_A V)^{\mathrm{Gal}_K} \to ((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$$

is an isomorphism.

Proof. Recall that, for $n \gg 0$, there is an injection $i_n : \widetilde{\mathbf{B}}^{\dagger,s_n}_{\log} \to \mathbf{B}^{+}_{\mathrm{dR}}$, where $s_n = p^n s_0 = p^{n-1}(p-1)$. Then i_n yields an injection

$$((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger,s_n}_{\log})\otimes_A V)^{\mathrm{Gal}_K} \to \mathbf{D}^{K}_{\mathrm{dR}}(V).$$

Note that V admits a free Gal_K-stable $\mathscr{A}$-submodule of rank d; under these hypotheses, we will show (in a noncircular way) in Proposition 4.3.2 that $\mathbf{D}^{K}_{\mathrm{dR}}(V)$ is a finite $A\otimes_{\mathbf{Q}_p} K$-module. Therefore, $((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger,s_n}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$ is a finite $A\otimes_{\mathbf{Q}_p} K_0$-module.

Further, we claim that there is some s_n such that

$$((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger,s_n}_{\log})\otimes_A V)^{\mathrm{Gal}_K} = ((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V)^{\mathrm{Gal}_K}.$$

Indeed, $A\otimes_{\mathbf{Q}_p} K_0 \cong \prod_i A_i$, where the A_i are a finite collection of $\mathbf{Q}_p$-Banach fields which are finite extensions of A (and isomorphic to each other, because $K_0/\mathbf{Q}_p$ is Galois), so that $\mathbf{D}^{K}_{\mathrm{dR}}(V) \cong \bigoplus_i (A_i \otimes_{K_0} K)^{\oplus d_i}$ for some integers $d_i \geq 0$. It follows that $((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger,s_n}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$ is an A-vector space of dimension at most $\sum_i d_i[K:K_0]\dim_A A_i$ for any n, so the same is true of $((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$. Therefore,

$$((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V)^{\mathrm{Gal}_K} := \bigcup_n ((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger,s_n}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$$

is a finite module over the noetherian ring A, so the conclusion follows.

Now, let $D := ((A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V)^{\mathrm{Gal}_K}$, let $D_i := D\otimes_A A_i$ be the factor of D over A_i, let $v_1,\dots,v_d$ be an A-basis of V, and let $w_1,\dots,w_{d'}$ be an A_i-basis of D_i. Then $v_1,\dots,v_d$ is an $A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}$-basis of $(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V$ and $w_j \in (A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})\otimes_A V$, so there is a matrix $M \in \mathrm{Mat}_{d\times d'}(A\,\widehat{\otimes}\,\widetilde{\mathbf{B}}^{\dagger}_{\log})$ whose j-th column is the coordinates of w_j with respect to $v_1,\dots,v_d$. Let $P \in \mathrm{GL}_{d'}(A_i)$ be the matrix of $\varphi^{[K_0:\mathbf{Q}_p]}$

with respect to $w_1, \dots, w_{d'}$. To justify this, recall that $\varphi : \widetilde{\mathbf{B}}^{\dagger}_{\log} \to \widetilde{\mathbf{B}}^{\dagger}_{\log}$ is a bijection, and note that φ cyclically permutes the D_i so that $\varphi^{[K_0:\mathbf{Q}_p]}$ carries D_i to itself. Then $MP = \varphi^{[K_0:\mathbf{Q}_p]}(M)$, since φ acts trivially on $v_1, \dots, v_d$, so that $M = \varphi^{-[K_0:\mathbf{Q}_p]}(M)\varphi^{-[K_0:\mathbf{Q}_p]}(P)$. Then, by Frobenius regularization, M has coefficients in $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{+}_{\log} \subset A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{+}_{\mathrm{st}}$, so we are done. □

Remark 4.2.17. The conclusion of Proposition 4.2.16 is used in the proof of [Berger and Colmez 2008, Proposition 6.2.4]. Since the proof requires some minor adjustments when A is not $\mathbf{Q}_p$-finite, we have written out the details here.

We can deduce the same result for Galois representations with affinoid coefficients, generalizing [Berger 2002]:

Corollary 4.2.18. *Let A be an E-affinoid algebra, and let V be a finite free A-module of rank d equipped with a continuous action of* Gal_K. *Then the natural map*

$$((A \mathbin{\widehat{\otimes}} \mathbf{B}^{+}_{\mathrm{st}}) \otimes_A V)^{\mathrm{Gal}_K} \to ((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_A V)^{\mathrm{Gal}_K}$$

is an isomorphism.

Proof. Let $A \to R = \prod_i R_i$ be a closed embedding into a finite product of Artin rings, with R_i a finite-dimensional vector space over a complete discretely valued field B_i with perfect residue field; this is possible by Lemma 2.1.2. Then we have an exact sequence of $\mathbf{Q}_p$-Banach spaces $0 \longrightarrow V \longrightarrow V_R \longrightarrow V_R/V \longrightarrow 0$. Since $\mathbf{Q}_p$ is discretely valued, this exact sequence admits a continuous $\mathbf{Q}_p$-linear splitting, and we have a commutative diagram of Fréchet spaces

$$\begin{array}{ccccccccc}
0 & \longrightarrow & (A \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_A V & \longrightarrow & (R \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_R V_R & \longrightarrow & \widetilde{\mathbf{B}}^{\dagger}_{\log} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} (V_R/V) & \longrightarrow & 0 \\
 & & \big\uparrow & & \big\uparrow & & \big\uparrow & & \\
0 & \longrightarrow & (A \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \widetilde{\mathbf{B}}^{+}_{\mathrm{st}}) \otimes_A V & \longrightarrow & (R \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \widetilde{\mathbf{B}}^{+}_{\mathrm{st}}) \otimes_R V_R & \longrightarrow & \widetilde{\mathbf{B}}^{+}_{\mathrm{st}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} (V_R/V) & \longrightarrow & 0
\end{array}$$

where the rows are exact and the vertical maps are injections. Moreover, the maps are Gal_K-equivariant, so we have a commutative diagram of Banach spaces

$$\begin{array}{ccccccc}
0 & \longrightarrow & (\widetilde{\mathbf{B}}^{\dagger}_{\log} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K} & \longrightarrow & (\widetilde{\mathbf{B}}^{\dagger}_{\log} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} V_R)^{\mathrm{Gal}_K} & \longrightarrow & (\widetilde{\mathbf{B}}^{\dagger}_{\log} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} (V_R/V))^{\mathrm{Gal}_K} \\
 & & \big\uparrow & & \big\uparrow & & \big\uparrow \\
0 & \longrightarrow & (\widetilde{\mathbf{B}}^{+}_{\mathrm{st}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K} & \longrightarrow & (\widetilde{\mathbf{B}}^{+}_{\mathrm{st}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} V_R)^{\mathrm{Gal}_K} & \longrightarrow & (\widetilde{\mathbf{B}}^{+}_{\mathrm{st}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} (V_R/V))^{\mathrm{Gal}_K}
\end{array}$$

where the rows are still exact and the vertical maps are still injections. For each idempotent factor R_i of R, we can view V_{R_i} as a finite-dimensional B_i-vector space and apply Proposition 4.2.16; we see that the inclusion $((R \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{+}_{\mathrm{st}}) \otimes_R V_R)^{\mathrm{Gal}_K} \subset ((R \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_R V_R)^{\mathrm{Gal}_K}$ is an equality. Then a diagram chase shows that the inclusion $((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{+}_{\mathrm{st}}) \otimes_A V)^{\mathrm{Gal}_K} \subset ((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_A V)^{\mathrm{Gal}_K}$ is also an equality. □

We now define

$$\mathbf{D}^{\dagger,s}_{\log,K}(V) := (A \mathbin{\widehat{\otimes}} \mathbf{B}^{\dagger,s}_{\log,K}) \otimes_{A \widehat{\otimes} \mathbf{B}^{\dagger,s}_{\mathrm{rig,K}}} \mathbf{D}^{\dagger,s}_{\mathrm{rig,K}}(V), \qquad \mathbf{D}^{\dagger}_{\log,K}(V) := \bigcup_s \mathbf{D}^{\dagger,s}_{\mathrm{rig,K}}(V),$$

as well as

$$\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V) := (A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger,s}_{\log,K}) \otimes_{A \widehat{\otimes} \mathbf{B}^{\dagger,s}_{\log,K}} \mathbf{D}^{\dagger,s}_{\log,K}(V), \quad \widetilde{\mathbf{D}}^{\dagger}_{\log,K}(V) := \bigcup_s \widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V).$$

Proposition 4.2.19. *Let A be a noetherian $\mathbf{Q}_p$-Banach algebra with valuation ring $\mathscr{A}$, and let V be a finite free A-module of rank d equipped with a continuous A-linear action of* Gal_K *such that V admits a free* Gal_K*-stable $\mathscr{A}$-submodule of rank d. Then the natural map*

$$(\mathbf{D}^{\dagger}_{\log,K}(V)[1/t])^{\Gamma_K} \to ((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log})[1/t] \otimes_A V)^{\mathrm{Gal}_K}$$

is an isomorphism.

Proof. It suffices to prove this with K replaced by a finite extension, so we may assume $\mathbf{D}^{\dagger}_{\log,K}(V)$ is free. After twisting V by some power of the cyclotomic character, we may assume that $(\mathbf{D}^{\dagger}_{\log,K}(V)[1/t])^{\Gamma_K} = (\mathbf{D}^{\dagger}_{\log,K}(V))^{\Gamma_K}$, and consider only the map

$$(\mathbf{D}^{\dagger}_{\log,K}(V))^{\Gamma_K} \to ((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log,K}) \otimes_A V)^{\mathrm{Gal}_K}.$$

Furthermore, we observe that

$$\begin{aligned}((A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}) \otimes_A V)^{H_K} &= (A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log,K}) \otimes_{A \widehat{\otimes} \mathbf{B}^{\dagger}_{\log,K}} \mathbf{D}^{\dagger}_{\log,K}(V)\\ &= \widetilde{\mathbf{D}}^{\dagger}_{\log,K}(V).\end{aligned}$$

Since $(\widetilde{\mathbf{D}}^{\dagger}_{\log,K}(V))^{\Gamma_K}$ and $(\mathbf{D}^{\dagger}_{\log,K}(V))^{\Gamma_K}$ are finite modules over the noetherian Banach algebra $A \otimes_{\mathbf{Q}_p} K_0$, we see that $(\mathbf{D}^{\dagger}_{\log,K}(V))^{\Gamma_K}$ is a closed submodule of $(\widetilde{\mathbf{D}}^{\dagger}_{\log,K}(V))^{\Gamma_K}$, by [Bosch et al. 1984, Proposition 3.7.3/1]. Thus, it suffices to show that $(\mathbf{D}^{\dagger}_{\log,K}(V))^{\Gamma_K}$ is dense in $(\widetilde{\mathbf{D}}^{\dagger}_{\log,K}(V))^{\Gamma_K}$.

We will actually do something slightly different. For $s \gg 0$ and any integer $k \geq 0$, we consider the $A \widehat{\otimes} \varphi^{-k}(\mathbf{B}^{\dagger,p^k s}_{\log,K})$-submodule $\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V)) \subset \widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V)$. Since $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$ is a finite A-module, it follows that $\big(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\big)^{\Gamma_K}$ is a closed submodule of $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$.

We claim that $\bigcup_k \big(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\big)^{\Gamma_K} = (\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$. If we choose a basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ of $\mathbf{D}^{\dagger,s}_{\log,K}(V)$, then, for any Γ_K-fixed element $m \in \widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V)$, we may write $m = a_1\boldsymbol{v}_1 + \cdots + a_d\boldsymbol{v}_d$. Recall that there are Γ_K-equivariant maps

$$R_k : \widetilde{\mathbf{B}}^{\dagger,s}_{\mathrm{rig,K}} \to \varphi^{-k}(\mathbf{B}^{\dagger,p^k s}_{\mathrm{rig,K}})$$

which are sections to the inclusions $\varphi^{-k}(\mathbf{B}^{\dagger,p^k s}_{\mathrm{rig,K}}) \subset \widetilde{\mathbf{B}}^{\dagger,s}_{\mathrm{rig,K}}$ and extend to maps

$$R_k : \widetilde{\mathbf{B}}^{\dagger,s}_{\log,K} \to \varphi^{-k}(\mathbf{B}^{\dagger,p^k s}_{\log,K}).$$

For each k, let $m_k = R_k(m) = R_k(a_1)\boldsymbol{v}_1 + \cdots + R_k(a_d)\boldsymbol{v}_d$. Then m_k is a Γ_K-fixed element of $\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))$, because R_k is Γ_K-equivariant. Since $\lim_{k\to\infty} R_k(a) = a$ for any $a \in \widetilde{\mathbf{B}}^{\dagger,s}_{\log,K}$, it follows that $\lim_{k\to\infty} m_k = m$. Thus, $\bigcup_k \bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr)^{\Gamma_K}$ is dense in $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$. Equality follows, since it is also closed in $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$ (as it is a submodule of a finite module over a noetherian Banach algebra).

Next, we note that

$$\varphi^{k+1}\bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr) \subset \varphi(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V)) \subset \mathbf{D}^{\dagger,p^{k+1} s}_{\log,K}(V).$$

This implies that $\bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr)^{\Gamma_K} \subset \bigl(\varphi^{-(k+1)}(\mathbf{D}^{\dagger,p^{k+1} s}_{\log,K}(V))\bigr)^{\Gamma_K}$, and therefore that $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K} = \bigcup_k \bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr)^{\Gamma_K}$ is a rising union. Since $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$ is A-finite, there is some k such that $\bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr)^{\Gamma_K}$ is equal to $(\widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V))^{\Gamma_K}$.

But we have A-linear isomorphisms

$$\varphi^k : \bigl(\varphi^{-k}(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))\bigr)^{\Gamma_K} \to (\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))^{\Gamma_K} \text{ and } \varphi^k : \widetilde{\mathbf{D}}^{\dagger,s}_{\log,K}(V) \xrightarrow{\sim} \widetilde{\mathbf{D}}^{\dagger,p^k s}_{\log,K}(V),$$

so we conclude that $(\mathbf{D}^{\dagger,p^k s}_{\log,K}(V))^{\Gamma_K} = (\widetilde{\mathbf{D}}^{\dagger,p^k s}_{\log,K}(V))^{\Gamma_K}$, as desired. □

Now we can prove Theorem 4.2.9:

Proof of Theorem 4.2.9. We may assume that $X = \mathrm{Sp}(A)$ for some E-affinoid algebra A, and that $V := \mathscr{V}(A)$ is A-free of rank d and admits a Gal_K-stable integral lattice. Then $\mathbf{D}^K_{\mathrm{st}}(V) = (\mathbf{D}_{\log,K}(V)[1/t])^{\Gamma_K}$ by Corollary 4.2.18 and Proposition 4.2.19. Since

$$\mathbf{D}_{\mathrm{cris}}(V) = \mathbf{D}_{\mathrm{st}}(V)^{N=0} \text{ and } \mathbf{D}_{\mathrm{rig,K}}(V) = \mathbf{D}_{\log,K}(V)^{N=0},$$

it follows that $\mathbf{D}^K_{\mathrm{cris}}(V) = (\mathbf{D}_{\mathrm{rig,K}}(V)[1/t])^{\Gamma_K}$. □

4.3. *Properties of* $\mathbf{D}_{\mathbf{B}_*}(V)$. Now we can combine Theorems 4.2.5 and 4.2.8 with "cohomology and base change" to deduce various useful properties of the functors $V \mapsto \mathbf{D}_{\mathrm{HT}}(V)$ and $V \mapsto \mathbf{D}_{\mathrm{dR}}(V)$.

Theorem 4.3.1. *Let X and $\mathscr{V}$ be as above.*

(1) *$\mathscr{D}^K_{\mathrm{HT}}(\mathscr{V})$ and $\mathscr{D}^K_{\mathrm{dR}}(\mathscr{V})$ are coherent sheaves of $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$-modules. More generally, their formation commutes with flat base change on X.*

(2) *$\mathscr{D}_{\mathrm{HT}}^K(\mathscr{V})$ and $\mathscr{D}_{\mathrm{dR}}^K(\mathscr{V})$ take values in the categories of graded coherent sheaves and filtered coherent sheaves, respectively. If $\mathscr{V}$ is $\mathbf{B}_{\mathrm{HT}}$-admissible, then $\mathscr{D}_{\mathrm{HT}}^K(\mathscr{V})$ is a graded vector bundle over $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$, and, if $\mathscr{V}$ is $\mathbf{B}_{\mathrm{dR}}$-admissible, then $\mathscr{D}_{\mathrm{HT}}^K(\mathscr{V})$ is a filtered vector bundle over $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$.*

As before, we reduce immediately to the case when $X = \mathrm{Sp}(A)$ and $V := \Gamma(X, \mathscr{V})$ is a free A-linear representation of Gal_K which admits a free Gal_K-stable $\mathscr{A}$-lattice V_0 of rank d.

Proposition 4.3.2. *Let A be a noetherian $\mathbf{Q}_p$-Banach algebra with valuation ring $\mathscr{A}$, let V_0 be a free $\mathscr{A}$-module of rank d equipped with a continuous $\mathscr{A}$-linear action of* Gal_K, *and let $V := V_0[1/p]$. Then $\mathbf{D}_{\mathrm{HT}}^K(V)$ and $\mathbf{D}_{\mathrm{dR}}^K(V)$ are finite $A \otimes_{\mathbf{Q}_p} K$-modules.*

Proof. Recall that

$$\mathbf{D}_{\mathrm{HT}}^K(V) = \bigoplus_{i \in \mathbf{Z}} (\mathbf{D}_{\mathrm{Sen}}^K(V) \cdot t^i)^{\Gamma_K}.$$

Now $(\mathbf{D}_{\mathrm{Sen}}^K(V) \cdot t^i)^{\Gamma_K} = (\mathbf{D}_{\mathrm{Sen}}^K(V))^{\Gamma_K = \chi^{-i}}$ for every $i \in \mathbf{Z}$, so $\mathbf{D}_{\mathrm{HT}}^K(V) \subset \mathbf{D}_{\mathrm{Sen}}^K(V)$. But $\mathbf{D}_{\mathrm{Sen}}^K(V)$ is a finite module over the noetherian ring $A \otimes_{\mathbf{Q}_p} K$, so $\mathbf{D}_{\mathrm{HT}}^K(V)$ is A-finite as well.

Moreover, we observe that the summands of $\mathbf{D}_{\mathrm{HT}}^K(V)$ have pairwise trivial intersection. Therefore, only finitely many of them are nonzero.

To see that $\mathbf{D}_{\mathrm{dR}}^K(V)$ is finite over $A \otimes_{\mathbf{Q}_p} K$, we observe that

$$\mathrm{gr}^\bullet \mathbf{D}_{\mathrm{dR}}^K(V) \hookrightarrow (\mathrm{gr}^\bullet((A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_A V))^{\mathrm{Gal}_K} = \mathbf{D}_{\mathrm{HT}}^K(V).$$

In fact, we claim that there exist integers i_0, i_1 such that $\mathrm{Fil}^i \mathbf{D}_{\mathrm{dR}}^K(V) = \mathrm{Fil}^{i_0} \mathbf{D}_{\mathrm{dR}}^K(V)$ for all $i \leq i_0$ and $\mathrm{Fil}^i \mathbf{D}_{\mathrm{dR}}^K(V) = \mathrm{Fil}^{i_1} \mathbf{D}_{\mathrm{dR}}^K(V) = 0$ for all $i \geq i_1$. Indeed,

$$\mathrm{gr}^i \mathbf{D}_{\mathrm{dR}}^K(V) \hookrightarrow (\mathrm{gr}^i((A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_A V))^{\mathrm{Gal}_K} = (\mathbf{D}_{\mathrm{Sen}}^K(V) \cdot t^i)^{\mathrm{Gal}_K}.$$

But the rightmost term is one of the summands of $\mathbf{D}_{\mathrm{HT}}^K(V)$, and only finitely many such summands are nonzero. Therefore, $\mathrm{gr}^i \mathbf{D}_{\mathrm{dR}}^K(V) = 0$ for $i \ll 0$ and $i \gg 0$, and $\mathbf{D}_{\mathrm{dR}}^K(V)$ is $A \otimes_{\mathbf{Q}_p} K$-finite. $\square$

Definition 4.3.3. The *Hodge–Tate weights* of V are those integers i such that

$$(\mathbf{D}_{\mathrm{Sen}}^K(V) \cdot t^i)^{\mathrm{Gal}_K} \neq 0.$$

Remark 4.3.4. This is a slight departure from the traditional definition of Hodge–Tate weights, which are usually only defined for representations which are Hodge–Tate. However, we will find this abuse of terminology convenient.

Proposition 4.3.5. *Let V be a free A-module of rank d, equipped with a continuous, A-linear action of G_K. Then $\mathrm{H}^1\big(\Gamma_K, \bigoplus_{k\in\mathbf{Z}} t^k\mathbf{D}_{\mathrm{Sen}}(V)\big)$ and $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))$ are A-finite if and only if there is a bounded interval $[a,b]$ containing all the Hodge–Tate weights of the fibral representations.*

Proof. We first reduce to the case where Γ_K is procyclic. In general, $\Gamma_K \cong \Delta\times\Gamma_K'$, where Δ is a finite abelian group and Γ_K' is procyclic. The statement about the fibral Hodge–Tate weights can be checked after restriction to a finite-index subgroup of Γ_K, and, in particular, after restriction to Γ_K'. On the other hand, taking Δ-invariants on $\mathbf{Q}_p$-vector spaces is an exact functor, so $\mathrm{H}^1\big(\Gamma_K', \bigoplus_{k\in\mathbf{Z}} t^k\mathbf{D}_{\mathrm{Sen}}(V)\big) = \mathrm{H}^1\big(\Gamma_K, \bigoplus_{k\in\mathbf{Z}} t^k\mathbf{D}_{\mathrm{Sen}}(V)\big)^\Delta$ and $\mathrm{H}^1(\Gamma_K', \mathbf{D}_{\mathrm{dif}}(V)) = \mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))^\Delta$. It follows that we can also check the finiteness of $\mathrm{H}^1\big(\Gamma_K, \bigoplus_{k\in\mathbf{Z}} t^k\mathbf{D}_{\mathrm{Sen}}(V)\big)$ and $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))$ after restriction to Γ_K'. We may therefore assume that Γ_K is torsion-free, and apply the results of Section 3.

The statement is clear for $\mathrm{H}^1\big(\Gamma_K, \bigoplus_{k\in\mathbf{Z}} t^k\mathbf{D}_{\mathrm{Sen}}(V)\big)$.

Suppose first that the fibral Hodge–Tate weights are contained in an interval $[a,b]$. The natural map

$$\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}^+(V)/t^{k+1}) \to \mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}^+(V)/t^k)$$

is a surjection for all $k\geq 0$, and its kernel is surjected onto by $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))$. But the formation of $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))$ commutes with arbitrary base change on A, so the hypothesis on the fibral Hodge–Tate weights implies that, if $k>b$, $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))$ is trivial when reduced modulo any power of any maximal ideal of A. Therefore, $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))$ is itself trivial, and $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}^+(V)) \cong \mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}^+(V)/t^{\max\{0,b\}+1})$, which is A-finite. This implies that, for any $k\geq 0$,

$$\mathrm{H}^1(\Gamma_K, t^{-k}\mathbf{D}_{\mathrm{dif}}^+(V)) \cong \mathrm{H}^1(\Gamma_K, t^{-k}\mathbf{D}_{\mathrm{dif}}^+(V)/t^{\max\{0,b\}+1})$$

is A-finite, as well. Further, the proof of Corollary 3.2.3(3) shows that, for any $k\in\mathbf{Z}$, the cokernel of the natural map

$$\mathrm{H}^1(\Gamma_K, t^{-k}\mathbf{D}_{\mathrm{dif}}^+(V)) \to \mathrm{H}^1(\Gamma_K, t^{-(k+1)}\mathbf{D}_{\mathrm{dif}}^+(V))$$

is $\mathrm{H}^1(\Gamma_K, t^{-(k+1)}\mathbf{D}_{\mathrm{Sen}}(V))$. But the hypothesis on the fibral Hodge–Tate weights implies that this is 0 for $k\geq -a$, so $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))$ is A-finite.

Conversely, suppose that $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))$ is A-finite. We need to show that $\mathrm{H}^1(\Gamma_K, t^k\cdot\mathbf{D}_{\mathrm{Sen}}(V)) = 0$ for $k\gg 0$ and $k\ll 0$. By Corollary 3.2.3(3), there exist integers N_0, N_0' such that, for $k\geq N_0$ or $k\leq N_0'$, the transition maps $\mathrm{H}^1(\Gamma_K, t^{k+1}\mathbf{D}_{\mathrm{dif}}^+(V)) \to \mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{dif}}^+(V))$ are injective. These transition maps moreover always have A-finite kernels and cokernels. Since A is noetherian, this implies that $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{dif}}^+(V))$ is finite for all $k\in\mathbf{Z}$.

Let $x\in \mathrm{Sp}(A)$ and let $\kappa(x)$ denote the residue field of A at x. By Proposition 3.2.1,

$$\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))\otimes_A \kappa(x)\cong \mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V)\,\widehat{\otimes}_A\,\kappa(x))$$

for all $k\geq 0$; it follows that there is some $N_{1,x}\geq 0$ such that, for all $k\geq 0$, $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))\otimes_A\kappa(x)=0$. Since $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))$ is a finite A-module for all k, there is some Zariski-open $U_x\subset \mathrm{Sp}(A)$ such that $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))|_{U_x}=0$ for all $k\geq N_{1,x}$. Since $\mathrm{Spec}\, A$ is quasicompact, it follows that there is some $N_1\gg 0$ such that $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))=0$ for all $k\geq N_1$. Thus, $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))=0$ for $k\geq N_1$.

Finally, the finiteness of $\mathrm{H}^1(\Gamma_K, \mathbf{D}_{\mathrm{dif}}(V))$ implies that there exists $N_1'\in\mathbf{Z}$ such that the transition map $\mathrm{H}^1(\Gamma_K, t^{k+1}\mathbf{D}^+_{\mathrm{dif}}(V))\to \mathrm{H}^1(\Gamma_K, t^k\mathbf{D}^+_{\mathrm{dif}}(V))$ has vanishing cokernel for $k\leq N_0'$. But this cokernel is $\mathrm{H}^1(\Gamma_K, t^k\mathbf{D}_{\mathrm{Sen}}(V))$, so it follows that the fibral Hodge–Tate weights are bounded. □

Now we can deduce that $\mathbf{D}^K_{\mathrm{cris}}(V)$ and $\mathbf{D}^K_{\mathrm{st}}(V)$ are finite modules, as well:

Corollary 4.3.6. *$\mathbf{D}^K_{\mathrm{cris}}(V)$ and $\mathbf{D}^K_{\mathrm{st}}(V)$ are finite $A\otimes_{\mathbf{Q}_p}K_0$-modules.*

Proof. Recall that there is an injection $\mathbf{B}_{\max}\to\mathbf{B}_{\mathrm{dR}}$. Since A is a Banach space over the discretely valued field $\mathbf{Q}_p$ (and therefore potentially orthonormalizable), this extends to an injection $A\,\widehat{\otimes}\,\mathbf{B}_{\max}\hookrightarrow A\,\widehat{\otimes}\,\mathbf{B}_{\mathrm{dR}}$. It follows that $\mathbf{D}^K_{\mathrm{cris}}(V)\hookrightarrow\mathbf{D}^K_{\mathrm{dR}}(V)$ and $\mathbf{D}^K_{\mathrm{cris}}(V)$ is A-finite.

Similarly, $\mathbf{B}_{\mathrm{st}}$ can be injected into $\mathbf{B}_{\mathrm{dR}}$ (although this depends on a choice of p-adic logarithm), so $\mathbf{D}^K_{\mathrm{st}}(V)\hookrightarrow\mathbf{D}^K_{\mathrm{dR}}(V)$. Thus, $\mathbf{D}^K_{\mathrm{st}}(V)$ is A-finite. □

Note that $\mathbf{D}^K_{\mathrm{cris}}(V)$ and $\mathbf{D}^K_{\mathrm{st}}(V)$ are equipped with semilinear actions of Frobenius φ (over $1\otimes\varphi$ on $A\otimes_{\mathbf{Q}_p}K_0$) coming from the coefficients, and $\mathbf{D}^K_{\mathrm{st}}(V)$ has a monodromy operator N coming from the coefficients and satisfying $N\circ\varphi=p\varphi\circ N$.

We turn to base change properties of the functors $\mathbf{D}_{\mathbf{B}_*}(V)$:

Proposition 4.3.7. *Let $f:A\to A'$ be a flat morphism of E-affinoid algebras.*

(1) $A'\otimes_A\mathbf{D}^K_{\mathrm{HT}}(V)\xrightarrow{\sim}\mathbf{D}^K_{\mathrm{HT}}(V\otimes_A A')$.

(2) $A'\otimes_A\mathbf{D}^K_{\mathrm{dR}}(V)\xrightarrow{\sim}\mathbf{D}^K_{\mathrm{dR}}(V\otimes_A A')$.

It follows that $U\mapsto\mathbf{D}^K_{\mathrm{HT}}(V_U)$ and $U\mapsto\mathbf{D}^K_{\mathrm{dR}}(V_U)$ are coherent sheaves on $\mathrm{Sp}(A)$.

Proof. (1) This follows by noting that $\mathbf{D}^{L_n}_{\mathrm{HT}}(V)=\varinjlim_{h\to\infty}\left(\bigoplus_{k=-h}^h t^k\mathbf{D}^{L_n}_{\mathrm{Sen}}(V)\right)^{\Gamma_{L_n}}$ and the formation of $\left(\bigoplus_{k=-h}^h t^k\mathbf{D}^{L_n}_{\mathrm{Sen}}(V)\right)^{\Gamma_{L_n}}$ commutes with flat base change, by Theorem 3.1.1. We are done, since $\mathbf{D}^{L_n}_{\mathrm{HT}}(V)=L_n\otimes_K\mathbf{D}^K_{\mathrm{HT}}(V)$.

(2) We apply Proposition 3.2.2 to $M=\mathbf{D}^{L_n,+}_{\mathrm{dif}}$ to see that

$$\mathbf{D}^{L_n,+}_{\mathrm{dR}}(V)=(\mathbf{D}^{L_n,+}_{\mathrm{dif}}(V)/t^N)^{\Gamma_{L_n}=1}$$

for some $N \geq 0$. Similarly, $\mathbf{D}_{\mathrm{dR}}^{L_n,+}(V \otimes_A A') = (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V \otimes_A A')/t^{N'})^{\Gamma_{L_n}=1}$. Since the $\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^k$ are finite A-modules,

$$\begin{aligned}\mathbf{D}_{\mathrm{dR}}^{L_n,+}(V \otimes_A A') &= \mathrm{H}^0(\Gamma_{L_n}, \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V \otimes_A A')/t^{\max\{N,N'\}})\\ &= \mathrm{H}^0(\Gamma_{L_n}, \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^{\max\{N,N'\}}) \otimes_A A' = \mathbf{D}_{\mathrm{dR}}^{L_n,+}(V) \otimes_A A',\end{aligned}$$

where the second equality again follows from Theorem 3.1.1. □

Conjecture 4.3.8. The formation of $\mathbf{D}_{\mathrm{cris}}^K(V)$ and $\mathbf{D}_{\mathrm{st}}^K(V)$ commutes with flat base change on A.

Remark 4.3.9. Conjecture 4.3.8 does not follow automatically from Theorem 3.1.1, because the (φ, Γ)-module $\mathbf{D}_{\mathrm{rig,K}}^{\dagger}(V)$ is not A-finite. It is also difficult, in general, to verify this conjecture for particular examples of families of Galois representations.

However, life is considerably better when considering a *trianguline* family, because one can compute with rank-1 families. Triangulation results for certain families of Galois representations arising from eigenvarieties are proved in [Hellmann 2012; Kedlaya et al. 2014; Liu 2014]. In particular, Kedlaya, Pottharst, and Xiao [Kedlaya et al. 2014, Theorem 6.3.9] show that if $\mathscr{V}$ is the family of Galois representations on the normalization X of the eigencurve, then away from the image of the θ^{k-1}-map, the associated family $\mathscr{D}_{\mathrm{rig}}^{\dagger}(\mathscr{V})$ has a global triangulation

$$0 \longrightarrow \mathscr{D}_1 \longrightarrow \mathscr{D}_{\mathrm{rig}}^{\dagger}(\mathscr{V}) \longrightarrow \mathscr{D}_2 \longrightarrow 0.$$

Here the $\mathscr{D}_i = \mathscr{R}_X(\delta_i) \otimes_X \mathscr{L}_i$ are rank-1 families of (φ, Γ)-modules over X. We have used the notation of [Kedlaya et al. 2014]: $\delta_i : \mathbf{Q}_p^\times \to \Gamma(X, \mathscr{O}_X^\times)$ are continuous characters, $\mathscr{R}_X(\delta_i)$ is the free rank-1 $(\varphi, \Gamma_{\mathbf{Q}_p})$-module with basis $\boldsymbol{e}$ such that $\varphi(\boldsymbol{e}) = \delta(p)\boldsymbol{e}$ and $\gamma(\boldsymbol{e}) = \delta(\chi(\gamma))\boldsymbol{e}$ for $\gamma \in \Gamma_{\mathbf{Q}_p}$, and $\mathscr{L}_i$ are line bundles on X with no action of φ or $\Gamma_{\mathbf{Q}_p}$.

In this case, $\delta_1|_{\mathbf{Z}_p^\times}$ is trivial, while $\delta_2|_{\mathbf{Z}_p^\times}$ is the weight-nebentypus character. As a result, for any affinoid $U \subset X$ which trivializes $\mathscr{L}_2$, we have $\mathscr{D}_2(U)[1/t]^{\Gamma_{\mathbf{Q}_p}} = 0$ and so, by Theorem 4.2.9,

$$\mathscr{D}_{\mathrm{cris}}(\mathscr{V})(U) = \mathscr{D}_1(U)[1/t]^{\Gamma_{\mathbf{Q}_p}}.$$

Moreover, by construction,

$$\mathscr{D}_1(U)[1/t]^{\Gamma_{\mathbf{Q}_p}} = \mathscr{D}_1(U)[1/t]^{\varphi=\delta(p),\Gamma_{\mathbf{Q}_p}=1} = \mathscr{D}_{\mathrm{cris}}(\mathscr{V})(U)^{\varphi=\delta(p)}.$$

It follows from [Kedlaya et al. 2014] or [Liu 2014] that $\mathscr{D}_{\mathrm{cris}}(\mathscr{V})^{\varphi=\delta(p)}$ is a coherent $\mathscr{O}_X$-module, and hence so is $\mathscr{D}_{\mathrm{cris}}(\mathscr{V})$. This is a very natural example of a family $\mathscr{V}$ of Galois representations such that $\mathscr{D}_{\mathrm{cris}}(\mathscr{V})$ is a coherent sheaf.

5. $\mathbf{B}_*$-admissible loci

5.1. ***Overview.*** In this section, we fix a family $\mathscr{V}$ of rank-d representations of Gal_K over an E-analytic space X, and we study the loci on X where $\mathscr{V}$ is $\mathbf{B}_*$-admissible for various period rings $\mathbf{B}_*$. We have the following theorem:

Theorem 5.1.1. *Let X and $\mathscr{V}$ be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$. Then there is a closed subspace $X_{\mathbf{B}_*}^{[a,b]} \hookrightarrow X$ such that for any E-finite Artin local ring B, a map $x : \mathrm{Sp}(B) \to X$ factors through $X_{\mathbf{B}_*}^{[a,b]}$ if and only if the induced B-linear Galois representation $\mathscr{V}_x$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$.*

If the family $\mathscr{V}$ is $\mathbf{B}_*$-admissible, then certainly for every morphism $f : X' \to X$ the base change $f^*\mathscr{V}$ is $\mathbf{B}_*$-admissible. We prove a converse theorem in two parts:

Theorem 5.1.2. *Let X and $\mathscr{V}$ be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$. Suppose that $\mathscr{V}_x$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$ for every morphism $x : \mathrm{Sp}(B) \to X$, where B is an E-finite Artin local ring.*

(1) *The sheaf $\mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V})$ is a sheaf of projective $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-modules of rank d.*

(2) *The formation of $\mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V})$ commutes with arbitrary base change on X.*

In each case, we then use the base change property to finish proving that $\mathscr{V}$ is a $\mathbf{B}_*$-admissible family of Galois representations:

Theorem 5.1.3. *Let X and $\mathscr{V}$ be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}, \mathrm{st}, \mathrm{cris}\}$. Suppose that $\mathscr{V}_x$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$ for every morphism $x : \mathrm{Sp}(B) \to X$, where B is an E-finite Artin local ring. Then the natural map $\mathscr{B}_{X,*} \otimes_{\mathscr{O}_X \otimes \mathbf{B}_*^{\mathrm{Gal}_K}} \mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V}) \to \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}$ is an isomorphism.*

Remark 5.1.4. We do not know whether assuming that $\mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V})$ is a sheaf of projective $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-modules of rank d implies that the formation of $\mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V})$ commutes with base change. This is why our definition of $\mathbf{B}_*$-admissibility of a family includes the condition that $\mathscr{B}_{X,*} \otimes_{\mathscr{O}_X \otimes \mathbf{B}_*^{\mathrm{Gal}_K}} \mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V}) \to \mathscr{B}_{X,*} \otimes_{\mathscr{O}_X} \mathscr{V}$ is an isomorphism.

If the natural base change morphism $B \otimes_{\mathscr{O}_X} \mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V}) \to \mathscr{D}_{\mathbf{B}_*}^K(\mathscr{V} \otimes_{\mathscr{O}_X} B)$ were injective for all morphisms $\mathrm{Sp}(B) \to X$, with B an E-finite Artin local ring, we could deduce that $\mathscr{V}$ is $\mathbf{B}_*$-admissible. However, the low-degree exact sequence in Corollary 3.1.3 shows that there is an obstruction to such injectivity when $* \in \{\mathrm{HT}, \mathrm{dR}\}$, at least a priori.

When $\mathbf{B}_* = \mathbf{B}_{\mathrm{HT}}$ or $\mathbf{B}_{\mathrm{dR}}$, we can prove finer results. Fix an interval $[a,b]$, and define

$$\mathscr{D}_{\mathrm{HT}}^{[a,b]}(\mathscr{V}) := \left(\left(\mathscr{O}_X \otimes_{\mathbf{Q}_p} \bigoplus_{i=a}^{b} \mathbf{C}_K \cdot t^i\right) \otimes_{\mathscr{O}_X} \mathscr{V}\right)^{\mathrm{Gal}_K},$$

$$\mathscr{D}_{\mathrm{dR}}^{[a,b]}(\mathscr{V}) := ((\mathscr{O}_X \otimes_{\mathbf{Q}_p} t^a \mathbf{B}_{\mathrm{dR}}/t^b) \otimes_{\mathscr{O}_X} \mathscr{V})^{\mathrm{Gal}_K}.$$

We think of these coherent sheaves as sheaves of "periods in Hodge–Tate weight $[a,b]$". Fix an integer $0 \leq d' \leq d$.

Definition 5.1.5. A morphism $f : X \to X'$ is a *Zariski-locally closed immersion* if there is a Zariski-open subspace $U \subset X'$ such that f factors through a Zariski-closed immersion $X \hookrightarrow U$.

Theorem 5.1.6. *Let X and $\mathscr{V}$ be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}\}$. There is a Zariski-locally closed immersion $X^{[a,b]}_{\mathbf{B}_*,d'} \hookrightarrow X$ such that $x : \mathrm{Sp}(B) \to X$ factors through $X^{[a,b]}_{\mathbf{B}_*,d'}$ if and only if $\mathscr{D}^{[a,b]}_{\mathbf{B}_*}(\mathscr{V}_x)$ is a free $B \otimes_{\mathbf{Q}_p} K$-module of rank d', where B is an E-finite Artin local ring and $\mathscr{V}_B := \mathscr{V} \otimes_A B$. In fact, $X^{[a,b]}_{\mathbf{B}_*,d'} = X^{[a,b]}_{\mathbf{B}_*,\leq d'} \cap X^{[a,b]}_{\mathbf{B}_*,\geq d'}$, where $X^{[a,b]}_{\mathbf{B}_*,\leq d'} \subset X$ is Zariski-open and $X^{[a,b]}_{\mathbf{B}_*,\geq d'} \hookrightarrow X$ is Zariski-closed.*

In fact, the $X^{[a,b]}_{\mathbf{B}_*,d'}$ give a stratification of X, in the sense that $X^{[a,b]}_{\mathbf{B}_*,\leq d'-1} = X^{[a,b]}_{\mathbf{B}_*,\leq d'} \setminus X^{[a,b]}_{\mathbf{B}_*,d'}$ and $X = X^{[a,b]}_{\mathbf{B}_*,\leq d}$.

Theorem 5.1.7. *Let X and $\mathscr{V}$ be as above, and let $* \in \{\mathrm{HT}, \mathrm{dR}\}$. Suppose that for every E-finite artinian point $x : A \to B$, the $B \otimes_{\mathbf{Q}_p} K$-module $\mathscr{D}^{[a,b]}_{\mathbf{B}_*}(\mathscr{V}_x)$ is free of rank d', where $0 \leq d' \leq d$.*

(1) *$\mathscr{D}^{[a,b]}_{\mathbf{B}_*}(\mathscr{V})$ is a rank-d' locally free $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$-module, and the $(t^k \cdot \mathscr{D}^{K_n}_{\mathrm{Sen}}(\mathscr{V}))^{\Gamma_K}$ are locally free $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$-modules.*

(2) *The formation of $\mathscr{D}^{[a,b]}_{\mathbf{B}_*}(\mathscr{V})$ commutes with arbitrary base change $f : X' \to X$.*

If $d' = d$, then:

(3) *$\mathscr{D}_{\mathbf{B}_*}(\mathscr{V}) = \mathscr{D}^{[a,b]}_{\mathbf{B}_*}(\mathscr{V})$.*

(4) *The natural morphism*

$$\alpha_{\mathscr{V}} : \mathscr{B}_* \otimes_{\mathscr{O}_X \otimes_{\mathbf{Q}_p} K} \mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V}) \to \mathscr{B}_* \otimes_{\mathscr{O}_X} \mathscr{V}$$

is an isomorphism.

Before we begin proving Theorems 5.1.1, 5.1.2, and 5.1.3 and their refinements, we discuss some consequences.

First of all, the subspaces $X^{[a,b]}_{\mathbf{B}_*} \hookrightarrow X$ have strong functorial properties:

Corollary 5.1.8. *Let X and $\mathscr{V}$ be as above, and let $f : X' \to X$ be a morphism of rigid analytic spaces.*

(1) *f factors through $X^{[a,b]}_{\mathbf{B}_*}$ if and only if $f^*\mathscr{V}$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$.*

(2) *The subspace $X'^{[a,b]}_{\mathbf{B}_*} \hookrightarrow X'$ is induced via base change from $X^{[a,b]}_{\mathbf{B}_*} \hookrightarrow X$.*

Proof. (1) By Theorems 5.1.1 and 5.1.2, $f^*\mathscr{V}$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$ if and only if $x^* f^* \mathscr{V}$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in the interval $[a,b]$ for every E-finite Artin local point $x : \mathrm{Sp}(B) \to X'$.

We may assume that X and X' are affinoid, with $X = \mathrm{Sp}(A)$ and $X' = \mathrm{Sp}(A')$ for E-affinoid algebras A and A'.

Suppose that $V_{A'}$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in $[a, b]$. Then, for every maximal ideal $\mathfrak{m}' \subset A'$ and every integer $n \geq 0$, the composition $A \to A' \to A'/\mathfrak{m}'^n$ factors through $A^{[a,b]}_{\mathbf{B}_*}$. In other words, if $x \in \ker(A \twoheadrightarrow A^{[a,b]}_{\mathbf{B}_*})$, then $f(x) \in \mathfrak{m}'^n$ for all $\mathfrak{m}'$ and all n. Since $\bigcap_{n,\mathfrak{m}'} \mathfrak{m}'^n = 0$, we see that $f(x) = 0$.

Suppose conversely that f factors through $A \twoheadrightarrow A^{[a,b]}_{\mathbf{B}_*}$, and consider an E-finite artinian point $x : A' \to B'$. By assumption, the composition $A \to A' \to B'$ factors through $A \twoheadrightarrow A^{[a,b]}_{\mathbf{B}_*}$, so the induced representation $V_{B'}$ is $\mathbf{B}_*$-admissible with Hodge–Tate weights in $[a, b]$. Thus, $V_{A'}$ is Hodge–Tate.

(2) This follows from the first part, and from the universal property of fiber products of rigid analytic spaces. □

Similarly, if $* \in \{\mathrm{HT}, \mathrm{dR}\}$, the subspaces $X^{[a,b]}_{\mathbf{B}_*,d'}$ are functorial for any $0 \leq d' \leq d$:

Corollary 5.1.9. *Let X and $\mathscr{V}$ be as above, and let $f : X' \to X$ be a morphism of rigid analytic spaces.*

(1) *f factors through $X^{[a,b]}_{\mathbf{B}_*,d'}$ if and only if $\mathscr{D}^{[a,b]}_{\mathbf{B}_*}(f^*\mathscr{V})$ is a locally free $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$-module of rank d'.*

(2) *The subspace $X'^{[a,b]}_{\mathbf{B}_*,d'} \hookrightarrow X'$ is induced via base change from $X^{[a,b]}_{\mathbf{B}_*,d'} \hookrightarrow X$.*

The proof proceeds identically to the proof of Corollary 5.1.8.

We can also refine our conclusions about the structure of $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$.

Corollary 5.1.10. *Let $\mathscr{V}$ be a family of Galois representations such that, for any E-finite artinian point $x : \mathrm{Sp}(B) \to X$, the specialization $\mathscr{V}_x$ is $\mathbf{B}_*$-admissible. Then the vector bundle $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ is X-locally free.*

Remark 5.1.11. The hypothesis on $\mathscr{V}$ is phrased in terms of a pointwise condition because this corollary is used in the proof of Theorem 5.1.3 when $* \in \{\mathrm{st}, \mathrm{cris}\}$.

Proof. It suffices to show that, for any point $x \in X$, the completed stalk $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})^\wedge_x$ is a free $\mathscr{O}^\wedge_{X,x} \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module. Since $\mathscr{V}$ is assumed to be $\mathbf{B}_*$-admissible, the formation of $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ commutes with arbitrary base change on X, and hence

$$\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})^\wedge_x = \varprojlim_n \mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V} \otimes_{\mathscr{O}_X} \mathscr{O}_{X,x}/\mathfrak{m}_x^n).$$

But $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V} \otimes_{\mathscr{O}_X} \mathscr{O}_{X,x}/\mathfrak{m}_x^n)$ is a free $\mathscr{O}_{X,x}/\mathfrak{m}_x^n \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module of rank d, again by the $\mathbf{B}_*$-admissibility of $\mathscr{V}$, and the transition maps are simply reduction modulo $\mathfrak{m}_x^n$. Therefore, $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})^\wedge_x$ is $\mathscr{O}^\wedge_{X,x} \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-free of rank d, as desired. □

If $\mathscr{V}$ is Hodge–Tate (resp. de Rham), the graded pieces $\mathrm{gr}^i \mathscr{D}^K_{\mathrm{HT}}(\mathscr{V})$ (resp. the submodules $\mathrm{Fil}^i \mathscr{D}^K_{\mathrm{dR}}(\mathscr{V})$) need not be X-locally free as $\mathscr{O}_X \otimes_{\mathbf{Q}_p} K$-modules; an example is given in [Breuil and Mézard 2002, Remarque 3.1.1.4]. However, we can use Theorem 5.1.1 to cut out the locus where $\mathscr{D}_{\mathrm{dR}}(\mathscr{V})$ is filtered by subbundles and has specified Hodge polygon.

Definition 5.1.12. Let A be an E-affinoid algebra, and let D be an A-locally free $A \otimes_{\mathbf{Q}_p} K$-module equipped with a separated exhaustive decreasing filtration $\mathrm{Fil}^{\bullet} D$ by A-locally free subbundles. Let $\{i_0 < i_2 < \cdots < i_k\}$ be the distinct i such that $\mathrm{gr}^i(D) \neq 0$. The *Hodge polygon* Δ_D of D is the convex polygon in the plane with leftmost endpoint $(0,0)$ and $\mathrm{rk}_{A\otimes_{\mathbf{Q}_p} K}\,\mathrm{gr}^{i_j}(D)$ segments of horizontal distance 1 and slope i_j for $0 \leq j \leq k$. The *Hodge number* t_H of D is the y-coordinate of the rightmost point of Δ_D, i.e., $\sum_j i_j \cdot \mathrm{rk}_{A\otimes_{\mathbf{Q}_p} K}\,\mathrm{gr}^{i_j}(D)$.

Corollary 5.1.13. *Let $\mathscr{V}$ be a rank-d family of* Gal_K*-representations, and let Δ be a convex polygon in the plane with leftmost endpoint* $(0,0)$ *and rightmost endpoint* (d, t_H) *for some $t_H \in \mathbf{Z}$. Then there is a closed immersion $X^{\Delta}_{\mathbf{B}_*} \hookrightarrow X$ such that, for any E-finite Artin local ring B, a map* $\mathrm{Sp}(B) \to X$ *factors through $X^{\Delta}_{\mathbf{B}_*}$ if and only if the induced B-linear Galois representation $\mathscr{V}_B$ is $\mathbf{B}_*$-admissible and the Hodge polygon of* $\mathbf{D}_{\mathrm{dR}}(\mathscr{V}_B)$ *is Δ. In fact, $X^{\Delta}_{\mathbf{B}_*}$ is a union of connected components of $X^{[a,b]}_{\mathbf{B}_*}$, where $a = i_0$ and $b = i_k$. If $X = X^{\Delta}_{\mathbf{B}_*}$, then the Hodge polygon of $\mathscr{D}_{\mathrm{dR}}(\mathscr{V})$ is Δ.*

Proof. We may assume that $X = X^{[a,b]}_{\mathbf{B}_*}$. For each $i \in [a,b]$, we have an exact sequence of vector bundles

$$0 \longrightarrow \mathrm{Fil}^{i+1}\mathbf{D}_{\mathrm{dR}}(\mathscr{V}) \longrightarrow \mathrm{Fil}^{i}\mathbf{D}_{\mathrm{dR}}(\mathscr{V}) \longrightarrow \mathrm{gr}^{i}\mathbf{D}_{\mathrm{dR}}(\mathscr{V}) \longrightarrow 0.$$

These vector bundles have locally constant rank on X (though not necessarily globally constant rank), and their formations commute with base change on X, by Theorem 5.1.2, so we take the union of the connected components where $\mathrm{gr}^i\mathbf{D}_{\mathrm{dR}}(\mathscr{V})$ has the correct dimension. $\square$

We can also stratify the spaces $X^{[a,b]}_{\mathbf{B}_{\mathrm{dR}},d'}$ by the Hodge polygon, although we do not get a decomposition into connected components, because we have no a priori interpretation of the graded pieces of $\mathscr{D}_{\mathrm{dR}}(\mathscr{V})$ when $\mathscr{V}$ is not de Rham.

Corollary 5.1.14. *Let $\mathscr{V}$ be a rank-d family of* Gal_K*-representations, let d' be an integer* $0 \leq d' \leq d$*, and let Δ be a convex polygon in the plane with leftmost endpoint* $(0,0)$ *and rightmost endpoint* (d', t_H) *for some $t_H \in \mathbf{Z}$. Then there is a Zariski-locally closed immersion $X^{\Delta}_{\mathrm{dR}} \hookrightarrow X$ such that, for any E-finite Artin local ring B, a map* $\mathrm{Sp}(B) \to X$ *factors through X^{Δ}_{dR} if and only if* $\mathrm{Fil}^i\mathbf{D}^{[a,b]}_{\mathrm{dR}}(V_B)$ *is projective for all i and* $\mathbf{D}^{[a,b]}_{\mathrm{dR}}(V_B)$ *has Hodge polygon Δ, where $a = i_0$ and $b = i_k$. If $X = X^{\Delta}_{\mathrm{dR}}$, then the Hodge polygon of $\mathscr{D}^{[a,b]}_{\mathrm{dR}}(\mathscr{V})$ is Δ.*

Proof. Let c_j be the x-coordinate of the right endpoint of the j-th segment of Δ (where we count segments starting with 0), and, for $i_j < i \leq i_{j+1}$, let $d_i = \sum_{\ell=j}^{k} c_\ell$. Then we are looking for the locus of E-finite Artin points $x : \mathrm{Sp}(B) \to X$ where for all $a \leq i \leq b$, if $i_j < i \leq i_{j+1}$ then $\mathbf{D}^{[i,b]}_{\mathrm{dR}}(\mathscr{V}_B)$ is a free $B \otimes_{\mathbf{Q}_p} K$-module of rank i.

We use Theorem 5.1.6 to construct the desired locus. Indeed, for every $i \in [a,b]$, Theorem 5.1.6 gives us a Zariski-open subspace $X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\le d_i}$ and a Zariski-closed subspace $X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\ge d_i}$ such that the intersection

$$X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},d_i} := X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\le d_i} \cap X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\ge d_i}$$

represents the condition "$\mathbf{D}^{[i,b]}_{\mathrm{dR}}(\mathscr{V}_B)$ is free of rank d_i". Then we put

$$X^{\Delta}_{\mathrm{dR}} := \Bigl(\bigcap_{i\in[a,b]} X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\le d_i}\Bigr) \cap \Bigl(\bigcap_{i\in[a,b]} X^{[i,b]}_{\mathbf{B}_{\mathrm{dR}},\ge d_i}\Bigr).$$

The second claim follows similarly, by repeated application of Theorem 5.1.7. □

Remark 5.1.15. It is possible to define an ordering on Hodge polygons so that the spaces X^{Δ}_{dR} yield a Zariski-stratification of $X^{[a,b]}_{\mathrm{dR}}$. This is done in [Shah 2013, §3].

We can also show that, when restricted to the category of $\mathbf{B}_*$-admissible representations, the functor $\mathscr{D}^K_{\mathbf{B}_*}$ is well-behaved with respect to exact sequences, tensors, and duals.

Corollary 5.1.16. (1) *Let* $\mathrm{Rep}^{\mathbf{B}_*}_X(\mathrm{Gal}_K)$ *be the category of* $\mathbf{B}_*$*-admissible families of representations of* Gal_K *over* X. *Then* $\mathscr{D}_{\mathbf{B}_*} : \mathrm{Rep}^{\mathbf{B}_*}_X(\mathrm{Gal}_K) \to \mathrm{Proj}_{\mathscr{O}_X\otimes_{\mathbf{Q}_p}\mathbf{B}_*^{\mathrm{Gal}_K}}$ *is exact and faithful, where* $\mathrm{Proj}_{\mathscr{O}_X\otimes_{\mathbf{Q}_p}\mathbf{B}_*^{\mathrm{Gal}_K}}$ *is the category of sheaves of projective* $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$*-modules, and any quotient or subrepresentation of a* $\mathbf{B}$*-admissible family of representations is itself* $\mathbf{B}$*-admissible.*

(2) *The subcategory* $\mathrm{Rep}^{\mathbf{B}_*}_X(\mathrm{Gal}_K)$ *is stable under formation of tensor products and duals, and the functor* $\mathscr{V} \mapsto \mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ *commutes with these operations when restricted to* $\mathrm{Rep}^{\mathbf{B}_*}_X(\mathrm{Gal}_K)$.

(3) *If* $\mathbf{B}_* = \mathbf{B}_{\mathrm{HT}}$ (*resp.* $\mathbf{B}_{\mathrm{dR}}$), *then the grading* (*resp. filtration*) *on* $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ *is also exact and tensor compatible.*

Proof. These statements all follow from the corresponding statements with artinian coefficients, because for $\mathscr{V}$ a $\mathbf{B}_*$-admissible family of representations of Gal_K, the formation of $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ commutes with arbitrary base change on X. Since $\mathscr{D}^K_{\mathbf{B}_*}(\mathscr{V})$ is a finite $\mathscr{O}_X \otimes_{\mathbf{Q}_p} \mathbf{B}_*^{\mathrm{Gal}_K}$-module, we can check isomorphisms on thickenings of closed points of X. □

Finally, we can use the existence of the closed subspace $X^{[a,b]}_{\mathbf{B}_*} \hookrightarrow X$ to deduce $\mathbf{B}_*$-admissibility on a Zariski-open neighborhood of $x \in X$ from $\mathbf{B}_*$-admissibility on infinitesimal neighborhoods of x.

Corollary 5.1.17. *Let* V *be a continuous* $\mathscr{O}_X$*-linear representation of* Gal_K *as above, and suppose* $x \in X$ *is a point such that* $V \otimes_{\mathscr{O}_X} \mathscr{O}_{X,x}/\mathfrak{m}_x^n$ *is* $\mathbf{B}_*$*-admissible with Hodge–Tate weights in* $[a,b]$ *for all* $n \ge 0$. *Then there is a Zariski-open neighborhood* U *of* x *such that* $V|_U$ *is* $\mathbf{B}_*$*-admissible with Hodge–Tate weights in* $[a,b]$.

Proof. We may assume that $X = \mathrm{Sp}(A)$ for some E-affinoid algebra A, so that $X_{\mathbf{B}_*}^{[a,b]} = \mathrm{Sp}(A_{\mathbf{B}_*}^{[a,b]})$ for some quotient $A \twoheadrightarrow A_{\mathbf{B}_*}^{[a,b]}$. The assumption on the infinitesimal neighborhoods of x implies that the natural map $A \to A_{\mathfrak{m}_x}^{\wedge}$ factors through $A_{\mathbf{B}_*}^{[a,b]}$. This implies that the complete local rings of A and $A_{\mathbf{B}_*}^{[a,b]}$ at $\mathfrak{m}_x$ are the same, which in turn implies that $\mathrm{Sp}(A_{\mathbf{B}_*}^{[a,b]})$ contains an Zariski-open neighborhood of x. □

Remark 5.1.18. In fact, since $X_{\mathbf{B}_*}^{[a,b]}$ is a closed subspace containing an admissible open neighborhood of x, if $X = \mathrm{Sp}(A)$ for some E-affinoid algebra A, then $X_{\mathbf{B}_*}^{[a,b]}$ contains all irreducible components of $\mathrm{Sp}(A)$ passing through x.

5.2. ***Hodge–Tate and de Rham loci.*** As above, we let X be a quasiseparated E-analytic space and we let $\mathscr{V}$ be a finite locally free $\mathscr{O}_X$-module of rank d, equipped with a continuous $\mathscr{O}_X$-linear action of Gal_K.

In order to construct quotients $\mathscr{O}_X \twoheadrightarrow \mathscr{O}_{X_{\mathrm{HT}}^{[a,b]}}$ and $\mathscr{O}_X \twoheadrightarrow \mathscr{O}_{X_{\mathrm{dR}}^{[a,b]}}$ for Theorem 5.1.1 for $\mathbf{B}_* = \mathbf{B}_{\mathrm{HT}}$ or $\mathbf{B}_{\mathrm{dR}}$, we work locally on X and construct a suitable coherent ideal sheaf. Similarly, we work locally on X to construct the Zariski-locally closed immersions $X_{\mathrm{HT},d'}^{[a,b]} \hookrightarrow X$ and $X_{\mathrm{dR},d'}^{[a,b]} \hookrightarrow X$ required by Theorem 5.1.6.

In order to prove Theorems 5.1.2 and 5.1.7 for $\mathbf{B}_* = \mathbf{B}_{\mathrm{HT}}$ and $\mathbf{B}_* = \mathbf{B}_{\mathrm{dR}}$, it likewise suffices to work locally on X. Thus, we may reduce to the case where $X = \mathrm{Sp}(A)$ for A an E-affinoid algebra and $V := \Gamma(X, \mathscr{V})$ is a finite free A-module of rank d equipped with a continuous A-linear action of Gal_K which admits a free Gal_K-stable $\mathscr{A}$-lattice V_0 for some formal $\mathscr{O}_E$-model $\mathscr{A}$ of A.

Before we begin, we prove a useful lemma.

Lemma 5.2.1. *Let R be an Artin ring, let M be a free R-module of rank r equipped with an endomorphism $T : M \to M$, and suppose that*

$$0 \longrightarrow M' \longrightarrow M \xrightarrow{T} M \longrightarrow M'' \longrightarrow 0$$

is exact. Then M' is free of rank d over R if and only if M'' is.

Proof. Since R is an Artin ring, it is semilocal. The assertions "M' is free of rank d" and "M'' is free of rank d" can each be checked by passing to local factors of R, so we may assume that R is a local ring with maximal ideal $\mathfrak{m}$.

It suffices to show that for an exact sequence $0 \longrightarrow M' \longrightarrow M \longrightarrow N \longrightarrow 0$ of R-modules (with M free of rank r), M' is free of rank d if and only if N is free of rank $r-d$.

If N is free of rank $r-d$, it is projective, so the exact sequence splits and M' is free of rank d (as R is local).

Conversely, suppose that M' is free of rank d. We will prove that M/M' is free of rank $r-d$ by induction on d. If $d = 0$, there is nothing to prove. So suppose that M' is free of rank d, and suppose we know the result for submodules of rank

$d-1$. Choose bases $\{e_1,\dots,e_d\}$ and $\{f_1,\dots,f_r\}$ for M' and M, respectively, and consider the image $a_1 f_1+\cdots+a_r f_r$ of e_1. At least one of the a_i is a unit in R, because otherwise, by injectivity of $M'\to M$, the element $e_1\in M'$ would be killed by $\mathrm{ann}_R(\mathfrak{m})\neq 0$ (which is impossible, as e_1 is part of a basis of M'). Thus, without loss of generality, we assume that a_1 is a unit. Then $\{e_1,f_2,\dots,f_r\}$ is a basis of M, and

$$0\longrightarrow M'/\langle e_1\rangle\longrightarrow M/\langle e_1\rangle\longrightarrow N\longrightarrow 0$$

is exact, so by the inductive hypothesis N is free of rank $(r-1)-(d-1)=r-d$. □

5.2.1. *Hodge–Tate locus.* Let L/K be a finite extension such that Gal_L acts trivially on $V_0/12pV_0$. Then for $n\gg 0$, $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ is finite free of rank d over $A\otimes_{\mathbf{Q}_p}L_n$ and carries a linear action of Γ_{L_n}. If necessary, we increase n so that Γ_{L_n} is procyclic, with topological generator γ. Moreover, the formation of $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$ commutes with arbitrary E-affinoid base change on A and $(\mathbf{D}_{\mathrm{Sen}}^{L_n}(V))^{\Gamma_{L_n}}=((\mathbf{C}_K\otimes A)\otimes_A V)^{\mathrm{Gal}_{L_n}}$. As a consequence,

$$\mathbf{D}_{\mathrm{HT}}^{L_n}(V)=\Big(\bigoplus_k t^k\cdot\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)\Big)^{\Gamma_{L_n}=1}.$$

Recall that we have defined $\mathbf{D}_{\mathrm{HT}}^{K,[a,b]}(V)$ to be $\big(\bigoplus_{i\in[a,b]}\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)\cdot t^i\big)^{\Gamma_K}$.

Theorem 5.2.2. *Let V be as above. Then, for every $0\le d'\le d$, there is a Zariski-locally closed immersion $X_{\mathrm{HT,d'}}^{[a,b]}\hookrightarrow X$ such that, for any E-finite Artin local ring B and $V_B:=V\otimes_A B$, $x:\mathrm{Sp}(B)\to X$ factors through $X_{\mathrm{HT,d'}}^{[a,b]}$ if and only if $\mathbf{D}_{\mathrm{HT}}^{K,[a,b]}(V_B)$ is a free $B\otimes_{\mathbf{Q}_p}K$-module of rank d'. In fact, $X_{\mathrm{HT,d'}}^{[a,b]}=X_{\mathrm{HT,\le d'}}^{[a,b]}\cap X_{\mathrm{HT,\ge d'}}^{[a,b]}$, where $X_{\mathrm{HT,\le d'}}^{[a,b]}\subset X$ is Zariski-open and $X_{\mathrm{HT,\ge d'}}^{[a,b]}\hookrightarrow X$ is Zariski-closed. If $d'=d$, there is a quotient $A\twoheadrightarrow A_{\mathrm{HT}}^{[a,b]}$ such that an E-finite artinian point $x:A\to B$ factors through $A_{\mathrm{HT}}^{[a,b]}$ if and only if V_B is Hodge–Tate with Hodge–Tate weights in the interval $[a,b]$.*

Proof. First, we note that because $\mathbf{D}_{\mathrm{HT}}^{L_n}(V)=L_n\otimes\mathbf{D}_{\mathrm{HT}}^{K}(V)$, it is enough consider V_B as a representation of Gal_{L_n}.

Next let $M=\bigoplus_{i=a}^b\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)\cdot t^i$. We note that $\mathrm{H}^0(\Gamma_{L_n},M\otimes_A B)$ is free over $B\otimes_{\mathbf{Q}_p}L_n$ of rank d' if and only if $\mathrm{H}^1(\Gamma_{L_n},M\otimes_A B)$ is, because the continuous Γ_{L_n}-cohomology of $M\otimes_A B$ is computed by the complex

$$0\longrightarrow M\otimes_A B\xrightarrow{\gamma-1}M\otimes_A B\longrightarrow 0,$$

and we can apply Lemma 5.2.1. Further, the formation of H^1 commutes with arbitrary base change on A, by Corollary 3.1.3. Since $\mathrm{H}^1(\Gamma_{L_n},M)$ is a coherent $A\otimes_{\mathbf{Q}_p}L_n$-module and

$$M\otimes_A B\xrightarrow{\gamma-1}M\otimes_A B\longrightarrow\mathrm{H}^1(\Gamma_{L_n},M\otimes_A B)\longrightarrow 0$$

is a finite presentation of $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$, it follows that

$$M \xrightarrow{T=\gamma-1} M \longrightarrow \mathrm{H}^1(\Gamma_{L_n}, M) \longrightarrow 0$$

is a finite presentation of $\mathrm{H}^1(\Gamma_{L_n}, M)$.

We use the theory of Fitting ideals to cut out the locus where $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ is free of rank d' over $B\otimes_{\mathbf{Q}_p} L_n$. By Proposition 20.8 of [Eisenbud 1995], $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ is projective of rank d' over $B\otimes_{\mathbf{Q}_p} L_n$ if and only if

$$\mathrm{Fitt}_{d'}(\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)) = B\otimes_{\mathbf{Q}_p} L_n \quad \text{and} \quad \mathrm{Fitt}_{d'-1}(\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)) = 0.$$

The latter is a closed condition defined by the $(d'-1)\times(d'-1)$-minors of T. The former is a Zariski-open condition defined by inverting the $d'\times d'$-minors of T.

If $d'=d$, we claim the open condition can be ignored on the complement of the zero locus of ideal generated by the $(d-1)\times(d-1)$-minors of T. First, we note that $\mathrm{Fitt}_d(\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)) = B\otimes_{\mathbf{Q}_p} L_n$ if and only if $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ can be generated by d elements, by [Eisenbud 1995, Proposition 20.6]. More precisely, $B\otimes_{\mathbf{Q}_p} L_n$ is a semilocal ring, while that proposition applies to modules over local rings. But $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ can be generated by d elements over $B\otimes_{\mathbf{Q}_p} L_n$ if and only if the same is true after passing to idempotent factors of $B\otimes_{\mathbf{Q}_p} L_n$. Similarly, since the formation of Fitting ideals commutes with base change, we can check that $\mathrm{Fitt}_d(\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)) = (1)$ after passing to idempotent factors of $B\otimes_{\mathbf{Q}_p} L_n$.

Moreover, the formation of $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ commutes with base change on B, so by Nakayama's lemma $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B)$ can be generated by lifts of generators of $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$. But if the Fitting ideal $\mathrm{Fitt}_{d-1}(\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B))$ vanishes, then $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ cannot be generated by $d-1$ elements at any point of $B/\mathfrak{m}_B\otimes_{\mathbf{Q}_p} L_n$. Therefore, the $\mathbf{Q}_p$-dimension of $\mathrm{H}^0(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ (which is the same as the $\mathbf{Q}_p$-dimension of $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$) is at least $d\cdot\dim_{\mathbf{Q}_p} B\cdot\dim_{\mathbf{Q}_p} L_n$. But then the formalism of admissible representations implies that the $\mathbf{Q}_p$-dimension of $\mathrm{H}^0(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ is exactly $d\cdot\dim_{\mathbf{Q}_p} B\cdot\dim_{\mathbf{Q}_p} L_n$, and Proposition 4.1.3 implies that $\mathrm{H}^0(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ is a free $B/\mathfrak{m}_B\otimes_{\mathbf{Q}_p} L_n$-module of rank d, so $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ is as well, by Lemma 5.2.1. Clearly this implies that $\mathrm{H}^1(\Gamma_{L_n}, M\otimes_A B/\mathfrak{m}_B)$ can be generated by d elements.

Thus, the condition that V_B be Hodge–Tate with Hodge–Tate weights in the appropriate range is cut out by the $(d-1)\times(d-1)$-minors of T. Since T has coefficients in $A\otimes_{\mathbf{Q}_p} L_n$ and $E\otimes_{\mathbf{Q}_p} L_n$ is finite free over E, we obtain the quotient of A we sought. □

We turn to the converse question:

Theorem 5.2.3. *Let V be as above. Assume that, for every E-finite artinian point $x : A \to B$, the $B \otimes_{\mathbf{Q}_p} K$-module $\mathbf{D}_{\mathrm{HT}}^{K,[a,b]}(V_x)$ is free of rank d', where $0 \leq d' \leq d$.*

(1) *$\mathbf{D}_{\mathrm{HT}}^{K,[a,b]}(V)$ is a rank-d' locally free $A \otimes_{\mathbf{Q}_p} K$-module, and the $(t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{K_n}(V))^{\Gamma_K}$ are locally free A-modules.*

(2) *the formation of $\mathbf{D}_{\mathrm{HT}}^{K,[a,b]}(V)$ commutes with arbitrary base change $f : A \to A'$.*

If $d' = d$, then:

(3) *$\mathbf{D}_{\mathrm{HT}}^{K}(V) = (\bigoplus_{k=a}^{b} t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{K_n}(V))^{\Gamma_K}$, and the formation of $\mathbf{D}_{\mathrm{HT}}^{K}(V)$ commutes with arbitrary base change $f : A \to A'$.*

(4) *The natural map $\alpha_V : (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{HT}}) \otimes_{A \otimes_{\mathbf{Q}_p} K} \mathbf{D}_{\mathrm{HT}}^{K}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{HT}}) \otimes_A V$ is a Galois-equivariant isomorphism of graded $A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{HT}}$-modules, and so V is Hodge–Tate.*

Proof. It is enough to consider V as a representation of Gal_{L_n}, since $\mathbf{D}_{\mathrm{HT}}^{L_n}(V) = L_n \otimes \mathbf{D}_{\mathrm{HT}}^{K}(V)$. Since $\mathbf{D}_{\mathrm{Sen}}^{K_n}(V) = (\mathbf{D}_{\mathrm{Sen}}^{L_n}(V))^{H_K}$, the decomposition of $\mathbf{D}_{\mathrm{HT}}^{K}(V)$ follows from the decomposition of $\mathbf{D}_{\mathrm{HT}}^{L_n}(V)$.

We will use the base change spectral sequence of Theorem 3.1.1 for the continuous Γ_{L_n}-cohomology of $M := \bigoplus_{i=a}^{b} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \cdot t^i$.

By assumption, $\mathrm{H}^1(\Gamma_{L_n}, M \otimes_A B)$ is a free $B \otimes_{\mathbf{Q}_p} L_n$-module of the same rank as $\mathrm{H}^0(\Gamma_{L_n}, M \otimes_A B)$. Since M is A-finite, Corollary 3.1.3 implies that the formation of $\mathrm{H}^1(\Gamma_{L_n}, M)$ commutes with base change $A \to B$. It follows that $\mathrm{H}^1(\Gamma_{L_n}, M)$ is a projective A-module and $\mathrm{Tor}_{-p}^{A}(\mathrm{H}^1(\Gamma_{L_n}, M), A')$ vanishes for all homomorphisms $f : A \to A'$ and all $p < 0$.

But if we consider the low-degree exact sequence of Corollary 3.1.3, the vanishing of the Tor terms shows that formation of $\mathrm{H}^0(\Gamma_{L_n}, M)$ commutes with arbitrary base change $A \to A'$, and in particular with the base change $x : A \to B$. Thus,

$$M^{\Gamma_{L_n}} \subset \mathbf{D}_{\mathrm{HT}}^{L_n}(V) = \left(\bigoplus_{i \in \mathbf{Z}} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \cdot t^i\right)^{\Gamma_{L_n}}$$

is a locally free $A \otimes_{\mathbf{Q}_p} L_n$-module of rank d'. This proves the first two parts.

Now we assume that $d' = d$. We claim that $M^{\Gamma_{L_n}} = \mathbf{D}_{\mathrm{HT}}^{L_n}(V)$. Suppose to the contrary that there is some nonzero $y \in (\mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \cdot t^i)^{\Gamma_{L_n}}$ for some $i \notin [0, h]$. Then there is some E-finite artinian point $x : A \to B$ such that y is nonzero in $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V_x) \cdot t^i$. But $M_x^{\Gamma_{L_n}} = \mathbf{D}_{\mathrm{HT}}^{L_n}(V_x)$ because $M_x^{\Gamma_{L_n}} \subset \mathbf{D}_{\mathrm{HT}}^{L_n}(V_x)$ and the L_n-dimensions of the two sides agree, contradicting the assumed Γ_{L_n}-invariance of y.

Let $f : A \to A'$ be a morphism of E-affinoid algebras. We have already seen that the formation of $\mathrm{H}^0(\Gamma_{L_n}, M)$ commutes with arbitrary affinoid base change on A, so $\mathrm{H}^0(\Gamma_{L_n}, M \otimes_A A') = \mathrm{H}^0(\Gamma_{L_n}, M) \otimes_A A'$. If $\mathbf{D}_{\mathrm{HT}}^{L_n}(V \otimes_A A') = \mathrm{H}^0(\Gamma_{L_n}, M \otimes_A A')$, we are done. But for any E-finite artinian point $x : A' \to B'$, the induced B'-linear representation $(x \circ f)^* V$ is Hodge–Tate with Hodge–Tate weights in $[a,b]$. Then we have just seen that $\mathbf{D}_{\mathrm{HT}}^{L_n}(V \otimes_A A') = \left(\bigoplus_{k=a}^{b} t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{K_n}(V \otimes_A A')\right)^{\Gamma_{L_n}}$, as desired.

Finally, we show that $(A \mathbin{\hat{\otimes}} \mathbf{B}_{\mathrm{HT}}) \otimes_{A \otimes_{\mathbf{Q}_p} K} \mathbf{D}_{\mathrm{HT}}^{K}(V) \to (A \mathbin{\hat{\otimes}} \mathbf{B}_{\mathrm{HT}}) \otimes_A V$ is a Galois-equivariant isomorphism. Since the natural map

$$(A \mathbin{\hat{\otimes}} \mathbf{C}_K) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \to (A \mathbin{\hat{\otimes}} \mathbf{C}_K) \otimes_A V$$

is a Galois-equivariant isomorphism, it suffices to show that the natural map

$$(A \otimes_{\mathbf{Q}_p} L_n[t, t^{-1}]) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{HT}}^{L_n}(V) \to \bigoplus_{i \in \mathbf{Z}} \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \cdot t^i$$

is a Galois-equivariant isomorphism of graded $A \otimes_{\mathbf{Q}_p} L_n$-modules. We may further reduce to checking that the natural map

$$\mathrm{gr}^i\left((A \otimes_{\mathbf{Q}_p} L_n[t, t^{-1}]) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{HT}}^{L_n}(V)\right) \to \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \cdot t^i$$

is a Galois-equivariant isomorphism of $A \otimes_{\mathbf{Q}_p} L_n$-modules for all i.

Now we have a map of A-finite modules, so it suffices to check our desired isomorphism modulo powers of maximal ideals of A. But when A is an E-finite Artin ring, this follows from the formalism of admissible representations, so we are done. □

5.2.2. *de Rham locus.* To treat the de Rham case, we work with $\mathbf{D}_{\mathrm{dif}}^{L_n}(V)$ instead of $\mathbf{D}_{\mathrm{Sen}}^{L_n}(V)$. Recall that we have defined

$$\mathbf{D}_{\mathrm{dR}}^{K,[a,b]}(V) = ((A \otimes_{\mathbf{Q}_p} t^a \mathbf{B}_{\mathrm{dR}}/t^b) \otimes_A V)^{\Gamma_K} = (t^a \mathbf{D}_{\mathrm{dif}}^{K,+}(V)/t^b \mathbf{D}_{\mathrm{dif}}^{K,+}(V))^{\Gamma_K}.$$

Theorem 5.2.4. *Let V be as above. Then, for every $0 \le d' \le d$, there is a Zariski-locally closed immersion $X_{\mathrm{dR},d'}^{[a,b]} \hookrightarrow X$ such that, for any E-finite Artin local ring B and $V_B := V \otimes_A B$, $x : \mathrm{Sp}(B) \to X$ factors through $X_{\mathrm{dR},d'}^{[a,b]}$ if and only if $\mathbf{D}_{\mathrm{dR}}^{K,[a,b]}(V)$ is a free $B \otimes_{\mathbf{Q}_p} K$-module of rank d'. In fact,*

$$X_{\mathrm{dR},d'}^{[a,b]} = X_{\mathrm{dR},\le d'}^{[a,b]} \cap X_{\mathrm{dR},\ge d'}^{[a,b]},$$

where $X_{\mathrm{dR},\le d'}^{[a,b]} \subset X$ is Zariski-open and $X_{\mathrm{dR},\ge d'}^{[a,b]} \hookrightarrow X$ is Zariski-closed. If $d' = d$, there is a quotient $A \twoheadrightarrow A_{\mathrm{dR}}^{[a,b]}$ such that an E-finite artinian point $x : A \to B$ factors through $A_{\mathrm{dR}}^{[a,b]}$ if and only if V_B is de Rham with Hodge–Tate weights in the interval $[a,b]$.

Proof. Because $\mathbf{D}_{\mathrm{dR}}^{L_n}(V) = L_n \otimes_K \mathbf{D}_{\mathrm{dR}}^{K}(V)$, it is enough to cut out the locus where V_B is de Rham as a representation of Gal_{L_n} (with weights in the appropriate range). Then the proof of Theorem 5.2.2 carries over verbatim with M redefined as $M = t^a \mathbf{D}_{\mathrm{dif}}^{L_n}(V)/t^b$, and we obtain the desired result. □

Now we treat the converse question:

Theorem 5.2.5. *Let V be as above. Assume that for every E-finite artinian point $x : A \to B$ the $B \otimes_{\mathbf{Q}_p} K$-module $\mathbf{D}_{\mathrm{dR}}^{K,[a,b]}(V_x)$ is free of rank d', where $0 \le d' \le d$.*

(1) *$\mathbf{D}_{\mathrm{dR}}^{K,[a,b]}(V)$ is a locally free $A \otimes_{\mathbf{Q}_p} K$-module of rank d'.*

(2) *The formation of $\mathbf{D}_{\mathrm{dR}}^{K,[a,b]}(V)$ commutes with arbitrary base change $f : A \to A'$.*

If $d' = d$, then:

(3) *$\mathbf{D}_{\mathrm{dR}}^{K}(V) = \mathbf{D}_{\mathrm{dR}}^{K,[a,b]}$, and the formation of $\mathbf{D}_{\mathrm{dR}}^{K}(V)$ commutes with arbitrary base change $f : A \to A'$.*

(4) *The natural map $\alpha_V : (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_{A \otimes_{\mathbf{Q}_p} K} \mathbf{D}_{\mathrm{dR}}^{K}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_A V$ is a Galois-equivariant isomorphism of graded $A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}$-modules, and so V is de Rham.*

Proof. It is enough to consider V as a representation of Gal_{L_n}, since $\mathbf{D}_{\mathrm{dR}}^{L_n}(V) = L_n \otimes_K \mathbf{D}_{\mathrm{dR}}^{K}(V)$. Then, by the same arguments as in the proof of Theorem 5.2.3 applied with $M = t^a \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^b$, we see that $M^{\Gamma_{L_n}}$ is a locally free $A \otimes_{\mathbf{Q}_p} L_n$-module of rank d'.

Now we assume $d' = d$. We claim that $M^{\Gamma_{L_n}} = \mathbf{D}_{\mathrm{dR}}^{L_n,+}(V) = \mathbf{D}_{\mathrm{dR}}^{L_n}(V)$. Since V is Hodge–Tate with Hodge–Tate weights in $[0,h]$, $(t^k \mathbf{D}_{\mathrm{Sen}}^{L_n}(V))^{\Gamma_{L_n}} = 0$ for $k \notin [0,h]$. Moreover, $\mathrm{H}^1(\Gamma_{L_n}, t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{L_n}(V \otimes_A B)) = 0$ for any artinian point $A \to B$ and $k \notin [0,h]$, so $\mathrm{H}^1(\Gamma_{L_n}, t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{L_n}(V)) = 0$ for $k \notin [0,h]$. Then for any $k > h$, the long exact sequence associated to the short exact sequence of Γ_{L_n}-modules

$$0 \longrightarrow t^k \cdot \mathbf{D}_{\mathrm{Sen}}^{L_n}(V) \longrightarrow \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^{k+1} \longrightarrow \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^k \longrightarrow 0$$

shows that $(\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^{k+1})^{\Gamma_{L_n}} \to (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^k)^{\Gamma_{L_n}}$ is an isomorphism. It follows that

$$(\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V))^{\Gamma_{L_n}} = \varprojlim_k (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^k)^{\Gamma_{L_n}} = (\mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^{h+1})^{\Gamma_{L_n}}.$$

Let $f : A \to A'$ be a morphism of E-affinoid algebras. Then we have seen that the formation of $\mathrm{H}^0(\Gamma_{L_n}, M)$ commutes with arbitrary base change on A, so $\mathrm{H}^0(\Gamma_{L_n}, M \otimes_A A') = \mathrm{H}^0(\Gamma_{L_n}, M) \otimes_A A'$. If $\mathbf{D}_{\mathrm{dR}}^{L_n}(V \otimes_A A') = \mathrm{H}^0(\Gamma_{L_n}, M \otimes_A A')$, we are done.

But, for any E-finite artinian point $x : A' \to B'$, the induced B'-linear representation $(x \circ f)^* V$ is de Rham with Hodge–Tate weights in $[a,b]$. Then we have just seen that

$$\mathbf{D}_{\mathrm{dR}}^{L_n}(V \otimes_A A') = (t^a \mathbf{D}_{\mathrm{dif}}^{K_n,+}(V \otimes_A A')/t^b)^{\Gamma_{L_n}},$$

as desired.

Finally, we show that V is de Rham. Recall that the natural map

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}^+) \otimes_{A \widehat{\otimes} L_n[\![t]\!]} \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}^+) \otimes_A V$$

is a Galois-equivariant isomorphism respecting the filtration on each side. Therefore, it suffices to show that the natural map

$$\mathrm{Fil}^0\big((A \mathbin{\widehat{\otimes}} L_n((t))) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{dR}}^{L_n,+}(V)\big) \to \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)$$

is a Galois-equivariant isomorphism respecting the filtrations. For this, it further suffices to show that, for every $k \geq 0$, the natural map

$$\mathrm{Fil}^0\big((A \mathbin{\widehat{\otimes}} L_n((t))) \otimes_{A \otimes_{\mathbf{Q}_p} L_n} \mathbf{D}_{\mathrm{dR}}^{L_n,+}(V)\big)/t^k \to \mathbf{D}_{\mathrm{dif}}^{L_n,+}(V)/t^k$$

is a Galois-equivariant isomorphism respecting the filtrations.

Now we have a morphism of A-finite modules, so it suffices to check this modulo every power of every maximal ideal of A. But when A is an E-finite Artin ring, this follows from the formalism of admissible representations, so we are done. □

5.3. ***Semistable and crystalline loci.*** We will produce similar quotients of A parametrizing the semistable and crystalline loci. However, our method of proof, which follows the argument in [Berger and Colmez 2008], is quite different. Instead of using cohomology and base change arguments, we will prove that "de Rham implies uniformly potentially semistable", and then deduce our desired results.

5.3.1. *Uniform potential semistability.* Suppose that $A = \mathbf{Q}_p$ and $\rho : \mathrm{Gal}_K \to \mathrm{GL}(V)$ is a de Rham representation of Gal_K. Then there is a finite extension L/K such that $\rho|_{\mathrm{Gal}_L}$ is a semistable representation. This is known as the p-adic local monodromy theorem, and it follows from a theorem of Berger, combined with Crew's conjecture, which was proved separately by André [2002], Mebkhout [2002], and Kedlaya [2004]. Berger [2002] then associated to any de Rham p-adic Galois representation a p-adic differential equation and characterized semistability of the representation in terms of unipotence of the associated differential equation. Crew's conjecture states that any p-adic differential equation becomes semistable over a finite extension.

Neither Berger's construction nor Crew's conjecture work naively when A is an affinoid algebra, so we cannot proceed directly. However, both pieces are known when the coefficients are a general complete discretely valued field with perfect residue field of characteristic $p > 0$.

The version of the p-adic monodromy theorem we need is the following:

Theorem 5.3.1 [Berger and Colmez 2008, Corollaire 6.2.5]. *Let B be a complete discretely valued field with perfect residue field of characteristic $p > 0$, and let V be a B-representation of Gal_K of dimension d which is de Rham. Then there is a finite extension L/K such that the $B \mathbin{\widehat{\otimes}} \widehat{\mathbf{Q}}_p^{\mathrm{nr}}$-module $((B \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_B V)^{I_L}$ is free of rank d and the map*

$$L \otimes_{L_0} ((B \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_B V)^{I_L} \to ((B \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_B V)^{I_L}$$

is an isomorphism.

We return to our general setup. Let A be an E-affinoid algebra, and let V be a finite free A-module equipped with a continuous A-linear action of Gal_K. We will embed A isometrically into a finite product $B = \prod_i B_i$ of Artin local rings B_i, and apply Theorem 5.3.1 to $V \otimes_A B$.

However, we need to be able to compare the semistability of V (an A-linear representation of Gal_K) and the semistability of $V_B = \prod_{i=1}^r V_{B_i}$ (as a B-linear representation of Gal_K).

Recall that we defined $\mathbf{B}_{\mathrm{st}}^+$ to be $\mathbf{B}_{\max}^+[\log([\pi])]$ rather than the usual semistable period ring. We further define $\mathbf{B}_{\mathrm{st}}^{+,h}$ to be $\bigoplus_{i=0}^h \mathbf{B}_{\max}^+ \log([\pi])^i$, so that $\mathbf{B}_{\mathrm{st}}^{+,h}$ is the kernel of N^{h+1} on $\mathbf{B}_{\mathrm{st}}^+$ (here N is the monodromy operator).

Further, by Remark A.1.5, there is an isomorphism of K-Fréchet spaces $\mathbf{B}_{\mathrm{dR}}^+ \xrightarrow{\sim} \mathbf{C}_K[\![T]\!]$ which defines compatible isomorphisms

$$\mathbf{B}_{\max}^+ \xrightarrow{\sim} \mathbf{C}_K\langle T\rangle \quad \text{and} \quad \mathbf{B}_{\mathrm{st}}^+ \xrightarrow{\sim} \mathbf{C}_K\langle T\rangle[\log(1+T)]$$

(as well as $\mathbf{B}_{\mathrm{st}}^{+,h} \xrightarrow{\sim} \bigoplus_{i=0}^h \mathbf{C}_K\langle T\rangle \log(1+T)^i$).

Proposition 5.3.2. *Let A be an E-affinoid algebra, and let $x : A \to B$ be a closed embedding of Banach algebras. Then if $a \in A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}^+$ and $x(a) \in B \mathbin{\widehat{\otimes}} (L \otimes_{L_0} \mathbf{B}_{\mathrm{st}}^{+,h})$, a is actually in $A \mathbin{\widehat{\otimes}} (L \otimes_{L_0} \mathbf{B}_{\mathrm{st}}^{+,h})$.*

This follows as in [Berger and Colmez 2008, Lemme 6.3.1].

Combined with the p-adic local monodromy theorem, we have the following:

Theorem 5.3.3. *Let A be an E-affinoid algebra, V an A-linear representation of Gal_K on a finite free A-module of rank d, and $[a, b]$ an interval such that V_x is a de Rham representation of Gal_K with Hodge–Tate weights in $[a, b]$ for every E-finite artinian point x of A. Then V is potentially semistable: there is a finite Galois extension L/K such that the $A \otimes_{\mathbf{Q}_p} L_0$-module $\mathbf{D}_{\mathrm{st}}^L(V)$ is locally free of rank d and*

$$(A \otimes_{\mathbf{Q}_p} L) \otimes_{A \otimes_{\mathbf{Q}_p} L_0} \mathbf{D}_{\mathrm{st}}^L(V) = \mathbf{D}_{\mathrm{dR}}^L(V).$$

In addition, for any homomorphism of E-affinoid algebras $A \to A'$, the natural map $A' \otimes_A \mathbf{D}_{\mathrm{st}}^L(V) \to \mathbf{D}_{\mathrm{st}}^L(V \otimes_A A')$ is an isomorphism.

When A is reduced, this is [Berger and Colmez 2008, Théorème 6.3.2].

Proof. We first apply Lemma 2.1.2 to find a closed embedding $A \to \prod_i B_i$ into a finite product of Artin rings. Here B_i is an E_i-finite algebra, where E_i is a complete discretely valued field with perfect residue field of characteristic p (and B_i is topologized as a finite-dimensional E_i-vector space).

Now we can apply the p-adic monodromy theorem to the representations V_{B_i}, because each of them can be viewed as a finite-dimensional E_i-representation. In other words, there is a finite extension L/K such that for each i the natural map

$$L \otimes_{L_0} ((B_i \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_{B_i} V_{B_i})^{I_L} \to ((B_i \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_{B_i} V_{R_i})^{I_L}$$

is an isomorphism. Theorem 5.3.1 only produces an isomorphism of $E_i \otimes_{\mathbf{Q}_p} L$-modules, but the natural map respects the B_i-linear structure on each side, and there is no kernel or cokernel (because it is an isomorphism of $E_i \otimes_{\mathbf{Q}_p} L$-modules), so it is actually an isomorphism of underlying $B_i \otimes_{\mathbf{Q}_p} L$-modules.

But we know by Theorem 5.1.2 applied to $\mathbf{B}_* = \mathbf{B}_{\mathrm{dR}}$ that $\mathbf{D}^L_{\mathrm{dR}}(V)$ is a locally free $A \otimes_{\mathbf{Q}_p} L$-module of rank d. We have an injective map

$$\mathbf{D}^L_{\mathrm{dR}}(V) \to ((B \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}) \otimes_A V)^{I_L}$$

with $B = \prod B_i$, and, for $y \in \mathbf{D}^L_{\mathrm{dR}}(V)$, we can write $y = \sum_{j=1}^d y_j \otimes v_j$ with $y_j \in A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{dR}}$ and $\{v_j\}$ an A-basis for V. By the isomorphism above, each of the y_j lies in $B \mathbin{\widehat{\otimes}} (L \otimes_{L_0} \mathbf{B}_{\mathrm{st}})$. Then by Proposition 5.3.2, each lies in $A \mathbin{\widehat{\otimes}} (L \otimes_{L_0} \mathbf{B}_{\mathrm{st}})$, so

$$\mathbf{D}^L_{\mathrm{dR}}(V) = (A \mathbin{\widehat{\otimes}} (L \otimes_{L_0} \mathbf{B}_{\mathrm{st}}) \otimes_A V)^{\mathrm{Gal}_L} = L \otimes_{L_0} \mathbf{D}^L_{\mathrm{st}}(V),$$

and hence $\mathbf{D}^L_{\mathrm{st}}(V)$ is locally free of rank d over $A \otimes_{\mathbf{Q}_p} L_0$.

For the last part, consider the natural map $A' \otimes_A \mathbf{D}^L_{\mathrm{st}}(V) \to \mathbf{D}^L_{\mathrm{st}}(V \otimes_A A')$. We can extend scalars from $A' \otimes_{\mathbf{Q}_p} L_0$ to $A' \otimes_{\mathbf{Q}_p} L$ to get a map

$$\begin{tikzcd}
(A' \otimes_{\mathbf{Q}_p} L) \otimes_{A' \otimes_{\mathbf{Q}_p} L_0} (A' \otimes_A \mathbf{D}^L_{\mathrm{st}}(V)) \arrow[r] \arrow[d, equal] & (A' \otimes_{\mathbf{Q}_p} L) \otimes_{A' \otimes_{\mathbf{Q}_p} L_0} (\mathbf{D}^L_{\mathrm{st}}(V \otimes_A A')) \arrow[d] \\
A' \otimes_A \mathbf{D}^L_{\mathrm{dR}}(V) & \mathbf{D}^L_{\mathrm{dR}}(V \otimes_A A')
\end{tikzcd}$$

Since $A' \otimes_{\mathbf{Q}_p} L_0 \to A' \otimes_{\mathbf{Q}_p} L$ is a faithfully flat extension and

$$A' \otimes_A \mathbf{D}^L_{\mathrm{dR}}(V) \xrightarrow{\sim} \mathbf{D}^L_{\mathrm{dR}}(V \otimes_A A')$$

is an isomorphism, our map must have been an isomorphism to begin with. □

5.3.2. *Semistability and crystallinity.* We will use uniform potential semistability of families of de Rham representations to define quotients $A^{[a,b]}_{\mathrm{dR}} \twoheadrightarrow A^{[a,b]}_{\mathrm{st}}$ and $A^{[a,b]}_{\mathrm{dR}} \twoheadrightarrow A^{[a,b]}_{\mathrm{cris}}$ cutting out the semistable and crystalline loci in $\mathrm{Sp}(A)$, respectively.

Note that if V becomes semistable over L, then we get a representation ρ of the inertia group $I_{L/K} \subset \mathrm{Gal}_{L/K}$ on $\mathbf{D}^L_{\mathrm{st}}(V)$, a locally free $A \otimes_{\mathbf{Q}_p} L_0$-module of rank d. If $A = \mathbf{Q}_p$, ρ is trivial precisely when V is semistable as a representation of Gal_K. This is because L_0K/K is an unramified extension, and $\mathbf{D}^{L_0K}_{\mathrm{st}}(V) = L_0 \otimes_{K_0} \mathbf{D}^K_{\mathrm{st}}(V)$. Thus, although the Galois group $\mathrm{Gal}_{L/K}$ acts semilinearly on $\mathbf{D}^L_{\mathrm{st}}(V)$ over L_0, only the L_0-linear action of $I_{L/K}$ matters for checking semistability.

Lemma 5.3.4. *Let B be an E-finite Artin local ring, with maximal ideal $\mathfrak{m}$ and residue field k, so that we may view B as a k-algebra. Let V be a free B-module of rank d, equipped with a B-linear action of a finite group G. Then V is isomorphic as a representation of G to $(V \otimes_B k) \otimes_k B$. In particular,* $\operatorname{Tr} V = \operatorname{Tr}(V \otimes_B k)$.

Proof. This follows from the fact that $\mathrm{H}^i(G, \mathfrak{m} \otimes_k \operatorname{ad}(V \otimes_E k)) = 0$ for $i > 0$, since G is finite and the coefficients are a characteristic-0 vector space. □

Proposition 5.3.5. *Let A be an E-affinoid algebra such that* $\operatorname{Sp}(A)$ *is connected. Let V be a free A-module of rank d, and let G be a finite group. Let $\rho : G \to \operatorname{GL}(V)$ be a representation of G on V. Then for any closed points x_1 and x_2 of X, there are finite étale extensions $i_1 : k(x_1) \hookrightarrow E'$, $i_2 : k(x_2) \hookrightarrow E'$ such that $\rho_{x_1} \otimes_{i_1} E'$ and $\rho_{x_2} \otimes_{i_2} E'$ are isomorphic as E'-valued representations of G.*

Proof. Fix an algebraic closure $\overline{E}/E$. There are only finitely many isomorphism classes of d-dimensional representations of G over $\overline{E}$, call them $\rho_1, \ldots, \rho_k$, so there is some subfield $F \subset \overline{\mathbf{Q}_p}$, finite over $\mathbf{Q}_p$, such that they are all defined over F.

Now consider $A' := A \otimes_E F$ and the representation $\rho_{A'} : G \to \operatorname{GL}(V_{A'})$. The conditions $\operatorname{Tr}(\rho) = \operatorname{Tr}(\rho_i)$ each define pairwise disjoint closed subspaces of X_F whose set-theoretic union is all of X_F. Therefore they are all open, as well, and the function $\operatorname{Tr}(\rho)$ is constant on connected components of X_F.

Now let X' be any connected component of X_F. It is finite étale over X, and, in particular, the map $X' \to X$ is surjective. This gives the desired result. □

In the situation of interest to us, Lemma 5.3.4 and Proposition 5.3.5 have the following consequence:

Theorem 5.3.6. *Let V be a de Rham representation of* Gal_K *on a projective A-module of rank d, and let L/K be the finite Galois extension provided by Theorem 5.3.3. Let $\tau : I_{L/K} \to \operatorname{GL}_d(\overline{E})$ be a representation of the inertia group of L/K. There is a quotient $A \twoheadrightarrow A_{\mathrm{dR},\emptyset}$ such that an E-finite artinian point $x : A \to B$ factors through $A_{\mathrm{dR},\emptyset}$ if and only if the representation of $I_{L/K}$ on $\mathbf{D}_{\mathrm{st}}^L(V_x)$ is equivalent to τ. In particular, there is a quotient $A \twoheadrightarrow A_{\mathrm{st}}$ corresponding to τ being the trivial representation.*

Remark 5.3.7. Since there are only finitely many isomorphism classes of representations $\tau : I_{L/K} \to \operatorname{GL}_d(\overline{E})$, $\operatorname{Sp}(A_{\mathrm{dR},\emptyset})$ is a union of connected components of $\operatorname{Sp}(A)$.

Corollary 5.3.8. *Let A and V be as above. If for every E-finite artinian point $x : A \to B$, V_x is semistable with Hodge–Tate weights in a fixed interval $[a, b]$, then $\mathbf{D}_{\mathrm{st}}^K(V)$ is a locally free $A \otimes_{\mathbf{Q}_p} K_0$-module of rank d.*

Proof. The assumptions imply that for every E-finite artinian point $x : A \to B$, V_x is de Rham with Hodge–Tate weights in the interval $[a, b]$. Then, by Theorem 5.1.2 applied to $\mathbf{B}_* = \mathbf{B}_{\mathrm{dR}}$, V is a de Rham representation.

Let L be the finite Galois extension provided by Theorem 5.3.3. Then the assumption on artinian points implies that $A \twoheadrightarrow A_{\mathrm{st}}$ is the identity map. For every closed point $x \in X$ and every $n \geq 0$,

$$\mathbf{D}_{\mathrm{st}}^{L}(V \otimes_A A/\mathfrak{m}_x^n) = L_0 \otimes_{K_0} \mathbf{D}_{\mathrm{st}}^{K}(V \otimes_A A/\mathfrak{m}_x^n),$$

and, in particular, every element $v \in \mathbf{D}_{\mathrm{st}}^{L}(V)$ is fixed by $g \in I_{L/K}$ modulo all $\mathfrak{m}_x^n$. But then v is actually fixed by all $g \in I_{L/K}$, so $I_{L/K}$ acts trivially on $\mathbf{D}_{\mathrm{st}}^{L}(V)$. We can then write $\mathbf{D}_{\mathrm{st}}^{L}(V) = (A \otimes_{\mathbf{Q}_p} L_0) \otimes_{A \otimes_{\mathbf{Q}_p} K_0} \mathbf{D}_{\mathrm{st}}^{K}(V)$, so $\mathbf{D}_{\mathrm{st}}^{K}(V)$ is a locally free $A \otimes_{\mathbf{Q}_p} K_0$-module of rank d. □

Corollary 5.3.9. *Let A and V be as above. If for every E-finite artinian point $x : A \to B$, V_x is semistable with Hodge–Tate weights in a fixed interval $[a, b]$, then the natural map $(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\log}^{\dagger})[1/t] \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{D}_{\log}^{\dagger})(V)[1/t]$ is an isomorphism.*

Proof. Since there is some s_n such that $\mathbf{D}_{\mathrm{st}}(V) = (\mathbf{D}_{\log,K}^{\dagger,s_n}(V)[1/t])^{\Gamma}$, we are reduced to showing that

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\log}^{\dagger,s_n})[1/t] \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to \mathbf{D}_{\log,K}^{\dagger,s_n}(V)[1/t]$$

is an isomorphism.

The left and right sides are each coherent modules over $A \mathbin{\widehat{\otimes}} \mathbf{B}_{\log,K}^{\dagger,s_n}[1/t] = A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{rig,K}}^{\dagger,s_n}[\log([\overline{\pi}]), 1/t]$, and the homomorphism respects the grading given by powers of $\log([\overline{\pi}])$. We are thus further reduced to considering homomorphisms of coherent modules over $A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{rig,K}}^{\dagger,s_n}[1/t]$, which is the ring of global sections of a quasi-Stein space. The quasi-Stein space in question is the product of $\mathrm{Sp}(A)$ with a half-open annulus (associated to $\mathbf{B}_{\mathrm{rig,K}}^{\dagger,s_n}$) minus the divisor of t, which we denote Y.

Since every point of $\mathrm{Sp}(A) \times Y$ sits over a point of $\mathrm{Sp}(A)$, to prove surjectivity of the desired map, it suffices to check on artinian thickenings of closed points of $\mathrm{Sp}(A)$. But this holds by [Berger 2002, Proposition 3.7]. Furthermore, since $\mathbf{D}_{\log,K}^{\dagger,s_n}(V)[1/t]$ is finite projective, we may check injectivity on points, as well, and this again follows from the same theorem. □

Corollary 5.3.10. *Let A and V be as above. If for every E-finite artinian point $x : A \to B$, V_x is semistable with Hodge–Tate weights in a fixed interval $[a, b]$, then the natural map $(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_{\log}^{\dagger})[1/t] \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_{\log}^{\dagger})[1/t] \otimes_A V$ is an isomorphism.*

Proof. Since $(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_{\log}^{\dagger})[1/t] \otimes_{A \widehat{\otimes} \mathbf{B}_{\mathrm{rig,K}}^{\dagger}} \mathbf{D}_{\mathrm{rig,K}}^{\dagger}(V) \cong (A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_{\log}^{\dagger})[1/t] \otimes_A V$, it suffices to show that the natural map

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\log}^{\dagger})[1/t] \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to \mathbf{D}_{\log,K}^{\dagger}(V)[1/t]$$

is an isomorphism, and follows from Corollary 5.3.9 □

Corollary 5.3.11. *Let A and V be as above. If for every E-finite artinian point $x : A \to B$, V_x is semistable with Hodge–Tate weights in a fixed interval $[a, b]$, then the natural map $(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_A V$ is an isomorphism.*

Proof. We first show that the natural morphism

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}^+_{\log})[1/t] \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}^+_{\log})[1/t] \otimes_A V$$

is an isomorphism. Injectivity follows from Corollary 5.3.10, since $A \mathbin{\widehat{\otimes}} \mathbf{B}^+_{\log} \to A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}$ is injective and $\mathbf{D}_{\mathrm{st}}(V)$ is flat. For surjectivity, we choose bases $\{v_i\}$ and $\{w_j\}$ of V and $\mathbf{D}_{\mathrm{st}}(V)$, respectively (since $\mathbf{D}_{\mathrm{st}}(V)$ is A-locally free, we may assume that both V and $\mathbf{D}_{\mathrm{st}}(V)$ are free). Let $M \in \mathrm{Mat}_{d\times d}(A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{\dagger}_{\log}[1/t])$ be the matrix whose j-th column is the coordinates of w_j with respect to $\{v_i\}$, and let $P \in \mathrm{GL}(A \otimes_{\mathbf{Q}_p} K_0)$ be the matrix of Frobenius on $\mathbf{D}_{\mathrm{st}}(V)$ with respect to $\{w_j\}$. As in the proof of Proposition 4.2.16, M and P satisfy the relation $MP = \varphi(M)$, so, by Corollary 4.2.14, $M \in \mathrm{GL}_d(\mathbf{A} \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}^{+}_{\log}[1/t])$.

Finally, since $\mathbf{B}^+_{\log}[1/t] \subset \mathbf{B}_{\mathrm{st}}$, we may extend scalars to see that

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{st}}) \otimes_A V$$

is an isomorphism. □

Remark 5.3.12. A similar isomorphism is proved in [Hartl and Hellmann 2013], using a different sheaf of semistable period rings.

Corollary 5.3.13. *Suppose V is semistable with Hodge–Tate weights in the interval $[a, b]$, and let $f : A \to A'$ be a homomorphism of E-affinoid algebras. Then $V \otimes_A A'$ is semistable with Hodge–Tate weights in the interval $[a, b]$.*

To handle crystalline representations, we note that $\mathbf{D}_{\mathrm{cris}}(V) = \mathbf{D}_{\mathrm{st}}(V)^{N=0}$. Then the results below follow easily:

Theorem 5.3.14. *Let A be an E-affinoid algebra, and let V be a free A-module of rank d equipped with a continuous A-linear action of Gal_K. Let $[a, b]$ be a finite interval. Then there is a quotient $A \twoheadrightarrow A^{[a,b]}_{\mathrm{cris}}$ such that an E-finite Artin point $x : A \to B$ factors through $A^{[a,b]}_{\mathrm{cris}}$ if and only if V_x is crystalline with Hodge–Tate weights in the interval $[a, b]$.*

Theorem 5.3.15. *Let A be an E-affinoid algebra, and let V be a free A-module of rank d equipped with a continuous A-linear action of Gal_K. Suppose that for every E-finite artinian point $x : A \to B$ the representation V_x is crystalline with Hodge–Tate weights in an interval $[a, b]$. Then $\mathbf{D}_{\mathrm{cris}}(V)$ is a locally free $A \otimes_{\mathbf{Q}_p} K_0$-module of rank d, the formation of $\mathbf{D}_{\mathrm{cris}}(V)$ commutes with base change on A, and the natural homomorphism*

$$(A \mathbin{\widehat{\otimes}} \mathbf{B}_{\max}) \otimes_{A \otimes K_0} \mathbf{D}_{\mathrm{st}}(V) \to (A \mathbin{\widehat{\otimes}} \mathbf{B}_{\max}) \otimes_A V$$

is an isomorphism.

Appendix: Rings of p-adic Hodge theory

Most of the definitions and properties of the rings we use are given in [Berger 2002], and we refer to it freely. However, we describe the construction and topologies of the period rings $\mathbf{B}_{\mathrm{HT}}$, $\mathbf{B}_{\mathrm{dR}}$, $\mathbf{B}_{\mathrm{st}}$, and $\mathbf{B}_{\max}$ with some care so that we can define sheaves of period rings in Section A.2.

A.1. *Period rings.* Let V be a finite-dimensional $\mathbf{Q}_p$-vector space equipped with a continuous $\mathbf{Q}_p$-linear action of Gal_K, for some finite extension $K/\mathbf{Q}_p$. Then for any period ring $\mathbf{B}_*$ listed above, we define $\mathbf{D}^K_{\mathbf{B}_*}(V) := (\mathbf{B}_* \otimes_{\mathbf{Q}_p} V)^{\mathrm{Gal}_K}$. For every choice of $\mathbf{B}_*$, $\mathbf{B}_*^{\mathrm{Gal}_K}$ is a field; we say that V is $\mathbf{B}_*$-*admissible* if $\dim_{\mathbf{B}_*^{\mathrm{Gal}_K}} \mathbf{D}^K_{\mathbf{B}_*}(V) = \dim_{\mathbf{Q}_p} V$.

Remark A.1.1. There is a general formalism of period rings developed in [Fontaine 1994, §1]. However, because of issues related to the topologies on various rings, it is not clear to us that this formalism generalizes in any meaningful way to the study of arithmetic families of Galois representations. Thus, we content ourselves with giving a list of period rings of interest to us.

The ring $\mathbf{B}_{\mathrm{HT}}$. We define $\mathbf{B}_{\mathrm{HT}}$ to be the polynomial ring $\mathbf{C}_p[t, t^{-1}]$. This ring is graded by powers of t. For any $K/\mathbf{Q}_p$, the Galois group Gal_K acts on $\mathbf{B}_{\mathrm{HT}}$ via the natural action on $\mathbf{C}_p$ and via $g \cdot t = \chi(g)t$.

The ring $\mathbf{B}_{\mathrm{dR}}$. The construction of $\mathbf{B}_{\mathrm{dR}}$ is more complicated. Recall the existence of a Galois-equivariant map $\theta : \widetilde{\mathbf{B}}^+ \to \mathbf{C}_K$ characterized by $\theta\left(\sum [c_n] p^n\right) = \sum c_n^{(0)} p^n$. It is continuous with respect to the weak topology on $\widetilde{\mathbf{B}}^+$ and the p-adic topology on $\mathbf{C}_K$, and its kernel is the principal ideal generated by $[\tilde{p}] - p$. Then $\mathbf{B}^+_{\mathrm{dR}}$ is by definition $\varprojlim \widetilde{\mathbf{B}}^+ / \ker(\theta)^h$.

We are grateful to Laurent Berger for providing the following definition of the topology on $\mathbf{B}^+_{\mathrm{dR}}$. Since θ is Galois-equivariant, the Galois action on $\widetilde{\mathbf{B}}^+$ induces a Galois action on $\mathbf{B}^+_{\mathrm{dR}}$. We want to topologize the quotients $\widetilde{\mathbf{B}}_h := \widetilde{\mathbf{B}}^+ / \ker(\theta)^h$ so that this action is continuous. We could make $\widetilde{\mathbf{B}}_h$ into a p-adic Banach space with unit ball $\mathbf{A}_h := \widetilde{\mathbf{A}}^+ / \ker(\theta)^h \cap \widetilde{\mathbf{A}}^+$, but then the action of Galois would not be obviously continuous, so instead we try to use the weak topology, i.e., the topology on the image of $\widetilde{\mathbf{A}}^+$ generated by the images of the $U_{k,n}$. For $n \geq h$, though,

$$[\tilde{p}]^n = (([\tilde{p}] - p) + p)^n = (([\tilde{p}] - p) + p)^h (([\tilde{p}] - p) + p)^{n-h} \in p\mathbf{A}_h.$$

In particular, this shows that $p\widetilde{\mathbf{A}}_h$ is an open ideal.

A priori, the p-adic topology on $\widetilde{\mathbf{A}}_h$ has more open sets than the weak topology does. But we have just shown that every open set of the p-adic topology is actually open in the weak topology, so the two topologies must be the same. In particular, the weak topology on $\widetilde{\mathbf{A}}_h$ is Hausdorff and complete.

The upshot is that $\widetilde{\mathbf{B}}_h = \bigcup_{i\geq 0} p^{-i}\widetilde{\mathbf{A}}_h$ has a natural structure of a p-adic Banach space with unit ball $\widetilde{\mathbf{A}}_h$, so $\mathbf{B}_{\mathrm{dR}}^+$ has a natural structure of a p-adic Fréchet space.

The ring $\mathbf{B}_{\max}$. We will use the ring $\mathbf{B}_{\max}$ to study crystalline representations, rather than $\mathbf{B}_{\mathrm{cris}}$, because the topology on $\mathbf{B}_{\max}$ is much nicer.

Remark A.1.2. There is a closely related ring $\widetilde{\mathbf{B}}_{\mathrm{rig}}^+$, defined as $\bigcap_n \varphi^n(\mathbf{B}_{\max}^+) = \bigcap_n \varphi^n(\mathbf{B}_{\mathrm{cris}})$; $\mathbf{B}_{\max}$ and $\widetilde{\mathbf{B}}_{\mathrm{rig}}^+[1/t]$ define the same functor from the category of $\mathbf{Q}_p$-representations to the category of filtered K_0-isocrystals. However, we prefer to work with $\mathbf{B}_{\max}^+$ because it is a Banach space, while $\widetilde{\mathbf{B}}_{\mathrm{rig}}^+$ is a Fréchet space.

The ring $\mathbf{B}_{\mathrm{st}}$. After we choose a value for $\log(p)$, the power series defining $\log(\bar{\pi}^{(0)}) + \log([\bar{\pi}]/\bar{\pi}^{(0)})$ converges in $\mathbf{B}_{\mathrm{dR}}^+$. We let $\mathbf{B}_{\mathrm{st}} := \mathbf{B}_{\max}[\log[\bar{\pi}]]$ and $\log(p) = 0$.

Remark A.1.3. This is not the standard definition of $\mathbf{B}_{\mathrm{st}}$, but it is the usage of [Berger 2002] and [Berger and Colmez 2008]. It defines the same functor on representations of Gal_K as the usual $\mathbf{B}_{\mathrm{st}}$, because $\mathbf{B}_{\mathrm{cris}}$ and $\mathbf{B}_{\max}$ define the same functor. If we define $\widetilde{\mathbf{B}}_{\log} := \widetilde{\mathbf{B}}_{\mathrm{rig}}[\log[\bar{\pi}]]$, then $\widetilde{\mathbf{B}}_{\log}$ defines the same functor as well.

Remark A.1.4. We have defined $\mathbf{B}_{\mathrm{st}}$ as a subring of $\mathbf{B}_{\mathrm{dR}}$ in terms of a choice of a branch of the p-adic logarithm. A different choice would lead to a different subring. It is also possible to define $\mathbf{B}_{\mathrm{st}}$ intrinsically as an abstract ring and use the choice of a p-adic logarithm to define an embedding of $\mathbf{B}_{\mathrm{st}}$ in $\mathbf{B}_{\mathrm{dR}}$. This approach makes clear that $\mathbf{B}_{\mathrm{st}}$-admissibility of a representation does not depend on any choices. We have not taken this approach here; for details about the construction of the usual $\mathbf{B}_{\mathrm{st}}$ as an extension of $\mathbf{B}_{\mathrm{cris}}$; see [Brinon and Conrad 2009, §9.2].

Remark A.1.5. Let L/K be a finite extension. Then we obtain a map

$$L \otimes_{L_0} \mathbf{B}_{\max}^+ \to \mathbf{B}_{\mathrm{dR}}^+$$

by extending the inclusion $\mathbf{B}_{\max}^+ \hookrightarrow \mathbf{B}_{\mathrm{dR}}^+$ by L-linearity. Then Colmez [2002, Proposition 7.14] has shown that this map is an injection. In the course of the proof, he showed that it is possible to write down an isomorphism of K-Fréchet spaces $\mathbf{B}_{\mathrm{dR}}^+ \cong \mathbf{C}_K[[t]]$ so that $L \otimes_{L_0} \mathbf{B}_{\max}^+$ is carried isomorphically to the Banach space $\mathbf{C}_K\langle T\rangle$ and $L \otimes_{L_0} \mathbf{B}_{\mathrm{st}}^+$ is carried isomorphically to $\mathbf{C}_K\langle T\rangle[\log(1+T)]$. Note that these are isomorphisms as vector spaces, not as rings!

A.2. ***Sheaves of period rings.*** As we wish to study p-adic families of Galois representations, we need to define versions of these rings with "coefficients" in Banach algebras, rather than simply $\mathbf{Q}_p$.

Let $(X, \mathscr{O}_X)$ be a quasicompact quasiseparated rigid space over a finite extension $E/\mathbf{Q}_p$.

Definition A.2.1. Let B be a $\mathbf{Q}_p$-Banach algebra. Then we define the presheaf $\mathscr{B}_X$ on X by setting

$$\mathscr{B}_X(U) := \mathscr{O}_X(U) \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B,$$

where U is an admissible affinoid open of X.

By [Kedlaya and Liu 2010, Lemma 3.3], $\mathscr{B}_X$ is actually a sheaf on U when U is affinoid. Therefore, it extends to a sheaf on X. We wish to extend this to Fréchet algebras in the role of B.

We first record some basic functional analysis results, which will be useful for checking that exactness properties are preserved under completed tensor products.

Let I be a set (not necessarily countable), and let N be a Fréchet space equipped with a countable family of seminorms $\{q_j\}$ (in particular, N could be a Banach space). We define the space $c_I(N)$ to be the set of functions $f : I \to N$ such that $\lim_{i\in I} q_j(f(i)) = 0$ for each seminorm q_j. That is, for each j and each $\varepsilon > 0$, the set $\{i \in I \mid q_j(f(i)) > \varepsilon\}$ is finite. We equip $c_I(N)$ with the seminorms $q_{j,\infty}$ defined by $q_{j,\infty}(f) := \sup_{i\in I} q_j(f(i))$, making $c_I(N)$ into a Fréchet space. Following [Buzzard 2007], we say that a Banach module N over a Banach algebra A is *potentially orthonormalizable* if there is some Banach norm on N making it is isomorphic to $c_I(A)$ for some index set I. For example, all Banach spaces over discretely valued fields are potentially orthonormalizable [Schneider 2002, Proposition 10.1].

Lemma A.2.2. *Let k be a nonarchimedean field, let M be a potentially orthonormalizable k-Banach space, and let N be a k-Fréchet space, with countable family of seminorms $\{q_j\}$. Write $M \cong c_I(k)$. Then the natural $\mathbf{Q}_p$-linear map $M \mathbin{\widehat{\otimes}}_k N \to c_I(N)$ is an isomorphism, functorially in N.*

Proof. The natural map $M \mathbin{\widehat{\otimes}} N \to c_I(N)$ is induced by the bilinear map $M \times N \to c_I(N)$ sending $(f : I \to k, b)$ to $\sum_{i\in I} f(i)b$. The sum converges because $q_j(f(i)b) \le |f(i)| \cdot q_j(b)$ and $\lim_{i\in I} a_i = 0$.

To construct a map in the other direction, we observe that any element $f \in c_I(N)$ can be written as the limit of elements of the form $f|_S$, where $S \subset I$ is a finite subset. More precisely, the set of finite subsets $S \subset I$ is a directed set, and $S \mapsto f|_S$ is a net converging to f. For any finite set $S \subset I$, we write $(\mathbf{1})_S \in c_I(k)$ for the characteristic function of S. Now, consider $\sum_{i\in S} \mathbf{1}_{\{i\}} \otimes f(i) \in c_I(k) \otimes_k N$. We have

$$q_j\Bigl(\sum_{i\in S} \mathbf{1}_{\{i\}} \otimes f(i)\Bigr) \le \max_{i\in S} q_j(f(i)),$$

so the net $S \mapsto \sum_{i\in S} \mathbf{1}_{\{i\}} \otimes f(i)$ converges in $c_I(k) \mathbin{\widehat{\otimes}} N$. The map $f \mapsto \lim_S \sum_{i\in S} \mathbf{1}_{\{i\}} \otimes f(i)$ provides an inverse to the map $M \mathbin{\widehat{\otimes}}_k N \to c_I(N)$. □

Corollary A.2.3. *Let k and M be as above, and let $N \to N'$ be a continuous injection of k-Fréchet spaces. Then the natural map $M \mathbin{\widehat{\otimes}} N \to M \mathbin{\widehat{\otimes}} N'$ is injective.*

Definition A.2.4. Let $B = \varprojlim_n B_n$ be a $\mathbf{Q}_p$-Fréchet algebra, where the B_n are $\mathbf{Q}_p$-Banach algebras. Then we define the presheaf $\mathscr{B}_X$ on X by setting

$$\mathscr{B}_X(U) := \mathscr{O}_X(U) \mathbin{\widehat{\otimes}} V = \varprojlim_n \mathscr{O}_X(U) \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B_n$$

when U is an admissible affinoid open of X.

Lemma A.2.5. *$\mathscr{B}_X$ is a sheaf on X.*

Proof. It suffices to prove this when $X = \mathrm{Sp}(A)$ is affinoid, for A some E-affinoid algebra. Further, it suffices to check the sheaf property on Laurent coverings of $\mathrm{Sp}(A)$. That is, we need to check that the sequence

$$0 \longrightarrow \varprojlim_n A \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B_n \longrightarrow \varprojlim_n A\langle f\rangle \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B_n \times \varprojlim_n A\langle f^{-1}\rangle \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B_n \longrightarrow \varprojlim_n A\langle f, f^{-1}\rangle \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} B_n$$

is exact. But

$$0 \longrightarrow A \longrightarrow A\langle f\rangle \times A\langle f^{-1}\rangle \longrightarrow A\langle f, f^{-1}\rangle \longrightarrow 0$$

is exact, and the quotient admits a section by [Schneider 2002, Proposition 10.5], since $\mathbf{Q}_p$-affinoid algebras are countable-type over $\mathbf{Q}_p$. It follows that

$$0 \longrightarrow A \widehat{\otimes}_{\mathbf{Q}_p} B_n \longrightarrow A\langle f\rangle \widehat{\otimes}_{\mathbf{Q}_p} B_n \times A\langle f^{-1}\rangle \widehat{\otimes}_{\mathbf{Q}_p} B_n \longrightarrow A\langle f, f^{-1}\rangle \widehat{\otimes}_{\mathbf{Q}_p} B_n \longrightarrow 0$$

is exact for each n, and inverse limits are left-exact. ☐

Thus, taking B to be $\widetilde{\mathbf{B}}_K^{\dagger,s}$, $\mathbf{B}_K^{\dagger,s}$, $\widetilde{\mathbf{B}}_{\mathrm{rig,K}}^{\dagger,s}$, $\mathbf{B}_{\mathrm{rig,K}}^{\dagger,s}$, $\mathbf{C}_K$, $\mathbf{B}_{\mathrm{dR}}^+$, or $\mathbf{B}_{\max}^+$, we get a sheaf of rings. Furthermore, since taking rising unions is exact, we see that if $U \subset X = \mathrm{Sp}(A)$ is an affinoid subdomain with coordinate ring A_U,

$$U \mapsto \bigcup_s A_U \mathbin{\widehat{\otimes}} \mathbf{B}_{\mathrm{rig,K}}^{\dagger,s}$$

is a sheaf on X. Similarly, we also get sheaves associated to $\mathbf{B}_{\mathrm{HT}} = \mathbf{C}_K[t, t^{-1}]$, $\mathbf{B}_{\mathrm{dR}} = \bigcup_i t^{-i}\mathbf{B}_{\mathrm{dR}}^+$, $\mathbf{B}_{\max} = \bigcup_i t^{-i}\mathbf{B}_{\max}^+$, and $\mathbf{B}_{\mathrm{st}} = \mathbf{B}_{\max}[\log[\overline{\pi}]]$. Each of these sheaves carries the additional structures, such as Galois action, Frobenius action, grading, filtration, or monodromy action, of the absolute ring.

Proposition A.2.6. (1) *$\mathscr{B}_{X,\mathrm{HT}}$ is a graded sheaf of rings over X, equipped with an action of* Gal_K, *and* $\mathscr{B}_{X,HT}^{\mathrm{Gal}_K} = \mathscr{O}_X \otimes_{\mathbf{Q}_p} K$.

(2) *$\mathscr{B}_{X,\mathrm{dR}}$ is a filtered sheaf of rings over X, equipped with an action of* Gal_K, *and* $\mathscr{B}_{X,\mathrm{dR}}^{\mathrm{Gal}_K} = \mathscr{O}_X \otimes_{\mathbf{Q}_p} K$.

(3) *$\mathscr{B}_{X,\max}$ is a sheaf of rings over X, equipped with an action of* Gal_K *and an action of φ, and* $\mathscr{B}_{X,\max}^{\mathrm{Gal}_K} = \mathscr{O}_X \otimes_{\mathbf{Q}_p} K_0$.

(4) *$\mathscr{B}_{X,\mathrm{st}}$ is a sheaf of rings over X, equipped with an action of* Gal_K *and an action of φ and N, and $\mathscr{B}_{X,\mathrm{st}}^{\mathrm{Gal}_K} = \mathscr{O}_X \otimes_{\mathbf{Q}_p} K_0$.*

Proof. For all of these, it suffices to consider the case when $X = \mathrm{Sp}(A)$ is affinoid for some E-affinoid algebra A. Then A is countable-type over $\mathbf{Q}_p$, so we can choose a Schauder basis for A with index set I. Under the resulting isomorphism $A \cong c_I(\mathbf{Q}_p)$, we have $A \mathbin{\widehat{\otimes}} B \cong c_I(B)$ for a $\mathbf{Q}_p$-Fréchet space B. The Galois action on $A \mathbin{\widehat{\otimes}} B$ is $g \cdot (b_i)_{i\in I} = (g \cdot b_i)_{i\in I}$, where I is the index set for the Schauder basis, so the assertions follow from the corresponding classical results. □

Lemma A.2.7. *Let A be a reduced $\mathbf{Q}_p$-affinoid algebra, and let $\mathbf{B}_*$ be a period ring. Then the natural map*

$$A \mathbin{\widehat{\otimes}} \mathbf{B}_* \to \prod_{x\in \mathrm{Sp}(A)} A/\mathfrak{m}_x \otimes_{\mathbf{Q}_p} \mathbf{B}_*$$

is injective.

Proof. If $\mathbf{B}_* = \mathbf{B}_{\mathrm{HT}}$, it suffices to show that $A \mathbin{\widehat{\otimes}} \mathbf{C}_p \to \prod_{x\in\mathrm{Sp}(A)} A/\mathfrak{m}_x \otimes_{\mathbf{Q}_p} \mathbf{C}_p$ is injective. If $\mathbf{B}_* = \mathbf{B}_{\mathrm{dR}}$, it suffices to show that $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_h \to \prod_{x\in\mathrm{Sp}(A)} A/\mathfrak{m}_x \otimes_{\mathbf{Q}_p} \widetilde{\mathbf{B}}_h$ is injective for all h. And, since $\mathbf{B}_{\mathrm{st}}$ is a polynomial algebra over $\mathbf{B}_{\max}$, it suffices to show that $A \mathbin{\widehat{\otimes}} \widetilde{\mathbf{B}}_{\max} \to \prod_{x\in\mathrm{Sp}(A)} A/\mathfrak{m}_x \otimes_{\mathbf{Q}_p} \widetilde{\mathbf{B}}_{\max}$ is injective. We are therefore reduced to showing that for any $\mathbf{Q}_p$-Banach space B, the natural map $A \mathbin{\widehat{\otimes}} B \to \prod_{x\in\mathrm{Sp}(A)} A/\mathfrak{m}_x \otimes_{\mathbf{Q}_p} B$ is injective.

Since B is a $\mathbf{Q}_p$-Banach space, it is potentially orthonormalizable in the sense of [Buzzard 2007]. That is, it admits a basis $\{e_i\}_{i\in I}$ such that

$$B \cong c_I(\mathbf{Q}_p) := \{f : I \to \mathbf{Q}_p \mid |f(i)| < \varepsilon \text{ for almost all } i \in I, \text{ for all } \varepsilon > 0\}$$

Then $A \mathbin{\widehat{\otimes}} B \cong c_I(A)$, and the desired injectivity follows from the injectivity of the natural map $A \to \prod_{x\in\mathrm{Sp}(A)} A/\mathfrak{m}_x$. □

Acknowledgments

The results in this paper form a portion of my Ph.D. thesis, and I would first like to thank my advisor, Brian Conrad, for suggesting the direction of this research and for numerous conversations, as well as for helpful comments on earlier drafts of this paper. I am indebted to Laurent Berger and Pierre Colmez for their mathematical contributions, and I am grateful to Kiran Kedlaya and Jay Pottharst for helpful conversations, as well as Toby Gee for helpful comments on earlier drafts. This work was partially supported by an NSF Graduate Research Fellowship.

References

[André 2002] Y. André, "Filtrations de type Hasse–Arf et monodromie p-adique", *Invent. Math.* **148**:2 (2002), 285–317. MR 2003k:12011 Zbl 1081.12003

[Bellovin 2013] R. Bellovin, *p-adic Hodge theory in rigid analytic families*, Ph.D. thesis, Stanford University, 2013, Available at http://www2.imperial.ac.uk/~rbellovi/thesis.pdf.

[Berger 2002] L. Berger, "Représentations p-adiques et équations différentielles", *Invent. Math.* **148**:2 (2002), 219–284. MR 2004a:14022 Zbl 1113.14016

[Berger and Colmez 2008] L. Berger and P. Colmez, "Familles de représentations de de Rham et monodromie p-adique", pp. 303–337 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et* (ϕ, Γ)*-modules*, edited by L. Berger et al., Astérisque **319**, 2008. MR 2010g:11091 Zbl 1168.11020

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR 91k:32038 Zbl 0715.14013

[Bosch and Lütkebohmert 1993] S. Bosch and W. Lütkebohmert, "Formal and rigid geometry, I: Rigid spaces", *Math. Ann.* **295**:2 (1993), 291–317. MR 94a:11090 Zbl 0808.14017

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis*, Grundlehren der Mathematischen Wissenschaften **261**, Springer, Berlin, 1984. MR 86b:32031 Zbl 0539.14017

[Breuil and Mézard 2002] C. Breuil and A. 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 (2002), 205–310. MR 2004i:11052 Zbl 1042.11030

[Brinon and Conrad 2009] O. Brinon and B. Conrad, "CMI Summer School notes on p-adic Hodge theory (preliminary version)", course notes, 2009, Available at http://math.stanford.edu/~conrad/papers/notes.pdf.

[Buzzard 2007] K. Buzzard, "Eigenvarieties", pp. 59–120 in *L-functions and Galois representations*, edited by D. Burns et al., London Math. Soc. Lecture Note Ser. **320**, Cambridge University Press, 2007. MR 2010g:11076 Zbl 1230.11054

[Chenevier 2009] G. Chenevier, "Une application des variétés de Hecke des groups unitaires", preprint, 2009, Available at http://gaetan.chenevier.perso.math.cnrs.fr/articles/famgal.pdf.

[Cherbonnier and Colmez 1998] F. Cherbonnier and P. Colmez, "Représentations p-adiques surconvergentes", *Invent. Math.* **133**:3 (1998), 581–611. MR 2000d:11146 Zbl 0928.11051

[Colmez 2002] P. Colmez, "Espaces de Banach de dimension finie", *J. Inst. Math. Jussieu* **1**:3 (2002), 331–439. MR 2004b:11160 Zbl 1044.11102

[Diao and Liu 2014] H. Diao and R. Liu, "The Eigencurve is proper", preprint, 2014. arXiv 1401.4871

[Eisenbud 1995] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, 1995. MR 97a:13001 Zbl 0819.13001

[Fontaine 1994] J.-M. Fontaine, "Représentations p-adiques semi-stables", pp. 113–184 in *Périodes p-adiques* (Séminaire de Bures-sur-Yvette, 1988), edited by J.-M. Fontaine, Astérisque **223**, 1994. MR 95g:14024 Zbl 0865.14009

[Fontaine 2004] J.-M. Fontaine, "Arithmétique des représentations galoisiennes p-adiques", pp. xi, 1–115 in *Cohomologies p-adiques et applications arithmétiques (III)*, edited by P. Berthelot et al., Astérisque **295**, 2004. MR 2005i:11074

[Hartl and Hellmann 2013] U. Hartl and E. Hellmann, "The universal family of semi-stable p-adic Galois representations", preprint, 2013. arXiv 1312.6371

[Hellmann 2012] E. Hellmann, "Families of trianguline representations and finite slope spaces", preprint, 2012. arXiv 1202.4408

[Kedlaya 2004] K. S. Kedlaya, "A p-adic local monodromy theorem", *Ann. of Math.* (2) **160**:1 (2004), 93–184. MR 2005k:14038 Zbl 1088.14005

[Kedlaya and Liu 2010] K. Kedlaya and R. Liu, "On families of (φ, Γ)-modules", *Algebra Number Theory* **4**:7 (2010), 943–967. MR 2012h:11081 Zbl 1278.11060

[Kedlaya et al. 2014] K. S. Kedlaya, J. Pottharst, and L. Xiao, "Cohomology of arithmetic families of (φ, Γ)-modules", *J. Amer. Math. Soc.* **27**:4 (2014), 1043–1115. MR 3230818

[Kiehl 1967] R. Kiehl, "Theorem A und Theorem B in der nichtarchimedischen Funktionentheorie", *Invent. Math.* **2** (1967), 256–273. MR 35 #1834 Zbl 0202.20201

[Kisin 2008] M. Kisin, "Potentially semi-stable deformation rings", *J. Amer. Math. Soc.* **21**:2 (2008), 513–546. MR 2009c:11194 Zbl 1205.11060

[Liu 2014] R. Liu, "Triangulation of refined families", preprint, 2014. arXiv 1202.2188

[Liu 2015] R. Liu, "Semistable periods of finite slope families", *Algebra Number Theory* **9**:2 (2015), 435–458.

[Matsumura 1989] H. Matsumura, *Commutative ring theory*, 2nd ed., Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, 1989. MR 90i:13001 Zbl 0666.13002

[Mebkhout 2002] Z. Mebkhout, "Analogue p-adique du théorème de Turrittin et le théorème de la monodromie p-adique", *Invent. Math.* **148**:2 (2002), 319–351. MR 2003k:14018 Zbl 1071.12004

[Pottharst 2013] J. Pottharst, "Analytic families of finite-slope Selmer groups", *Algebra Number Theory* **7**:7 (2013), 1571–1612. MR 3117501 Zbl 06226679

[Schneider 2002] P. Schneider, *Nonarchimedean functional analysis*, Springer, Berlin, 2002. MR 2003a:46106 Zbl 0998.46044

[Schneider and Teitelbaum 2003] P. Schneider and J. Teitelbaum, "Algebras of p-adic distributions and admissible representations", *Invent. Math.* **153**:1 (2003), 145–196. MR 2004g:22015 Zbl 1028.11070

[Sen 1973] S. Sen, "Lie algebras of Galois groups arising from Hodge–Tate modules", *Ann. of Math.* (2) **97** (1973), 160–170. MR 47 #3403 Zbl 0258.12009

[Shah 2013] S. Shah, "Interpolating periods", preprint, 2013. arXiv 1305.2872

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR 95f:18001 Zbl 0797.18001

[Wintenberger 1983] J.-P. Wintenberger, "Le corps des normes de certaines extensions infinies de corps locaux; applications", *Ann. Sci. École Norm. Sup.* (4) **16**:1 (1983), 59–89. MR 85e:11098 Zbl 0516.12015

Communicated by Mark Kisin
Received 2014-01-23 Revised 2014-11-21 Accepted 2014-12-25

rmb@math.berkeley.edu *Department of Mathematics, University of California, Berkeley, Evans Hall, Berkeley, CA 94720, United States*

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.435

Semistable periods of finite slope families

Ruochuan Liu

We introduce the notion of finite slope families to encode the local properties of the p-adic families of Galois representations appearing in the work of Harris, Lan, Taylor and Thorne on the construction of Galois representations for (non-self-dual) regular algebraic cuspidal automorphic representations of $\mathrm{GL}(n)$ over CM fields. Our main result is to prove the analytic continuation of semistable (and crystalline) periods for such families.

Introduction and notation

Skinner and Urban [2006], in their ICM talk, outlined a program to connect the order of vanishing of the L-functions of certain polarized regular motives with the rank of the associated Bloch–Kato Selmer groups. Their strategy is to deform those motives along certain p-adic families, the so-called eigenfamilies, to construct the expected extensions. To this end, they introduced the notion of *finite slope families* of p-adic representations to encode the local properties of the Galois representations arising from those p-adic families. One may view the finite slope families as a generalization of the p-adic families of Galois representations arising from the Coleman–Mazur eigencurve. Bellaïche and Chenevier [2009] introduced the notion of *weakly refined families* of p-adic representations to encode the local properties of the latter. More precisely, a family of weakly refined p-adic representations is

MSC2010: 11F80.
Keywords: finite slope families, semistable, (φ, Γ)-modules.

a family of p-adic representations[1] over a rigid analytic space with a Zariski-dense subset of crystalline points which have crystalline periods of a prescribed Frobenius eigenvalue and constant Hodge–Tate weight. Moreover, this constant weight is the largest one[2] of all Hodge–Tate weights, and the difference of this constant weight with any other weight is unbounded over the base. For example, in the case of the eigencurve, one can take the subset of all classical eigenforms, and the prescribed Frobenius eigenvalue and constant Hodge–Tate weight are the function of U_p-eigenvalues and 0 respectively. On the other hand, finite slope families generalize weakly refined families in the way that allows *multiple* prescribed Frobenius eigenvalues and constant Hodge–Tate weights. Skinner and Urban then used the (unproven) analytic continuation of crystalline periods of finite slope families to deduce that the extensions constructed by p-adic deformations lie in the Selmer groups.

Most recently, Harris, Lan, Taylor and Thorne [Harris et al. 2014] (and Scholze independently) constructed Galois representations for (non-self-dual) regular algebraic cuspidal automorphic representations of $\mathrm{GL}(n)$ over CM fields. It turns out that these Galois representations emerge from certain p-adic families whose local properties generalize Skinner and Urban's finite slope families by allowing prescribed semistable periods. Therefore, to show that the Galois representations constructed by Harris, Lan, Taylor and Thorne have the expected properties at p, one needs to show the analytic continuation of semistable periods for those p-adic families.

In this paper, we make use of the notion of finite slope families to encode the local properties of the p-adic families of Galois representations appearing in the work of Harris, Lan, Taylor and Thorne; this generalizes the original definition of Skinner and Urban. Our main result is then to prove the analytic continuation of semistable periods for such families. This will provide a necessary ingredient in Skinner and Urban's ICM program. Besides, we recently learned from Taylor that Ila Varma, as part of an ongoing project, will establish the expected properties of those Galois representations based on the results of this paper. We also note that recently Shah [2013] proved some results about interpolating Hodge–Tate and de Rham periods in families of p-adic Galois representations which may be applied to some related situations.

In the following, we state our main results precisely. We fix a finite extension K of $\mathbb{Q}_p$. Let K_0 be the maximal unramified subextension of K, and let $f=[K_0:\mathbb{Q}_p]$. We also fix a finite extension F of $\mathbb{Q}_p$ contained in $\overline{\mathbb{Q}}_p$ such that $\mathrm{Hom}(K,F)=\mathrm{Hom}(K,\overline{\mathbb{Q}}_p)$; here Hom denotes the set of $\mathbb{Q}_p$-algebra homomorphisms.

[1] Strictly speaking, Bellaïche–Chenevier used pseudorepresentations rather than genuine representations in their definition of weakly refined families.

[2] We normalize the Hodge–Tate weight such that the p-adic cyclotomic character has Hodge–Tate weight 1.

Definition 0.1. Let X be a reduced rigid analytic space over F. A *finite slope family* of p-adic representations of dimension d over X is a locally free coherent $\mathscr{O}_X$-module V_X of rank d equipped with a continuous G_K-action, together with the following data:

(1) Positive integers b, c.

(2) A monic polynomial $Q(T) \in \mathscr{O}_X(X)[T]$ of degree m with unit constant term.

(3) A subset Z of X such that, for all z in Z, V_z is semistable with nonpositive Hodge–Tate weights, and, for all $B \in \mathbb{Z}$, the set of z in Z such that V_z has $d-c$ Hodge–Tate weights less than B is Zariski-dense in X.

(4) For $z \in Z$, a $K_0 \otimes_{\mathbb{Q}_p} k(z)$-direct summand $\mathscr{F}_z$ of $D^+_{\mathrm{st}}(V_z)$ which is free of rank c and stable under φ and N such that φ^f has characteristic polynomial $Q(z)(T)$ and all Hodge–Tate weights of $\mathscr{F}_z$ lie in $[-b, 0]$.

We also need to extend the functors D^+_{crys} and D^+_{st} to families of p-adic representations over rigid analytic spaces.

Definition 0.2. Let X be a rigid analytic space over $\mathbb{Q}_p$, and let V_X be locally free coherent $\mathscr{O}_X$-module equipped with a continuous G_K-action. Define $D^+_{\mathrm{crys}}(V_X)$ and $D^+_{\mathrm{st}}(V_X)$ to be the presheaves

$$M(S) \mapsto D^+_{\mathrm{crys}}(V_S) = (V_S \mathbin{\widehat{\otimes}}_{\mathbb{Q}_p} \boldsymbol{B}^+_{\mathrm{crys}})^{G_K}$$

and

$$M(S) \mapsto D^+_{\mathrm{st}}(V_S) = (V_S \mathbin{\widehat{\otimes}}_{\mathbb{Q}_p} \boldsymbol{B}^+_{\mathrm{st}})^{G_K},$$

respectively, where $M(S)$ runs through all admissible affinoid subdomain of X; here V_S is the restriction of V_X on $M(S)$.

Now we can state our main result precisely:

Theorem 0.3. *Let V_X be a finite slope family over X. Then there exists a surjective proper morphism $X' \to X$ so that $(K \otimes_{K_0} D^+_{\mathrm{st}}(V_{X'}))^{Q(\varphi^f)=0}$ has a rank-c locally free coherent $K_0 \otimes_{\mathbb{Q}_p} \mathscr{O}_{X'}$-submodule which specializes to a rank-c free $K_0 \otimes_{\mathbb{Q}_p} k(x)$-submodule in $D^+_{\mathrm{st}}(V_x)$ for any $x \in X'$. As a consequence, for any $x \in X$, $D^+_{\mathrm{st}}(V_x)^{Q(x)(\varphi^f)=0}$ has a free $K_0 \otimes_{\mathbb{Q}_p} k(x)$-submodule of rank c.*

The next result follows immediately:

Corollary 0.4. *Let V_X be a finite slope family over X. If V_z is crystalline for any $z \in Z$, then there exists a surjective proper morphism $X' \to X$ so that $(K \otimes_{K_0} D^+_{\mathrm{crys}}(V_{X'}))^{Q(\varphi^f)=0}$ has a rank-c locally free coherent $K_0 \otimes_{\mathbb{Q}_p} \mathscr{O}_{X'}$-submodule which specializes to a rank-c free $K_0 \otimes_{\mathbb{Q}_p} k(x)$-submodule in $D^+_{\mathrm{crys}}(V_x)$ for every $x \in X'$. As a consequence, $D^+_{\mathrm{crys}}(V_x)^{Q(x)(\varphi^f)=0}$ has a free $K_0 \otimes_{\mathbb{Q}_p} k(x)$-submodule of rank c for every $x \in X$.*

Since the families of p-adic representations arising from the Coleman–Mazur eigencurve are special cases of finite slope families, Theorem 0.3 generalizes the famous result of Kisin [2003] on the analytic continuation of crystalline periods over the eigencurve. However, even in the case that the prescribed periods are crystalline, our method is completely different from his. In fact, in the work of Kisin as well as our recent enhancement [Liu 2014], one crucially uses the fact that the families of p-adic representations arising from the eigencurve have only one constant Hodge–Tate weight. On the other hand, our strategy and techniques are largely inspired by the work of Berger and Colmez [2008] on families of de Rham representations with bounded Hodge–Tate weights, and Kedlaya, Pottharst and Xiao [Kedlaya et al. 2014] on the cohomology of families of (φ, Γ)-modules. In fact, for a finite slope family, we first construct a subfamily of (φ, Γ)-modules interpolating the prescribed semistable periods, after making a proper and surjective base change. This is achieved by adapting some techniques of [Kedlaya et al. 2014]. This subfamily of (φ, Γ)-modules is expected to be semistable and produce the desired semistable periods. However, we are unable to prove this directly due to some technical obstacles. Instead, we first show that this subfamily of (φ, Γ)-modules is de Rham, using the fact that it is de Rham at a Zariski-dense subset of the base. To this end, we develop a theory of families of Hodge–Tate and de Rham (φ, Γ)-modules with bounded Hodge–Tate weights, which generalizes the theory of families of Hodge–Tate and de Rham representations with bounded Hodge–Tate weights developed in [Berger and Colmez 2008]. Then we prove the p-adic local monodromy for the restrictions of families of de Rham (φ, Γ)-modules with bounded Hodge–Tate weights on their Shilov boundaries by mimicking the proof for families of de Rham representations with bounded Hodge–Tate weights given in [loc. cit.]. This implies that the de Rham periods of this subfamily of (φ, Γ)-modules become potentially semistable after restricting on the Shilov boundary. Finally, we use a key lemma due to Berger and Colmez [2008] to conclude that these de Rham periods are actually semistable.

Notation

We choose a compatible sequence of primitive p-power roots of unity $(\varepsilon_n)_{n\geq 0}$, i.e., each $\varepsilon_n \in \overline{\mathbb{Q}}_p$ is a primitive p^n-th root of 1, and they satisfy $\varepsilon_{n+1}^p = \varepsilon_n$ for all $n \geq 0$. Fix $\varepsilon = (\varepsilon_0, \varepsilon_1, \dots)$, and let $t = \log[\varepsilon]$ be Fontaine's p-adic $2\pi i$. For a finite extension L of $\mathbb{Q}_p$ in $\mathbb{C}_p$, let $L_n = L(\varepsilon_n)$ for $n \geq 1$, and let $L_\infty = \bigcup_{n\in\mathbb{N}} L_n$. Let L'_0 be the maximal unramified extension of $\mathbb{Q}_p$ in L_∞. Let $\Gamma_L = \mathrm{Gal}(L_\infty/L)$ and $\Gamma_{L_n} = \mathrm{Gal}(L_\infty/L_n)$ for $n \geq 1$. For simplicity, denote Γ_K and Γ_{K_n} by Γ and Γ_n respectively. Let χ denote the p-adic cyclotomic character. For a p-adic representation V of G_K and $n \in \mathbb{Z}$, we set $V(n) = V \otimes \chi^n$. For $n \geq 0$, let $r_n = p^{n-1}(p-1)$. For $s > 0$, let $n(s)$ be the maximal integer n such that $r_n \leq s$.

1. Families of (φ, Γ)-modules

In this section we recall the notion of families of (φ, Γ)-modules over rigid analytic spaces. For the period rings involved in this paper, we follow the notation introduced in [Berger 2002], and we refer the reader to that paper for precise definitions. Note that this is different from the "Robba ring" type notation used in [Kedlaya et al. 2014]. A good dictionary for these two types of notation is given in [Berger 2008a, §1].

Definition 1.1. Let A be a Banach algebra over $\mathbb{Q}_p$. For $s > 0$, a *φ-module* over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ is a finite projective $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$-module D^s_A equipped with an isomorphism

$$\varphi^* D^s_A \cong D^s_A \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A} \boldsymbol{B}^{\dagger,ps}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A.$$

A *φ-module* D_A over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ is the base change to $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ of a φ-module D^s_A over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ for some $s > 0$. A *(φ, Γ)-module* over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ is a φ-module D^s_A over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ equipped with a commuting $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K}$-semilinear and A-linear continuous action of Γ. A *(φ, Γ)-module* D_A over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ is the base change to $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ of a (φ, Γ)-module D^s_A over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ for some $s > 0$.

Notation 1.2. For a morphism $A \to B$ of Banach algebras over $\mathbb{Q}_p$, we denote by D^s_B and D_B the base changes of D^s_A and D_A to $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} B$ and $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} B$, respectively. When $A = S$ is an affinoid algebra over $\mathbb{Q}_p$ and $x \in M(S)$, we denote $D^s_{k(x)}$ and $D_{k(x)}$ by D^s_x and D_x instead.

To define (φ, Γ)-modules over general rigid analytic spaces, one needs to show that φ-modules over affinoid spaces satisfy the gluing property. To this end, we recall the notion of φ-bundles introduced in [Kedlaya et al. 2014]. Let S be an affinoid algebra over $\mathbb{Q}_p$. For $0 < s_1 < s_2$ which are sufficiently large, a vector bundle over $\boldsymbol{B}^{[s_1,s_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S$ is a finite projective module $D^{[s_1,s_2]}_S$ over $\boldsymbol{B}^{[s_1,s_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S$. By the identification of $\boldsymbol{B}^{[s_1,s_2]}_K$ with the ring of rigid analytic functions over the closed annulus $s_1 \le v_p(T) \le s_2$ over K'_0, one may identify $D^{[s_1,s_2]}_S$ with a locally free coherent sheaf over the product of the annulus $s_1 \le v_p(T) \le s_2$ over K'_0 with $M(S)$ in the category of rigid analytic spaces over $\mathbb{Q}_p$. It then follows that vector bundles over $\boldsymbol{B}^{[s_1,s_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S$ satisfy the gluing property for the weak G-topology of $M(S)$. For sufficiently large s, a vector bundle over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ consists of one vector bundle $D^{[s_1,s_2]}_S$ over each ring $\boldsymbol{B}^{[s_1,s_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S$ with $s \le s_1 \le s_2$, together with isomorphisms

$$D^{[s_1,s_2]}_S \otimes_{\boldsymbol{B}^{[s_1,s_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S} \boldsymbol{B}^{[s'_1,s'_2]}_K \widehat{\otimes}_{\mathbb{Q}_p} S \cong D^{[s'_1,s'_2]}_S$$

for all $s \le s'_1 \le s_1 \le s_2 \le s'_2$ satisfying the cocycle conditions. A φ-bundle over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ is a vector bundle $(D^{[s_1,s_2]}_S)_{s \le s_1 \le s_2}$ over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ equipped with isomorphisms $\varphi^* D^{[s_1,s_2]}_S \cong D^{[ps_1,ps_2]}_S$ for all $s \le s_1 \le s_2$ satisfying the obvious compatibility conditions. When s is sufficiently large, by [Kedlaya et al.

2014, Proposition 2.2.7], the natural functor from the category of φ-modules over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ to the category of φ-bundles over $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ is an equivalence of categories. Note that by the gluing property of vector bundles, one can glue φ-bundles $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ over $M(S)$. Therefore this equivalence of categories enables us to glue φ-modules over affinoid spaces.

Definition 1.3. Let X be a rigid analytic space over $\mathbb{Q}_p$. A family of (φ,Γ)-modules D_X over X is a compatible family of (φ,Γ)-modules D_S over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ for each affinoid subdomain $M(S)$ of X. By the gluing property of φ-modules over affinoid spaces, one may view D_X as a sheaf over X for the weak G-topology (which hence extends uniquely to the strong G-topology).

Theorem 1.4. *Let A be a Banach algebra over $\mathbb{Q}_p$, and let V_A be a finite locally free A-linear representation of G_K. Then there is a (φ,Γ)-module $\mathrm{D}^{\dagger}_{\mathrm{rig}}(V_A)$ over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ functorially associated to V_A. The rule $V_A \mapsto \mathrm{D}^{\dagger}_{\mathrm{rig}}(V_A)$ is fully faithful and exact, and it commutes with base change in A.*

Proof. See [Kedlaya and Liu 2010, Theorem 3.11], which generalizes [Berger and Colmez 2008, Théorème 4.2.9]. Note that both the results do not really verify the φ-module condition. This gap is fixed by [Liu 2014, Theorem 1.1.4]. □

Let A be a Banach algebra over K_0. Recall that one has a canonical decomposition

$$A \otimes_{\mathbb{Q}_p} K_0 \cong \prod_{\sigma \in \mathrm{Gal}(K_0/\mathbb{Q}_p)} A_\sigma,$$

where each A_σ is the base change of A by the automorphism σ. Furthermore, the $\mathrm{Gal}(K_0/\mathbb{Q}_p)$-action permutes all the A_σ such that $\tau(A_\sigma) = A_{\tau\sigma}$. For any $a \in A^\times$, we equip $A \otimes_{\mathbb{Q}_p} K_0$ with a $1 \otimes \varphi$-semilinear action φ by setting

$$\varphi((x_1, x_\varphi, \ldots, x_{\varphi^{f-1}})) = (a x_{\varphi^{f-1}}, x_1, \ldots, x_{\varphi^{f-2}}),$$

where φ is the geometric Frobenius and $x_{\varphi^i} \in A_{\varphi^i}$ for each $0 \le i \le f-1$; we denote this φ-module by D_a. It is clear that the φ-action on D_a satisfies $\varphi^f = 1 \otimes a$.

We fix a uniformizer π_K of K.

Definition 1.5. For any continuous character $\delta : K^\times \to A^\times$, we define a rank-1 (φ,Γ)-module $(\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A)(\delta)$ over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A$ as follows. If $\delta|_{\mathcal{O}_K^\times} = 1$, we set

$$(\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A)(\delta) = (\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A) \otimes_{A \otimes_{\mathbb{Q}_p} K_0} D_{\delta(\pi_K)},$$

where we equip $D_{\delta(\pi_K)}$ with the trivial Γ-action. For general δ, we write $\delta = \delta'\delta''$ such that $\delta'(\pi_K) = 1$ and $\delta''|_{\mathcal{O}_K^\times} = \mathrm{id}$. We view δ' as an A-valued character of W_K via the local reciprocity map, and extend it to a character of G_K continuously. We then set

$$(\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A)(\delta) = \mathrm{D}^{\dagger}_{\mathrm{rig}}(\delta') \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A} (\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} A)(\delta'').$$

For any (φ,Γ)-module D_A over $\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} A$, put

$$D_A(\delta)=D_A\otimes_{\boldsymbol{B}^\dagger_{\mathrm{rig},K}\widehat{\otimes}_{\mathbb{Q}_p}A}(\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} A)(\delta).$$

Let X be a rigid analytic space over $\mathbb{Q}_p$. For a continuous character $\delta: K^\times\to\mathscr{O}(X)^\times$ and a family of (φ,Γ)-module D_X over X, we define the families of (φ,Γ)-modules $(\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p}\mathscr{O}_X)(\delta)$ and $D_X(\delta)$ by gluing $(\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S)(\delta)$ and $D_S(\delta)$ for all affinoid subdomains $M(S)$, respectively.

2. Cohomology of families of (φ,Γ)-modules

Let Δ_K be the p-torsion subgroup of Γ. Choose γ_K in Γ_K whose image in Γ/Δ_K is a topological generator.

Definition 2.1. For a (φ,Γ)-module D_S over $\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$, we define the Herr complex $C^\bullet_{\varphi,\gamma_K}(D_S)$ of D_S concentrated in degree $[0,2]$ as

$$C^\bullet_{\varphi,\gamma_K}(D_S)=[D_S^{\Delta_K}\xrightarrow{d_1}D_S^{\Delta_K}\oplus D_S^{\Delta_K}\xrightarrow{d_2}D_S^{\Delta_K}],$$

with $d_1(x)=((\gamma_K-1)x,(\varphi-1)x)$ and $d_2(x,y)=(\varphi-1)x-(\gamma_K-1)y$. One shows that this complex is independent of the choice of γ_K up to canonical S-linear isomorphisms: the isomorphism $C^\bullet_{\varphi,\gamma_K}(D_S)\xrightarrow{\sim}C^\bullet_{\varphi,\gamma'_K}(D_S)$ is given by $[1,1\oplus(\gamma'_K-1)/(\gamma_K-1),(\gamma'_K-1)/(\gamma_K-1)]$. We will denote the cohomology of $C^\bullet_{\varphi,\gamma_K}(D_S)$ by $H^\bullet(D_S)$.

By the main result of [Kedlaya et al. 2014], one knows that $H^i(D_S)$ is a finitely generated S-module. It therefore follows that $H^i(D_S)$ commutes with flat base change in S. That is, if $S\to S'$ is flat, then $H^i(D_S)\otimes_S S'\cong H^i(D_{S'})$. This enables a cohomology theory for families of (φ,Γ)-modules over general rigid analytic spaces.

Definition 2.2. Let X be a rigid analytic space over $\mathbb{Q}_p$, and let D_X be a family of (φ,Γ)-modules over X. We define $H^\bullet(D_X)$ to be the cohomology of the complex of sheaves

$$C^\bullet_{\varphi,\gamma_K}(D_X)=[D_X^{\Delta_K}\xrightarrow{d_1}D_X^{\Delta_K}\oplus D_X^{\Delta_K}\xrightarrow{d_2}D_X^{\Delta_K}]$$

in the category of presheaves over X, with $d_1(x)=((\gamma_K-1)x,(\varphi-1)x)$ and $d_2(x,y)=(\varphi-1)x-(\gamma_K-1)y$. For each affinoid subdomain $M(S)$ of X and $0\le i\le 2$, the module of sections of $H^i(D_X)$ on $M(S)$ is canonically isomorphic to $H^i(D_S)$. Hence $H^i(D_X)$ forms a coherent $\mathscr{O}_X$-module by the flat base change property of $H^i(D_S)$.

As a consequence of finiteness of the cohomology of families of (φ, Γ)-modules, by a standard argument we see that, locally on X, the complex $C^\bullet_{\varphi,\gamma_K}(D_X)$ is quasi-isomorphic to a complex of locally free coherent sheaves concentrated in degree $[0, 2]$. This enables us to flatten the cohomology of families of (φ, Γ)-modules by blowing up the base X. The following lemma is a rearrangement of some arguments in [Kedlaya et al. 2014, §6.3]:

Lemma 2.3. *Let X be a reduced and irreducible rigid analytic space over F, and let D_X be a family of (φ, Γ)-modules of rank d over X. Then the following statements are true*:

(1) *There exists a proper birational morphism $\pi : X' \to X$ of reduced rigid analytic spaces over F so that $H^0(D_{X'})$ is flat and $H^i(D_{X'})$ has Tor-dimension ≤ 1 for each $i = 1, 2$.*

(2) *Suppose that D'_X is a family of (φ, Γ)-modules over X of rank d', and that $\lambda : D'_X \to D_X$ is a morphism between them so that for any $x \in X$ the image of λ_x is a (φ, Γ)-submodule of rank d of D_x. Then there exists a proper birational morphism $\pi : X' \to X$ of reduced rigid analytic spaces over F so that the cokernel of $\pi^*\lambda$ has Tor-dimension ≤ 1.*

Proof. The upshot is that for a bounded complex $(C^\bullet, d^\bullet)$ of locally free coherent sheaves on X, there exists a blow-up $\pi : X' \to X$, which depends only on the quasi-isomorphism class of $(C^\bullet, d^\bullet)$, so that $\pi^* d^i$ has flat image for each i. Furthermore, the construction of X' commutes with dominant base change in X (see [Kedlaya et al. 2014, Corollary 6.3.6] for more details). Thus for (1), we can construct X' locally and then glue. For (2), let Q_X denote the cokernel of λ. For any $x \in X$, since the image of λ_x is a (φ, Γ)-submodule of rank d, by [Liu 2014, Lemma 5.3.1], we get that Q_x is killed by a power of t. Now let $M(S)$ be an affinoid subdomain of X, and suppose that D^s_S and D'^s_S are defined for some suitable $s > 0$. For $r > s$, set $Q^{[s,r]}_S$ to be $D^{[s,r]}_S / \lambda(D'^{[s,r]}_S)$. Since for any $x \in M(S)$ the fiber of $Q^{[s,r]}_S$ at x is killed by a power of t, we get that $Q^{[s,r]}_S$ is killed by t^k for some $k > 0$. This yields that $Q^{[s,r]}_S$ is a finite S-module. Now we apply [Kedlaya et al. 2014, Corollary 6.3.6] to a finite presentation of $Q^{[s,ps]}_S$ to get a blow-up Y of $M(S)$ so that the pullback of $Q^{[s,ps]}_S$ has Tor-dimension ≤ 1. Using the fact that $(\varphi^n)^* Q^{[s,ps]}_S \cong Q^{[p^n s, p^{n+1} s]}_S$, we see that Y is also the blow-up obtained by applying [Kedlaya et al. 2014, Corollary 6.3.6] to a finite presentation of $Q^{[s,p^{n+1}s]}_S$ for any positive integer n. It therefore follows that for any $r > s$ the pullback of $Q^{[s,r]}_S$ has Tor-dimension ≤ 1; hence the pullback of Q_S has Tor-dimension ≤ 1. Furthermore, the blow-ups for all affinoid subdomains $M(S)$ glue to form a blow-up X' of X which satisfies the desired condition. $\square$

Lemma 2.4. *Let X be a reduced and irreducible rigid analytic space over F. Let D'_X and D_X be families of (φ,Γ)-modules over X of ranks d' and d respectively, and let $\lambda : D'_X \to D_X$ be a morphism between them. Suppose that for any $x \in X$ the image of λ_x is a (φ,Γ)-submodule of rank d of D_x. Then there exists a proper birational morphism $\pi : X' \to X$ of reduced rigid analytic spaces over F such that the kernel of $\pi^*\lambda$ is a family of (φ,Γ)-modules of rank $d'-d$ over X', and there exists a Zariski-open dense subset $U \subset X'$ such that $(\ker(\pi^*\lambda))_x = \ker((\pi^*\lambda)_x)$ for any $x \in U$.*

Proof. Let Q_X be the cokernel of λ. By Lemma 2.3, we may suppose that Q_X has Tor-dimension ≤ 1 after adapting X. Now let P_X denote the kernel of λ. For any $x \in X$, the Tor spectral sequence computing the cohomology of the complex $[D_X \xrightarrow[\lambda]{} D'_X] \otimes^{\boldsymbol{L}}_{\mathscr{O}_X} k(x)$ gives rise to a short exact sequence

$$0 \longrightarrow P_x \longrightarrow \ker(\lambda_x) \longrightarrow \mathrm{Tor}_1(Q_X, k(x)) \longrightarrow 0.$$

Since the image of λ_x is a (φ,Γ)-module of rank d, $\ker(\lambda_x)$ is a (φ,Γ)-module of rank $d'-d$. Since Q_X is killed by a power of t locally on X, we get that the last term of the exact sequence is killed by a power of t. This yields that P_x is a (φ,Γ)-module of rank $d'-d$. We therefore conclude that P_X is a family of (φ,Γ)-modules of rank $d'-d$ over X by [Kedlaya et al. 2014, Corollary 2.1.9]. Furthermore, since Q_X has Tor-dimension ≤ 1, by [Kedlaya et al. 2014, Lemma 6.3.7] we get that the set of $x \in X$ for which $\mathrm{Tor}_1(Q_X, k(x)) \neq 0$ forms a nowhere-dense Zariski-closed subset of X; this yields the rest of the lemma. ☐

The following proposition modifies part of [Kedlaya et al. 2014, Theorem 6.3.9]:

Proposition 2.5. *Let X be a reduced and irreducible rigid analytic space over F. Let D_X be a family of (φ,Γ)-modules of rank d over X, and let $\delta : K^\times \to \mathscr{O}(X)^\times$ be a continuous character. Suppose that there exists a Zariski-dense subset Z of closed points of X and a positive integer $c \leq d$ such that, for every $z \in Z$, $H^0(D_z^\vee(\delta_z))$ is a c-dimensional $k(z)$-vector space. Then there exists a proper birational morphism $\pi : X' \to X$ of reduced rigid analytic spaces over F and a morphism $\lambda : D_{X'} \to M_{X'} = (\boldsymbol{B}^\dagger_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} \mathscr{O}_{X'})(\delta) \otimes_{\mathscr{O}_{X'}} L$ of (φ,Γ)-modules, where L is a locally free coherent $\mathscr{O}_{X'}$-module of rank c equipped with trivial (φ,Γ)-actions, such that:*

(1) *For any $x \in X'$, the image of λ_x is a (φ,Γ)-submodule of rank c.*

(2) *The kernel of λ is a family of (φ,Γ)-modules of rank $d-c$ over X', and there exists a Zariski-open dense subset $U \subset X'$ such that $(\ker\lambda)_x = \ker(\lambda_x)$ for any $x \in U$.*

Proof. Using Lemma 2.3, we first choose a proper birational morphism $\pi : X' \to X$ with X' reduced such that $N_{X'} = \pi^*(D_X^\vee(\delta))$ satisfies the conditions that $H^0(N_{X'})$ is flat and $H^i(N_{X'})$ has Tor-dimension ≤ 1 for each $i = 1, 2$. Then, for any $x \in X'$, the base change spectral sequence $E_2^{i,j} = \mathrm{Tor}_{-i}(H^j(N_{X'}), k(x)) \Rightarrow H^{i+j}(N_x)$ gives a short exact sequence (using that $H^1(N_{X'})$ has Tor-dimension ≤ 1 and $N_{X'}$ is flat)

$$0 \longrightarrow H^0(N_{X'}) \otimes_{\mathbb{O}_{X'}} k(x) \longrightarrow H^0(N_x) \longrightarrow \mathrm{Tor}_1(H^1(N_{X'}), k(x)) \longrightarrow 0.$$

Since $H^1(N_{X'})$ has Tor-dimension ≤ 1, by [Kedlaya et al. 2014, Lemma 6.3.7] the set of $x \in X'$ for which the last term of the above exact sequence does not vanish forms a nowhere-dense Zariski-closed subset V. For any $z \in \pi^{-1}(Z) \setminus V$, we deduce that $H^0(N_{X'}) \otimes_{\mathbb{O}_{X'}} k(z)$ is a c-dimensional $k(z)$-vector space. Since $H^0(N_{X'})$ is flat and $\pi^{-1}(Z) \setminus V$ is a Zariski-dense subset of X', we get that $H^0(N_{X'})$ is locally free of constant rank c. Let L be its dual coherent $\mathbb{O}_{X'}$-module. Then the natural map $(\boldsymbol{B}^\dagger_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} \mathbb{O}_{X'})H^0(N_{X'}) \to N_{X'}$ gives rise to a map

$$\lambda : D_{X'} \to M_{X'} = (\boldsymbol{B}^\dagger_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} \mathbb{O}_{X'})(\delta) \otimes_{\mathbb{O}_{X'}} L.$$

For any $x \in X'$, since the map $H^0(N_{X'}) \otimes_{\mathbb{O}_{X'}} k(x) \to H^0(N_x)$ is injective, we get that the image of λ_x is a rank-c (φ, Γ)-submodule of M_x. We thus conclude the proposition using Lemma 2.4. □

3. Families of Hodge–Tate (φ, Γ)-modules

From now on, let S be an affinoid algebra over F. Recall that for any $n \geq n(s)$ there is a continuous Γ-equivariant injective map

$$\iota_n : \boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \to K_n[\![t]\!].$$

It is defined as the composite

$$\boldsymbol{B}^{\dagger,s}_K \subset \widetilde{\boldsymbol{B}}^{\dagger,s} \xrightarrow{\varphi^{-n}} \widetilde{\boldsymbol{B}}^{\dagger,p^{-n}s} \subset \widetilde{\boldsymbol{B}}^+ \subset \boldsymbol{B}^+_{\mathrm{dR}},$$

and it factors through $K_n[\![t]\!]$ (see [Berger 2002, §2] for more details about ι_n). In particular, we have $\iota_{n+1} \circ \varphi = \iota_n$. The map ι_n induces a continuous Γ-equivariant map

$$\iota_n : \boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S \to K_n[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S.$$

Definition 3.1. Let D_S be a (φ, Γ)-module of rank d over $\boldsymbol{B}^\dagger_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$. For any positive integer n, if $D_S^{r_n}$ is defined, then for any $0 < s \leq r_n$ we set

$$\mathrm{D}^{+,K_n}_{\mathrm{dif}}(D_S) = D_S^s \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S, \iota_n} (K_n[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S)$$

and

$$\mathrm{D}^{K_n}_{\mathrm{dif}}(D_S) = \mathrm{D}^{+,K_n}_{\mathrm{dif}}(D_S)[1/t].$$

We also denote the natural map

$$D_S^s \to \mathrm{D}_{\mathrm{dif}}^{+,K_n}(D_S)$$

by ι_n, and call it the *localization map*. Define $\mathrm{D}_{\mathrm{Sen}}^{K_n}(D_S) = \mathrm{D}_{\mathrm{dif}}^{+,K_n}(D_S)/(t)$. For $D_S = \mathrm{D}_{\mathrm{rig}}^{\dagger}(V_S)$ coming from a finite locally free S-linear representation V_S, we write $\mathrm{D}_{\mathrm{dif}}^{+,K_n}(V_S)$ and $\mathrm{D}_{\mathrm{Sen}}^{K_n}(V_S)$ for $\mathrm{D}_{\mathrm{dif}}^{+,K_n}(D_S)$ and $\mathrm{D}_{\mathrm{Sen}}^{K_n}(D_S)$ respectively. When the base field is clear from the context, we write $\mathrm{D}_{\mathrm{dif}}^{+,n}(D_S)$ and $\mathrm{D}_{\mathrm{Sen}}^{n}(D_S)$ instead of $\mathrm{D}_{\mathrm{dif}}^{+,K_n}(D_S)$ and $\mathrm{D}_{\mathrm{Sen}}^{K_n}(D_S)$ for simplicity.

Definition 3.2. We call D_S *Hodge–Tate with Hodge–Tate weights in* $[a,b]$ if there exists a positive integer n such that the natural map

$$\Big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))\Big)^{\Gamma} \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S)[t,t^{-1}] \longrightarrow \bigoplus_{i\in\mathbb{Z}} \mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i)) \tag{3.2.1}$$

is an isomorphism. We denote by $h_{\mathrm{HT}}(D_S)$ the smallest n which satisfies this condition, and we define $D_{\mathrm{HT}}(D_S) = \big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i))\big)^{\Gamma}$.

Lemma 3.3. *Let D_S be a Hodge–Tate (φ,Γ)-module over $\boldsymbol{B}_{\mathrm{rig},K}^{\dagger} \widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a,b]$. Then for any $n \ge h_{\mathrm{HT}}(D_S)$ (3.2.1) is an isomorphism and $\mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))^{\Gamma} = \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i))^{\Gamma}$ for any $i \in [a,b]$. As a consequence, we have $\big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))\big)^{\Gamma} = D_{\mathrm{HT}}(D_S)$.*

Proof. Tensoring with $K_n \otimes_{\mathbb{Q}_p} S[t,1/t]$ on both sides of the map

$$\Big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i))\Big)^{\Gamma} \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_{h_{\mathrm{HT}}(D_S)} \otimes_{\mathbb{Q}_p} S)[t,t^{-1}] \longrightarrow \bigoplus_{i\in\mathbb{Z}} \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i)),$$

we get that the natural map

$$\Big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i))\Big)^{\Gamma} \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S)[t,t^{-1}] \longrightarrow \bigoplus_{i\in\mathbb{Z}} \mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))$$

is an isomorphism. Taking Γ-invariants on both sides, we get

$$\Big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{h_{\mathrm{HT}}(D_S)}(D_S(-i))\Big)^{\Gamma} = \Big(\bigoplus_{a\le i\le b} \mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))\Big)^{\Gamma}. \qquad \square$$

Remark 3.4. If D_S is Hodge–Tate with weights in $[a,b]$, then, by taking Γ-invariants on both sides of (3.2.1), we see that $\mathrm{D}_{\mathrm{Sen}}^{n}(D_S(-i))^{\Gamma} = 0$ for any $n \ge h_{\mathrm{HT}}(D_S)$ and $i \notin [a,b]$.

Lemma 3.5. *If D_S is a Hodge–Tate (φ,Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a,b]$, then for any morphism $S\to R$ of affinoid algebras over K, D_R is Hodge–Tate with weights in $[a,b]$ and $h_{\mathrm{HT}}(D_R)\le h_{\mathrm{HT}}(D_S)$. Furthermore, the natural map*

$$\mathrm{D}^n_{\mathrm{Sen}}(D_S(i))^{\Gamma}\otimes_S R\to \mathrm{D}^n_{\mathrm{Sen}}(D_R(i))^{\Gamma}$$

is an isomorphism for any $i\in\mathbb{Z}$ and $n\ge h_{\mathrm{HT}}(D_S)$. As a consequence, the natural map $D_{\mathrm{HT}}(D_S)\otimes_S R\to D_{\mathrm{HT}}(D_R)$ is an isomorphism.

Proof. Let $n\ge h_{\mathrm{HT}}(D_S)$. Tensoring with R over S on both sides of (3.2.1), we get that the natural map

$$\Bigl(\bigoplus_{a\le i\le b}\mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma}\otimes_S R\Bigr)\otimes_{K\otimes_{\mathbb{Q}_p}R}(K_n\otimes_{\mathbb{Q}_p}R)[t,t^{-1}]\longrightarrow\bigoplus_{i\in\mathbb{Z}}\mathrm{D}^n_{\mathrm{Sen}}(D_R(-i))$$

is an isomorphism. Comparing Γ-invariants on both sides, we get that the natural map

$$\mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma}\otimes_S R\to \mathrm{D}^n_{\mathrm{Sen}}(D_R(-i))^{\Gamma}$$

is an isomorphism for any $a\le i\le b$. This implies that the natural map

$$\Bigl(\bigoplus_{a\le i\le b}\mathrm{D}^n_{\mathrm{Sen}}(D_R(-i))^{\Gamma}\Bigr)\otimes_{K\otimes_{\mathbb{Q}_p}R}(K_n\otimes_{\mathbb{Q}_p}R)[t,t^{-1}]\longrightarrow\bigoplus_{i\in\mathbb{Z}}\mathrm{D}^n_{\mathrm{Sen}}(D_R(-i)).$$

is an isomorphism. ☐

Corollary 3.6. *If D_S is a Hodge–Tate (φ,Γ)-module of rank d over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$, then $D_{\mathrm{HT}}(D_S)$ is a locally free coherent $K\otimes_{\mathbb{Q}_p}S$-module of rank d.*

Proof. By the previous lemma, it suffices to treat the case that S is a finite extension of K; this is clear from the isomorphism (3.2.1). ☐

Definition 3.7. Let X be a rigid analytic space over F, and let D_X be a family of (φ,Γ)-modules of rank d over X. We call D_X *Hodge–Tate* with weights in $[a,b]$ if for some (hence any) admissible cover $\{M(S_i)\}_{i\in I}$ of X, D_{S_i} is Hodge–Tate with weights in $[a,b]$ for any $i\in I$. We define $D_{\mathrm{HT}}(D_X)$ to be the gluing of all the $D_{\mathrm{HT}}(D_{S_i})$.

Lemma 3.8. *Let D_S be a (φ,Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$. Then (3.2.1) is an isomorphism if and only if the natural map*

$$\bigoplus_{a\le i\le b}\mathrm{D}^n_{\mathrm{Sen}}(D_S)^{\Gamma_n=\chi^i}\longrightarrow \mathrm{D}^n_{\mathrm{Sen}}(D_S) \tag{3.8.1}$$

is an isomorphism. Furthermore, if this is the case, then (3.2.1) holds for n.

Proof. For the "$\Rightarrow$" part, since (3.2.1) is an isomorphism, we deduce that

$$\mathrm{D}^n_{\mathrm{Sen}}(D_S) = \bigoplus_{a \le i \le b} t^i \cdot \mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma} \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S). \tag{3.8.2}$$

Note that $t^i \cdot \mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma} \subseteq \mathrm{D}^n_{\mathrm{Sen}}(D_S)^{\Gamma_n = \chi^i}$. Hence (3.8.2) implies that (3.8.1) is surjective. On the other hand, it is clear that (3.2.1) is injective; hence it is an isomorphism. Conversely, suppose that (3.8.1) is an isomorphism. Note that

$$\begin{aligned} \mathrm{D}^n_{\mathrm{Sen}}(D_S)^{\Gamma_n = \chi^i} &= t^i \cdot \mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma_n} \\ &= (t^i \cdot \mathrm{D}^n_{\mathrm{Sen}}(D_S(-i))^{\Gamma}) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S), \end{aligned}$$

where the latter equality follows from [Berger and Colmez 2008, Proposition 2.2.1]. This implies that D_S satisfies (3.8.2), yielding that D_S satisfies (3.2.1). ☐

Proposition 3.9. *Let S be reduced, and let D_S be a (φ, Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes} S$. Suppose that there exists a Zariski-dense subset $Z \subset M(S)$ such that D_z is Hodge–Tate with weights in $[a, b]$ for any $z \in Z$ and $\sup_{z \in Z}\{h_{\mathrm{HT}}(D_z)\} < \infty$. Then D_S is Hodge–Tate with weights in $[a, b]$.*

Proof. Let $n \ge \sup_{z \in Z}\{h_{\mathrm{HT}}(D_z)\}$ such that D^n_S is defined, and let γ be a topological generator of Γ_n. For any $a \le i \le b$, let p_i denote the operator

$$\prod_{a \le j \le b, j \ne i} \frac{\gamma - \chi^j(\gamma)}{\chi^i(\gamma) - \chi^j(\gamma)},$$

and let $M_i = p_i(\mathrm{D}^n_{\mathrm{Sen}}(D_S))$. It is clear that p_i is the identity on $\mathrm{D}^n_{\mathrm{Sen}}(D_S)^{\Gamma_n = \chi^i}$; hence $\mathrm{D}^n_{\mathrm{Sen}}(D_S)^{\Gamma_n = \chi^i} \subseteq M_i$. On the other hand, for any $z \in Z$, since D_z is Hodge–Tate with weights in $[a, b]$ and $h_{\mathrm{HT}}(D_z) \le n$, we deduce from Lemma 3.8 that $p_i(\mathrm{D}^n_{\mathrm{Sen}}(D_z)) = \mathrm{D}^n_{\mathrm{Sen}}(D_z)^{\Gamma_n = \chi^i}$. This implies that M_i maps onto $\mathrm{D}^n_{\mathrm{Sen}}(D_z)^{\Gamma_n = \chi^i}$ under the specialization $\mathrm{D}^n_{\mathrm{Sen}}(D_S) \to \mathrm{D}^n_{\mathrm{Sen}}(D_z)$. Since S is reduced and Z is Zariski-dense, we obtain $M_i \subseteq \mathrm{D}^n_{\mathrm{Sen}}(D)^{\Gamma_n = \chi^i}$; hence $M_i = \mathrm{D}^n_{\mathrm{Sen}}(D)^{\Gamma_n = \chi^i}$. Let $M = \bigoplus_{a \le i \le b} M_i$. We claim that the natural inclusion $M \subseteq \mathrm{D}^n_{\mathrm{Sen}}(D_S)$ is an isomorphism. In fact, for any $z \in Z$, since $\mathrm{D}^n_{\mathrm{Sen}}(D_z) = \bigoplus_{a \le i \le b} \mathrm{D}^n_{\mathrm{Sen}}(D_z)^{\Gamma_n = \chi^i}$, we have that M maps onto $\mathrm{D}^n_{\mathrm{Sen}}(D_z)$. Thus $\mathrm{D}^n_{\mathrm{Sen}}(D_S)/M$ vanishes at z. We therefore conclude that $\mathrm{D}^n_{\mathrm{Sen}}(D_S)/M = 0$ because S is reduced and Z is Zariski-dense. By Lemma 3.8 and the claim, we conclude that D_S is Hodge–Tate with weights in $[a, b]$. ☐

4. Families of de Rham (φ, Γ)-modules

Definition 4.1. Let D_S be a (φ, Γ)-module of rank d over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$. For any positive integer n, if $D^{r_n}_S$ is defined, then we equip $\mathrm{D}^n_{\mathrm{dif}}(D_S)$ with the filtration

$\mathrm{Fil}^i\,\mathrm{D}^n_{\mathrm{dif}}(D_S) = t^i\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$. We call D_S *de Rham with weights in* $[a,b]$ if there exists a positive integer n such that:

(1) The natural map

$$\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S)[\![t]\!][1/t] \longrightarrow \mathrm{D}^n_{\mathrm{dif}}(D_S) \tag{4.1.1}$$

is an isomorphism.

(2) $\mathrm{Fil}^{-b}(\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma) = D_S$ and $\mathrm{Fil}^{-a+1}(\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma) = 0$, where $\mathrm{Fil}^i(\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma)$ is the induced filtration on $\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma$.

We denote the smallest n satisfying these conditions by $h_{\mathrm{dR}}(D_S)$, and we set $D_{\mathrm{dR}}(D_S) = \mathrm{D}^{h_{\mathrm{dR}}(D_S)}_{\mathrm{dif}}(D_S)^\Gamma$.

Lemma 4.2. *Let D be a de Rham (φ,Γ)-module over $\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$. Then, for any $n \geq h_{\mathrm{dR}}(D_S)$, $\mathrm{D}^n_{\mathrm{dif}}(D_S)^\Gamma = D_{\mathrm{dR}}(D_S)$.*

Proof. We tensor with $K_{n+1} \otimes_{\mathbb{Q}_p} S[\![t]\!][1/t]$ on both sides of the map

$$\mathrm{D}^{h_{\mathrm{dR}}(D_S)}_{\mathrm{dif}}(D_S)^\Gamma \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_{h_{\mathrm{dR}}(D_S)} \otimes_{\mathbb{Q}_p} S)[\![t]\!][1/t] \longrightarrow \mathrm{D}^{h_{\mathrm{dR}}(D_S)}_{\mathrm{dif}}(D_S),$$

yielding that the map

$$\mathrm{D}^{h_{\mathrm{dR}}(D_S)}_{\mathrm{dif}}(D_S)^\Gamma \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S)[\![t]\!][1/t] \longrightarrow \mathrm{D}^n_{\mathrm{dif}}(D_S)$$

is an isomorphism. Comparing Γ-invariants on both sides, we get the desired result. □

Lemma 4.3. *If D is a de Rham (φ,Γ)-module of rank d over $\boldsymbol{B}^\dagger_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a,b]$, then D is Hodge–Tate with weights in $[a,b]$ and $h_{\mathrm{HT}}(D_S) \leq h_{\mathrm{dR}}(D_S)$. Furthermore, we have $\mathrm{Gr}^i\, D_{\mathrm{dR}}(D_S) = \mathrm{D}^n_{\mathrm{Sen}}(D_S(i))^\Gamma$ under the identification $\mathrm{Gr}^i\,\mathrm{D}^n_{\mathrm{dif}}(D_S) = \mathrm{D}^n_{\mathrm{Sen}}(D_S(i))$ for any $n \geq h_{\mathrm{dR}}(D_S)$.*

Proof. Let $n \geq h_{\mathrm{dR}}(D_S)$. Since (4.1.1) is an isomorphism, we deduce that the natural map of graded modules

$$\bigoplus_{i\in\mathbb{Z}} \mathrm{Gr}^i\, D_{\mathrm{dR}}(D_S) \otimes_{K\otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} S)[t,t^{-1}] \longrightarrow \bigoplus_{i\in\mathbb{Z}} \mathrm{D}^n_{\mathrm{Sen}}(D_S(i)) \tag{4.3.1}$$

is surjective. On the other hand, since $t^i \cdot \mathrm{Gr}^{-i}\, D_{\mathrm{dR}}(D_S) \subset \mathrm{D}^n_{\mathrm{Sen}}(D_S)$, we have that the natural map

$$\bigoplus_{a\leq i\leq b} t^i \cdot \mathrm{Gr}^{-i}\, D_{\mathrm{dR}}(D_S) \longrightarrow \mathrm{D}^n_{\mathrm{Sen}}(D_S)$$

is injective. This implies that (4.3.1) is injective; hence it is an isomorphism. Comparing Γ-invariants on both sides, we get $\mathrm{Gr}^i\, D_{\mathrm{dR}}(D_S) = \mathrm{D}^n_{\mathrm{Sen}}(D_S(i))^\Gamma$ for each $i \in \mathbb{Z}$. □

Lemma 4.4. *If D_S is a de Rham (φ,Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$, then, for any morphism $S\to R$ of affinoid algebras over K, D_R is de Rham with weights in $[a,b]$ and $h_{\mathrm{dR}}(D_R)\le h_{\mathrm{dR}}(D_S)$. Furthermore, the natural maps $\mathrm{Fil}^i\, \mathrm{D}_{\mathrm{dR}}(D_S)\otimes_S R\to \mathrm{Fil}^i\, \mathrm{D}_{\mathrm{dR}}(D_R)$ are isomorphisms for all $i\in\mathbb{Z}$.*

Proof. Let $n\ge h_{\mathrm{dR}}(D_S)$. Tensoring with $(K_n\otimes_{\mathbb{Q}_p} R)[[t]][1/t]$ on both sides of (4.1.1), we get that the natural map

$$(\mathrm{D}^n_{\mathrm{dif}}(D_S)^{\Gamma}\otimes_S R)\otimes_{K\otimes_{\mathbb{Q}_p}R}(K_n\otimes_{\mathbb{Q}_p} R)[[t]][1/t]\longrightarrow \mathrm{D}^n_{\mathrm{dif}}(D_R). \tag{4.4.1}$$

is an isomorphism. Comparing Γ-invariants on both sides of (4.4.1), we get that the natural map $\mathrm{D}^n_{\mathrm{dif}}(D_S)^{\Gamma}\otimes_S R\to \mathrm{D}^n_{\mathrm{dif}}(D_R)^{\Gamma}$ is an isomorphism; hence D_R is de Rham. Then, by Lemmas 3.5 and 4.3, we deduce that the natural map $\mathrm{Gr}^i(\mathrm{D}_{\mathrm{dR}}(D_S))\otimes_S R\to \mathrm{Gr}^i(\mathrm{D}_{\mathrm{dR}}(D_R))$ is an isomorphism. This implies the rest of the lemma. $\square$

Corollary 4.5. *If D_S is a de Rham (φ,Γ)-module of rank d over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$, then $\mathrm{D}_{\mathrm{dR}}(D_S)$ is a locally free coherent $K\otimes_{\mathbb{Q}_p} S$-module of rank d.*

Proof. We first note that, for each $i\in\mathbb{Z}$, $\mathrm{Gr}^i(\mathrm{D}_{\mathrm{dR}}(D_S))$, which is isomorphic to $\mathrm{D}^n_{\mathrm{Sen}}(D_S(i))^{\Gamma}$ by Lemma 4.3, is a coherent $K\otimes_{\mathbb{Q}_p} S$-module. We then deduce that $\mathrm{D}_{\mathrm{dR}}(D_S)$ is a coherent $K\otimes_{\mathbb{Q}_p} S$-module. Using Lemma 4.4, it then suffices to treat the case that S is a finite extension of K; this follows easily from the isomorphism (4.1.1). $\square$

Definition 4.6. Let X be a rigid analytic space over F, and let D_X be a family of (φ,Γ)-modules of rank d over X. We call D_X *de Rham* if for some (hence any) admissible cover $\{M(S_i)\}_{i\in I}$ of X, D_{S_i} is de Rham with weights in $[a,b]$ for any $i\in I$. We define $\mathrm{D}_{\mathrm{dR}}(D_X)$ to be the gluing of all the $\mathrm{D}_{\mathrm{dR}}(D_{S_i})$.

Lemma 4.7. *If D_S is a de Rham (φ,Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$ of rank d with weights in $[a,b]$, then*

$$t^{-a}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)\subset \mathrm{D}_{\mathrm{dR}}(D_S)\otimes_{K\otimes_{\mathbb{Q}_p}S}(K_n\otimes_{\mathbb{Q}_p} S)[[t]]\subset t^{-b}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$$

for any $n\ge h_{\mathrm{dR}}(D_S)$.

Proof. Since $\mathrm{Gr}^{-b}\,\mathrm{D}_{\mathrm{dR}}(D_S)=\mathrm{D}_{\mathrm{dR}}(D_S)$, we get $\mathrm{D}_{\mathrm{dR}}(D_S)\subset t^{-b}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$; hence $\mathrm{D}_{\mathrm{dR}}(D_S)\otimes_{K\otimes_{\mathbb{Q}_p}S}(K_n\otimes_{\mathbb{Q}_p} S)[[t]]\subset t^{-b}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$. By the proof of Lemma 4.3, we know that the natural map (4.3.1) is an isomorphism of graded modules. By the facts that $\mathrm{Gr}^i\,\mathrm{D}_{\mathrm{dR}}(D_S)=0$ for $i\ge -a+1$ and $\mathrm{Fil}^i\,\mathrm{D}^n_{\mathrm{dif}}(D_S)$ is t-adically complete, we thus deduce that $t^{-a}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)\subset \mathrm{D}_{\mathrm{dR}}(D_S)\otimes_{K\otimes_{\mathbb{Q}_p}S}(K_n\otimes_{\mathbb{Q}_p} S)[[t]]$. $\square$

Lemma 4.8. *Let D_S be a Hodge–Tate (φ,Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\,\widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a,b]$. Then, for any $k\ge b-a+1$, $i\in[a,b]$, $n\ge h_{\mathrm{HT}}(D_S)$ and $\gamma\in\Gamma_n$, the map $\gamma-\chi^i(\gamma): t^k\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)\to t^k\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ is bijective.*

Proof. Since $\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ is t-adically complete, it suffices to show that

$$\gamma-1 : t^k\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)/t^{k+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \to t^k\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)/t^{k+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$$

is bijective for any $k \geq b-a+1$. Note that $t^k\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)/t^{k+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ is isomorphic to $\mathrm{D}^n_{\mathrm{Sen}}(D_S(k))$ as a Γ-module. Furthermore, note that $\mathrm{D}^n_{\mathrm{Sen}}(D_S(k)) = \bigoplus_{a\leq j\leq b}(\mathrm{D}^n_{\mathrm{Sen}}(D_S))^{\Gamma_n=\chi^{j+k}}$ by Lemma 3.8. Since $j+k \geq b+1$ for all $j \in [a,b]$, we deduce that $\gamma-\chi^i(\gamma)$ is bijective on $\mathrm{D}^n_{\mathrm{Sen}}(D_S(k))$. □

Lemma 4.9. *Let D_S be a Hodge–Tate (φ,Γ)-module over $\boldsymbol{B}^\dagger_{\mathrm{rig},K}\widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a,b]$. Then D_S is de Rham if and only if there exists a positive integer $n \geq h_{\mathrm{HT}}(D_S)$ such that $\prod_{i=a}^{2b-a}(\gamma-\chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$. Furthermore, if this is the case, then (4.1.1) holds for n.*

Proof. Suppose that D_S is de Rham. Let $n \geq h_{\mathrm{dR}}(D_S)$, and put

$$N = D_{\mathrm{dR}}(D_S)\otimes_{K\otimes_{\mathbb{Q}_p}S}(K_n\otimes_{\mathbb{Q}_p}S)[\![t]\!].$$

Since D has weights in $[a,b]$, by Lemma 4.7 we have $t^{-a}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \subset N \subset t^{-b}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$. On the other hand, by the construction of N, it is clear that $(\gamma-1)N \subset tN$. It therefore follows that

$$\begin{aligned}\prod_{i=a}^{2b-a}(\gamma-\chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) &\subset \prod_{i=a}^{2b-a}(\gamma-\chi(\gamma)^i)(t^aN)\\ &\subset t^{2b-a+1}N \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S).\end{aligned}$$

Now suppose $\prod_{i=a}^{2b-a}(\gamma-\chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ for some $n \geq h_{\mathrm{HT}}(D_S)$. We claim that for any $j \in [a,b]$ and $a \in (\mathrm{D}^n_{\mathrm{Sen}}(D_S))^{\Gamma_n=\chi^j}$, we can lift a to an element in $(\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S))^{\Gamma_n=\chi^j}$. In fact, let $\tilde{a}$ be any lift of a in $\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$, and let

$$\tilde{b} = \prod_{\substack{a\leq i\leq 2b-a\\ i\neq j}} \frac{\gamma-\chi^i(\gamma)}{\chi^j(\gamma)-\chi^i(\gamma)}\tilde{a},$$

where γ is a topological generator of Γ_n; it is clear that $\tilde{b}$ is also a lift of a. Furthermore, by assumption, we have $(\gamma-\chi^j(\gamma))(\tilde{b}) \in \prod_{i=a}^{2b-a}(\gamma-\chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$. By the previous lemma, we choose some $\tilde{c} \in t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ satisfying $(\gamma-\chi^j(\gamma))(\tilde{b}) = (\gamma-\chi^j(\gamma))(\tilde{c})$. It is then clear that $\tilde{b}-\tilde{c}$ is a desired lift of a. Since $\mathrm{D}^n_{\mathrm{Sen}}(D_S) = \bigoplus_{a\leq i\leq b}(\mathrm{D}^n_{\mathrm{Sen}}(D_S))^{\Gamma_n=\chi^i}$, we have that $(\mathrm{D}^n_{\mathrm{Sen}}(D_S))^{\Gamma_n=\chi^i}$ is locally free for each $i \in [a,b]$. By shrinking $M(S)$, we may further suppose that each $(\mathrm{D}^n_{\mathrm{Sen}}(D_S))^{\Gamma_n=\chi^i}$ is free. We then deduce from the claim that there exists a free $K_n\otimes_{\mathbb{Q}_p}S$-module $M \subseteq (\mathrm{D}^n_{\mathrm{dif}}(D_S))^{\Gamma_n}$ such that the natural map

$$M\otimes_{K_n\otimes_{\mathbb{Q}_p}S}(K_n\otimes_{\mathbb{Q}_p}S)[\![t]\!][1/t] \longrightarrow \mathrm{D}^n_{\mathrm{dif}}(D_S)$$

is an isomorphism. It follows that the natural map

$$M^{\Gamma} \otimes_{K\otimes_{\mathbb{Q}_p}S} (K_n \otimes_{\mathbb{Q}_p} S)[\![t]\!][1/t] \longrightarrow \mathrm{D}^{n}_{\mathrm{dif}}(D_S)$$

is an isomorphism because $M = M^{\Gamma} \otimes_{K\otimes_{\mathbb{Q}_p}S} (K_n \otimes_{\mathbb{Q}_p} S)$ by [Berger and Colmez 2008, Proposition 2.2.1]. Taking Γ-invariants on both sides, we get $M^{\Gamma} = (\mathrm{D}^{n}_{\mathrm{dif}}(D_S))^{\Gamma}$. This implies that D_S is de Rham. □

Proposition 4.10. *Let S be reduced, and let D_S be a (φ, Γ)-module over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K}\widehat{\otimes}S$. Suppose that there exists a Zariski-dense subset $Z \subset M(S)$ such that D_z is de Rham with weights in $[a, b]$ for any $z \in Z$ and $\sup_{z\in Z}\{h_{\mathrm{dR}}(D_z)\} < \infty$. Then D_S is de Rham with weights in $[a, b]$.*

Proof. By Proposition 3.9, we first have that D_S is Hodge–Tate with weights in $[a, b]$. Let $n \geq \max\{h_{\mathrm{HT}}(D_S), \sup_{z\in Z}\{h_{\mathrm{dR}}(D_z)\}\}$. By Lemma 4.9, we have

$$\prod_{i=a}^{2b-a} (\gamma - \chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_z) \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_z)$$

for any $z \in Z$. This implies $\prod_{i=a}^{2b-a} (\gamma - \chi(\gamma)^i)\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S) \subset t^{b-a+1}\mathrm{D}^{+,n}_{\mathrm{dif}}(D_S)$ because S is reduced and Z is Zariski-dense. Hence D_S is de Rham by Lemma 4.9 again. □

Remark 4.11. The work presented in this paper was finished in the summer of 2012 and made public at the beginning of 2013. Later that year came the preprint of [Bellovin 2015], in which the author built up a more robust theory of families of Hodge–Tate and de Rham representations over rigid analytic spaces. First of all, she generalized Berger's dictionary, which relates Fontaine's functors to (φ, Γ)-modules, to families of p-adic representations [Bellovin 2015, Theorem 1.1.1]. This result implies that our theory of families of Hodge–Tate and de Rham (φ, Γ)-modules with bounded Hodge–Tate weights developed in §3 and §4 can be viewed as a generalization of Berger and Colmez's theory of families of Hodge–Tate and de Rham representations with bounded Hodge–Tate weights. Moreover, she developed a theory of families of "partial" Hodge–Tate and de Rham representations with bounded Hodge–Tate weights. That is, the periods of the fibers are assumed to be of some constant rank which is not necessarily equal to the rank of the family. In addition, she removes the "reduced" assumption on the base by considering all artinian points. We refer the reader to [Bellovin 2015] for more results and details.

5. p-adic local monodromy for families of de Rham (φ, Γ)-modules

The main goal of this section is to prove the p-adic local monodromy for the restrictions of families of de Rham (φ, Γ)-modules with bounded Hodge–Tate weights on their Shilov boundary. The proof is modeled on Berger and Colmez's proof of the p-adic local monodromy for families of de Rham representations with bounded Hodge–Tate weights [Berger and Colmez 2008, §6]. Recall that $\nabla = \log(\gamma)/\log(\chi(\gamma))$, which is independent of the choice of $\gamma \in \Gamma$. This gives rise to an action of the Lie algebra of Γ on (φ, Γ)-modules and their localizations. In the following, we fix E to be a finite extension of the products of the complete residue fields of the Shilov boundary of $M(S)$.

Proposition 5.1. *Let D_S be a de Rham (φ, Γ)-module of rank d over $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} S$ with weights in $[a, b]$. For any $s > 0$ such that $n(s) \geq h_{\mathrm{dR}}(D_S)$, let*

$$N_s(D_E) = \big\{ y \in t^{-b} D_E^s \text{ such that } \iota_n(y) \in D_{\mathrm{dR}}(D_S) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!] \\ \text{for each } n \geq n(s) \big\}.$$

Then the following are true:

(1) *The $\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E$-module $N_s(D_E)$ is free of rank d and stable under Γ.*

(2) *For each $n \geq n(s)$, we have*

$$N_s(D_E) \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E, \iota_n} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!] = D_{\mathrm{dR}}(D_S) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!].$$

Furthermore, if we put $N_{\mathrm{dR}}(D_E) = N_s(D_E) \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E$, then the following are true:

(3) *The $\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E$-module $N_{\mathrm{dR}}(D_E)$ is free of rank d, stable under Γ, and independent of the choice of s.*

(4) *We have $\varphi^*(N_{\mathrm{dR}}(D_E)) = N_{\mathrm{dR}}(D_E)$ and $\nabla(N_{\mathrm{dR}}(D_E)) \subset t \cdot N_{\mathrm{dR}}(D_E)$.*

Proof. First, note that the sequence of $K_n \otimes_{\mathbb{Q}_p} E[\![t]\!]$-modules

$$\{D_{\mathrm{dR}}(D_S) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!]\}_{n \geq n(s)}$$

is φ-compatible in the sense of [Berger 2008b, Définition II.1.1]. Then by the proof of [Berger 2008b, Théorème II.1.2] (using the fact that E is a finite product of p-adic local fields which are endowed with discrete valuations extending the standard one on $\mathbb{Q}_p$), we see that $N_{\mathrm{dR}}(D_E)$ is the unique (φ, Γ)-module M_E contained in $D_E[1/t]$ such that

$$M_E^s \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E, \iota_n} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!] = D_{\mathrm{dR}}(D_S) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!]$$

for any $n \geq n(s)$. Furthermore, the proof of [Berger 2008b, Théorème II.1.2] implies all of the proposition except the second half of (4). To see that part, note that

$$\iota_n\big(\nabla(N_s(D_E))\big) = \nabla\big(\iota_n(N_s(D_E))\big) \subset t D_{\mathrm{dR}}(D_S) \otimes_{K \otimes_{\mathbb{Q}_p} S} (K_n \otimes_{\mathbb{Q}_p} E)[\![t]\!].$$

This yields that $\nabla(N_s(D_E)) \subset t N_s(D_E)$ for all s. Thus $\nabla(N_{\mathrm{dR}}(D_E)) \subset t N_{\mathrm{dR}}(D_E)$, as $N_{\mathrm{dR}}(D_E)$ is equal to the union of all $N_s(D_E)$. □

Proposition 5.2. *Keep the notation of Proposition 5.1. Then there exists a finite extension L over K such that*

$$M = \big(N_{\mathrm{dR}}(D_E) \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\log,L} \widehat{\otimes}_{\mathbb{Q}_p} E\big)^{I_L}$$

is a free $L'_0 \otimes_{\mathbb{Q}_p} E$-module of rank d and the natural map

$$M \otimes_{L'_0 \otimes_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\log,L} \widehat{\otimes}_{\mathbb{Q}_p} E \longrightarrow N_{\mathrm{dR}}(D_E) \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\log,L} \widehat{\otimes}_{\mathbb{Q}_p} E$$

is an isomorphism.

Proof. Let $f' = [K'_0 : \mathbb{Q}_p]$. Note that there is a canonical decomposition

$$\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E \cong \prod_{i=0}^{f'-1} \mathcal{R}^{(i)}_E,$$

where each $\mathcal{R}^{(i)}_E$ is isomorphic to $\mathcal{R}_E$ and stable under Γ_K, and satisfies $\varphi(\mathcal{R}^{(i)}_E) \subset \mathcal{R}^{(i+1)}_E$ (define $\mathcal{R}^{(f')}_E = \mathcal{R}^{(0)}_E$). Let

$$N^{(i)}_{\mathrm{dR}}(D_E) = N_{\mathrm{dR}}(D_E) \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \mathcal{R}^{(i)}_E.$$

It follows that each $N^{(i)}_{\mathrm{dR}}(D_E)$ is stable under $\partial = \nabla/t$ and $\varphi^{f'}$; hence it is a p-adic differential equation with a Frobenius structure. By the versions of the p-adic local monodromy theorem proved by André [2002] or Mebkhout [2002], we conclude that each $N^{(i)}_{\mathrm{dR}}(D_E)$ is potentially unipotent. This yields the proposition using the argument of Proposition 6.2.2 and Corollaire 6.2.3 of [Berger and Colmez 2008]. □

Lemma 5.3. *Keep notation as in Proposition 5.2, and let*

$$M = \big(N_s(D_E) \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger,s}_{\log,K} \widehat{\otimes}_{\mathbb{Q}_p} E\big)^{I_L}$$

for sufficiently large s. Then, for any $n \geq n(s)$, we have

$$L \otimes_{L_0} \iota_n(M) = \big(\mathrm{D}_{\mathrm{dif}}\big(D_E \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\mathrm{rig},L} \widehat{\otimes}_{\mathbb{Q}_p} E\big)\big)^{I_L}. \tag{5.3.1}$$

Proof. By the previous proposition, the left side of (5.3.1) is a free $L \otimes_{L_0} L'_0 \otimes_{\mathbb{Q}_p} E$-module of rank d. On the other hand, since $((L_n \otimes_{\mathbb{Q}_p} E)[\![t]\!][1/t])^{I_L} = L \otimes_{L_0} L'_0 \otimes_{\mathbb{Q}_p} E$, we deduce that the right side of (5.3.1), which obviously contains the left side, is an $L \otimes_{L_0} L'_0 \otimes_{\mathbb{Q}_p} E$-module generated by at most d elements. Using the fact that $L \otimes_{L_0} L'_0 \otimes_{\mathbb{Q}_p} E$ is a product of fields, we deduce the desired identity. □

6. Proof of the main theorem

Now let V_X be a finite slope family of dimension d over a reduced rigid analytic space X over F. We start by making some preliminary reductions. After a finite surjective base change of X, we may assume that $Q(T)$ factors as $\prod_{i=1}^{m}(T-F_i)$. By reordering the F_i and throwing away some points of Z, we may further assume that, for all $z\in Z$, $v_p(F_i(z))\geq v_p(F_j(z))$ if $i>j$ and $F_i(z)\neq F_j(z)$ if $F_i\neq F_j$. We then set $\mathscr{F}_{i,z}=D^+_{\mathrm{st}}(V_z)^{(\varphi^f-F_1(z))\cdots(\varphi^f-F_i(z))=0}$ for all $z\in Z$ and $1\leq i\leq m$. Using Definition 0.1(3), we may suppose that $\mathscr{F}_{i,z}\subseteq\mathscr{F}_z$ for all $z\in Z$ and $1\leq i\leq m$ by shrinking Z. Furthermore, by the fact that $N\varphi=p\varphi N$ and the condition that $v_p(F_i(z))\geq v_p(F_j(z))$ if $i>j$, we see that $N=0$ on each graded piece $\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}$. Let $c_{i,z}$ be the rank of $\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}$ over $K_0\otimes k(z)$, and partition Z into finitely many subsets according to the sequence $c_{i,z}$. Note that at least one of these subsets of Z has to be Zariski-dense. Replace Z by this subset, and set $c_i=c_{i,z}$ for $z\in Z$.

For $z\in Z$, we will inductively define (φ,Γ)-submodules $\mathrm{Fil}_{i,z}\subset\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)$ for $1\leq i\leq m$ such that $D^+_{\mathrm{st}}(\mathrm{Fil}_{i,z})=\mathscr{F}_{i,z}$. For $i=1$, since V_z has nonpositive Hodge–Tate weights and $N(\mathscr{F}_{1,z})=0$, we have

$$\mathscr{F}_{1,z}=D^+_{\mathrm{st}}(V_z)^{\varphi^f=F_1(z),N=0}=D^+_{\mathrm{crys}}(V_z)^{\varphi^f=F_1(z)}=\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)^{\Gamma=1,\varphi^f=F_z(z)},$$

using Berger's dictionary [2002, Théorème 3.6]. Let $\mathrm{Fil}_{1,z}$ be the saturation of the (φ,Γ)-submodule of $\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)$ generated by $\mathscr{F}_{1,z}$. It is then clear that $D^+_{\mathrm{st}}(\mathrm{Fil}_{1,z})=D^+_{\mathrm{crys}}(\mathrm{Fil}_{1,z})=\mathscr{F}_{1,z}$. Now suppose we have defined $\mathrm{Fil}_{i-1,z}$ for some $i\geq 2$ such that $D^+_{\mathrm{st}}(\mathrm{Fil}_{i-1,z})=\mathscr{F}_{i-1,z}$. It follows that

$$D^+_{\mathrm{st}}(\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)/\mathrm{Fil}_{i-1,z})=D^+_{\mathrm{st}}(V_z)/\mathscr{F}_{i-1,z}.$$

Note that

$$\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}=(D^+_{\mathrm{st}}(V_z)/\mathscr{F}_{i-1,z})^{\varphi^f=F_i(z),N=0}.$$

Hence

$$\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}=D^+_{\mathrm{crys}}(\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)/\mathrm{Fil}_{i,z})^{\varphi^f=F_i(z)}\subset(\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)/\mathrm{Fil}_{i-1,z})^{\Gamma}.$$

We set $\mathrm{Fil}_{i,z}$ to be the preimage of the saturation of the (φ,Γ)-submodule of $\mathrm{D}^\dagger_{\mathrm{rig}}(V_z)/\mathrm{Fil}_{i-1,z}$ generated by $\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}$. Now, for each $1\leq i\leq m$, we define the character $\delta_i:K^\times\to\mathscr{O}(X)^\times$ by setting $\delta_i(p)=F_i^{-1}$ and $\delta_i(\mathscr{O}_K^\times)=1$. Let $D_X=\mathrm{D}^\dagger_{\mathrm{rig}}(V_X)^\vee$.

Lemma 6.1. *Suppose that X is irreducible. Then, for each $0\leq i\leq m$, there exists a proper birational morphism $\pi:X'\to X$ and a subfamily of (φ,Γ)-modules $D^{(i)}_{X'}\subset D_{X'}$ over X' of rank $d-c_1-\cdots-c_i$ such that:*

(1) *For any $x\in X'$, the natural map $D^{(i)}_x\to D_x$ is injective.*

(2) *There exists a Zariski-open dense subset U of X' such that for any $z \in Z' = \pi^{-1}(Z) \cap U$, the natural map $D_z^{(i)} \to D_z$ is the dual of the projection* $\mathrm{D}_{\mathrm{rig}}^{\dagger}(V_{\pi(z)}) \to \mathrm{D}_{\mathrm{rig}}^{\dagger}(V_{\pi(z)})/\operatorname{Fil}_{i,\pi(z)}$.

Proof. We proceed by induction on i. The initial case is trivial. Suppose that, for some $1 \le i \le m$, the lemma is true for $i-1$. Note that $\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z}$ maps into $\mathrm{D}_{\mathrm{rig}}^{\dagger}(V_z)/\operatorname{Fil}_{i,z}$ for any $z \in Z$. Since $\mathscr{F}_{i,z}/\mathscr{F}_{i-1,z} = (D_{\mathrm{crys}}^{+}(V_z)/\mathscr{F}_{i-1,z})^{\varphi^f = F_i(z)}$, we get that $(D_z^{(i)})^{\vee}(\pi^*(\delta_i)(z))$ has $k(z)$-dimension c_i for any $z \in Z'$. Since Z' is Zariski-dense in X', by Proposition 2.5 after adapting X' and U, we may find a subfamily of (φ, Γ)-modules $D_{X'}^{(i)}$ of $D_{X'}^{(i-1)}$ with rank $d - c_1 - \cdots - c_i$ such that:

(1') $D_x^{(i)} \to D_x^{(i-1)}$ is injective for any $x \in X'$.

(2') For any $z \in \pi^{-1}(Z) \cap U$, $D_z^{(i)}$ is the kernel of the dual of the map

$$(\boldsymbol{B}_{\mathrm{rig},K}^{\dagger} \otimes_{\mathbb{Q}_p} k(z)) \cdot (\mathscr{F}_{i,\pi(z)}/\mathscr{F}_{i-1,\pi(z)}) \to \mathrm{D}_{\mathrm{rig}}^{\dagger}(V_{\pi(z)})/\operatorname{Fil}_{i,\pi(z)}.$$

It is clear that (1') and (2') imply (1) and (2), respectively; this finishes the inductive step. □

To prove Theorem 0.3, we also need the following lemma.

Lemma 6.2. *Let V_S be a free S-linear representation of G_K of rank d. Then there exists a positive integer $m(V_S)$ such that for any $x \in M(S)$ and $a \in \mathrm{D}_{\mathrm{dif}}^{+}(V_x)$, if a is Γ-invariant, then $a \in \mathrm{D}_{\mathrm{dif}}^{+,m(V_S)}(V_x)$.*

Proof. This is a consequence of the Tate–Sen method. Using [Berger and Colmez 2008, Théorème 4.2.9], we first choose a finite extension L over K and some positive integer m so that $\mathrm{D}_{\mathrm{rig},L}^{\dagger,r_m}(V_S)$ is a free $\boldsymbol{B}_{\mathrm{rig},L}^{\dagger,r_m} \widehat{\otimes}_{\mathbb{Q}_p} S$-module with a basis $\boldsymbol{e} = (e_1, \ldots, e_d)$. Let γ be a topological generator of Γ_{L_m} and write $\gamma(e) = eG$ for some $G \in \mathrm{GL}_d(\boldsymbol{B}_{\mathrm{rig},L}^{\dagger,r_m} \widehat{\otimes}_{\mathbb{Q}_p} S)$. Recall that by the classical work [Tate 1967] we know that there exists a constant $c > 0$ such that $v_p((\gamma-1)x) \le v_p(x) + c$ for any nonzero $x \in (1 - R_{L,m})\widehat{L}_{\infty}$, where $R_{L,m} : \widehat{L}_{\infty} \to L_m$ is Tate's normalized trace map. Since the localization map $\iota_m : \boldsymbol{B}_{\mathrm{rig},L}^{\dagger,r_m} \to L_m[\![t]\!]$ is continuous, by enlarging m we may suppose that the constant term of $\iota_m(G) - 1$ has norm less than p^{-c}. We fix some $m_0 \in \mathbb{N}$ such that $K_{\infty} \cap L_m = K_{m_0} \cap L_m$.

Now let $a \in \mathrm{D}_{\mathrm{dif}}^{+,K_n}(V_x)^{\Gamma}$ for some $x \in X$ and $n \ge m$. We will show that $a \in \mathrm{D}_{\mathrm{dif}}^{+,K_{m_0}}(V_x)^{\Gamma}$. Since $\iota_m(\boldsymbol{e})$ forms a basis of $\mathrm{D}_{\mathrm{dif}}^{+,L_n}(V_S)$, we may write $a = \iota_m(\boldsymbol{e})(x)A$ for some

$$A \in \mathrm{M}_{d\times 1}((L_n \otimes_{\mathbb{Q}_p} k(x))[\![t]\!]).$$

The Γ-invariance of a implies $\iota_m(G(x))\gamma(A) = A$; thus

$$(1 - R_{L,m})\iota_m(G(x))\gamma(A) = (1 - R_{L,m})A.$$

Note that $\iota_m(G(x))$ has entries in $(L_m \otimes_{\mathbb{Q}_p} k(x))[\![t]\!]$. It follows that $(G(x)-1)B = (1-\gamma^{-1})B$, where $B = (1-R_{L,m})A$. Let B_0 be the constant term of B. If $B_0 \neq 0$, then the constant term of $(\iota_m(G(x))-1)B$ has valuation

$$\geq v(\iota_m(G(x))-1) + v(B_0) > v(B_0) + c,$$

whereas the constant term $(1-\gamma^{-1})B_0$ of $(1-\gamma^{-1})B$ has valuation $\leq v(B_0)+c$; this yields a contradiction. Hence $B_0 = 0$. Iterating this argument, we get $B = 0$. Hence $a \in \mathrm{D}^{+,L_m}_{\mathrm{dif}}(V_x) \cap \mathrm{D}^{+,K_n}_{\mathrm{dif}}(V_x) \subset \mathrm{D}^{+,K_{m_0}}_{\mathrm{dif}}(V_x)$. Thus we may choose $m(V_S) = m_0$. □

Remark 6.3. Although we do not need it in this paper, it is worthwhile to point out that the argument of Lemma 6.2 works equally well for families of (φ, Γ)-modules and even a sequence of φ-compatible $K_n[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S$-modules $\{M_n\}_n$ in the vein of [Berger 2008b, Définition II.1.1]. That is, each M_n is a finite projective $K_n[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S$-module equipped with a continuous $K_n[\![t]\!]$-semilinear and S-linear Γ-action, and satisfies $M_n \otimes_{K_n[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S} K_{n+1}[\![t]\!] \widehat{\otimes}_{\mathbb{Q}_p} S \cong M_{n+1}$.

Proof of Theorem 0.3. We retain the notation above. By passing to irreducible components, we may suppose that X is irreducible. We then apply Lemma 6.1 to V_X. Note that $V_{X'}$ is again a finite slope family over X' with the Zariski-dense set of semistable points $\pi^{-1}(Z)$. We may suppose that $X' = X$. Let $\lambda : \mathrm{D}^{\dagger}_{\mathrm{rig}}(V_X) = D_X^{\vee} \to (D_X^{(m)})^{\vee}$ be the dual of $D_X^{(m)} \to D_X$, and let $P_X = \ker(\lambda)$. For any $x \in X$, since $D_x^{(m)} \to D_x$ is injective, we get that the image of λ_x is a (φ, Γ)-submodule of rank $d - c_1 - \cdots - c_m$. Thus, by Lemma 2.4, after adapting X we may assume that P_X is a family of (φ, Γ)-modules of rank $c_1 + \cdots + c_m$, and there exists a Zariski-open dense subset $U \subset X$ such that $P_x = \ker(\lambda_x)$ for any $x \in U$. Note that $\ker(\lambda_z) = \mathrm{Fil}_{i,z}$ for any $z \in Z$. Thus, by replacing Z with $Z \cap U$, we may assume that $P_z = \mathrm{Fil}_{i,z}$ for any $z \in Z$. We claim that P_X is de Rham with weights in $[-b, 0]$. To do so, we set Y to be the set of $x \in X$ for which P_x is de Rham with weights in $[a, b]$. By the previous lemma, we see that for any affinoid subdomain $M(S) \subset X$, there exists an integer $m(V_S)$ such that if P_x is de Rham for some $x \in M(S)$, then $h_{\mathrm{dR}}(P_x) \leq m(V_S)$. We then deduce from Proposition 4.10 that $Y \cap M(S)$ is a Zariski-closed subset of $M(S)$. Hence Y is a Zariski-closed subset of X. On the other hand, since P_z is de Rham with weights in $[-b, 0]$, we get $Z \subset Y$; thus $Y = X$ by the Zariski density of Z. Furthermore, using Proposition 4.10 and the previous lemma again, we deduce that P_X is de Rham with weights in $[-b, 0]$. As a consequence, we obtain a locally free coherent $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K$-module $D_{\mathrm{dR}}(P_X)$ of rank $c_1 + \cdots + c_m$.

The next step is to show that $D_{\mathrm{dR}}(P_x)$ is contained in $D^+_{\mathrm{st}}(V_x) \otimes_{K_0} K$ for any $x \in X$. Let Y be the set of $x \in X$ satisfying this condition. We first show that Y is a

Zariski-closed subset of X. For this, it suffices to show that $Y \cap M(S)$ is a Zariski-closed subset of $M(S)$ for any affinoid subdomain $M(S)$ of X. To show this, we employ the p-adic local monodromy for families of de Rham (φ, Γ)-modules. As in §5, let E be the product of the complete residue fields of the Shilov boundary of $M(S)$. Since P_S is a family of de Rham (φ, Γ)-modules with weights in $[-b, 0]$, by Lemma 5.3 there exists a finite extension L of K such that for sufficiently large s and $n \geq n(s)$, we have

$$L \otimes_{L_0} \iota_n(M) = \big(\mathrm{D}_{\mathrm{dif}}\big(P_E \otimes_{\boldsymbol{B}^{\dagger}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger}_{\mathrm{rig},L} \widehat{\otimes}_{\mathbb{Q}_p} E\big)\big)^{I_L}$$

for $M = \big(N_s(P_E) \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger,s}_{\log,K} \widehat{\otimes}_{\mathbb{Q}_p} E\big)^{I_L}$; furthermore, $N_s(P_E) \subset P^s_E$. Thus

$$\begin{aligned}\iota_n(M) &\subset \iota_n\big(P_E \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger,s}_{\log,K} \widehat{\otimes}_{\mathbb{Q}_p} E\big) \\ &\subset \iota_n\big(\mathrm{D}^{\dagger}_{\mathrm{rig}}(V_E) \otimes_{\boldsymbol{B}^{\dagger,s}_{\mathrm{rig},K} \widehat{\otimes}_{\mathbb{Q}_p} E} \boldsymbol{B}^{\dagger,s}_{\log,K} \widehat{\otimes}_{\mathbb{Q}_p} E\big) \subset \boldsymbol{B}^{+}_{\mathrm{st}} \widehat{\otimes}_{\mathbb{Q}_p} V_E.\end{aligned}$$

Note that $D_{\mathrm{dR}}(P_E) \subset \mathrm{D}^+_{\mathrm{dif}}(P_E) \subset \mathrm{D}^+_{\mathrm{dif}}(V_E) \subset \boldsymbol{B}^+_{\mathrm{dR}} \widehat{\otimes}_{\mathbb{Q}_p} V_E$. This yields

$$D_{\mathrm{dR}}(P_E) \subset (\boldsymbol{B}^+_{\mathrm{st}} \widehat{\otimes}_{\mathbb{Q}_p} V_E) \otimes_{L_0} L \cap \boldsymbol{B}^+_{\mathrm{dR}} \widehat{\otimes}_{\mathbb{Q}_p} V_E = (\boldsymbol{B}^+_{\mathrm{st}} \widehat{\otimes}_{\mathbb{Q}_p} V_E) \otimes_{L_0} L.$$

We therefore deduce from [Berger and Colmez 2008, Lemme 6.3.1] that

$$D_{\mathrm{dR}}(P_S) \subset (\boldsymbol{B}^+_{\mathrm{st}} \widehat{\otimes}_{\mathbb{Q}_p} V_E) \otimes_{L_0} L \cap \boldsymbol{B}^+_{\mathrm{dR}} \widehat{\otimes}_{\mathbb{Q}_p} V_S = (\boldsymbol{B}^+_{\mathrm{st}} \widehat{\otimes}_{\mathbb{Q}_p} V_S) \otimes_{L_0} L.$$

It follows that $Y \cap M(S)$, which is the set of $x \in M(S)$ such that $D_{\mathrm{dR}}(P_x) \subset (\boldsymbol{B}^+_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V_x) \otimes_{K_0} K$, is Zariski-closed in $M(S)$.

To conclude the proof of the theorem, it then suffices to show that, for any $x \in X$, $D_{\mathrm{dR}}(P_x) \subset (D^+_{\mathrm{st}}(V_x) \otimes_{K_0} K)^{Q(\varphi^f)(x)=0}$; here, we K-linearly extend the φ^f-action to $D^+_{\mathrm{st}}(V_x) \otimes_{K_0} K$. Note that $\mathrm{Fil}_{m,z}$ is semistable with $D_{\mathrm{st}}(\mathrm{Fil}_{m,z}) = \mathscr{F}_{m,z}$. This implies that $Q(\varphi^f)(D_{\mathrm{dR}}(P_X))$ vanishes at z, yielding that $Q(\varphi^f)(D_{\mathrm{dR}}(P_X)) = 0$ by the Zariski density of Z. □

Acknowledgements

Thanks to Christopher Skinner, Richard Taylor and Ila Varma for useful communications. We especially thank Richard Taylor for suggesting a more concise definition of finite slope families.

References

[André 2002] Y. André, "Filtrations de type Hasse–Arf et monodromie p-adique", *Invent. Math.* **148**:2 (2002), 285–317. MR 2003k:12011 Zbl 1081.12003

[Bellaïche and Chenevier 2009] J. Bellaïche and G. Chenevier, *Families of Galois representations and Selmer groups*, Astérisque **324**, 2009. MR 2011m:11105 Zbl 1192.11035

[Bellovin 2015] R. Bellovin, "p-adic Hodge theory in rigid analytic families", *Algebra Number Theory* **9**:2 (2015), 371–433.

[Berger 2002] L. Berger, "Représentations p-adiques et équations différentielles", *Invent. Math.* **148**:2 (2002), 219–284. MR 2004a:14022 Zbl 1113.14016

[Berger 2008a] L. Berger, "Construction de (ϕ, Γ)-modules: représentations p-adiques et B-paires", *Algebra Number Theory* **2**:1 (2008), 91–120. MR 2009j:14025 Zbl 1219.11078

[Berger 2008b] L. Berger, "Équations différentielles p-adiques et (ϕ, N)-modules filtrés", pp. 13–38 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et* (ϕ, Γ)-*modules*, edited by L. Berger et al., Astérisque **319**, 2008. MR 2010d:11056 Zbl 1168.11019

[Berger and Colmez 2008] L. Berger and P. Colmez, "Familles de représentations de de Rham et monodromie p-adique", pp. 303–337 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et* (ϕ, Γ)-*modules*, edited by L. Berger et al., Astérisque **319**, 2008. MR 2010g:11091 Zbl 1168.11020

[Harris et al. 2014] M. Harris, K.-W. Lan, R. Taylor, and J. Thorne, "On the rigid cohomology of certain Shimura varieties", preprint, 2014. arXiv 1411.6717

[Kedlaya and Liu 2010] K. Kedlaya and R. Liu, "On families of (ϕ, Γ)-modules", *Algebra Number Theory* **4**:7 (2010), 943–967. MR 2012h:11081 Zbl 1278.11060

[Kedlaya et al. 2014] K. S. Kedlaya, J. Pottharst, and L. Xiao, "Cohomology of arithmetic families of (φ, Γ)-modules", *J. Amer. Math. Soc.* **27**:4 (2014), 1043–1115. MR 3230818 Zbl 06355531

[Kisin 2003] M. Kisin, "Overconvergent modular forms and the Fontaine–Mazur conjecture", *Invent. Math.* **153**:2 (2003), 373–454. MR 2004f:11053 Zbl 1045.11029

[Liu 2014] R. Liu, "Triangulation of refined families", preprint, 2014. arXiv 1202.2188

[Mebkhout 2002] Z. Mebkhout, "Analogue p-adique du théorème de Turrittin et le théorème de la monodromie p-adique", *Invent. Math.* **148**:2 (2002), 319–351. MR 2003k:14018 Zbl 1071.12004

[Shah 2013] S. Shah, "Interpolating periods", preprint, 2013. arXiv 1305.2872

[Skinner and Urban 2006] C. Skinner and E. Urban, "Vanishing of L-functions and ranks of Selmer groups", pp. 473–500 in *International Congress of Mathematicians, Volume II*, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR 2008a:11063 Zbl 1157.11020

[Tate 1967] J. T. Tate, "p-divisible groups", pp. 158–183 in *Proceedings of a Conference on Local Fields* (Driebergen, 1966), edited by T. Springer, Springer, Berlin, 1967. MR 38 #155 Zbl 0157.27601

Communicated by Brian Conrad
Received 2014-02-11 Revised 2014-12-09 Accepted 2015-01-14

liuruochuan@math.pku.edu.cn *Beijing International Center for Mathematical Research, Peking University, Beijing, 100871, China*

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.459

The Picard rank conjecture for the Hurwitz spaces of degree up to five

Anand Deopurkar and Anand Patel

We prove that the rational Picard group of the simple Hurwitz space $\mathcal{H}_{d,g}$ is trivial for d up to five. We also relate the rational Picard groups of the Hurwitz spaces to the rational Picard groups of the Severi varieties of nodal curves on Hirzebruch surfaces.

Introduction

Let $\mathcal{H}_{d,g}$ be the simple Hurwitz space which parametrizes isomorphism classes of simply branched degree-d covers of genus-zero curves by genus-g curves. Although $\mathcal{H}_{d,g}$ has been studied classically, many fundamental questions about its geometry are still unanswered. The goal of this paper is to address one such question: the question of its Picard group. It is conjectured (for example, [Diaz and Edidin 1996]) that the rational Picard group $\operatorname{Pic}_{\mathbb{Q}}(\mathcal{H}_{d,g})$ is trivial. We call this the *Picard rank conjecture* for $\mathcal{H}_{d,g}$. Our main result is a proof of this conjecture for $d \leq 5$.

Theorem A. *The rational Picard group of $\mathcal{H}_{d,g}$ is trivial for $d \leq 5$.*

In the main text, Theorem A is divided into the case of degree 3 (Proposition 3.3), degree 4 (Proposition 4.10), and degree 5 (Proposition 5.4).

The Picard rank conjecture was known for $d = 2$ and 3. For $d = 2$, it was proved by Cornalba and Harris [1988, Lemma 4.5], and for $d = 3$ by Stankova-Frenkel [Stankova-Frenkel 2000, §12.2]. In these cases, there are now more refined results about the moduli stacks; see [Cornalba 2007] for $d = 2$ and [Bolognesi and Vistoli 2012; Bolognesi and Lönne 2014] for $d = 3$.

The conjecture is also known for $d > 2g-2$. In this range, the map $\mathcal{H}_{d,g} \to \mathcal{M}_g$ is a fibration, where $\mathcal{M}_g$ is the moduli space of smooth curves of genus g. An analysis of this fibration shows that $\operatorname{Pic}_{\mathbb{Q}}(\mathcal{H}_{d,g}) = 0$ if and only if $\operatorname{Pic}_{\mathbb{Q}}(\mathcal{M}_g) \cong \mathbb{Q}$ (see, for example, [Mochizuki 1995] or [Diaz and Edidin 1996, §3]). Thus, the conjecture for $d > 2g-2$ follows from Harer's theorem [1983].

MSC2010: primary 14H10; secondary 14C22.
Keywords: Hurwitz space, Picard group.

We briefly explain the rationale behind the conjecture. Let us blur the distinction between the coarse moduli spaces and the fine moduli stacks. This is harmless, since we are concerned with the rational Picard group. Let us also take $d \geq 4$ (the discussion holds for $d = 2, 3$ with minor modifications). Denote by $\widetilde{\mathcal{H}}_{d,g}$ the partial compactification of $\mathcal{H}_{d,g}$ that parametrizes covers $[\alpha : C \to \mathbb{P}^1]$ where C is allowed to be nodal, but still irreducible, and α need not be simply branched. Let $\alpha : \mathcal{C} \to \mathcal{P}$ be the universal family over $\widetilde{\mathcal{H}}_{d,g}$, where $\rho : \mathcal{C} \to \widetilde{\mathcal{H}}_{d,g}$ is a family of irreducible, at worst nodal curves of arithmetic genus g, and $\pi : \mathcal{P} \to \widetilde{\mathcal{H}}_{d,g}$ a family of smooth curves of genus 0. From this data, we can construct three "tautological" divisor classes on $\widetilde{\mathcal{H}}_{d,g}$, given by

$$\rho_*(c_1(\omega_\rho)^2), \quad \rho_*(c_1(\omega_\rho)\alpha^* c_1(\omega_\pi)), \quad \text{and} \quad \rho_*([\delta_\rho]).$$

Here ω stands for the relative dualizing sheaf and δ for the singular locus. It is easy to check that the three tautological classes are $\mathbb{Q}$-linearly independent. On the other hand, $\widetilde{\mathcal{H}}_{d,g} \setminus \mathcal{H}_{d,g}$ is a union of three irreducible divisors, namely, the locus Δ where C is singular, the locus T where α has a higher order ramification point, and the locus D where α has two ramification points over a branch point. It is also easy to check that the classes of Δ, T, and D are $\mathbb{Q}$-linearly independent. Thus, $\operatorname{Pic}_{\mathbb{Q}}(\mathcal{H}_{d,g}) = 0$ is equivalent to $\operatorname{Pic}_{\mathbb{Q}}(\widetilde{\mathcal{H}}_{d,g})$ being generated by the tautological classes. The Picard rank conjecture thus expresses the often-satisfied expectation that there are no other divisor classes than the tautological ones.

We now outline our strategy for proving Theorem A. Let $\alpha : C \to \mathbb{P}^1$ be a degree-d cover. Then C embeds in a $\mathbb{P}^{d-2}$-bundle over $\mathbb{P}^1$, which we denote by $\mathbb{P}E \to \mathbb{P}^1$. Thanks to the work of Casnati and Ekedahl, the resolution of the ideal of C in $\mathbb{P}E$ can be described explicitly. The terms in this resolution involve (twists of) vector bundles on $\mathbb{P}^1$ [Casnati and Ekedahl 1996]. Let $U \subset \widetilde{\mathcal{H}}_{d,g}$ be the open locus where these vector bundles are the most generic. The key steps in our proof are the following:

(1) Identify the divisorial components of $\widetilde{\mathcal{H}}_{d,g} \setminus U$.

(2) Express U as a (successive) quotient of an open subset of an affine space by actions of linear algebraic groups.

(3) Use the previous two steps to get a bound on the Picard rank of $\widetilde{\mathcal{H}}_{d,g}$, and in turn, the Picard rank of $\mathcal{H}_{d,g}$.

Needless to say, we are able to carry out all three steps only for $d \leq 5$. However, we can carry out parts of step (1) in general. For step (2), we highlight that it remains unknown in general whether one can dominate $\widetilde{\mathcal{H}}_{d,g}$ by an affine space for $d \geq 6$.

To analyze $\widetilde{\mathcal{H}}_{d,g} \setminus U$, we must analyze the loci in $\widetilde{\mathcal{H}}_{d,g}$ where the bundle E and the vector bundles appearing in the resolution of C are unbalanced. We call these loci the Maroni loci and the Casnati–Ekedahl loci, respectively. We spend significant effort on understanding the decomposition of $\widetilde{\mathcal{H}}_{d,g}$ into these loci. Contained in Section 2, the results of this analysis may be of independent interest.

A key tool in our analysis is a construction that relates the Maroni loci to the Severi varieties of Hirzebruch surfaces. Originally due to Ohbuchi [1997], this "associated scroll construction" allows us to get the required dimension estimates. The key input here is a theorem of Tyomkin [2007] that guarantees that the Severi varieties of Hirzebruch surfaces are irreducible of the expected dimension.

The associated scroll construction also lets us relate the Picard ranks of the Hurwitz spaces to the Picard ranks of the Severi varieties. To state our result, let us denote by $\mathcal{U}_g(\mathbb{F}_m, d\tau)$ the space of irreducible nodal curves of geometric genus g in the linear system $|d\tau|$ on the Hirzebruch surface $\mathbb{F}_m$, where τ is the section with self-intersection m.

Theorem B. *Let* $m \geq \lfloor (g+d-1)/(d-1) \rfloor$. *Then* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{U}_g(\mathbb{F}_m, d\tau) = 0$ *implies* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{d,g} = 0$.

Let $m \geq \lceil 2(g+d-1)/(d-1) \rceil$. *Then* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{U}_g(\mathbb{F}_m, d\tau) = 0$ *if and only if* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{d,g} = 0$.

In the main text, Theorem B is Theorem 6.7.

Notation. We work with a few different versions of the Hurwitz spaces. We assemble their definitions here. We work over the field $\mathbb{C}$ of complex numbers. By a curve, we mean a connected, proper, reduced scheme of finite type over $\mathbb{C}$. Throughout, assume that $g \geq 3$.

$\mathcal{H}_{d,g}$: This is the coarse moduli space of $[\alpha : C \to \mathbb{P}^1]$, where C is a smooth curve of genus g and α a finite map of degree d with simple branching (that is, the branch divisor of α is supported at $2g+2d-2$ distinct points). Two such covers $[\alpha_1 : C_1 \to \mathbb{P}^1]$ and $[\alpha_2 : C_2 \to \mathbb{P}^1]$ are considered isomorphic if there are isomorphisms $\phi : C_1 \to C_2$ and $\psi : \mathbb{P}^1 \to \mathbb{P}^1$ such that $\alpha_2 \circ \phi = \psi \circ \alpha_1$.

$\widetilde{\mathcal{H}}_{d,g}$: This is the coarse moduli space of $[\alpha : C \to \mathbb{P}^1]$, where C is an irreducible, at worst nodal curve of arithmetic genus g, and α a finite map of degree d. The isomorphism condition is the same as that for $\mathcal{H}_{d,g}$.

$\mathcal{H}^{\dagger}_{d,g}$: This is like $\mathcal{H}_{d,g}$, but with "framed" target $\mathbb{P}^1$. The objects it parametrizes are $[\alpha : C \to \mathbb{P}^1]$ as in the description of $\mathcal{H}_{d,g}$, but $[\alpha_1 : C_1 \to \mathbb{P}^1]$ and $[\alpha_2 : C_2 \to \mathbb{P}^1]$ are considered isomorphic if there is an isomorphism $\phi : C_1 \to C_2$ such that $\alpha_2 \circ \phi = \alpha_1$.

$\widetilde{\mathcal{H}}^{\dagger}_{d,g}$: This is like $\widetilde{\mathcal{H}}_{d,g}$, but with framed target $\mathbb{P}^1$.

All four are irreducible quasiprojective varieties with at worst quotient singularities. In particular, they are normal and $\mathbb{Q}$-factorial. The group $\operatorname{Aut} \mathbb{P}^1 = \mathrm{PGL}_2$ acts on the framed versions. The unframed versions are the quotients by this action in the sense that the fibers of the morphism from the framed space to the unframed space are precisely the PGL_2 orbits. We have

$$\dim \mathcal{H}_{d,g} = \dim \widetilde{\mathcal{H}}_{d,g} = 2g + 2d - 5,$$

and

$$\dim \mathcal{H}^{\dagger}_{d,g} = \dim \widetilde{\mathcal{H}}^{\dagger}_{d,g} = 2g + 2d - 2.$$

In addition, we work with the following Severi varieties:

$\mathcal{U}_g(\mathbb{F}_m, d\tau)$: This is the locus of irreducible nodal curves of geometric genus g in the linear series $|d\tau|$ in the Hirzebruch surface $\mathbb{F}_m$. Here $\tau \subset \mathbb{F}_m$ is the section of self-intersection m.

$\mathcal{V}_g(\mathbb{F}_m, d\tau)$: This is the closure of $\mathcal{U}_g(\mathbb{F}_m, d\tau)$ in the projective space $|d\tau|$.

$\mathcal{V}_g^{\mathrm{irr}}(\mathbb{F}_m, d\tau)$: This is the open subset of reduced irreducible curves in $\mathcal{V}_g(\mathbb{F}_m, d\tau)$.

We do not distinguish between a vector bundle and the corresponding locally free sheaf. Note that the vector bundle associated to the locally free sheaf F is the relative Spec of the symmetric algebra on $F^{\vee}$.

1. Preliminaries

In this expository section, we recall two key results. The first describes the Picard group of the quotient of a variety by a group action. The second is a structure theorem for finite covers which enables us to describe a large open subset of the Hurwitz space as such a quotient.

1A. ***Picard groups of quotients.*** Let G be a linear algebraic group acting on a variety X. Denote by $\operatorname{Pic}_G X$ the group of G-linearized line bundles on X. Forgetting the G-linearization gives a homomorphism $\operatorname{Pic}_G X \to \operatorname{Pic} X$.

Proposition 1.1 [Knop et al. 1989, Lemma 2.2, Proposition 2.3]. *For a connected linear algebraic group G acting on an irreducible variety X, we have an exact sequence*

$$\chi(G) \longrightarrow \operatorname{Pic}_G X \longrightarrow \operatorname{Pic} X,$$

where $\chi(G)$ is the group of (algebraic) characters of G. Furthermore, if X is normal, then the sequence has an extension by a homomorphism $\operatorname{Pic} X \to \operatorname{Pic} G$.

Let $\pi : X \to Y$ be a morphism that is equivariant with the trivial G action on Y. Let L be a line bundle on Y. The pullback $\pi^* L$ carries a natural G-linearization. We thus have a homomorphism $\operatorname{Pic} Y \to \operatorname{Pic}_G X$.

Proposition 1.2. *Let X and Y be irreducible normal varieties, G a linear algebraic group acting on X, and $\pi : X \to Y$ a surjective morphism, equivariant with the trivial action on Y. Suppose the fibers of π consist of single G-orbits. Then the map* $\operatorname{Pic} Y \to \operatorname{Pic}_G X$ *is injective and we have*

$$\operatorname{rk}\operatorname{Pic} Y \le \operatorname{rk}\chi(G) + \operatorname{rk}\operatorname{Pic} X.$$

Furthermore, if G is reductive and the stabilizers G_x are finite, then we have an isomorphism

$$\operatorname{Pic} Y \otimes \mathbb{Q} \xrightarrow{\sim} \operatorname{Pic}_G X \otimes \mathbb{Q}.$$

Proof. Suppose L is a line bundle on Y such that π^*L is trivial as a G-linearized line bundle. Then π^*L has a G-invariant nowhere-vanishing section. We claim that such a section descends to a nowhere-vanishing section of L on Y. The crucial point is that in our setup, Y is a geometric quotient of X [Mumford et al. 1994, Proposition 0.2]. That is, for every open $U \subset Y$, the preimage $\pi^{-1}U$ is open and the functions on U are the invariant functions on $\pi^{-1}U$:

$$\Gamma(U, \mathcal{O}_Y) = \Gamma(\pi^{-1}U, \mathcal{O}_X)^G.$$

It follows that the sections of L on U are the invariant sections of π^*L on $\pi^{-1}(U)$:

$$\Gamma(U, L) = \Gamma(\pi^{-1}U, \pi^*L)^G.$$

Thus, a G-invariant section σ of π^*L on X gives a section $\overline{\sigma}$ of L on Y. It is easy to check that if σ is nowhere-vanishing, so is $\overline{\sigma}$.

The bound on $\operatorname{rk}\operatorname{Pic} Y$ follows from injectivity and Proposition 1.1. For the last statement, we use the characterization of the image of $\operatorname{Pic} Y \to \operatorname{Pic}_G X$ from [Knop et al. 1989, Proposition 4.2]: a G-linearized line bundle L is in the image if and only if for every $x \in X$, the stabilizer group G_x acts trivially on the fiber L_x. Since the stabilizers are finite, we can arrange this by passing to a large enough power of L. □

We end with a simple application:

Proposition 1.3. *Let $U \subset \widetilde{\mathcal{H}}_{d,g}$ be any open subset and $U^\dagger$ its preimage under $\widetilde{\mathcal{H}}^\dagger_{d,g} \to \widetilde{\mathcal{H}}_{d,g}$. Then*

$$\operatorname{rk}\operatorname{Pic} U = \operatorname{rk}\operatorname{Pic} U^\dagger.$$

Proof. Apply Propositions 1.1 and 1.2 with $G = \mathrm{PGL}_2$, $X = U^\dagger$, and $Y = U$. □

1B. ***The Casnati–Ekedahl structure theorem.*** Let X and Y be integral schemes and $\alpha : X \to Y$ a finite flat Gorenstein morphism of degree $d \ge 3$. The map α gives an exact sequence

$$0 \longrightarrow \mathcal{O}_Y \longrightarrow \alpha_*\mathcal{O}_X \longrightarrow E_\alpha{}^\vee \longrightarrow 0, \tag{1-1}$$

where $E = E_\alpha$ is a vector bundle of rank $d-1$ on Y, called the *Tschirnhausen bundle* of α. Denote by ω_α the dualizing sheaf of α. Applying $\mathrm{Hom}_Y(-, \mathcal{O}_Y)$ to (1-1), we get

$$0 \longrightarrow E \longrightarrow \alpha_*\omega_\alpha \longrightarrow \mathcal{O}_Y \longrightarrow 0. \tag{1-2}$$

The map $E \to \alpha_*\omega_\alpha$ induces a map $\alpha^* E \to \omega_\alpha$.

Theorem 1.4 [Casnati and Ekedahl 1996, Theorem 2.1]. *In the above setup, $\alpha^* E \to \omega_\alpha$ gives an embedding $\iota : X \to \mathbb{P}E$ with $\alpha = \pi \circ \iota$, where $\pi : \mathbb{P}E \to Y$ is the projection. Moreover, the subscheme $X \subset \mathbb{P}E$ can be described as follows*:

(1) *The resolution of $\mathcal{O}_X$ as an $\mathcal{O}_{\mathbb{P}E}$-module has the form*

$$\begin{aligned} 0 \longrightarrow \pi^* N_{d-2}(-d) \longrightarrow \pi^* N_{d-3}(-d+2) \longrightarrow \pi^* N_{d-4}(-d+3) \\ \longrightarrow \cdots \longrightarrow \pi^* N_2(-3) \longrightarrow \pi^* N_1(-2) \longrightarrow \mathcal{O}_{\mathbb{P}E} \longrightarrow \mathcal{O}_X \longrightarrow 0, \end{aligned} \tag{1-3}$$

where the N_i are vector bundles on Y. Restricted to a point $y \in Y$, this sequence is the minimal free resolution of $X_y \subset \mathbb{P}E_y$.

(2) *The ranks of the N_i are given by*

$$\mathrm{rk}\, N_i = \frac{i(d-2-i)}{d-1}\binom{d}{i+1}.$$

(3) *We have $N_{d-2} \cong \pi^* \det E$. Furthermore, the resolution is symmetric, that is, isomorphic to the resolution obtained by applying* $\mathrm{Hom}_{\mathcal{O}_{\mathbb{P}E}}(-, N_{d-2}(-d))$.

The branch divisor of $\alpha : X \to Y$ is given by a section of $(\det E)^{\otimes 2}$. In particular, if X is a curve of (arithmetic) genus g, α has degree d, and $Y = \mathbb{P}^1$, then

$$\mathrm{rk}\, E = d-1 \text{ and } \deg E = g+d-1. \tag{1-4}$$

2. The Maroni and Casnati–Ekedahl loci

Consider a cover $\alpha : C \to \mathbb{P}^1$ and the relative canonical embedding $C \subset \mathbb{P}E_\alpha$. Since vector bundles on $\mathbb{P}^1$ split as direct sums of line bundles, the vector bundle E_α, and the higher syzygy bundles N_i appearing in Theorem 1.4, are discrete invariants of α. We thus get a decomposition of the Hurwitz space into locally closed subsets where the isomorphism type of the bundles E_α and N_i are constant. This section is devoted to the analysis of some of these locally closed subvarieties, particularly their dimensions. We only consider the bundle E_α and $F_\alpha := N_1$. Note that

$$E_\alpha = \ker(\alpha_*\omega_\alpha \to \mathcal{O}_Y) \text{ and } F_\alpha = \alpha_* I_C(2),$$

where $I_C \subset \mathcal{O}_{\mathbb{P}E_\alpha}$ is the ideal sheaf of C.

Definition 2.1. For vector bundles E and F on $\mathbb{P}^1$, define closed subvarieties of $\mathcal{H}^{\dagger}_{d,g}$

$$\begin{aligned} M(E,F) &:= \overline{\{[\alpha : C \to \mathbb{P}^1] \mid E_\alpha \cong E \text{ and } F_\alpha \cong F\}}, \\ M(E) &:= \overline{\{[\alpha : C \to \mathbb{P}^1] \mid E_\alpha \cong E\}}, \\ C(F) &:= \overline{\{[\alpha : C \to \mathbb{P}^1] \mid F_\alpha \cong F\}}. \end{aligned}$$

Call $M(E)$ the *Maroni loci* and $C(F)$ the *Casnati–Ekedahl loci*. Define subvarieties $\widetilde{M}(E,F)$, $\widetilde{M}(E)$, and $\widetilde{C}(F)$ of $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$ analogously.

Abusing notation, we denote the images of these loci in the unframed versions $\mathcal{H}_{d,g}$ and $\widetilde{\mathcal{H}}_{d,g}$ by the same letters. The framed versus unframed setting is usually clear by context, and sometimes irrelevant, for example in discussing the codimensions. We caution the reader that these loci are not necessarily irreducible or of expected dimension (Examples 4.3, 4.4). Even determining whether they are nonempty remains a challenge in full generality.

2A. ***The associated scroll construction.*** To analyze the Maroni loci $M(E)$, we associate to a cover of $\mathbb{P}^1$ a curve on a Hirzebruch surface. The construction is originally due to Ohbuchi [1997]. Let C be an irreducible curve of arithmetic genus g and $\alpha : C \to \mathbb{P}^1$ a finite cover of degree d. Let ζ be a global section of $\mathcal{O}_C(m) = \alpha^*\mathcal{O}_{\mathbb{P}^1}(m)$ that projects to a nonzero section of $E_\alpha^\vee(m)$. In other words, let $\zeta \in H^0(C, \mathcal{O}_C(m))$ be an element not contained in $\alpha^* H^0(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(m))$. The section ζ gives a map from C to the total space of the line bundle $\mathcal{O}(m)$ over $\mathbb{P}^1$. Let $\mathbb{F}_m = \operatorname{Proj}(\mathcal{O} \oplus \mathcal{O}(-m))$ be the Hirzebruch surface that compactifies this total space. We thus get the diagram

$$\begin{tikzcd} C \arrow[r, "\nu"] \arrow[dr, "\alpha"'] & \mathbb{F}_m \arrow[d, "\pi"] \\ & \mathbb{P}^1 \end{tikzcd}$$

Let $\sigma \subset \mathbb{F}_m$ be the directrix (the unique section of $\mathbb{F}_m \to \mathbb{P}^1$ of negative self-intersection) and $\tau \subset \mathbb{F}_m$ the section disjoint from σ (so that $\sigma^2 = -m$ and $\tau^2 = m$). By construction, $\nu(C) \subset \mathbb{F}_m$ avoids the directrix σ. Suppose C is smooth and $\alpha : C \to \mathbb{P}^1$ does not factor nontrivially. Then ν is birational onto its image, and therefore $\nu(C)$ is a reduced and irreducible element of the linear system $|d\tau|$. By the following proposition, $\nu(C)$ is a point in the Severi variety $\mathcal{V}_g(\mathbb{F}_m, d\tau)$.

Proposition 2.2. *A reduced and irreducible curve on $\mathbb{F}_m$ of geometric genus g in the linear system $|d\tau|$ is a flat limit of irreducible nodal curves of geometric genus g.*

Proof. Let $\overline{C} \subset \mathbb{F}_m$ be such a reduced and irreducible curve. Let $C \to \overline{C}$ be the normalization map. Denote by ν the composite map $\nu : C \to \mathbb{F}_m$. Let $\mathcal{M}$ be a component of the Kontsevich space of maps $\mathcal{M}_g(\mathbb{F}_m, d\tau)$ containing ν. Let N_ν be the normal sheaf of ν; this is the cokernel of $T_C \to \nu^* T_{\mathbb{F}_m}$. Then, we have a lower bound $\dim \mathcal{M} \geq \chi(N_\nu)$. Since

$$\chi(N_\nu) = \chi(\nu^* T_{\mathbb{F}_m}) - \chi(T_C) = g - \deg(K_{\mathbb{F}_m} \cdot \overline{C}) - 1,$$

we get

$$\dim \mathcal{M} \geq g - \deg(K_{\mathbb{F}_m} \cdot \overline{C}) - 1.$$

By [Harris 1986, Proposition 2.2], a general $\nu_{\text{gen}} : C_{\text{gen}} \to \mathbb{F}_m$ in $\mathcal{M}$ is birational onto its image and the image has only nodes as singularities. □

We can make the construction in a family. Let M be a reduced scheme, $\rho : C \to M$ a generically smooth family of reduced and irreducible curves of genus g, and $\alpha : C \to \mathbb{P}^1 \times M$ a finite flat M-morphism of degree d. Set $\mathcal{O}_C(m) = \alpha^* \mathcal{O}(m)$. Assume that none of the fibers $\alpha_t : C_t \to \mathbb{P}^1$ factor nontrivially and that $H^0(C_t, \mathcal{O}_{C_t}(m))$ has constant rank. Then $\rho_* \mathcal{O}_C(m)$ is a vector bundle on M. The trivial subbundle $H^0(\mathbb{P}^1, \mathcal{O}(m)) \otimes \mathcal{O}_M$ maps injectively to $\rho_* \mathcal{O}_C(m)$. Let U be the complement of the image of this map in the total space of $\rho_* \mathcal{O}_C(m)$. Fiberwise, the sections of U correspond to the sections ζ which project nontrivially onto $E_\alpha^\vee(m)$. Then the associated scroll construction gives a morphism

$$U \to \mathcal{V}_g(\mathbb{F}_m, d\tau).$$

We will use this construction where M is a Maroni locus. As described, the construction depends on the existence of a universal family, and thus gives a morphism from the fine moduli stack. But since $\mathcal{V}_g(\mathbb{F}_m, d\tau)$ is a scheme, we get a canonical induced map from the coarse space.

The following crucial result makes the above construction useful:

Theorem 2.3 [Tyomkin 2007]. *All Severi varieties parametrizing irreducible curves on Hirzebruch surfaces are irreducible and of expected dimension. In particular, the variety $\mathcal{V}_g(\mathbb{F}_m, d\tau)$ is irreducible of dimension $dm + 2d + g - 1$.*

We also need the following result, which we prove for lack of a reference:

Proposition 2.4. *Let $\overline{C} \subset \mathbb{F}_m$ be a general point of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$ and $C \to \overline{C}$ the normalization. Then the composite $C \to \mathbb{P}^1$ is simply branched.*

Proof. In light of Theorem 2.3, it suffices to exhibit a particular $\overline{C}$ of geometric genus g in $\mathcal{V}_g(\mathbb{F}_m, d\tau)$ whose normalization is simply branched over $\mathbb{P}^1$. One way is to start with $X = \mathbb{P}^1$ and $\alpha : X \to \mathbb{P}^1$ a simply branched cover of degree d. Then $E_\alpha = \mathcal{O}(1)^{\oplus(d-1)}$. Choosing a general section of $E_\alpha^\vee(m)$ gives $\nu : X \to \mathbb{F}_m$ such that $\nu(X)$ is nodal. It is easy to see that $\nu(X)$ is in the closure of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$. Indeed, since the set of nodes of $\nu(X)$ impose independent conditions on $|K_{\mathbb{F}_m} + d\tau|$, they automatically impose independent conditions on $|d\tau|$ as well, and hence we may smooth out the required number of nodes of $\nu(X)$ to deform to a curve of geometric genus g. A general fiber of such a smoothing is the required $\overline{C}$. □

Remark 2.5. We can realize the associated scroll construction geometrically as follows. The choice of a general global section ζ of $\mathcal{O}_C(m)$ can be thought of as a choice of a geometric section $\sigma : \mathbb{P}^1 \to \mathbb{P}E$. In the $\mathbb{P}^{d-2}$ fibers of $\pi : \mathbb{P}E \to \mathbb{P}^1$, we now have $d+1$ points: d points coming from the fibers of the map $\alpha : C \to \mathbb{P}^1$, and one more point provided by the section σ. For general $t \in \mathbb{P}^1$, these $d+1$ points will be in general position, and so will define a unique rational normal curve $R_t \subset \mathbb{P}E$. Consider the birationally ruled surface $S \subset \mathbb{P}E$ defined as the closure of the union of the R_t. S contains both σ and C, and is fibered over $\mathbb{P}^1$. We contract all components of the fibers of the projection $\pi : S \to \mathbb{P}^1$ which do not meet the directrix σ. The resulting surface is $\mathbb{F}_m$, with σ being the directrix. The image of C under the contraction $S \to \mathbb{F}_m$ is the associated scroll construction.

For a vector bundle $E = \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_n)$ on $\mathbb{P}^1$, set

$$\lfloor E \rfloor = \min\{a_i\} \quad \text{and} \quad \lceil E \rceil = \max\{a_i\}.$$

Given a cover $\alpha : C \to \mathbb{P}^1$, the associated scroll construction $\nu : C \to \mathbb{F}_m$ can be made for $m \geq \lfloor E_\alpha \rfloor$. Conversely, given a point $\overline{C} \in \mathcal{V}_g^{\mathrm{irr}}(\mathbb{F}_m, d\tau)$, let $C \to \overline{C}$ be the normalization. Then the induced cover $\alpha : C \to \mathbb{P}^1$ has $\lfloor E_\alpha \rfloor \leq m$.

Proposition 2.6. *If $\widetilde{M}(E)$ is nonempty, then*

$$\lceil E \rceil \leq \frac{2g+2d-2}{d}. \tag{2-1}$$

Furthermore, if E_α comes from a cover $[\alpha : C \to \mathbb{P}^1]$, with C irreducible, and where α does not factor nontrivially, then

$$\frac{g+d-1}{\binom{d}{2}} \leq \lfloor E_\alpha \rfloor \leq \frac{g+d-1}{d-1}. \tag{2-2}$$

Proof. The resolution of $\mathcal{O}_C$ in Theorem 1.4 tells us that $C \subset \mathbb{P}E_\alpha$ is not contained in any hyperplane divisor. Let h denote the hyperplane divisor class associated to $\mathcal{O}_{\mathbb{P}E_\alpha}(1)$, and let f denote the class of the fiber of $\pi : \mathbb{P}E \to \mathbb{P}^1$. Set $N := \lceil E_\alpha \rceil$. Then the divisor class $h - Nf$ is effective. Since C is irreducible and does not lie in $(h-Nf)$, it intersects $(h-Nf)$ nonnegatively. Since $h \cdot [C] = 2g+2d-2$ and $f \cdot [C] = d$, we conclude that $N \leq (2d+2g-2)/d$.

For the second inequality, we appeal to the associated scroll construction. Let $n := \lfloor E_\alpha \rfloor$. Since α does not factor, $\nu : C \to \mathbb{F}_n$ must be birational onto its image.

Adjunction on $\mathbb{F}_n$ gives

$$p_a(\nu(C)) = \binom{d}{2}n - (d-1).$$

The second statement now follows from the inequality $g \leq p_a(\nu(C))$. □

The following result places a strong restriction on a large class of Tschirnhausen bundles E.

Proposition 2.7 [Ohbuchi 1997]. *Let $\alpha : C \to \mathbb{P}^1$ be a cover of degree d, with C irreducible, and where α does not factor nontrivially. Write $E_\alpha = \mathcal{O}(a_1) \oplus \cdots \oplus \mathcal{O}(a_{d-1})$, where $\lfloor E_\alpha \rfloor = a_1 \leq a_2 \leq \cdots \leq a_{d-1} = \lceil E_\alpha \rceil$. Then*

$$a_{i+1} - a_i \leq \lfloor E_\alpha \rfloor \quad \textit{for} \quad 1 \leq i \leq d-2. \tag{2-3}$$

Remark 2.8. Proposition 2.7 implies the second inequality in Proposition 2.6.

Definition 2.9. We call a vector bundle E on $\mathbb{P}^1$ of rank $d-1$ and degree $g+d-1$ *tame* if it satisfies inequalities (2-1), (2-2), and (2-3).

Notice that Proposition 2.6 and Proposition 2.7 imply that E_α is tame in the following two cases: α is simply branched, or d is prime. Indeed, in either case, the cover cannot factor nontrivially.

Denote by $\rightsquigarrow$ the partial order on isomorphism classes of vector bundles on $\mathbb{P}^1$ given by $E \rightsquigarrow E'$ if E specializes to E' in a flat family. Note that isomorphism classes of vector bundles of rank r and degree n on $\mathbb{P}^1$ can be identified with unordered partitions of n with r parts. Then the order $\rightsquigarrow$ is the usual dominance order of partitions. For example, we have $(2,3,4) \rightsquigarrow (2,2,5)$ and $(2,3,4) \rightsquigarrow (1,4,4)$, but $(2,2,5)$ and $(1,4,4)$ are incomparable.

Define the finite set $\mathcal{T}[m]$ by

$$\mathcal{T}[m] := \big\{\text{Isomorphism classes of tame bundles } E \text{ of rank } d-1, \\ \text{degree } g+d-1, \text{ and } \lfloor E \rfloor = m\big\}.$$

Observe that $\mathcal{T}[m]$ contains an element $E[m]$ such that $E[m] \rightsquigarrow E$ for all $E \in \mathcal{T}[m]$. In other words, $E[m]$ is the most generic among all the bundles in $\mathcal{T}[m]$.

Theorem 2.10. *Let $m \in \mathbb{Z}$ satisfy $(g+d-1)/\binom{d}{2} \leq m \leq (g+d-1)/(d-1)$.*

(1) *If $M(E)$ is nonempty, then E is a tame bundle.*

(2) *If $\lfloor E \rfloor \leq m$ then $M(E) \subset M(E[m])$.*

(3) *$M(E[m]) \subset M(E[m+1])$ for all m.*

(4) *$M(E[m])$ is an irreducible subvariety of $\mathcal{H}^\dagger_{d,g}$ of codimension $g-(d-1)m+1$ unless $m = \lfloor (g+d-1)/(d-1) \rfloor$, in which case $M(E[m]) = \mathcal{H}^\dagger_{d,g}$.*

(5) *If d is prime, then all the statements above hold with $M(-)$ replaced by $\widetilde{M}(-)$ and $\mathcal{H}^{\dagger}_{d,g}$ replaced by $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$.*

In the proof, we use the following result (restated here for our setup):

Theorem 2.11 [Coppens 1999]. *For all m satisfying $(g+d-1)/\binom{d}{2} \le m \le (g+d-1)/(d-1)$, there is a genus-$g$ and degree-d cover $C \to \mathbb{P}^1$ with Tschirnhausen bundle $E[m]$. Moreover, C is birational onto its image under the associated scroll construction $C \to \mathbb{F}_m$.*

Proof of Theorem 2.10. The first statement follows from Propositions 2.6 and 2.7.

Before we proceed, we make two observation about the normalization C of a general point $[\overline{C}]$ of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$. First, $C \to \mathbb{P}^1$ is simply branched. Second, $C \to \mathbb{P}^1$ has Tschirnhausen bundle $E[m]$. Indeed, both are open conditions on $\mathcal{V}_g(\mathbb{F}_m, d\tau)$. By Proposition 2.4, there is a point satisfying the first condition. By Theorem 2.11, there is a point satisfying the second condition. By the irreducibility of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$, a generic point satisfies both conditions.

For the second statement, suppose $\lfloor E \rfloor \le m$ and let $[\alpha : C \to \mathbb{P}^1]$ be a point of $M(E)$. The associated scroll construction gives $\nu : C \to \mathbb{F}_m$; let $\overline{C} \subset \mathbb{F}_m$ be the image. Since α is simply branched, $\nu : C \to \overline{C}$ is birational. By the previous paragraph, we know that $[\overline{C}] \in \mathcal{V}_g(\mathbb{F}_m, d\tau)$ is the limit of an arc in $\mathcal{V}_g(\mathbb{F}_m, d\tau)$ whose general point corresponds to a curve with Tschirnhausen bundle $E[m]$. More precisely, we know that over a germ of a smooth curve (or the spectrum of a DVR) Δ there exists $\overline{\mathcal{C}} \subset \mathbb{F}_m \times \Delta$ such that:

- $\overline{\mathcal{C}} \to \Delta$ is a family of reduced and irreducible curves of geometric genus g.
- The fibers $\overline{\mathcal{C}}_t \subset \mathbb{F}_m$ are in the linear system $|d\tau|$.
- The special fiber $\overline{\mathcal{C}}_0$ is $\overline{C}$.
- The general fiber $\overline{\mathcal{C}}_t$ has the property that $(\overline{\mathcal{C}}_t)^{\nu} \to \mathbb{P}^1$ has Tschirnhausen bundle $E[m]$, where the superscript ν denotes normalization.

Let $\mathcal{C} \to \overline{\mathcal{C}}$ be the normalization of the total space of this family. The main theorem of [Teissier 1980] says that the fibers of $\mathcal{C} \to \Delta$ are the normalizations of the corresponding fibers of $\overline{\mathcal{C}} \to \Delta$. Considering the composition $\mathcal{C} \to \mathbb{P}^1$ of the sequence of maps $\mathcal{C} \to \overline{\mathcal{C}} \to \mathbb{F}_m \to \mathbb{P}^1$, we see that $\alpha : C \to \mathbb{P}^1$ is the limit of covers $\mathcal{C}_t \to \mathbb{P}^1$ which have Tschirnhausen bundle $E[m]$. The second statement follows.

The third statement is a corollary of the second statement.

For the fourth statement, suppose $m = \lfloor (g+d-1)/(d-1) \rfloor$. Then $E[m]$ is balanced, so $M(E[m]) = \mathcal{H}^{\dagger}_{d,g}$. Suppose $m < \lfloor (g+d-1)/(d-1) \rfloor$. Let $U \subset \mathcal{V}_g(\mathbb{F}_m, d\tau)$ be the locus of nodal curves of geometric genus g whose normalization is simply branched over $\mathbb{P}^1$. Then U is a smooth open subset of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$. Normalization of the universal family of curves in $\mathbb{F}_m$ of geometric genus g gives a family of smooth curves of genus g with a simply branched map of degree d to $\mathbb{P}^1$

(induced from $\mathbb{F}_m \to \mathbb{P}^1$). By definition, the image is in $M(E[m])$. We thus get a dominant map

$$q : U \to M(E[m]).$$

The fiber of q over $[\alpha : C \to \mathbb{P}^1]$ corresponds to the global sections of $\mathcal{O}_C(m)$ that project nontrivially onto $E^\vee(m)$. For general $\alpha \in M(E[m])$, we have $E_\alpha = E[m]$. Also, since $m < \lfloor (g+d-1)/(d-1) \rfloor$, the bundle $E[m]$ has a unique $\mathcal{O}(m)$ summand and all other summands have degree greater than m. Therefore, the general fiber of q has dimension $m+2$. From the dimension of $\mathcal{V}_g(\mathbb{F}_m, d\tau)$, we get

$$\dim M(E[m]) = \dim \mathcal{V}_g(\mathbb{F}_m, d\tau) - (m+2) = (d-1)m + g + 2d - 3.$$

Since $\dim \mathcal{H}^\dagger_{d,g} = 2g + 2d - 2$, the fourth statement follows.

For the last statement, all the arguments hold for $\widetilde{M}(E)$ if d is prime, since the associated scroll construction $\nu : C \to \mathbb{F}_m$ is automatically birational onto its image. □

Theorem 2.10 gives us good control on the dimensions of the Maroni loci for E based on the minimal summand of E. We must now consider those E which are nongeneric, but nonetheless have the same minimal summand as the generic Tschirnhausen bundle. Set $k = \lfloor (g+d-1)/(d-1) \rfloor$. Then

$$E[k] = \mathcal{O}(k)^{\oplus r} \oplus \mathcal{O}(k+1)^{\oplus d-r-1},$$

where $0 < r \leq d-1$. A general cover $\alpha \in \mathcal{H}^\dagger_{d,g}$ has $E[k]$ as its Tschirnhausen bundle. Let E' be any tame bundle, and set $s := h^0(E'^\vee(k))$. Upper-semicontinuity implies $s \geq r$. Suppose $s > r$. Define

$$M^\circ(E') = \{\alpha \in \mathcal{H}^\dagger_{d,g} \mid E_\alpha \cong E'\}.$$

Then $M^\circ(E')$ is locally closed, and $\overline{M^\circ(E')} = M(E')$.

Lemma 2.12. *Under the assumptions above, let $Z \subset M^\circ(E')$ be any irreducible component. Then the codimension of $\overline{Z}$ in $\mathcal{H}^\dagger_{d,g}$ is at least $(s-r)+1$.*

Proof. Let $z = \dim Z$. We use the associated scroll construction over Z. We have an open subset U of a vector bundle of rank $s+k+1$ over Z and a morphism $U \to \mathcal{V}_g(\mathbb{F}_k, d\tau)$. Since $E' \neq E[k]$, the closure of the image of U is a proper subvariety of $\mathcal{V}_g(\mathbb{F}_k, \tau)$. In particular, we have $\dim U < \dim \mathcal{V}_g(\mathbb{F}_k, d\tau) = dk + 2d + g - 1$. The lemma follows from this inequality. □

We now have the tools to determine all the Maroni divisors.

Proposition 2.13. *The Maroni locus $M(E) \subset \mathcal{H}_{d,g}$ is a divisor if and only if $g = (k-1)(d-1)$ for some integer $k \geq 1$, and*

$$E = E[k-1] = \mathcal{O}(k-1) \oplus \mathcal{O}(k)^{\oplus d-3} \oplus \mathcal{O}(k+1).$$

Furthermore, in this situation, $M(E[k-1])$ is irreducible.

Proof. If $\lfloor E \rfloor = k = \lfloor (g+d-1)/(d-1) \rfloor$, then the statement follows by applying Lemma 2.12. If, on the other hand, $\lfloor E \rfloor < \lfloor (g+d-1)/(d-1) \rfloor$, then the statement follows from statement (4) of Theorem 2.10. ☐

We record a particularly interesting case of the irreducibility of the Maroni divisor:

Corollary 2.14. *Let $g = 2(d-1)$. Then $M(E[2]) \subset \mathcal{H}_{d,g}$ is irreducible, and it is the ramification locus of the generically finite and dominant forgetful map $\mu : \mathcal{H}_{d,g} \to \mathcal{M}_g$.*

Proof. The irreducibility statement follows from Theorem 2.10. To show that $M(E[2])$ is the ramification locus of μ, consider $[\alpha : C \to \mathbb{P}^1] \in \mathcal{H}_{d,g}$ and the map of sheaves

$$0 \longrightarrow \alpha^*(T_{\mathbb{P}^1}) \longrightarrow T_C \longrightarrow N_\alpha \longrightarrow 0.$$

The tangent space to $\mathcal{H}_{d,g}$ at α is $H^0(C, N_\alpha)/\alpha^* H^0(\mathbb{P}^1, T_{\mathbb{P}^1})$ and the tangent space to $\mathcal{M}_g$ at C is $H^1(C, T_C)$. The map

$$d\mu : H^0(C, N_\alpha)/\alpha^* H^0(\mathbb{P}^1, T_{\mathbb{P}^1}) \to H^1(C, T_C)$$

fails to be surjective precisely when $H^1(C, \alpha^* T_{\mathbb{P}^1}) \neq 0$, i.e., when $\alpha \in M(E[2])$. ☐

2B. ***Linear independence of T, D, and Δ.*** In this section, we prove that the divisorial components of the boundary of $\widetilde{\mathcal{H}}_{d,g}$ are linearly independent. Define the closed loci T, D, Δ in $\widetilde{\mathcal{H}}_{d,g}$ by

$$T = \overline{\{[\alpha : C \to \mathbb{P}^1] \mid \alpha^{-1}(q) = 3p_1 + p_2 + \cdots + p_{d-2} \text{ for some } q \text{ and distinct } p_i.\}}$$

$$D = \overline{\{[\alpha : C \to \mathbb{P}^1] \mid \alpha^{-1}(q) = 2p_1 + 2p_2 + p_3 + \cdots + p_{d-2}} \\ \overline{\text{for some } q \text{ and distinct } p_i.\}}$$

$$\Delta = \overline{\{[\alpha : C \to \mathbb{P}^1] \mid C \text{ is singular.}\}}$$

These three loci correspond to the three possibilities of the limit when two branch points of a branched cover come together. Note that T, D, and Δ are irreducible and their union is the complement of $\mathcal{H}_{d,g}$ in $\widetilde{\mathcal{H}}_{d,g}$.

Proposition 2.15. *For $d \geq 4$, the classes of T, D, and Δ are linearly independent in $\mathrm{Pic}_{\mathbb{Q}}(\widetilde{\mathcal{H}}_{d,g})$. For $d \geq 3$, the same is true for the classes of T and Δ.*

Proof. We construct curves with nonsingular intersection matrix with our divisors. For this, a slight enlargement of $\widetilde{\mathcal{H}}_{d,g}$ is more convenient. Define $\widetilde{\mathcal{H}}^{ns}_{d,g}$ as the moduli space of $[\alpha : C \to \mathbb{P}^1]$, where C is an at worst nodal curve of arithmetic genus g, not necessarily irreducible, but without any separating nodes, and α is a map of degree d. The target $\mathbb{P}^1$ is taken to be unframed. It is easy to see that $\widetilde{\mathcal{H}}_{d,g}$ is a dense open subset of $\widetilde{\mathcal{H}}^{ns}_{d,g}$ with codimension-2 complement. Abusing notation,

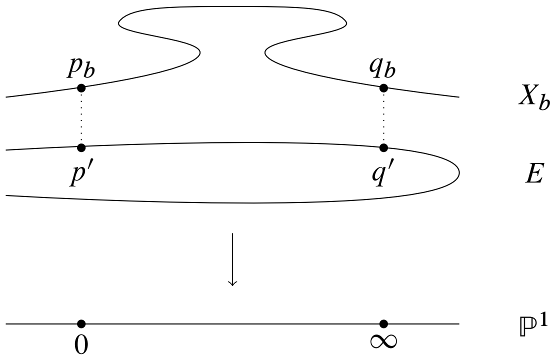


Figure 1. We construct families of covers parametrized by $b \in B$ by attaching a variable family of covers $\alpha_b : X_b \to \mathbb{P}^1$ to a fixed cover $\beta : E \to \mathbb{P}^1$.

we denote the closures of T, D, and Δ in $\widetilde{\mathcal{H}}^{ns}_{d,g}$ by the same letters. It suffices to prove the proposition for $\widetilde{\mathcal{H}}^{ns}_{d,g}$.

We now construct test curves in $\widetilde{\mathcal{H}}^{ns}_{d,g}$. Pick nonnegative integers g_1 and g_2 with $g_1 + g_2 = g - 1$ and positive integers d_1 and d_2 with $d_1 + d_2 = d$. Take a family $\alpha_b : X_b \to \mathbb{P}^1$ of covers of degree d_1 and genus g_1, where b denotes a parameter on a smooth complete curve B. Assume that we have two sections $p, q : B \to X$ with $\alpha_b(p_b) = 0$ and $\alpha_b(q_b) = \infty$ for all $b \in B$. Take $\beta : E \to \mathbb{P}^1$ to be a fixed simply branched cover of degree d_2 and genus g_2, unramified over 0 and ∞, and let $p', q' \in E$ be two points over 0 and ∞ respectively. Our test curve in $\widetilde{\mathcal{H}}^{ns}_{d,g}$ is given by the family $\gamma_b : C_b \to \mathbb{P}^1$, where C_b is obtained by gluing (X_b, p_b, q_b) to the constant family (E, p', q'), and $\gamma_b : C_b \to \mathbb{P}^1$ is induced from $\alpha : X_b \to \mathbb{P}^1$ and $\beta : E \to \mathbb{P}^1$. The construction is depicted in Figure 1.

Let T_α, D_α, and Δ_α denote the pullbacks of the divisor classes T, D, and Δ along the map from B to $\widetilde{\mathcal{H}}_{d_1,g_1}$ given by α_b. Define T_γ, D_γ, and Δ_γ likewise. Let e be the intersection number of $\mathrm{Br}(\alpha)$ with a horizontal section of $\mathbb{P}^1 \times B$. Denote by $[p]$ and $[q]$ respectively the classes of $p(B)$ and $q(B)$ on X.

Claim. *With the notation above, we have*

$$\begin{aligned}
\deg T_\gamma &= \deg T_\alpha + 3([p]+[q]) \cdot \mathrm{Ram}(\alpha),\\
\deg D_\gamma &= \deg D_\alpha + (2g_2 + 2d_2 - 2)e + 4e - 4([p]+[q]) \cdot \mathrm{Ram}(\alpha), \ \textit{and}\\
\deg \Delta_\gamma &= \deg \Delta_\alpha + [p]^2 + [q]^2.
\end{aligned}$$

Proof of the claim. The pullback of the line bundle $\mathcal{O}(\Delta)$ from $\widetilde{\mathcal{H}}^{ns}_{d,g}$ to B is given by

$$(N_{p/X} \otimes N_{p'/E}) \otimes (N_{q/E} \otimes N_{q'/E}) \otimes \mathcal{O}_B(\Delta_\alpha),$$

where $N_{p/X}$ denotes the normal bundle of p in X, and so on. The third equation follows.

For a generic $b \in B$, the point of $\widetilde{\mathcal{H}}^{ns}_{d,g}$ given by $\gamma_b : C_b \to \mathbb{P}^1$ does not lie in T or D. We have the following specializations:

(1) $\alpha_b : X_b \to \mathbb{P}^1$ has a fiber of the form $3p_1 + p_2 + \cdots$. Such b are precisely the points of T_α, each contributing 1 to $\deg T_\gamma$.

(2) $\alpha_b : X_b \to \mathbb{P}^1$ has a fiber of the form $2p_1 + 2p_2 + p_3 + \cdots$. Such b are precisely the points of D_α, each contributing 1 to $\deg D_\gamma$

(3) A branch point of $\alpha_b : X_b \to \mathbb{P}^1$ coincides with a branch point of $\beta : E \to \mathbb{P}^1$. There are $(2g_2 + 2d_2 - 2)e$ such b, each contributing 1 to $\deg D_\gamma$.

(4) p_b (resp. q_b) is a ramification point of α_b. We compute the intersection multiplicity of B with T and D at such a point by looking at a versal deformation space of γ_b. We may restrict γ_b over an analytic neighborhood U of 0 (resp. ∞). Let x be a coordinate on U. Then $\gamma_b^{-1}(U) \to U$ has the form

$$U[y]/(y^3 - xy) \sqcup U \sqcup \cdots \sqcup U \to U.$$

A versal deformation of this cover is given over $\operatorname{Spec} \mathbb{C}[s,t]$ by

$$U[y]/(y^3 - xy - sx - t) \sqcup U \sqcup \cdots \sqcup U \to U.$$

In $\operatorname{Spec} \mathbb{C}[s,t]$, the divisor D does not contain the origin, and hence the intersection number of B with D at b is 0. The divisor $T \subset \operatorname{Spec} \mathbb{C}[s,t]$ is defined by $t = 0$. The curve B approaches the origin along the locus where $U[y]/(y^3 - xy - sx - t)$ is singular, namely along $s^3 + t = 0$. We deduce that the intersection number of B with T at b is 3. There are $[p] \cdot \operatorname{Ram}(\alpha)$ (resp. $[q] \cdot \operatorname{Ram}(\alpha)$) such b, each contributing 3 to $\deg T_\gamma$.

(5) p_b (resp. q_b) is not a ramification point of α_b, but lies over a branch point. Again, we look at a versal deformation of γ_b. In this case, $\gamma_b^{-1}(U) \to U$ has the form

$$U[y]/(y^2 - x) \sqcup U[z]/(z^2 - x^2) \sqcup U \sqcup \cdots \sqcup U \to U.$$

A versal deformation of this cover is given over $\operatorname{Spec} \mathbb{C}[s,t]$ by

$$U[y]/(y^2 - x) \sqcup U[z]/(z^2 - x^2 - sx - t) \sqcup U \sqcup \cdots \sqcup U \to U.$$

In $\operatorname{Spec} \mathbb{C}[s,t]$, the divisor T does not contain the origin, and hence the intersection number of B with T at b is 0. The divisor $D \subset \operatorname{Spec} \mathbb{C}[s,t]$ is defined by $t = 0$. The curve B approaches the origin along the locus where $U[z]/(z^2 - x^2 - sx - t)$ is singular, namely along $s^2 - 4t = 0$. We deduce that the intersection number of B with D at b is 2. Let us count the number of such points, first for p_b, and analogously for q_b. The points b for which p_b is not a ramification point but lies over a branch point correspond to the intersection points of $\operatorname{Br}(\alpha) \cap \{0\} \times B$ which are not the images of the points of $\mathrm{Ram}(\alpha) \cap p(B)$. Note, however, that the image of a point of $\mathrm{Ram}(\alpha) \cap p(B)$ is actually a point of tangency of $\mathrm{Br}(\alpha)$ with $\{0\} \times B$, and hence contributes 2 to the intersection number $e = \mathrm{Br}(\alpha) \cdot \{0\} \times B$. The remaining count, which we want, is therefore $e - 2[p] \cdot \mathrm{Ram}(\alpha)$. Similarly, the count for q_b is $e - 2[q] \cdot \mathrm{Ram}(\alpha)$.

The expressions for T_γ and D_γ follow from combining the above contributions. □

Returning to the proof of the proposition, consider the following three particular test curves for $d \geq 4$.

B_1: Take $\alpha_b : X_b \to \mathbb{P}^1$ to be a family of hyperelliptic curves of genus $g-1$ obtained by taking a double cover $X \to \mathbb{P}^1 \times \mathbb{P}^1$ branched along a curve of type $(2g, 2)$. To have sections p and q of X over $\{0\} \times \mathbb{P}^1$ and $\{\infty\} \times \mathbb{P}^1$, let the branch divisor be tangent to $\{0\} \times \mathbb{P}^1$ and $\{\infty\} \times \mathbb{P}^1$. Take E to be a smooth rational curve and $\gamma : E \to \mathbb{P}^1$ a generic cover of degree $d-2$.

B_2: Take $\alpha_b : X_b \to \mathbb{P}^1$ to be a family of trigonal curves of genus $g-1$ obtained by taking a general pencil on $\mathbb{F}_0$ in the linear system $|((g+1)/2, 3)|$ if g is odd, or on $\mathbb{F}_1$ in the linear system $|3 \cdot \text{directrix} + (g/2+2) \cdot \text{fiber}|$ if g is even. Two base-points give p_b and q_b. Take E to be a rational curve and $\gamma : E \to \mathbb{P}^1$ a general cover of degree $d-3$.

B_3: Take $\alpha_b : X_b \to \mathbb{P}^1$ to be a family of hyperelliptic curves of genus $g-2$ as in B_1. Take E to be a smooth genus-1 curve and $\gamma : E \to \mathbb{P}^1$ a generic cover of degree $d-2$. This curve exists only for $d \geq 4$.

Using the claim, we get the following nonsingular intersection matrix:

	T	D	Δ
B_1	6	$4d-12$	$8g-6$
B_2	$3g+9$	$8d-24$	$7g-3$
B_3	6	$4d-8$	$8g-14$

For $d=3$, we take a pencil in $\mathbb{F}_0$ or $\mathbb{F}_1$ as in B_1, but of trigonal curves of genus g, without any E. Then the middle column vanishes, and the second row becomes $(3g+6, 0, 7g+6)$, which is linearly independent from the first row. □

3. Degree 3

Let C be a curve of genus g and $\alpha : C \to \mathbb{P}^1$ a map of degree 3. The relative canonical map embeds C as a divisor in a $\mathbb{P}^1$-bundle $\mathbb{P}E$ over $\mathbb{P}^1$, where E is a vector bundle of rank 2 and degree $g+2$.

Let

$$E^{\text{gen}} = \mathcal{O}\left(\left\lfloor \frac{g+2}{2} \right\rfloor\right) \oplus \mathcal{O}\left(\left\lceil \frac{g+2}{2} \right\rceil\right)$$

be the most generic vector bundle on $\mathbb{P}^1$ of rank 2 and degree $g+2$. Set

$$U_{E^{\text{gen}}} := \{\alpha \in \widetilde{\mathcal{H}}_{3,g} \mid E_\alpha \cong E^{\text{gen}}\}.$$

Note that $U_{E^{\text{gen}}}$ is an open subset of $\widetilde{\mathcal{H}}_{3,g}$.

Proposition 3.1. *The complement of $U_{E^{\text{gen}}}$ in $\widetilde{\mathcal{H}}_{3,g}$ is a divisor if and only if g is even, in which case it is irreducible.*

Proof. This is the degree-3 case of Proposition 2.13. □

Let $\pi : \mathbb{P}E^{\text{gen}} \to \mathbb{P}^1$ be the projection. Set

$$V = H^0(\mathbb{P}^1, \operatorname{Sym}^3 E^{\text{gen}} \otimes \det E^{\text{gen}\vee}).$$

Elements of $\mathbb{P}_{\text{sub}} V$ correspond to divisors in the linear series of the line bundle $\mathcal{O}_{\mathbb{P}E^{\text{gen}}}(3) \otimes \pi^*(\det E^{\text{gen}})^\vee$ on $\mathbb{P}E^{\text{gen}}$. Let $C_v \subset \mathbb{P}E^{\text{gen}}$ be the divisor corresponding to $v \in V$. Let $V^\circ \subset \mathbb{P}_{\text{sub}} V$ be the open locus consisting of $v \in V^\circ$ for which C_v is irreducible and at worst nodal. Let $G := \operatorname{Aut}(\pi)$ be the group of automorphisms of $\mathbb{P}E^{\text{gen}}$ over $\mathbb{P}^1$. Then G acts on V°. The assignment

$$v \mapsto [\pi : C_v \to \mathbb{P}^1]$$

gives a map

$$q : V^\circ \to \widetilde{\mathcal{H}}^\dagger_{3,g}.$$

Denote by $U^\dagger_{E^{\text{gen}}}$ the preimage of $U_{E^{\text{gen}}}$ under $\widetilde{\mathcal{H}}^\dagger_{3,g} \to \widetilde{\mathcal{H}}_{3,g}$.

Proposition 3.2. *The image of q is $U^\dagger_{E^{\text{gen}}}$. The fibers of q consist of single G-orbits.*

Proof. For brevity, set $E = E^{\text{gen}}$. For $v \in V^\circ$, consider the sequence

$$0 \longrightarrow \mathcal{O}_{\mathbb{P}E}(-3) \otimes \pi^* \det E \longrightarrow \mathcal{O}_{\mathbb{P}E} \longrightarrow \mathcal{O}_{C_v} \longrightarrow 0.$$

Applying $R\pi_*$, we get

$$0 \longrightarrow \mathcal{O}_{\mathbb{P}^1} \longrightarrow \pi_* \mathcal{O}_{C_u} \longrightarrow E^\vee \longrightarrow 0, \tag{3-1}$$

which says that the Tschirnhausen bundle of $C_u \to \mathbb{P}^1$ is E. Conversely, from the Casnati–Ekedahl resolution, it follows that every point of $U^\dagger_{E^{\text{gen}}}$ is in the image of q.

Let $u, v \in U^\dagger_{E^{\text{gen}}}$ be in a fiber of q. Then there is an isomorphism $C_u \to C_v$ over the identity of $\mathbb{P}^1$. The sequence (3-1) for C_u and C_v shows that such an isomorphism induces an isomorphism $E \to E$. The induced automorphism of $\mathbb{P}E$ over $\mathbb{P}^1$ takes C_u to C_v and hence u to v. □

Proposition 3.3 (Picard rank conjecture for degree 3). *We have* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{3,g} = 0$.

Proof. Retain the notation introduced above. For brevity, set $U = U_{E^{\mathrm{gen}}}$ and $U^\dagger = U^\dagger_{E^{\mathrm{gen}}}$. Then $V^\circ \to U^\dagger$ is a quotient by G and $U^\dagger \to U$ is a quotient by PGL_2. By Proposition 1.2 and Proposition 3.2, we have

$$\begin{aligned}\operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} U \leq \operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} U^\dagger + \operatorname{rk}\chi(\mathrm{PGL}_2) &= \operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} U^\dagger \\ &\leq \operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} V^\circ + \operatorname{rk}\chi(G) \leq 1 + \operatorname{rk}\chi(G).\end{aligned}$$

The final inequality follows because V° is an open subset of a projective space. Let e be the number of divisorial components of $\widetilde{\mathcal{H}}_{3,g} \setminus U$. We then get the bound

$$\operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{3,g} \leq \operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} U + e \leq 1 + \operatorname{rk}\chi(G) + e.$$

If g is even, then

$$\begin{aligned} G &= \mathrm{PGL}_2, \\ \operatorname{rk}\chi(G) &= 0, \\ e &= 1 \quad \text{by Proposition 3.1.}\end{aligned}$$

If g is odd, then

$$\begin{aligned} G &= \left\{ \begin{pmatrix} a & l \\ 0 & b \end{pmatrix} \,\middle|\, a, b \in \mathbb{C}^*,\ l \in H^0(\mathbb{P}^1, \mathcal{O}(1)) \right\} \Big/ \mathbb{C}^*, \\ \operatorname{rk}\chi(G) &= 1, \\ e &= 0 \quad \text{by Proposition 3.1.}\end{aligned}$$

In either case, we have

$$\operatorname{rk}\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{3,g} \leq 2.$$

By Proposition 2.15, the classes in $\operatorname{Pic}_{\mathbb{Q}}(\widetilde{\mathcal{H}}_{3,g})$ of the two components of $\widetilde{\mathcal{H}}_{3,g} \setminus \mathcal{H}_{3,g}$ are linearly independent. Therefore, we get $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{3,g} = 0$ as desired. □

4. Degree 4

Let C be a curve of genus g and $\alpha : C \to \mathbb{P}^1$ a map of degree 4. The relative canonical map embeds C into a $\mathbb{P}^2$-bundle $\mathbb{P}E$ over $\mathbb{P}^1$, where E is a vector bundle of rank 3 and degree $g+3$. The Casnati–Ekedahl structure theorem provides the following resolution of $\mathcal{O}_C$:

$$0 \longrightarrow \pi^* \det E(-4) \longrightarrow \pi^* F(-2) \longrightarrow \mathcal{O}_{\mathbb{P}E} \longrightarrow \mathcal{O}_C \longrightarrow 0,$$

where F is a vector bundle of rank 2 and degree $g+3$.

Explicitly, we can describe $C \subset \mathbb{P}E$ as follows. Write $F = \mathcal{O}(a) \oplus \mathcal{O}(b)$, where $a + b = g + 3$ and $a \leq b$. Let h denote the divisor class associated to $\mathcal{O}_{\mathbb{P}E}(1)$ on $\mathbb{P}E$ and f the class of the fiber of the projection $\pi : \mathbb{P}E \to \mathbb{P}^1$. Then the curve C

is the complete intersection of two divisors:

$$C = Q_a \cap Q_b,$$

where $[Q_a] = 2h - af$ and $[Q_b] = 2h - bf$.

We can describe the equations of Q_a and Q_b even more explicitly as follows. Write $E = \mathcal{O}(m_1) \oplus \mathcal{O}(m_2) \oplus \mathcal{O}(m_3)$. Over an open set $U \subset \mathbb{P}^1$, let X, Y, and Z denote the relative coordinates on $\mathbb{P}E|_U$ corresponding to the three summands of E. Assume that $m_1 \leq m_2 \leq m_3$. Over U, the divisor Q_a is the zero locus of a form

$$p_{1,1}X^2 + p_{1,2}XY + p_{1,3}XZ + p_{2,2}Y^2 + p_{2,3}YZ + p_{3,3}Z^2, \tag{4-1}$$

where $p_{i,j}$ is the restriction to U of a global section of $\mathcal{O}(m_i + m_j - a)$. Similarly, over U, the divisor Q_b is the zero locus of a form

$$q_{1,1}X^2 + q_{1,2}XY + q_{1,3}XZ + q_{2,2}Y^2 + q_{2,3}YZ + q_{3,3}Z^2, \tag{4-2}$$

where $q_{i,j}$ is the restriction to U of a global section of $\mathcal{O}(m_i + m_j - b)$.

The irreducibility of C puts some restrictions on the possible (E, F). Indeed, if $p_{1,1} = q_{1,1} = 0$, then the section $[X : Y : Z] = [1 : 0 : 0]$ of $\mathbb{P}E$ is contained in both Q_a and Q_b, making $C = Q_a \cap Q_b$ reducible. An irreducible C thus forces

$$2m_1 \geq a. \tag{4-3}$$

Proposition 4.1. *Let E be a vector bundle of rank 3 and degree $g+3$ and F a vector bundle of rank 2 and degree $g+3$. If the locus $M(E, F)$ is nonempty, then it is irreducible and unirational.*

Proof. Consider the dense open subset $M^\circ(E, F) \subset M(E, F)$ corresponding to those $\alpha \in \mathcal{H}_{4,g}$ that have $E_\alpha \cong E$ and $F_\alpha \cong F$. It suffices to prove the statement for $M^\circ(E, F)$. Consider the vector space

$$V := H^0(\mathbb{P}^1, F^\vee \otimes \operatorname{Sym}^2 E).$$

Elements of V correspond to maps $\pi^* F(-2) \to \mathcal{O}_{\mathbb{P}E}$. Let $V^\circ \subset V$ be the open subset where the ideal generated by the image of $\pi^* F(-2)$ defines a smooth curve, simply branched over $\mathbb{P}^1$. Then V° surjects onto $M^\circ(E, F)$. □

Remark 4.2. From the dominant map $V^\circ \to M(E, F)$ in the proof of Proposition 4.1, it is easy to compute the codimension of $M(E, F)$ in $\mathcal{H}_{4,g}$, which is

$$\operatorname{codim} M(E, F) = \dim \operatorname{Ext}^1(E, E) + \dim \operatorname{Ext}^1(F, F) - \dim \operatorname{Ext}^1(F, \operatorname{Sym}^2 F).$$

We may think of $\dim \operatorname{Ext}^1(E, E) + \dim \operatorname{Ext}^1(F, F)$ as the "expected codimension". The next example shows that the actual codimension is not always the expected codimension.

Example 4.3. Let $E = \mathcal{O}(m) \oplus \mathcal{O}(2m) \oplus \mathcal{O}(g+3-3m)$, where $\lceil (g+3)/6 \rceil \leq m < (g+3)/5$. To get an irreducible curve C, the only possibility for F is $F = \mathcal{O}(2m) \oplus \mathcal{O}(g+3-2m)$, by (4-3). The resulting locus $M(E, F)$ is not of expected codimension because $\dim \operatorname{Ext}^1(F, \operatorname{Sym}^2 E)$ is nonzero.

Example 4.4. The Maroni locus $M(E)$ may be reducible. Let $g = 12$, and consider the bundle $E = \mathcal{O}(3) \oplus \mathcal{O}(5) \oplus \mathcal{O}(7)$. Then the reader can easily check (using Bertini's theorem) that $M(E, F)$ and $M(E, F')$ are nonempty and of *equal* codimension $\dim \operatorname{Ext}^1(E, E)$ for the bundles $F = \mathcal{O}(6) \oplus \mathcal{O}(9)$ and $F' = \mathcal{O}(5) \oplus \mathcal{O}(10)$. Therefore $M(E, F)$ and $M(E, F')$ are two components of $M(E)$. It is easy to see by analyzing the explicit equations that these are the only components of $M(E)$.

Let E^{gen} (resp. F^{gen}) be the most generic vector bundle on $\mathbb{P}^1$ of rank 3 (resp. 2) and degree $g+3$. Define

$$\begin{aligned} U_{E^{\text{gen}}} &:= \{\alpha \in \widetilde{\mathcal{H}}_{4,g} \mid E_\alpha \cong E^{\text{gen}}\}, \\ U_{F^{\text{gen}}} &:= \{\alpha \in \widetilde{\mathcal{H}}_{4,g} \mid F_\alpha \cong F^{\text{gen}}\}, \\ U_{E^{\text{gen}}, F^{\text{gen}}} &:= U_{E^{\text{gen}}} \cap U_{F^{\text{gen}}}. \end{aligned}$$

It is easy to see that these are open subsets of $\widetilde{\mathcal{H}}_{d,g}$. Our next task is to identify the divisorial components of their complements.

Proposition 4.5. *The subvariety $M := \widetilde{\mathcal{H}}_{4,g} \setminus U_{E^{\text{gen}}}$ is a divisor if and only if g is divisible by* 3, *in which case it is irreducible.*

Proof. This is the degree-4 case of Proposition 2.13. □

For the complement of $U_{F^{\text{gen}}}$, we could do a careful analysis of the defining equations of C in $\mathbb{P}E$, as we will have to do for the next case of $d = 5$. But we can take a more geometric approach using the *resolvent cubic construction*. Originally due to Recillas [1973], the construction can be described as follows. For simplicity, we give an informal description, restricting to simply branched covers. See [Casnati 1998] for a detailed account. Consider a point $[\alpha : C \to \mathbb{P}^1]$ of $\mathcal{H}_{4,g}$. The resolution of $\mathcal{O}_C$ as an $\mathcal{O}_{\mathbb{P}E_\alpha}$-module shows that $C \subset \mathbb{P}E_\alpha$ is the complete intersection of two relative quadrics. A fiber of $\mathbb{P}F_\alpha \to \mathbb{P}^1$ naturally corresponds to the pencil of conics in the corresponding fiber of $\mathbb{P}E_\alpha \to \mathbb{P}^1$ containing the corresponding fiber of $C \to \mathbb{P}^1$. Each such pencil contains three singular conics, counted with multiplicity. The total locus of these singular conics forms a trigonal curve $R(C) \subset \mathbb{P}F_\alpha$. Let $R(\alpha) : R(C) \to \mathbb{P}^1$ be the projection. We call $R(\alpha)$ the *resolvent cubic* of α. Using that $C \to \mathbb{P}^1$ is simply branched, it is easy to check that $R(C)$ is smooth and the branch divisor of $R(\alpha)$ coincides with the branch divisor of α. In particular, $R(C)$ has genus $g+1$. The association $\alpha \to R(\alpha)$ defines a map

$$R : \mathcal{H}_{4,g} \to \mathcal{H}_{3,g+1},$$

which we call the *resolvent cubic map*. The fiber of R over a point $[D \to \mathbb{P}^1]$ in $\mathcal{H}_{3,g+1}$ corresponds bijectively to the set of étale double covers $D' \to D$ (see [Recillas 1973], [Casnati 1998, Theorem 6.5], or [Donagi 1981]). In particular, R is a finite morphism.

Proposition 4.6. *Let F be a vector bundle of rank 2 and degree $g+3$ on $\mathbb{P}^1$. The Casnati–Ekedahl locus $C(F) \subset \mathcal{H}_{4,g}$ is nonempty if and only if $\lfloor F \rfloor \geq \lceil (g+3)/3 \rceil$. In this case, it is of the expected codimension* $\dim \operatorname{Ext}^1(F, F)$.

Proof. Consider a point $[\alpha : C \to \mathbb{P}^1]$ of $\mathcal{H}_{4,g}$ and its resolvent cubic $R(\alpha) : R(C) \to \mathbb{P}^1$. Since $R(C) \subset \mathbb{P}F_\alpha$, and F_α is a vector bundle of rank 2 and degree $(g+1)+2$, it must be the Tschirnhausen bundle of $R(C)$. That is, we have $E_{R(\alpha)} = F_\alpha$. By [Recillas 1973], the map R is finite, and hence $C(F) = R^{-1}(M(F))$. Both statements about $C(F)$ now follow from the corresponding statements about $M(F)$. $\square$

Proposition 4.7. *Let $g \geq 4$. The subvariety $CE := \mathcal{H}_{4,g} \setminus U_{F^{\text{gen}}}$ has codimension at least 2 if g is even and is an irreducible divisor if g is odd.*

Proof. The image $R(U_{F^{\text{gen}}}) \subset \mathcal{H}_{3,g+1}$ is the open locus of trigonal covers having F^{gen} as their Tschirnhausen bundle. The complement $Z := \mathcal{H}_{3,g+1} \setminus R(U_{F^{\text{gen}}})$ has codimension at least 2 if $g+1$ is odd, and it is the Maroni divisor if $g+1$ is even (Proposition 3.1). The complement $\mathcal{H}_{4,g} \setminus U_{F^{\text{gen}}}$ is the preimage $R^{-1}(Z)$. Therefore, the statements about the codimension follow from the finiteness of R.

For the question of reducibility, let $F = \mathcal{O}(k-1) \oplus \mathcal{O}(k+1)$ with $k = (g+3)/2 \geq 3$. The claim is that $C(F)$ is irreducible when $g > 3$, and has two components when $g = 3$. We have

$$C(F) = \bigcup_E M(E, F).$$

By Proposition 4.1, the varieties $M(E, F)$ are irreducible. Therefore, every component of $C(F)$ must be of the form $M(E, F)$ for some E.

Let $g > 3$ and suppose $E \neq E^{\text{gen}}$. The inclusion $M(E, F) \subset M(E)$ and Proposition 2.13 imply that $M(E, F)$ is a divisor if and only if $M(E, F) = M(E)$ and $E = \mathcal{O}(m-1) \oplus \mathcal{O}(m) \oplus \mathcal{O}(m+1)$. By choosing two generic quadrics as in (4-1) and (4-2), we can explicitly construct a curve in $M(E, F^{\text{gen}})$, showing that $M(E, F) \neq M(E)$. Thus, it follows that the only component of $C(F)$ is $M(E^{\text{gen}}, F)$. $\square$

Example 4.8. The divisor $\mathcal{H}_{4,g} \setminus U_{F^{\text{gen}}}$ is not irreducible for $g = 3$. Indeed, take $F = \mathcal{O}(2) \oplus \mathcal{O}(4)$. Then $M(E^{\text{gen}}, F)$ is an irreducible component. Now consider the only other possibility for E, namely $E = \mathcal{O}(1) \oplus \mathcal{O}(2) \oplus \mathcal{O}(3)$. By (4-3), a cover in $M(E)$ must have $F = \mathcal{O}(2) \oplus \mathcal{O}(4)$. Furthermore, for this E and F, we can choose the two quadrics generically and see that $M(E, F)$ is nonempty. Therefore, $M(E) = M(E, F)$ is another component of $\mathcal{H}_{4,g} \setminus U_{F^{\text{gen}}}$.

Our next goal is to exhibit $U_{E^{\mathrm{gen}},F^{\mathrm{gen}}}$ as a quotient. Let $\pi : \mathbb{P}E^{\mathrm{gen}} \to \mathbb{P}^1$ be the projection. For brevity, set $E = E^{\mathrm{gen}}$ and $F = F^{\mathrm{gen}}$. Set

$$V := H^0(\mathbb{P}^1, F^\vee \otimes \mathrm{Sym}^2 E).$$

An element $v \in \mathbb{P}_{\mathrm{sub}} V$ corresponds to a map $\pi^* F(-2) \to \mathcal{O}_{\mathbb{P}E}$. Let C_v be the zero locus of the image of this map. Let $V^\circ \subset \mathbb{P}_{\mathrm{sub}} V$ be the open locus consisting of $v \in \mathbb{P}_{\mathrm{sub}} V$ for which C_v is irreducible and at worst nodal. Let $G_F := \mathrm{Aut}(\mathbb{P}F/\mathbb{P}^1)$ and $G_E := \mathrm{Aut}(\mathbb{P}E/\mathbb{P}^1)$. Then $G_F \times G_E$ acts on V°. The assignment

$$v \mapsto [\pi : C_v \to \mathbb{P}^1]$$

defines a map

$$q : V^\circ \to \widetilde{\mathcal{H}}^\dagger_{4,g}.$$

Denote by $U^\dagger_{E,F}$ the preimage of $U_{E,F}$ under $\widetilde{\mathcal{H}}^\dagger_{4,g} \to \widetilde{\mathcal{H}}_{4,g}$.

Proposition 4.9. *The image of q is $U^\dagger_{E^{\mathrm{gen}},F^{\mathrm{gen}}}$. The fibers of q consist of single G-orbits.*

Proof. The proof is exactly analogous to the proof of Proposition 3.2. □

Proposition 4.10 (Picard rank conjecture for degree 4). *We have* $\mathrm{Pic}_{\mathbb{Q}} \mathcal{H}_{4,g} = 0$.

Proof. Retain the notation introduced above. For brevity, set $U = U_{E^{\mathrm{gen}},F^{\mathrm{gen}}}$ and $U^\dagger = U^\dagger_{E^{\mathrm{gen}},F^{\mathrm{gen}}}$. By Proposition 1.2 and Proposition 4.9, we have

$$\begin{aligned}\mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} U &\leq \mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} U^\dagger + \mathrm{rk}\,\chi(\mathrm{PGL}_2) = \mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} U^\dagger \\ &\leq \mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} V^\circ + \mathrm{rk}\,\chi(G) \leq 1 + \mathrm{rk}\,\chi(G).\end{aligned}$$

The final inequality follows because V° is an open subset of a projective space. Let e be the number of divisorial components of $\widetilde{\mathcal{H}}_{3,g} \setminus U$. We then get the bound

$$\mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{4,g} \leq \mathrm{rk}\,\mathrm{Pic}_{\mathbb{Q}} U + e \leq 1 + \mathrm{rk}\,\chi(G) + e.$$

Recall that $G = G_{F^{\mathrm{gen}}} \times G_{E^{\mathrm{gen}}}$.

If g is an odd multiple of 3, then

$G = \mathrm{PGL}_2 \times \mathrm{PGL}_3,$

$\mathrm{rk}\,\chi(G) = 0,$

$e = 2$ corresponding to M in Proposition 4.5 and CE in Proposition 4.7.

If g is odd, but not divisible by 3, then

$$G = \mathrm{PGL}_2 \times G_E$$

$$G_E = \left\{ \begin{pmatrix} a & b & l_1 \\ c & d & l_2 \\ 0 & 0 & e \end{pmatrix} \,\middle|\, a,b,c,d,e \in \mathbb{C}, e(ad-bc) \in \mathbb{C}^*, l_i \in H^0(\mathbb{P}^1, \mathcal{O}(1)) \right\} \Big/ \mathbb{C}^*,$$

$$\operatorname{rk} \chi(G) = 1,$$

$$e = 1 \quad \text{corresponding to } CE \text{ in Proposition 4.7.}$$

If g is even and divisible by 3, then

$$G = G_F \times \mathrm{PGL}_2,$$

$$G_F = \left\{ \begin{pmatrix} a & l \\ 0 & b \end{pmatrix} \,\middle|\, a, b \in \mathbb{C}^*, l \in H^0(\mathbb{P}^1, \mathcal{O}(1)) \right\} \Big/ \mathbb{C}^*,$$

$$\operatorname{rk} \chi(G) = 1,$$

$$e = 1 \quad \text{corresponding to } M \text{ in Proposition 4.5.}$$

If g is even and not divisible by 3, then

$$G = G_F \times G_E, \text{ where } G_F \text{ and } G_E \text{ are as in the previous two cases,}$$

$$\operatorname{rk} \chi(G) = 2,$$

$$e = 0.$$

In all cases, we get

$$\operatorname{rk} \operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{4,g} \leq 3.$$

By Proposition 2.15, the classes in $\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{4,g}$ of the three components of $\widetilde{\mathcal{H}}_{4,g} \setminus \mathcal{H}_{4,g}$ are linearly independent. Therefore, we get $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{4,g} = 0$, as desired. □

5. Degree 5

Let C be a curve of genus g and $\alpha : C \to \mathbb{P}^1$ a map of degree 5. The relative canonical map embeds C into a $\mathbb{P}^3$ bundle $\mathbb{P}E$ over $\mathbb{P}^1$, where E is a vector bundle of rank 4 and degree $g+4$. The Casnati–Ekedahl structure theorem provides the following resolution of $\mathcal{O}_C$:

$$0 \longrightarrow \pi^* \det E(-5) \longrightarrow \pi^*(F^\vee(\det E))(-3) \longrightarrow \pi^* F(-2) \longrightarrow \mathcal{O}_{\mathbb{P}E} \longrightarrow \mathcal{O}_C \longrightarrow 0,$$

where F is a vector bundle of rank 3 and degree $2g+8$.

Explicitly, we can describe $C \subset \mathbb{P}E$ as follows. The resolution is determined completely by the middle map

$$w : \pi^*(F^\vee(\det E))(-3) \to \pi^* F(-2).$$

We can view this map as an element of the vector space $H^0(\mathbb{P}^1, F \otimes F \otimes E(-\det E))$. Due to a theorem of [Casnati 1996], w can be taken to be antisymmetric, that is, in the subspace

$$V := H^0\big(\mathbb{P}^1, \textstyle\bigwedge^2 F \otimes E \otimes \det E^\vee\big).$$

Even more explicitly, we can describe the defining equations of C as follows. Let

$$F = \mathcal{O}(n_1) \oplus \cdots \oplus \mathcal{O}(n_5), \quad \text{where } n_1 \le \cdots \le n_5, \text{ and}$$
$$E = \mathcal{O}(m_1) \oplus \cdots \oplus \mathcal{O}(m_4), \quad \text{where } m_1 \le \cdots \le m_4.$$

We represent an element $w \in V$ by a skew symmetric matrix of forms

$$M_w = \begin{pmatrix} 0 & L_{1,2} & L_{1,3} & L_{1,4} & L_{1,5} \\ -L_{1,2} & 0 & L_{2,3} & L_{2,4} & L_{2,5} \\ -L_{1,3} & -L_{2,3} & 0 & L_{3,4} & L_{3,5} \\ -L_{1,4} & -L_{2,4} & -L_{3,4} & 0 & L_{4,5} \\ -L_{1,5} & -L_{2,5} & -L_{3,5} & -L_{4,5} & 0 \end{pmatrix}, \tag{5-1}$$

where $L_{i,j} \in H^0(\mathbb{P}^1, E \otimes \det E^\vee \otimes \mathcal{O}(n_i + n_j))$. In $\mathbb{P}E$, the curve C_w is cut out by the 4×4 sub-Pfaffians of the matrix M_w.

The irreducibility of C puts some restrictions on the possible matrices. Indeed, suppose

$$L_{1,2} = L_{1,3} = 0.$$

Then the Pfaffian Q_5 of the submatrix obtained by eliminating the fifth row and column is

$$Q_5 = L_{1,2}L_{3,4} - L_{1,3}L_{2,4} + L_{2,3}L_{1,4} = L_{2,3}L_{1,4}.$$

Since Q_5 is reducible, C_w is forced to be reducible.

Suppose further that $E = \mathcal{O}(k)^r \oplus \mathcal{O}(k+1)^{4-r}$, where $0 \le r \le 3$. Then the observation above implies that the maximum of the degrees of the summands of $E \otimes (\det E^\vee) \otimes \mathcal{O}(n_1 + n_3)$ must be nonnegative, meaning

$$n_1 + n_3 + k - (g+4) \ge -1.$$

Since the n_i are increasing, we get the inequalities

$$n_i + n_j + (k+1) - (g+4) \ge 0 \quad \text{for every } (i,j) \text{ with } i \ne j \text{ except } (i,j) = (1,2). \tag{5-2}$$

Let E^{gen} (resp. F^{gen}) be the most generic vector bundle on $\mathbb{P}^1$ of rank 4 (resp. 5) and degree $g+4$ (resp. $2g+8$). Define $U_{E^{\text{gen}}}$, $U_{F^{\text{gen}}}$, and $U_{E^{\text{gen}},F^{\text{gen}}}$ as before. These are the open subsets of $\widetilde{\mathcal{H}}_{5,g}$ consisting of covers α for which E_α, F_α, and both E_α and F_α are the most generic.

Proposition 5.1. *The subvariety* $M := \widetilde{\mathcal{H}}_{5,g} \setminus U_{E^{\mathrm{gen}}}$ *has codimension at least* 2 *if* g *is not divisible by* 4, *and has a unique divisorial component if* g *is divisible by* 4.

Proof. This is the degree-5 case of Proposition 2.13. □

For the complement of $U_{F^{\mathrm{gen}}}$, we must analyze the defining equations of C in $\mathbb{P}E$.

Proposition 5.2. *The subvariety* $CE := \mathcal{H}_{5,g} \setminus U_{F^{\mathrm{gen}}}$ *has codimension at least* 2 *if* $g+4$ *is not a multiple of* 5 (*with the exception of* $g = 3$, *in which case the complement parametrizes hyperelliptic curves*), *and contains a unique divisorial component if* $g+4$ *is a multiple of* 5.

Proof. We must characterize the Casnati–Ekedahl loci $C(F)$ which are divisorial. We have

$$C(F) = \bigcup_{E} M(E, F).$$

The loci $M(E, F)$ are irreducible by the same argument as in Proposition 4.1 — in the proof, just take $V = H^0\big(\mathbb{P}^1, \bigwedge^2 F \otimes E \otimes \det E^{\vee}\big)$. Therefore, any component of $C(F)$ must be of the form $M(E, F)$. From the explicit description of degree-5 covers above, it is straightforward to compute that

$$\operatorname{codim} M(E, F) = \dim \operatorname{Ext}^1(E, E) + \dim \operatorname{Ext}^1(F, F) - h^1\big(\textstyle\bigwedge^2 F \otimes E \otimes \det E^{\vee}\big).$$

Suppose $E \neq E^{\mathrm{gen}}$. Then $M(E, F) \subset M(E)$. By Proposition 2.13, $M(E)$ has codimension at least 2 unless $E = \mathcal{O}(k) \oplus \mathcal{O}(k+1)^{\oplus d-3} \oplus \mathcal{O}(k+2)$. In this case, using the explicit description of degree-5 covers, it is easy to construct covers α with $E_\alpha = E$ and $F_\alpha = F^{\mathrm{gen}}$. Thus, $M(E, F) \neq M(E)$, and, since $M(E)$ is irreducible, $M(E, F) \subset M(E)$ has codimension at least 1. Therefore, $M(E, F) \subset \mathcal{H}_{4,g}$ has codimension at least 2. Therefore, for $M(E, F)$ to be divisorial, we must have $E = E^{\mathrm{gen}}$. In this case, we have

$$\operatorname{codim} M(E, F) = \dim \operatorname{Ext}^1(F, F) - h^1\big(\textstyle\bigwedge^2 F \otimes E \otimes \det E^{\vee}\big).$$

Suppose $h^1\big(\bigwedge^2 F \otimes E \otimes \det E^{\vee}\big) = 0$. Note that $\dim \operatorname{Ext}^1(F, F) = 1$ if and only if

$$F = \mathcal{O}(n-1) \oplus \mathcal{O}(n) \oplus \mathcal{O}(n) \oplus \mathcal{O}(n) \oplus \mathcal{O}(n+1).$$

In this case $5n = 2(g+4)$, and hence 5 divides $g+4$.

We are thus reduced to showing that $M(E, F)$ is not a divisor when $E = E^{\mathrm{gen}}$ and

$$h^1\big(\textstyle\bigwedge^2 F \otimes E(-\det E)\big) > 0,$$

with the exception of $g = 3$. Write

$$E = \mathcal{O}(k)^{\oplus r} \oplus \mathcal{O}(k+1)^{\oplus 4-r}, \quad \text{where } 0 \le r \le 3,$$

and

$$F = \mathcal{O}(n_1) \oplus \mathcal{O}(n_2) \oplus \mathcal{O}(n_3) \oplus \mathcal{O}(n_4) \oplus \mathcal{O}(n_5), \quad \text{where } n_1 \le \cdots \le n_5.$$

Consider an antisymmetric matrix

$$M_w = (L_{i,j}), \quad 1 \le i, j \le 5,$$

as in (5-1), representing an element of $H^0\left(\bigwedge^2 F \otimes E \otimes \det E^\vee\right)$. Inequality (5-2) implies that any contribution to $h^1\left(\bigwedge^2 F \otimes E \otimes \det E^\vee\right)$ must come from the $L_{1,2}$ entry. In other words, we have

$$h^1\left(\bigwedge^2 F \otimes E(-\det E)\right) = h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)).$$

Since $E = E^{\text{gen}}$, we have $h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)) > 0$ if and only if

$$n_1 + n_2 + (k+1) - (g+4) < 0.$$

Hence, we get

$$\begin{aligned} h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)) &= 4(-(n_1+n_2+k-(g+4))-1)-(4-r) \\ &= 4g - 4(n_1+n_2+k) + r + 8. \end{aligned}$$

Equation (5-2) tells us that $n_1 + n_3 + (k+1) - (g+4) \ge 0$, which implies $n_2 < n_3$. Therefore,

$$\dim \operatorname{Ext}^1(F, F) \ge (2n_5 + 2n_4 + 2n_3) - 3(n_1+n_2) - 6.$$

Combining the two, we get

$$\begin{aligned} &\dim \operatorname{Ext}^1(F, F) - h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)) \\ &\qquad \ge 2n_5 + 2n_4 + 2n_3 + n_1 + n_2 - 3(g+4) - 2. \end{aligned}$$

Using $n_1 + \cdots + n_5 = 2(g+4)$, the above inequality becomes

$$\dim \operatorname{Ext}^1(F, F) - h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)) \ge (g+4) - (n_1+n_2) - 2.$$

Finally, by using the assumption $n_1 + n_2 + (k+1) - (g+4) < 0$, we conclude that

$$\operatorname{codim} M(E^{\text{gen}}, F) = \dim \operatorname{Ext}^1(F, F) - h^1(E \otimes \det E^\vee \otimes \mathcal{O}(n_1+n_2)) > k - 1.$$

If $k > 1$, then we get $\operatorname{codim} M(E^{\text{gen}}, F) > 1$ as desired. We consider the cases where $k = 1$ on an individual basis. These cases correspond to $0 \le g \le 4$.

Case $g = 4$. Then $E^{\text{gen}} = \mathcal{O}(2)^{\oplus 4}$ and $F^{\text{gen}} = \mathcal{O}(3)^{\oplus 4} \oplus \mathcal{O}(4)$. The relative canonical map embeds C in $\mathbb{P}E^{\text{gen}} \simeq \mathbb{P}^3 \times \mathbb{P}^1$. The projection to $\mathbb{P}^3$ restricts to the canonical map on C. Therefore, if C is nonhyperelliptic, then there is only one quadric in $\mathbb{P}^3$ containing the canonical model of C. This means that the bundle F has exactly one $\mathcal{O}(4)$ summand, and hence $F \cong F^{\text{gen}}$. The locus where C is hyperelliptic is easily seen to be codimension-2 in $\mathcal{H}_{5,4}$. This exhausts all possibilities in this case.

Case $g = 3$. Then $E^{\text{gen}} = \mathcal{O}(1) \oplus \mathcal{O}(2)^{\oplus 3}$ and $F^{\text{gen}} = \mathcal{O}(2) \oplus \mathcal{O}(3)^{\oplus 4}$. Consider the special bundle $F = \mathcal{O}(2) \oplus \mathcal{O}(2) \oplus \mathcal{O}(3)^{\oplus 2} \oplus \mathcal{O}(4)$. Then

$$\dim \operatorname{Ext}^1(F, F) - h^1\big(\textstyle\bigwedge^2 F \otimes E \otimes \det E^\vee\big) = 1.$$

Now consider a general $[\alpha : C \to \mathbb{P}^1] \in M(E, F) \subset \mathcal{H}_{5,3}$. Let $[X : Y : Z : W]$ denote the homogeneous coordinates (locally over $\mathbb{P}^1$) on $\mathbb{P}E$ corresponding to the summands of E. As usual, denote by h the class of $\mathcal{O}_{\mathbb{P}E}(1)$ and by f the class of the fiber of $\mathbb{P}E \to \mathbb{P}^1$. Since $\mathcal{O}(4)$ is a summand of F, there exists a unique effective divisor Q of class $2h - 4f$ on $\mathbb{P}E$ which contains C. The quadric Q may be written as the zero locus of a form

$$c_0 Y^2 + c_1 YZ + \cdots + c_5 W^2,$$

where the c_i are constants. Let $p : \mathbb{P}E \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^1$ be the projection from the section $[1 : 0 : 0 : 0]$, and $g : \mathbb{P}E \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^1 \to \mathbb{P}^2$ the composition with the projection onto the first factor. Then the rational map g is given by the linear system $|h - 2f|$ on $\mathbb{P}E$, which restricts to the canonical series on C. However, the fact that C lies on the relative quadric Q means that the image $g(C)$ is exactly the conic defined by the equation for Q. Thus, C is hyperelliptic.

Given the above geometric understanding of the $\mathcal{O}(4)$ summand of F, it is easy to show that if we begin with a hyperelliptic curve C, and a degree-5 map $\alpha : C \to \mathbb{P}^1$, then F_α must contain a unique $\mathcal{O}(4)$ summand. By the inequalities in (5-2), there are no other choices for F.

Case $g = 1, 2$. In these cases, we leave it to the reader to see that there are no nontrivial Casnati–Ekedahl or Maroni loci. □

As before, we now exhibit $U_{E^{\text{gen}}, F^{\text{gen}}}$ as a quotient. For brevity, set $E = E^{\text{gen}}$ and $F = F^{\text{gen}}$. Set

$$V := H^0\big(\mathbb{P}^1, \textstyle\bigwedge^2 F \otimes E \otimes \det E\big).$$

An element $v \in \mathbb{P}_{\text{sub}} V$ defines an antisymmetric matrix as in (5-1). Let C_v be the zero locus of the 4×4 sub-Pfaffians of this matrix. Let $V^\circ \subset \mathbb{P}_{\text{sub}} V$ be the open locus consisting of v for which C_v is irreducible and at worst nodal. Let $G_F := \operatorname{Aut}(\mathbb{P}F/\mathbb{P}^1)$ and $G_E := \operatorname{Aut}(\mathbb{P}E/\mathbb{P}^1)$. Then $G := G_F \times G_E$ acts on V°.

The assignment $v \mapsto [\pi : C_v \to \mathbb{P}^1]$ defines a map

$$q : V^\circ \to \widetilde{\mathcal{H}}^\dagger_{5,g}.$$

Let $U^\dagger_{E,F}$ be the preimage of $U_{E^{\text{gen}},F^{\text{gen}}}$ under $\widetilde{\mathcal{H}}^\dagger_{5,g} \to \widetilde{\mathcal{H}}_{5,g}$.

Proposition 5.3. *The image of q is $U^\dagger_{E^{\text{gen}},F^{\text{gen}}}$. The fibers of q consist of single G-orbits.*

Proof. The proof is exactly analogous to that of Proposition 3.2. □

Proposition 5.4 (Picard rank conjecture for degree 5). *We have* $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{5,g} = 0$.

Proof. The proof is entirely analogous to the proof of Proposition 4.10. We indicate only the major steps. Set $U = U_{E^{\text{gen}},F^{\text{gen}}}$, and $U^\dagger = U^\dagger_{E^{\text{gen}},F^{\text{gen}}}$. Applying Proposition 1.2 to $V^\circ \to U^\dagger$ and $U^\dagger \to U$, we get

$$\operatorname{rk} \operatorname{Pic}_{\mathbb{Q}} U \leq 1 + \operatorname{rk} \chi(G).$$

Let e be the number of divisorial components of $\widetilde{\mathcal{H}}_{5,g} \setminus U$. We then get

$$\operatorname{rk} \operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{5,g} \leq 1 + \operatorname{rk} \chi(G) + e.$$

Both G and e depend on g modulo 4 and 5. Using Propositions 5.1 and 5.2, we get the following possibilities:

	$\operatorname{rk} \chi(G) = \operatorname{rk} \chi(G_E) + \operatorname{rk} \chi(G_F)$	e
$4 \mid g,\ 5 \mid g+4$	$0 = 0 + 0$	2 (M and CE)
$4 \mid g,\ 5 \nmid g+4$	$1 = 0 + 1$	1 (M)
$4 \nmid g,\ 5 \mid g+4$	$1 = 1 + 0$	1 (CE)
$4 \nmid g,\ 5 \nmid g+4$	$2 = 1 + 1$	0

In all the cases, we have $\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}_{5,g} \leq 3$. Combined with Proposition 2.15, this gives $\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{5,g} = 0$. □

6. From Hurwitz spaces to Severi varieties

The associated scroll construction in Section 2A lets us relate the Picard rank of a Hurwitz space to the Picard rank of a Severi variety. In this section, we work out this relation.

Recall the notation $\mathcal{U}_g(\mathbb{F}_m, d\tau)$, $\mathcal{V}_g(\mathbb{F}_m, d\tau)$, and $\mathcal{V}^{\text{irr}}_g(\mathbb{F}_m, d\tau)$ from page 462. When confusion is unlikely, we abbreviate them by $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{V}^{\text{irr}}$. Following [Diaz and Harris 1988a], we enlarge $\mathcal{U}$ by including the irreducible curves of geometric genus g having a cusp, a tacnode, a triple point, and irreducible nodal curves of geometric genus $(g-1)$ (that is, curves having an "additional" node). Note that the resulting enlargement of $\mathcal{U}$ is a partial compactification of $\mathcal{U}$ in the linear system

$|d\tau|$. Although it does not include extremely singular degenerations of nodal curves, it does include all codimension-1 degenerations (Proposition 6.3). Denote by $\widetilde{\mathcal{U}}$ the normalization of this partial compactification. The local analysis from [Diaz and Harris 1988a, §1] of the Severi variety at points corresponding to cusps, tacnodes, triple points, and an additional node shows that $\widetilde{\mathcal{U}}$ is smooth. Since $\widetilde{\mathcal{U}}$ maps to the linear series $|d\tau|$, it carries over it a family of (singular) curves. The normalization of the total space of this family gives a family $\mathcal{C} \to \widetilde{\mathcal{U}}$ of curves of arithmetic genus g. A generic fiber of $\mathcal{C} \to \widetilde{\mathcal{U}}$ is the normalization the corresponding curve on $\mathbb{F}_m$.

Using the universal family, we can construct tautological divisor classes on $\widetilde{\mathcal{U}}$ as follows. Consider the diagram

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\nu} & \mathbb{F}_m \\ {\scriptstyle\rho}\big\downarrow & & \\ \widetilde{\mathcal{U}} & & \end{array}$$

Define five tautological divisor classes on $\widetilde{\mathcal{U}}$ (The subscript s stands for "Severi"):

(1) $\lambda_s := c_1(\rho_*\omega_\rho)$.

(2) $\kappa_s := \rho_*(c_1(\omega_\rho)^2)$.

(3) $\xi_s := \rho_*(\nu^*(f) \cdot c_1(\omega_\rho))$.

(4) $\theta_s := \rho_*(\nu^*(\sigma) \cdot c_1(\omega_\rho))$.

(5) $\psi_s := \rho_*(\nu^*[\text{point}])$.

Since the irreducible curves in the linear system $|d\tau|$ avoid the directrix σ, we get $\theta_s = \psi_s = 0$. Therefore, it is natural to conjecture:

Conjecture 6.1. The rational Picard group of $\widetilde{\mathcal{U}}$ is tautological, that is,

$$\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{U}} = \mathbb{Q}\langle \lambda_s, \kappa_s, \xi_s \rangle.$$

Denote by CU, TN, TP, and Δ the closures in $\mathcal{V}^{\text{irr}}$ of the locus curves with a cusp, tacnode, triple point, or an additional node, respectively. Abusing notation, denote their preimages in $\widetilde{\mathcal{U}}$ by the same letters.

Remark 6.2. It is not hard to check that the classes in $\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{U}}$ of CU, TN, TP, and Δ can be expressed as $\mathbb{Q}$-linear combinations of λ_s, κ_s, and ξ_s and vice versa. Conjecture 6.1 is therefore equivalent to

$$\operatorname{Pic}_{\mathbb{Q}} \mathcal{U} = 0.$$

Proposition 6.3. *The only divisorial components of* $\mathcal{V}^{\text{irr}} \backslash \mathcal{U}$ *are* CU, TN, TP, *and* Δ.

Proof. It suffices to show that the codimension-1 components of $\mathcal{V} \setminus \mathcal{U}$ are the loci of curves with cusps, tacnodes, triple points or an additional node. This follows by the same proof as for Theorem 1.4 in [Diaz and Harris 1988b]. The critical ingredient of the argument is provided by Lemma 6.4. ☐

Lemma 6.4. *Let $D \in |d\tau|$ be a reduced irreducible curve on the Hirzebruch surface $\mathbb{F}_m$. Denote by A the conductor ideal of the singularities of D. Then A imposes independent conditions on $H^0(\mathbb{F}_m, \mathcal{O}(d\tau))$.*

Proof. Let $K = K_{\mathbb{F}_m}$ be the canonical class. The anticanonical class $-K$ is effective. Furthermore, the fixed component of $-K$ is the directrix σ, and $-K$ separates points away from σ. It is classical that A imposes independent conditions on the adjoint linear system $|K + D|$. Let $Z = V(A)$ be the zero-dimensional scheme defined by the ideal sheaf A. Then the restriction map

$$H^0(\mathcal{O}(K + D)) \to H^0(\mathcal{O}_Z(K + D))$$

is surjective. Therefore, we can conclude the same for

$$H^0(\mathcal{O}(D)) \to H^0(\mathcal{O}_Z(D))$$

by multiplying the previous restriction map by a general section of $\mathcal{O}(-K)$. ☐

We now rephrase the Picard rank conjecture for Hurwitz spaces in a manner similar to Conjecture 6.1. Consider the diagram

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\alpha} & \mathbb{P}^1 \\ {\scriptstyle f}\big\downarrow & & \\ \widetilde{\mathcal{H}}^{\dagger}_{d,g} & & \end{array}$$

Define the following tautological divisor classes on $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$ (The subscript "h" stands for "Hurwitz"):

(1) $\lambda_h := c_1(f_*\omega_f)$.

(2) $\kappa_h := f_*(c_1(\omega_f)^2)$.

(3) $\xi_h := f_*(\alpha^*[\text{point}] \cdot c_1(\omega_f))$.

Conjecture 6.5. The rational Picard group of $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$ is tautological, that is,

$$\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}^{\dagger}_{d,g} = \mathbb{Q}\langle \lambda_h, \kappa_h, \xi_h \rangle.$$

Remark 6.6. It is easy to see that the classes of T, D, and Δ can be expressed as $\mathbb{Q}$-linear combinations of λ_h, κ_h, and ξ_h and vice versa. Also, by Proposition 1.3, the framed/unframed distinction is irrelevant. Therefore, Conjecture 6.5 is equivalent to the Picard rank conjecture stated in the introduction, namely, that

$$\operatorname{Pic}_{\mathbb{Q}} \mathcal{H}_{d,g} = 0.$$

We now state the main theorem of this section:

Theorem 6.7. *If $m \geq \lfloor (g+d-1)/(d-1) \rfloor$, then Conjecture 6.1 for $\widetilde{\mathcal{U}}_g(\mathbb{F}_m, d\tau)$ implies Conjecture 6.5 for $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$. If $m \geq \lceil 2(g+d-1)/d \rceil$, then Conjecture 6.1 for $\widetilde{\mathcal{U}}_g(\mathbb{F}_m, d\tau)$ is equivalent to Conjecture 6.5 for $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$.*

Proof. Let $m \geq \lfloor (g+d-1)/(d-1) \rfloor$. Retain the notation introduced in this section. In particular, abbreviate $\mathcal{U}_g(\mathbb{F}_m, d\tau)$ by $\mathcal{U}$, and so on. Let $\pi : \mathbb{F}_m \to \mathbb{P}^1$ be the projection and $\sigma \subset \mathbb{F}_m$ the directrix. Fix a section $\zeta \in H^0(\mathbb{F}_m, \pi^*\mathcal{O}(m))$ corresponding to a smooth element of the linear series $|\tau|$. We view $\mathbb{F}_m \setminus \sigma$ as the total space of the line bundle $\mathcal{O}(m)$ on $\mathbb{P}^1$ and ζ as the tautological section of $\pi^*\mathcal{O}(m)$ on this total space.

Let $\phi : \mathcal{C} \to \mathbb{P}^1$ be the composition $\phi = \pi \circ \nu$. Let $Z \subset \widetilde{\mathcal{U}}$ be the open subset consisting of the u where $h^0(\mathcal{C}_u, \phi^*\mathcal{O}(m))$ is minimal. Likewise, let $W \subset \widetilde{\mathcal{H}}^{\dagger}_{d,g}$ be the subset consisting of $[\alpha : C \to \mathbb{P}^1]$ where $h^0(C, \alpha^*\mathcal{O}(m))$ is minimal. By Proposition 2.13, the complement of W in $\widetilde{\mathcal{H}}^{\dagger}_{d,g}$ has codimension at least 2. Let V be the total space of the vector bundle $f_*\alpha^*\mathcal{O}(m)|_W$ over W.

We have a birational morphism $q : Z \to V$ defined as follows. A point $u \in \widetilde{\mathcal{U}}$ is mapped to $[\phi_u : \mathcal{C}_u \to \mathbb{P}^1, v]$, where $v \in H^0(\mathcal{C}_u, \phi_u^*\mathcal{O}(m))$ is the restriction of ζ. To define the inverse, we must restrict to an open subset of V. Let $X \subset V$ be the open subset consisting of $([\alpha : C \to \mathbb{P}^1], v)$, where $v \in H^0(C, \alpha^*\mathcal{O}(m))$ is such that the lift of $C \to \mathbb{P}^1$ to $C \to \mathbb{F}_m$ defined by v is birational onto its image. We then get a morphism $p : X \to \mathcal{V}^{\mathrm{irr}}$ which is quasifinite and generically one-to-one. Let $Y \subset X$ be the open subset consisting of points whose associated element in $\mathcal{V}^{\mathrm{irr}}$ has at worst a cusp, a tacnode, a triple point, or an additional node. By Proposition 6.3 and the quasifiniteness of p, the complement of Y in X has codimension at least 2. Since Y is normal, we get a morphism $p : Y \to Z \subset \widetilde{\mathcal{U}}$, inverse to q. We summarize the spaces we have defined and their relationships in the following diagram:

$$\begin{tikzcd}
& V \arrow[d, "\star"] & X \arrow[l, hook'] & Y \arrow[l, hook', "\star"'] \arrow[r, bend left, "p"] & Z \arrow[l, bend left, "q"] \arrow[d, hook] \\
\widetilde{\mathcal{H}}^{\dagger}_{d,g} & W \arrow[l, hook', "\star"'] & & & \widetilde{\mathcal{U}}
\end{tikzcd} \tag{6-1}$$

The inclusions are open inclusions. Y and Z are isomorphic via p and q. The maps marked by $\star$ induce isomorphisms on Picard groups. For the open inclusions, this is because the complements have codimension at least 2. For $V \to W$, this is because it is a vector bundle.

Denote the pullbacks of λ_h, κ_h, and ξ_h to W, V, X, and Y by the same letters. Then, we have

$$\begin{aligned} p^*\lambda_s &= \lambda_h, & p^*\kappa_s &= \kappa_h, & p^*\xi_s &= \xi_h, \\ q^*\lambda_h &= \lambda_s, & q^*\kappa_h &= \kappa_s, & q^*\xi_h &= \xi_s. \end{aligned}$$

We may thus drop the subscripts and use λ, κ, and ξ to denote the corresponding divisors on any of the spaces in (6-1).

Before we proceed, we must comment on the inclusion $X \hookrightarrow V$. The complement consists of $([\alpha : C \to \mathbb{P}^1], v)$, where $v \in H^0(C, \alpha^*\mathcal{O}(m))$ does not give a birational map to $\mathbb{F}_m$. Let us disregard the α that factor nontrivially (such α form a set of codimension at least 2). Then the only such v are the pullbacks of the sections in $H^0(\mathbb{P}^1, \mathcal{O}(m))$. The locus $([\alpha : C \to \mathbb{P}^1], v)$ where $v \in \alpha^* H^0(\mathbb{P}^1, \mathcal{O}(m))$ has codimension at least 2 except in the case $g \equiv -1 \pmod{(d-1)}$ and $m = \lfloor (g+d-1)/(d-1) \rfloor$, that is, when the generic splitting of $\alpha_*\mathcal{O}_C$ is

$$\alpha_*\mathcal{O}_C = \mathcal{O} \oplus \mathcal{O}(-m) \oplus \mathcal{O}(-m-1) \oplus \cdots \oplus \mathcal{O}(-m-1).$$

In this case, the complement of X in V has a divisorial component given by the image of the constant vector bundle $H^0(\mathbb{P}^1, \mathcal{O}(m)) \otimes \mathcal{O}_W$. However, the class of this divisor in $\operatorname{Pic}_{\mathbb{Q}} V \cong \operatorname{Pic}_{\mathbb{Q}} W$ is in the span of λ, κ, and ξ. Therefore, in any case, $\operatorname{Pic}_{\mathbb{Q}} V$ is spanned by λ, κ, and ξ if and only if $\operatorname{Pic}_{\mathbb{Q}} X$ is.

Assume that Conjecture 6.1 holds. From diagram (6-1), we see that $\operatorname{Pic}_{\mathbb{Q}} X$ is spanned by λ, κ, and ξ. By the comment about $X \hookrightarrow V$ above, this implies that $\operatorname{Pic}_{\mathbb{Q}} V$, and in turn $\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{H}}^\dagger_{d,g}$, is spanned by λ, κ, and ξ. Hence Conjecture 6.5 holds.

Assume that $m \geq \lceil 2(g+d-1)/(d-1) \rceil$ and Conjecture 6.5 holds. Then, by Proposition 2.6 the inclusion $Z \hookrightarrow \widetilde{\mathcal{U}}$ is in fact an isomorphism. Again, diagram (6-1) shows that $\operatorname{Pic}_{\mathbb{Q}} \widetilde{\mathcal{U}}$ is spanned by λ, κ, and ξ. Hence Conjecture 6.1 holds. □

Acknowledgments

We are grateful to our advisor Joe Harris for suggesting this problem and for providing constant support and encouragement. We also thank Gabriel Bujokas, Dawei Chen, Maksym Fedorchuk, and Ravi Vakil for helpful comments and conversations.

References

[Bolognesi and Lönne 2014] M. Bolognesi and M. Lönne, "Mapping Class Groups of Trigonal Loci", preprint, 2014. arXiv 1403.7399

[Bolognesi and Vistoli 2012] M. Bolognesi and A. Vistoli, "Stacks of trigonal curves", *Trans. Amer. Math. Soc.* **364**:7 (2012), 3365–3393. MR 2901217 Zbl 06188646

[Casnati 1996] G. Casnati, "Covers of algebraic varieties, II: Covers of degree 5 and construction of surfaces", *J. Algebraic Geom.* **5**:3 (1996), 461–477. MR 97c:14015 Zbl 0921.14006

[Casnati 1998] G. Casnati, "Covers of algebraic varieties, III: The discriminant of a cover of degree 4 and the trigonal construction", *Trans. Amer. Math. Soc.* **350**:4 (1998), 1359–1378. MR 98i:14021

[Casnati and Ekedahl 1996] G. Casnati and T. Ekedahl, "Covers of algebraic varieties, I: A general structure theorem, covers of degree 3, 4 and Enriques surfaces", *J. Algebraic Geom.* **5**:3 (1996), 439–460. MR 97c:14014 Zbl 0866.14009

[Coppens 1999] M. Coppens, "Existence of pencils with prescribed scrollar invariants of some general type", *Osaka J. Math.* **36**:4 (1999), 1049–1057. MR 2001g:14006 Zbl 1047.14500

[Cornalba 2007] M. Cornalba, "The Picard group of the moduli stack of stable hyperelliptic curves", *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **18**:1 (2007), 109–115. MR 2008b:14017 Zbl 1222.14054

[Cornalba and Harris 1988] M. Cornalba and J. Harris, "Divisor classes associated to families of stable varieties, with applications to the moduli space of curves", *Ann. Sci. École Norm. Sup.* (4) **21**:3 (1988), 455–475. MR 89j:14019 Zbl 0674.14006

[Diaz and Edidin 1996] S. Diaz and D. Edidin, "Towards the homology of Hurwitz spaces", *J. Differential Geom.* **43**:1 (1996), 66–98. MR 98e:14028 Zbl 0848.14005

[Diaz and Harris 1988a] S. Diaz and J. Harris, "Geometry of the Severi variety", *Trans. Amer. Math. Soc.* **309**:1 (1988), 1–34. MR 89i:14018 Zbl 0677.14003

[Diaz and Harris 1988b] S. Diaz and J. Harris, "Ideals associated to deformations of singular plane curves", *Trans. Amer. Math. Soc.* **309**:2 (1988), 433–468. MR 89m:14003 Zbl 0707.14022

[Donagi 1981] R. Donagi, "The tetragonal construction", *Bull. Amer. Math. Soc.* (*N.S.*) **4**:2 (1981), 181–185. MR 82a:14009 Zbl 0491.14016

[Harer 1983] J. Harer, "The second homology group of the mapping class group of an orientable surface", *Invent. Math.* **72**:2 (1983), 221–239. MR 84g:57006 Zbl 0533.57003

[Harris 1986] J. Harris, "On the Severi problem", *Invent. Math.* **84**:3 (1986), 445–461. MR 87f:14012 Zbl 0596.14017

[Knop et al. 1989] F. Knop, H. Kraft, and T. Vust, "The Picard group of a G-variety", pp. 77–87 in *Algebraische Transformationsgruppen und Invariantentheorie*, edited by H. Kraft et al., DMV Sem. **13**, Birkhäuser, Basel, 1989. MR 1044586 Zbl 0705.14005

[Mochizuki 1995] S. Mochizuki, "The geometry of the compactification of the Hurwitz scheme", *Publ. Res. Inst. Math. Sci.* **31**:3 (1995), 355–441. MR 96j:14017 Zbl 0866.14013

[Mumford et al. 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)] **34**, Springer, Berlin, 1994. MR 95m:14012 Zbl 0797.14004

[Ohbuchi 1997] A. Ohbuchi, "On some numerical relations of d-gonal linear systems", *J. Math. Tokushima Univ.* **31** (1997), 7–10. MR 99b:14005 Zbl 0938.14011

[Recillas 1973] S. Recillas, "Maps between Hurwitz spaces", *Bol. Soc. Mat. Mexicana* (2) **18** (1973), 59–63. MR 50 #13050 Zbl 0313.14003

[Stankova-Frenkel 2000] Z. E. Stankova-Frenkel, "Moduli of trigonal curves", *J. Algebraic Geom.* **9**:4 (2000), 607–662. MR 2001h:14031 Zbl 1001.14007

[Teissier 1980] B. Teissier, "Résolution simultanée, I: Familles de courbes", pp. viii+339 in *Séminaire sur les singularités des surfaces* (Palaiseau, 1976–1977), edited by M. Demazure et al., Lecture Notes in Mathematics **777**, Springer, Berlin, 1980. MR 82d:14021 Zbl 0464.14005

[Tyomkin 2007] I. Tyomkin, "On Severi varieties on Hirzebruch surfaces", *Int. Math. Res. Not.* **2007**:23 (2007), 31. MR 2008m:14048 Zbl 1135.14026

Communicated by Ravi Vakil
Received 2014-07-27 Revised 2014-11-04 Accepted 2014-12-11

anandrd@math.columbia.edu Department of Mathematics, Columbia University, 2990 Broadway, New York, NY 10027, United States

anand.patel@bc.edu Mathematics Department, Boston College, Carney Hall, Chestnut Hill, MA 02467, United States

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.493

Finite-dimensional quotients of Hecke algebras

Ivan Losev

Let W be a complex reflection group. We prove that there is a maximal finite-dimensional quotient of the Hecke algebra $\mathcal{H}_q(W)$ of W, and that the dimension of this quotient coincides with $|W|$. This is a weak version of a 1998 Broué–Malle–Rouquier conjecture. The proof is based on the categories $\mathcal{O}$ for rational Cherednik algebras.

1. Introduction

Let W be a complex reflection group. Recall that such groups are fully classified; see [Shephard and Todd 1954]. In this context, one can also define the braid group B_W. Namely, let $\mathfrak{h}$ denote the reflection representation of W. Inside $\mathfrak{h}$, one considers the open subset of regular points $\mathfrak{h}^{\mathrm{reg}} := \{x \in \mathfrak{h} \mid W_x = \{1\}\}$, so that W is the Galois group of the cover $\mathfrak{h}^{\mathrm{reg}} \twoheadrightarrow \mathfrak{h}^{\mathrm{reg}}/W$. By definition, the braid group B_W is the fundamental group $\pi_1(\mathfrak{h}^{\mathrm{reg}}/W)$.

If W is a Coxeter group, one considers a flat deformation of $\mathbb{C}W$, called a *Hecke algebra*. These algebras are of importance in representation theory (e.g., that of finite groups of Lie type) and beyond (e.g., in knot theory). One can define Hecke algebras for complex reflection groups as well. In the general case, this was done in [Broué et al. 1998, Section 4C]. To recall the definition, we need some more notation. Namely, let $\mathfrak{H}$ denote the set of reflection hyperplanes for W. For $\Gamma \in \mathfrak{H}$, let W_Γ denote the pointwise stabilizer of Γ; this is a cyclic group. Set $\ell_\Gamma := |W_\Gamma|$. The group B_W is generated by elements T_Γ, $\Gamma \in \mathfrak{H}$, where, roughly speaking, T_Γ is the rotation around Γ by $2\pi/\ell_\Gamma$; see [loc. cit., Section 2]. Now, pick independent variables $u_{\Gamma,i}$, $i = 0, 1, \ldots, \ell_\Gamma - 1$ with $u_{\Gamma,i} = u_{\Gamma',i}$ for W-conjugate Γ, Γ'. Set $\boldsymbol{u} := (u_{\Gamma,i})$. By definition ([loc. cit., Definition 4.21]), the Hecke algebra $\mathcal{H}_{\boldsymbol{u}}(W)$ is the quotient of $\mathbb{Z}[\boldsymbol{u}^{\pm 1}]B_W$ by the relations

$$\prod_{i=0}^{\ell_\Gamma - 1} (T_\Gamma - u_{\Gamma,i}) = 0.$$

MSC2010: primary 20C08; secondary 20F55, 16G99.
Keywords: Hecke algebras, rational Cherednik algebras, categories $\mathcal{O}$, KZ functor.

Broué, Malle and Rouquier [Broué et al. 1998, Section 4C] conjectured that $\mathcal{H}_{\boldsymbol{u}}(W)$ is a free $\mathbb{Z}[\boldsymbol{u}^{\pm 1}]$-module generated by $|W|$ elements. Currently, the proof is missing in the case of several exceptional complex reflection groups. In this paper, we are going to prove a weaker version of this conjecture.

First of all, we are dealing with specializations to $\mathbb{C}$. For a collection of nonzero complex numbers $(q_{\Gamma,i})$, where $\Gamma \in \mathfrak{H}/W$ and $i \in \{0, 1, \dots, \ell_\Gamma - 1\}$, consider the $\mathbb{C}$-algebra $\mathcal{H}_q(W)$, the specialization of $\mathcal{H}_{\boldsymbol{u}}(W)$ with $u_{\Gamma,i} \mapsto q_{\Gamma,i}$. Note that replacing the collection $q_{\Gamma,i}$ with $(\alpha_\Gamma q_{\Gamma,i})$ for $\alpha_\Gamma \in \mathbb{C}^\times$, we get isomorphic algebras; see, e.g., [Rouquier 2008, Section 3.3.3]. So the number of parameters actually equals $|S/W|$, where S denotes the set of complex reflections in W. Note that if $q_{\Gamma,j} = \exp(2\pi\sqrt{-1}\,j/\ell_\Gamma)$ for all Γ and j, we just have $\mathcal{H}_q(W) = \mathbb{C}W$. In general, however, it is even unclear whether the algebra $\mathcal{H}_q(W)$ is finite-dimensional or not. In a way, the infinite dimension is the only obstruction to the equality $\dim \mathcal{H}_q(W) = |W|$. More precisely, we have the following theorem, which is the main result of this paper:

Theorem 1.1. *There is a minimal two-sided ideal $I \subset \mathcal{H}_q(W)$ such that $\mathcal{H}_q(W)/I$ is finite-dimensional. Moreover, we have* $\dim \mathcal{H}_q(W)/I = |W|$.

Other results towards the Broué–Malle–Rouquier conjecture were known before; see [Marin 2014] for a review. One advantage of our approach is that it is fully conceptual and does not involve any case-by-case arguments.

The key idea of the proof is to use categories $\mathcal{O}$ for rational Cherednik algebras $H_c(W)$, introduced in [Ginzburg et al. 2003]. By definition, the algebra $H_c(W)$ is the subalgebra in the skew-group algebra $D(\mathfrak{h}^{\mathrm{reg}})\#W$ generated by $\mathbb{C}[\mathfrak{h}]$, $\mathbb{C}W$ and the so-called Dunkl operators D_a, $a \in \mathfrak{h}$. These are differential operators with first-order poles along the reflection hyperplanes Γ. We have a triangular decomposition $H_c(W) = \mathbb{C}[\mathfrak{h}] \otimes \mathbb{C}W \otimes S(\mathfrak{h})$ that allows to define the category $\mathcal{O}$. This is the category of all $H_c(W)$-modules that are finitely generated over $\mathbb{C}[\mathfrak{h}]$ and have a locally nilpotent action of $\mathfrak{h} \subset S(\mathfrak{h})$.

Let us pick $M \in \mathcal{O}$. Its restriction to $\mathfrak{h}^{\mathrm{reg}}$ is a W-equivariant local system on $\mathfrak{h}^{\mathrm{reg}}$. So the fiber M_x carries a monodromy representation of B_W. It was shown in [Ginzburg et al. 2003] that the $\mathbb{C}B_W$-action on M_x factors though a certain quotient of $\mathcal{H}_q(W)$ that has dimension $|W|$. We will show that every finite-dimensional $\mathcal{H}_q(W)$-module can be represented in the form M_x for some $M \in \mathcal{O}$.

The assumption $\dim \mathcal{H}_q(W) = |W|$ is actually important for the representation theory of $H_c(W)$. Theorem 1.1 should make it possible to remove this assumption, but we are not going to elaborate on that.

The paper is organized as follows: In Section 2 we gather various generalities on the rational Cherednik algebras and their categories $\mathcal{O}$. Then in Section 3 we prove the main theorem.

2. Generalities

2A. *Rational Cherednik algebras.* Rational Cherednik algebras were introduced by Etingof and Ginzburg [2002]. In this subsection we recall their definition.

Let W be a complex reflection group and $\mathfrak{h}$ be its reflection representation. We denote the subset of W consisting of complex reflections by S. For $s \in S$, pick an eigenvector $\alpha_s \in \mathfrak{h}^*$ for s with eigenvalue $\lambda_s \neq 1$. We fix a W-invariant function $c : S \to \mathbb{C}$. Using this function, for $a \in \mathfrak{h}$, we can define the Dunkl operator $D_a \in D(\mathfrak{h}^{\mathrm{reg}})\#W$ by

$$D_a = \partial_a + \sum_{s \in S} \frac{2c(s)}{1-\lambda_s} \frac{\langle \alpha_s, a\rangle}{\alpha_s}(s-1).$$

The rational Cherednik algebra $H_c(W)$ is the subalgebra in $D(\mathfrak{h}^{\mathrm{reg}})\#W$ generated by $\mathbb{C}[\mathfrak{h}], \mathbb{C}W$ and the Dunkl operators $D_a, a \in \mathfrak{h}$. Alternatively, one can present $H_c(W)$ by generators and relations: $H_c(W)$ is the quotient of $T(\mathfrak{h} \oplus \mathfrak{h}^*)\#W$ by the relations

$$[x,x'] = [y,y'] = 0, [y,x] = \langle y,x\rangle - \sum_{s \in S} c(s)\langle \alpha_s, y\rangle\langle \alpha_s^\vee, x\rangle s, \quad x,x' \in \mathfrak{h}^*,\ y,y' \in \mathfrak{h}.$$

Here, we write $\alpha_s^\vee$ for the eigenvector of s in $\mathfrak{h}$ with eigenvalue λ_s^{-1} and $\langle \alpha_s^\vee, \alpha_s\rangle = 2$. To get from the second definition to the first one, we use the homomorphism $H_c(W) \to D(\mathfrak{h}^{\mathrm{reg}})\#W$ given by $x \mapsto x, w \mapsto w, y \mapsto D_y$. Set $\delta := \prod_{s \in S} \alpha_s^{\ell_s}$, where ℓ_s stands for the order of s (note that this is slightly different from the usual definition). This is a W-invariant element, and the operator $\operatorname{ad}\delta : H_c(W) \to H_c(W)$ is locally nilpotent, so the localization $H_c(W)[\delta^{-1}]$ is defined. The homomorphism $H_c(W) \to D(\mathfrak{h}^{\mathrm{reg}})\#W$ extends to an isomorphism $H_c(W)[\delta^{-1}] \xrightarrow{\sim} D(\mathfrak{h}^{\mathrm{reg}})\#W$.

The algebra $H_c(W)$ admits a triangular decomposition: a natural map $S(\mathfrak{h}^*) \otimes \mathbb{C}W \otimes S(\mathfrak{h}) \to H_c(W)$ is an isomorphism. Also $H_c(W)$ is graded with $\deg x = 1$, $\deg w = 0$, $\deg y = -1$, $x \in \mathfrak{h}^*$, $w \in W$, and $y \in \mathfrak{h}$. We call this grading the Euler grading. It is inner: it is given by the eigenvalues of $\operatorname{ad} h$, where

$$h = \sum_{i=1}^{n} x_i y_i + \frac{n}{2} - \sum_{s \in S} \frac{2c(s)}{1-\lambda_s} s. \tag{2-1}$$

Now let us discuss base change for $H_c(W)$. Let U be an affine algebraic variety equipped with an étale map $U \to \mathfrak{h}/W$. Then $\mathbb{C}[U] \otimes_{\mathbb{C}[\mathfrak{h}]^W} H_c(W)$ has a natural algebra structure; it is a subalgebra in $D(U \times_{\mathfrak{h}/W} \mathfrak{h}^{\mathrm{reg}})\#W$ generated by $\mathbb{C}[U] \otimes_{\mathbb{C}[\mathfrak{h}/W]} \mathbb{C}[\mathfrak{h}], \mathbb{C}W$ and the Dunkl operators. Similarly, if U is a Stein complex analytic manifold (again equipped with an étale map $U \to \mathfrak{h}/W$), then $\mathbb{C}_{\mathrm{an}}[U] \otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W)$ is an algebra. Here and below $\mathbb{C}_{\mathrm{an}}[U]$ denotes the algebra of analytic functions on U.

2B. ***Category $\mathcal{O}$ and* KZ *functor.*** The category $\mathcal{O}$ for $H_c(W)$ was defined in [Ginzburg et al. 2003]. By definition, it consists of all $H_c(W)$-modules M that are finitely generated over $S(\mathfrak{h}^*) = \mathbb{C}[\mathfrak{h}]$ and where the action of $\mathfrak{h}$ is locally nilpotent. Equivalently, $\mathcal{O}$ consists of all $H_c(W)$-modules M that are finitely generated over $\mathbb{C}[\mathfrak{h}]$ and that can be graded. This category $\mathcal{O}$ will be denoted by $\mathcal{O}_c(W)$.

Let us proceed to the KZ functor introduced in [loc. cit., Section 5]. Pick $M \in \mathcal{O}_c(W)$. Then $M[\delta^{-1}]$ is a W-equivariant local system on $D(\mathfrak{h}^{\mathrm{reg}})$ with regular singularities. The category of such local systems is equivalent to the category B_W-mod$_{\mathrm{fin}}$ of finite-dimensional B_W-modules; to a local system M' one assigns its fiber (or, more precisely, the fiber of its descent to $\mathfrak{h}^{\mathrm{reg}}/W$) equipped with the monodromy representation. It turns out that the monodromy representation associated to $M[\delta^{-1}]$ factors through $\mathscr{H}_q(W)$ [loc. cit., Section 5.3], where the parameter q is computed as follows: For a reflection hyperplane Γ, set

$$h_{\Gamma,i} = \frac{1}{\ell_\Gamma} \sum_{s \in W_\Gamma \setminus \{1\}} \frac{2c(s)}{\lambda_s - 1} \lambda_s^{-i}, \tag{2-2}$$

$$q_{\Gamma,i} = \exp(2\pi\sqrt{-1}(h_{\Gamma,j} + j/\ell_H)). \tag{2-3}$$

So we get an exact functor $\mathrm{KZ} : \mathcal{O}_c(W) \to \mathscr{H}_q(W)$-mod$_{\mathrm{fin}}$. This functor is given by $\mathrm{Hom}_{\mathcal{O}_c(W)}(P_{\mathrm{KZ}}, \bullet)$, where P_{KZ} is a projective object such that $\dim \mathrm{End}_{\mathcal{O}_c(W)}(P_{\mathrm{KZ}})$ is equal to $|W|$, equipped with a homomorphism $\mathscr{H}_q(W) \to \mathrm{End}_{\mathcal{O}_c(W)}(P_{\mathrm{KZ}})^{\mathrm{opp}}$. The proof of [loc. cit., Theorem 5.15] shows that this homomorphism is surjective.

Theorem 1.1 will follow if we show that the functor KZ is essentially surjective.

2C. ***Isomorphisms of étale lifts.*** Here we are going to recall some results of [Bezrukavnikov and Etingof 2009] regarding isomorphisms of completions.

Let $W' \subset W$ be a parabolic subgroup, i.e., the stabilizer of a point in $\mathfrak{h}$. Set $\mathfrak{h}^{\mathrm{reg}-W'} := \{b \in \mathfrak{h} \mid W_b \subset W'\}$. The complement of $\mathfrak{h}^{\mathrm{reg}-W'}$ in $\mathfrak{h}$ is the union of the hyperplanes $\ker \alpha_s$ for $s \notin W'$. So $\mathfrak{h}^{\mathrm{reg}-W'}$ is a principal open subset of $\mathfrak{h}$. Note that the natural morphism $\mathfrak{h}^{\mathrm{reg}-W'}/W' \to \mathfrak{h}/W$ is étale (and $\mathfrak{h}^{\mathrm{reg}-W'}/W'$ is precisely the unramified locus of $\mathfrak{h}/W' \to \mathfrak{h}/W$).

Consider the space $H_c(W)_{\mathrm{reg}-W'} := \mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}]^{W'} \otimes_{\mathbb{C}[\mathfrak{h}]^W} H_c(W)$. As was mentioned in the end of Section 2A, $H_c(W)_{\mathrm{reg}-W'}$ is actually an algebra. Bezrukavnikov and Etingof [2009, Section 3.3] essentially found an alternative description of this algebra. Namely, consider the Cherednik algebra $H_c(W', \mathfrak{h})$ defined for the pair $W', \mathfrak{h}$; it decomposes into the tensor product $H_c(W', \mathfrak{h}) = D(\mathfrak{h}^{W'}) \otimes H_c(W')$ (here we abuse notation and write c for the restriction of c to $S \cap W'$; $D(\mathfrak{h}^{W'})$ stands for the algebra of differential operators on $\mathfrak{h}^{W'}$). Then consider its localization $H_c(W', \mathfrak{h})_{\mathrm{reg}-W'} := \mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}]^{W'} \otimes_{\mathbb{C}[\mathfrak{h}]^{W'}} H_c(W', \mathfrak{h})$. Then, following [loc. cit., Section 3.2], we can form the centralizer algebra $Z(W, W', H_c(W', \mathfrak{h})_{\mathrm{reg}-W'})$. Recall that, by definition, for an algebra A equipped with a homomorphism $\mathbb{C}W' \to A$, one defines the centralizer algebra $Z(W, W', A)$ as the endomorphism algebra of the right A-module $\mathrm{Map}_{W'}(W, A) = \{f : W \to A \mid f(w'w) = w' f(w)\}$. Choosing representatives of the left W'-cosets in W, we get an identification $Z(W, W', A) \cong \mathrm{Mat}_{|W/W'|}(A)$. The algebra A can be recovered from $Z(W, W', A)$ as follows: Consider the element $e(W') \in Z(W, W', A)$ given by $e(W')f(u) = f(u)$ if $u \in W'$ and 0 otherwise. Then $e(W')Z(W, W', A)e(W')$ is naturally identified with A.

The following is essentially [loc. cit., Theorem 3.2] (there, the authors considered completions instead of étale lifts, but the proof works in our situation as well).

Lemma 2.1. *There is a unique isomorphism*

$$\theta : H_c(W)_{\mathrm{reg}-W'} \xrightarrow{\sim} Z(W, W', H_c(W', \mathfrak{h})_{\mathrm{reg}-W'}) \tag{2-4}$$

such that the following hold for any $f \in \mathrm{Map}_{W'}(W, A)$ and any $u \in W$:

$$\begin{aligned}
&[\theta(g) f](u) = g f(u), \quad g \in \mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}]^{W'},\\
&[\theta(\alpha) f](u) = (u\alpha) f(u), \quad \alpha \in \mathfrak{h}^*,\\
&[\theta(w) f](u) = f(uw), \quad w \in W,\\
&[\theta(a) f](u) = (ua) f(u) + \sum_{s \in S \setminus W'} \frac{2c(s)}{1-\lambda_s} \frac{\langle \alpha_s, ua\rangle}{\alpha_s}(f(su) - f(u)), \quad a \in \mathfrak{h}.
\end{aligned}$$

Note that the algebras in (2-4) come equipped with $\mathbb{C}^\times$-actions by algebra automorphisms. For example, the action of $H_c(W)_{\mathrm{reg}-W'}$ comes from the action on $H_c(W)$ induced from the Euler grading and the action on $\mathfrak{h}^{\mathrm{reg}-W'}/W'$ induced from the scaling action ($t.x = t^{-1}x$) on $\mathfrak{h}$. The isomorphism θ is $\mathbb{C}^\times$-equivariant.

We can further restrict θ to some analytic submanifolds or formal subschemes of $\mathfrak{h}^{\mathrm{reg}-W'}/W'$. Choose a little disk $Y \subset \mathfrak{h}^{\mathrm{reg}-W'} \cap \mathfrak{h}^{W'}$ and also a little disk D around 0 in $\mathfrak{h}_{W'}/W'$, where $\mathfrak{h}_{W'}$ denotes the unique W'-stable complement to $\mathfrak{h}^{W'}$ in $\mathfrak{h}$. We set $\widehat{Y} := Y \times D$; this is an open submanifold in $\mathfrak{h}^{\mathrm{reg}-W'}/W'$, in $(\mathfrak{h}^{\mathrm{reg}-W'} \cap \mathfrak{h}^{W'}) \times \mathfrak{h}_{W'}/W'$ or in $\mathfrak{h}/W$ (under the natural morphism $\mathfrak{h}^{\mathrm{reg}-W'}/W' \to \mathfrak{h}/W$).

So we get an isomorphism

$$\theta_Y : \mathbb{C}_{\mathrm{an}}[\widehat{Y}] \otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W) \xrightarrow{\sim} Z(W, W', \mathbb{C}_{\mathrm{an}}[\widehat{Y}] \otimes_{\mathbb{C}[\mathfrak{h}/W']} H_c(W', \mathfrak{h})). \tag{2-5}$$

Note that this isomorphism is compatible with the Euler derivations.

We can restrict even further. Pick a point $b \in Y$, and consider the completion $\mathbb{C}[\mathfrak{h}/W]^{\wedge b}$ of $\mathbb{C}[\mathfrak{h}/W]$ with respect to the maximal ideal defined by b. Then θ_Y induces

$$\begin{aligned}
\theta_b : \mathbb{C}[\mathfrak{h}/W]^{\wedge b} &\otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W)\\
&\xrightarrow{\sim} Z(W, W', \mathbb{C}[\mathfrak{h}/W']^{\wedge b} \otimes_{\mathbb{C}[\mathfrak{h}/W']} H_c(W', \mathfrak{h})).
\end{aligned} \tag{2-6}$$

This isomorphism was originally constructed in [Bezrukavnikov and Etingof 2009].

3. Proof of the main theorem

3A. ***Scheme of the proof.*** Let V be a finite-dimensional $\mathcal{H}_q(W)$-module and let N denote the corresponding W-equivariant D-module on $\mathfrak{h}^{\mathrm{reg}}$. Our goal is to show that there is an $M \in \mathcal{O}_c(W)$ such that $M[\delta^{-1}] \cong N$. This consists of two steps:

(I) Set $\mathfrak{h}^{\mathrm{sr}} := \{b \in W \mid \dim \mathfrak{h}^{W_b} \geqslant \dim \mathfrak{h} - 1\}$. This is an open subset that coincides with $\bigcup_{\Gamma \in \mathfrak{H}} \mathfrak{h}^{\mathrm{reg}-W_\Gamma}$; the codimension of its complement is bigger than 1. We will see that there is an $\mathcal{O}_{\mathfrak{h}^{\mathrm{sr}}/W}$-coherent $\mathcal{O}_{\mathfrak{h}^{\mathrm{sr}}/W} \otimes_{\mathbb{C}[\mathfrak{h}]^W} H_c(W)$-module $\widetilde{N}$ whose restriction to $\mathfrak{h}^{\mathrm{reg}}/W$ is isomorphic to N and that carries a locally finite derivation compatible with the Euler derivation of $H_c(W)$.

(II) We will see that $\widetilde{N}$ is a vector bundle. From here we will deduce that the global sections of $\widetilde{N}$ are finitely generated and hence lie in $\mathcal{O}_c(W)$. Then we take $M := \Gamma(\widetilde{N})$.

Let us elaborate on how we are going to achieve (I). First, in Section 3B we will check that the Euler vector field acts on N locally finitely. This will eventually prove that $\widetilde{N}$ comes equipped with a locally finite derivation that is compatible with the Euler one on $H_c(W)$.

Now let us explain how we produce $\widetilde{N}$; this is done in Section 3C. Take $W' = W_\Gamma$ and let $\widehat{Y}$ have the same meaning as in Section 2C. Set $\widehat{Y}^\times := \widehat{Y} \setminus Y$. Consider $N_\Gamma := e(W')(\mathbb{C}_{\mathrm{an}}[\widehat{Y}^\times] \otimes_{\mathbb{C}[\mathfrak{h}^{\mathrm{reg}}/W]} N)$. This is a vector bundle on $\widehat{Y}^\times$ with a meromorphic connection that has pole of order 1 on Y (since N_Γ is obtained by restricting an algebraic vector bundle $e(W')\eta_\Gamma^* N$, where η_Γ is a natural morphism $\mathfrak{h}/W_\Gamma \to \mathfrak{h}/W$, it makes sense to speak about sections of N_Γ with poles on Y; here and below η_Γ denotes the projection $\mathfrak{h}/W_\Gamma \to \mathfrak{h}/W$). Our first step will be to see that N_Γ is obtained by restricting a $\mathbb{C}_{\mathrm{an}}[\widehat{Y}] \otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W)$-module M_Γ. Then we will see that $[e(W')\eta_\Gamma^* N] \cap M_\Gamma$ (the intersection of subspaces in N_Γ) is finitely generated over $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}/W']$. We will get $\widetilde{N}$, roughly speaking, by taking the intersection of N and $\eta_\Gamma^* N \cap M_\Gamma$ over all possible Γ.[1]

3B. ***Locally finite derivation.*** Our goal here is to show that the Euler vector field acts on N locally finitely. Recall that N is a local system on $\mathfrak{h}^{\mathrm{reg}}/W$ with regular singularities. Our claim is a consequence of the following general result.

Lemma 3.1. *Let X be the complement to a $\mathbb{C}^\times$-stable divisor in $\mathbb{C}^d$, and let N be a local system with regular singularities on X. Then the Euler vector field* eu *acts on N locally finitely, meaning that every $n \in N$ is included into a finite-dimensional* eu*-stable subspace.*

[1] After this paper was written, I learned from Etingof that most of the proof is already contained in some form in [Wilcox 2011]. Lemmas 5.7 and 5.8 there are similar to what is done in Section 3C, while the main result of Section 3D has a somewhat easier proof in [Wilcox 2011, Lemma 3.6].

This claim should be standard, but we provide the proof for the sake of completeness. For a different proof, see [Wilcox 2011, Lemma 3.2].

Proof. If $N' \subset N$ is a $D(X)$-submodule and eu acts locally finitely on N/N' and N', then the same is true for N. So it is enough to assume that N (and hence V) is irreducible. Consider the element $\eta \in \pi_1(X)$ given by the loop $\exp(2\pi\sqrt{-1}t)x_0$, $t \in [0,1]$, where x_0 denotes the base point. The element η is central and hence has to act on V by a scalar. Under the Riemann–Hilbert correspondence, this translates to the claim that N is twisted equivariant with respect to the $\mathbb{C}^\times$-action. This implies our claim. □

3C. *Extension to codimension* 1. We start by constructing M_Γ.

Set $\widehat{Y}^\times := Y \times D^\times = \widehat{Y} \setminus Y$. Consider the category $\mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$ of meromorphic local system on $\widehat{Y}^\times$ with regular singularities on Y (so an object in $\mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$ comes equipped with a lattice over the ring of meromorphic differential operators on $\widehat{Y}^\times$, and a morphism in the category is supposed to preserve such lattices). The category $\mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$ is equivalent to the category $\mathbb{C}[T^{\pm 1}]$-mod of finite-dimensional $\mathbb{C}[T^{\pm 1}]$-modules via taking the monodromy representation, because $\pi_1(\widehat{Y}^\times) = \pi_1(D^\times) = \mathbb{Z}$ (here we use the regular singularities condition). Under the equivalence $\mathbb{C}[T^{\pm 1}]\text{-mod} \cong \mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$, the KZ functor becomes

$$\mathbb{O}_c(W') \to \mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y), \quad M \mapsto \mathbb{C}_{\mathrm{an}}[\widehat{Y}^\times] \otimes_{\mathbb{C}[\mathfrak{h}_{W'}/W']} M$$

(with meromorphic lattice $\mathbb{C}_{\mathrm{an}}[\widehat{Y}][\nu^{-1}] \otimes_{\mathbb{C}[\mathfrak{h}_{W'}/W']} M$, where ν denotes a coordinate on $\mathfrak{h}_{W'}/W'$). The right adjoint KZ^* sends $N' \in \mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$ to the subspace of N' of all meromorphic elements annihilated by the vector fields on Y and lying in the generalized eigenspace for $\mathfrak{h}_{W'} \subset H_c(W')$ with eigenvalue 0.

Now we can produce a $\mathbb{C}[\widehat{Y}] \otimes_{\mathbb{C}[\mathfrak{h}/W']} H_c(W', \mathfrak{h})$-module $M_\Gamma \in \mathbb{O}_c(W', \widehat{Y})$. Set $N_\Gamma = e(W')(\mathbb{C}_{\mathrm{an}}[\widehat{Y}^\times] \otimes_{\mathbb{C}[\mathfrak{h}^{\mathrm{reg}}/W]} N)$. This is an object in $\mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y)$. Note that, under the equivalence $\mathrm{Loc}_{\mathrm{rs}}(\widehat{Y}, Y) \cong \mathbb{C}[T^{\pm 1}]$-mod, we have $N_\Gamma \in \mathcal{H}_q(W')$-mod. Now set

$$M_\Gamma := \mathbb{C}_{\mathrm{an}}[\widehat{Y}] \otimes_{\mathbb{C}[\mathfrak{h}_{W'}/W']} \mathrm{KZ}^*(N_\Gamma). \tag{3-1}$$

Note that the description of $\mathrm{KZ}^*(N_\Gamma) \subset N_\Gamma$ implies that it is stable under the Euler vector field on N_Γ. So $M_\Gamma \subset N_\Gamma$ is also stable under the Euler vector field.

Let $\widetilde{N}_\Gamma := M_\Gamma \cap e(W')\eta_\Gamma^* N$ (the intersection is taken inside N_Γ). This is a submodule in the $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}]^{W'} \otimes_{\mathbb{C}[\mathfrak{h}/W']} H_c(W')$-module $e(W')\eta_\Gamma^* N$ which is stable under the Euler vector field.

Lemma 3.2. *The module $\widetilde{N}_\Gamma$ is finitely generated over $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}]^{W'}$ and satisfies $\widetilde{N}_\Gamma[\nu^{-1}] = e(W')\eta_\Gamma^* N$, where ν is a coordinate on $\mathfrak{h}_{W'}/W' \cong \mathbb{C}$.*

Proof. Note that the epimorphism $\mathcal{H}_q(W') \to \operatorname{End}_{\mathcal{O}_c(W')}(P_{\mathrm{KZ}})^{\mathrm{opp}}$ is an isomorphism. Let $N_\Gamma^{\mathrm{mer}} \subset N_\Gamma$ denote the meromorphic lattice. Then M_Γ is contained in N_Γ^{mer} and is a $\mathbb{C}_{\mathrm{an}}[\widehat{Y}]$-lattice there. Indeed, it is enough to show this fiberwise (i.e., at any point of Y), where this is clear (to prove that M_Γ is a lattice we use the observation that N_Γ is the image of M_Γ under the KZ functor). For any other lattice M', we have $\nu^d M' \subset M_\Gamma \subset \nu^{-d} M'$ for some $d > 0$. So it is enough to show that $e(W')\eta_\Gamma^* N \cap M'$ is finitely generated for some lattice M'. Let us construct such an M'.

As a $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}}/W]$-module, N is projective and hence is a direct summand in a free module, say $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}}/W]^{\oplus r}$. So N_Γ^{mer} is a direct summand in $\mathbb{C}_{\mathrm{an}}[\widehat{Y}][\nu^{-1}]^{\oplus r}$. The intersection $M' := N_\Gamma^{\mathrm{mer}} \cap \mathbb{C}_{\mathrm{an}}[\widehat{Y}]^{\oplus r}$ (inside $\mathbb{C}_{\mathrm{an}}[\widehat{Y}][\nu^{-1}]^{\oplus r}$) is clearly a lattice in N_Γ^{mer}. Further, the intersection $M' \cap e(W')\eta_\Gamma^* N$ coincides with

$$e(W')\eta_\Gamma^* N \cap \mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W'}/W']^{\oplus r},$$

and hence is finitely generated (and clearly is a lattice in $e(W')\eta_\Gamma^* N$). □

Now we are ready to define a module $\widetilde{N}$ over $\mathcal{O}_{\mathfrak{h}^{\mathrm{sr}}/W} \otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W)$. Abusing notation, we will write $\widetilde{N}_\Gamma$ for the corresponding (under the equivalence) $\mathbb{C}[\mathfrak{h}^{\mathrm{reg}-W_\Gamma}/W_\Gamma] \otimes_{\mathbb{C}[\mathfrak{h}/W]} H_c(W)$-module. The restriction of $\widetilde{N}_\Gamma$ to $\eta_\Gamma^{-1}(\mathfrak{h}^{\mathrm{reg}}/W)$ coincides with $\eta_\Gamma^* N$ by construction. Let $\iota_\Gamma : \mathfrak{h}^{\mathrm{reg}-W_\Gamma} \hookrightarrow \mathfrak{h}$ be the inclusion and $\pi_\Gamma : \mathfrak{h}^{\mathrm{reg}-W_\Gamma} \twoheadrightarrow \mathfrak{h}^{\mathrm{reg}-W_\Gamma}/W_\Gamma$ be the quotient morphism. Also let $\pi : \mathfrak{h} \to \mathfrak{h}/W$ denote the quotient morphism and $\iota : \mathfrak{h}^{\mathrm{reg}} \hookrightarrow \mathfrak{h}^{\mathrm{sr}}$ the inclusion. Note that, by the construction, $\iota_{\Gamma *}\pi_\Gamma^* \widetilde{N}_\Gamma \subset \iota_* \pi^* N$ (recall that we view N as a coherent sheaf on $\mathfrak{h}^{\mathrm{reg}}/W$). The intersection $\widehat{N} := \bigcap_\Gamma \iota_{\Gamma *}\pi_\Gamma^* \widetilde{N}_\Gamma$ is a coherent sheaf on $\mathfrak{h}^{\mathrm{sr}}$ because of Lemma 3.2 and the equality $\mathfrak{h}^{\mathrm{sr}} = \bigcup_\Gamma \mathfrak{h}^{\mathrm{reg}-W_\Gamma}$. The intersection is stable under the Euler vector field because all the $\widetilde{N}_\Gamma$ are. Also $\widehat{N}$ is W-stable; this is because $w\pi_\Gamma^* \widetilde{N}_\Gamma = \pi_{w\Gamma}^* \widetilde{N}_{w\Gamma}$. Now set $\widetilde{N} := \pi_*(\widehat{N})^W = \pi_* \widehat{N} \cap \iota'_* N$ (where $\iota' : \mathfrak{h}^{\mathrm{sr}}/W \hookrightarrow \mathfrak{h}/W$ denotes the inclusion). This is a coherent sheaf on $\mathfrak{h}^{\mathrm{sr}}/W$, stable under the Euler vector field on $\iota'_* N$. It remains to show that $\widetilde{N} \subset \iota'_* N$ is stable under $H_c(W)$. But this follows from the equality

$$\widetilde{N} = \iota'_* N \cap \bigcap_\Gamma \eta_{\Gamma *} \widetilde{N}_\Gamma, \tag{3-2}$$

where now we view η_Γ as a morphism $\mathfrak{h}^{\mathrm{reg}-W'}/W' \to \mathfrak{h}^{\mathrm{sr}}/W$. Equation (3-2) follows from the observation that $\widetilde{N}_\Gamma = \pi_{\Gamma *}(\pi_\Gamma^* \widetilde{N}_\Gamma)^{W_\Gamma}$ and $\pi = \eta_\Gamma \circ \pi_\Gamma$. Since all sheaves in the right-hand side of (3-2) are stable under $H_c(W)$, we see that $\widetilde{N}$ is stable as well. It follows from the construction that $\widetilde{N}|_{\mathfrak{h}^{\mathrm{reg}}/W} \cong N$.

3D. *Global sections.*

Lemma 3.3. *The sheaf $\widetilde{N}$ is a vector bundle on $\mathfrak{h}^{\mathrm{sr}}/W$.*

Proof. The proof is inspired by [Etingof et al. 2013, Section 3.2]. We need to show that $\tilde{N}$ is maximal Cohen–Macaulay when viewed as a coherent sheaf on $\mathfrak{h}^{\mathrm{sr}}/W$. Let Z denote the non-CM locus of $\tilde{N}$ in $\mathfrak{h}^{\mathrm{sr}}/W$ and let d be the codimension of Z in $\mathfrak{h}^{\mathrm{sr}}/W$. Pick an open $\mathbb{C}^{\times}$-stable affine subvariety U of $\mathfrak{h}^{\mathrm{sr}}/W$ that intersects Z (or, more precisely, an irreducible component of maximal dimension in Z). Consider $H^i_{U\cap Z}(U,\tilde{N})$ for $i<d$. As in [Etingof et al. 2013, Section 3.2], all these groups are $\mathbb{C}[U]\otimes_{\mathbb{C}[\mathfrak{h}/W]}H_c(W)$-modules, finitely generated over $\mathbb{C}[U]$ (this follows from [SGA 2 1968, Exposé VIII, Corollary 2.3] using equivalence of (ii) and (iii) there; note that the depth used in (ii) coincides with the codimension thanks to the choice of Z). Moreover, by the choice of Z, one of these modules is nonzero, as in [Etingof et al. 2013, Section 3.2]. The support of $R:=H^i_{U\cap Z}(U,\tilde{N})$ is contained in $Z\cap U$.

Pick $b\in\mathfrak{h}$ lying over the support of R. Recall the isomorphism $\theta_b: H_c(W)^{\wedge_b}\cong Z(W,W',H_c(W',\mathfrak{h})^{\wedge_b})$, where we take W' to be W_b. So we get a nonzero $H_c(W',\mathfrak{h})^{\wedge_b}$-module $e(W')\theta_{b*}(R^{\wedge_b})$. This module is finitely generated over $\mathbb{C}[\mathfrak{h}]^{\wedge_b}$. So it is of the form $\mathbb{C}[\mathfrak{h}_{W'}]^{\wedge_b}\otimes R_0^{\wedge 0}$ for $R_0\in\mathcal{O}_c(W')$. It follows that $d=1$ and that $R=\Gamma_{Z\cap U}(U,\tilde{N})$. But, by construction, $\Gamma(U,\tilde{N})$ is embedded into $\Gamma(U\cap\mathfrak{h}^{\mathrm{reg}}/W,N)$ and so $\Gamma(U,\tilde{N})$ has no torsion $\mathbb{C}[U]$-submodules. We get a contradiction, showing that $\tilde{N}$ is Cohen–Macaulay. Since $\tilde{N}$ is torsion-free, we see that it is maximal Cohen–Macaulay, and hence is a vector bundle. □

Now we can use [SGA 2 1968, Exposé VIII, Corollary 2.3(iv)] (applied to an extension of $\tilde{N}$ to a coherent sheaf on $\mathfrak{h}/W$) to see that $M:=\Gamma(\mathfrak{h}^{\mathrm{sr}}/W,\tilde{N})$ is finitely generated over $\mathbb{C}[\mathfrak{h}]^W$. Let us show that the $H_c(W)$-module M lies in $\mathcal{O}_c(W)$. By construction, M carries a locally finite derivation compatible with the derivation $\operatorname{ad}h$ of $H_c(W)$. It follows that M is gradable and hence lies in $\mathcal{O}$. Also, by construction, $M[\delta^{-1}]=N$. This completes the proof.

Acknowledgements

This project was sparked by a discussion with Raphael Rouquier in April 2014. I would like to thank him and Roman Bezrukavnikov, Pavel Etingof and Dennis Gaitsgory for stimulating discussions. Furthermore, I am grateful to Pavel Etingof for his comments on a preliminary version of this paper and for pointing out the reference [Wilcox 2011], and Daniel Thompson for his careful reading of the paper. My work was supported by the NSF under grant DMS-1161584.

References

[Bezrukavnikov and Etingof 2009] R. Bezrukavnikov and P. Etingof, "Parabolic induction and restriction functors for rational Cherednik algebras", *Selecta Math.* (*N.S.*) **14**:3-4 (2009), 397–425. MR 2010e:20007 Zbl 1226.20002

[Broué et al. 1998] M. Broué, G. Malle, and R. Rouquier, "Complex reflection groups, braid groups, Hecke algebras", *J. Reine Angew. Math.* **500** (1998), 127–190. MR 99m:20088 Zbl 0921.20046

[Etingof and Ginzburg 2002] P. Etingof and V. Ginzburg, "Symplectic reflection algebras, Calogero–Moser space, and deformed Harish-Chandra homomorphism", *Invent. Math.* **147**:2 (2002), 243–348. MR 2003b:16021 Zbl 1061.16032

[Etingof et al. 2013] P. Etingof, E. Gorsky, and I. Losev, "Representations of rational Cherednik algebras with minimal support and torus knots", preprint, 2013. arXiv 1304.3412

[Ginzburg et al. 2003] V. Ginzburg, N. Guay, E. Opdam, and R. Rouquier, "On the category $\mathcal{O}$ for rational Cherednik algebras", *Invent. Math.* **154**:3 (2003), 617–651. MR 2005f:20010 Zbl 1071.20005

[Marin 2014] I. Marin, "The freeness conjecture for Hecke algebras of complex reflection groups, and the case of the Hessian group G_{26}", *J. Pure Appl. Algebra* **218**:4 (2014), 704–720. MR 3133700 Zbl 1291.20042

[Rouquier 2008] R. Rouquier, "q-Schur algebras and complex reflection groups", *Mosc. Math. J.* **8**:1 (2008), 119–158. MR 2010b:20081 Zbl 1213.20007

[SGA 2 1968] A. Grothendieck, *Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux* (Séminaire de Géométrie Algébrique du Bois Marie 1962 = SGA 2), Advanced Stud. in Pure Math. **2**, North-Holland, Amsterdam, 1968. MR 57 #16294 Zbl 0197.47202

[Shephard and Todd 1954] G. C. Shephard and J. A. Todd, "Finite unitary reflection groups", *Canadian J. Math.* **6** (1954), 274–304. MR 15,600b Zbl 0055.14305

[Wilcox 2011] S. Wilcox, "Supports of representations of the rational Cherednik algebra of type A", preprint, 2011. arXiv 1012.2585

Communicated by J. Toby Stafford
Received 2014-08-13 Accepted 2015-02-18

i.loseu@neu.edu *Department of Mathematics, Northeastern University, Boston, MA 02115, United States*

ALGEBRA AND NUMBER THEORY 9:2 (2015)
dx.doi.org/10.2140/ant.2015.9.503

Semiample invertible sheaves with semipositive continuous hermitian metrics

Atsushi Moriwaki

Let (L, h) be a pair of a semiample invertible sheaf and a semipositive continuous hermitian metric on a proper algebraic variety over $\mathbb{C}$. In this paper, we prove that (L, h) is semiample metrized, answering a generalization of a question of S. Zhang.

Introduction

Let X be a proper algebraic variety over $\mathbb{C}$. Let L be an invertible sheaf on X, and let h be a continuous hermitian metric of L. We say that (L, h) is *semiample metrized* if, for any $\epsilon > 0$, there is $n > 0$ such that, for any $x \in X(\mathbb{C})$, we can find $l \in H^0(X, L^{\otimes n}) \setminus \{0\}$ with

$$\sup\{h^{\otimes n}(l, l)(w) \mid w \in X(\mathbb{C})\} \le e^{\epsilon n} h^{\otimes n}(l, l)(x).$$

Shouwu Zhang proposed the following question:

Question 0.1 [Zhang 1995, Question 3.6]. If L is ample and h is smooth and semipositive, does it follow that (L, h) is semiample metrized?

Theorem 3.5 of the same reference gives an affirmative answer in the case where X is smooth over $\mathbb{C}$. The purpose of this paper is to give an answer for a generalization of the above question. First of all, we fix some notation: We say that L is *semiample* if there is a positive integer n_0 such that $L^{\otimes n_0}$ is generated by global sections. Moreover, h is said to be *semipositive* (or we say that (L, h) is semipositive) if, for any point $x \in X(\mathbb{C})$ and a local basis s of L on a neighborhood of x, $-\log h(s, s)$ is plurisubharmonic around x (for the definition of plurisubharmonicity on a singular variety, see Section 1). Note that h is not necessarily smooth. By using the recent work of Coman, Guedj and Zeriahi [Coman et al. 2013], we have the following answer:

Theorem 0.2. *If L is semiample and h is continuous and semipositive, then (L, h) is semiample metrized.*

MSC2010: primary 14C20; secondary 32U05, 14G40.
Keywords: semiample metrized, semipositive.

1. Plurisubharmonic functions on singular complex analytic spaces

Let T be a reduced complex analytic space. An upper-semicontinuous function

$$\varphi : T \to \mathbb{R} \cup \{-\infty\}$$

is said to be *plurisubharmonic* if $\varphi \not\equiv -\infty$ and, for each $x \in T$, there is an analytic closed embedding $\iota_x : U_x \hookrightarrow W_x$ of an open neighborhood U_x of x into an open set W_x of $\mathbb{C}^{n_x}$ together with a plurisubharmonic function Φ_x on W_x such that $\varphi|_{U_x} = \iota_x^*(\Phi_x)$. For an analytic map $f : T' \to T$ of reduced complex analytic spaces and a plurisubharmonic function φ on T, it is easy to see that $\varphi \circ f$ is either identically $-\infty$ or plurisubharmonic on T'. By [Fornæss and Narasimhan 1980, Theorem 5.3.1], an upper-semicontinuous function $\varphi : T \to \mathbb{R} \cup \{-\infty\}$ is plurisubharmonic if and only if, for any analytic map $\varrho : \mathbb{D} \to T$, $\varphi \circ \varrho$ is either identically $-\infty$ or subharmonic on $\mathbb{D}$, where $\mathbb{D} := \{z \in \mathbb{C} \mid |z| < 1\}$. Moreover, if T is compact and φ is plurisubharmonic on T, then φ is locally constant.

Let ω be a smooth $(1, 1)$-form on T, that is, in the same way as in the definition of plurisubharmonic functions, ω is a smooth $(1, 1)$-form on the regular part of T and, for each $x \in T$, there is an analytic closed embedding $\iota_x : U_x \hookrightarrow W_x$ of an open neighborhood U_x of x into an open set W_x of $\mathbb{C}^{n_x}$ together with a smooth $(1, 1)$-form Ω_x on W_x such that $\omega|_{U_x} = \iota_x^*(\Omega_x)$. We assume that ω is locally given by $dd^c(u)$ for some smooth function u on a neighborhood of x. Let ϕ be a *quasiplurisubharmonic function* on T; that is, for each $x \in T$, ϕ can be locally written as the sum of a smooth function and a plurisubharmonic function around x. We say that ϕ is *ω-plurisubharmonic* if there is an open covering $T = \bigcup_\lambda U_\lambda$, together with a smooth function u_λ on U_λ for each λ, such that $\omega|_{U_\lambda} = dd^c(u_\lambda)$ and $\phi|_{U_\lambda} + u_\lambda$ is plurisubharmonic on U_λ. The condition for ω-plurisubharmonicity is often denoted by $dd^c([\phi]) + \omega \geq 0$.

Here we consider the following lemma:

Lemma 1.1. *Let $f : X \to Y$ be a surjective and proper morphism of algebraic varieties over $\mathbb{C}$. Let φ be a real-valued function on $Y(\mathbb{C})$.*

(1) *φ is continuous if and only if $\varphi \circ f$ is continuous.*

(2) *Assume that φ is continuous. Then φ is plurisubharmonic if and only if $\varphi \circ f$ is plurisubharmonic.*

Proof. (1) It is sufficient to see that if $\varphi \circ f$ is continuous, then φ is continuous. Otherwise, there are $y \in Y(\mathbb{C})$, $\epsilon_0 > 0$ and a sequence $\{y_n\}$ on $Y(\mathbb{C})$ such that $\lim_{n\to\infty} y_n = y$ and $|\varphi(y_n) - \varphi(y)| \geq \epsilon_0$ for all n. We choose $x_n \in X(\mathbb{C})$ such that $f(x_n) = y_n$. As $f : X \to Y$ is proper, we can find a subsequence $\{x_{n_i}\}$ of $\{x_n\}$ such that $x := \lim_{i\to\infty} x_{n_i}$ exists in $X(\mathbb{C})$. Note that

$$f(x) = \lim_{i\to\infty} f(x_{n_i}) = \lim_{i\to\infty} y_{n_i} = y,$$

so that, as $\varphi \circ f$ is continuous,

$$\varphi(y) = (\varphi \circ f)(x) = \lim_{i\to\infty} (\varphi \circ f)(x_{n_i}) = \lim_{i\to\infty} \varphi(f(x_{n_i})) = \lim_{i\to\infty} \varphi(y_{n_i}),$$

which is a contradiction, so that φ is continuous.

(2) We need to check that if $\varphi \circ f$ is plurisubharmonic, then φ is plurisubharmonic. By using Chow's lemma, we may assume that $f : X \to Y$ is projective. Moreover, since the assertion is local with respect to Y, we may further assume that there is a closed embedding $\iota : X \hookrightarrow Y \times \mathbb{P}^N$ such that $p \circ \iota = f$, where $p : Y \times \mathbb{P}^n \to Y$ is the projection to the first factor. The remaining proof is same as the last part of the proof of [Demailly 1985, Theorem 1.7]. Let $g : (\mathbb{D}, 0) \to (Y, y)$ be a germ of an analytic map. By the theorem of Fornæss and Narasimhan, it is sufficient to show that $\varphi \circ g$ is subharmonic. Clearly we may assume that g is given by the normalization of a 1-dimensional irreducible germ (C, y) in (Y, y). Using hyperplanes in $\mathbb{P}^N$, we can find $x \in X$ and a 1-dimensional irreducible germ (C', x) in (X, x) such that (C', x) lies over (C, y). Let $g' : (\mathbb{D}, 0) \to (X, x)$ be the germ of an analytic map given by the normalization of (C', x). Then we have an analytic map $\sigma : (\mathbb{D}, 0) \to (\mathbb{D}, 0)$ with $g \circ \sigma = f \circ g'$:

$$\begin{array}{ccc} (\mathbb{D}, 0) & \xrightarrow{g'} & (X, x) \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle f} \\ (\mathbb{D}, 0) & \xrightarrow{g} & (Y, y) \end{array}$$

Changing a variable of $(\mathbb{D}, 0)$, we may assume that σ is given by $\sigma(z) = z^m$ for some positive integer m. Then $\varphi \circ g \circ \sigma$ is subharmonic because $\varphi \circ f$ is plurisubharmonic. Therefore, as σ is étale over the outside of 0, $\varphi \circ g$ is subharmonic on the outside of 0, and hence $\varphi \circ g$ is subharmonic on $(\mathbb{D}, 0)$ by the removability of singularities of subharmonic functions. □

2. Descent of a semipositive continuous hermitian metric

Here, we consider a descent problem of a semipositive continuous hermitian metric.

Theorem 2.1. *Let $f : X \to Y$ be a surjective and proper morphism of algebraic varieties over $\mathbb{C}$ with $f_*\mathcal{O}_X = \mathcal{O}_Y$. Let L be an invertible sheaf on Y. If h' is a semipositive continuous hermitian metric of $f^*(L)$, then there is a semipositive continuous hermitian metric h of L such that $h' = f^*(h)$.*

Proof. Let h_0 be a continuous hermitian metric of L on Y. There is a continuous function ϕ on $X(\mathbb{C})$ such that $h' = \exp(\phi) f^*(h_0)$. Let F be a subvariety of X such that F is an irreducible component of a fiber of $f : X \to Y$. Then, as

$$(f^*(L), h')|_F \simeq (\mathcal{O}_F, \exp(\phi|_F)),$$

we can see that $-\phi|_F$ is plurisubharmonic, so that $\phi|_F$ is constant. Therefore, for any point $y \in Y(\mathbb{C})$, $\phi|_{\mu^{-1}(y)}$ is constant because $\mu^{-1}(y)$ is connected, and hence there is a function ψ on $Y(\mathbb{C})$ such that $\psi \circ f = \phi$. By Lemma 1.1(1), ψ is continuous, so that, if we set $h := \exp(\psi)h_0$, then h is continuous on $Y(\mathbb{C})$ and $h' = f^*(h)$.

Finally, let us see that h is semipositive. As this is a local question on Y, we may assume that there is a local basis s of L over Y. If we set $\varphi = -\log h(s,s)$, then $\varphi \circ f$ is plurisubharmonic because h' is semipositive. Therefore, by Lemma 1.1(2), φ is plurisubharmonic, as required $\square$

3. The proof of Theorem 0.2

In the case where X is smooth over $\mathbb{C}$, L is ample and h is smooth, this theorem was proved by Zhang [1995, Theorem 3.5]. First we assume that L is ample. Then there are a positive integer n_0 and a closed embedding $X \hookrightarrow \mathbb{P}^N$ such that $\mathbb{O}_{\mathbb{P}^N}(1)|_X \simeq L^{\otimes n_0}$. Let h_{FS} be the Fubini–Study metric of $\mathbb{O}_{\mathbb{P}^n}(1)$. Let ϕ be the continuous function on $X(\mathbb{C})$ given by $h^{\otimes n_0} = \exp(-\phi)h_{\mathrm{FS}}|_X$. We set $\omega = c_1(\mathbb{O}_{\mathbb{P}^N}(1), h_{\mathrm{FS}})$. Then ϕ is $(\omega|_X)$-plurisubharmonic. Therefore, by [Coman et al. 2013, Corollary C], there is a sequence $\{\varphi_i\}$ of smooth functions on $\mathbb{P}^N(\mathbb{C})$ with the following properties:

(1) φ_i is ω-plurisubharmonic for all i.

(2) $\varphi_i \geq \varphi_{i+1}$ for all i.

(3) For $x \in X(\mathbb{C})$, $\lim_{i\to\infty} \varphi_i(x) = \phi(x)$.

Since X is compact and ϕ is continuous, (3) implies that the sequence $\{\varphi_i\}$ converges to ϕ uniformly on $X(\mathbb{C})$. We choose i such that $|\phi(x) - \varphi_i(x)| \leq \epsilon n_0/2$ for all $x \in X$. We set $h_i = \exp(-\varphi_i)h_{\mathrm{FS}}$. Then h_i is a semipositive smooth hermitian metric of $\mathbb{O}_{\mathbb{P}^N}(1)$. Therefore, there is a positive integer n_1 such that, for $x \in \mathbb{P}^N(\mathbb{C})$, we can find $l \in H^0(\mathbb{P}^N, \mathbb{O}_{\mathbb{P}^N}(n_1)) \setminus \{0\}$ with

$$\sup\{h_i^{\otimes n_1}(l,l)(w) \mid w \in \mathbb{P}^N(\mathbb{C})\} \leq e^{n_1(\epsilon n_0/2)} h_i^{\otimes n_1}(l,l)(x).$$

In particular, if $x \in X(\mathbb{C})$, then $l(x) \neq 0$ (so that $l|_X \neq 0$) and

$$\sup\{h_i^{\otimes n_1}(l,l)(w) \mid w \in X(\mathbb{C})\} \leq e^{\epsilon n_0 n_1/2} h_i^{\otimes n_1}(l,l)(x).$$

Note that

$$h^{\otimes n_0} e^{-\epsilon n_0/2} \leq h_i \leq h^{\otimes n_0} \tag{3-1}$$

on $X(\mathbb{C})$, because $h_i = h^{\otimes n_0}\exp(\phi - \varphi_i)$ and $-\epsilon n_0/2 \leq \phi - \varphi_i \leq 0$ on $X(\mathbb{C})$. Therefore,

$$\sup\{h^{\otimes n_0 n_1}(l,l)(w) \mid w \in X(\mathbb{C})\} e^{-n_0 n_1 \epsilon/2} \leq \sup\{h_i^{\otimes n_1}(l,l)(w) \mid w \in X(\mathbb{C})\}$$

and

$$h_i^{\otimes n_1}(l,l)(x) \le h^{\otimes n_0 n_1}(l,l)(x),$$

and hence

$$\sup\{h^{\otimes n_0 n_1}(l,l)(w) \mid w \in X(\mathbb{C})\} \le e^{n_1 n_0 \epsilon} h^{\otimes n_0 n_1}(l,l)(x),$$

as required.

In general, as L is semiample, there are a positive integer n_2, a projective algebraic variety Y over $\mathbb{C}$, a morphism $f : X \to Y$ and an ample invertible sheaf A on Y such that $f_*\mathcal{O}_X = \mathcal{O}_Y$ and $f^*(A) \simeq L^{\otimes n_2}$. Thus, by Theorem 2.1, there is a semipositive continuous hermitian metric k of A such that $(f^*(A), f^*(k)) \simeq (L^{\otimes n_2}, h^{\otimes n_2})$. Therefore, the assertion of the theorem follows from the previous observation.

4. A variant of Theorem 0.2

The following theorem is a consequence of Theorem 0.2 together with the arguments in [Zhang 1995, Theorem 3.3]. However, we can give a direct proof using ideas in the proof of Theorem 0.2.

Theorem 4.1. *Let X be a projective algebraic variety over $\mathbb{C}$. Let L be an ample invertible sheaf on X and let h be a semipositive continuous hermitian metric of L. Let us fix a reduced subscheme Y of X, $l \in H^0(Y, L|_Y)$ and a positive number ϵ. Then, for the given X, L, h, Y, l and ϵ, there is a positive integer n_1 such that, for all $n \ge n_1$, we can find $l' \in H^0(X, L^{\otimes n})$ with $l'|_Y = l^{\otimes n}$ and*

$$\sup\{h^{\otimes n}(l',l')(w) \mid w \in X(\mathbb{C})\} \le e^{n\epsilon} \sup\{h(l,l)(w) \mid w \in Y(\mathbb{C})\}^n.$$

Proof. In the case where X is smooth over $\mathbb{C}$ and h is smooth and positive, the assertion of the theorem follows from [Zhang 1995, Theorem 2.2], in which Y is actually assumed to be a subvariety of X. However, the proof works well under the assumption that Y is a reduced subscheme. First of all, let us see the theorem in the case where X is smooth over $\mathbb{C}$ and h is smooth and semipositive. As L is ample, there is a positive smooth hermitian metric t of L with $t \le h$. Let us choose a positive integer m such that $e^{-\epsilon/2} \le (t/h)^{1/m} \le 1$ on $X(\mathbb{C})$. If we set $t_m = h^{1-1/m} t^{1/m}$, then t_m is smooth and positive, so that, for a sufficiently large integer n, there is $l' \in H^0(X, L^{\otimes n})$ such that $l'|_Y = l^{\otimes n}$ and

$$\sup\{t_m^{\otimes n}(l',l')(w) \mid w \in X(\mathbb{C})\} \le e^{n\epsilon/2} \sup\{t_m(l,l)(w) \mid w \in Y(\mathbb{C})\}^n,$$

and hence the assertion follows because $e^{-\epsilon/2} h \le t_m \le h$ on $X(\mathbb{C})$.

For a general case, we use the same symbols n_0, $X \hookrightarrow \mathbb{P}^N$, h_{FS}, ϕ, ω and $\{\varphi_i\}$ as in the proof of Theorem 0.2. Clearly we may assume that $l \neq 0$. Since L is ample, if a_0 is a sufficiently large integer, then, for each $j = 0, \dots, n_0 - 1$, there is

$l_j \in H^0(X, L^{\otimes n_0 a_0 + j})$ with $l_j|_Y = l^{\otimes n_0 a_0 + j}$. Let us fix a positive number A such that

$$\sup\{h^{\otimes n_0 a_0 + j}(l_j, l_j)(w) \mid w \in X(\mathbb{C})\} \le e^A \sup\{h(l, l)(w) \mid w \in Y(\mathbb{C})\}^{n_0 a_0 + j} \quad (4\text{-}1)$$

for $j = 0, \dots, n_0 - 1$. We choose i with $|\phi(x) - \varphi_i(x)| \le \epsilon n_0/2$ for all $x \in X$, and we set $h_i = \exp(-\varphi_i) h_{\mathrm{FS}}$. As h_i is smooth and semipositive, for the given $\mathbb{P}^N$, $\mathcal{O}_{\mathbb{P}^N}(1)$, h_i, Y, $l^{\otimes n_0}$ (as an element of $H^0(Y, \mathcal{O}_{\mathbb{P}^N}(1)|_Y)$) and $n_0\epsilon/4$, there is a positive integer a_1 such that the assertion of the theorem holds for all $a \ge a_1$. We put

$$n_1 := n_0 \max\left\{a_1 + a_0 + 1,\ \frac{4A}{n_0 \epsilon} - 3a_0 + 1\right\}.$$

Let n be an integer with $n \ge n_1$. If we set $n = n_0(a + a_0) + j$ $(0 \le j \le n_0 - 1)$, then

$$a \ge a_1 \quad \text{and} \quad a \ge \frac{4A}{n_0 \epsilon} - 4a_0,$$

so that we can find $l'' \in H^0(\mathbb{P}^N, \mathcal{O}_{\mathbb{P}^N}(a))$ with $l''|_Y = l^{\otimes n_0 a}$ and

$$\sup\{h_i^{\otimes a}(l'', l'')(w) \mid w \in \mathbb{P}^N(\mathbb{C})\} \le e^{a(n_0\epsilon/4)} \sup\{h_i(l^{\otimes n_0}, l^{\otimes n_0})(w) \mid w \in Y(\mathbb{C})\}^a,$$

which implies that

$$\sup\{h^{\otimes n_0 a}(l'', l'')(w) \mid w \in X(\mathbb{C})\} \le e^{(3/4) n_0 a \epsilon} \sup\{h(l, l)(w) \mid w \in Y(\mathbb{C})\}^{n_0 a}, \quad (4\text{-}2)$$

because of (3-1). Here we set $l' = l'' \otimes l_j$. Then, $l'|_Y = l^{\otimes n}$ and, using (4-1) and (4-2), we have

$$\begin{aligned}
&\sup\{h^{\otimes n}(l', l')(w) \mid w \in X(\mathbb{C})\} \\
&\qquad \le \sup\{h^{\otimes n_0 a}(l'', l'')(w) \mid w \in X(\mathbb{C})\} \sup\{h^{\otimes n_0 a_0 + j}(l_j, l_j)(w) \mid w \in X(\mathbb{C})\} \\
&\qquad \le e^{(3/4) n_0 a \epsilon + A} \sup\{h(l, l)(w) \mid w \in Y(\mathbb{C})\}^n,
\end{aligned}$$

which implies the assertion because $(3/4) n_0 a \epsilon + A \le \epsilon n$. □

5. Arithmetic application

As an application of Theorem 0.2, we have the following generalization of the arithmetic Nakai–Moishezon criterion (see [Zhang 1995, Corollary 4.8]).

Corollary 5.1. *Let $\mathscr{X}$ be a projective and flat integral scheme over $\mathbb{Z}$. Let $\mathscr{L}$ be an invertible sheaf on $\mathscr{X}$ such that $\mathscr{L}$ is nef on every fiber of $\mathscr{X} \to \mathbb{Z}$. Let h be an F_∞-invariant semipositive continuous hermitian metric of $\mathscr{L}$, where F_∞ is the complex conjugation map $\mathscr{X}(\mathbb{C}) \to \mathscr{X}(\mathbb{C})$. If $\widehat{\deg}(\hat{c}_1((\mathscr{L}, h)|_{\mathscr{Y}})^{\dim \mathscr{Y}}) > 0$ for all horizontal integral subschemes $\mathscr{Y}$ of $\mathscr{X}$, then, for an F_∞-invariant continuous hermitian invertible sheaf $(\mathscr{M}, k)$ on $\mathscr{X}$, $H^0(\mathscr{X}, \mathscr{L}^{\otimes n} \otimes \mathscr{M})$ has a basis consisting of strictly small sections for a sufficiently large integer n.*

Proof. Let X be the generic fiber of $\mathscr{X} \to \operatorname{Spec}(\mathbb{Z})$ and let Y be a subvariety of X. Let $\mathscr{Y}$ be the Zariski closure of Y in $\mathscr{X}$. As

$$\widehat{\deg}(\hat{c}_1((\mathscr{L}, h)|_{\mathscr{Y}})^{\dim \mathscr{Y}}) > 0,$$

$(\mathscr{L}, h)|_{\mathscr{Y}}$ is big by [Moriwaki 2012, Theorem 6.6.1], so that $H^0(\mathscr{Y}, \mathscr{L}^{\otimes n_0}|_{\mathscr{Y}}) \setminus \{0\}$ has a strictly small section for a sufficiently large integer n_0. Moreover, if we set $L = \mathscr{L}|_X$, then $L|_Y$ is big, and hence $\deg(L^{\dim Y} \cdot Y) > 0$ because L is nef. Therefore, L is ample by the Nakai–Moishezon criterion for ampleness. In particular, by Theorem 0.2, h is semiample metrized. Thus the assertion follows from the arguments in [Zhang 1995, Theorem 4.2]. □

Acknowledgements

I would like to thank Professor Zhang and the referee for their comments and suggestions.

References

[Coman et al. 2013] D. Coman, V. Guedj, and A. Zeriahi, "Extension of plurisubharmonic functions with growth control", *J. Reine Angew. Math.* **676** (2013), 33–49. MR 3028754 Zbl 1269.32018

[Demailly 1985] J.-P. Demailly, *Mesures de Monge–Ampère et caractérisation géométrique des variétés algébriques affines*, Mém. Soc. Math. France (N.S.) **19**, 1985. MR 87g:32030 Zbl 0579.32012

[Fornæss and Narasimhan 1980] J. E. Fornæss and R. Narasimhan, "The Levi problem on complex spaces with singularities", *Math. Ann.* **248**:1 (1980), 47–72. MR 81f:32020 Zbl 0411.32011

[Moriwaki 2012] A. Moriwaki, "Zariski decompositions on arithmetic surfaces", *Publ. Res. Inst. Math. Sci.* **48**:4 (2012), 799–898. MR 2999543 Zbl 1281.14017

[Zhang 1995] S. Zhang, "Positive line bundles on arithmetic varieties", *J. Amer. Math. Soc.* **8**:1 (1995), 187–221. MR 95c:14020 Zbl 0861.14018

Communicated by Shou-Wu Zhang
Received 2014-11-09 Revised 2015-01-01 Accepted 2015-02-16

moriwaki@math.kyoto-u.ac.jp *Department of Mathematics, Faculty of Science, Kyoto University, Kyoto 606-8502, Japan*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 9 No. 2 2015