Algebra & Number Theory

Volume 7

2013

No. 7

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2013 is US $200/year for the electronic version, and $350/year (+$40, 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 Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1535

Weil representation and transfer factor

Teruji Thomas

This paper concerns the Weil representation of the semidirect product of the metaplectic and Heisenberg groups. First we present a canonical construction of the metaplectic group as a central extension of the symplectic group by a subquotient of the Witt group. This leads to simple formulas for the character, for the inverse Weyl transform, and for the transfer factor appearing in J. Adams's work on character lifting. Along the way, we give formulas for outer automorphisms of the metaplectic group induced by symplectic similitudes. The approach works uniformly for finite and local fields.

1. Introduction

1.1. This paper presents some calculations related to the character of the Weil representation. This representation has a fundamental role in the representation theory of the symplectic group and in many related contexts. Before explaining the results, let us recall the classical theory as explained by Lion and Vergne [1980].

Let V be a finite-dimensional vector space, with symplectic form ω. The ground field may be any finite or local field F of characteristic not 2; for example, most classically, F could be the real numbers. Let $\mathrm{Sp}(V)$ be the corresponding symplectic group, that is, the group of automorphisms of V preserving ω. Choose a nontrivial, continuous group homomorphism $\psi : F \to U(1) \subset \mathbb{C}^\times$; for example, in the case of the real numbers, one may take $\psi(x) = e^{ix}$. Choose also a Lagrangian subspace $\ell \subset V$. From the data (ψ, ℓ), one constructs a central extension

$$1 \to Z_F \to \mathrm{Mp}^{\psi,\ell}(V) \to \mathrm{Sp}(V) \to 1. \tag{1}$$

$\mathrm{Mp}^{\psi,\ell}(V)$ is known as "the" metaplectic group; as we will see, it is essentially independent of ψ and ℓ. In the special case when F is $\mathbb{C}$ or a finite field,[1] the central factor Z_F is trivial, so that $\mathrm{Mp}^{\psi,\ell}(V)$ is nothing but the symplectic group $\mathrm{Sp}(V)$; in all other cases, $Z_F = Z_2 = \{\pm 1\}$, and the extension is nontrivial. For example, when $F = \mathbb{R}$, $\mathrm{Mp}^{\psi,\ell}(V)$ is the unique connected double cover of $\mathrm{Sp}(V)$.

MSC2010: primary 11F27; secondary 20C15.

Keywords: metaplectic group, Weil representation, Weyl transform, transfer factor, Cayley transform, Maslov index.

[1] See Section 1.5.2 for more simplifications in these cases.

The construction of $\mathrm{Mp}^{\psi,\ell}(V)$ goes hand-in-hand with the construction of a unitary representation $\rho^{\psi,\ell}_{\mathrm{Mp}}$, known as the Weil representation (also as the oscillator or metaplectic representation). One starts from the Heisenberg group $H(V)$, which is a central extension of V by F as additive groups; thus

$$H(V) = V \times F \quad \text{(as a set)}.$$

Associated to the data (ψ, ℓ) is an irreducible unitary representation $\rho^{\psi,\ell}_H$ of $H(V)$ whose restriction to the center $F \subset H(V)$ is ψ (it is, up to nonunique isomorphism, the unique such representation, but its construction depends also on ℓ). Meanwhile, the natural action of $\mathrm{Sp}(V)$ on V defines a semidirect product $\mathrm{Sp}(V) \ltimes H(V)$. The central extension $\mathrm{Mp}^{\psi,\ell}(V)$ is defined so that $\rho^{\psi,\ell}_H$ naturally extends to a representation $\rho^{\psi,\ell}$ of the covering group $\mathrm{Mp}^{\psi,\ell}(V) \ltimes H(V)$. Its restriction to $\mathrm{Mp}^{\psi,\ell}(V)$ is the Weil representation $\rho^{\psi,\ell}_{\mathrm{Mp}}$.

1.2. A number of people have recently studied the character $\operatorname{Tr} \rho^{\psi,\ell}$, defined to be the generalized function on $\mathrm{Mp}^{\psi,\ell}(V) \ltimes H(V)$ whose integral against any smooth, compactly supported measure h on $\mathrm{Mp}^{\psi,\ell}(V) \ltimes H(V)$ is

$$\int h \cdot \operatorname{Tr} \rho^{\psi,\ell} = \operatorname{Tr}\left(\int h \cdot \rho^{\psi,\ell}\right). \tag{2}$$

(The right-hand side is the trace of a trace-class operator — see Remark 5.3.1.) The studies mentioned make some restrictions, focusing on $\mathrm{Mp}^{\psi,\ell}(V)$ (e.g., [Thomas 2008]), or on some open subset (e.g., [Maktouf 1999; Gurevich and Hadani 2007]), and/or making a particular choice of field (e.g., [de Gosson and Luef 2009] for the reals, [Gurevich and Hadani 2007; Prasad 2009] for finite fields). This article completes the project in the following ways.

(A) The different metaplectic groups $\mathrm{Mp}^{\psi,\ell}$ corresponding to varying data (ψ, ℓ) are canonically isomorphic. The first task is to construct an extension

$$1 \to Z_F \to \mathrm{Mp}(V) \to \mathrm{Sp}(V) \to 1 \tag{3}$$

isomorphic to (1), but defined without any reference to ψ and ℓ. Using this canonical construction, we give explicit formulas for the isomorphisms between the various groups $\mathrm{Mp}^{\psi,\ell}(V)$. As a by-product, we find explicit formulas for the conjugation action of $\mathrm{GSp}(V)$ on $\mathrm{Mp}(V)$ and $\mathrm{Mp}^{\psi,\ell}(V)$.

(B) Because of (A), every Weil representation $\rho^{\psi,\ell}$ can be considered as a representation of the single group $\mathrm{Mp}(V) \ltimes H(V)$. We give a formula for the character $\operatorname{Tr} \rho^{\psi,\ell}$ as a generalized function on $\mathrm{Mp}(V) \ltimes H(V)$. The isomorphisms described in (A) allow easy translation of this character formula to other versions of the metaplectic group.

(C) The answer to (B) also yields explicit formulas for the "invariant presentation", or inverse Weyl transform, of $\rho_{\mathrm{Mp}}^{\psi,\ell}$; this is (roughly speaking) a homomorphism from $\mathrm{Mp}(V)$ into the ψ-coinvariant group algebra of $H(V)$.

(D) Writing $\rho_{\mathrm{Mp}}^{\psi,\ell} = \rho_+^{\psi,\ell} \oplus \rho_-^{\psi,\ell}$ as the direct sum of two irreducibles, we calculate the character of the virtual representation $\rho_+^{\psi,\ell} - \rho_-^{\psi,\ell}$ (which then determines the characters of $\rho_+^{\psi,\ell}$ and $\rho_-^{\psi,\ell}$ separately). This is a generalized function on $\mathrm{Mp}(V)$. Over a finite field, the method leads naturally to a "geometric" version of this virtual character, in the sense of Grothendieck's sheaf-function dictionary.

The virtual character in (D) plays a key role in Jeff Adams's theory [1998] of character lifting between metaplectic and orthogonal groups, which provides one of my main motivations for studying this subject.

Remark 1.2.1. The method for (B) is closely related to Roger Howe's wonderful unpublished notes [1973], and some similar ideas have been exploited by Gurevich and Hadani [2007] over finite fields, and de Gosson and Luef [2009] over the reals. In particular, the work of de Gosson ([op. cit.] and references therein) gives a very nice, and closely related, character formula in terms of the Conley–Zehnder index of paths in the real symplectic group.

1.3. ***Results.*** **(A)** The construction of the canonical metaplectic extension (3) proceeds in two steps, which make sense for *any* field F of characteristic not 2. The details are given in Section 2; here we outline the basic features, to fix our notation. First we define a central extension

$$0 \to W(F)/I^3 \to M(V) \to \mathrm{Sp}(V) \to 1$$

where $W(F)$ is the Witt ring of quadratic spaces over F, and $I \subset W(F)$ is the ideal of even-dimensional quadratic spaces (see Section A.1 in the Appendix). This construction is by means of a cocycle, so that

$$M(V) = \mathrm{Sp}(V) \times W(F)/I^3 \quad \text{as a set.}$$

Second, we define $\mathrm{Mp}(V)$ to be a certain subgroup of $M(V)$. In short, $\mathrm{Mp}(V)$ is the unique subgroup extending $\mathrm{Sp}(V)$ by I^2/I^3:

$$0 \to I^2/I^3 \to \mathrm{Mp}(V) \to \mathrm{Sp}(V) \to 1. \tag{4}$$

It turns out (see Theorem A.2) that, for a finite or local field, we can identify I^2/I^3 with the group Z_F, thus obtaining (3) as a special case. Concretely, for each $g \in \mathrm{Sp}(V)$, define a bilinear form σ_g on $(g-1)V$ by the formula

$$\sigma_g((g-1)x, (g-1)y) = \omega(x, (g-1)y) \quad \text{for all } x, y \in V.$$

Then σ_g is nondegenerate as a bilinear form, but, in general, asymmetric. It nonetheless has a rank $\dim \sigma_g = \dim(g-1)V$ and discriminant $\det \sigma_g \in F^\times/(F^\times)^2$. This is enough to determine a class $[\sigma_g]$ in $W(F)/I^2$ — the class of quadratic spaces with the same rank modulo 2 and the same signed discriminant as σ_g (see Section A.1). The definition of $\mathrm{Mp}(V)$ is as follows:

$$\mathrm{Mp}(V) = \{(g, q) \in M(V) \mid q = [\sigma_g] \bmod I^2/I^3\}.$$

In Proposition 2.4 we show that this definition makes $\mathrm{Mp}(V)$ into a subgroup of $M(V)$, and therefore obviously an extension of $\mathrm{Sp}(V)$ by I^2/I^3.

In Section 2.6 we also recall the construction of $\mathrm{Mp}^{\psi,\ell}(V)$ from [Lion and Vergne 1980] — this construction requires F to be finite or local. In Section 3 we describe canonical isomorphisms $\mathrm{Mp}(V) \to \mathrm{Mp}^{\psi,\ell}(V)$. They are "canonical" in the sense of being unique as isomorphisms of central extensions; see Section 3.1.

Remark 1.3.1. The idea of constructing an extension by I^2/I^3 comes from [Parimala et al. 2000] (using, however, a choice of Lagrangian $\ell \subset V$; see Section 2.7.1 for a synopsis). It also follows from the work of Suslin [1987] that these extensions can be characterized by a universal property; see Remark 2.5.1.

Remark 1.3.2. The Weil representation (which, again, is defined only when F is a finite or local field) can be extended very naturally to a representation of $M(V)$ rather than $\mathrm{Mp}(V)$, and practically all the results stated herein for $\mathrm{Mp}(V)$ hold also for $M(V)$. However, we will continue to refer primarily to $\mathrm{Mp}(V)$, to connect better with the literature.

(B) For the rest of this introduction, we take F to be a finite or local field, so that $\rho^{\psi,\ell}$ is defined (we recall the definition in Section 4). We consider it as a representation of $\mathrm{Mp}(V) \ltimes H(V)$. To describe its character, we need some further notation.

Notation. Let $\gamma_\psi : W(F)/I^3 \to \mathbb{C}^\times$ be the *Weil index* (see Section A.3, especially A.4.1(d)). For any $g \in \mathrm{Sp}(V)$, let Q_g be the associated *Cayley form*: it is a symmetric, usually degenerate, bilinear form on $(g-1)V$ defined by

$$Q_g((g-1)x, (g-1)y) := \tfrac{1}{2}\omega((g+1)x, (g-1)y) \quad \text{for all } x, y \in V.$$

Some further comments about the Cayley form are given in Section A.6.

Finally, let μ_{σ_g} be the Haar measure on $(g-1)V$ self-dual with respect to $\psi \circ \sigma_g$, and μ_V the Haar measure on V self-dual with respect to $\psi \circ \omega$ (see Section A.3.1 for conventions on measures). Define a generalized function D_g^ψ on V by the equation

$$\int_V f D_g^\psi \, \mu_V = \int_{(g-1)V} f \, \mu_{\sigma_g} \tag{5}$$

for all compactly supported, smooth functions f on V.

If F is a finite field, then this definition amounts to the following: D_g^ψ is the function on V supported on $(g-1)V$ and equal there to the constant $\sqrt{\#\ker(g-1)}$. When F is infinite, we just have $D_g^\psi(v) = \|\det(g-1)\|^{-1/2}$ if $\det(g-1) \neq 0$ (and, as standard, we choose the norm $\|\cdot\|$ on $F^\times$ such that $d(ax) = \|a\|\,dx$ for any translation-invariant measure dx on F).

Theorem B (character formula). *For fixed $(g, q) \in \mathrm{Mp}(V)$, the character*

$$T^\psi_{(g,q)}(v, t) := \operatorname{Tr} \rho^{\psi,\ell}(g, q; v, t)$$

is a well-defined generalized function of $(v, t) \in H(V)$, supported on $(g-1)V \times F$, and given by

$$T^\psi_{(g,q)}(v, t) = \psi\big(\tfrac{1}{2} Q_g(v, v)\big) \cdot D_g^\psi(v) \cdot \gamma_\psi(q) \cdot \psi(t).$$

The main part of the proof, using the Weyl transform, is given in Section 5. Note that the right-hand side is manifestly independent of ℓ, reflecting the independence of $\rho^{\psi,\ell}$ up to nonunique isomorphism.

Theorem B can be read as a formula for a locally integrable function[2] on

$$\mathrm{Mp}(V) \ltimes H(V)$$

representing $\operatorname{Tr} \rho^{\psi,\ell}$, but it says something more precise. The point is that, when F is infinite, $\operatorname{Tr} \rho^{\psi,\ell}$ is smooth almost everywhere, but “blows up” on the locus where $\det(g-1) = 0$. Theorem B gives a natural extension of $\operatorname{Tr} \rho^{\psi,\ell}$ to that singular locus — “natural” in the sense that it satisfies Theorem C below.

If we are only interested in the representation $\rho^{\psi,\ell}_{\mathrm{Mp}}$ of $\mathrm{Mp}(V)$ then Theorem B takes on the following simple form. Let

$$D^0(g) := \sqrt{\# V^g} \quad \text{or} \quad D^0(g) := \|\det(g-1)\|^{-1/2}$$

depending on whether F is finite or infinite. Here $V^g := \ker(g-1)$.

Corollary 1.4 (restriction to $\mathrm{Mp}(V)$). *As generalized functions of $(g, q) \in \mathrm{Mp}(V)$,*

$$\operatorname{Tr} \rho^{\psi,\ell}_{\mathrm{Mp}}(g, q) = D^0(g) \cdot \gamma_\psi(q).$$

The extreme simplicity of this formula suggests that the cocycle we have used to define $\mathrm{Mp}(V)$ is the natural one in this context. In particular, it is much better than the formula we developed in [Thomas 2008]. (In Remark 2.8.2 we explain how the thing called $\mathrm{Mp}(V)$ in that work is related to the present one.)

---

[2] That is, for F infinite, $D_g^\psi(v) = \|\det(g-1)\|^{-1/2}$ for almost all $(g, v) \in \mathrm{Sp}(V) \times V$, and $(g, q, v, t) \mapsto \psi(\frac{1}{2} Q_g(v, v)) \cdot \|\det(g-1)\|^{-1/2} \cdot \gamma_\psi(q) \cdot \psi(t)$ is locally integrable on $\mathrm{Mp}(V) \ltimes H(V)$: the modulus is just $\|\det(g-1)\|^{-1/2}$, so that the singularities of order $k/2$ lie in subspaces of codimension at least k.

(C) The formula of Theorem B also makes explicit the "invariant presentation" of the Weil representation emphasized, for example, in [Gurevich and Hadani 2007]. Let us recall that description. Let $\mathscr{A}_\psi$ be the L^2-completion of the ψ-coinvariant group-algebra of $H(V)$. In more detail, we consider functions on $H(V)$ that transform by ψ under the action of the center $F \subset H(V)$; these can be identified (by restriction) with functions on V. With that in mind, we define $\mathscr{A}_\psi$ to consist of all complex L^2 functions on V, equipped with the "convolution" multiplication induced by the multiplication on $H(V)$:

$$(f_1 \star f_2)(x) := \int_{v\in V} f_1(v)\, \psi\big(\tfrac{1}{2}\omega(v,x)\big)\, f_2(x-v)\, \mu_V.$$

(Here μ_V again denotes the Haar measure on μ_V that is self-dual with respect to $\psi \circ \omega$.) It is well known, and we prove in Proposition 5.2, that there is an isomorphism $W^{\psi,\ell}$ from $\mathscr{A}_\psi$ to the algebra of Hilbert–Schmidt operators on the representation space of $\rho^{\psi,\ell}$. This $W^{\psi,\ell}$ is called the *Weyl transform*.

Theorem C. *For any $f \in \mathscr{A}_\psi$, the convolution $T^\psi_{(g,q)} \star f$ is well-defined and lies in $\mathscr{A}_\psi$, and*

$$W^{\psi,\ell}(T^\psi_{(g,q)} \star f) = \rho^{\psi,\ell}(g,q) \circ W^{\psi,\ell}(f).$$

Theorem C may be restated more transparently when F is a finite field: it says that the map $(g,q) \mapsto T^\psi_{(g,q)}$ is a multiplicative homomorphism $\mathrm{Mp}(V) \to \mathscr{A}_\psi$, and $W^{\psi,\ell}(T^\psi_{(g,q)}) = \rho^{\psi,\ell}(g,q)$.

Versions of Theorem C are well known (see for example [Gurevich and Hadani 2007, §1.2; Howe 1973, Theorem 2.9]), so the new aspect is the explicit formula provided by Theorem B; nonetheless, we will find it convenient and easy to prove Theorem C in Section 6.

(D) The representation space of $\rho^{\psi,\ell}$ can be understood as the space of L^2 functions on V/ℓ. One has a decomposition

$$\rho^{\psi,\ell}_{\mathrm{Mp}} = \rho^{\psi,\ell}_{+} \oplus \rho^{\psi,\ell}_{-}$$

into irreducibles, where $\rho^{\psi,\ell}_{+}$ acts on the subspace of even functions, and $\rho^{\psi,\ell}_{-}$ on the subspace of odd ones. In Section 7 we give two proofs of the following result.

Theorem D. *As generalized functions of $(g,q) \in \mathrm{Mp}(V)$,*

$$\mathrm{Tr}\big(\rho^{\psi,\ell}_{+} - \rho^{\psi,\ell}_{-}\big)(g,q) = \gamma_\psi(Q_g) \cdot \mathrm{Tr}\, \rho^{\psi,\ell}_{\mathrm{Mp}}(-g,q).$$

Again, the right-hand side in Theorem D is manifestly independent of ℓ.

Remark 1.4.1. One knows on general grounds that the characters $\mathrm{Tr}\, \rho^{\psi,\ell}_{\pm}$ are well-defined (see [Harish-Chandra 1954] for the real case and [Sliman 1984, Theorem 1.2.3] for admissibility in the nonarchimedean case).

Geometrization. Suppose that $F = \mathbb{F}_q$ is a finite field.[3] In this situation, the central extension (7) is split, so that we may consider $\rho^{\psi,\ell}$ as a representation of $\mathrm{Sp}(V) \ltimes H(V)$. We can also consider $\mathrm{Sp}(V) \ltimes H(V)$ as the $\mathbb{F}_q$-points of a group scheme $\mathbb{G} = \mathbb{S}\mathrm{p}(V) \ltimes \mathbb{H}(V)$. Gurevich and Hadani [2007] have constructed an irreducible perverse sheaf $\mathcal{K}$ on $\mathbb{G}$ corresponding (under Grothendieck's sheaf-function dictionary) to the character $\mathrm{Tr}\,\rho^{\psi,\ell}$. The proof of Theorem D (specifically (33)) shows that there is, as well, an irreducible perverse sheaf $\mathcal{K}'$ on $\mathbb{G}$ whose pull-back to $\mathbb{S}\mathrm{p}(V)$ corresponds to the virtual character $\mathrm{Tr}\big(\rho_+^{\psi,\ell} - \rho_-^{\psi,\ell}\big)$; namely, $\mathcal{K}'$ is just the Fourier–Deligne transform of $\mathcal{K}$ along V with respect to the pairing $\psi \circ \frac{1}{2}\omega$.

Remark 1.4.2. The fact (33) that $\mathrm{Tr}\big(\rho_+^{\psi,\ell} - \rho_-^{\psi,\ell}\big)$ is related to $\mathrm{Tr}\,\rho^{\psi,\ell}$ by a Fourier transform explains the relationship between Theorem B and Theorem D: recall (Theorem A.4) that the $\gamma_\psi(Q_g)$ appearing in Theorem D is itself related by Fourier transform to the $\psi \circ \frac{1}{2} Q_g$ appearing in Theorem B.

1.5. *Remarks.*

1.5.1. *Dependence on ψ.* Let us briefly clarify the dependence of our results on the character ψ. For any chosen ψ, any other nontrivial additive character is uniquely of the form $\psi_a(x) = \psi(ax)$, with $a \in F^\times$. The isomorphism class of $\rho_{\mathrm{Mp}}^{\psi_a,\ell}$ depends only on the class of a modulo $(F^\times)^2$. For $(g, q) \in \mathrm{Mp}(V)$, we have

$$\gamma_{\psi_a}(q) = \gamma_\psi(q) \cdot (\gamma_\psi(a)/\gamma_\psi(1))^{\dim(g-1)V} \, (a, \det \sigma_g)_H$$

where $(\cdot\,,\cdot)_H$ is the Hilbert symbol (see Lemma 3.13 and Section A.1.2). Moreover, $D_g^{\psi_a} = D_g^{\psi} \cdot \|a\|^{-(\dim V^g)/2}$ (see Section A.3.1).

1.5.2. *Special fields.* The framework presented here gives a uniform treatment for any choice of field F. However, some simplifications are possible, case by case.

When $F = \mathbb{C}$, the central factor Z_F is trivial, and both γ_ψ and the Hilbert symbol always equal 1. When F is finite, Z_F is again trivial. This means that for each g, there is a unique $q \in W(F)/I^3$ with $(g, q) \in \mathrm{Mp}(V)$. One has

$$\gamma_\psi(q) = \gamma_\psi(1)^{\dim(g-1)V-1} \gamma_\psi(\det \sigma_g).$$

Moreover, the Hilbert symbol always equals 1, γ_ψ takes values in the fourth roots of unity Z_4 (or even Z_2 if -1 is a square), and the common expression $\gamma_\psi(a)/\gamma_\psi(1)$ equals 1 if a is a square, and -1 if not.

2. Metaplectic cocycles

In this section we construct the canonical metaplectic extension (4), which exists for any field of characteristic not 2. We also recall the traditional construction (1)

[3] Lafforgue and Lysenko [2009] have also considered a geometric version of the even part of the Weil representation over a *local* field $\mathbb{F}_q((t))$.

in Section 2.6, which makes sense only for a finite or local field, and depends on the choice of a Lagrangian ℓ and a character ψ. In Section 2.7 we examine these choices more closely. This will allow us to give explicit isomorphisms between all these various incarnations of the metaplectic group in Section 3.

The key tools are the Maslov index τ and the Weil index γ_ψ. The relevant facts and notation concerning these objects are recalled in Appendix A.

2.1. ***Generalities.*** Suppose that G is a group and A an abelian group, written additively; by a 2-cocycle $c : G \times G \to A$ we mean a function such that

$$c(g, g') - c(g, g'g'') + c(gg', g'') - c(g', g'') = 0 \quad \text{and} \quad c(1, 1) = 0. \qquad (6)$$

Given such a 2-cocycle, define $\tilde{G} = G \times A$ as a set, with a multiplication operation

$$(g, a)(g', a') := (gg', a + a' + c(g, g')).$$

Then it follows from (6) that $\tilde{G}$ is a group, with A as a central subgroup, and $G = \tilde{G}/A$. In other words, we have constructed a central extension

$$0 \to A \to \tilde{G} \to G \to 1.$$

Now let us apply this construction to various 2-cocycles, with $G = \mathrm{Sp}(V)$.

2.2. ***The canonical cocycle.*** Here we allow F to be any field (but always of characteristic not 2). Let $\overline{V}$ be the symplectic vector space $(V, -\omega)$. Then for each $g \in \mathrm{Sp}(V)$, the graph $\Gamma_g = \{(x, gx) \in \overline{V} \oplus V\}$ is a Lagrangian subspace of $\overline{V} \oplus V$. Define

$$c(g, h) = \tau(\Gamma_1, \Gamma_g, \Gamma_{gh})$$

for $g, h \in \mathrm{Sp}(V)$.

Lemma 2.3. *The function $c : G \times G \to W(F)$ is a 2-cocycle.*

Proof. The left-hand side of (6) is

$$\tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}) - \tau(\Gamma_1, \Gamma_g, \Gamma_{gg'g''}) + \tau(\Gamma_1, \Gamma_{gg'}, \Gamma_{gg'g''}) - \tau(\Gamma_1, \Gamma_{g'}, \Gamma_{g'g''}).$$

The last term is $-\tau(\Gamma_g, \Gamma_{gg'}, \Gamma_{gg'g''})$, applying A.5(d) to $1 \oplus g \in \mathrm{GL}(\overline{V} \oplus V)$. Thus the sum is a sum over the faces of the following tetrahedron, with each face contributing the Maslov index of its vertices, in the manner explained in Section A.5.2.

$$\begin{tikzcd} \Gamma_1 \arrow[r, dash] \arrow[d, dash] \arrow[dr, dash] & \Gamma_g \arrow[d, dash] \arrow[dl, dash, dotted] \\ \Gamma_{gg'} \arrow[r, dash] & \Gamma_{gg'g''}. \end{tikzcd}$$

The sum therefore vanishes. $\square$

From now on we reduce the values of c modulo I^3, where (as explained in Section A.1), $I \subset W(F)$ is the ideal of even-dimensional quadratic spaces.[4] Thus we obtain the following definition.

Definition 2.3.1. Let $M(V)$ be the central extension

$$0 \to W(F)/I^3 \to M(V) \to \mathrm{Sp}(V) \to 1 \tag{7}$$

defined by the cocycle c.

Remark 2.3.2. When F is a local field, $M(V)$ has a natural topology, as described in Remark 2.8.3 below.

2.3.3. *Reduction to I^2/I^3.* We now construct $\mathrm{Mp}(V)$ as a subgroup of $M(V)$, fitting into a central extension

$$0 \to I^2/I^3 \to \mathrm{Mp}(V) \to \mathrm{Sp}(V) \to 1. \tag{8}$$

When F is a finite or local field, $I^2/I^3 = Z_F$ (see Theorem A.2), yielding the central extension (3).

Definition 2.3.4. Let σ_g be the nondegenerate bilinear form on $(g-1)V$ defined[5] by

$$\sigma_g((g-1)x, (g-1)y) = \omega(x, (g-1)y) \quad \text{for all } x, y \in V.$$

Let $[\sigma_g]$ be the class in $W(F)/I^2$ generated by quadratic spaces with the same rank mod 2 and the same signed discriminant as σ_g; see Remark A.1.1. Let $\mathrm{Mp}(V) \subset M(V)$ be the subset of pairs (g, q) such that $q = [\sigma_g] \bmod I^2/I^3$.

We will have constructed a central extension (8) if we can prove this:

Proposition 2.4. $\mathrm{Mp}(V)$ *is a subgroup of* $M(V)$.

Proof. We use the calculation of the rank and discriminant of the Maslov index described in Section A.5.1. Write $\alpha_g = (1, g) : \Gamma_1 \to \Gamma_g$. Choose a nonzero $o_1 \in \det(\Gamma_1)$, and let $o_g = \alpha_g(o_1) \in \det(\Gamma_g)$. Let us calculate $Q(\Gamma_g, o_g; \Gamma_1, o_1)$, as defined in Section A.5.1. Using $\alpha = \alpha_g^{-1}$, this is the class in $W(F)/I^2$ of the bilinear form

$$q(x, gx; y, gy) = \omega(x, gy) - \omega(x, y) = \omega(x, (g-1)y)$$

[4] The reduction modulo I^3 is not crucial. We could deal with extensions of $\mathrm{Sp}(V)$ by $W(F)$ and I^2 rather than $W(F)/I^3$ and (as below) I^2/I^3. However, it is convenient that for finite and local fields, we can identify I^2/I^3 with the group Z_F (see Theorem A.2; in fact, $I^3 = 0$ for all finite or local fields other than $\mathbb{R}$). The reduction modulo I^3 is also necessary for Proposition 3.16.

[5] To see that σ_g is well defined, suppose that $(g-1)x = 0$. The claim is that $\omega(x, (g-1)y) = 0$. By direct calculation, $\omega(x, (g-1)y) = -\omega((g-1)x, gy) = -\omega(0, gy) = 0$. To see that σ_g is nondegenerate, observe that if, for some $(g-1)y$ and all $(g-1)x$, $\sigma_g((g-1)x, (g-1)y) = 0$, then $\omega(x, (g-1)y) = 0$ for all x, whence $(g-1)y = 0$ by the nondegeneracy of ω.

pairing (x, gx) and $(y, gy) \in \Gamma_g / \Gamma_g \cap \Gamma_1$. But $(x, gx) \mapsto (g-1)x$ is an isometry between $(\Gamma_g / \Gamma_g \cap \Gamma_1, q)$ and $((g-1)V, \sigma_g)$. Therefore

$$Q(\Gamma_g, o_g; \Gamma_1, o_1) = [\sigma_g] \in W(F)/I^2.$$

Now, according to (34) and the preceding discussion,

$$\begin{aligned}\tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}) &= Q(\Gamma_{gg'}, o_{gg'}; \Gamma_1, o_1) + Q(\Gamma_1, o_1; \Gamma_g, o_g) + Q(\Gamma_g, o_g; \Gamma_{gg'}, o_{gg'}) \\ &= Q(\Gamma_{gg'}, o_{gg'}; \Gamma_1, o_1) - Q(\Gamma_g, o_g; \Gamma_1, o_1) - Q(\Gamma_{g'}, o_{g'}; \Gamma_1, o_1)\end{aligned}$$

(all modulo I^2) and therefore, by our calculation,

$$\tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}) = [\sigma_{gg'}] - [\sigma_{g'}] - [\sigma_g] \mod I^2. \tag{9}$$

This is exactly the condition for $\mathrm{Mp}(V)$ to be closed under multiplication. □

2.4.1. *Uniqueness.* Before proceeding, note that in fact $\mathrm{Mp}(V)$ is the *unique* subgroup of $M(V)$ such that the projection to $\mathrm{Sp}(V)$ makes it a central extension of $\mathrm{Sp}(V)$ by I^2/I^3. Indeed, the following general statement applies.

Lemma 2.5. *Suppose that $\tilde{G}$ is a central extension of* $\mathrm{Sp}(V)$ *by an abelian group A. For any subgroup $B \subset A$ such that A/B has no* 3*-torsion, there is at most one subgroup $\tilde{G}' \subset \tilde{G}$ such that the given projection $\tilde{G}' \to$* $\mathrm{Sp}(V)$ *is surjective with kernel B.*

In our case, $A = W(F)/I^3$ and $B = I^2/I^3$; the lemma applies because $A/B = W(F)/I^2$ has only 2-primary torsion (being isomorphic to the group $W_0(F)$ described in Section A.1). In fact, $W(F)$ itself, and therefore any subquotient, has only 2-primary torsion; see [Lam 2005, Chapter 8, Theorem 3.2].

Proof of Lemma 2.5. Suppose that $\tilde{G}'$ and $\tilde{G}''$ are two such subgroups. Then for each $g \in \mathrm{Sp}(V)$ there exists $f(g) \in A$ such that $(g, a) \in \tilde{G}' \iff (g, a + f(g)) \in \tilde{G}''$. Moreover, $f(g)$ is unique modulo B, and f is a homomorphism $\mathrm{Sp}(V) \to A/B$. Thus it is enough to prove that there is no nontrivial homomorphism $\mathrm{Sp}(V) \to A/B$. In fact, $\mathrm{Sp}(V)$ is perfect unless $V \cong \mathbb{F}_3^2$; see [Grove 2001, Propositions 3.7–3.8]. In that exceptional case, the abelianization of $\mathrm{Sp}(V)$ is cyclic of order 3 (one can compute that $\mathrm{Sp}(V) \cong \mathrm{SL}_2(\mathbb{F}_3)$ has 24 elements, and that the commutator subgroup is the unique subgroup of order 8). Since, by assumption, A/B has no 3-torsion, any homomorphism $\mathrm{Sp}(V) \to A$ is trivial. □

Remark 2.5.1. The metaplectic extension $\mathrm{Mp}(V)$ of $\mathrm{Sp}(V)$ by I^2/I^3 also has a universal property, which can be deduced from the work of Suslin [1987]. Namely, the metaplectic extension of $\mathrm{Sp}_{2n}(F)$ is the universal central extension that extends to $\mathrm{SL}_{2n}(F)$ and splits over $\mathrm{SL}_n(F)$.

2.6. ***The traditional cocycle.*** Now we assume that F is finite or local, which allows us to use the Weil index γ_ψ (see Section A.3).

That is, for chosen Lagrangian subspace $\ell \subset V$ and nontrivial additive character $\psi : F \to \mathbb{C}^\times$, define

$$c_{\psi,\ell}(g, g') = \gamma_\psi(\tau(\ell, g\ell, gg'\ell)).$$

Then $c_{\psi,\ell}$ is a 2-cocycle with values in the group $Z_8 \subset \mathbb{C}^\times$ of eighth roots of unity (as can be proved in parallel to Lemma 2.3).

Definition 2.6.1. Define a central extension

$$1 \to Z_8 \to M^{\psi,\ell}(V) \to \mathrm{Sp}(V) \to 1 \tag{10}$$

using the cocycle $c_{\psi,\ell}$.

2.6.2. *Reduction to Z_F.* We now construct $\mathrm{Mp}^{\psi,\ell}(V)$ as a subgroup of $M^{\psi,\ell}(V)$, fitting into a central extension

$$1 \to Z_F \to \mathrm{Mp}^{\psi,\ell}(V) \to \mathrm{Sp}(V) \to 1. \tag{11}$$

We use the notation of Section A.5.1. Choose an orientation $o \in \det(\ell)$, and, for each $g \in \mathrm{Sp}(V)$, let go be the corresponding orientation of $g\ell$. The class $Q(g\ell, go; \ell, o) \in W(F)/I^2$ is independent of the choice of o.

Definition 2.6.3. Let $\mathrm{Mp}^{\psi,\ell}(V) \subset M^{\psi,\ell}(V)$ be the subset of pairs (g, ξ) with

$$\xi = \gamma_\psi(Q(g\ell, go; \ell, o)) \bmod Z_F.$$

(Recall that $Q(g\ell, go; \ell, o)$ is defined modulo I^2, and that $\gamma_\psi(I^2) = Z_F$; see Property A.4.1(d).)

It follows easily from (34) that $\mathrm{Mp}^{\psi,\ell}(V)$ is a subgroup of $M^{\psi,\ell}(V)$; indeed, by Lemma 2.5, it is the unique subgroup yielding a central extension of $\mathrm{Sp}(V)$ by Z_F.

Remark 2.6.4. The definition of $\mathrm{Mp}^{\psi,\ell}(V)$ can be unwound a bit to give a standard formula, as follows. For each $g \in \mathrm{Sp}(V)$, choose a basis $(q_1, \dots, q_n)$ of ℓ and a basis $(p_1, \dots, p_m, q_{m+1}, \dots, q_n)$ of $g\ell$, such that $(q_{m+1}, \dots, q_n)$ is a basis for $\ell \cap g\ell$ and $\omega(p_i, q_j) = \delta_{ij}$. Let $\theta^\ell(g) \in F^\times$ be the scalar such that

$$gq_q \wedge \cdots \wedge gq_n = \theta^\ell(g)(p_1 \wedge \cdots \wedge p_m \wedge q_{m+1} \wedge \cdots \wedge q_n)$$

in $\det(g\ell)$. The class of $\theta^\ell(g)$ in $F^\times/(F^\times)^2$ is independent of the bases. Then $\mathrm{Mp}^{\psi,\ell}(V) \subset M^{\psi,\ell}(V)$ is the subset of pairs (g, ξ) with

$$\xi = \gamma_\psi(1)^{\dim(\ell/\ell\cap g\ell)-1}\gamma_\psi(\theta^\ell(g)) \mod Z_F.$$

Indeed, this follows from Section A.4.1(c): $\dim(\ell/\ell \cap g\ell)$ and $\theta^\ell(g)$ are just the rank and discriminant of the quadratic form used to define $Q(g\ell, go; \ell, o)$ in Section A.5.1.

Remark 2.6.5. For a brief history of this construction of the metaplectic group and the related calculation of the cocycle of the Weil representation, see the bibliographical note in [Lion and Vergne 1980].

2.7. ***Intermediate cocycles.*** The transition from $\mathrm{Mp}(V)$ to $\mathrm{Mp}^{\psi,\ell}(V)$ involves two choices: that of the Lagrangian $\ell \subset V$, and that of the character ψ. To clarify the relationship between the different versions of the metaplectic group, we now examine these choices separately.

2.7.1. *Choice of Lagrangian.* The definitions follow the same pattern as before, and make sense for any F.

Definition. Let $M^\ell(V)$ be the central extension

$$0 \to W(F)/I^3 \to M^\ell(V) \to \mathrm{Sp}(V) \to 1 \tag{12}$$

defined by the cocycle

$$c_\ell(g,h) = \tau(\ell, g\ell, gh\ell).$$

Definition [Parimala et al. 2000]. Let $\mathrm{Mp}^\ell(V) \subset M^\ell(V)$ be the subset of pairs (g,q) such that $q = Q(g\ell, go; \ell, o) \bmod I^2$ (in the notation of Definition 2.6.3). In other words, q has rank $n := \dim(\ell/\ell \cap g\ell) \bmod 2$ and signed discriminant $(-1)^{n(n-1)/2}\theta^\ell(g)$ (in the notation of Remark 2.6.4).

With this definition, one can show that $\mathrm{Mp}^\ell(V)$ is a subgroup of $M^\ell(V)$, and, indeed, it is the unique (cf. Section 2.4.1) subgroup of $M^\ell(V)$ yielding a central extension

$$0 \to I^2/I^3 \to \mathrm{Mp}^\ell(V) \to \mathrm{Sp}(V) \to 1. \tag{13}$$

Remark 2.7.2. The following relationship is crucial to the proof of Theorem B. As in Section 2.2, let $\overline{V}$ be the symplectic vector space $(V, -\omega)$. Then the map $M(V) \to M^{\Gamma_1}(\overline{V} \oplus V)$ given by $(g,q) \mapsto (1 \oplus g, q)$ is a homomorphic embedding (and, by Section 2.4.1, it embeds $\mathrm{Mp}(V)$ into $\mathrm{Mp}^{\Gamma_1}(\overline{V} \oplus V)$). All of what we have said about $\mathrm{Mp}(V)$ can thereby be reduced to facts about $\mathrm{Mp}^{\Gamma_1}(\overline{V} \oplus V)$.

2.7.3. *Choice of an additive character.* Here we assume that F is finite or local.

Definition. Define a central extension

$$1 \to Z_8 \to M^\psi(V) \to \mathrm{Sp}(V) \to 1 \tag{14}$$

using the cocycle

$$c_\psi(g, g') = \gamma_\psi(\tau(\Gamma_1, \Gamma_g, \Gamma_{gg'})).$$

We again construct a subgroup $\mathrm{Mp}^\psi(V) \subset M^\psi(V)$ fitting into a central extension

$$1 \to Z_F \to \mathrm{Mp}^\psi(V) \to \mathrm{Sp}(V) \to 1 \tag{15}$$

and this subgroup is again unique, by Lemma 2.5.

Definition. Let $\mathrm{Mp}^{\psi}(V) \subset M^{\psi}(V)$ be the subgroup consisting of pairs (g, ξ) with $\xi = \gamma_{\psi}([\sigma_g]) \bmod Z_F$. Equivalently (using Section A.4.1(c)), the requirement is that

$$\xi = \gamma_{\psi}(1)^{\dim(g-1)V-1}\gamma_{\psi}(\det\sigma_g) \mod Z_F.$$

2.8. *Remarks.*

Remark 2.8.1. Given the existence of a unique isomorphism $I^2/I^3 \to Z_F$ (Theorem A.2) when F is finite or local, the introduction of a character ψ may seem entirely extraneous to the construction of the metaplectic group. Indeed, its use is motivated by the Weil representation, which may be considered as a representation of $M^{\psi}(V)$ (or $M^{\psi,\ell}(V)$) in which the central factor Z_8 acts by scalar multiplication.

Remark 2.8.2. Let us explain the relationship between the present constructions and the version of the metaplectic group used in [Thomas 2008] (which considered only finite and local fields). Let $\mathrm{Gr}(V)$ be the set of all Lagrangian subspaces $\ell \subset V$. As we explain in the next section, there is a canonical isomorphism $\delta^{\psi}_{\ell\ell'} : \mathrm{Mp}^{\psi,\ell}(V) \to \mathrm{Mp}^{\psi,\ell'}(V)$ for every pair $\ell, \ell' \in \mathrm{Gr}(V)$. Then

$$G = \left\{ (g_\ell) \in \prod_{\ell\in\mathrm{Gr}(V)} \mathrm{Mp}^{\psi,\ell}(V) \;\middle|\; \delta^{\psi}_{\ell\ell'}(g_\ell) = g_{\ell'} \text{ for all } \ell, \ell' \in \mathrm{Gr}(V) \right\}$$

is a group under component-wise multiplication, with the obvious projections making G isomorphic to each $\mathrm{Mp}^{\psi,\ell}(V)$. This G is essentially what was called $\mathrm{Mp}(V)$ in [Thomas 2008, Definition 5.2]. By construction, it does not depend on any *particular* choice of $\ell \in \mathrm{Gr}(V)$; one could, of course, remove the apparent dependence on ψ by a similar trick.

Remark 2.8.3. Suppose that F is a local field. It is well-known that $\mathrm{Mp}^{\psi,\ell}(V)$ is naturally a topological covering group of $\mathrm{Sp}(V)$ — the topology is the one that makes the Weil representation continuous. Since, as explained in the next section, $\mathrm{Mp}(V)$ and $\mathrm{Mp}^{\psi,\ell}(V)$ are canonically isomorphic, this defines a topology on $\mathrm{Mp}(V)$, which can be extended in a unique way to $M(V)$, making $M(V)$ a covering group of $\mathrm{Sp}(V)$ as well. It is interesting to describe this topology more explicitly, by giving an open neighborhood U of the identity $(1, 0) \in M(V)$ that maps homeomorphically onto its image in $\mathrm{Sp}(V)$. It turns out we can take

$$U = \{(g, q) \in M(V) \mid \ker(g+1) = 0,\ q = -Q_g \bmod I^3\}.$$

For example, this means that the formula in Theorem D is continuous at $(1, 0)$. For an analogous description of the topology of $\mathrm{Mp}^{\psi,\ell}(V)$, see [Thomas 2008, Proposition 5.3].

3. Isomorphisms between metaplectic groups

In this section, we describe isomorphisms between the different versions of the metaplectic group that were introduced in Section 2. First we consider the choice of Lagrangian, describing canonical (see Section 3.1) isomorphisms that fit into a commutative diagram (omitting V from the notation):

$$
\begin{tikzcd}
 & M^{\ell} \arrow[r, dotted, "\alpha^{\ell}_{\psi}"] \arrow[d, "\delta_{\ell\ell'}"] & M^{\psi,\ell} \arrow[d, "\delta^{\psi}_{\ell\ell'}"'] & \\
M \arrow[ur, "\alpha_{\ell}"] \arrow[dr, "\alpha_{\ell'}"] \arrow[rrr, dotted, bend right=50, "\alpha_{\psi}"'] & & & M^{\psi}. \arrow[ul, "\alpha^{\psi}_{\ell}"'] \arrow[dl, "\alpha^{\psi}_{\ell'}"'] \\
 & M^{\ell'} \arrow[r, dotted, "\alpha^{\ell'}_{\psi}"] & M^{\psi,\ell'} &
\end{tikzcd}
$$

(The dotted arrows are homomorphisms, not isomorphisms, but all the maps shown restrict to isomorphisms between the various groups $\mathrm{Mp}^{\bullet}(V)$.) Next we consider the choice of additive character, describing a commutative diagram of canonical isomorphisms:

$$
\begin{tikzcd}
 & M^{\psi} \arrow[r, "\alpha^{\psi}_{\ell}"] \arrow[d, "\delta_{\psi\psi'}"] & M^{\psi,\ell} \arrow[d, "\delta^{\ell}_{\psi\psi'}"'] & \\
M \arrow[ur, dotted, "\alpha_{\psi}"] \arrow[dr, dotted, "\alpha_{\psi'}"] \arrow[rrr, bend right=50, "\alpha_{\ell}"'] & & & M^{\ell} \arrow[ul, dotted, "\alpha^{\ell}_{\psi}"'] \arrow[dl, dotted, "\alpha^{\ell}_{\psi'}"'] \\
 & M^{\psi'} \arrow[r, "\alpha^{\psi'}_{\ell}"] & M^{\psi',\ell}. &
\end{tikzcd}
$$

Finally, we describe canonical actions of $\mathrm{GSp}(V)$ on $M(V)$ and $M^{\psi,\ell}(V)$ that cover the action by conjugation on $\mathrm{Sp}(V)$.

As in Section 2, objects labeled by the character ψ are defined only when F is a finite or local field; objects that do not involve ψ make sense more generally.

3.1. In the above overview, we used the word "canonical" to mean "unique" in the following sense. If $\tilde{G}$ and $\tilde{G}'$ are central extensions of a group G by an abelian group A, then "an isomorphism of central extensions" is an isomorphism $\tilde{G} \to \tilde{G}'$ which covers the identity $G \to G$ and restricts to the identity $A \to A$. The claim is that all the isomorphisms are unique as isomorphisms of central extensions. This uniqueness is guaranteed by the following lemma.

Lemma 3.2. *Let $\tilde{G}$ and $\tilde{G}'$ be central extensions of* $\mathrm{Sp}(V)$ *by an abelian group A with no* 3*-torsion. Then there exists at most one isomorphism $\tilde{G} \to \tilde{G}'$ of central extensions.*

Proof. If $f_1, f_2 : \tilde{G} \to \tilde{G}'$ are isomorphisms of central extensions, then

$$(g, a) \mapsto f_1(g, a) \cdot f_2(g, a)^{-1}$$

is given by a homomorphism $\mathrm{Sp}(V) = \tilde{G}/A \to A \subset \tilde{G}'$. But, as explained in the proof of Lemma 2.5, any such homomorphism is trivial. □

As we noted after Lemma 2.5, the Witt group $W(F)$ has only 2-primary torsion, so Lemma 3.2 applies to all the central extensions of interest.

3.2.1. *Coboundary description.* We will repeatedly use the following basic observation. If $\tilde{G}$ and $\tilde{G}'$ are defined by 2-cocycles c and c', then an isomorphism $f : \tilde{G} \to \tilde{G}'$ of central extensions is equivalent to giving a function $s : G \to A$ such that

$$c'(g, g') - c(g, g') = s(gg') - s(g) - s(g').$$

(This expresses $c' - c$ as the coboundary of s.) Namely, $f(g, a) = (g, a + s(g))$.

3.3. *Choice of Lagrangian.*

Proposition 3.4. *There is a unique isomorphism* $\alpha_\ell : M(V) \to M^\ell(V)$ *of central extensions, and it is given by*

$$\alpha_\ell(g, q) = (g, q + \tau(\ell \oplus \ell, \Gamma_1, \Gamma_g, \ell \oplus g\ell)). \tag{16}$$

It restricts to an isomorphism $\mathrm{Mp}(V) \to \mathrm{Mp}^\ell(V)$, *also unique.*

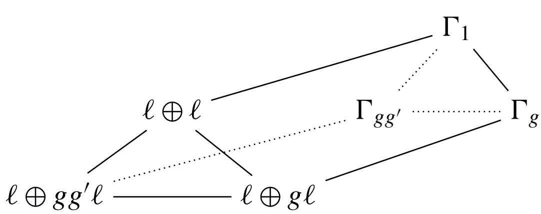
Proof. For α_ℓ to be an isomorphism, it suffices, by Section 3.2.1, to check

$$c_\ell(g, g') - c(g, g') + s(g) + s(g') - s(gg') = 0 \tag{17}$$

where $s(g) := \tau(\ell \oplus \ell, \Gamma_1, \Gamma_g, \ell \oplus g\ell)$. Observe that $\tau(\ell, \ell, \ell) = 0$: according to Section A.5(e), it is represented by the zero bilinear form on ℓ. Therefore

$$c_\ell(g, g') = \tau(\ell, g\ell, gg'\ell) = \tau(\ell \oplus g\ell, \ell \oplus gg'\ell, \ell \oplus \ell)$$

by A.5(c). Moreover, $s(g') = \tau(\ell \oplus g\ell, \Gamma_g, \Gamma_{gg'}, \ell \oplus gg'\ell)$ by A.5(d) applied to $(1, g) \in \mathrm{GL}(\overline{V} \oplus V)$. Graphically, then, (17) is a sum over the faces of the polyhedron



and therefore vanishes, as explained in Section A.5.2.

The fact that α_ℓ maps $\mathrm{Mp}(V)$ to $\mathrm{Mp}^\ell(V)$ follows from the uniqueness property of $\mathrm{Mp}^\ell(V)$ (Section 2.4.1), or by direct computation, using (34); the uniqueness of α_ℓ follows from Lemma 3.2. □

Corollary 3.5. *There is a unique isomorphism* $\alpha_\ell^\psi : M^\psi(V) \to M^{\psi,\ell}(V)$ *of central extensions, and it is given by*

$$\alpha_\ell^\psi(g,\xi) = (g, \xi \cdot \gamma_\psi(\tau(\ell \oplus \ell, \Gamma_1, \Gamma_g, \ell \oplus g\ell))). \tag{18}$$

It restricts to an isomorphism $\mathrm{Mp}^\psi(V) \to \mathrm{Mp}^{\psi,\ell}(V)$, *also unique.*

3.6. ***Change of Lagrangian.***

Proposition 3.7. *There is a unique isomorphism* $\delta_{\ell\ell'} : M^\ell(V) \to M^{\ell'}(V)$ *of central extensions, given by*

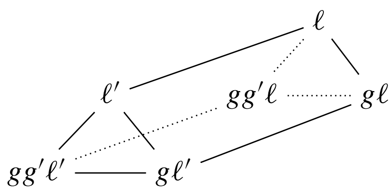
$$\delta_{\ell\ell'}(g,q) = (g, q + \tau(\ell, g\ell, g\ell', \ell')).$$

It restricts to an isomorphism $\mathrm{Mp}^\ell(V) \to \mathrm{Mp}^{\ell'}(V)$, *also unique.*

Proof. The proof is very similar to that of Proposition 3.4. The main difference is that we must now show

$$c_{\ell'}(g,g') - c_\ell(g,g') + s(g) + s(g') - s(gg') = 0 \tag{19}$$

where now $s(g) := \tau(\ell, g\ell, g\ell', \ell')$. Observe that $s(g') = \tau(g\ell, gg'\ell, gg'\ell', g\ell')$ by Section A.5(d). Thus (19) is a sum over the faces of the polyhedron



and again vanishes by Section A.5.2. $\square$

Corollary 3.8. *There is a unique isomorphism* $\delta_{\ell\ell'}^\psi : M^{\psi,\ell}(V) \to M^{\psi,\ell'}(V)$ *of central extensions, given by*

$$\delta_{\ell\ell'}^\psi(g,\xi) = (g, \xi \cdot \gamma_\psi(\tau(\ell, g\ell, g\ell', \ell'))).$$

It restricts to an isomorphism $\mathrm{Mp}^{\psi,\ell}(V) \to \mathrm{Mp}^{\psi,\ell'}(V)$, *also unique.*

3.9. ***Choice of additive character.*** There are obvious homomorphisms

$$\alpha_\psi : M(V) \to M^\psi(V), \quad \alpha_\psi^\ell : M^\ell(V) \to M^{\psi,\ell}(V),$$

each given by $(g,q) \mapsto (g, \gamma_\psi(q))$.

Proposition 3.10. *The maps* α_ψ, α_ψ^ℓ *are the unique homomorphisms that cover the identity on* $\mathrm{Sp}(V)$ *and restrict to* $\gamma_\psi : W(F)/I^3 \to Z_8$. *Moreover, they restrict to isomorphisms*

$$\alpha_\psi : \mathrm{Mp}(V) \to \mathrm{Mp}^\psi(V), \quad \alpha_\psi^\ell : \mathrm{Mp}^\ell(V) \to \mathrm{Mp}^{\psi,\ell}(V),$$

that are unique as isomorphisms of central extensions.

Proof. Uniqueness is a simple variation on Lemma 3.2. The fact that $\mathrm{Mp}(V)$ and $\mathrm{Mp}^{\ell}(V)$ map to $\mathrm{Mp}^{\psi}(V)$ and $\mathrm{Mp}^{\psi,\ell}(V)$ is immediate from the definitions. The fact that the restricted maps are isomorphisms follows from the fact that

$$\gamma_{\psi} : I^2/I^3 \to Z_F$$

is an isomorphism (Section A.4.1(d)). □

3.11. *Change of additive character.* Suppose that ψ, ψ' are nontrivial additive characters of F. Let $a \in F^{\times}$ be the unique scalar such that $\psi'(x) = \psi(ax)$ for all $x \in F$. In the next proposition, $(\cdot,\cdot)_H : F^{\times} \otimes_{\mathbb{Z}} F^{\times} \to Z_F$ is the Hilbert symbol (defined in Section A.1.2).

Proposition 3.12. *There is a unique isomorphism $\delta_{\psi\psi'} : M^{\psi}(V) \to M^{\psi'}(V)$ of central extensions, and it is given by*

$$\delta_{\psi\psi'}(g,\xi) = (g, r_a(g)\xi)$$

where $r_a(g) := (\gamma_{\psi}(a)/\gamma_{\psi}(1))^{\dim(g-1)V}\,(a,\det\sigma_g)_H$. It restricts to an isomorphism $\mathrm{Mp}^{\psi}(V) \to \mathrm{Mp}^{\psi'}(V)$, also unique.

To prove Proposition 3.12, we first study the dependence of γ_{ψ} on ψ.

Lemma 3.13. *For any quadratic space (A,q),*

$$\gamma_{\psi'}(q) = \gamma_{\psi}(q)\,(\gamma_{\psi}(a)/\gamma_{\psi}(1))^{\dim A}(a,\det q)_H.$$

Proof. Both sides of the equation define homomorphisms $W(F) \to \mathbb{C}^{\times}$. Since any quadratic space is the perpendicular sum of one-dimensional ones, we can reduce to the case where $A = F$ and $q(x,y) = bxy$. Then $\gamma_{\psi'}(q) = \gamma_{\psi}(ab)$ and the statement amounts to the standard formula Section A.4.1(b). □

Proof of Proposition 3.12. To get an isomorphism, by Section 3.2.1 we must check

$$\gamma_{\psi'}(\tau(\Gamma_1,\Gamma_g,\Gamma_{gg'})) = \gamma_{\psi}(\tau(\Gamma_1,\Gamma_g,\Gamma_{gg'})) \cdot \frac{r_a(gg')}{r_a(g)r_a(g')}.$$

The right-hand side simplifies to

$$\gamma_{\psi}(\tau(\Gamma_1,\Gamma_g,\Gamma_{gg'})) \cdot (\gamma_{\psi}(a)/\gamma_{\psi}(1))^d (a,\delta)_H,$$

where

$$d = \dim(gg'-1)V - \dim(g-1)V - \dim(g'-1)V \quad \text{and} \quad \delta = \det\sigma_{gg'}/(\det\sigma_g \det\sigma_{g'}).$$

Comparing this to Lemma 3.13, we are reduced to checking that $\tau(\Gamma_1,\Gamma_g,\Gamma_{gg'})$ has rank d mod 2 and signed discriminant $(-1)^{d(d-1)/2}\delta$. This is equivalent to (9).

We therefore have an isomorphism; uniqueness follows from Lemma 3.2, and the fact that $\mathrm{Mp}^{\psi}(V)$ maps to $\mathrm{Mp}^{\psi'}(V)$ follows from Lemma 2.5. □

Here is the analogue of Proposition 3.12 for $M^{\psi,\ell}(V)$.

Proposition 3.14. *There is a unique isomorphism* $\delta^{\ell}_{\psi\psi'} : M^{\psi,\ell}(V) \to M^{\psi',\ell}(V)$ *of central extensions, and it is given by*

$$\delta^{\ell}_{\psi\psi'}(g,\xi) = (g, r^{\ell}_a(g)\xi)$$

where $r^{\ell}_a(g) := (\gamma_\psi(a)/\gamma_\psi(1))^{\dim(\ell/\ell\cap g\ell)}\,(a, \theta^{\ell}(g))_H$. *It restricts to an isomorphism* $\mathrm{Mp}^{\psi,\ell}(V) \to \mathrm{Mp}^{\psi',\ell}(V)$, *also unique.*

3.15. ***Outer automorphisms.*** We return to the general setting where F is any field of characteristic not 2. Let $\mathrm{GSp}(V) \subset \mathrm{GL}(V)$ be the group of symplectic similitudes, that is, linear transformations $f \in \mathrm{GL}(V)$ such that there exists $\lambda(f) \in F^\times$ satisfying $\omega(fx, fy) = \lambda(f)\omega(x, y)$ for all $x, y \in V$. Then $\mathrm{GSp}(V)$ contains $\mathrm{Sp}(V)$ as a normal subgroup, and so acts on it by conjugation. (In fact, according to [Hua 1948], any automorphism of $\mathrm{Sp}(V)$ can be written as a composition $\varphi \circ \mathrm{Ad}\, f$ with $f \in \mathrm{GSp}(V)$ and φ a field automorphism of F.)

The goal of this section is to describe explicitly an action of $\mathrm{GSp}(V)$ on the metaplectic group, lifting the conjugation action on $\mathrm{Sp}(V)$. This lifting is unique.

First let us define a function

$$\mathrm{Sp}(V) \times F^\times \to W(F)/I^3.$$

Given $(g, a) \in \mathrm{Sp}(V) \times F^\times$, let $b_g \in W(F)$ be represented by a quadratic space of rank $\dim(g-1)V$ and discriminant $\det \sigma_g$ (thus $b_g = [\sigma_g]$ modulo I^2). Now let $q_{g,a} = (q_a - 1) \otimes b_g$. The class of $q_{g,a}$ in $W(F)/I^3$ is independent of choices.

Proposition 3.16. *For any* $f \in \mathrm{GSp}(V)$ *there is a unique automorphism* N_f *of* $M(V)$ *covering* $\mathrm{Ad}\, f$ *and restricting to the identity on* $W(F)/I^3$. *It is given by* $N_f(g, q) = (\mathrm{Ad}\, f(g), q + q_{g,\lambda(f)})$.

Proof. Simple variations on Lemma 3.2 and Section 3.2.1 show that N_f will be a unique isomorphism so long as

$$\tau(\Gamma_1, \Gamma_{\mathrm{Ad}\, f(g)}, \Gamma_{\mathrm{Ad}\, f(gg')}) - \tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}) = q_{gg',\lambda(f)} - q_{g,\lambda(f)} - q_{g',\lambda(f)} \quad (20)$$

modulo I^3. Now,

$$\Gamma_{\mathrm{Ad}\, f(g)} = \{(v, fgf^{-1}v)\} = \{(fv, fgv)\} = (f, f)\cdot\Gamma_g \subset \overline{V} \oplus V.$$

This and Section A.5(d) imply that

$$\tau(\Gamma_1, \Gamma_{\mathrm{Ad}\, f(g)}, \Gamma_{\mathrm{Ad}\, f(gg')}) = q_{\lambda(f)} \otimes \tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}). \quad (21)$$

Thus the left-hand side of (20) is $(q_{\lambda(f)} - 1) \otimes \tau(\Gamma_1, \Gamma_g, \Gamma_{gg'})$. By the definition of $q_{g,\lambda(f)}$, to establish (20), it suffices to show that

$$\tau(\Gamma_1, \Gamma_g, \Gamma_{gg'}) = b_{gg'} - b_g - b_{g'} \quad \mod I^2.$$

But this is equivalent to (9). □

Remark 3.16.1. Proposition 3.16 is stated for $M(V)$, but the uniqueness of $\mathrm{Mp}(V)$ (Section 2.4.1) implies that N_f restricts to an automorphism of that subgroup, which is again the unique automorphism covering $\mathrm{Ad}\, f$.

A description of the automorphisms of $M^{\ell}(V)$, $M^{\psi}(V)$, and $M^{\psi,\ell}(V)$ covering the action of $\mathrm{GSp}(V)$ is easily deduced in parallel to Proposition 3.16, using the isomorphisms of Sections 3.6–3.11. For example, we have:

Proposition 3.17. *For any $f \in \mathrm{GSp}(V)$ there is a unique automorphism $N_f^{\psi,\ell}$ of $M^{\psi,\ell}(V)$ covering $\mathrm{Ad}\, f$ and restricting to the identity on Z_8. It is given by*

$$N_f^{\psi,\ell}(g,\xi) = (\mathrm{Ad}\, f(g), \gamma_\psi(\tau(\ell, g\ell, gf^{-1}\ell, f^{-1}\ell)) \cdot r_{\lambda(f)}^{f^{-1}\ell}(g) \cdot \xi).$$

Proof. Put $a := \lambda(f)$, $\psi'(x) = \psi(ax)$, and $\ell' = f^{-1}\ell$. By Section A.5(d), we have

$$\begin{aligned} c_{\psi,\ell}(\mathrm{Ad}\, f(g), \mathrm{Ad}\, f(g')) &= \gamma_\psi(\tau(\ell, fgf^{-1}\ell, fgg'f^{-1}\ell) \\ &= \gamma_{\psi'}(\tau(\ell', g\ell', gg'\ell')) = c_{\psi',\ell'}(g, g'). \end{aligned}$$

It follows that $s : (g, \xi) \mapsto (\mathrm{Ad}\, f(g), \xi)$ is an isomorphism $M^{\psi',\ell'}(V) \to M^{\psi,\ell}(V)$ and thence that $N_f^{\psi,\ell}(g,\xi) = s \circ \delta_{\psi\psi'}^{\ell'} \circ \delta_{\ell\ell'}^{\psi}$ is an automorphism of $M^{\psi,\ell}(V)$ of the required kind. □

Remark 3.17.1. Proposition 3.17 is related to Proposition 3.16 in the sense that we must have $N_f^{\psi,\ell} \circ \alpha_\psi^\ell \circ \alpha_\ell = \alpha_\psi^\ell \circ \alpha_\ell \circ N_f$. (One can even use this to deduce Proposition 3.17 from Proposition 3.16, but the proof we have presented is much easier, given what we already know.)

4. Heisenberg group and Weil representation

Henceforth F is a finite or local field with characteristic not 2.

In this section we recall the definition and basic properties of the Weil representation $\rho^{\psi,\ell}$. A more detailed exposition can be found in [Lion and Vergne 1980, §1.2–1.4 and Appendix].

4.1. ***Hilbert spaces and norms.*** In describing representations, we use natural Hilbert spaces of half-densities, with the notation laid out in Section A.3.1. Thus if X is a finite-dimensional vector space over F then $L^2(X)$ denotes the space of L^2 functions $X \to \Omega_{1/2}(X)$.

4.2. ***The Heisenberg group.*** The Heisenberg group $H(V)$ based on V is, as a set, the direct product $H(V) = V \times F$, equipped with the multiplication

$$(v,s)(w,t) = \big(v + w,\, s + t + \tfrac{1}{2}\omega(v,w)\big).$$

The center of $H(V)$ is the factor F. We are interested in representations of $H(V)$ with fixed central character ψ (so again ψ is a continuous homomorphism

$F \to U(1)$). To avoid always writing the action of the center, note that such a representation ρ is determined by the family of operators $\{\rho(v)\}_{v\in V}$, which satisfy

$$\rho(v)\rho(w) = \psi\big(\tfrac{1}{2}\omega(v,w)\big)\cdot\rho(v+w).$$

Theorem 4.3 (Stone and von Neumann). *$H(V)$ has, for each nontrivial central character ψ, a unique isomorphism class of unitary, continuous, irreducible representations.* (The notion of continuity is that of the strong operator topology.)

The proof over $\mathbb{R}$ can be found in [Lion and Vergne 1980, §1.3], and a general exposition is in [Prasad 2011]. The main step is Proposition 5.2(a) below.

4.4. ***Formulas for its representation.*** For chosen $\ell \in \mathrm{Lagr}(V)$, the representation from Theorem 4.3 is realized by

$$\rho_H^{\psi,\ell} := \mathrm{Ind}_{\ell\times F}^{H}(\tilde{\psi})$$

where $\tilde{\psi}$ is the composition $\ell\times F \to F \xrightarrow{\psi} \mathbb{C}^\times$. One has the following explicit description of the corresponding Hilbert space $\mathcal{H}^{\psi,\ell}$. It is the completion of the space of smooth functions $\phi : V \to \Omega_{1/2}(V/\ell)$ that satisfy

$$\phi(v+w) = \phi(v)\,\psi\big(\tfrac{1}{2}\omega(v,w)\big) \quad \text{for all } w\in\ell \tag{22}$$

and that are finite under the norm

$$|\phi|^2 := \int_{v\in V/\ell} \overline{\phi(v)}\,\phi(v).$$

The action of $H(V)$ on $\mathcal{H}^{\psi,\ell}$ is given, for $\phi\in\mathcal{H}^{\psi,\ell}$ and $v\in V$, by

$$\rho_H^{\psi,\ell}(v)\phi(x) = \phi(x-v)\,\psi\big(\tfrac{1}{2}\omega(v,x)\big). \tag{23}$$

4.4.1. *Transverse Lagrangians.* For any Lagrangian ℓ' transverse to ℓ, the isomorphism $V/\ell \to \ell'$ yields an isometry

$$\mathrm{Res}_{\ell'} : \mathcal{H}^{\psi,\ell} \to L^2(\ell').$$

The action of $H(V)$ on $L^2(\ell')$ is described by the formula

$$\big(\mathrm{Res}_{\ell'}\circ\rho_H^{\psi,\ell}(v+v')\circ\mathrm{Res}_{\ell'}^{-1}\big)(\phi)(x') = \phi(x'-v')\cdot\psi\big(\omega\big(v, x'-\tfrac{1}{2}v'\big)\big) \tag{24}$$

for all $v\in\ell$ and $v', x'\in\ell'$.

4.5. ***The Weil representation.*** Since $\mathrm{Sp}(V)$ is the group of automorphisms of $H(V)$ preserving the center, one obtains a *projective* representation $\rho_{\mathrm{Sp}}^{\psi,\ell}$ of $\mathrm{Sp}(V)$ acting on $\mathcal{H}^{\psi,\ell}$, characterized by

$$\rho_{\mathrm{Sp}}^{\psi,\ell}(g)\circ\rho_H^{\psi,\ell}(v)\circ\rho_{\mathrm{Sp}}^{\psi,\ell}(g)^{-1} = \rho_H^{\psi,\ell}(gv).$$

In detail, $(\rho_H^{\psi,\ell})^g : v \mapsto \rho_H^{\psi,\ell}(gv)$ defines a representation of $H(V)$ on $\mathcal{H}^{\psi,\ell}$ with central character ψ. By Theorem 4.3, there is a unique-up-to-scale operator $\rho_{\mathrm{Sp}}^{\psi,\ell}(g)$ on $\mathcal{H}^{\psi,\ell}$ intertwining from $\rho_H^{\psi,\ell}$ to $(\rho_H^{\psi,\ell})^g$.

The next result is due to Lion and Perrin.

Theorem 4.6 [Perrin 1981]. *There is a true representation $\rho_{\mathrm{Mp}}^{\psi,\ell}$ of $\mathrm{Mp}^{\psi,\ell}(V)$, uniquely characterized by the formulas*

$$\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)\circ\rho_H^{\psi,\ell}(v)\circ\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)^{-1}=\rho_H^{\psi,\ell}(gv),\quad \rho_{\mathrm{Mp}}^{\psi,\ell}(1,\xi)=\xi\cdot\mathrm{id}.$$

The operators $\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi):\mathcal{H}^{\psi,\ell}\to\mathcal{H}^{\psi,\ell}$ are given on Schwartz functions ϕ by

$$\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)\phi(x):=\xi\cdot\int_{y\in(g^{-1}\ell)/(\ell\cap g^{-1}\ell)}\phi(g^{-1}x+y)\,\psi\big(\tfrac{1}{2}\omega(y,g^{-1}x)\big)\,\mu_g^{\psi,\ell}$$

where $\mu_g^{\psi,\ell}\in\Omega_1\big((g^{-1}\ell)/(\ell\cap g^{-1}\ell)\big)$ is the unique invariant measure such that $\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)$ is unitary.

Remark 4.6.1. More concretely, $\mu_g^{\psi,\ell}$ is characterized by the following property. First, $g^{-1}\ell/(\ell\cap g^{-1}\ell)$ and $\ell/(\ell\cap g^{-1}\ell)$ are Pontryagin-dual abelian groups under the pairing $\psi\circ\omega$. Let μ be the measure on $\ell/(\ell\cap g^{-1}\ell)$ dual to $\mu_g^{\psi,\ell}$. Choose a measure μ_0 on $\ell\cap g^{-1}\ell$. Then $\mu_g^{\psi,\ell}\otimes\mu_0$ and $\mu\otimes\mu_0$ are measures on $g^{-1}\ell$ and ℓ, respectively. The property is that these measures correspond under the isomorphism $g:g^{-1}\ell\to\ell$.

4.7. ***Definition.*** Let $\rho^{\psi,\ell}$ be the representation of $\mathrm{Mp}^{\psi,\ell}(V)\ltimes H(V)$ defined by

$$\rho^{\psi,\ell}(g,\xi;v,t)=\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)\circ\rho_H^{\psi,\ell}(v,t).$$

We also use $\rho^{\psi,\ell}$ to denote the corresponding representation of $\mathrm{Mp}(V)\ltimes H(V)$, defined using the canonical isomorphism $\alpha_\psi^\ell\circ\alpha_\ell=\alpha_\ell^\psi\circ\alpha_\psi:\mathrm{Mp}(V)\to\mathrm{Mp}^{\psi,\ell}(V)$. Thus for $q\in W(F)/I^3$,

$$\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q):=\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi),\quad \text{with } \xi:=\gamma_\psi(q+\tau(\ell\oplus\ell,\Gamma_1,\Gamma_g,\ell\oplus g\ell))\in Z_8.$$

5. The character: Proof of Theorem B

The goal of this section is to prove Theorem B. There are two main ideas involved: first, the Weyl transform, developed in Section 5.1, and second, the homomorphism $\mathrm{Sp}(V)\to\mathrm{Sp}(\overline{V}\oplus V)$, studied in Section 5.4. We conclude the proof of Theorem B in Section 5.6.

5.1. ***Weyl transform.*** Let $\mathcal{S}(V) \subset L^2(V)$ be the subspace of Schwartz-class half-densities.[6] Let $\operatorname{End}_0 \mathcal{H}^{\psi,\ell} \cong \mathcal{S}(V/\ell \times V/\ell)$ be the algebra of operators on the Hilbert space $\mathcal{H}^{\psi,\ell} \cong L^2(V/\ell)$ that can be represented by Schwartz-class integral kernels. It is dense in the algebra $\operatorname{End}\mathcal{H}^{\psi,\ell} \cong L^2(V/\ell \times V/\ell)$ of Hilbert–Schmidt operators (that is, those with L^2 integral kernels).

The following proposition is well-known (it is the heart of the Stone–von Neumann Theorem, 4.3). As usual, μ_V denotes the measure on V self-dual with respect to $\psi \circ \omega$.

Proposition 5.2. *For $h \in \mathcal{S}(V)$, let $W^{\psi,\ell}(h)$ be the operator on $\mathcal{H}^{\psi,\ell}$ defined by*

$$W^{\psi,\ell}(h)(\phi)(x) = \int_{v \in V} \rho_H^{\psi,\ell}(v)\phi(x) \cdot h(v)\, \mu_V^{1/2}. \tag{25}$$

(a) *$W^{\psi,\ell}$ is an isomorphism $\mathcal{S}(V) \to \operatorname{End}_0(\mathcal{H}^{\psi,\ell})$ and extends to an isometry*

$$W^{\psi,\ell} : L^2(V) \to \operatorname{End}(\mathcal{H}^{\psi,\ell}).$$

(b) *If we equip $L^2(V)$ with the multiplication*

$$(f_1 \star f_2)(x) := \int_{v \in V} f_1(v)\, \psi\big(\tfrac{1}{2}\omega(v,x)\big)\, f_2(x-v)\, \mu_V^{1/2}$$

then $W^{\psi,\ell}$ becomes an algebra isomorphism $W^{\psi,\ell} : L^2(V) \to \operatorname{End}(\mathcal{H}^{\psi,\ell})$.

(c) *For $h \in \mathcal{S}(V)$, the operator $W^{\psi,\ell}(h)$ is trace class, and*

$$\operatorname{Tr} W^{\psi,\ell}(h) \cdot \mu_V^{1/2} = h(0).$$

Proof. Choose ℓ' transverse to ℓ, and identify $L^2(V) = L^2(\ell \times \ell')$. Let $\mathfrak{F}_0$ be the Fourier transform $L^2(\ell) \to L^2(\ell')$ with respect to the pairing $\psi \circ \frac{1}{2}\omega$:

$$\mathfrak{F}_0 f(a') := \|2\|^{-\frac{\dim V}{4}} \int_{a \in \ell} f(a)\psi\big(\tfrac{1}{2}\omega(a,a')\big)\, \mu_V^{1/2}.$$

(There is a canonical isomorphism $\Omega_{1/2}(\ell) \otimes \Omega_{1/2}(V) = \Omega_1(\ell) \otimes \Omega_{1/2}(\ell')$ which allows us to interpret $\mathfrak{F}_0$ as a map from half-densities on ℓ to half-densities on ℓ'.) Let $A \in \mathrm{GL}(\ell' \times \ell')$ be the isomorphism $A(a',x') = (x'+a', x'-a')$. Write A^* for the corresponding isometry $f \mapsto \|2\|^{(\dim V)/4}\,(f \circ A)$ of $L^2(\ell \times \ell')$.

Lemma 5.3. *$W^{\psi,\ell}$ factors as a composition of isometries*

$$L^2(V) = L^2(\ell \times \ell') \xrightarrow{\mathfrak{F}_0 \otimes \mathrm{id}} L^2(\ell' \times \ell') \xrightarrow{A^*} L^2(\ell' \times \ell') = \operatorname{End}(\mathcal{H}^{\psi,\ell}).$$

[6] Our exposition here differs slightly from the sketch in Section 1.3(C) in that we use half-densities rather than complex-valued functions; the square root $\mu_V^{1/2}$ of the self-dual measure for $\psi \circ \omega$ can be used to pass between the two.

Proof. By (25) and (24), we have

$$W^{\psi,\ell}(h)\phi(x') = \int_{(a,a')\in V} \phi(x'-a')\cdot\psi\big(\omega\big(a, x'-\tfrac{1}{2}a'\big)\big)\cdot h(a,a')\,\mu_V^{1/2}$$
$$= \int_{a'\in\ell'} \phi(a') \int_{a\in\ell} \psi\big(\tfrac{1}{2}\omega(a, x'+a')\big)\cdot h(a, x'-a')\,\mu_V^{1/2} \tag{26}$$

with a change of variables $a'\mapsto x'-a'$; this is exactly what the lemma claims. □

Part (a) of the proposition follows from the fact that Fourier transforms preserve the Schwartz class. In part (b), the $\star$-product is just the product induced on $L^2(V)$ by viewing it as the ψ-coinvariants of the group algebra $L^2(H(V))$; thus the fact that $W^{\psi,\ell}$ is a homomorphism is just due to the fact that $\rho_H^{\psi,\ell}$ is a representation.

As for part (c), formula (26) expresses $W^{\psi,\ell}(h)$ as a smooth integral kernel; we calculate the trace by integrating along the diagonal $x'=a'$ to find

$$\operatorname{Tr} W^{\psi,\ell}(h)\cdot\mu_V^{1/2} = \int_{a'\in\ell'}\int_{a\in\ell} \psi(\omega(a,a'))\cdot h(a,0)\,\mu_V = h(0),$$

the last equality being Fourier inversion. □

Remark 5.3.1. Since trace-class operators form an ideal among bounded operators, we conclude from Proposition 5.2(c) that for any $(g,q)\in\mathrm{Mp}(V)$ and any h smooth and compactly supported (or even Schwartz) on V, the composed operator $\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q)\circ W^{\psi,\ell}(h)$ is also trace-class; its trace is the integral of $T^{\psi}_{(g,q)}$ against h (this is the defining property of $T^{\psi}_{(g,q)}$ in Theorem B). Moreover, if h is now compactly supported on $\mathrm{Mp}^{\psi,\ell}(V)\ltimes H(V)$, we can see why $\operatorname{Tr}\rho^{\psi,\ell}(h)$ — that is, the right-hand side of (2) — is well-defined. For let $h_{g,\xi,t}$ be the restriction of h to

$$\{(g,\xi)\}\times V\times\{t\}\subset \mathrm{Mp}^{\psi,\ell}(V)\ltimes H(V).$$

Then $(g,\xi,t)\mapsto\psi(t)\rho_{\mathrm{Mp}}^{\psi,\ell}(g,\xi)\circ W^{\psi,\ell}(h_{g,\xi,t})$ is a continuous, compactly supported, hence integrable function from $\mathrm{Mp}^{\psi,\ell}(V)\times F$ to trace-class operators, and the trace of its integral is $\operatorname{Tr}\rho^{\psi,\ell}(h)$.

5.4. ***Doubling.*** The metaplectic group $\mathrm{Mp}(V)$ acts on $L^2(V)$ in two ways. First we have a representation A_1,

$$A_1(g,q)(h) := (W^{\psi,\ell})^{-1}(\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q)\circ W^{\psi,\ell}(h)).$$

(The right-hand side makes sense — $\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q)\circ W^{\psi,\ell}(h)$ is in the image of $W^{\psi,\ell}$ — because Hilbert–Schmidt operators form an ideal.) An integral formula for A_1 will be given in Proposition 6.2. Second, let us identify $\mathrm{Sp}(V)$ with the subgroup of

$\mathrm{Sp}(\overline{V} \oplus V)$ acting trivially on $\overline{V}$. The subgroup of $\mathrm{Mp}^{\Gamma_1}(\overline{V} \oplus V)$ over $\mathrm{Sp}(V)$ is precisely $\mathrm{Mp}(V)$ (see Remark 2.7.2). We have an isomorphism

$$b : V \to \Gamma_{-1}, \quad b(x) = (-x/2, x/2) \tag{27}$$

and the restriction map $\mathrm{Res}_{\Gamma_{-1}} : \mathscr{H}^{\psi,\Gamma_1} \to L^2(\Gamma_{-1})$ as in Section 4.4.1. Define

$$R : \mathscr{H}^{\psi,\Gamma_1} \to L^2(V), \quad R := b^* \circ \mathrm{Res}_{\Gamma_{-1}} \tag{28}$$

so that $\mathrm{Mp}(V)$ acts on $L^2(V)$ by $A_2(g, q) := R \circ \rho^{\psi,\Gamma_1}_{\mathrm{Mp}}(g, q) \circ R^{-1}$.

Proposition 5.5. $A_1 = A_2$.

Proof. Consider the representations B_1, B_2 of $H(\overline{V} \oplus V)$ on $L^2(V)$ defined by

$$\begin{aligned} B_1(\bar{v}, v)h(x) &= (W^{\psi,\ell})^{-1}(\rho^{\psi,\ell}_H(v) \circ W^{\psi,\ell}(h) \circ \rho^{\psi,\ell}_H(\bar{v})^{-1}) \\ B_2(\bar{v}, v)h(x) &= R \circ \rho^{\psi,\Gamma_1}_H(\bar{v}, v) \circ R^{-1}(h)(x) \end{aligned}$$

for all $(\bar{v}, v) \in \overline{V} \oplus V$. We have

$$A_i(g, q) \circ B_i(\bar{v}, v) \circ A_i(g, q)^{-1} = B_i(\bar{v}, gv), \quad B_i(1, q) = \gamma_\psi(q) \cdot \mathrm{id}$$

for $i = 1, 2$, and, as in Theorem 4.6, A_2 is uniquely characterized by these equations. We show that in fact $B_1 = B_2$, from which it follows that $A_1 = A_2$.

Write $b'(v) := (v/2, v/2)$ for $v \in V$, so that $(\bar{v}, v) = b(v - \bar{v}) + b'(v + \bar{v})$. Then

$$\begin{aligned} B_2(\bar{v}, v)h(x) &= \big(R \circ \rho^{\psi,\Gamma_1}_H(\bar{v}, v) \circ R^{-1}\big)(h)(x) \\ &= \big(\rho^{\psi,\Gamma_1}_H(b(v - \bar{v}) + b'(v + \bar{v})) \circ R^{-1}\big)(h)(b(x)) \\ &= \big(R^{-1}(h)\big)(b(x) - b(v - \bar{v})) \cdot \psi(\omega(b'(v + \bar{v}), b(x - (v - \bar{v})/2))) \\ &= h(x + \bar{v} - v) \cdot \psi\big(\tfrac{1}{2}\omega(v + \bar{v}, x + \bar{v})\big) \end{aligned}$$

using (24) for the third equality. On the other hand,

$$\begin{aligned} &\rho^{\psi,\ell}_H(v) \circ W^{\psi,\ell}(h) \circ \rho^{\psi,\ell}_H(\bar{v})^{-1} \\ &\qquad = \int_{x \in V} h(x) \rho^{\psi,\ell}_H(v) \rho^{\psi,\ell}_H(x) \rho^{\psi,\ell}_H(\bar{v})^{-1} \, \mu_V^{1/2} \\ &\qquad = \int_{x \in V} h(x) \psi\big(\tfrac{1}{2}\omega(v + \bar{v}, x + v)\big) \rho^{\psi,\ell}_H(v + x - \bar{v}) \, \mu_V^{1/2} \\ &\qquad = \int_{x \in V} h(x + \bar{v} - v) \psi\big(\tfrac{1}{2}\omega(v + \bar{v}, x + \bar{v})\big) \rho^{\psi,\ell}_H(x) \, \mu_V^{1/2} \end{aligned}$$

using the multiplication law of $H(V)$ and then a change of variables. It follows that $B_1(v)h(x) = h(x + \bar{v} - v) \cdot \psi\big(\tfrac{1}{2}\omega(v + \bar{v}, x + \bar{v})\big) = B_2(v)h(x)$ as claimed. □

5.6. ***Proof of Theorem B.*** By the definition of $T^{\psi}_{(g,q)}$, we have

$$\int_V T^{\psi}_{(g,q)}\, h\, \mu_V^{1/2} = \mathrm{Tr}(\rho^{\psi,\ell}_{\mathrm{Mp}}(g,q) \circ W^{\psi,\ell}(h))\mu_V^{1/2}$$

for any $h \in \mathcal{S}(V)$. According to Proposition 5.2(c), the right-hand side equals $A_1(g,q)h(0)$. Therefore, by Proposition 5.5 and Theorem 4.6, we have

$$\begin{aligned}\int_V T^{\psi}_{(g,q)}\, h\, \mu_V^{1/2} &= (R \circ \rho^{\psi,\Gamma_1}_{\mathrm{Mp}}(g,q) \circ R^{-1})(h)(0)\\ &= \gamma_\psi(q) \cdot \int_{y\in\Gamma} (R^{-1}h)(y)\, \mu_g^{\psi,\Gamma_1} \qquad (29)\end{aligned}$$

where, for brevity, $\Gamma := \Gamma_{g^{-1}}/\Gamma_1 \cap \Gamma_{g^{-1}}$. As in the proof of Lemma A.7, define $P : \overline{V} \oplus V \to V$ by $P(v,w) = w - v$; it restricts to an isomorphism

$$P : \Gamma \to (g^{-1}-1)V = (g-1)V, \quad P(x, g^{-1}x) := (g^{-1}-1)x = (g-1)(-g^{-1}x).$$

We use P to rewrite (29) as an integral over $(g-1)V$.

Let $p : \Gamma \to \Gamma_{-1}$ be the projection along Γ_1, and $b : V \to \Gamma_{-1}$ as in (27). Then $P = b^{-1} \circ p$. By (28) and (22) we have, for $y \in \Gamma$,

$$\begin{aligned}(R^{-1}h)(y) &= (\mathrm{Res}^{-1}_{\Gamma_{-1}} \circ (b^*)^{-1}h)(y) = h(P(y))\psi\big(\tfrac{1}{2}\omega(p(y),\, y - p(y))\big)\\ &= h(P(y))\psi\big(\tfrac{1}{2}\omega(p(y), y)\big).\end{aligned}$$

Now (36) gives $\omega(p(y), y) = -Q_{g^{-1}}(P(y), P(y))$. Moreover, it is easy to verify from the definition (35) that $-Q_{g^{-1}} = Q_g$. We therefore have

$$\int_V T^{\psi}_{(g,q)}\, h\, \mu_V^{1/2} = \gamma_\psi(q) \cdot \int_{v\in(g-1)V} h(v)\, \psi\big(\tfrac{1}{2}Q_g(v,v)\big)\, P_*\mu_g^{\psi,\Gamma_1}$$

and it only remains to argue that $P_*\mu_g^{\psi,\Gamma_1} = \mu_{\sigma_g}$.

To do so, note that the natural action of g on (the second factor of) $\overline{V} \oplus V$ fixes $\Gamma_{g^{-1}} \cap \Gamma_1$ point-wise. Therefore, following Remark 4.6.1, we conclude that μ_g^{ψ,Γ_1} is the measure on Γ that is self-dual with respect to $\psi \circ q$, where q is the bilinear form $q(x,y) = \omega(x, gy)$. On the other hand, it is elementary to check that P intertwines the forms q and σ_g, that is, $\sigma_g(P(x), P(y)) = q(x,y)$. Since μ_{σ_g} is self-dual for $\psi \circ \sigma_g$, we must have $P_*\mu_{g,\Gamma_1} = \mu_{\sigma_g}$ as desired.

6. Invariant presentation: Proof of Theorem C

6.1. Now we deduce Theorem C. Here is a reformulation of it, in terms of the representation A_1 of $\mathrm{Mp}(V)$ on $L^2(V)$ defined in Section 5.4. (As noted in footnote 6, we continue to deal with Hilbert spaces of half-densities rather than functions.)

Proposition 6.2. *For any* $(g,q) \in \mathrm{Mp}(V)$ *and* $h \in \mathcal{S}(V)$,

$$A_1(g,q)(h)(x) = \int_{v\in V} T^{\psi}_{(g,q)}(v)\, \psi\big(\tfrac{1}{2}\omega(v,x)\big)\, h(x-v)\, \mu_V^{1/2}$$
$$=: \big(T^{\psi}_{(g,q)}\mu_V^{1/2} \star h\big)(x). \tag{30}$$

Proof. Suppose h is Schwartz. Setting $\tilde{h} := A_1(g,q)(h)$, we want to calculate $\tilde{h}(x)$. For any $f \in \mathcal{S}(V)$ one has $W^{\psi,\ell}(f) \circ \rho_H^{\psi,\ell}(x) = W^{\psi,\ell}(f_x)$, where

$$f_x(v) := f(v-x)\psi\big(\tfrac{1}{2}\omega(v,x)\big).$$

According to Proposition 5.2(c),

$$\tilde{h}(x) = \tilde{h}_{-x}(0) = \mathrm{Tr}(W^{\psi,\ell}(\tilde{h}_{-x})) \cdot \mu_V^{1/2}.$$

Unraveling the definitions, we find

$$\begin{aligned}\tilde{h}(x) &= \mathrm{Tr}(W^{\psi,\ell}(\tilde{h}) \circ \rho_H^{\psi,\ell}(-x)) \cdot \mu_V^{1/2}\\ &= \mathrm{Tr}(\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q) \circ W^{\psi,\ell}(h) \circ \rho_H^{\psi,\ell}(-x)) \cdot \mu_V^{1/2}\\ &= \int_V T^{\psi}_{(g,q)}\, h_{-x}\, \mu_V^{1/2} = \int_{v\in V} T^{\psi}_{(g,q)}(v)\, \psi(\tfrac{1}{2}\omega(v,-x))\, h(v+x)\, \mu_V^{1/2}.\end{aligned}$$

Since $T^{\psi}_{(g,q)}$ is an even function on V, we obtain the right-hand side of (30). □

7. Transfer factor: Proof of Theorem D

7.1. First, in Section 7.2, we give a purely algebraic proof, using the central characters to distinguish between $\rho_+^{\psi,\ell}$ and $\rho_-^{\psi,\ell}$. Then, in Section 7.4, we sketch an alternative argument, because it emphasizes the structure of the Weyl transform, and leads naturally to the geometrization mentioned in Section 1.3. Both methods rely on the following observation.

The decomposition $\rho_{\mathrm{Mp}}^{\psi,\ell} = \rho_+^{\psi,\ell} \oplus \rho_-^{\psi,\ell}$ into irreducible representations corresponds to the decomposition of the representation space $\mathcal{H}^{\psi,\ell} \cong L^2(\ell')$ into even and odd functions. Let $\Pi : \mathcal{H}^{\psi,\ell} \to \mathcal{H}^{\psi,\ell}$ be the parity operator defined by

$$(\Pi f)(x) = f(-x).$$

Then, as generalized functions on $\mathrm{Mp}(V)$,

$$\mathrm{Tr}\,\rho_{\pm}^{\psi,\ell}(g,q) = \tfrac{1}{2}\,\mathrm{Tr}(\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q) \pm \rho_{\mathrm{Mp}}^{\psi,\ell}(g,q) \circ \Pi),$$

whence

$$\mathrm{Tr}(\rho_+^{\psi,\ell} - \rho_-^{\psi,\ell})(g,q) = \mathrm{Tr}\big(\rho_{\mathrm{Mp}}^{\psi,\ell}(g,q) \circ \Pi\big). \tag{31}$$

7.2. *"Algebraic" proof.* The representations $\rho_+^{\psi,\ell}$ and $\rho_-^{\psi,\ell}$ have different central characters, and this can be used to distinguish them. Concretely, the central element $(-1, 1) \in \mathrm{Mp}^{\psi,\ell}(V)$ acts as Π on $\mathscr{H}^{\psi,\ell}$. Given $(g, \xi) \in \mathrm{Mp}^{\psi,\ell}(V)$, one has $(g, \xi)(-1, 1) = (-g, \xi) \in \mathrm{Mp}^{\psi,\ell}(V)$, and therefore

$$\bigl(\mathrm{Tr}\,\rho_+^{\psi,\ell} - \mathrm{Tr}\,\rho_-^{\psi,\ell}\bigr)(g, \xi) = \mathrm{Tr}\,\rho_{\mathrm{Mp}}^{\psi,\ell}(-g, \xi).$$

On the other hand, if in the notation of Section 3.3 we have $\alpha_\ell(g, q) = (g, \xi)$, then $\alpha_\ell^{-1}(-g, \xi) = (-g, q + \epsilon_g)$ as elements of $\mathrm{Mp}(V)$, where

$$\epsilon_g := \tau(\ell \oplus \ell, \Gamma_1, \Gamma_g, \ell \oplus g\ell) - \tau(\ell \oplus \ell, \Gamma_1, \Gamma_{-g}, \ell \oplus (-g)\ell).$$

Since the central factor $W(F)/I^3 \subset \mathrm{Mp}(V)$ acts through γ_ψ, we have

$$(\mathrm{Tr}\,\rho_+^{\psi,\ell} - \mathrm{Tr}\,\rho_-^{\psi,\ell})(g, q) = \mathrm{Tr}\,\rho(-g, q + \epsilon_g) = \mathrm{Tr}\,\rho^{\psi,\ell}(-g, q) \cdot \gamma_\psi(\epsilon_g).$$

Thus it remains to prove the following lemma, which relies on the combinatorics of the Maslov index.

Lemma 7.3. *One has $\epsilon_g = Q_g$ in $W(F)$.*

Proof. Consider the polyhedron with two triangular and two quadrilateral faces:

$$\begin{tikzcd}[arrows=dash]
 & \ell \oplus \ell \arrow[dl] \arrow[dr] & \\
\Gamma_1 \arrow[d] \arrow[drr] & & \ell \oplus g\ell \arrow[d] \\
\Gamma_{-g} \arrow[rr] \arrow[urr, dotted] & & \Gamma_g
\end{tikzcd}$$

As explained in Section A.5.2, the sum of the Maslov indices of the faces vanishes. The sum over the two quadrilateral faces is ϵ_g (note that $(-g)\ell = g\ell$); therefore

$$\epsilon_g = \tau(\Gamma_{-g}, \Gamma_1, \Gamma_g) + \tau(\Gamma_g, \ell \oplus g\ell, \Gamma_{-g}).$$

The second term must vanish, since

$$\tau(\Gamma_g, \ell \oplus g\ell, \Gamma_{-g}) = -\tau(\Gamma_{-g}, \ell \oplus g\ell, \Gamma_g) = -\tau(\Gamma_g, \ell \oplus g\ell, \Gamma_{-g})$$

by Section A.5(a) and (d) applied to

$$1 \oplus (-1) \in \mathrm{GL}(\overline{V} \oplus V).$$

The first term $\tau(\Gamma_{-g}, \Gamma_1, \Gamma_g)$ equals $\tau(\Gamma_1, \Gamma_g, \Gamma_{-1})$ by Section A.5(d) applied to $(x, y) \mapsto (g^{-1}y, x)$, with $\lambda = -1$; but Lemma A.7 says that $\tau(\Gamma_1, \Gamma_g, \Gamma_{-1})$ is the class of Q_g. $\square$

7.4. ***"Analytic" proof.***

Lemma 7.5. *For any $h \in \mathscr{S}(V)$, we have $W^{\psi,\ell}(h) \circ \Pi = W^{\psi,\ell}(\mathfrak{F}h)$, where*

$$\mathfrak{F} : L^2(V) \to L^2(V)$$

is the Fourier transform

$$(\mathfrak{F}h)(x) := \|2\|^{-\frac{\dim V}{2}} \int_{v \in V} h(v)\, \psi\big(\tfrac{1}{2}\omega(v,x)\big)\, \mu_V .$$

Moreover, $\Pi \circ W^{\psi,\ell}(h) \circ \Pi = W^{\psi,\ell}(\Pi h)$ where $\Pi h(v) := h(-v)$.

Proof. The last statement follows directly from (26). From there, too, one sees that $W^{\psi,\ell}(h) \circ \Pi$ is represented by the kernel $A^* \circ B^* \circ (\mathfrak{F}_0 \otimes \mathrm{id})(h)$, where $B(a,b) = (b,a)$. The result then follows from the commutativity of the diagram

$$\begin{tikzcd}
L^2(V) \arrow[r, equal] \arrow[d, "W^{\psi,\ell}\circ\mathfrak{F}"] & L^2(\ell \times \ell') \arrow[r, "\mathfrak{F}_0\otimes\mathrm{id}"] & L^2(\ell' \times \ell') \arrow[r, "\mathrm{id}\otimes\mathfrak{F}_0^{-1}"] \arrow[d, "B^*"] & L^2(\ell' \times \ell) = L^2(V) \arrow[d, equal] \\
\mathrm{End}(\mathscr{H}^{\psi,\ell}) = L^2(\ell' \times \ell') & & L^2(\ell' \times \ell') \arrow[ll, "A^*"'] & L^2(\ell \times \ell') = L^2(V). \arrow[l, "\mathfrak{F}_0\otimes\mathrm{id}"']
\end{tikzcd}$$

Here the top row composes to $\mathfrak{F}$ and the bottom row to $W^{\psi,\ell}$ by Lemma 5.3. □

Now to deduce Theorem D. For brevity, we detail only the case when F is finite, but the infinite case is parallel. Applying the formula for $\mathrm{Tr}\,\rho^{\psi,\ell}_{\mathrm{Mp}}$ from Corollary 1.4, the claim is that

$$\mathrm{Tr}(\rho^{\psi,\ell}_{+} - \rho^{\psi,\ell}_{-})(g,q) = \sqrt{\#V^{-g}} \cdot \gamma_\psi(q)\gamma_\psi(Q_g). \tag{32}$$

By Theorem C, $\rho^{\psi,\ell}_{\mathrm{Mp}}(g,q)$ is the Weyl transform $W^{\psi,\ell}(T^{\psi}_{(g,q)}\, \mu_V^{1/2})$, so (31), Lemma 7.5, and Proposition 5.2(c) give

$$\begin{aligned}
\mathrm{Tr}\big(\rho^{\psi,\ell}_{+} - \rho^{\psi,\ell}_{-}\big)(g,q) &= \mathrm{Tr}\big(\rho^{\psi,\ell}_{\mathrm{Mp}}(g,q) \circ \Pi\big) \\
&= \mathrm{Tr}\, W^{\psi,\ell}(\mathfrak{F}(T^{\psi}_{(g,q)}\mu_V^{1/2})) = \mathfrak{F}(T^{\psi}_{(g,q)})(0).
\end{aligned} \tag{33}$$

The result now follows from Theorem B and the definition of γ_ψ in Section A.3. In detail:

$$\begin{aligned}
\mathfrak{F}(T^{\psi}_{(g,q)})(0) &= \gamma_\psi(q) \cdot \int_{v \in V} \psi\big(\tfrac{1}{2}Q_g(v,v)\big) \cdot D^{\psi}_{g} \cdot \mu_V && \text{(by Thm B)} \\
&= \gamma_\psi(q) \cdot \int_{v \in (g-1)V} \psi\big(\tfrac{1}{2}Q_g(v,v)\big) \cdot \mu_{\sigma_g} && \text{(by def. of } D^{\psi}_{g}\text{)} \\
&= \gamma_\psi(q) \cdot M \int_{v \in (g-1)V/V^{-g}} \psi\big(\tfrac{1}{2}Q_g(v,v)\big) \cdot \mu_{Q_g} && \text{(see below)} \\
&= M\gamma_\psi(q)\gamma_\psi(Q_g) && \text{(by def. of } \gamma_\psi\text{)}.
\end{aligned}$$

To explain the third line, there is a unique measure μ on V^{-g} such that μ_{σ_g} is a product measure $\mu_{\sigma_g} = \mu \otimes \mu_{Q_g}$, and then $M := \int_{V^{-g}} \mu$. However, a self-dual measure on a vector space X is always $1/\sqrt{\#X}$ times counting measure; this implies that $M = \sqrt{\#V^{-g}}$, and the proof of (32) is complete.

Appendix: Witt, Weil, Maslov, Cayley

A.1. ***Witt group.*** (The basic reference is [Lam 2005].) Let F be a field of characteristic not 2. A *quadratic space* is a pair (W, q), where W is a finite-dimensional vector space over F and $q : W \otimes W \to F$ is a nondegenerate symmetric bilinear form. The perpendicular direct sum and the tensor product of two quadratic spaces can be defined in an obvious way. With these operations, the set of isomorphism classes of quadratic spaces forms a commutative semiring. The Witt group (or ring) $W(F)$ is the commutative ring defined by imposing the relation

$$(W, q) + (W, -q) = 0.$$

The *dimension* (or rank) of a quadratic space (W, q) is $\dim W \in \mathbb{Z}$. The *discriminant* of (W, q) is defined as follows. First, q defines a symmetric map $\Phi : W \to W^*$ such that $q(x, y) = \Phi(x)(y)$. Suppose $e_1, \ldots, e_n$ is a basis for W, and $e_1^*, \ldots, e_n^*$ the dual basis for W^*: $e_i^*(e_j) = \delta_{ij}$. Then $\det q \in F$ is the scalar such that

$$\Phi e_1 \wedge \cdots \wedge \Phi e_n = (\det q)(e_1^* \wedge \cdots \wedge e_n^*) \in \wedge^n W^*.$$

The class of $\det q$ in $F^\times/(F^\times)^2$ is well defined, and is called the *discriminant* of (W, q). The *signed discriminant* $\operatorname{sdet} q$ of (W, q) is $(-1)^{n(n-1)/2} \det q \in F^\times/(F^\times)^2$.

Define a commutative ring $W_0(F)$ to be $\mathbb{Z}/2\mathbb{Z} \times F^\times/(F^\times)^2$ as a set, with the operations

$$(d_1, \Delta_1) + (d_2, \Delta_2) := (d_1 + d_2, (-1)^{d_1 d_2} \Delta_1 \Delta_2),$$
$$(d_1, \Delta_1)(d_2, \Delta_2) := (d_1 d_2, \Delta_1^{d_2} \Delta_2^{d_1}).$$

The dimension and signed discriminant together define a surjective homomorphism

$$\widetilde{Q} = (\dim, \operatorname{sdet}) : W(F) \to W_0(F).$$

Let $I \subset W(F)$ be the kernel $I = \ker(\dim)$. Then $\ker \widetilde{Q} = I^2$; see [Lam 2005, Chapter 2, Proposition 2.1]. In other words, $\widetilde{Q}$ identifies $W(F)/I^2$ with $W_0(F)$.

Remark A.1.1. The dimension and signed discriminant make sense for any nondegenerate bilinear form, symmetric or not. Such a form q therefore defines a class $[q]$ in $W_0(F) = W(F)/I^2$.

A.1.2. *Finite and local fields.* We want to describe $W(F)/I^3$, in case F is a finite or local field. For $a, b \in F^\times$, the *Hilbert symbol* $(a, b)_H$ is defined to equal 1 if a is a norm from $F(\sqrt{b})$, and to equal -1 if not. Let Z_F be the image of the Hilbert symbol; it is either $Z_F = \{\pm 1\}$ (when F is real or nonarchimedean) or $Z_F = \{1\}$ (when F is finite or complex). The Hasse invariant $s(q) \in \{\pm 1\}$ of a quadratic space (W, q) over F can be defined inductively by $s(q \oplus q') = s(q)s(q')(\det q, \det q')_H$, and $s(q) = 1$ if $\dim q = 1$.

Theorem A.2. *Let F be any finite or local field of characteristic not 2. Two classes in $W(F)$ are equal modulo I if and only if they can be represented by quadratic spaces of the same rank. Two quadratic spaces of the same rank have the same class modulo I^2 if and only if they have the same discriminant. Two quadratic spaces of the same rank and discriminant have the same class modulo I^3 if and only if they have the same Hasse invariant; moreover, I^2/I^3 is canonically isomorphic to Z_F.*

Proof. For the first statement, every class in $W(F)$ is represented by *some* quadratic space; see, e.g., [Lam 2005, Chapter 2, Proposition 1.4(1)]. If our two classes are represented by (W, q) and (W', q'), with $\dim W - \dim W' = 2m \geq 0$, let (W_0, q_0) be any quadratic space of rank m. Then $q' \oplus q_0 \oplus (-q_0)$ has the same class as q' and the same rank as q. The second statement follows from the isomorphism $\widetilde{Q} : W(F)/I^2 \to W_0(F)$. (The argument so far does not use the assumption that F is finite or local.)

For the third statement, we use the fact that two quadratic spaces of the same dimension have the same class in $W(F)$ if and only if they are isometric [Lam 2005, Chapter 2, Proposition 1.4(3)]. There are four cases.

First, suppose F is nonarchimedean local. Then two quadratic spaces are isometric if and only if they have the same rank, discriminant, and Hasse invariant [Lam 2005, Chapter 6, Theorem 2.12]; moreover, $I^3 = 0$ [Lam 2005, Chapter 6, Corollary 2.15]. So two quadratic spaces of the same rank have the same class in $W(F) = W(F)/I^3$ if and only if they have the same discriminant and Hasse invariant.

Second, suppose $F = \mathbb{F}_q$. This time quadratic spaces are isometric if and only if they have the same rank and discriminant [Lam 2005, Chapter 2, Thdorem 3.5]; the Hasse invariant (like the Hilbert symbol) always equals 1. From this it follows that $I^3 = I^2 = 0$, and we can argue as for the nonarchimedean local case.

Third, suppose $F = \mathbb{C}$. Now two quadratic spaces are isometric if and only if they have the same rank; the discriminant and Hasse invariant (like the Hilbert symbol) always equal 1. This time $I^3 = I = 0$, and we can argue as before.

Fourth, suppose $F = \mathbb{R}$. Isomorphism classes of quadratic spaces are classified by pairs (n_+, n_-) of nonnegative integers, $n_\pm$ being the dimension of the largest positive/negative-definite subspace. The "signature" $\operatorname{sig} : (n_+, n_-) \mapsto n_+ - n_-$

defines an isomorphism $W(F) \to \mathbb{Z}$, identifying I with $2\mathbb{Z}$ and I^3 with $8\mathbb{Z}$. (For all this see [Lam 2005, Chapter 2, Proposition 3.2]. One finds that the rank is $\dim(n_+, n_-) = n_+ + n_-$, $\det(n_+, n_-) = (-1)^{n_-}$, and $s(n_+, n_-) = (-1)^{n_-(n_- -1)/2}$. It follows that the rank and signed discriminant determine the signature mod 4, and that for fixed rank and discriminant, the two choices of Hasse invariant correspond to the two choices of signature mod 8.

For the last statement, it is formally only necessary to show that I^2/I^3 and Z_F have the same number of elements, which follows from the above considerations; however, we will explain the isomorphism using the Weil index — see A.4.1(d) below. □

A.3. *Weil index.* In this section, let F be a finite or local field of characteristic not 2. The Weil index is a homomorphism $\gamma_\psi : W(F) \to Z_8$, where $Z_8 \subset \mathbb{C}^\times$ is the group of eighth roots of unity. It is defined using Fourier transforms.

A.3.1. *Densities and measures.* First let us recall some facts about measures and densities that will be useful both here and in the main text. A nice introduction to densities can be found in [Woodhouse 1980, §5.9].

For $s \in \mathbb{R}$, and X any finite-dimensional vector space over F, let $\Omega_s(X)$ denote the space of complex translation-invariant s-densities on X; it is a one-dimensional complex vector space, the complexification of the space of real translation-invariant s-densities. In particular, there is a canonical isomorphism

$$\Omega_{1/2}(X) \otimes_{\mathbb{C}} \Omega_{1/2}(X) \to \Omega_1(X),$$

and every positive invariant density (that is, Haar measure) $\mu \in \Omega_1(X)$ has a canonical square root $\mu^{1/2} \in \Omega_{1/2}(X)$. The space of functions $X \to \Omega_{1/2}(X)$ has a natural Hermitian inner product:

$$(f_1, f_2) := \int_X \overline{f_1} f_2$$

considering $\overline{f_1} f_2 : X \to \Omega_{1/2}(X) \otimes \Omega_{1/2}(X) = \Omega_1(X)$ as a density on X. Let $L^2(X)$ denote the corresponding Hilbert space.

A perfect pairing $B : X \otimes_F Y \to U(1)$ (making X the Pontryagin dual of Y) associates to each nonzero $\mu \in \Omega_1(X)$ a *dual measure* $\mu^* \in \Omega_1(Y)$. It can be usefully characterized by the *Fourier inversion* formula $(\mathfrak{F}^{\mu^*}_{B^*} \mathfrak{F}^{\mu}_B f)(z) = f(-z)$ for all Schwartz functions $f : X \to \mathbb{C}$. Here

$$(\mathfrak{F}^{\mu}_B f)(y) = \int_{x \in X} f(x) B(x, y)\, \mu$$

and $B^*(y, x) := B(x, y)$ for all $(x, y) \in X \times Y$.

If $Y = X$ then there is a unique *self-dual* $\mu \in \Omega_1(X)$ such that $\mu^* = \mu$. Of particular interest is the situation where $B = B^{\psi}_q := \psi \circ q$ for some nontrivial, continuous homomorphism $\psi : F \to U(1)$ and some nondegenerate bilinear form $q : X \otimes_F X \to F$. It is easy to see from the Fourier inversion formula that if μ_q^ψ is self-dual for B_q^ψ, then the measure that is self-dual for B_{aq}^ψ, $a \in F^\times$, is $\mu_{aq}^\psi = \|a\|^{(\dim X)/2} \mu_q^\psi$.

A.3.2. *Definition.* Suppose now that (X, q) is a quadratic space (that is, q is from now on symmetric). We fix a nontrivial, continuous homomorphism $\psi : F \to U(1)$ and write f_q^ψ for the function $f_q^\psi(x) = \psi\big(\frac{1}{2}q(x, x)\big)$.

Theorem A.4 [Weil 1964, Theorem 2 and Proposition 3]. *There exists a number $\gamma_\psi(q) \in Z_8$ such that*

$$\mathfrak{F}_{B_q^\psi}^{\mu_q^\psi} f_q^\psi = \gamma_\psi(q) \cdot f_{-q}^\psi$$

as generalized functions on X. Moreover, $(X, q) \mapsto \gamma_\psi(q)$ defines a character $\gamma_\psi \psi : W(F) \to Z_8$.

Note that f_q^ψ is not Schwartz, but its Fourier transform can be defined in the sense of distributions.

A.4.1. *Properties.* The following properties of γ_ψ are used in this paper, and go back to [Weil 1964]. For $a \in F^\times$, let q_a be the bilinear form $q_a(x, y) = axy$ on F, and write $\gamma_\psi(a) := \gamma_\psi(q_a)$. We again write $(\cdot\,,\cdot)_H$ for the Hilbert symbol, Z_F for its image, and $s(q) \in Z_F$ for the Hasse invariant of any quadratic space (W, q) (see Section A.1.2).

(a) If $\psi'(x) = \psi(ax)$, then $\gamma_\psi(q_a \otimes q) = \gamma_{\psi'}(q)$.

(b) $\gamma_\psi(a)\,\gamma_\psi(b) = \gamma_\psi(1)\,\gamma_\psi(ab)\,(a, b)_H$.

(c) $\gamma_\psi(q) = \gamma_\psi(1)^{\dim q - 1}\gamma_\psi(\det q)s(q)$.

(d) γ_ψ is trivial on $I^3 \subset W(F)$, and γ_ψ restricts to an isomorphism $I^2/I^3 \to Z_F$.

Proofs. Statement (a) follows easily from the definition of γ_ψ in Theorem A.4 (note that $f_{q_a\otimes q}^\psi = f_q^{\psi'}$, $B_{q_a\otimes q}^\psi = B_q^{\psi'}$, $\mu_{q_a\otimes q}^\psi = \mu_q^{\psi'}$). Statement (b) is equivalent to the last formula on p. 176 of [Weil 1964]. Statement (c) follows from (b) by induction on the dimension (that is, if we decompose q as a perpendicular sum of two smaller spaces). The first part of statement (d) follows from Theorem A.2: if two classes in $W(F)$ are equal modulo I^3, then they can be represented by spaces of the same rank, discriminant, and Hasse invariant, and so by (c) have the same Weil index. For the second part of (d), set $q_{a,b} = (q_1 \oplus q_{-a}) \otimes (q_1 \oplus q_{-b}) = q_1 \oplus q_{-a} \oplus q_{-b} \oplus q_{ab}$, for any $a, b \in F^\times$; I^2 is generated by forms of this type [Lam 2005, Chapter 2, Proposition 1.2]. By (b), $\gamma_\psi(q_{a,b}) = (a, b)_H$, so indeed $\gamma_\psi(I^2) = Z_F$. To see that I^3 is the kernel of γ_ψ on I^2, recall from Theorem A.2 that any two classes in I^2 can be represented by quadratic spaces (W, q), (W', q') of the same rank and discriminant; according to (c), $\gamma_\psi(q) = \gamma_\psi(q')$ if and only if $s(q) = s(q')$, in other words (again according to Theorem A.2) if and only if $q = q' \mod I^3$. □

A.5. ***Maslov index.*** In this section, let F be any field of characteristic not 2. Let (V, ω) be a finite-dimensional symplectic vector space over F. The Maslov index τ associates to each arbitrary sequence $\ell_1, \dots, \ell_n \subset V$ of Lagrangian subspaces, a class $\tau(\ell_1, \dots, \ell_n)$ in $W(F)$. It is characterized by the following properties:

(a) Dihedral symmetry:

$$\tau(\ell_1, \dots, \ell_n) = -\tau(\ell_n, \dots, \ell_1) = \tau(\ell_n, \ell_1, \dots, \ell_{n-1}).$$

(b) Chain condition: For any j, $1 < j < n$,

$$\tau(\ell_1, \ell_2, \dots, \ell_j) + \tau(\ell_1, \ell_j, \dots, \ell_n) = \tau(\ell_1, \ell_2, \dots, \ell_n).$$

(c) Additivity: If V, V' are symplectic spaces, $\ell_1, \dots, \ell_n \in \mathrm{Lagr}(V)$, $\ell'_1, \dots, \ell'_n \in \mathrm{Lagr}(V')$, so that $\ell_i \oplus \ell'_i \in \mathrm{Lagr}(V \oplus V')$, then we have

$$\tau(\ell_1 \oplus \ell'_1, \dots, \ell_n \oplus \ell'_n) = \tau(\ell_1, \dots, \ell_n) + \tau(\ell'_1, \dots, \ell'_n).$$

(d) Invariance: Suppose $g \in \mathrm{GL}(V)$ satisfies $\omega(gx, gy) = \lambda\omega(x, y)$ for all $x, y \in V$. Then

$$\tau(g\ell_1, \dots, g\ell_n) = q_\lambda \otimes \tau(\ell_1, \dots, \ell_n)$$

where $q_\lambda \in W(F)$ is the bilinear form on F defined by $(x, y) \mapsto \lambda xy$.

(e) $\tau(\ell_1, \ell_2, \ell_3)$ can be represented by the (possibly degenerate) bilinear form on $\ell_2 \cap (\ell_1 + \ell_3)$ given by $(x, y) \mapsto \omega(x, y_3)$ (where $y = y_1 + y_3$ with $y_i \in \ell_i$).

For a definition and proofs of (a) and (b), see [Thomas 2006]; (c), (d), and (e) are simple consequences of the definition given there.

A.5.1. *Rank and discriminant.* The rank and discriminant were calculated in [Parimala et al. 2000, Proposition 2.1], with the following result. For each Lagrangian ℓ, choose an "orientation" o, that is, a nonzero element of $\det(\ell)$, the top exterior power of ℓ. Given (ℓ, o), (ℓ', o'), choose an isomorphism $\alpha : \ell \to \ell'$ such that α is the identity on $\ell \cap \ell'$, and $\alpha_*(o) = o'$. Consider the nondegenerate bilinear form $q(x, y) = \omega(\alpha(x), y)$ on $\ell/\ell \cap \ell'$. Set

$$Q(\ell, o; \ell', o') = [q] \in W(F)/I^2$$

(in the notation of Remark A.1.1). It is easy to check that $Q(\ell, o; \ell', o')$, unlike q, is independent of the choice of α; moreover, $Q(\ell', o'; \ell, o) = -Q(\ell, o; \ell', o')$. What [Parimala et al. 2000] show is that, for any choice of orientations $o_i \in \det \ell_i$,

$$\tau(\ell_1, \dots, l_n) = \sum_{i \in \mathbb{Z}/n\mathbb{Z}} Q(\ell_i, o_i; \ell_{i+1}, o_{i+1}) \mod I^2. \tag{34}$$

A.5.2. *Polygons and polyhedra.* Properties (a) and (b) deserve further comment. Suppose given an oriented n-sided polygon F with vertices $\ell_1, \dots, \ell_n$. The dihedral symmetry (a) allows us to unambiguously define $\tau(F) = \tau(\ell_1, \dots, \ell_n)$; reversing the orientation of the polygon reverses the sign of $\tau(F)$. The chain condition (b) has the following interpretation: suppose that P is a closed, oriented polyhedron with vertices $\ell_1, \dots, \ell_n$. Then (b) implies that

$$\sum_F \tau(F) = 0$$

where the sum is over the faces F of P.

A.6. ***Cayley transform.*** We continue with any field F of characteristic not 2. Let (V, ω) be a finite-dimensional symplectic vector space over F.

A.6.1. *Formulas.* For all $g \in \mathrm{Sp}(V)$ there is a symmetric form Q on V given by

$$Q(x, y) = \tfrac{1}{2}\omega((g+1)x, (g-1)y).$$

The kernel is $V^g + V^{-g}$ (a direct sum in V). The corresponding map

$$\mathrm{Sp}(V) \to \mathrm{Sym}^2(V^*) = \mathfrak{sp}(V)$$

is *the Cayley transform* (usually defined without the factor $\frac{1}{2}$); it is traditionally formulated [Cayley 1846] as a bijection between the open subsets of $\mathrm{Sp}(V)$ and $\mathfrak{sp}(V)$ defined (in both cases) by the condition $\det(g-1) \neq 0$.

The canonical isomorphism $V/V^g \to (g-1)V$ transfers Q to a symmetric form Q_g on $(g-1)V$, with kernel V^{-g}. This is the form used in the main text:

$$Q_g((g-1)x, (g-1)y) := \tfrac{1}{2}\omega((g+1)x, (g-1)y) \quad \text{for all } x, y \in V. \tag{35}$$

It is easy to check that $Q_g = -Q_{g^{-1}} = Q_{-g^{-1}}$.

A.6.2. *The Cayley form as a Maslov index.* Let $\overline{V}$ be the same vector space V, but equipped with symplectic form $-\omega$. For $g \in \mathrm{Sp}(V)$, we write Γ_g for the graph $\Gamma_g = \{(v, gv) \mid v \in V\}$ considered as a Lagrangian subspace of the symplectic vector space $\overline{V} \oplus V$.

Lemma A.7. *The class of Q_g in the Witt group $W(F)$ equals the Maslov index $\tau(\Gamma_1, \Gamma_g, \Gamma_{-1})$.*

Proof. Let $p : \overline{V} \oplus V \to \Gamma_{-1}$ be the projection along Γ_1. According to Section A.5(e), $\tau(\Gamma_1, \Gamma_g, \Gamma_{-1})$ can be represented by the degenerate symmetric bilinear form on Γ_g defined by

$$q(x, y) = \omega(x, p(y)).$$

Now consider the map $P : \overline{V} \oplus V \to V$ given by $P(v, w) = w - v$. We have the following more precise claim, which is easy to check: *P induces an isomorphism $\Gamma_g / \Gamma_1 \cap \Gamma_g \to (g-1)V$ that is an isometry between q and Q_g*. In particular,

$$Q_g(P(x), P(y)) = \omega(x, p(y)). \tag{36}$$

This concludes the proof. □

Acknowledgments

I am grateful to Jeff Adams for discussing his work with me and posing some interesting questions, to Vladimir Drinfeld for emphasizing to me the importance of the Weyl transform, to Maurice de Gosson and Amritanshu Prasad for sharing their results with me, and to Masoud Kamgarpour for many useful discussions. While preparing this paper I was supported by a Seggie Brown Fellowship from the University of Edinburgh.

References

[Adams 1998] J. Adams, "Lifting of characters on orthogonal and metaplectic groups", *Duke Math. J.* **92**:1 (1998), 129–178. MR 99h:22014 Zbl 0983.11025

[Cayley 1846] A. Cayley, "Sur quelques propriétés des déterminants gauches", *J. Reine Angew. Math.* **32** (1846), 119–123.

[de Gosson and Luef 2009] M. de Gosson and F. Luef, "The pseudo-character of the Weil representation and its relation with the Conley–Zehnder Index", preprint, 2009. arXiv 0909.1437

[Grove 2001] L. C. Grove, *Classical groups and geometric algebra*, Graduate Studies in Mathematics **39**, American Mathematical Society, Providence, RI, 2001. MR 2002m:20071 Zbl 0990.20001

[Gurevich and Hadani 2007] S. Gurevich and R. Hadani, "The geometric Weil representation", *Selecta Math.* (*N.S.*) **13**:3 (2007), 465–481. MR 2009e:11078 Zbl 1163.22004

[Harish-Chandra 1954] Harish-Chandra, "Representations of semisimple Lie groups. III", *Trans. Amer. Math. Soc.* **76** (1954), 234–253. MR 16,11e Zbl 0055.34002

[Howe 1973] R. Howe, "Invariant theory and duality for classical groups over finite fields", preprint, Yale University, 1973.

[Hua 1948] L.-K. Hua, "On the automorphisms of the symplectic group over any field", *Ann. of Math.* (2) **49** (1948), 739–759. MR 10,352e

[Lafforgue and Lysenko 2009] V. Lafforgue and S. Lysenko, "Geometric Weil representation: local field case", *Compos. Math.* **145**:1 (2009), 56–88. MR 2010c:22024 Zbl 1220.22015

[Lam 2005] T. Y. Lam, *Introduction to quadratic forms over fields*, Graduate Studies in Mathematics **67**, American Mathematical Society, Providence, RI, 2005. MR 2005h:11075 Zbl 1068.11023

[Lion and Vergne 1980] G. Lion and M. Vergne, *The Weil representation, Maslov index and theta series*, Progress in Mathematics **6**, Birkhäuser, Mass., 1980. MR 81j:58075 Zbl 0444.22005

[Maktouf 1999] K. Maktouf, "Le caractère de la représentation métaplectique et la formule du caractère pour certaines représentations d'un groupe de Lie presque algébrique sur un corps p-adique", *J. Funct. Anal.* **164**:2 (1999), 249–339. MR 2000f:22022 Zbl 0948.22017

[Parimala et al. 2000] R. Parimala, R. Preeti, and R. Sridharan, "Maslov index and a central extension of the symplectic group", *K-Theory* **19**:1 (2000), 29–45. MR 2001c:11053a Zbl 1037.11026

[Perrin 1981] P. Perrin, "Représentations de Schrödinger, indice de Maslov et groupe metaplectique", pp. 370–407 in *Noncommutative harmonic analysis and Lie groups* (Marseille, 1980), edited by J. Carmona and M. Vergne, Lecture Notes in Math. **880**, Springer, Berlin, 1981. MR 83m:22027 Zbl 0462.22008

[Prasad 2009] A. Prasad, "On character values and decomposition of the Weil representation associated to a finite abelian group", *J. Anal.* **17** (2009), 73–85. MR 2012a:11053 Zbl 05924869

[Prasad 2011] A. Prasad, "An easy proof of the Stone–von Neumann–Mackey theorem", *Expo. Math.* **29**:1 (2011), 110–118. MR 2012f:22003 Zbl 1227.43009

[Sliman 1984] M. H. Sliman, *Théorie de Mackey pour les groupes adéliques*, Astérisque **115**, Société Mathématique de France, Paris, 1984. MR 86b:22033 Zbl 0552.22001

[Suslin 1987] A. A. Suslin, "Torsion in K_2 of fields", *K-Theory* **1**:1 (1987), 5–29. MR 89a:11123 Zbl 0635.12015

[Thomas 2006] T. Thomas, "The Maslov index as a quadratic space", *Math. Res. Lett.* **13**:5-6 (2006), 985–999. MR 2007j:53093 arXiv 0505561 (expanded version)

[Thomas 2008] T. Thomas, "The character of the Weil representation", *J. Lond. Math. Soc.* (2) **77**:1 (2008), 221–239. MR 2008k:11049 Zbl 1195.11058

[Weil 1964] A. Weil, "Sur certains groupes d'opérateurs unitaires", *Acta Math.* **111** (1964), 143–211. MR 29 #2324 Zbl 0203.03305

[Woodhouse 1980] N. Woodhouse, *Geometric quantization*, Oxford Mathematical Monographs, The Clarendon Press, New York, 1980. MR 84j:58058 Zbl 0458.58003

Communicated by Marie-France Vignéras
Received 2011-08-14 Revised 2012-06-05 Accepted 2012-07-05

jtthomas@uchicago.edu *School of Mathematics, University of Edinburgh, JCMB, The King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ, United Kingdom*

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1571

Analytic families of finite-slope Selmer groups

Jonathan Pottharst

We develop a theory of Selmer groups for analytic families of Galois representations, which are only assumed "ordinary" on the level of their underlying (φ, Γ)-modules. Our approach brings the finite-slope nonordinary case of Iwasawa theory onto an equal footing with ordinary cases in which p is inverted.

Introduction

This paper provides new foundations for the algebraic side of Iwasawa theory. We develop a theory of Galois representations and Galois cohomology over p-adic analytic spaces. In the classical case, where one works over a complete Noetherian local ring, this amounts to passing to the generic fiber of the associated formal scheme (or, what amounts to the same here, inverting p). Moreover, we develop a parallel theory for (φ, Γ)-modules varying in families of the same type. The upshot is that we may mimic Greenberg's "ordinary" Iwasawa theory for Galois representations that look ordinary only on the level of their associated (φ, Γ)-module. Although our work was originally motivated by examples coming from eigenvarieties, we have more recently found significant applications even to the classical case of cyclotomic deformations, to be explained in [Pottharst 2012].

This article is rather technical by nature, so we must remain imprecise in the following outline of our results. First, we develop a theory of group cohomology of a profinite group G with coefficients in families of representations over a p-adic analytic space X over $\mathbf{Q}_p$. By a *family of G-representations* over X, we mean a locally finitely generated, flat $\mathcal{O}_X$-module M, equipped with a continuous map $G \to \mathrm{Aut}^{\mathrm{cont}}_{\mathcal{O}_X}(M)$. In Section 1, we prove the following results:

Theorem. *Assume G has finite cohomology on all discrete G-modules of finite, p-power order, vanishing in degrees greater than e. Then the continuous cohomology with values in $\Gamma(Y, M)$, where $Y \subseteq X$ ranges over affinoid subdomains, gives rise to a perfect complex of coherent $\mathcal{O}_X$-modules, vanishing in degrees greater*

MSC2010: primary 11R23; secondary 11R34, 12G05.
Keywords: Iwasawa theory, Selmer group, families of Galois representations, (φ, Γ)-modules.

than e. If $f : X' \to X$ is a morphism, then there is a canonical isomorphism $\mathbf{L}f^*\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \xrightarrow{\sim} \mathbf{R}\Gamma_{\mathrm{cont}}(G, f^*M)$.

In the case where X is quasi-Stein, we show that $\Gamma(X, \mathbf{R}\Gamma_{\mathrm{cont}}(G, M))$ is computed by $\mathbf{R}\Gamma_{\mathrm{cont}}(G, \Gamma(X, M))$. In the case where X is the generic fiber of the formal scheme $\mathrm{Spf}(A, I)$ and M is the analytification of the I-adic G-representation $\mathfrak{M}$, we show that $\mathbf{R}\Gamma_{\mathrm{cont}}(G, \mathfrak{M}) \otimes_A^{\mathbf{L}} \mathcal{O}_X \xrightarrow{\sim} \mathbf{R}\Gamma_{\mathrm{cont}}(G, M)$, thus providing the link with the classical case.

With these tools in hand, it is straightforward to translate into our context the theory of Selmer complexes (and hence Selmer groups) and show that our theory receives the analytification of the classical theory as in the preceding paragraph.

In Section 2, we turn to the case where $G = G_K$ is the absolute Galois group of a finite extension K of $\mathbf{Q}_p$. We formulate a notion of families of (φ, Γ)-modules over X as above, define their Galois cohomology, and give their basic functorial properties. One important ingredient was conspicuously lacking in a prior version of this paper: we did not know that the $\mathrm{H}^i(G_K, D)$ are finitely generated. This has recently been proven; see [Kedlaya et al. 2012]. As for the relation to the cohomology of Galois representations, we prove the following result:

Theorem. *There is a functorial isomorphism* $\mathbf{R}\Gamma_{\mathrm{cont}}(G_K, M) \xrightarrow{\sim} \mathbf{R}\Gamma(G_K, \mathbf{D}(M))$, *where* $\mathbf{D}(M)$ *is the family of* (φ, Γ_K)*-modules associated to* M.

The essential image of the functor $\mathbf{D}$ is poorly understood at present (see [Hellmann 2012a] for an example of the nontrivial complications that arise), so we note:

Corollary. *Let* $0 \to D' \to E \to D \to 0$ *be a short exact sequence of families of* (φ, Γ_K)*-modules over X as in the preceding theorem. If D and D' arise from families of Galois representations, then so does E.*

In Section 3, we study the p-adic Hodge theory of (φ, Γ_K)-modules, extending to them well-known notions and results for Galois representations. We define *ordinary* (φ, Γ_K)-modules and formulate the (*strict*) *ordinary local condition* in their Galois cohomology and then compare the latter to the Bloch–Kato local conditions. We note that our notion of ordinariness is extremely broad; for example, in the case of modular forms, it includes all cases "of finite slope" (that is, having nonzero U_p-eigenvalue up to p-stabilization and twisting) or, on the automorphic side, having local Weil–Deligne representation at p that is nonsupercuspidal and of nonscalar Frobenius (the latter condition being conjecturally automatic).

We conclude with a semicontinuity result on the ranks of Selmer groups in an ordinary (in our sense) family, which was also observed by Bellaïche [2012]. Various recent works [Hellmann 2012b; Kedlaya et al. 2012; Liu 2012] show that ordinary families are abundant: in particular, any family that is *refined* in the sense of Bellaïche and Chenevier [2009], such as an eigenvariety, is automatically ordinary away from a proper Zariski-closed subset. Thus, our hypotheses are not very restrictive.

Background. The practice of p-adically interpolating Selmer groups goes back to the seminal work of Mazur [1972], where abelian varieties at good ordinary primes are treated; it is suggested there that "the situation is remarkably different" for nonordinary primes. That one definitively cannot integrally interpolate the usual Selmer groups in a naïve way was confirmed by work of Schneider [1987]. For motives satisfying an "ordinary" hypothesis, Greenberg [1989; 1994a] found a purely Galois-cohomological replacement for the p-adic Hodge-theoretic local conditions that is amenable to interpolation. The latter approach was axiomatized by Nekovář [2006].

The above left open the question of what happens in the nonordinary setting. Work of Amice and Vélu [1975] and Višik [1976] showed that the analytic p-adic L-functions of modular forms belong to $\mathcal{O}_X$, and not to $\Lambda[1/p]$ as in the ordinary case, where Λ is the Iwasawa algebra and X is the generic fiber of $\mathrm{Spf}(\Lambda, \mathfrak{m})$. Then, heavily using Fontaine's tools of p-adic Hodge theory, Perrin-Riou [1994b; 2000] constructed algebraic p-adic L-functions (i.e., would-be characteristic ideals), also belonging to Λ_∞. Somewhat surprisingly, her construction eschewed the Selmer groups with finer local conditions although it recovered their characteristic ideals in the ordinary case. Using her language, in the case of modular forms, Kato [2004] used his Euler system to prove a divisibility in an Iwasawa main conjecture and in the ordinary case deduced a statement about Selmer groups.

The next advance came when Kisin [2003] made a Galois-theoretic study of the eigencurve, identifying the relevant two-dimensional p-adic Galois representations as those admitting a crystalline period after twist. Colmez [2005; 2008; 2010] followed the analogy between these representations and principal series, reformulated Kisin's condition in terms of (φ, Γ)-modules (terming it "trianguline"), and made a rigorous p-adic local Langlands correspondence for them. These two works have influenced, e.g., Bellaïche and Chenevier [2009], who refine the methods to make a detailed study of Selmer groups in the infinitesimal neighborhoods of classical points on eigenvarieties.

We briefly mention that somewhat recently there has been progress in nonordinary cyclotomic Iwasawa theory employing similar tools to ours but resulting in mysteriously different outputs. The theory was initiated by R. Pollack and S. Kobayashi (building on work of M. Kurihara) and generalized by F. Sprung, A. Lei, D. Loeffler, and S. Zerbes. See [Pottharst 2012] for references and more commentary on this direction.

Future directions. Our theory is incomplete in that we have direct access to no integral information, having chosen to exchange it for major simplifications in p-adic Hodge theory when working only up-to-isogeny. The remedy for this is likely to be the use of Euler–Poincaré formulas to construct integral isogeny invariants, following [Bloch and Kato 1990; Fontaine and Perrin-Riou 1994; Perrin-Riou 2000]. Still, the theory has several applications. It essentially subsumes Perrin-Riou's cyclotomic Iwasawa theory, as explained in [Pottharst 2012]. Nekovář's work on the parity of Selmer groups in families, as well as the parity conjecture for ordinary Hilbert modular forms of parallel weight, readily generalizes to our setting. In joint work with K. S. Kedlaya, L. Xiao, and the author [Kedlaya et al. 2012], a perfectness and duality result for the Galois cohomology of families of (φ, Γ_K)-modules is applied to give a general construction of triangulations of eigenvarieties as well as a classification of rank-one families of (φ, Γ_K)-modules; for other recent progress on triangulations, see [Chenevier 2010; Hellmann 2012b; Kedlaya et al. 2012; Liu 2012]. Bellaïche has also used our Selmer groups to prove an Iwasawa main conjecture for Eisenstein series using their nonordinary choice of p-stabilization (personal communication). Finally, Benois [2011; 2009] has used methods similar to ours to study $\mathscr{L}$-invariants of Perrin-Riou's Iwasawa L-functions.

This paper is intended as the first step of an Iwasawa theory within the p-adic Langlands program. Namely, Galois-theoretic eigenvarieties for reductive groups H over $\mathbf{Q}$ should be moduli of ordinary filtrations on the (φ, Γ)-modules of universal Galois deformations with values in ${}^{\mathrm{L}}H$. For each $\iota : {}^{\mathrm{L}}H \to \mathrm{GL}_d$ preserving the ordinariness of the filtration, the Galois-theoretic eigenvariety will then have a natural Selmer module. First steps in this direction have been made by Chenevier [2010] and Hellmann [2010; 2012b]. The automorphic (i.e., usual) eigenvariety associated to H will map to the Galois-theoretic one by virtue of its family of Galois representations. A generalization à la Kisin of the ordinary "$R = T$" conjectures would predict this map to be an isomorphism, and Iwasawa theory would relate the ι-Selmer module to the p-adic L-function interpolating the ι-L-values of automorphic representations on H.

Notation. Throughout, we fix a prime p and a finite extension E of $\mathbf{Q}_p$ with ring of integers $\mathbb{O}_E$.

Let $a \le b$ be integers. For $* \in \{[a, b], \mathrm{b}, +, -, \varnothing\}$, we say that a complex or graded module is $*$-*bounded* if it is, respectively, concentrated in degrees $[a, b]$, bounded, bounded above, bounded below, or is arbitrary.

If $(Z, \mathbb{O}_Z)$ denotes a ringed topos and $? \in \{\varnothing, \mathrm{ft}\}$, we write $\mathbf{K}^*_?(Z)$ for the category of complexes of $\mathbb{O}_Z$-modules, each of whose cohomologies is $*$-bounded and, if $? = \mathrm{ft}$, satisfies a finiteness condition to be made precise as it arises. We write $\mathbf{D}^*_?(Z)$ for its derived category, and we write $\mathbf{Gr}^*_?(Z)$ for the category of $*$-bounded graded $\mathbb{O}_Z$-modules, each of whose components satisfies ?. Denote by $[\,\cdot\,] : \mathbf{K}^*_?(Z) \to \mathbf{D}^*_?(Z)$ and $\mathrm{H}^* : \mathbf{D}^*_?(Z) \to \mathbf{Gr}^*_?(Z)$ the obvious functors. Denote also $\mathbf{D}^{[a,b]}_{\mathrm{perf}}(Z) \subseteq \mathbf{D}^{[a,b]}_{\mathrm{ft}}(Z)$ the strictly full subcategory consisting of objects X quasi-isomorphic to complexes $C^\bullet$ concentrated in degrees $[a, b]$ consisting of ft and flat modules. On the latter category, $X \mapsto X^* = \mathbf{R}\operatorname{Hom}_{\mathbb{O}_Z}(X, \mathbb{O}_Z)$ is in each case under consideration an anti-involution, and for $X \cong [C^\bullet]$ as above, X^* is represented by $\operatorname{Hom}_{\mathbb{O}_Z}(C^\bullet, \mathbb{O}_Z)$.

1. Group cohomology

1A. ***Continuous cochains: local calculations.*** We will say that an $\mathcal{O}_E$-module M is *linearly topologized* if it is equipped with a topology with basis around the identity consisting of a decreasing sequence M_n of $\mathcal{O}_E$-submodules. We will say that an $\mathcal{O}_E$-algebra R is *linearly topologized* (*as an algebra*) if the system of submodules R_n can be chosen so that $R_n \cdot R_n \subseteq R_n$. For such R, we will say that an R-module M is *linearly topologized compatibly with* R if the systems R_n and M_n of open submodules can be chosen so that $R_n \cdot M_n \subseteq M_n$; in particular, the multiplication map $R \times M \to M$ is bicontinuous.

Let G be a profinite group. A *continuous G-module* is a linearly topologized $\mathcal{O}_E$-module M endowed with a continuous map $G \to \mathrm{Aut}^{\mathrm{cont}}_{\mathcal{O}_E}(M)$, the latter equipped with the compact-open topology. Given a linearly topologized $\mathcal{O}_E$-algebra R, a *continuous $R[G]$-module* is an R-module M that is linearly topologized compatibly with R endowed with a continuous map $G \to \mathrm{Aut}^{\mathrm{cont}}_R(M)$, the latter again equipped with the compact-open topology. We define the complex $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \in \mathbf{K}^+(R)$ of *continuous cochains* on G with values in M to be $\mathrm{C}^i_{\mathrm{cont}}(G, M) = \mathrm{Map}^{\mathrm{cont}}(G^i, M)$ with the usual differential (see, e.g., [Nekovář 2006, 3.4.1.2]). We denote its image in $\mathbf{D}^+(R)$ by $\mathbf{R}\Gamma_{\mathrm{cont}}(G, M)$ and its cohomology by $\mathrm{H}^*_{\mathrm{cont}}(G, M) \in \mathbf{Gr}^+(R)$. The latter defines a functor that, of course, turns short exact sequences into long exact sequences provided the usual existence of continuous (though not necessarily group-theoretic) sections. The reader may check that, under our hypotheses below, one always has the necessary continuous sections for turning short exact sequences into long exact sequences.

In order to get reasonable behavior, we will need to impose some hypotheses. The following are sufficient for our applications:

Hypotheses A. (1) G is a profinite group having finite p-cohomological dimension e, and $\#\mathrm{H}^i_{\mathrm{cont}}(G, T) < \infty$ for all finite discrete $\mathbf{F}_p[G]$-modules T and all $i \geq 0$.

(2) A is a Noetherian $\mathcal{O}_E$-algebra, separated and complete with respect to a proper ideal I containing a power of p and equipped with the I-adic topology.

(3) M is a finite-type A-module, considered with its I-adic topology and equipped with a continuous $A[G]$-module structure.

(4) The A-module M is flat.

We say that an A-module satisfies the condition "ft" if it is of finite type.

Note that, under Hypotheses A(1)–(3), since A/I^N has the discrete topology for $N > 0$, the stabilizer of any element of M/I^N is open in G; since M is finitely generated, we see that G acts on M/I^N through a finite quotient.

The following is the main result of this section:

Theorem 1.1. *Assume Hypotheses A.*

(1) *The complexes* $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M/I^N)$ *and* $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M)$ *consist of flat* A/I^N*-modules and* A*-modules, respectively.*

(2) *The inverse system* $\{\mathrm{H}^*_{\mathrm{cont}}(G, M/I^N)\}_N$ *satisfies Mittag-Leffler.*

(3) *The natural map* $\mathrm{H}^*_{\mathrm{cont}}(G, M) \to \varprojlim_N \mathrm{H}^*_{\mathrm{cont}}(G, M/I^N)$ *is an isomorphism.*

(4) *The* $\mathrm{H}^i_{\mathrm{cont}}(G, M)$ *are finitely generated* A*-modules and vanish for* $i > e$*.*

The above theorem shows that $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \in \mathbf{K}^{[0,e]}_{\mathrm{ft}}(A)$. In fact, copying the proof of [Nekovář 2006, 4.2.9] verbatim (in the case $a = b = 0$, $\mathcal{S} = \{1\}$), one obtains the following strengthening:

Corollary 1.2. *Assume Hypotheses A. Then* $\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \in \mathbf{D}^{[0,e]}_{\mathrm{perf}}(A)$.

Lemma 1.3. *Assume Hypotheses A(1)–(3).*

(1) *For any compact topological space* X*, the natural maps*

$$\mathrm{Map}^{\mathrm{cont}}(X, M)/I^N \to \mathrm{Map}^{\mathrm{cont}}(X, M/I^N)$$

are isomorphisms.

(2) *The natural maps*

$$\mathrm{C}^\bullet_{\mathrm{cont}}(G, M)/I^N \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M/I^N)$$

are isomorphisms of complexes.

(3) *One has* $\mathrm{H}^i_{\mathrm{cont}}(G, M) = 0$ *for* $i > e$*.*

(4) *If* M *is annihilated by a power of* I*, then* $\mathrm{H}^*_{\mathrm{cont}}(G, M)$ *is a finitely generated* A*-module.*

(5) *If Hypothesis (4) holds too, then for each* $N > 0$ *the complexes* $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M/I^N)$ *and* $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M)$ *consist of flat* A/I^N*-modules and* A*-modules, respectively.*

Proof. For (1), the main fact we will use is that the continuous maps from a compact topological space to a discrete topological space are precisely the locally constant ones.

The map of the claim is surjective by the discreteness of M/I^N; that it is injective amounts to the claim that the natural map $I^N \cdot \mathrm{Map}^{\mathrm{cont}}(X, M) \to \mathrm{Map}^{\mathrm{cont}}(X, I^N M)$ is surjective. As the source of this map is complete and $I^N M \xrightarrow{\sim} \varprojlim_k I^N M/I^{N+k}M$, it suffices to show that the maps $I^N \cdot \mathrm{Map}^{\mathrm{cont}}(X, M) \to \mathrm{Map}^{\mathrm{cont}}(X, I^N M/I^{N+k}M)$ for $k \geq 0$ are surjective. But the $I^N M/I^{N+k}M$ are discrete, so the latter claim is obvious.

For (2), the maps are clearly compatible with differentials, so we must check that they are isomorphisms term-by-term. For the i-th term, this results from (1) with $X = G^i$.

To prove (3), observe that by the universal property of the inverse limit one has $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) = \varprojlim_N \mathrm{C}^\bullet_{\mathrm{cont}}(G, M/I^N)$ with surjective transition maps; hence, for each i, we have short exact sequences

$$0 \to \mathbf{R}^1 \varprojlim_N \mathrm{H}^{i-1}_{\mathrm{cont}}(G, M/I^N) \to \mathrm{H}^i_{\mathrm{cont}}(G, M) \to \varprojlim_N \mathrm{H}^i_{\mathrm{cont}}(G, M/I^N) \to 0.$$

Since each M/I^N is a discrete p-primary G-module, the claim for $i > e+1$ follows. For $i = e+1$, the inverse system $\{\mathrm{H}^{i-1}_{\mathrm{cont}}(G, M/I^N)\}_N$ has surjective maps because each $\mathrm{H}^i_{\mathrm{cont}}(G, I^N M/I^{N+k} M) = 0$. This implies that the $\mathbf{R}^1 \varprojlim_N$-term vanishes, giving the claim in this case too.

Next we treat (4). We may perform a dévissage to reduce to the case where $(I, p)M = 0$, and then we choose an open normal subgroup $H \subseteq G$ with $M^H = M$. Since M is $\mathbf{F}_p$-flat, the natural map $\mathrm{H}^*_{\mathrm{cont}}(H, \mathbf{F}_p) \otimes_{\mathbf{F}_p} M \to \mathrm{H}^*_{\mathrm{cont}}(H, M)$ is an isomorphism. By Hypothesis (1), the term $\mathrm{H}^*_{\mathrm{cont}}(H, \mathbf{F}_p)$ is finite; hence, $\mathrm{H}^*_{\mathrm{cont}}(H, M)$ is finitely generated over A. Now consider the spectral sequence

$$\mathrm{H}^i_{\mathrm{cont}}(G/H, \mathrm{H}^j_{\mathrm{cont}}(H, M)) \Longrightarrow \mathrm{H}^{i+j}_{\mathrm{cont}}(G, M).$$

The terms on the left are finitely generated over A because G/H is finite. This forces $\mathrm{H}^*_{\mathrm{cont}}(G, M)$ to be finitely generated.

Finally, for (5), it follows from (2) that for any ideal J of A one has

$$J \otimes_A \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \xrightarrow{\sim} J \otimes_A \varprojlim_N \mathrm{C}^\bullet_{\mathrm{cont}}(G, M)/I^N.$$

But A is Noetherian, so J is finitely presented, and $\otimes_A J$ commutes with taking inverse limits of surjective systems. Therefore, one has

$$J \otimes_A \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \xrightarrow{\sim} \varprojlim_N (J \otimes_A \mathrm{C}^\bullet_{\mathrm{cont}}(G, M))/I^N,$$

and the right-hand side is clearly I-adically separated. This verifies that $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M)$ is "I-adically ideal separated". Hence, by the well-known local criterion for flatness, e.g., [Matsumura 1989, Theorem 22.3(1, 5)], this complex consists of flat A-modules if and only if each of its respective reductions modulo I^N is A/I^N-flat. But since M is A-flat, it follows that the M/I^N are A/I^N-flat, and indeed so are the complexes $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M/I^N) = \varinjlim_{H \trianglelefteq G \text{ open}} (M/I^N)^{\oplus (G/H)^\bullet}$. We conclude by again using (2). □

Proof of Theorem 1.1. The claim (1) follows immediately from Lemma 1.3(5), and the last part of the claim (4) is Lemma 1.3(3).

The remaining claims follow from the results of [Berthelot and Ogus 1978, Appendix B], whose conventions for handling inverse limits in the derived category we now recall. Namely, we consider the poset $\mathbf{N}$ of nonnegative integers as a Grothendieck site with the discrete topology: only identity maps are coverings. A sheaf on $\mathbf{N}$ is merely an inverse system, and the condition that a sheaf be flasque simply amounts to the Mittag-Leffler condition. We equip $\mathbf{N}$ with the sheaf $A_{\cdot}$ of rings given by $N \mapsto A_N = A/I^{N+1}$ and work in the derived category $\mathbf{D}(\mathbf{N}, A_{\cdot})$ of $A_{\cdot}$-modules on $\mathbf{N}$.

Write $M_N = M \otimes_A A_N$; the $\mathrm{C}^i_{\mathrm{cont}}(G, M_N)$ determine a complex $C^\bullet_{\cdot}$ of sheaves of $A_{\cdot}$-modules. Lemma 1.3(3) shows that $[C^\bullet_{\cdot}] \in \mathbf{D}^-(\mathbf{N}, A_{\cdot})$, and Lemma 1.3(5) implies that $[C^\bullet_{N+1}] \otimes^{\mathbf{L}}_{A_{N+1}} A_N$ is represented by the complex $C^\bullet_{N+1} \otimes_{A_{N+1}} A_N$, which by Lemma 1.3(2) is isomorphic to $C^\bullet_N$. The latter two claims mean that $C^\bullet_{\cdot}$ is what in [Berthelot and Ogus 1978] is called a *quasiconsistent* complex. Lemma 1.3(4) shows (in particular) that $[C^\bullet_0] \in \mathbf{D}^-_{\mathrm{ft}}(A_0)$, and hence, the main finiteness result [Berthelot and Ogus 1978, Proposition B.7][1] applies to $C^\bullet_{\cdot}$ and $\mathbf{R}\varprojlim_N [C^\bullet_{\cdot}]$. Finally, Lemma 1.3(2) shows that $C^\bullet_{\cdot}$ is a complex of flasque sheaves, and therefore, $\mathbf{R}\varprojlim_N [C^\bullet_{\cdot}] = [\varprojlim_N C^\bullet_N] = [\mathrm{C}^\bullet_{\mathrm{cont}}(G, M)]$, allowing us to rephrase the finiteness result as claims (2)–(4). □

We turn to base-changing properties. We say that B is an *I-adic A-algebra* if it is a Noetherian A-algebra that is (IB)-adically separated and complete and equipped with the (IB)-adic topology; when no confusion may arise, we abusively denote by I the ideal IB of B.

Theorem 1.4. *Assume Hypotheses A hold, and let B be an I-adic A-algebra.*

(1) *The natural map*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \otimes^{\mathbf{L}}_A B \to \mathbf{R}\Gamma_{\mathrm{cont}}(G, M \otimes_A B)$$

is an isomorphism in $\mathbf{D}^{[0,e]}_{\mathrm{perf}}(B)$.

(2) *There is a canonical spectral sequence*

$$\mathrm{E}_2^{ij} = \mathrm{Tor}^A_{-i}(\mathrm{H}^j_{\mathrm{cont}}(G, M), B) \Longrightarrow \mathrm{H}^{i+j}_{\mathrm{cont}}(G, M \otimes_A B).$$

(3) *If B is flat over A or becomes flat after inverting p, then the natural map*

$$\mathrm{H}^*_{\mathrm{cont}}(G, M) \otimes_A B \to \mathrm{H}^*_{\mathrm{cont}}(G, M \otimes_A B)$$

is an isomorphism or becomes an isomorphism after inverting p, respectively.

Lemma 1.5. *Assume Hypotheses A hold, and let B be an I-adic A-algebra.*

(1) *For any compact topological space X, the natural map*

$$\mathrm{Map}^{\mathrm{cont}}(X, M) \widehat{\otimes}_A B \to \mathrm{Map}^{\mathrm{cont}}(X, M \widehat{\otimes}_A B)$$

is an isomorphism.

[1] In the statement of this result, the hypothesis “$D_0 \in \mathbf{D}^-_{\mathrm{ft}}(\mathbf{N}, A_{\cdot})$” should read “$D_0 \in \mathbf{D}^-_{\mathrm{ft}}(A_0)$”.

(2) *The natural map*

$$\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \mathbin{\widehat{\otimes}}_A B \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M \otimes_A B)$$

is an isomorphism of complexes.

(3) *The natural map*

$$\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \otimes_A B \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \mathbin{\widehat{\otimes}}_A B$$

is a quasi-isomorphism.

Proof. We treat (1). One has natural maps

$$\begin{aligned}\mathrm{Map}^{\mathrm{cont}}(X, M) \otimes_A B/I^N &= \mathrm{Map}^{\mathrm{cont}}(X, M)/I^N \otimes_A B/I^N \\ &\xrightarrow{\alpha} \mathrm{Map}^{\mathrm{cont}}(X, M/I^N) \otimes_A B/I^N \\ &\xrightarrow{\beta} \mathrm{Map}^{\mathrm{cont}}(X, M \otimes_A B/I^N).\end{aligned}$$

The map α is an isomorphism by Lemma 1.3(2). One easily deduces that β is an isomorphism from the fact that all of M/I^N, B/I^N, and $M \otimes_A B/I^N$ are discrete. Thus, we deduce the claim by passing to the inverse limit over N.

For (2), the map is clearly compatible with differentials, so we must check that it is an isomorphism term-by-term. For the i-th term, this results from applying (1) with $X = G^i$ and noting that $M \mathbin{\widehat{\otimes}}_A B = M \otimes_A B$ because M is finitely generated over A.

We now show (3). Choose a quasi-isomorphism $D^\bullet \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \otimes_A B$ with $D^\bullet$ a bounded-above complex of finitely generated, flat B-modules. Because both $D^\bullet$ and $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \otimes_A B$ are B-flat, the induced maps

$$D^\bullet \otimes_B B/I^N \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \otimes_A B/I^N$$

are quasi-isomorphisms. On the other hand, both systems satisfy Mittag-Leffler so that, applying $\varprojlim_N$, the induced map $\widehat{D}^\bullet \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \mathbin{\widehat{\otimes}}_A B$ is also a quasi-isomorphism. We conclude by tracing around the commutative diagram

$$\begin{tikzcd} D^\bullet \arrow[r, "\sim"] \arrow[d, "\wr"] & \widehat{D}^\bullet \arrow[d, "\wr"] \\ \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \otimes_A B \arrow[r] & \mathrm{C}^\bullet_{\mathrm{cont}}(G, M) \mathbin{\widehat{\otimes}}_A B \end{tikzcd}$$

where the top horizontal isomorphism is because the D^i are finitely generated. □

Proof of Theorem 1.4. The claim (1) follows from Lemmas 1.3(5) and 1.5(2)–(3), and (2) follows from (1) formally. To see (3), consider the spectral sequence of (2). After inverting p if necessary, all terms with $i \neq 0$ vanish, yielding degeneracy at E_2, whence the desired isomorphism. □

Often, we are only given a finitely generated $A[1/p]$-module M equipped with the structure of a continuous $A[1/p][G]$-module. If there exists a finitely generated, flat, G-stable A-submodule $M_0 \subseteq M$ such that $M_0[1/p] = M$, then $M_0 \otimes_A B$ plays the same role inside $M \otimes_{A[1/p]} B[1/p]$, and the preceding finiteness, perfectness, and base-change results apply with $A[1/p]$, $B[1/p]$, and M in place of A, B, and M_0, respectively, and all objects occurring in it are independent of the choice of M_0.

For any strictly E-affinoid space Y, we write $\mathcal{A}_Y$ for the set of unit balls of Banach algebra norms on $\Gamma(Y, \mathcal{O}_Y)$. Any $A \in \mathcal{A}_Y$, equipped with $I = (p)$, satisfies Hypothesis (2). If M is a finitely generated, flat $\Gamma(Y, \mathcal{O}_Y)$-module equipped with the structure of a continuous $\Gamma(Y, \mathcal{O}_Y)[G]$-module, then by [Chenevier 2009, Lemme 3.18] one has $M = M_0[1/p]$ for M_0 a finitely generated, flat, G-stable A-submodule, for some $A \in \mathcal{A}_Y$, and $\mathbf{R}\Gamma(G, M) = \mathbf{R}\Gamma(G, M_0)[1/p]$ belongs to $\mathbf{D}^{[0,e]}_{\mathrm{perf}}(\Gamma(Y, \mathcal{O}_Y))$. If $f : Y' \to Y$ is a morphism of affinoid spaces, then the image of A in $\Gamma(Y', \mathcal{O}_{Y'})$ is contained in some $B \in \mathcal{A}_{Y'}$, and any such B is a p-adic A-algebra. Thus, by the preceding paragraph, we may apply Theorem 1.4 with A and B replaced by $\Gamma(Y, \mathcal{O}_Y)$ and $\Gamma(Y', \mathcal{O}_{Y'})$. In particular, if Y' is an affinoid subdomain of Y, then $B[1/p] = \Gamma(Y', \mathcal{O}_{Y'})$ is flat over $A[1/p] = \Gamma(Y, \mathcal{O}_Y)$, so Theorem 1.4(3) applies.

1B. ***General p-adic analytic spaces.*** Let X be a p-adic analytic space over E, and let $\mathcal{U}$ be an admissible affinoid covering that is *quasiclosed under intersection*, meaning that whenever $Y, Y' \in \mathcal{U}$ then $Y \cap Y'$ has an admissible cover consisting of elements of $\mathcal{U}$. We consider $\mathcal{U}$ as a poset category and equip it with the discrete Grothendieck topology: only identity maps are coverings, and all presheaves are sheaves. Thus, homological algebra of sheaves on $\mathcal{U}$ is essentially carried out independently over each affinoid. It is a ringed site via the rule $\Gamma(Y, \mathcal{O}_{\mathcal{U}}) = \Gamma(Y, \mathcal{O}_Y)$, and an $\mathcal{O}_{\mathcal{U}}$-module M consists of the data of a $\Gamma(Y, \mathcal{O}_Y)$-module $\Gamma(Y, M)$ for each $Y \in \mathcal{U}$ together with a morphism $\Gamma(Y, M) \to \Gamma(Y', M)$ of $\Gamma(Y, \mathcal{O}_Y)$-modules for each $Y, Y' \in \mathcal{U}$ with Y' contained in Y, satisfying the obvious compatibility law for $Y'' \subseteq Y' \subseteq Y$. Similarly to the situation in [Berthelot and Ogus 1978, Appendix B], we say that a complex $C^\bullet$ of $\mathcal{O}_{\mathcal{U}}$-modules is *quasiconsistent* if each induced map

$$\Gamma(Y, C^\bullet) \otimes_{\Gamma(Y,\mathcal{O}_Y)} \Gamma(Y', \mathcal{O}_{Y'}) \to \Gamma(Y', C^\bullet)$$

is a quasi-isomorphism. Since $\Gamma(Y', \mathcal{O}_{Y'})$ is $\Gamma(Y, \mathcal{O}_Y)$-flat, we in fact have an isomorphism

$$[\Gamma(Y, C^\bullet)] \otimes^{\mathbf{L}}_{\Gamma(Y,\mathcal{O}_Y)} \Gamma(Y', \mathcal{O}_{Y'}) \xrightarrow{\sim} [\Gamma(Y', C^\bullet)]$$

in $\mathbf{D}(\Gamma(Y', \mathcal{O}_{Y'}))$. We say that a quasiconsistent complex of $\mathcal{O}_{\mathcal{U}}$-modules $C^\bullet$ is *of finite type, or satisfies the condition* "ft"*,* or is *flat* if, for all $i \in \mathbf{Z}$ and for all $Y \in \mathcal{U}$, the $\Gamma(Y, \mathcal{O}_Y)$-module $\Gamma(Y, C^i)$ is of finite type or flat, respectively. Quasiconsistent $\mathcal{O}_{\mathcal{U}}$-modules form an abelian subcategory of all $\mathcal{O}_{\mathcal{U}}$-modules that is closed under extensions. Note that a complex of quasiconsistent $\mathcal{O}_{\mathcal{U}}$-modules is a quasiconsistent complex of $\mathcal{O}_{\mathcal{U}}$-modules but not in general conversely.

From now on, by a *complex of* $\mathcal{O}_{\mathfrak{U}}$*-modules*, we implicitly mean *quasiconsistent complex of* $\mathcal{O}_{\mathfrak{U}}$*-modules* unless said otherwise. Especially, for $* \in \{+, -, \mathrm{b}, \varnothing\}$, the notations $\mathbf{K}^*_{\mathrm{ft}}(\mathfrak{U})$, $\mathbf{D}^*_{\mathrm{ft}}(\mathfrak{U})$, $\mathbf{D}^{[a,b]}_{\mathrm{perf}}(\mathfrak{U})$, and $\mathbf{Gr}^*_{\mathrm{ft}}(\mathfrak{U})$ denote categories of *quasiconsistent* complexes or graded modules of finite type. We have the obvious commutative diagram, where the vertical arrows denote taking sections over $Y \in \mathfrak{U}$:

$$\begin{array}{ccccc} \mathbf{K}^*_{\mathrm{ft}}(\mathfrak{U}) & \xrightarrow{[\,\cdot\,]} & \mathbf{D}^*_{\mathrm{ft}}(\mathfrak{U}) & \xrightarrow{\mathrm{H}^*} & \mathbf{Gr}^*_{\mathrm{ft}}(\mathfrak{U}) \\ \downarrow & & \downarrow & & \downarrow \\ \mathbf{K}^*_{\mathrm{ft}}(\Gamma(Y,\mathcal{O})) & \xrightarrow{[\,\cdot\,]} & \mathbf{D}^*_{\mathrm{ft}}(\Gamma(Y,\mathcal{O})) & \xrightarrow{\mathrm{H}^*} & \mathbf{Gr}^*_{\mathrm{ft}}(\Gamma(Y,\mathcal{O})) \end{array} \tag{1-1}$$

By a *family of G-representations* over X, we mean a locally finitely generated, flat $\mathcal{O}_X$-module M, equipped with a continuous map $G \to \mathrm{Aut}^{\mathrm{cont}}_{\mathcal{O}_X}(M)$. Then M determines by restriction a finitely generated, flat $\mathcal{O}_{\mathfrak{U}}$-module, which we also denote by M, whose group of sections over each $Y \in \mathfrak{U}$ is a continuous $\Gamma(Y, \mathcal{O})[G]$-module. By the discussion at the end of Section 1A, it follows from Theorem 1.4 that the $\mathcal{O}_{\mathfrak{U}}$-module determined by the rule $Y \mapsto \mathbf{C}^\bullet_{\mathrm{cont}}(G, \Gamma(Y, M))$ is quasiconsistent; combined with Theorem 1.1(4), this rule hence determines an object of $\mathbf{K}^{[0,e]}_{\mathrm{ft}}(\mathfrak{U})$, which we denote by $\mathbf{C}^\bullet_{\mathrm{cont}}(G, M)$. Its class $\mathbf{R}\Gamma_{\mathrm{cont}}(G, M)$ in the derived category belongs to $\mathbf{D}^{[0,e]}_{\mathrm{perf}}(\mathfrak{U})$, and its cohomology $\mathrm{H}^*_{\mathrm{cont}}(G, M)$ belongs to $\mathbf{Gr}^{[0,e]}_{\mathrm{ft}}(\mathfrak{U})$. For any $Y \in \mathfrak{U}$, we have

$$\Gamma(Y, \mathbf{C}^\bullet_{\mathrm{cont}}(G, M)) = \mathbf{C}^\bullet_{\mathrm{cont}}(G, \Gamma(Y, M)), \tag{1-2}$$

and it follows from the commutativity of the diagram (1-1) that also

$$\Gamma(Y, \mathbf{R}\Gamma_{\mathrm{cont}}(G, M)) = \mathbf{R}\Gamma_{\mathrm{cont}}(G, \Gamma(Y, M)),$$
$$\Gamma(Y, \mathrm{H}^*_{\mathrm{cont}}(G, M)) = \mathrm{H}^*_{\mathrm{cont}}(G, \Gamma(Y, M)),$$

so all we really have is a compatible family of cohomology data over the affinoids in question. Using Kiehl's theorem, we may identify $\mathrm{H}^*_{\mathrm{cont}}(G, M)$ to an object of $\mathbf{Gr}^{[0,e]}_{\mathrm{coh}}(X)$, in which the subscript denotes coherent $\mathcal{O}_X$-modules. Since the latter is invariant under passing between $\mathfrak{U}$ and a refinement, we canonically associate to M a coherent analytic sheaf on X whose sections over any affinoid domain Y give the continuous cohomology of G with coefficients in the sections of M over Y.

We now state a general base-change theorem for group cohomology. Suppose we are given a morphism $f : X' \to X$ of p-adic analytic spaces over E, and let $\mathfrak{U}'$ be an admissible affinoid covering of X' that is quasiclosed under intersection with the property that for each $Y' \in \mathfrak{U}'$ there exists $Y \in \mathfrak{U}$ with $f(Y') \subseteq Y$. Any object $C^\bullet$ of $\mathbf{K}^-_{\mathrm{ft}}(\mathfrak{U})$ gives rise to an object $\mathbf{L}f^*C^\bullet$ of $\mathbf{D}^-_{\mathrm{ft}}(\mathfrak{U}')$ by the usual recipe.

Theorem 1.6. (1) *The natural map*

$$\mathbf{L}f^*\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \to \mathbf{R}\Gamma_{\mathrm{cont}}(G, f^*M)$$

is an isomorphism in $\mathbf{D}^{[0,e]}_{\mathrm{perf}}(\mathcal{U}')$.

(2) *There exists a spectral sequence in coherent* $\mathcal{O}_{X'}$*-modules*

$$\mathrm{E}_2^{ij} = \mathrm{Tor}_{-i}^{f^{-1}\mathcal{O}_X}(f^{-1}\mathrm{H}^j_{\mathrm{cont}}(G, M), \mathcal{O}_{X'}) \Longrightarrow \mathrm{H}^{i+j}_{\mathrm{cont}}(G, f^*M).$$

(3) *If f is flat, then the natural map*

$$f^*\mathrm{H}^*_{\mathrm{cont}}(G, M) \to \mathrm{H}^*_{\mathrm{cont}}(G, f^*M)$$

is an isomorphism in $\mathbf{Gr}^{[0,e]}_{\mathrm{coh}}(X')$.

Proof. After clearing away the abstract nonsense using (1-2), this is just an application of Theorem 1.4 to $A[1/p] = \Gamma(Y, \mathcal{O}_Y)$ and $B[1/p] = \Gamma(Y', \mathcal{O}_{Y'})$ for each pair of affinoids $Y \in \mathcal{U}$ and $Y' \in \mathcal{U}'$ with $f(Y') \subseteq Y$, using Kiehl's theorem to patch back up. □

1C. *Quasi-Stein spaces.* Continuing with X and M as in Section 1B, assume that X is *quasi-Stein*: it admits an admissible covering $\mathcal{U}$ by an increasing union of strictly E-affinoid subdomains $Y_1 \subseteq Y_2 \subseteq \cdots$, each of whose restriction maps $\Gamma(Y_{n+1}, \mathcal{O}_{Y_{n+1}}) \to \Gamma(Y_n, \mathcal{O}_{Y_n})$ has dense image.

Let M_n be the $\Gamma(Y_n, \mathcal{O}_{Y_n})$-module $\Gamma(Y_n, M)$. The ring $A_\infty = \varprojlim_n \Gamma(Y_n, \mathcal{O}_{Y_n}) = \Gamma(X, \mathcal{O}_X)$ is a commutative *Fréchet–Stein algebra*, and $M_\infty = \varprojlim_n M_n = \Gamma(X, M)$ is a *coadmissible* A_∞-module in the sense of [Schneider and Teitelbaum 2003, §3]. "Theorem A" for such modules states that the natural maps $M_\infty \to M_n$ have dense image; they induce isomorphisms $M_\infty \otimes_{A_\infty} \Gamma(Y_n, \mathcal{O}_{Y_n}) \xrightarrow{\sim} M_n$. Also, this denseness suffices for Mittag-Leffler considerations, whence "Theorem B" states that for all $i > 0$ one has $\mathbf{R}^i \varprojlim_n M_n = 0$. We obtain an exact equivalence between coherent sheaves on X and coadmissible modules over A_∞. In particular, the subcategory of all A_∞-modules consisting of the coadmissible ones forms an abelian subcategory that is closed under extensions. We say that an A_∞-module satisfies condition "ft" if its coadmissible.

Turning to cohomology, each of the maps $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M_{n+1}) \to \mathrm{C}^\bullet_{\mathrm{cont}}(G, M_n)$ has dense image, and $\mathrm{C}^\bullet_{\mathrm{cont}}(G, M_\infty) = \varprojlim_n \mathrm{C}^\bullet_{\mathrm{cont}}(G, M_n)$ by the definition of the inverse limit, so Mittag-Leffler gives $\mathrm{H}^*_{\mathrm{cont}}(G, M_\infty) = \varprojlim_n \mathrm{H}^*_{\mathrm{cont}}(G, M_n)$. Hence, $\mathrm{C}^*_{\mathrm{cont}}(G, M_\infty)$ is an object of $\mathbf{K}^{[0,e]}_{\mathrm{ft}}(A_\infty)$, where the subscript means that the cohomology modules are required to be coadmissible.

The following theorem follows easily from the preceding discussion:

Theorem 1.7. *The natural maps*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G, M_\infty) \to \mathbf{R}\varprojlim_n[\mathrm{C}^\bullet_{\mathrm{cont}}(G, M_n)],$$

$$\mathrm{H}^*_{\mathrm{cont}}(G, M_\infty) \to \varprojlim_n \mathrm{H}^*_{\mathrm{cont}}(G, M_n),$$

$$\mathrm{H}^*_{\mathrm{cont}}(G, M_\infty) \otimes_{A_\infty} \Gamma(Y_n, \mathcal{O}) \to \mathrm{H}^*_{\mathrm{cont}}(G, M_n)$$

are isomorphisms.

Remark 1.8. The usual explicit construction shows that, for any M and M' and $n \leq \infty$, the group cohomology $\mathrm{H}^1_{\mathrm{cont}}(G, \mathrm{Hom}_{\Gamma(Y_n,\mathbb{O})}(M_n, M'_n))$ is in canonical bijection with the Yoneda group $\mathrm{Ext}^1_{\Gamma(Y_n,\mathbb{O})[G]\text{-cont}}(M_n, M'_n)$ of extensions of M'_n by M_n in the category of continuous $\Gamma(Y_n, \mathbb{O})[G]$-modules.

1D. ***Generic fibers of formal spectra.*** In this section, we let A and M satisfy Hypotheses A. Assume for this discussion that A/I is a finitely generated $\mathbb{O}_E$-algebra and that A is p-torsion-free. We now compare the group cohomology of M to that of its generic fiber.

Let $A_n^0 = A[I^n/p]$, the A-subalgebra of $A[1/p]$ generated by all i/p with $i \in I^n$, and let A_n be its p-adic completion. Each $A_n[1/p]$ is a strictly E-affinoid algebra, and maps $A_{n+1}[1/p] \to A_n[1/p]$ arising from the inclusions $A_{n+1}^0 \subseteq A_n^0$ correspond to inclusions $Y_n \subseteq Y_{n+1}$ of affinoid subdomains. The increasing system $Y_1 \subseteq Y_2 \subseteq \cdots$ forms an admissible affinoid covering of its union X. (See [de Jong 1995, §7.1] for details.) It is clear that X is a quasi-Stein space so that Section 1C applies.

The powers of the ideals $pA_n^0, IA_n^0 \subseteq A_n^0$ are cofinal so that A_n is also the I-adic completion of A_n^0. In particular, each A_n is an I-adic A-algebra. Each $A_n[1/p]$ is flat over $A[1/p]$. The following base-changing theorem now follows easily from the preceding work:

Theorem 1.9. *For $n \leq \infty$, the natural maps*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \otimes^{\mathbf{L}}_A A_n \to \mathbf{R}\Gamma_{\mathrm{cont}}(G, M \otimes_A A_n),$$
$$\mathrm{H}^*_{\mathrm{cont}}(G, M) \otimes_A A_n[1/p] \to \mathrm{H}^*_{\mathrm{cont}}(G, M \otimes_A A_n[1/p])$$

are isomorphisms.

Remark 1.10. By Remark 1.8 and Theorem 1.9, for $n \leq \infty$, the map on Yoneda groups $\mathrm{Ext}^1_{A[G]\text{-cont}}(M, M') \to \mathrm{Ext}_{A_n[1/p][G]\text{-cont}}(M \otimes_A A_n[1/p], M' \otimes_A A_n[1/p])$ determined by applying $\otimes_A A_n[1/p]$ to an extension class $0 \to M' \to E \to M \to 0$ induces an isomorphism

$$\mathrm{Ext}^1_{A[G]\text{-cont}}(M, M') \otimes_A A_n \xrightarrow{\sim} \mathrm{Ext}_{A_n[1/p][G]\text{-cont}}(M \otimes_A A_n[1/p], M' \otimes_A A_n[1/p]).$$

Morally, $A[1/p] \subseteq A_\infty$ consists of the p-adically bounded functions on X, so calling the image of $\mathrm{Ext}^1_{A[G]\text{-cont}}(M, M')[1/p]$ in $\mathrm{Ext}^1_{A_\infty[G]\text{-cont}}(M \otimes_A A_\infty, M' \otimes_A A_\infty)$ the *bounded* extension classes is reasonable.

1E. ***Selmer complexes.*** We now copy ideas of Nekovář [2006] into the context of the preceding sections.

In the preceding sections, we describe a variety of situations all of the following sort: one is given a profinite group G satisfying Hypothesis (1), a ringed topos Z built from p-adic rings, and a locally finitely generated flat $\mathbb{O}_Z$-module M with continuous $\mathscr{O}_Z[G]$-module structure. In each situation, we show that the continuous cohomology objects satisfy $\mathbf{C}^\bullet_{\mathrm{cont}}(G, M) \in \mathbf{K}^{[0,e]}_{\mathrm{ft}}(Z)$, $\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \in \mathbf{D}^{[0,e]}_{\mathrm{perf}}(Z)$, and $\mathrm{H}^*_{\mathrm{cont}}(G, M) \in \mathbf{Gr}^{[0,e]}_{\mathrm{ft}}(Z)$ and that their formation commutes with (derived) pullback along appropriate morphisms of topoi $f : Z' \to Z$ and in a certain case pushforward as well. Specifically, we treat the following cases, using a tilde to denote the associated ringed topos:

(1) (A, I) and M satisfy Hypotheses A(2)–(4), and B is an I-adic A-algebra. We define $Z = (\operatorname{Spec} A)^\sim$ and $Z' = (\operatorname{Spec} B)^\sim$, and let f be the induced morphism.

(2) $M = M_0[1/p]$ where (A, I) and M_0 satisfy Hypotheses A (2)–(4), and B is an I-adic A-algebra. We define $Z = (\operatorname{Spec} A[1/p])^\sim$ and $Z' = (\operatorname{Spec} B[1/p])^\sim$, and let f be the induced morphism.

(3) X is a p-adic analytic space over E, M is a family of G-representations over X, $\mathscr{U}$ is an admissible affinoid covering of X that is quasiclosed under intersection, $f_0 : X' \to X$ is a morphism of p-adic analytic spaces over E, and $\mathscr{U}'$ is an admissible affinoid covering of X' that is quasiclosed under intersection with the property that for each $Y' \in \mathscr{U}'$ there exists $Y \in \mathscr{U}$ with $f_0(Y') \subseteq Y$. We define $Z = (\mathscr{U}, \mathscr{O}_{\mathscr{U}})^\sim$ and $Z' = (\mathscr{U}', \mathscr{O}_{\mathscr{U}'})^\sim$, and let $f = \tilde{f}_0$.

(4) X is a quasi-Stein p-adic analytic space over E, M is a family of G-representations over X, and $\mathscr{U} = \{Y_n\}_{n \geq 1}$ is an increasing admissible affinoid covering as in Section 1C. We define $Z = (\operatorname{Spec} \Gamma(X, \mathscr{O}_X))^\sim$ and $Z' = (\mathscr{U}, \mathscr{O}_{\mathscr{U}})^\sim$, and let f be the induced morphism, interpreting "finite type" over Z to mean "coadmissible". In this case, the formation of cohomology also commutes with (derived) pushforward along f.

(5) (A, I) and M satisfy Hypotheses A(2)–(4), and moreover, A is p-torsion-free and A/I is a finitely generated $\mathscr{O}_E$-algebra. We define $Z = (\operatorname{Spec} A)^\sim$, the rings A_n for $n \leq \infty$ as in Section 1D, and $Z' = (\operatorname{Spec} A_n[1/p])^\sim$, and let f be the induced morphism, in the case $n = \infty$ interpreting "finite type" over Z' to mean "coadmissible".

In this section, we specialize to the case of the group $G = G_{K,S}$ defined in the next paragraph and give analogous results in each of these scenarios where the continuous cohomology objects have been replaced by Selmer complexes relative to appropriate local conditions. We also give variants of Tate's local and Poitou–Tate's global arithmetic duality theorems for these objects.

We let K be a finite extension of $\mathbf{Q}$, and we let S be a finite set of finite places of K containing all v dividing p. We choose a maximal algebraic extension K_S of K unramified outside $S \cup \{\infty\}$, and put $G_{K,S} = \mathrm{Gal}(K_S/K)$. We also choose, for each $v \in S$, an algebraic closure K_v^{alg} of K_v together with a K-algebra embedding $K_S \hookrightarrow K_v^{\mathrm{alg}}$, and put $G_v = \mathrm{Gal}(K_v^{\mathrm{alg}}/K_v)$. Denote by $I_v \subset G_v$ the inertia subgroup.

We write res_v for the map $G_v \to G_{K,S}$ given by restriction along our chosen embedding as well as for the map $\mathrm{C}^\bullet_{\mathrm{cont}}(G_{K,S}, M) \to \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M)$ it induces, via pullback, on cocycles. Assume for simplicity that $p \neq 2$ or K is totally complex. By arithmetic duality theory, the groups $G_{K,S}$ and G_v for $v \in S$ satisfy Hypothesis (1) with $e = 2$ (see [Neukirch et al. 2008, 8.3.10, 8.3.17–19, and 7.1.8]). In the exceptional case where $p = 2$ and K is not totally complex, one can get a similar theory working with a little more care; see [Nekovář 2006, 5.7], and note that the complication is annihilated by inverting 2 anyway.

We place ourselves in one of the scenarios (1)–(5) above, where the hypotheses are made relative to the group $G = G_{K,S}$. By a ($*$-*bounded*) *local condition* Δ_v at $v \in S$, we mean the data of an object $U_v^\bullet \in \mathbf{K}^*_{\mathrm{ft}}(Z)$ and a morphism $i_v : U_v^\bullet \to \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M)$. Assume we are given the data $\Delta = \{\Delta_v\}_{v\in S}$ of a local condition Δ_v for each $v \in S$. We define the *Selmer complex* $\widetilde{\mathrm{C}}^\bullet_{\mathrm{f}}(G_{K,S}, M; \Delta)$ of M with respect to Δ to be the complex

$$\mathrm{Cone}\left[\mathrm{C}^\bullet_{\mathrm{cont}}(G_{K,S}, M) \oplus \bigoplus_{v\in S} U_v^\bullet \xrightarrow{\bigoplus_{v\in S}(\mathrm{res}_v - i_v)} \bigoplus_{v\in S} \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M)\right][-1].$$

We denote by $\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta)$ the image of the Selmer complex in the derived category, and we denote its cohomology groups, which we call the (*extended*) *Selmer groups*, by $\widetilde{\mathrm{H}}^*_{\mathrm{f}}(G_{K,S}, M; \Delta)$. For brevity, we usually suppress the dependence on Δ from the notation. By the definition of the extended Selmer groups in terms of a mapping cone, one has an exact triangle

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta) \to \mathbf{R}\Gamma_{\mathrm{cont}}(G_{K,S}, M) \to \bigoplus_{v\in S} E_v, \tag{1-3}$$

where the objects $E_v = \mathrm{Cone}(i_v)$ sit in exact triangles

$$U_v \xrightarrow{i_v} \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M) \xrightarrow{j_v} E_v. \tag{1-4}$$

Thus, the image of the extended Selmer group in $\mathrm{H}^i_{\mathrm{cont}}(G_{K,S}, M)$ consists of those classes that everywhere locally live in the image of the $\mathrm{H}^i(i_v)$; this image is what one more traditionally encounters in the literature, so we call it the *nonextended Selmer group* $\mathrm{H}^i_{\mathrm{f}}(G_{K,S}, M; \Delta)$.

The following finiteness theorem is just an application to each of $G = G_{K,S}$ and $G = G_v$ of the finiteness theorems of the preceding sections in light of the exact triangles (1-3) and (1-4):

Theorem 1.11. *The complex* $\widetilde{\mathrm{C}}^\bullet_{\mathrm{f}}(G_{K,S}, M; \Delta)$ *belongs to* $\mathbf{K}^*_{\mathrm{ft}}(Z)$. *In particular*, $\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta) \in \mathbf{D}^*_{\mathrm{ft}}(Z)$ *and* $\widetilde{\mathrm{H}}^*_{\mathrm{f}}(G_{K,S}, M; \Delta) \in \mathbf{Gr}^*_{\mathrm{ft}}(Z)$. *If for each* $v \in S$ *one has* $[U_v] \in \mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$, *then* $\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta)$ *belongs to* $\mathbf{D}^{[0,3]}_{\mathrm{perf}}(Z)$.

We turn to treat base change. If the local conditions are bounded above (up to quasi-isomorphism), on the one hand we may form $\mathbf{L}f^*\widetilde{\mathrm{C}}^\bullet_{\mathrm{f}}(G_{K,S}, M; \Delta)$. On the other hand, for $v \in S$, we may form the local condition $f^*\Delta_v$ for f^*M by choosing a representative in $\mathbf{K}^-_{\mathrm{ft}}(Z')$ of the morphism

$$\mathbf{L}f^*U_v \xrightarrow{\mathbf{L}f^*i_v} \mathbf{L}f^*\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M) \xrightarrow{\sim} \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, f^*M),$$

and we write $f^*\Delta = \{f^*\Delta_v\}_{v\in S}$. The following theorem is similarly deduced from the finiteness and base-changing theorems of the preceding sections:

Theorem 1.12. *In the situations (1)–(5) above, assume the local conditions Δ are bounded above. Then the natural map*

$$\mathbf{L}f^*\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta) \to \mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, f^*M; f^*\Delta)$$

is an isomorphism in $\mathbf{D}^-_{\mathrm{ft}}(Z')$.

If for each $v \in S$ one has $[U_v] \in \mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$, then the isomorphism takes place in $\mathbf{D}^{[0,3]}_{\mathrm{perf}}(Z')$.

In the case (4) above, so X is quasi-Stein with increasing admissible affinoid covering $\mathcal{U} = \{Y_n\}_{n\geq 1}$, there is also a pushforward result. Assume we are given, instead of local conditions on the global sections $\Gamma(X, M)$ as in the preceding theorem, local conditions on Δ' on M considered as a sheaf on $\mathcal{U}$. Thus, we have a quasiconsistent family of morphisms $\Gamma(Y_n, U'^\bullet_v) \to \Gamma(Y_n, \mathrm{C}^\bullet_{\mathrm{cont}}(G, M))$ for varying n. Assume that the maps $\Gamma(Y_{n+1}, U'^\bullet_v) \to \Gamma(Y_n, U'^\bullet_v)$ have *dense image*. We form local conditions Δ for $\Gamma(X, M)$ using the morphisms

$$i^\bullet_v : U^\bullet_v = \varprojlim_n \Gamma(Y_n, U'^\bullet_v) \to \varprojlim_n \Gamma(Y_n, \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M)) = \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, \Gamma(X, M))$$

for $v \in S$. Then $\mathrm{H}^*(U^\bullet_v) = \varprojlim_n \mathrm{H}^*\Gamma(Y_n, U'^\bullet_v)$ by our dense image assumption and Mittag-Leffler, and it follows that the $\Gamma(X, \mathcal{O}_X)$-modules on the left-hand side are coadmissible. The following theorem is again a consequence of the finiteness and pushforward results of Section 1C:

Theorem 1.13. *In the situation (4) above, assume the local conditions Δ' are bounded above with transition maps having dense image. Then the natural map*

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, \Gamma(X, M); \Delta) \to \mathbf{R}f_*\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta')$$

is an isomorphism in $\mathbf{D}^-_{\mathrm{coadm}}(Z)$.

If for each $v \in S$ one has $[U_v] \in \mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z')$, then the isomorphism takes place in $\mathbf{D}^{[0,3]}_{\mathrm{perf}}(Z)$.

Next we treat arithmetic duality. Recall the anti-involution on perfect complexes $X \mapsto X^* = \mathbf{R}\operatorname{Hom}_{\mathcal{O}_Z}(X, \mathcal{O}_Z) = \operatorname{Hom}_{\mathcal{O}_Z}(X, \mathcal{O}_Z)$. What follows is the basic local result, a variant in families of Tate's local duality theorem.

Theorem 1.14. (1) *For any $v \in S$, there is a canonical isomorphism*

$$\tau_{\geq 2}\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, \mathbb{O}_Z(1)) \xrightarrow{\sim} \mathbb{O}_Z[-2]$$

given by base change of the local trace map.

(2) *For any $v \in S$, the duality morphism*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M^*(1)) \to \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M)^*[-2]$$

adjoint to the pairing

$$\begin{aligned}\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M^*(1)) \otimes^{\mathbf{L}}_{\mathbb{O}_Z} \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M) &\xrightarrow{\cup} \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M^*(1) \otimes_{\mathbb{O}_Z} M)\\ &\to \tau_{\geq 2}\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, \mathbb{O}_Z(1)) \xrightarrow{\sim} \mathbb{O}_Z[-2],\end{aligned}$$

given by cup product, evaluation and truncation, and (1) above, is an isomorphism in $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$.

Proof. To see (1), we note that, because G_v satisfies $e = 2$, the Tor-spectral sequence shows that the rule $M \mapsto \mathrm{H}^2_{\mathrm{cont}}(G_v, M)$ commutes with arbitrary base change in Z. Thus, it suffices to take the composition

$$\begin{aligned}\tau_{\geq 2}\mathbf{R}\Gamma_{\mathrm{cont}}(G_v, \mathbb{O}_Z(1)) &\cong \mathrm{H}^2_{\mathrm{cont}}(G_v, \mathbb{O}_Z(1))[-2]\\ &\cong (\mathrm{H}^2_{\mathrm{cont}}(G_v, \mathbf{Z}_p(1)) \otimes_{\mathbf{Z}_p} \mathbb{O}_Z)[-2]\\ &\cong (\mathbf{Z}_p \otimes_{\mathbf{Z}_p} \mathbb{O}_Z)[-2] = \mathbb{O}_Z[-2],\end{aligned}$$

the last identification coming from the trace isomorphism of local class field theory.

To treat (2), we observe that the formation of the duality morphism commutes with arbitrary derived base change in Z, and all cases under consideration can be reduced to local scenarios that are pullbacks of the situation (1) on page 1584; hence, it suffices to assume we are in that specific case with (A, I) the ring in question. A morphism of perfect complexes is a quasi-isomorphism if and only if it becomes a quasi-isomorphism after applying $\otimes^{\mathbf{L}}_A A/\mathfrak{m}$ for any maximal ideal $\mathfrak{m}$ of A. By executing this base change and noting by [Matsumura 1989, Theorem 8.2(i)] that $I \subseteq \mathfrak{m}$, we are reduced to the case where A is a field of characteristic p with the discrete topology. But then G_v acts on M via a finite quotient, and the situation arises as the base change of a situation with coefficients in a finite field, where the result is known. □

The global result is more complicated to state. As a first approximation, we introduce *compactly supported cochains* $\mathrm{C}^\bullet_{\mathrm{cont,c}}(G_{K,S}, M)$ as the complex

$$\mathrm{Cone}\left[\mathrm{C}^\bullet_{\mathrm{cont}}(G_{K,S}, M) \xrightarrow{\bigoplus_{v \in S} \mathrm{res}_v} \bigoplus_{v \in S} \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M)\right][-1],$$

denoting its image in the derived category by $\mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M) \in \mathbf{D}^{[1,3]}_{\mathrm{perf}}(Z)$ and its cohomology by $\mathrm{H}^*_{\mathrm{cont,c}}(G_{K,S}, M) \in \mathbf{K}^{[1,3]}_{\mathrm{ft}}(Z)$. Following Nekovář [2006, 5.3.3.3], one constructs cup product pairings

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G_{K,S}, M) \otimes^{\mathbf{L}}_{\mathbb{O}_Z} \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M') \to \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M \otimes_{\mathbb{O}_Z} M'),$$

and the following theorem is proved in the exact same way as the preceding one:

Theorem 1.15. (1) *There is a canonical isomorphism*

$$\tau_{\geq 3}\mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, \mathbb{O}_Z(1)) \xrightarrow{\sim} \mathbb{O}_Z[-3]$$

given by base change of the global trace map.

(2) *For any $v \in S$, the duality morphism*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G_{K,S}, M^*(1)) \to \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M)^*[-3]$$

adjoint to the pairing

$$\begin{aligned}\mathbf{R}\Gamma_{\mathrm{cont}}(G_{K,S}, M^*(1)) \otimes^{\mathbf{L}}_{\mathbb{O}_Z} \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M)& \\ \xrightarrow{\cup} \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M^*(1) \otimes_{\mathbb{O}_Z} M)& \\ \to \tau_{\geq 3}\mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, \mathbb{O}_Z(1)) \xrightarrow{\sim} \mathbb{O}_Z[-3]&,\end{aligned}$$

given by cup product, evaluation and truncation, and (1) above, is an isomorphism in $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$.

To treat duality of Selmer complexes, assume the local conditions Δ satisfy $[U_v] \in \mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$ for all $v \in S$, and equip $M^*(1)$ with local conditions $\Delta^*(1)$ given for $v \in S$ by choosing a representative in $\mathbf{K}^{[0,2]}_{\mathrm{perf}}(Z)$ of the morphism

$$E^*_v[-2] \xrightarrow{j^*_v[-2]} \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M)^*[-2] \cong \mathbf{R}\Gamma_{\mathrm{cont}}(G_v, M^*(1)).$$

Then as in [Nekovář 2006, 6.3], one constructs cup product pairings

$$\begin{aligned}\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M^*(1); \Delta^*(1)) \otimes^{\mathbf{L}}_{\mathbb{O}_Z} \mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta)& \\ \to \mathbf{R}\Gamma_{\mathrm{cont,c}}(G_{K,S}, M^*(1) \otimes_{\mathbb{O}_Z} M)&,\end{aligned}$$

which, followed by evaluation and truncation, followed by the global trace map, gives rise via adjoint to a duality morphism

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M^*(1); \Delta^*(1)) \to \mathbf{R}\widetilde{\Gamma}_{\mathrm{f}}(G_{K,S}, M; \Delta)^*[-3].$$

The general global result is as follows, with the same proof.

Theorem 1.16. *Assume, for all $v \in S$, the local conditions satisfy $[U_v] \in \mathbf{D}^{[0,2]}_{\mathrm{perf}}(Z)$. Then the duality morphism above is an isomorphism in $\mathbf{D}^{[0,3]}_{\mathrm{perf}}(Z)$.*

Remark 1.17. The preceding base-changing and duality theorems are cheating because we have taken a narrow definition of "local condition". Rather than a particular morphism i_v, the phrase usually means a *rule* that associates such a morphism to a given M (perhaps equipped with additional data) as in the examples below. The true base-changing and duality theorems are reduced by the above theorems to the claim that the formation of the local conditions, as determined by the rule, commutes with base change and duality, respectively.

Example 1.18. We give some examples of useful local conditions. Let $v \in S$.

The *empty* local condition means taking i_v to be the identity map and results in no modification being made to the cohomology. The *full* local condition, taking i_v to be the map from the zero object, results in cohomology that is "compactly supported" at v. The formation of these conditions clearly commutes with arbitrary base change. Excepting these two, a local condition often has the property that $\mathrm{H}^0(i_v)$ is an isomorphism, $\mathrm{H}^1(i_v)$ is an injection, and $\mathrm{H}^n(U_v) = 0$ for $n \neq 0, 1$. In fact, given a subspace $\mathscr{L} \subseteq \mathrm{H}^1_{\mathrm{cont}}(G_v, M)$, there is a standard construction of such a local condition with $\operatorname{img} \mathrm{H}^1(i_v) = \mathscr{L}$, called the local condition *associated to* $\mathscr{L}$, namely by setting $U_v^0 = \mathrm{C}^0_{\mathrm{cont}}(G_v, M)$,

$$U_v^1 = \{c \in \mathrm{C}^1_{\mathrm{cont}}(G_v, M) \mid dc = 0,\ [c] \in \mathscr{L}\},$$

and $U_v^n = 0$ for $n \neq 0, 1$ (and taking i_v to be the inclusion map of complexes). On the other hand, when $\mathrm{H}^1(i_v)$ fails to be injective, its kernel is considered a local contribution to an "exceptional zero" of the related p-adic L-function.

For $v \in S$ not dividing p, the *unramified* local condition is given by inflation

$$i_v : U_v^\bullet = \mathrm{C}^\bullet_{\mathrm{cont}}(G_v/I_v, M^{I_v}) \to \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M).$$

(Note that $G_v/I_v \approx \widehat{\mathbf{Z}}$ satisfies Hypothesis (1) with $e = 1$.) It is isomorphic in the derived category to the local condition associated to the image of inflation $\mathrm{H}^1_{\mathrm{cont}}(G_v/I_v, M^{I_v}) \hookrightarrow \mathrm{H}^1_{\mathrm{cont}}(G_v, M)$. Whenever M^{I_v} is flat over $\mathscr{O}_Z$, it obeys the necessary hypotheses as a continuous $\mathscr{O}_Z[G_v/I_v]$-module. If $f^*(M^{I_v}) \xrightarrow{\sim} (f^*M)^{I_v}$, then the formation of the unramified local condition commutes with (derived) base change. When both M^{I_v} and $M^*(1)^{I_v}$ are flat over $\mathscr{O}_Z$, it makes sense to ask whether the unramified local conditions for M and $M^*(1)$ are self-dual, and this seems to be the case only generically: see, for example, [Nekovář 2006, 7.6 and 7.6.7(iii)] for closely related statements in the situation (1) on page 1584 with I a maximal ideal (where our *-operation has been replaced by Grothendieck duality, which is not a change when A is Gorenstein).

We warn the reader that in general M^{I_v} need *not* be flat, in which case its finiteness and base-changing properties become much more subtle. There is no problem for families of twists of a fixed global Galois representation; for example, in Example 1.19(1) below, one has $M^{I_v} = T^{I_v} \otimes_{\mathbf{Z}_p} A$. But the complication does arise in Hida theory, say in Example 1.19(2) below, when a p-ordinary eigenform admits level lowering modulo p at a prime $v = \ell \neq p$. Then there are ℓ-old and ℓ-new branches of the Hida family, over which M^{I_v} has generic rank, respectively, two and strictly less than two. Since geometrically these branches meet in the special fiber of Spec $\mathfrak{h}_\infty^{\mathrm{ord}}$, the module M^{I_v} cannot be locally free. This phenomenon is related to the appearance of p in the Tamagawa number at ℓ because the latter occurrence can be used to detect level-lowering.

For v dividing p, most local conditions are rather complicated; even the determination of a meaningful subspace $\mathscr{L}$ to which to associate them is rather delicate. The best-behaved notion is the (*strict*) *ordinary* one: one assumes given a G_v-stable locally direct summand $M_v^+ \subseteq M$ and takes Δ_v to be the data of the natural map

$$i_v : U_v^\bullet = \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M_v^+) \to \mathrm{C}^\bullet_{\mathrm{cont}}(G_v, M).$$

This local condition appears in work of Nekovář [2006, 6.7] as a derived variant of the subspace $\mathscr{L}$ given by the image of $\mathrm{H}^1_{\mathrm{cont}}(G_v, M_v^+) \to \mathrm{H}^1_{\mathrm{cont}}(G_v, M)$ used by Greenberg [1989; 1994a; 1994b]; they are frequently isomorphic in the derived category for example as in Proposition 3.7(3) below. The formation of the ordinary local condition commutes with base change and is dual to the ordinary local condition formed with the annihilator of M_v^+ in $M^*(1)$. The abstract base change and duality theorems recover Greenberg's control and duality theorems in the situation (1), with A local with finite residue field, and gives analogues of it in all cases. All this presupposes the existence of a useful choice of M_v^+ to begin with; the key observation of this article is that, after replacing the G_v-module M by its associated (φ, Γ_{K_v})-module, one can still form an ordinary local condition, and one gains access to subobjects of $\mathbf{D}_{\mathrm{pst}}(M)$ that are not necessarily weakly admissible.

Example 1.19. Take $K = \mathbf{Q}$ for simplicity. The present results, notably situation (5), apply to the following settings. Each A is local, we take I to be its maximal ideal, and A/I is a finite field. We write X for the generic fiber of $\mathfrak{X} = \mathrm{Spf}(A, I)$.

(1) A is the Iwasawa algebra $\mathbf{Z}_p[\![\Gamma]\!]$, where $\Gamma = \mathrm{Gal}(\mathbf{Q}_\infty/\mathbf{Q})$ is the Galois group of the cyclotomic $\mathbf{Z}_p$-extension $\mathbf{Q}_\infty/\mathbf{Q}$, and $M = T \otimes_{\mathbf{Z}_p} A$ with diagonal $G_{\mathbf{Q},S}$-action, where T is a continuous $\mathbf{Z}_p[G_{\mathbf{Q},S}]$-module that is free over $\mathbf{Z}_p$. Thus, M is the cyclotomic deformation of T, and the parameter space X is commonly called the weight space.

(2) A is a Hida–Hecke algebra $\mathfrak{h}_\infty^{\mathrm{ord}}$, and M is Hida's Λ-adic Galois representation (with A assumed Gorenstein so that M is flat). Thus, X is commonly called the ordinary locus of the Coleman–Mazur eigencurve.

(3) A is any $\mathbf{Z}_p$-flat quotient of a Galois deformation ring $R^{\mathrm{univ}}(\bar{\rho})$, where $\bar{\rho}$ is an absolutely irreducible mod p representation of $G_{\mathbf{Q},S}$, and M is deduced from the universal deformation of $\bar{\rho}$.

Traditionally, the algebraic part of Iwasawa theory concerns the study of Selmer groups over A in the above examples. Theorem 1.9 translates this study to X, with the loss of only p-torsion information, as $\Gamma(X, \mathbb{O}_X)$ is faithfully flat over $A[1/p]$. Since (φ, Γ)-modules over the Robba ring (which we use to generalize Greenberg's "ordinary" theory) form families over X and not over $\mathfrak{X}$, this change of view makes it now possible to treat nonordinary situations.

2. (φ, Γ)-modules and Galois cohomology

In this section, we fix a finite extension K of $\mathbf{Q}_p$ with ring of integers $\mathbb{O}_K$ and residue field k. We choose an algebraic closure K^{alg} of K and set $G = G_K = \mathrm{Gal}(K^{\mathrm{alg}}/K)$. (Several of the techniques discussed below will be valid if K is replaced by a general complete, discretely valued field of mixed characteristic $(0, p)$ and perfect residue field, but beware that the group G satisfies Hypothesis (1) only when K is finite over $\mathbf{Q}_p$, which gets in the way of certain base-change arguments.)

We let A' be a Noetherian commutative E-Banach algebra, having A as its unit ball, so that $A' = A[1/p]$ and A satisfies Hypothesis (2) when equipped with the ideal $I = (p)$. Finally, we let M satisfy Hypotheses A(3)–(4).

2A. *Recall of φ- and (φ, Γ)-modules.* There are several variants of (φ, Γ_K)-modules, so we must recall several base rings. We only do this minimally since they are defined in many places now (see, e.g., [Berger 2002]). For any field F, write $F_n = F(\mu_{p^n})$ for $n \leq \infty$.

Let $F = \mathrm{Frac}\, \mathrm{W}(k)$. If k' denotes the residue field of K_∞, define $F' = \mathrm{Frac}\, \mathrm{W}(k')$ and $K' = K.F'$. We set $H = H_K = \mathrm{Gal}(K^{\mathrm{alg}}/K_\infty)$ and $\Gamma = \Gamma_K = \mathrm{Gal}(K_\infty/K)$. The group Γ is either procyclic or of the form $\{\pm 1\} \times$ (procyclic); we let $\Delta = \Delta_K \subset \Gamma$ be trivial in the first case and $\{\pm 1\}$ in the second case.

There is an increasing system of p-torsion-free, p-adically separated, and complete $\mathrm{W}(k^{\mathrm{alg}})$-algebras $\widetilde{\mathrm{A}}^{\dagger,s}$ equipped with a compatible action of G, indexed by real numbers $s > 0$. This family is also equipped with an automorphism φ that commutes with G and takes $\widetilde{\mathrm{A}}^{\dagger,s}$ onto $\widetilde{\mathrm{A}}^{\dagger,ps}$. Each of the p-adic Banach algebras $\widetilde{\mathrm{B}}^{\dagger,s} = \widetilde{\mathrm{A}}^{\dagger,s}[1/p]$ admits a certain Fréchet completion $\widetilde{\mathrm{B}}^{\dagger,s}_{\mathrm{rig}}$, to which both actions extend uniquely by continuity, and these latter rings also fit into an increasing system. One defines the systems $\widetilde{\mathrm{A}}^{\dagger,s}_K = (\widetilde{\mathrm{A}}^{\dagger,s})^H$, $\widetilde{\mathrm{B}}^{\dagger,s}_K = (\widetilde{\mathrm{B}}^{\dagger,s})^H$, and $\widetilde{\mathrm{B}}^{\dagger,s}_{\mathrm{rig},K} = (\widetilde{\mathrm{B}}^{\dagger,s}_{\mathrm{rig}})^H$ with the induced topologies and actions of φ and Γ.

The theory of the field of norms allows one to make a choice of a sort of indeterminate π_K belonging to all the $\widetilde{\mathrm{A}}^{\dagger,s}_K$ and associates to K a constant $\mathrm{e}_K > 0$. When $K = F$, there is an almost canonical choice that is written π, and one can calculate that $\mathrm{e}_F = p/(p-1)$. For $s > 0$, one has subrings

$$\mathrm{B}_K^{\dagger,s} = \left\{ f(\pi_K) = \sum_{n\in\mathbf{Z}} a_n \pi_K^n \;\middle|\; \begin{array}{c} a_n \in F',\ \{|a_n|\} \text{ bounded}, \\ f(X) \text{ convergent for } 0 < \mathrm{ord}_p(X) < 1/\mathrm{e}_K s \end{array} \right\},$$

$$\mathrm{B}_{\mathrm{rig},K}^{\dagger,s} = \left\{ f(\pi_K) = \sum_{n\in\mathbf{Z}} a_n \pi_K^n \;\middle|\; \begin{array}{c} a_n \in F', \\ f(X) \text{ convergent for } 0 < \mathrm{ord}_p(X) < 1/\mathrm{e}_K s \end{array} \right\}$$

of $\widetilde{\mathrm{B}}_K^{\dagger,s}$ and $\widetilde{\mathrm{B}}_{\mathrm{rig},K}^{\dagger,s}$, respectively, that do not depend on the choice of π_K for $s \gg 0$. They inherit topologies, and for $s \gg 0$, they are stable under Γ, and φ sends $\mathrm{B}_K^{\dagger,s}$ into $\mathrm{B}_K^{\dagger,ps}$ and $\mathrm{B}_{\mathrm{rig},K}^{\dagger,s}$ into $\mathrm{B}_{\mathrm{rig},K}^{\dagger,ps}$. One knows that φ acts by Witt functoriality on $a_n \in F'$, and Γ acts on a_n through its quotient $\Gamma_K/\Gamma_{K'} = \mathrm{Gal}(F'/F)$. The action on π_K is generally not explicitly given (especially since there is some choice in π_K) except when $K = F$, in which case $\varphi(\pi) = (1+\pi)^p - 1$ and $\gamma \in \Gamma$ obeys $\gamma(\pi) = (1+\pi)^{\chi_{\mathrm{cycl}}(\gamma)} - 1$. In any case, φ now induces a finite free algebra extension of degree p for $s \gg 0$ instead of being an isomorphism. For such s, we obtain a left inverse $\psi : \mathrm{B}_{(\mathrm{rig},)K}^{\dagger,ps} \to \mathrm{B}_{(\mathrm{rig},)K}^{\dagger,s}$ to φ by the formula $p^{-1}\varphi^{-1} \circ \mathrm{Tr}_{\mathrm{B}_{(\mathrm{rig},)K}^{\dagger,ps}/\varphi \mathrm{B}_{(\mathrm{rig},)K}^{\dagger,s}}$.

For any of the above families of rings, we denote the result of applying $\varinjlim_s$ by omitting the index s, e.g., $\mathrm{B}_{\mathrm{rig},K}^{\dagger} = \varinjlim_s \mathrm{B}_{\mathrm{rig},K}^{\dagger,s}$. The result inherits the direct limit topology, an action of G, a ring endomorphism φ, and a left inverse ψ. The ring $\mathrm{B}_K^{\dagger}$ is the fraction field of a Henselian, mixed-characteristic discrete valuation ring with imperfect residue field (although we never make use of the topology it would provide). Although the rings $\mathrm{B}_{\mathrm{rig},K}^{\dagger(,s)}$ are non-Noetherian, they are Bézout domains. For brevity, we will often denote an unspecified one of the rings $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}_K^{\dagger(,s)}$, $A' \widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}_K^{\dagger(,s)}$, or $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}_{\mathrm{rig},K}^{\dagger(,s)}$ simply by $\mathrm{B}^{(s)}$, and when we must emphasize its dependence on K, we will write $\mathrm{B}_K^{(s)}$.

If L/K is a finite Galois extension inside K^{alg}, then one can arrange for π_L to satisfy an Eisenstein polynomial over a subring of $\mathrm{B}_K^{\dagger} \otimes_{F'} F'_L$ with respect to a suitable π_K-adic valuation. (The term F'_L is the maximal absolutely unramified subfield of L_∞ analogous to F'.) The constants e_K and e_L are normalized so that the growth conditions on power series coincide. For $s \gg 0$, one gets functorial embeddings $\mathrm{B}_K^{(s)} \hookrightarrow \mathrm{B}_L^{(s)}$, which are finite free ring extensions (for $s \gg 0$) compatible with the actions of φ and Γ_L, and thus an action of H_K/H_L on $\mathrm{B}_L^{(s)}$ with $(\mathrm{B}_L^{(s)})^{H_K} = \mathrm{B}_K^{(s)}$.

The series $\log(1+\pi) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \pi^n$ converges in $\mathrm{B}_{\mathrm{rig},F}^{\dagger,s}$ for every $s > 0$, and we call its limit t. By means of the above embedding process, t is an element of every $\mathrm{B}_{\mathrm{rig},K}^{\dagger,s}$. One has $\varphi(t) = pt$ and $\gamma(t) = \chi_{\mathrm{cycl}}(\gamma)t$ for all $\gamma \in \Gamma$.

Given a B^s-module D^s, write $D^{(s')} = D \otimes_{\mathrm{B}^s} \mathrm{B}^{(s')}$ for $s' \geq s$. A *φ-module over* $\mathrm{B}^{(s)}$ is a finitely presented, projective $\mathrm{B}^{(s)}$-module $D^{(s)}$ equipped with a semilinear map $\varphi : D^{(s)} \to D^{(ps)}$ such that the associated linear map $\varphi' : \mathrm{B}^{(ps)}\, {}_{\varphi}\!\otimes_{\mathrm{B}^{(s)}} D^{(s)} \to D^{(ps)}$ is an isomorphism. We write $\mathbf{M}(\varphi)_{/\mathrm{B}^{(s)}}$ or $\mathbf{M}(\varphi)$ for the exact category of φ-modules over $\mathrm{B}^{(s)}$. A *(φ, Γ)-module over* $\mathrm{B}^{(s)}$ is a φ-module $D^{(s)}$ over $\mathrm{B}^{(s)}$ equipped with a semilinear action of Γ that commutes with φ and is continuous for varying $\gamma \in \Gamma$. We write $\mathbf{M}(\varphi, \Gamma)_{/\mathrm{B}^{(s)}}$ or $\mathbf{M}(\varphi, \Gamma)$ for the exact category of (φ, Γ)-modules over $\mathrm{B}^{(s)}$; it has tensor products and internal homs. Ultimately, we are concerned with (φ, Γ)-modules over $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger}_{\mathrm{rig},K}$ because it is over this ring that the link to p-adic Hodge theory is direct, and a finiteness theorem is known for Galois cohomology in complete generality. But for technical reasons, we must make use of the other variants at times; especially, the construction of the functor from Galois representations to (φ, Γ)-modules is documented in the literature in terms of $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger}_{K}$, and our proof of this functor's compatibility of Galois cohomology makes use of $A' \widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s}_{K}$.

Suppose that $s \gg 0$ so that $\psi : \mathrm{B}^{\dagger,ps}_{(\mathrm{rig},)K} \to \mathrm{B}^{\dagger,s}_{(\mathrm{rig},)K}$ is defined and hence also a left inverse $1 \widehat{\otimes} \psi : \mathrm{B}^{ps} \to \mathrm{B}^{s}$ to $1 \widehat{\otimes} \varphi : \mathrm{B}^{s} \to \mathrm{B}^{ps}$, and let D^s be a φ-module over $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger,s}_{(\mathrm{rig},)\mathbf{Q}_p}$. Then we obtain a left inverse to φ on D^s by the rule

$$D^{ps} \xleftarrow{\sim} \mathrm{B}^{ps} \otimes_{\varphi,\mathrm{B}^s} D^s \xrightarrow{(1\widehat{\otimes}\psi)\otimes 1} \mathrm{B}^s \otimes_{\mathrm{B}^s} D^s = D^s.$$

Upon taking $\varinjlim_s$, one gets a map $\psi : D \to D$ that is also left inverse to φ.

We will make use of the slope theory for φ-modules as in the following:

Theorem 2.1 [Kedlaya 2008]. *There is a homomorphism* $\deg : (\mathrm{B}^{\dagger}_{\mathrm{rig},K})^{\times} \to \mathbf{Q}$ *extending* ord_p *with the property that the rule* $\deg(D) = \deg(\varphi \mid \det D)$ *gives rise to a theory of Harder–Narasimhan filtrations on* $\mathbf{M}(\varphi)_{/\mathrm{B}^{\dagger}_{\mathrm{rig},K}}$.

One calls a φ-module over $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$ *étale* if its only slope is 0. The full subcategory of étale φ-modules is denoted by $\mathbf{M}^{\text{ét}}(\varphi) \subset \mathbf{M}(\varphi)$. A (φ, Γ)-module D is called *étale* if its underlying φ-module is; the full subcategory of these is written $\mathbf{M}^{\text{ét}}(\varphi, \Gamma) \subset \mathbf{M}(\varphi, \Gamma)$. Since the slope filtration is unique, it is Γ-stable.

We write $\mathbf{Rep}_A(G)$ and $\mathbf{Rep}_{A'}(G)$ for the category of finitely generated, flat A-modules and A'-modules, respectively, equipped with a continuous, linear action of G.

Theorem 2.2. *Let* $M \in \mathbf{Rep}_{A'}(G)$. *For* $s \gg 0$, *there exists a canonical* φ*- and* G*-stable* $A \widehat{\otimes}_{\mathbf{Z}_p} \mathrm{B}^{\dagger,s}_{K}$*-submodule*

$$\mathbf{D}^{\dagger,s}(M) \subseteq (M \otimes_{A'} (A' \widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s}))^{H}$$

that is projective of the same rank as M *such that the natural map*

$$(A' \widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s}) \otimes_{(A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger,s}_{K})} \mathbf{D}^{\dagger,s}(M) \to M \otimes_{A'} (A' \widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s}) \tag{2-1}$$

is an isomorphism; $\mathbf{D}^{\dagger,s}(M)$ *is a* (φ, Γ)*-module over* $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger,s}_{K}$. *(By "*$\varphi$*-stable", we mean that* $\varphi \mathbf{D}^{\dagger,s}(M) \subseteq (A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger,ps}_{K}) \cdot \mathbf{D}^{\dagger,s}(M)$.*) If* M *is of the form* $M_0[1/p]$ *where* $M_0 \in \mathbf{Rep}_A(G)$ *is free and such that* G *acts trivially on* $M_0/12p$, *then*

$\mathbf{D}^{\dagger,s}(M)$ is in fact free over $\mathrm{B}_K^{\dagger,s}$. If B is a p-adic A-algebra and $B' = B[1/p]$, then the natural maps

$$\mathbf{D}^{(s)}(M) \mathbin{\widehat{\otimes}}_{A'} B' \to \mathbf{D}^{(s)}(M \otimes_{A'} B')$$

are isomorphisms, provided $s \gg 0$ so that both sides exist.

Proof. Except for the final claim of compatibility with base change along $A' \to B'$, this follows from [Berger and Colmez 2008, Proposition 4.2.8 and Théorème 4.2.9] when M is free and from [Kedlaya and Liu 2010, Theorem 3.11 and Definition 3.12] in general. The final claim follows directly from the constructions although this is never explicitly stated in either reference. □

For M as in the theorem, we define

$$\widetilde{\mathbf{D}}^{\dagger(,s)}(M) = \mathbf{D}^{\dagger,s}(M) \otimes_{(A \widehat{\otimes}_{\mathbf{Z}_p} \mathrm{B}_K^{\dagger,s})} (A \mathbin{\widehat{\otimes}}_{\mathbf{Z}_p} \widetilde{\mathrm{B}}_K^{\dagger(,s)}),$$

$$\mathbf{D}_{\mathrm{rig}}^{\dagger(,s)}(M) = \mathbf{D}^{\dagger,s}(M) \otimes_{(A \widehat{\otimes}_{\mathbf{Z}_p} \mathrm{B}_K^{\dagger,s})} (A \mathbin{\widehat{\otimes}}_{\mathbf{Z}_p} \mathrm{B}_{\mathrm{rig},K}^{\dagger(,s)}).$$

For brevity, we will often denote the above associated module corresponding to the ring $\mathrm{B}^{(s)}$ by $\mathbf{D}^{(s)}(M)$ and a general $\mathrm{B}^{(s)}$-module by $D^{(s)}$.

Theorem 2.3. *For each ring $\mathrm{B}^{(s)}$, the rule $M \mapsto \mathbf{D}(M)$ determines an exact functor $\mathbf{Rep}_{A'}(G) \to \mathbf{M}(\varphi, \Gamma)_{/\mathrm{B}}$ respecting tensor and internal hom structures. Assuming additionally that A' is an affinoid algebra, this functor is fully faithful. When A is finite over $\mathbb{O}_E$, the essential image of $\mathbf{D}_{\mathrm{rig}}^{\dagger}$ is $\mathbf{M}^{\text{ét}}(\varphi, \Gamma)_{/\mathrm{B}}$.*

Proof. The full faithfulness of $\mathbf{D}^{\dagger}$ is given by [Kedlaya and Liu 2010, Proposition 2.7] and the comment of [ibid., Definition 3.12], and the full faithfulness of $\widetilde{\mathbf{D}}^{\dagger}$ follows by the same argument. The full faithfulness of $\mathbf{D}_{\mathrm{rig}}^{\dagger}$ is given by [ibid., Proposition 6.5]. The remaining claims are straightforward. □

Let L/K be a finite Galois extension inside K^{alg}. For a $\mathrm{B}_K^{(s)}$-module $D^{(s)}$, we use the shorthand $D_L^{(s)} = D^{(s)} \otimes_{\mathrm{B}_K^{(s)}} \mathrm{B}_L^{(s)}$. If $D^{(s)}$ has a φ-action, so does $D_L^{(s)}$. If $D^{(s)}$ has a Γ_K-action, then $D_L^{(s)}$ has a Γ_L-action. For $M \in \mathbf{Rep}_{A'}(G_K)$, one has $\mathbf{D}^{(s)}(M|_{G_L}) = \mathbf{D}^{(s)}(M)_L$.

The above results suggest the following (ad hoc) formalism. Let X be a p-adic analytic space over E, and let $\mathcal{U}$ be an admissible affinoid covering that is quasiclosed under intersections. For each choice of ring $\mathrm{B}^{(s)} = \mathrm{B}_K^{\dagger(,s)}, \widetilde{\mathrm{B}}_K^{\dagger(,s)}, \mathrm{B}_{\mathrm{rig},K}^{\dagger(,s)}$, denote by $\mathbb{O}_{\mathcal{U}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}^{(s)}$ the sheaf of rings on $\mathcal{U}$ (as always with the discrete Grothendieck topology) determined by the rule

$$\Gamma(Y, \mathbb{O}_{\mathcal{U}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}^{(s)}) = \Gamma(Y, \mathbb{O}_Y) \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}^{(s)},$$

equipped with the obvious actions of φ and Γ. If D^s is a $(\mathbb{O}_{\mathcal{U}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}^s)$-module, then $D^{(s')}$, interpreted in the obvious manner, is naturally an $\mathbb{O}_{\mathcal{U}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}^{(s')}$-module for any $s' > s$. However, we warn that an $\mathbb{O}_{\mathcal{U}} \mathbin{\widehat{\otimes}}_{\mathbf{Q}_p} \mathrm{B}$-module need not conversely arise from any $\mathscr{O}_{\mathscr{U}} \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^s$-module because the necessary s might not be bounded above for varying $Y \in \mathscr{U}$. Similarly, the natural injection

$$\varinjlim_s \Gamma(\mathscr{U}, \mathscr{O}_{\mathscr{U}} \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^s) \to \Gamma(\mathscr{U}, \mathscr{O}_{\mathscr{U}} \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B})$$

need not be a bijection. By a *family of φ-modules* over X (of type $\mathrm{B}^{(s)}$), we mean a quasiconsistent sheaf $D^{(s)}$ of finitely presented flat $(\mathscr{O}_{\mathscr{U}} \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{(s)})$-modules equipped with a semilinear action of φ such that for each $Y \in \mathscr{U}$ the associated linear map

$$\varphi' : (\Gamma(Y, \mathscr{O}_Y) \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{(ps)}) \otimes_{1 \widehat{\otimes} \varphi, \Gamma(Y, \mathscr{O}_Y) \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{(s)}} \Gamma(Y, D^{(s)}) \to \Gamma(Y, D^{(ps)})$$

is an isomorphism. By a *family of (φ, Γ)-modules* over X (of type $\mathrm{B}^{(s)}$), we mean a family of φ-modules $D^{(s)}$ over X (of type $\mathrm{B}^{(s)}$) equipped with a semilinear action of Γ that commutes with φ and is continuous for varying $\gamma \in \Gamma$. Given a family M of G-representations over X, the rule that associates to $Y \in \mathscr{U}$ the (φ, Γ)-module $\mathbf{D}(\Gamma(Y, M))$ determines a family $\mathbf{D}(M)$ of (φ, Γ)-modules over X.

2B. *Galois cohomology of (φ, Γ)-modules.* For $D^{(s)}$ a (φ, Γ)-module over $\mathrm{B}^{(s)}$, we define its *Herr complex* [Herr 1998] or *Galois cochain complex* to be the object

$$\begin{aligned} \mathbf{R}\Gamma(G, D^{(s)}) &= \mathbf{R}\Gamma_{\mathrm{cont}}\big(\Gamma, \mathrm{Cone}\big[D^{(s)} \xrightarrow{\varphi-1} D^{(ps)}\big][-1]\big) \\ &\cong \mathrm{Cone}\big[\mathbf{R}\Gamma_{\mathrm{cont}}(\Gamma, D^{(s)}) \xrightarrow{\varphi-1} \mathbf{R}\Gamma_{\mathrm{cont}}(\Gamma, D^{(ps)})\big][-1] \end{aligned}$$

of $\mathbf{D}^{\mathrm{b}}(A')$ and its *Galois cohomology* $\mathrm{H}^*(G, D^{(s)})$ to be the associated graded in $\mathbf{Gr}^{\mathrm{b}}(A')$. This object can be made explicit: Γ/Δ is procyclic, say topologically generated by the image of $\gamma \in \Gamma$, and $\mathbf{R}\Gamma_{\mathrm{cont}}(G, D^{(s)})$ is represented by the complex

$$\mathrm{C}^{\bullet}_{\varphi,\gamma} : \big[(D^{(s)})^{\Delta} \xrightarrow{(\varphi-1, \gamma-1)} (D^{(ps)})^{\Delta} \oplus (D^{(s)})^{\Delta} \xrightarrow{(1-\gamma, \varphi-1)} (D^{(ps)})^{\Delta}\big] \qquad \text{(2-2)}$$

in $\mathbf{K}^{\mathrm{b}}(A')$ concentrated in degrees 0, 1, and 2.

Remark 2.4. It is easy to check that $\mathrm{H}^i(G, \mathrm{Hom}_{\mathrm{B}^{(s)}}(D^{(s)}, D'^{(s)}))$ computes the Yoneda group $\mathrm{Ext}^i_{\mathbf{M}(\varphi,\Gamma)_{/\mathrm{B}^{(s)}}}(D^{(s)}, D'^{(s)})$, where $D^{(s)}$ and $D'^{(s)}$ are any two (φ, Γ)-modules over $\mathrm{B}^{(s)}$, for $i \leq 1$.

For two (φ, Γ)-modules $D^{(s)}$ and $D'^{(s)}$ over $\mathrm{B}^{(s)}$, we define cup products on cochains as in [Liu 2007]. In the representation $\mathrm{C}^{\bullet}_{\varphi,\gamma}$, the map

$$\mathrm{C}^i_{\varphi,\gamma}(D^{(s)}) \otimes_{A'} \mathrm{C}^j_{\varphi,\gamma}(D'^{(s)}) \to \mathrm{C}^{i+j}_{\varphi,\gamma}(D^{(s)} \otimes_{\mathrm{B}^{(s)}} D'^{(s)})$$

is the obvious multiplication when $i = 0$ or $j = 0$, and otherwise, we have

$$\begin{aligned} \mathrm{C}^1_{\varphi,\gamma}(D^{(s)}) \otimes_{A'} \mathrm{C}^1_{\varphi,\gamma}(D'^{(s)}) &\to \mathrm{C}^2_{\varphi,\gamma}(D^{(s)} \otimes_{\mathrm{B}^{(s)}} D'^{(s)}), \\ (d_1, d_2) \otimes (d'_1, d'_2) &\mapsto d_2 \otimes \gamma(d'_1) - d_1 \otimes \varphi(d'_2). \end{aligned}$$

The finite generation of the Galois cohomology of (φ, Γ)-modules is now known thanks to [Kedlaya et al. 2012], which appeared after the writing of this paper. We state the result here and henceforth refer to it as the finiteness theorem when it is invoked. We stress that $K/\mathbf{Q}_p$ is assumed to be finite.

Theorem 2.5. *For the* (φ, Γ)*-module* D *over* $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger}_{\mathrm{rig},K}$, *the Galois cohomology* $\mathbf{R}\Gamma(G, D)$ *belongs to* $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(A')$, *and the morphism*

$$\mathbf{R}\Gamma(G, D^*(1)) \to \mathbf{R}\Gamma(G, D)^*[-2]$$

adjoint to the pairing given by cup product and evaluation, comparison (*see Theorem 2.8 below*) *and truncation, and the local trace map, namely*

$$\begin{aligned}\mathbf{R}\Gamma(G, D^*(1)) \otimes^{\mathbf{L}}_{A'} \mathbf{R}\Gamma(G, D) \to \mathbf{R}\Gamma(G, A' \widehat{\otimes}_{\mathbf{Q}_p} \mathbf{B}^{\dagger}_{\mathrm{rig},K}(1)) &\\ \to \tau_{\geq 2}\mathbf{R}\Gamma_{\mathrm{cont}}(G, A'(1)) \cong A',&\end{aligned}$$

is an isomorphism.

The following proposition is proved without using the finiteness theorem and in fact is an ingredient in its proof:

Proposition 2.6. *Let* $D^{(s)}$ *be a* (φ, Γ)*-module over* $\mathrm{B}^{(s)}$.

(1) $D^{(s)}$ *is a flat* A'*-module.*

(2) *Suppose* B' *is an affinoid* A'*-algebra. Then the natural map in* $\mathbf{D}^b(B')$

$$\mathbf{R}\Gamma(G, D^{(s)}) \otimes^{\mathbf{L}}_{A'} B' \to \mathbf{R}\Gamma(G, D^{(s)} \widehat{\otimes}_{A'} B')$$

is an isomorphism if B' *is a finite* A' *algebra or if the modules* $\mathrm{H}^*(G, D^{(s)})$ *and* $\mathrm{H}^*(G, D^{(s)} \widehat{\otimes}_{A'} B')$ *are finitely generated over* A' *and* B', *respectively.*

Proof. Since $D^{(s)}$ is a projective $\mathrm{B}^{(s)}$-module and each appropriate ring $\mathrm{B}^{(s)}$ is a flat A'-algebra, (1) follows.

For (2), we use (1) to identify our map with

$$h : [\mathrm{C}^{\bullet}_{\varphi,\gamma}(D^{(s)}) \otimes_{A'} B'] \to [\mathrm{C}^{\bullet}_{\varphi,\gamma}(D^{(s)}) \widehat{\otimes}_{A'} B'].$$

Under the first condition, we have $\otimes_{A'} B' = \widehat{\otimes}_{A'} B'$, so the result is trivial. Under the second condition, each $\mathrm{H}^i(h)$ is a map of finitely generated B'-modules, so it suffices to show that the induced map $\mathrm{H}^i(h) \otimes_{B'} B'/\mathfrak{m}^n$ is an isomorphism for each maximal ideal $\mathfrak{m} \subset B$ and $n \geq 0$. One has a morphism of spectral sequences

$$\begin{array}{ccccc}
E_2^{i,j} = & \mathrm{Tor}^{B'}_{-i}(\mathrm{H}^j(G, D \otimes_{A'} B'), B'/\mathfrak{m}^n) & \Rightarrow & \mathrm{H}^{i+j}(G, D \otimes_{A'} B'/\mathfrak{m}^n) \\
\downarrow & \downarrow & & \downarrow \wr \\
\widehat{E}_2^{i,j} = & \mathrm{Tor}^{B'}_{-i}(\mathrm{H}^j(G, D \widehat{\otimes}_{A'} B'), B'/\mathfrak{m}^n) & \Rightarrow & \mathrm{H}^{i+j}(G, D \widehat{\otimes}_{A'} B'/\mathfrak{m}^n)
\end{array}$$

where the downward isomorphisms on the abutments are due to the fact that $B'/\mathfrak{m}^n$ is a finite A'-algebra. Using that $\mathrm{H}^j = 0$ for $j > 2$ throughout, one deduces immediately that $\mathrm{H}^2(h)$ is an isomorphism and then proceeds by repeated application of the five lemma to show that $\mathrm{H}^1(h)$, and then $\mathrm{H}^0(h)$, are isomorphisms. □

Given a finite Galois extension L/K inside K^{alg}, for any (φ, Γ)-module $D^{(s)}$, we leave it to the reader to define restriction and corestriction maps

$$\mathrm{res}_{L/K} : \mathbf{R}\Gamma(G_K, D^{(s)}) \to \mathbf{R}\Gamma(G_L, D_L^{(s)}),$$
$$\mathrm{cores}_{L/K} : \mathbf{R}\Gamma(G_L, D_L^{(s)}) \to \mathbf{R}\Gamma(G_K, D^{(s)}),$$

whose composition $\mathrm{cores}_{L/K} \circ \mathrm{res}_{L/K}$ induces multiplication by $[L : K]$ on cohomology. It follows that $\mathrm{H}^*(G_K, D^{(s)})$ is functorially a direct summand of $\mathrm{H}^*(G_K, D_L^{(s)})$ and that, when $\mathrm{B}^{(s)} = A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}_K^{\dagger(,s)}$, this decomposition is respected by the maps induced by $D^{(s)} \to \widetilde{D}^{\dagger(,s)}$ and $D^{(s)} \to D_{\mathrm{rig}}^{\dagger(,s)}$.

If X is a p-adic analytic space over E and if D is a family of (φ, Γ)-modules over X of type $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$, then by the finiteness theorem and Proposition 2.6(2), the rule $\Gamma(Y, \mathrm{C}^{\bullet}_{\varphi,\gamma}(D)) = \mathrm{C}^{\bullet}_{\varphi,\gamma}(\Gamma(Y, D))$ determines an object $\mathrm{C}^{\bullet}_{\varphi,\gamma}(D)$ of $\mathbf{K}^{\mathrm{b}}_{\mathrm{ft}}(\mathcal{U})$, whose class $\mathbf{R}\Gamma(G, D)$ in the derived category belongs to $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(\mathcal{U})$.

2C. ***Galois cohomology of Galois representations.*** Throughout this section, let $M \in \mathbf{Rep}_{A'}(G)$, and assume that A' is an affinoid algebra.

Proposition 2.7. *The natural maps*

$$\mathbf{R}\Gamma(G, \mathbf{D}^{\dagger}(M)) \to \mathbf{R}\Gamma(G, \widetilde{\mathbf{D}}^{\dagger}(M)),$$
$$\mathbf{R}\Gamma(G, \mathbf{D}^{\dagger(,s)}(M)) \to \mathbf{R}\Gamma(G, \mathbf{D}_{\mathrm{rig}}^{\dagger(,s)}(M))$$

are isomorphisms in $\mathbf{D}^{\mathrm{b}}(A')$.

Proof. In order to check whether the induced maps on cohomology are isomorphisms, it suffices to check whether they become isomorphisms when restricted to the members of an affinoid covering of A'. Thus, we reduce to the case where M is free over A'. By replacing A by a different unit ball subalgebra of A', we may assume that $M = M_0[1/p]$ for a finitely generated, free A-lattice M_0 that is G-stable. Choose a finite Galois extension L/K inside K^{alg} such that G_L acts trivially on $M_0/12p$. Since the morphisms in question respect the direct sum decompositions of the Galois cohomology over L coming from inflation and restriction relative to L/K, it suffices to prove the theorem with K replaced by L, and thus, we may assume that $\mathbf{D}^{\dagger,s}(M)$ is a free module.

Consider the first map. It suffices to show that the natural morphism

$$\mathbf{R}\Gamma_{\mathrm{cont}}(\Gamma, \mathbf{D}^{\dagger}(M)) \to \mathbf{R}\Gamma_{\mathrm{cont}}(\Gamma, \widetilde{\mathbf{D}}^{\dagger}(M))$$

is an isomorphism in $\mathbf{D}(A)$. A standard fact in the Tate–Sen theory of (φ, Γ)-modules is that $\widetilde{\mathbf{D}}^{\dagger}(M)$ admits a Galois-stable topological $(A \,\widehat{\otimes}_{\mathbf{Z}_p} \mathbf{B}^{\dagger}_K)$-direct sum decomposition as $\mathbf{D}^{\dagger}(M) \oplus X$ such that $\gamma - 1$ acts bijectively on X with continuous inverse. (See [Andreatta and Iovita 2008, Theorem 7.16] for an explanation of the method, taking $d = 0$ everywhere, and generalize it to A-valued M as in [Berger and Colmez 2008, §3].) Since $\mathbf{R}\Gamma_{\mathrm{cont}}(\Gamma, X) \cong 0$, the claim follows.

For the second map, in the case with superscripts, one copies the proof of [Kedlaya 2008, Proposition 1.2.6] verbatim to obtain that the natural morphism of complexes

$$\big[\mathbf{D}^{\dagger,s}(M) \xrightarrow{\varphi-1} \mathbf{D}^{\dagger,ps}(M)\big] \to \big[\mathbf{D}^{\dagger,s}_{\mathrm{rig}}(M) \xrightarrow{\varphi-1} \mathbf{D}^{\dagger,ps}_{\mathrm{rig}}(M)\big]$$

is a quasi-isomorphism; the claim follows from this and the definitions. One obtains the case without superscripts from the former by taking $\varinjlim_s$. □

When A is a finite $\mathbf{Z}_p$-algebra, the following main result is due to Liu [2007]:

Theorem 2.8. *There is a functorial isomorphism*

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \xrightarrow{\sim} \mathbf{R}\Gamma(G, \mathbf{D}(M))$$

in $\mathbf{D}^{\mathrm{b}}(A')$, *which is compatible with cup products and in degrees* $i \le 1$ *agrees with applying* $\mathbf{D}$ *to Yoneda extension classes.*

The key to the proof is the following:

Lemma 2.9. *The obvious maps*

$$A' \to \mathrm{Cone}\big[A' \,\widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s} \xrightarrow{\varphi-1} A' \,\widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,ps}\big][-1],$$
$$A' \,\widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s}_K \to \mathbf{R}\Gamma_{\mathrm{cont}}(H, A' \,\widehat{\otimes}_{\mathbf{Q}_p} \widetilde{\mathrm{B}}^{\dagger,s})$$

are isomorphisms in $\mathbf{D}(A')$.

Proof. In the case $A' = \mathbf{Q}_p$, the lemma is well-known, so we have exact sequences

$$0 \to \mathbf{Q}_p \to \widetilde{\mathrm{B}}^{\dagger,s} \xrightarrow{\varphi-1} \widetilde{\mathrm{B}}^{\dagger,ps} \to 0,$$
$$0 \to \widetilde{\mathrm{B}}^{\dagger,s}_K \to \mathrm{C}^0_{\mathrm{cont}}(H, \widetilde{\mathrm{B}}^{\dagger,s}) \to \mathrm{C}^1_{\mathrm{cont}}(H, \widetilde{\mathrm{B}}^{\dagger,s}) \to \cdots .$$

To deduce the result, we simply note that the functor $\widehat{\otimes}_{\mathbf{Q}_p} S$ preserves exact sequences of $\mathbf{Q}_p$-Banach spaces whenever S is *potentially orthonormalizable* in the sense of [Buzzard 2007, §2] (even though this functor does *not* commute with formation of cohomology in general) and that any affinoid algebra has the latter property. □

Proof of Theorem 2.8. By Proposition 2.7, it suffices to give a functorial isomorphism

$$\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \to \mathbf{R}\Gamma(G, \widetilde{\mathbf{D}}^{\dagger,s}(M)).$$

The compatibility with cup products and operations on Yoneda extensions follows from a routine trace through the definitions and hence is omitted.

It is easy to deduce from the preceding lemma a canonical isomorphism

$$\mathbf{R}\Gamma_{\mathrm{cont}}(H, M) \cong \mathrm{Cone}\big[\widetilde{\mathbf{D}}^{\dagger,s}(M) \xrightarrow{\varphi-1} \widetilde{\mathbf{D}}^{\dagger,ps}(M)\big][-1].$$

Combining this isomorphism with a standard argument involving the Hochschild–Serre spectral sequence, the desired result follows. □

We may now shed some light on the essential image of $\mathbf{D}^{\dagger}_{\mathrm{rig}}$ on families of Galois representations, which at present is mysterious.

Corollary 2.10. *Let* $0 \to D' \to E \to D \to 0$ *be a short exact sequence of* (φ, Γ)*-modules over* $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^{\dagger,s}_{\mathrm{rig},K}$*. If* D *and* D' *arise from* A'*-valued Galois representations, then so does* E*.*

Proof. Note that $\mathbf{D}^{\dagger,s}_{\mathrm{rig}}(\mathrm{Hom}_{A'}(M, M')) = \mathrm{Hom}_{A'\widehat{\otimes}_{\mathbf{Q}_p}\mathrm{B}^{\dagger,s}_{\mathrm{rig},K}}(D, D')$ with $D = \mathbf{D}^{\dagger,s}_{\mathrm{rig}}(M)$ and $D' = \mathbf{D}^{\dagger,s}_{\mathrm{rig}}(M')$. By Remarks 1.8 and 2.4, to see the claim, it suffices to apply Theorem 2.8 to $\mathrm{Hom}_{A'}(M, M')$ and take H^1 of the result. □

Let X be a p-adic analytic space over E, let M be a family of G-representations over X, and let $\mathcal{U}$ be an admissible affinoid covering of X that is quasiclosed under intersection. Then the functoriality of the isomorphisms in Theorem 2.8 gives rise to a canonical isomorphism $\mathbf{R}\Gamma_{\mathrm{cont}}(G, M) \cong \mathbf{R}\Gamma(G, \mathbf{D}(M))$, which takes place in $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(\mathcal{U})$ by the finiteness theorem.

3. Ordinary (φ, Γ)-modules and Selmer groups

In this section, we put ourselves in the situation of Section 2, specializing to the case where $A' = \mathbf{Q}_p$ and $\mathrm{B}^{(s)} = \mathrm{B}^{\dagger(,s)}_{\mathrm{rig},K}$ unless otherwise specified.

3A. ***p-adic Hodge theory of*** (φ, Γ)***-modules.*** This section describes the p-adic Hodge theory of (φ, Γ)-modules over the Robba ring. All the constructions and results to be found here are extensions of well-known ones for p-adic Galois representations, and many have explicitly appeared elsewhere; see, especially, works of Benois [2011] and Bellaïche and Chenevier [2009, §2.2]. The reader may check that every construction in this subsection holds with K replaced by a general complete discretely valued field with the exception of the Euler–Poincaré formula.

There exists a sequence of Γ-equivariant maps $\iota_n : \mathrm{B}^{\dagger,p^n}_{\mathrm{rig},K} \to K_n[\![t]\!]$ for all $n \geq n(K)$, where t is the element of $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$ defined in Section 2A, such that $\iota_{n+1} \circ \varphi = \iota_n$. Given a φ-module D, we put $D^+_{\mathrm{dif}} = D^s \otimes_{\mathrm{B}^{\dagger,s}_{\mathrm{rig},K},\iota_n} K_\infty[\![t]\!]$, and $D_{\mathrm{dif}} = D^s \otimes_{\mathrm{B}^{\dagger,s}_{\mathrm{rig},K},\iota_n} K_\infty((t))$, where D^s is uniquely determined for $s > s(D)$ by [Berger 2008, Théorème I.3.3] and is a model of D over $\mathrm{B}^{\dagger,s}_{\mathrm{rig},K}$. Using the φ-structure, one shows that these rules are independent of s and n satisfying $s > s(D)$ and $p^n \geq \max(p^{n(K)}, s)$. The rules $D \mapsto D^{(+)}_{\mathrm{dif}}$ are functorial and exact in D.

If D is actually a (φ,Γ)-module, then the $D_{\mathrm{dif}}^{(+)}$ admit Γ-actions (by perhaps enlarging the s used), and we define $D_{\mathrm{dR}}^{(+)} = (D_{\mathrm{dif}}^{(+)})^{\Gamma}$. These are K-vector spaces of dimension at most rank D, and they carry a decreasing, separated, and exhaustive filtration induced by the t-adic filtration on $K_\infty((t))$. One says that D is *de Rham* if $\dim_K D_{\mathrm{dR}} = \operatorname{rank} D$ and denotes by $\mathbf{M}^{\mathrm{dR}}(\varphi,\Gamma) \subset \mathbf{M}(\varphi,\Gamma)$ the full subcategory of de Rham objects. For such D, we define its *Hodge–Tate weights* to be the $h \in \mathbf{Z}$ with $\mathrm{Gr}^h D_{\mathrm{dR}} \neq 0$ with respective multiplicities $\dim_K \mathrm{Gr}^h D_{\mathrm{dR}}$. (There is no standard convention for the sign of the Hodge–Tate weight of the cyclotomic character; in this paper, it is -1.)

We write for brevity

$$D[t^{-1}] = D \otimes_{\mathrm{B}^\dagger_{\mathrm{rig},K}} \mathrm{B}^\dagger_{\mathrm{rig},K}[t^{-1}],$$
$$D[\log\pi, t^{-1}] = D \otimes_{\mathrm{B}^\dagger_{\mathrm{rig},K}} \mathrm{B}^\dagger_{\mathrm{rig},K}[\log\pi, t^{-1}],$$

where the element $\log\pi$ is a free variable over $\mathrm{B}^\dagger_{\mathrm{rig},K}$ equipped with actions of φ and Γ by the formulas

$$\varphi(\log\pi) = p\log\pi + \log(\varphi(\pi)/\pi^p) \quad \text{and} \quad \gamma(\log\pi) = \log\pi + \log(\gamma(\pi)/\pi),$$

the series $\log(\varphi(\pi)/\pi^p)$ and $\log(\gamma(\pi)/\pi)$ being convergent in $\mathrm{B}^\dagger_{\mathrm{rig},\mathbf{Q}_p}$. We associate to D the modules

$$D^+_{\mathrm{crys}} = D^{\Gamma}, \quad D_{\mathrm{crys}} = D[t^{-1}]^{\Gamma}, \quad \text{and} \quad D_{\mathrm{st}} = D[\log\pi, t^{-1}]^{\Gamma}.$$

These three modules are semilinear φ-modules over F of dimension at most rank D. The latter two are related via the so-called monodromy operator N. Namely, consider the unique $\mathrm{B}^\dagger_{\mathrm{rig},K}$-derivation $N : \mathrm{B}^\dagger_{\mathrm{rig},K}[\log\pi] \to \mathrm{B}^\dagger_{\mathrm{rig},K}[\log\pi]$ satisfying $N(\log\pi) = -\frac{p}{p-1}$. It satisfies $N\varphi = p\varphi N$ and commutes with Γ and thus gives rise to a nilpotent operator N on D_{st} with the property that $D_{\mathrm{crys}} = D_{\mathrm{st}}^{N=0}$.

We say that D is *crystalline* or *semistable* if D_{crys} or D_{st} has the maximal F-dimension, namely $\dim_F D_{\mathrm{crys}} = \operatorname{rank} D$ or $\dim_F D_{\mathrm{st}} = \operatorname{rank} D$, respectively. Upon fixing a uniformizer for K, we can construct a canonical embedding $D_{\mathrm{st}} \otimes_F K \hookrightarrow D_{\mathrm{dR}}$ so that D being semistable implies D being de Rham. We call D *potentially crystalline* or *potentially semistable* if there exists a finite extension L/K inside K^{alg} such that D_L is crystalline or semistable, respectively, when considered as a (φ,Γ_L)-module. The following statement is known as Berger's p-adic local monodromy theorem:

Theorem 3.1 [Berger 2002]. *Every de Rham (φ,Γ)-module is potentially semistable.*

Given a de Rham D, let L/K be a finite Galois extension inside K^{alg} such that D_L is semistable. Then $(D_L)_{\mathrm{st}}$ is a (φ, N)-module over the maximal absolutely unramified subfield F_L of L, and $(D_L)_{\mathrm{st}} \otimes_{F_L} L = (D_L)_{\mathrm{dR}}$ is a filtered L-vector space.

Essentially because these data arise via base change from K, they are naturally equipped with a semilinear action of $\mathrm{Gal}(L/K)$ that commutes with φ and N and preserves the filtration. Such an object is called a *filtered* $(\varphi, N, \mathrm{Gal}(L/K))$*-module*. Given two extensions L_i and filtered $(\varphi, N, \mathrm{Gal}(L_i/K))$-modules D_i (for $i = 1, 2$), we consider them equivalent if there exists an extension L containing the L_i such that the $(D_i)_L$ are isomorphic. When we consider objects only up to this equivalence, we call them *filtered* (φ, N, G)*-modules*. We point out that if D becomes semistable over both L_1 and L_2, then $(D_{L_1})_{\mathrm{st}}$ and $(D_{L_2})_{\mathrm{st}}$ are equivalent, and we call this equivalence class D_{pst}. The rigid exact $\mathbf{Q}_p$-linear tensor category of filtered (φ, N, G)-modules is denoted $\mathbf{MF}(\varphi, N, G)$.

The objects M of the category $\mathbf{MF}(\varphi, N, G)$ admit a notion of degree, namely the Newton slope minus the Hodge–Tate weight of $M^{\wedge \operatorname{rank}(M)}$, which gives rise to a Harder–Narasimhan theory. One calls M (*weakly*) *admissible* if it is semistable of slope 0 in the sense of Harder–Narasimhan theory. (See [Berger 2008, §I.1] for details.)

Theorem 3.2 [Colmez and Fontaine 2000; Berger 2008]. *The functor $D \mapsto D_{\mathrm{pst}}$ is an exact equivalence of categories*

$$\mathbf{M}^{\mathrm{dR}}(\varphi, \Gamma) \xrightarrow{\sim} \mathbf{MF}(\varphi, N, G)$$

that matches their Harder–Narasimhan theories. In particular, a de Rham (φ, Γ)-module D is étale if and only if D_{pst} is (*weakly*) *admissible.*

Comparing notions of image and coimage, one deduces that the t-saturated (φ, Γ)-stable $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$-submodules of D are in a functorial, order-preserving correspondence with subspaces of D_{pst} that are stable under the (φ, N, G)-actions (equipped with the filtration induced from D_{pst}). Furthermore, a t-saturated (φ, Γ)-stable $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$-submodule is actually a $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$-direct summand.

The following immediate consequence of the p-adic monodromy theorem is usually stated for Galois representations, but the proof carries over without change for (φ, Γ)-modules. (As pointed out in [Berger 2002], the étale case was first proved by O. Hyodo without use of the p-adic monodromy theorem, but the proof cited below works for arbitrary complete discretely valued K.)

Corollary 3.3 [Berger 2002, Théorème 6.2]. *Let $0 \to D' \to D \to D'' \to 0$ be a short exact sequence of (φ, Γ)-modules. If D' and D'' are semistable and D is de Rham, then D is semistable.*

As in Remark 2.4, the cohomology groups $\mathrm{H}^1(G, D)$ coincide with Yoneda groups: to every $c \in \mathrm{H}^1(G, D)$, there corresponds a class of extensions

$$0 \to D \to E_c \to \mathbf{1} \to 0,$$

where $\mathbf{1}$ denotes the unit (φ, Γ)-module. The rule $[E_c] \mapsto [(E_c)_{\mathrm{dif}}]$ determines a map $\mathrm{H}^1(G, D) \to \mathrm{H}^1_{\mathrm{cont}}(\Gamma, D_{\mathrm{dif}})$, and the *Bloch–Kato "g" local subspace* is given by

$$\mathrm{H}^1_{\mathrm{g}}(G, D) = \ker[\mathrm{H}^1(G, D) \to \mathrm{H}^1_{\mathrm{cont}}(\Gamma, D_{\mathrm{dif}})].$$

When D is de Rham, one has

$$\begin{aligned}\mathrm{H}^1_{\mathrm{g}}(G, D) &= \{c \in \mathrm{H}^1(G, D) \mid E_c \text{ is de Rham}\} \\ &= \{c \in \mathrm{H}^1(G, D) \mid E_c \text{ is potentially semistable}\} \\ &= \mathrm{Ext}^1_{\mathbf{MF}(\varphi, N, G)}(\mathbf{1}, D_{\mathrm{pst}}).\end{aligned} \tag{3-1}$$

Similarly, a map $\mathrm{H}^1(G, D) \to \mathrm{H}^1_{\mathrm{cont}}(\Gamma, D[1/t])$ is determined by forgetting φ-structures and inverting t, and we define the *Bloch–Kato "f" local subspace* to be

$$\mathrm{H}^1_{\mathrm{f}}(G, D) = \ker[\mathrm{H}^1(G, D) \to \mathrm{H}^1_{\mathrm{cont}}(\Gamma, D[t^{-1}])].$$

When D is crystalline, one has

$$\mathrm{H}^1_{\mathrm{f}}(G, D) = \{c \in \mathrm{H}^1(G, D) \mid E_c \text{ is crystalline}\}.$$

If D is de Rham, then under the isomorphism (3-1) one can compute $\mathrm{H}^1_{\mathrm{f}}(G, D)$ as certain extensions of filtered (φ, N, G)-modules, obtaining the exact sequence

$$0 \to \mathrm{H}^0(G, D) \to D_{\mathrm{crys}} \xrightarrow{(1-\varphi, 1)} D_{\mathrm{crys}} \oplus D_{\mathrm{dR}}/D^+_{\mathrm{dR}} \to \mathrm{H}^1_{\mathrm{f}}(G, D) \to 0.$$

This computation can be enhanced to show that the local condition associated to the subspace $\mathrm{H}^1_{\mathrm{f}}(G, D)$ is isomorphic in the derived category to the complex

$$\begin{aligned}\mathrm{C}^\bullet_{\mathrm{f}}(G, D) &= \mathrm{Cone}\big[D_{\mathrm{crys}} \xrightarrow{(1-\varphi, 1)} D_{\mathrm{crys}} \oplus D_{\mathrm{dR}}/D^+_{\mathrm{dR}}\big][-1], \\ \mathbf{R}\Gamma_{\mathrm{f}}(G, D) &= [\mathrm{C}^\bullet_{\mathrm{f}}(G, D)],\end{aligned}$$

and one obtains the "Euler–Poincaré" formula

$$\dim_{\mathbf{Q}_p} \mathrm{H}^1_{\mathrm{f}}(G, D) = \dim_{\mathbf{Q}_p} \mathrm{H}^0(G, D) + \dim_{\mathbf{Q}_p} D_{\mathrm{dR}}/D^+_{\mathrm{dR}}. \tag{3-2}$$

The next result follows from the (elementary) computation of Γ-cohomology of $t^n K_\infty[\![t]\!]$ and that $K_\infty[\![t]\!]$ is a PID.

Proposition 3.4. *Let $0 \to D' \to D \to D'' \to 0$ be a short exact sequence of (φ, Γ)-modules with D' and D'' de Rham. If all the Hodge–Tate weights of D' are strictly less than all the Hodge–Tate weights of D'', then the short exact sequence of Γ-modules*

$$0 \to (D')^+_{\mathrm{dif}} \to D^+_{\mathrm{dif}} \to (D'')^+_{\mathrm{dif}} \to 0$$

is split. In particular, D is de Rham.

3B. ***The* (*strict*) *ordinary local condition.*** We briefly relax our hypothesis that $A' = \mathbf{Q}_p$; instead, A' can be any E-affinoid algebra.

Let D be a (φ, Γ)-module over $A' \widehat{\otimes}_{\mathbf{Q}_p} \mathrm{B}^\dagger_{\mathrm{rig},K}$, or a family of (φ, Γ)-modules of type $\mathrm{B}^\dagger_{\mathrm{rig},K}$ over a p-adic analytic space X, endowed with an admissible affinoid covering $\mathfrak{U}$ that is quasiclosed under intersection. By a *nearly ordinary filtration on* D, we mean a decreasing partial flag $F^* \subseteq D$, consisting of sub-(φ, Γ)-modules that are module-direct summands, or subfamilies of (φ, Γ)-modules that are direct summands over each $Y \in \mathfrak{U}$, respectively, such that each Gr^α_F has constant rank.

If M is either an object of $\mathbf{Rep}_{A'}(G)$ or a family of G-representations over X, then a *nearly ordinary filtration* for M is by definition one for $\mathbf{D}^\dagger_{\mathrm{rig}}(M)$, and we call it *classically nearly ordinary* if it arises from a partial flag of M consisting of G-stable direct summands.

Example 3.5. When $A' = \mathbf{Q}_p$, being nearly ordinary with a complete flag means being (split) trianguline.

Given a sub-(φ, Γ)-module $F^+ \subseteq D$ that is a module-direct summand or a subfamily of (φ, Γ)-modules that are module-direct summands over each $Y \in \mathfrak{U}$, we recall from Example 1.18 the (*strict*) *ordinary local condition* given by the morphism

$$\mathbf{R}\Gamma_{\mathrm{str}}(G, D) = \mathbf{R}\Gamma(G, F^+) \to \mathbf{R}\Gamma(G, D).$$

In the case $D = \mathbf{D}^\dagger_{\mathrm{rig}}(M)$, we get the local condition for M,

$$\mathbf{R}\Gamma_{\mathrm{str}}(G, D) = \mathbf{R}\Gamma(G, F^+) \to \mathbf{R}\Gamma(G, D) \cong \mathbf{R}\Gamma_{\mathrm{cont}}(G, M).$$

By the finiteness theorem, both the domain and codomain belong to $\mathbf{D}^{[0,2]}_{\mathrm{perf}}(A')$, and the formation of this local condition commutes with arbitrary base change by Proposition 2.6(2) as well as with duality (in the same sense as does the classical strict ordinary local condition as in Example 1.18) again by the finiteness theorem. The image of the local condition in cohomology is clearly

$$\mathrm{img}[\mathrm{H}^1(G, F^+) \to \mathrm{H}^1(G, D)] = \ker[\mathrm{H}^1(G, D) \to \mathrm{H}^1(G, D/F^+)],$$

which is a generalization of the (*strict*) *ordinary local subspace* studied by Greenberg [1994b] in conjunction with the *nonstrict ordinary local subspace* (introduced earlier in [Greenberg 1989; 1994a])

$$\ker\bigl[\mathrm{H}^1(G_K, D) \to \mathrm{H}^1\bigl(G_{\widehat{K^{\mathrm{unr}}}}, (D/F^+) \otimes_{\mathrm{B}^\dagger_{\mathrm{rig},K}} \mathrm{B}^\dagger_{\mathrm{rig},\widehat{K^{\mathrm{unr}}}}\bigr)\bigr].$$

Although the nonstrict local subspace appears more often in the literature, we will not use it essentially because a derived analogue, like $\mathbf{R}\Gamma_{\mathrm{str}}(G, D)$ in the strict case, would involve Galois cohomology for the group $G_{\widehat{K^{\mathrm{unr}}}}$, which does not satisfy p-cohomological finiteness, rendering the derived analogue pathological.

We now resume our assumption that $A' = \mathbf{Q}_p$.

Let L/K be a finite Galois extension inside K^{alg}. We say a (φ, Γ)-module D over $\mathrm{B}^{\dagger}_{\mathrm{rig},K}$ is *L-ordinary* if it admits a nearly ordinary flag $F^* \subseteq D$ with the properties that each $(\mathrm{Gr}^{\alpha}_F)_L$ is semistable and if $\alpha < \beta$ then all Hodge–Tate weights of Gr^{α}_F are strictly greater than all Hodge–Tate weights of Gr^{β}_F. We say that D is *ordinary* if there exists some L for which it is L-ordinary. If V is a Galois representation, we say that V is (L-)*ordinary* if $\mathbf{D}^{\dagger}_{\mathrm{rig}}(V)$ is and *classically* (L-)*ordinary* if there exists a partial flag of V by G-stable direct summands giving rise via $\mathbf{D}^{\dagger}_{\mathrm{rig}}$ to the desired filtration.

Example 3.6. We will see in Section 3C that Greenberg's notion of "ordinary" translates to our classically K-ordinary with each Gr^{α}_F having only one Hodge–Tate weight (possibly with multiplicity). Thereafter, we will see examples of ordinary but not classically ordinary Galois representations. It is possible that D be ordinary with respect to more than one filtration even though ordinary filtrations in Greenberg's sense are unique when they exist.

Here are the main properties of ordinary (φ, Γ)-modules:

Proposition 3.7. *Let L/K be a finite Galois extension inside K^{alg}, and let D be a (φ, Γ)-module that is L-ordinary with filtration F^*.*

(1) *D becomes semistable over L and hence is potentially semistable.*

Suppose moreover that there exists a filtration step F^+ with the property that all the Hodge–Tate weights of F^+ are negative and all the Hodge–Tate weights of D/F^+ are nonnegative.

(2) *One has* $\mathrm{H}^0(G, D/F^+) = (D/F^+)^{\varphi=1}_{\mathrm{crys}}$ *and* $\mathrm{H}^0(G, (F^+)^*(1)) = ((F^+)^*(1))^{\varphi=1}_{\mathrm{crys}}$.

(3) *Suppose that all the spaces mentioned in part (2) vanish. Then the canonical maps in the derived category are isomorphisms*:

$$\mathbf{R}\Gamma(G, F^+) \xleftarrow{\sim} \mathbf{R}\Gamma_{\mathrm{f}}(G, F^+) \xrightarrow{\sim} \mathbf{R}\Gamma_{\mathrm{f}}(G, D),$$

hence the local conditions

$$\mathbf{R}\Gamma_{\mathrm{str}}(G, D) \cong \mathbf{R}\Gamma_{\mathrm{f}}(G, D) \quad \textit{and} \quad \mathrm{H}^1(G, F^+) \hookrightarrow \mathrm{H}^1(G, D).$$

Remark 3.8. A variant of the proposition can be formulated for general K complete discretely valued, but we omit it for brevity. Suppose $D = \mathbf{D}^{\dagger}_{\mathrm{rig}}(V)$ with V ordinary in the sense of Greenberg. When $K/\mathbf{Q}_p$ is finite, the claim (1) for V is due to Fontaine [Perrin-Riou 1994a], and for general K, it is due to Berger [2002, Corollaire 6.3]. When $K = \mathbf{Q}_p$, the claim (3) for V is essentially a result of Flach [1990, Lemma 2]. Our formulation of parts (2)–(3) closely follows that of Fukaya and Kato [2006, Lemma 4.1.7].

Proof. For part (1), by restriction, we immediately reduce to the case where D is K-ordinary. The first claim now follows by induction on the length of the filtration, the case of length 1 being trivial and the inductive step being given by Proposition 3.4 and Corollary 3.3.

Part (2) follows from noting that both D/F and $F^*(1)$ have only nonnegative Hodge–Tate weights and applying to them the claim that for any (φ, Γ)-module D one has $\mathrm{H}^0(G, D) = D_{\mathrm{crys}}^{+,\varphi=1}$.

We now turn to part (3). For the first arrow, we have equality of cohomology outside degrees 1 and 2. For degree 1, we compute that

$$\begin{aligned}\dim_{\mathbf{Q}_p} \mathrm{H}^1(G, F^+) &= \dim_{\mathbf{Q}_p} \mathrm{H}^0(G, F^+) + \dim_{\mathbf{Q}_p} \mathrm{H}^2(G, F^+) + [K : \mathbf{Q}_p] \operatorname{rank} F^+ \\ &= \dim_{\mathbf{Q}_p} \mathrm{H}^0(G, F^+) + \dim_{\mathbf{Q}_p} (F^+)_{\mathrm{dR}}/(F^+)^+_{\mathrm{dR}} \\ &= \dim_{\mathbf{Q}_p} \mathrm{H}^1_{\mathrm{f}}(G, F^+),\end{aligned}$$

using the local Euler–Poincaré formulas of [Liu 2007] and Equation (3-2) for F^+, the computation $\mathrm{H}^2(G, F^+) = \mathrm{H}^0(G, (F^+)^*(1))^* = 0$ by local duality for (φ, Γ)-modules [Liu 2007], and that F^+ has only negative Hodge–Tate weights. The same local duality computation takes care of degree 2. For the second arrow, we have equality of cohomology outside degrees 0 and 1. The long exact cohomology sequence and the vanishing of $\mathrm{H}^0(G, D/F^+)$ give the result in degree 0 and the injectivity of $\mathrm{H}^1(G, F^+) \to \mathrm{H}^1(G, D)$ and hence also of $\mathrm{H}^1_{\mathrm{f}}(G, F^+) \to \mathrm{H}^1_{\mathrm{f}}(G, D)$. To conclude in degree 1, it suffices to compare the Euler–Poincaré formulas for the two dimensions, noting that $(F^+)_{\mathrm{dR}}/(F^+)^+_{\mathrm{dR}} \xleftarrow{\sim} (F^+)_{\mathrm{dR}} \xrightarrow{\sim} D_{\mathrm{dR}}/D^+_{\mathrm{dR}}$. $\square$

3C. ***Examples of ordinary representations.*** When discussing examples, the following equivalent formulation is helpful:

Alternate definition 3.9. We remind the reader that by Proposition 3.7(1), every ordinary (φ, Γ)-module is de Rham, so we assume this is the case from the outset.

Given a de Rham (φ, Γ)-module D, by the discussion of Section 3A, the ordinary filtrations $F^* \subseteq D$ are in a natural correspondence with filtrations $F^* \subseteq D_{\mathrm{pst}}$ by (φ, N, G)-stable subspaces (each equipped with its Hodge filtration induced by D_{pst}) such that for $\alpha < \beta$ all the jump indices of the induced Hodge filtration on $(\mathrm{Gr}^\alpha_F)_{\mathrm{dR}}$ are strictly greater than all the jump indices of the induced Hodge filtration on $(\mathrm{Gr}^\beta_F)_{\mathrm{dR}}$. Note the reversal of order of the jump indices: this feature is independent of one's normalizations of Hodge–Tate weights.

Example 3.10 (Greenberg's ordinary representations). Let us see how ordinary representations, defined by Greenberg [1989] when $K = \mathbf{Q}_p$, fit into our context. We are given a Galois representation V so that $D = \mathbf{D}^\dagger_{\mathrm{rig}}(V)$ is étale. Greenberg's ordinary hypothesis is that V admits a decreasing filtration $F^* \subseteq V$ by G-stable subspaces such that for each α the representation $\chi_{\mathrm{cycl}}^{-n_\alpha} \otimes \mathrm{Gr}^\alpha_F V$ is unramified for some integer n_α, and the n_α are strictly increasing. This means precisely that each $\mathrm{Gr}_F^\alpha V$ is crystalline of all Hodge–Tate weights equal to $-n_\alpha$. Thus, Greenberg's ordinary hypothesis is a strengthening of our classically K-ordinary hypothesis to require that each of the graded pieces be of a single Hodge–Tate weight. In the language of filtered (φ, N, G)-modules, a filtration $F^* \subseteq D_{\mathrm{pst}}$ corresponds to a Greenberg-ordinary filtration on V precisely when V is semistable, and each $\mathrm{Gr}_F^\alpha D_{\mathrm{pst}}$ is (weakly) admissibly filtered of a single Hodge–Tate weight, with the weights strictly decreasing, which means here that each $\mathrm{Gr}_F^\alpha D_{\mathrm{pst}}$ is of pure φ-slope $-n_\alpha$ and Hodge–Tate weight $-n_\alpha$ and satisfies $N = 0$ and that $n_\alpha > n_\beta$ for $\alpha < \beta$.

The reader will notice in the examples below that although V admits at most one Greenberg-ordinary filtration it may admit many different (φ, Γ)-ordinary filtrations. This is complementary to the existence of many p-adic L-functions.

Example 3.11 (Abelian varieties). Take an abelian variety B/K of dimension $d \geq 1$ with semistable reduction over $\mathcal{O}_K$, and consider $D = \mathbf{D}^\dagger_{\mathrm{rig}}(V)$ with $V = \mathrm{T}_p B \otimes \mathbf{Q}$ the p-adic Tate module up to isogeny. The Hodge filtration $\mathrm{Hodge}^* \subseteq D_{\mathrm{dR}}$ satisfies $\dim_K \mathrm{Gr}^0_{\mathrm{Hodge}} = \dim_K \mathrm{Gr}^{-1}_{\mathrm{Hodge}} = d$, and its Frobenius slopes h satisfy $-1 \leq h \leq 0$. By weak admissibility, the φ-eigenspaces with nonzero slopes do not meet Hodge^0. A nontrivial ordinary filtration thus consists of a (φ, N)-stable subspace $F \subseteq D_{\mathrm{st}}$ of rank d such that F_{dR} is complementary to Hodge^0 in D_{dR}.

Example 3.12 (Elliptic modular eigenforms). This case is treated in detail in [Pottharst 2012], the upshot being as follows. Let $p > 2$, and let f be a normalized elliptic modular cuspidal new eigenform of weight $k \geq 2$ with associated cohomological p-adic Galois representation V_f. If necessary, extend the scalars of V_f to contain the eigenvalues of φ on $\mathbf{D}_{\mathrm{pst}}(V_f)$. Then V_f is ordinary in our sense, often with two distinct ordinary filtrations, provided f *has finite slope*: the matrix of φ on $\mathbf{D}_{\mathrm{pst}}(V_f)$ is nonscalar. For example, when φ on $\mathbf{D}_{\mathrm{pst}}(V_f)$ is semisimple (as is conjecturally always the case), this is equivalent to there being some twist $f \otimes \varepsilon$ by a Dirichlet character ε that has an associated U_p-eigenform with nonzero U_p-eigenvalue. By contrast, f is "ordinary" in the parlance of p-adic modular forms if this condition is satisfied with $\varepsilon = 1$ and the U_p-eigenvalue a p-adic unit. Proposition 3.7(3) computes the Bloch–Kato local condition entering into the Selmer group of each of the Tate twists $V_f(n)$ corresponding to critical L-values except perhaps where exceptional zeroes (as in [Mazur et al. 1986]) occur.

3D. ***Ranks in families.*** We resume the assumptions of Section 1E, i.e., that K is a finite extension of $\mathbf{Q}$, that K^{alg} is a fixed algebraic closure, and that S is a fixed finite set of places v of K containing all v dividing p. Let M be a family of $G_{K,S}$-representations over a p-adic analytic space X, endowed with an admissible affinoid covering $\mathcal{U}$ that is quasiclosed under intersection. (For example, one can have X affinoid with algebra A' and M arising from a flat A'-module.) Suppose that, for each place $v \in S$ dividing p, the restriction $M|_{G_v}$ is equipped with a nearly ordinary filtration $F_v^* \subseteq D_v = \mathbf{D}_{\mathrm{rig}}^{\dagger}(M|_{G_v})$, and we have a distinguished index α_v.

In order to get a reasonable theory, we must assume that for each $v \in S$ *not* dividing p the subobject M^{I_v} is flat.

Using the strict ordinary local condition given by the $F_v^{\alpha_v}$ at places $v \in S$ dividing p and the unramified local condition at places $v \in S$ not dividing p, we build as in situation (3) of Section 1E the *strict ordinary Selmer complex*, denoted

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M) \in \mathbf{D}_{\mathrm{perf}}^{[0,3]}(\mathscr{U}).$$

One can check that it is invariant under enlarging S and hence is independent of S, so we may omit it from the notation.

We wish to compare the above complex to those associated to the members of the family. Namely, let $x \in X(E_x)$ be a point with residue field E_x, and let f_x denote the inclusion of the point x. We set $M_x = f_x^* M$, $D_{v,x} = f_x^* D_v$, and $F_{v,x}^* = f_x^* F_v^*$ for $v \in S$, and we use the strict ordinary local condition determined by the $F_{v,x}^{\alpha_v}$ at $v \in S$ above p, and the unramified local condition at $v \in S$ not dividing p, to construct in the same way the Selmer complex

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M_x) \in \mathbf{D}_{\mathrm{perf}}^{[0,3]}(E_x).$$

We study the natural specialization morphism

$$s_x : \mathbf{L} f_x^* \mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M) \to \mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M_x) \tag{3-3}$$

and in particular $\mathrm{H}^2(s_x)$.

It follows from the finiteness theorem and Proposition 2.6(2) that the formation of the strict ordinary local conditions commutes with $\mathbf{L} f_x^*$. We assume that $f_x^*(M^{I_v}) \xrightarrow{\sim} (M_x)^{I_v}$ so that the formation of the unramified local conditions commutes with $\mathbf{L} f_x^*$. Then the base-change theorem, situation (3) of Theorem 1.12, shows the morphism (3-3) to be an isomorphism, giving rise to a short exact sequence

$$0 \to f_x^* \widetilde{\mathrm{H}}_{\mathrm{str}}^2(G_K, M) \xrightarrow{\mathrm{H}^2(s_x)} \widetilde{\mathrm{H}}_{\mathrm{str}}^2(G_K, M_x) \to \mathrm{Tor}_1^{\mathscr{O}_X}(\widetilde{\mathrm{H}}_{\mathrm{str}}^3(G_K, M), E_x) \to 0.$$

Thus, $\mathrm{H}^2(s_x)$ is an isomorphism precisely when x avoids the support of the torsion in $\widetilde{\mathrm{H}}_{\mathrm{str}}^3(G_K, M)$.

On the other hand, we may relate $\widetilde{\mathrm{H}}_{\mathrm{str}}^2(G_K, M_x)$ to an extended Selmer group in degree 1 via duality. Namely, we equip $M_x^*(1)$ with the strict ordinary local conditions at $v \in S$ lying over p built from the $(F_{v,x}^{\alpha_v})^{\perp}$, and the unramified local conditions at $v \in S$ not dividing p, to construct the Selmer complex $\mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M_x^*(1))$. The local conditions for M_x and $M_x^*(1)$ are dual to one another. At places $v \in S$ lying over p, this follows from the duality of Galois cohomology of (φ, Γ)-modules contained in the finiteness theorem as previously mentioned. At places $v \in S$ not dividing p, there is a possible error in the integral self-duality of the unramified local conditions (coming from nontrivial Tamagawa numbers), but this contribution disappears after inverting p; see [Nekovář 2006, 7.6.7(iii)]. Thus, Theorem 1.16 gives

$$\mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G_K, M_x) \cong \mathbf{R}\widetilde{\Gamma}_{\mathrm{str}}(G, M_x^*(1))^*[-3]$$

in $\mathbf{D}^{[0,3]}_{\mathrm{perf}}(E_x)$ and, in particular,

$$\widetilde{\mathrm{H}}^2_{\mathrm{str}}(G_K, M_x) \cong \widetilde{\mathrm{H}}^1_{\mathrm{str}}(G_K, M_x^*(1))^*.$$

Now we relate the $\widetilde{\mathrm{H}}^1_{\mathrm{str}}(G_K, M_x^*(1))$ to Bloch–Kato Selmer groups for $M_x^*(1)$ by computing the complexes $E_v = \mathrm{Cone}(i_v)$ appearing in the exact triangles (1-3) and (1-4). For $v \in S$ dividing p, one has $E_v \cong \mathbf{R}\Gamma(G_v, (F^{\alpha_v}_{v,x})^*(1))$, whereas for $v \in S$ not dividing p one has

$$\mathrm{H}^0 E_v = 0 \quad \text{and} \quad \mathrm{H}^1 E_v \cong \mathrm{H}^1_{\mathrm{cont}}(I_v, M_x^*(1))^{G_v}.$$

Thus, one has an exact sequence

$$\cdots \to \bigoplus_{\substack{v \in S \\ \text{dividing } p}} \mathrm{H}^0(G_v, (F^{\alpha_v}_{v,x})^*(1)) \to \widetilde{\mathrm{H}}^1_{\mathrm{str}}(G_K, M_x^*(1)) \to \mathrm{H}^1_{\mathrm{cont}}(G_{K,S}, M_x^*(1))$$
$$\to \bigoplus_{\substack{v \in S \\ \text{dividing } p}} \mathrm{H}^1(G_v, (F^{\alpha_v}_{x,v})^*(1)) \oplus \bigoplus_{\substack{v \in S \\ \text{not dividing } p}} \mathrm{H}^1_{\mathrm{cont}}(I_v, M_x^*(1))^{G_v} \to \cdots,$$

and the image of $\widetilde{\mathrm{H}}^1_{\mathrm{str}}(G_K, M_x^*(1))$ in $\mathrm{H}^1_{\mathrm{cont}}(G_{K,S}, M_x^*(1))$ is identified to a subgroup cut out by local subspaces. For v not dividing p, these are the usual unramified local subspaces, and for v dividing p, these are strict ordinary local spaces in the sense of Section 3B. Therefore, assuming for each $v \in S$ dividing p that the hypotheses of Proposition 3.7(3) for $M_x^*(1)|_{G_v}$ hold, and in particular that each $\mathrm{H}^0(G_v, (F^{\alpha_v}_{x,v})^*(1)) = 0$, we have $\widetilde{\mathrm{H}}^1_{\mathrm{str}}(G_K, M_x^*(1)) \xrightarrow{\sim} \mathrm{H}^1_{\mathrm{f}}(G_K, M_x^*(1))$, where the right-hand side is Bloch–Kato's Selmer group for $M_x^*(1)$.

Because the $\widetilde{\mathrm{H}}^i_{\mathrm{str}}(G_K, M)$ are coherent sheaves on X, they have a reasonable structure theory. It follows that for x varying over all points of X with values in finite extensions E_x of E, the number $\dim_{E_x} f_x^* \widetilde{\mathrm{H}}^i_{\mathrm{str}}(G_K, M)$ is constant at its minimum value outside of a locally (i.e., over each affinoid) Zariski-closed proper subset. Further throwing away the support of the torsion in $\widetilde{\mathrm{H}}^3_{\mathrm{str}}(G_K, M)$, which only increases the resulting rank of $\widetilde{\mathrm{H}}^2_{\mathrm{str}}(G_K, M_x)$, we thus obtain the following:

Theorem 3.13. *Let M be a family of $G_{K,S}$-representations over X that is nearly ordinary at each $v \in S$ dividing p and for each $v \in S$ not dividing p that M^{I_v} is flat.*

Let X_0 be the set of points x with values in finite extensions E_x of E, for which

- $f_x^*(M^{I_v}) \xrightarrow{\sim} (M_x)^{I_v}$ *if* v *does not divide* p *and*
- *the hypotheses of Proposition 3.7(3) for* $M_x^*(1)|_{G_v}$ *hold if* v *divides* p.

Then the Bloch–Kato Selmer groups of the $M_x^*(1)$ *at the* $x \in X_0$ *have* E_x*-dimensions that are equal to their minimum except possibly on a locally Zariski-closed proper subset.*

Remark 3.14. In the case where the family is over a reduced affinoid of dimension 1, this statement is more or less equivalent to [Bellaïche 2012, Theorem 1].

Acknowledgments

Since this work grew out of investigations in my Ph.D. dissertation, I first thank my advisor, Barry Mazur, for his enthusiasm and encouragement. It was out of my desire to make rigorous many of his intuitions that I came upon this line of work. I benefited greatly from the works of Jan Nekovář and Ruochuan Liu, and I thank them heartily for their mathematical contributions. My recent collaboration with Kiran S. Kedlaya and Liang Xiao improved my understanding of many issues in this paper, and I am grateful to them. I am also indebted to Joël Bellaïche, Laurent Berger, Gaëtan Chenevier, Olivier Fouquet, Ben Howard, and Robert Pollack for numerous helpful conversations as well as the referee for suggesting several improvements in the exposition. I thank Tetsushi Ito for his repeated invitations to Japan, around which a great deal of progress on this project occurred, as well as Jan Nekovář for arranging my stay in Paris and l'Institut de Mathématiques de Jussieu for its hospitality. This work was supported by an NSF Graduate Research Fellowship.

References

[Amice and Vélu 1975] Y. Amice and J. Vélu, "Distributions p-adiques associées aux séries de Hecke", pp. 119–131 in *Journées Arithmétiques de Bordeaux* (Bordeaux, 1974), Astérisque **24/25**, Société Mathématique de France, Paris, 1975. MR 51 #12709 Zbl 0332.14010

[Andreatta and Iovita 2008] F. Andreatta and A. Iovita, "Global applications of relative (ϕ, Γ)-modules, I", pp. 339–420 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et (ϕ, Γ)-modules*, Astérisque **319**, Société Mathématique de France, Paris, 2008. MR 2010g:14020 Zbl 1163.11051

[Bellaïche 2012] J. Bellaïche, "Ranks of Selmer groups in an analytic family", *Trans. Amer. Math. Soc* **364** (2012), 4735–4761.

[Bellaïche and Chenevier 2009] J. Bellaïche and G. Chenevier, *Families of Galois representations and Selmer groups*, Astérisque **324**, Société Mathématique de France, Paris, 2009. MR 2011m:11105 Zbl 1192.11035

[Benois 2009] D. Benois, "On trivial zeroes of Perrin–Riou's L-functions", preprint, 2009. arXiv 0906.2862

[Benois 2011] D. Benois, "A generalization of Greenberg's $\mathscr{L}$-invariant", *Am. J. Math.* **133**:6 (2011), 1573–1632. MR 2863371 Zbl 5994421

[Berger 2002] L. Berger, "Représentations p-adiques et équations différentielles", *Invent. Math.* **148**:2 (2002), 219–284. MR 2004a:14022 Zbl 1113.14016

[Berger 2008] 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*, Astérisque **319**, Société Mathématique de France, Paris, 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*, Astérisque **319**, Société Mathématique de France, Paris, 2008. MR 2010g:11091 Zbl 1168.11020

[Berthelot and Ogus 1978] P. Berthelot and A. Ogus, *Notes on crystalline cohomology*, Princeton University Press, 1978. MR 58 #10908 Zbl 0383.14010

[Bloch and Kato 1990] S. Bloch and K. Kato, "L-functions and Tamagawa numbers of motives", pp. 333–400 in *The Grothendieck Festschrift*, vol. 1, edited by P. Cartier et al., Progr. Math. **86**, Birkhäuser, Boston, MA, 1990. MR 92g:11063 Zbl 0768.14001

[Buzzard 2007] K. Buzzard, "Eigenvarieties", pp. 59–120 in *L-functions and Galois representations* (Durham, 2004), 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 groupes unitaires", preprint, École Polytechnique, Paris, 2009, http://www.math.polytechnique.fr/~chenevier/articles/famgal.pdf.

[Chenevier 2010] G. Chenevier, "Sur la densité des représentations cristallines du groupe de galois absolu de $\mathbb{Q}_p$", preprint, 2010. arXiv 1012.2852

[Colmez 2005] P. Colmez, "Série principale unitaire pour $\mathrm{GL}_2(\mathbb{Q}_p)$ et représentations triangulines de dimension 2", preprint, 2005, http://www.math.jussieu.fr/~colmez/trianguline.pdf.

[Colmez 2008] P. Colmez, "Représentations triangulines de dimension 2", pp. 213–258 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et* (ϕ, Γ)*-modules*, edited by L. Berger et al., Astérisque **319**, Société Mathématique de France, Paris, 2008. MR 2010f:11173 Zbl 1168.11022

[Colmez 2010] P. Colmez, "La série principale unitaire de $\mathrm{GL}_2(\mathbf{Q}_p)$", pp. 213–262 in *Représentations p-adiques de groupes p-adiques, II: Représentations de* $\mathrm{GL}_2(\mathbf{Q}_p)$ *et* (ϕ, Γ)*-modules]*, edited by L. Berger et al., Astérisque **330**, Société Mathématique de France, Paris, 2010. MR 2011g:22026 Zbl 1242.11095

[Colmez and Fontaine 2000] P. Colmez and J.-M. Fontaine, "Construction des représentations p-adiques semi-stables", *Invent. Math.* **140**:1 (2000), 1–43. MR 2001g:11184 Zbl 1010.14004

[Flach 1990] M. Flach, "A generalisation of the Cassels–Tate pairing", *J. Reine Angew. Math.* **412** (1990), 113–127. MR 92b:11037 Zbl 0711.14001

[Fontaine and Perrin-Riou 1994] J.-M. Fontaine and B. Perrin-Riou, "Autour des conjectures de Bloch et Kato: cohomologie galoisienne et valeurs de fonctions L", pp. 599–706 in *Motives* (Seattle, WA, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, American Mathematical Society, Providence, RI, 1994. MR 95j:11046 Zbl 0821.14013

[Fukaya and Kato 2006] T. Fukaya and K. Kato, "A formulation of conjectures on p-adic zeta functions in noncommutative Iwasawa theory", pp. 1–85 in *Proceedings of the St. Petersburg Mathematical Society*, vol. 12, edited by N. N. Uraltseva, Amer. Math. Soc. Transl. Ser. 2 **219**, American Mathematical Society, Providence, RI, 2006. MR 2007k:11200 Zbl 1238.11105

[Greenberg 1989] R. Greenberg, “Iwasawa theory for p-adic representations”, pp. 97–137 in *Algebraic number theory*, edited by J. Coates et al., Adv. Stud. Pure Math. **17**, Academic Press, Boston, MA, 1989. MR 92c:11116 Zbl 0739.11045

[Greenberg 1994a] R. Greenberg, “Iwasawa theory and p-adic deformations of motives”, pp. 193–223 in *Motives* (Seattle, WA, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, American Mathematical Society, Providence, RI, 1994. MR 95i:11053 Zbl 0819.11046

[Greenberg 1994b] R. Greenberg, “Trivial zeros of p-adic L-functions”, pp. 149–174 in *p-adic monodromy and the Birch and Swinnerton–Dyer conjecture* (Boston, MA, 1991), edited by B. Mazur and G. Stevens, Contemp. Math. **165**, American Mathematical Society, Providence, RI, 1994. MR 95h:11063 Zbl 0838.11070

[Hellmann 2010] E. Hellmann, “On arithmetic families of filtered phi-modules and crystalline representations”, preprint, 2010. arXiv 1010.4577

[Hellmann 2012a] E. Hellmann, “Families of p-adic Galois representations and (φ, Γ)-modules”, preprint, 2012. arXiv 1202.3413

[Hellmann 2012b] E. Hellmann, “Families of trianguline representations and finite slope subspaces”, preprint, 2012. arXiv 1202.4408

[Herr 1998] L. Herr, “Sur la cohomologie galoisienne des corps p-adiques”, *Bull. Soc. Math. France* **126**:4 (1998), 563–600. MR 2000m:11118 Zbl 0967.11050

[de Jong 1995] A. J. de Jong, “Crystalline Dieudonné module theory via formal and rigid geometry”, *Inst. Hautes Études Sci. Publ. Math.* **82** (1995), 5–96. MR 97f:14047 Zbl 0864.14009

[Kato 2004] K. Kato, “p-adic Hodge theory and values of zeta functions of modular forms”, pp. 117–290 in *Cohomologies p-adiques et applications arithmétiques*, vol. 3, edited by P. Berthelot et al., Astérisque **295**, 2004. MR 2006b:11051 Zbl 1142.11336

[Kedlaya 2008] K. S. Kedlaya, “Slope filtrations for relative Frobenius”, pp. 259–301 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 2010c:14024 Zbl 1168.11053

[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 05852025

[Kedlaya et al. 2012] S. Kedlaya, Kiran, J. Pottharst, and L. Xiao, “Cohomology of arithmetic families of (φ, Γ)-modules”, preprint, 2012. arXiv 1203.5718

[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 2007] R. Liu, “Cohomology and duality for (ϕ, Γ)-modules over the Robba ring”, *Int. Math. Res. Not.* **2007**:3 (2007), Art. ID rnm150. MR 2009e:11222 Zbl 1248.11093

[Liu 2012] R. Liu, “Triangulation of refined families”, preprint, 2012. arXiv 1202.2188

[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

[Mazur 1972] B. Mazur, “Rational points of abelian varieties with values in towers of number fields”, *Invent. Math.* **18** (1972), 183–266. MR 56 #3020 Zbl 0245.14015

[Mazur et al. 1986] B. Mazur, J. Tate, and J. Teitelbaum, “On p-adic analogues of the conjectures of Birch and Swinnerton–Dyer”, *Invent. Math.* **84**:1 (1986), 1–48. MR 87e:11076 Zbl 0699.14028

[Nekovář 2006] J. Nekovář, *Selmer complexes*, Astérisque **310**, 2006. MR 2009c:11176 Zbl 1211 11120

[Neukirch et al. 2008] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, 2nd ed., Grundlehren Math. Wiss. **323**, Springer, Berlin, 2008. MR 2008m:11223 Zbl 1136.11001

[Perrin-Riou 1994a] B. Perrin-Riou, "Représentations p-adiques ordinaires", pp. 185–220 in *Périodes p-adiques* (Bures-sur-Yvette, 1988), Astérisque **223**, 1994. MR 96h:11063 Zbl 1043.11532

[Perrin-Riou 1994b] B. Perrin-Riou, "Théorie d'Iwasawa des représentations p-adiques sur un corps local", *Invent. Math.* **115**:1 (1994), 81–161. MR 95c:11082 Zbl 0838.11071

[Perrin-Riou 2000] B. Perrin-Riou, *p-adic L-functions and p-adic representations*, SMF/AMS Texts and Monographs **3**, American Mathematical Society, Providence, RI, 2000. MR 2000k:11077 Zbl 0988.11055

[Pottharst 2012] J. Pottharst, "Cyclotomic Iwasawa theory of motives", preprint, 2012, http://math.bu.edu/people/potthars/writings/cyc.pdf.

[Schneider 1987] P. Schneider, "Arithmetic of formal groups and applications, I: Universal norm subgroups", *Invent. Math.* **87**:3 (1987), 587–602. MR 88e:11044 Zbl 0608.14034

[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

[Višik 1976] M. M. Višik, "Nonarchimedean measures associated with Dirichlet series", *Mat. Sb. (N.S.)* **99(141)**:2 (1976), 248–260. In Russian; translated in *Math. USSR Sb.* **28**:2 (1976), 216–228. MR 54 #243 Zbl 0358.14014

Communicated by John Henry Coates
Received 2011-08-17 Revised 2012-07-28 Accepted 2012-12-17

potthars@bu.edu *Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA, 02215, United States*

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1613

Multiplicative excellent families of elliptic surfaces of type E_7 or E_8

Abhinav Kumar and Tetsuji Shioda

We describe explicit multiplicative excellent families of rational elliptic surfaces with Galois group isomorphic to the Weyl group of the root lattices E_7 or E_8. The Weierstrass coefficients of each family are related by an invertible polynomial transformation to the generators of the multiplicative invariant ring of the associated Weyl group, given by the fundamental characters of the corresponding Lie group. As an application, we give examples of elliptic surfaces with multiplicative reduction and all sections defined over $\mathbb{Q}$ for most of the entries of fiber configurations and Mordell–Weil lattices described by Oguiso and Shioda, as well as examples of explicit polynomials with Galois group $W(E_7)$ or $W(E_8)$.

1. Introduction

For an elliptic curve E over a field K, determining its Mordell–Weil group is a fundamental problem in algebraic geometry and number theory. When $K = k(t)$ is a rational function field in one variable, this problem becomes a geometrical one of understanding sections of an elliptic surface with section. Lattice theoretic methods of attack were described in [Shioda 1990]. In particular, when $\mathscr{E} \to \mathbb{P}^1_t$ is a rational elliptic surface given as a minimal proper model of

$$y^2 + a_1(t)xy + a_3(t)y = x^3 + a_2(t)x^2 + a_4(t)x + a_6(t)$$

with $a_i(t) \in k[t]$ of degree at most i, the possible configurations (types) of bad fibers and Mordell–Weil groups were analyzed by Oguiso and Shioda [1991].

In [Shioda 1991a], the second author studied sections for some families of elliptic surfaces with an additive fiber, by means of the specialization map, and obtained a relation between the coefficients of the Weierstrass equation and the fundamental invariants of the corresponding Weyl groups. Shioda and Usui [1992] expanded

Kumar was supported in part by NSF Career grant DMS-0952486, and by a grant from the Solomon Buchsbaum Research Fund. Shioda was partially supported by JSPS Grant-in-Aid for Scientific Research (C)20540051.
MSC2010: primary 14J27; secondary 11G05, 12F10, 13A50.
Keywords: rational elliptic surfaces, multiplicative invariants, inverse Galois problem, Weyl group, Mordell–Weil group.

this by studying families with a bad fiber of additive reduction more exhaustively. They defined the formal notion of an excellent family (see Section 2) and found excellent families for many of the "admissible" types.

The analysis of rational elliptic surfaces of high Mordell–Weil rank, but with a fiber of multiplicative reduction, is much more challenging. However, understanding this situation is arguably more fundamental, since if we write down a "random" elliptic surface, then with probability close to 1 it will have Mordell–Weil lattice E_8 and twelve nodal fibers (that is, of multiplicative reduction). To be more precise, if we choose Weierstrass coefficients $a_i(t)$ of degree i, with coefficients chosen uniformly at random from among rational numbers (say) of height at most N, then as $N \to \infty$ the surface will satisfy the condition above with probability approaching 1. One can make a similar statement for rational elliptic surfaces chosen to have Mordell–Weil lattice E_7^*, E_6^*, etc.

In [Shioda 2012], this study was carried out for elliptic surfaces with a fiber of type I_3 and Mordell–Weil lattice isometric to E_6^*, through a "multiplicative excellent family" of type E_6. We will describe this case briefly in Section 3. The main result of this paper shows that two explicitly described families of rational elliptic surfaces with Mordell–Weil lattices E_7^* or E_8 are multiplicative excellent. The proof involves a surprising connection with representation theory of the corresponding Lie groups, and in particular, their fundamental characters. In particular, we deduce that the Weierstrass coefficients give another natural set of generators for the multiplicative invariants of the respective Weyl groups, as a polynomial ring. Similar formulas were derived by Eguchi and Sakai [2003] using calculations from string theory and mirror symmetry.

The idea of an excellent family is quite useful and important in number theory. An excellent family of algebraic varieties leads to a Galois extension $F(\mu)/F(\lambda)$ of two purely transcendental extensions of a number field F (say $\mathbb{Q}$), with Galois group a desired finite group G. This setup has an immediate number-theoretic application, since one may specialize the parameters λ and apply Hilbert's irreducibility theorem to obtain Galois extensions over $\mathbb{Q}$ with the same Galois group. Furthermore, we can make the construction effective if appropriate properties of the group G are known (see Examples 8 and 19 for the case $G = W(E_7)$ or $W(E_8)$). At the same time, an excellent family will give rise to a split situation very easily, by specializing the parameters μ instead. For examples, in the situation considered in our paper, we obtain elliptic curves over $\mathbb{Q}(t)$ with Mordell–Weil rank 7 or 8 together with explicit generators for the Mordell–Weil group (see Examples 7 and 18). There are also applications to geometric specialization or degeneration of the family. Therefore, it is desirable (but quite nontrivial) to construct explicit excellent families of algebraic varieties. Such a situation is quite rare in general: theoretically, any finite reflection group is a candidate, but it is not generally neatly related to an algebraic geometric family. Hilbert treated the case of the symmetric group S_n, corresponding to families of zero-dimensional varieties. Not many examples were known before the (additive) excellent families for the Weyl groups of the exceptional Lie groups E_6, E_7 and E_8 were given in [Shioda 1991a], using the theory of Mordell–Weil lattices. Here, we finish the story for the multiplicative excellent families for these Weyl groups.

2. Mordell–Weil lattices and excellent families

Let $X \xrightarrow{\pi} \mathbb{P}^1$ be an elliptic surface with section $\sigma : \mathbb{P}^1 \to X$, that is, a proper relatively minimal model of its generic fiber, which is an elliptic curve. We denote the image of σ by O, which we take to be the zero section of the Néron model. We let F be the class of a fiber in $\mathrm{Pic}(X) \cong \mathrm{NS}(X)$, and let the reducible fibers of π lie over $v_1, \ldots, v_k \in \mathbb{P}^1$. The nonidentity components of $\pi^{-1}(v_i)$ give rise to a sublattice T_i of $\mathrm{NS}(X)$, which is (the negative of) a root lattice (see [Kodaira 1963a; 1963b; Tate 1975]). The *trivial lattice* T is $\mathbb{Z}O \oplus \mathbb{Z}F \oplus (\bigoplus T_i)$, and we have the isomorphism $\mathrm{MW}(X/\mathbb{P}^1) \cong \mathrm{NS}(X)/T$, which describes the Mordell–Weil group. In fact, one can induce a positive definite pairing on the Mordell–Weil group modulo torsion, by inducing it from the negative of the intersection pairing on $\mathrm{NS}(X)$. We refer the reader to [Shioda 1990] for more details. In this paper, we will call $\bigoplus T_i$ the *fibral lattice*.

Next, we recall from [Shioda and Usui 1992] the notion of an *excellent family* with Galois group G. Suppose $X \to \mathbb{A}^n$ is a family of algebraic varieties, varying with respect to n parameters $\lambda_1, \ldots, \lambda_n$. The generic member of this family X_λ is a variety over the rational function field $k_0 = \mathbb{Q}(\lambda)$. Let $k = \overline{k_0}$ be the algebraic closure, and suppose that $\mathscr{C}(X_\lambda)$ is a group of algebraic cycles on X_λ over the field k (in other words, it is a group of algebraic cycles on $X_\lambda \times_{k_0} k$). Suppose in addition that there is an isomorphism $\phi_\lambda : \mathscr{C}(X_\lambda) \otimes \mathbb{Q} \cong V$ for a fixed vector space V, and $\mathscr{C}(X_\lambda)$ is preserved by the Galois group $\mathrm{Gal}(k/k_0)$. Then we have the Galois representation

$$\rho_\lambda : \mathrm{Gal}(k/k_0) \to \mathrm{Aut}(\mathscr{C}(X_\lambda)) \to \mathrm{Aut}(V).$$

We let k_λ be the fixed field of the kernel of ρ_λ, that is, it is the smallest extension of k_0 over which the cycles of $\mathscr{C}(\lambda)$ are defined. We call it the *splitting field* of $\mathscr{C}(X_\lambda)$.

Now let G be a finite reflection group acting on the space V.

Definition 1. We say $\{X_\lambda\}$ is an *excellent family* with Galois group G if the following conditions hold:

(1) The image of ρ_λ is equal to G.

(2) There is a $\mathrm{Gal}(k/k_0)$-equivariant evaluation map $s : \mathscr{C}(X_\lambda) \to k$.

(3) There exists a basis $\{Z_1, \ldots, Z_n\}$ of $\mathscr{C}(X_\lambda)$ such that if we set $u_i = s(Z_i)$, then $u_1, \ldots, u_n$ are algebraically independent over $\mathbb{Q}$.

(4) $\mathbb{Q}[u_1, \ldots, u_n]^G = \mathbb{Q}[\lambda_1, \ldots, \lambda_n]$.

As an example, for $G = W(E_8)$, consider the following family of rational elliptic surfaces over $k_0 = \mathbb{Q}(\lambda)$:

$$y^2 = x^3 + x\Big(\sum_{i=0}^{3} p_{20-6i} t^i\Big) + \Big(\sum_{j=0}^{3} p_{30-6j} t^j + t^5\Big),$$

with $\lambda = (p_2, p_8, p_{12}, p_{14}, p_{18}, p_{20}, p_{24}, p_{30})$. Shioda [1991a] shows that this is an excellent family with Galois group G. The p_i are related to the fundamental invariants of the Weyl group of E_8, as is suggested by their degrees (subscripts).

We now define the notion of a *multiplicative excellent family* for a group G. As before, $X \to \mathbb{A}^n$ is a family of algebraic varieties, varying with respect to $\lambda = (\lambda_1, \ldots, \lambda_n)$, and $\mathscr{C}(X_\lambda)$ is a group of algebraic cycles on X_λ, isomorphic (via a fixed isomorphism) to a fixed abelian group M. The fields k_0 and k are as before, and we have a Galois representation

$$\rho_\lambda : \mathrm{Gal}(k/k_0) \to \mathrm{Aut}(\mathscr{C}(X_\lambda)) \to \mathrm{Aut}(M).$$

Suppose that G is a group acting on M.

Definition 2. We say $\{X_\lambda\}$ is a *multiplicative excellent family* with Galois group G if the following conditions hold:

(1) The image of ρ_λ is equal to G.

(2) There is a $\mathrm{Gal}(k/k_0)$-equivariant evaluation map $s : \mathscr{C}(X_\lambda) \to k^*$.

(3) There exists a basis $\{Z_1, \ldots, Z_n\}$ of $\mathscr{C}(X_\lambda)$ such that if we set $u_i = s(Z_i)$, then $u_1, \ldots, u_n$ are algebraically independent over $\mathbb{Q}$.

(4) $\mathbb{Q}[u_1, \ldots, u_n, u_1^{-1}, \ldots, u_n^{-1}]^G = \mathbb{Q}[\lambda_1, \ldots, \lambda_n]$.

Remark 3. Though we use similar notation, the specialization map s and the u_i in the multiplicative case are quite different from the ones in the additive case. Intuitively, one may think of these as exponentiated versions of the corresponding objects in the additive case. However, we want the specialization map to be an algebraic morphism, and so in general (additive) excellent families specified by Definition 1 will be very different from multiplicative excellent families specified by Definition 2.

In our examples, G will be a finite reflection group acting on a lattice in Euclidean space, which will be our choice for M. However, what is relevant here is not the ring of (additive) invariants of G on the vector space spanned by M. Instead, note that the action of G on M gives rise to a “multiplicative” or “monomial” action of G on the group algebra $\mathbb{Q}[M]$, and we will be interested in the polynomials on this space that are invariant under G. This is the subject of *multiplicative invariant theory* (see, for example, [Lorenz 2005]). In the case when G is the automorphism group of a root lattice or root system, multiplicative invariants were classically studied by using the terminology of "exponentiated" roots e^{α} (for instance, see [Bourbaki 1968, Section VI.3]).

3. The E_6 case

We now sketch the construction of multiplicative excellent family in [Shioda 2012]. Consider the family of rational elliptic surfaces S_λ with Weierstrass equation

$$y^2 + txy = x^3 + (p_0 + p_1 t + p_2 t^2)\,x + q_0 + q_1 t + q_2 t^2 + t^3$$

with parameter $\lambda = (p_0, p_1, p_2, q_0, q_1, q_2)$. The surface S_λ generically only has one reducible fiber at $t = \infty$, of type I_3. Therefore, the Mordell–Weil lattice M_λ of S_λ is isomorphic to E_6^*. There are 54 minimal sections of height $4/3$, and exactly half of them have the property that x and y are linear in t. If we have

$$x = at + b \quad \text{and} \quad y = ct + d,$$

then substituting these back in to the Weierstrass equation, we get a system of equations, and we may easily eliminate b, c, d from the system to get a monic equation of degree 27 (subject to a genericity assumption), which we write as $\Phi_\lambda(a) = 0$. Also, note that the specialization of a section of height $4/3$ to the fiber at ∞ gives us a point on one of the two nonidentity components of the special fiber of the Néron model (the same component for all the 27 sections). Identifying the smooth points of this component with $\mathbb{G}_m \times \{1\} \subset \mathbb{G}_m \times (\mathbb{Z}/3\mathbb{Z})$, the specialization map s takes the section to $(-1/a, 1)$. Let the specializations be $s_i = -1/a_i$ for $1 \le i \le 27$. We have

$$\begin{aligned}\Phi_\lambda(X) &= \prod_{i=1}^{27}(X - a_i) = \prod_{i=1}^{27}(X + 1/s_i)\\ &= X^{27} + \epsilon_{-1}X^{26} + \epsilon_{-2}X^{25} + \cdots + \epsilon_4 X^4 + \epsilon_3 X^3 + \epsilon_2 X^2 + \epsilon_1 X + 1.\end{aligned}$$

Here ϵ_i is the i-th elementary symmetric polynomial of the s_i and ϵ_{-i} that of the $1/s_i$. The coefficients of $\Phi_\lambda(X)$ are polynomials in the coordinates of λ, and we may use the equations for $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_{-1}$ and ϵ_{-2} to solve for $p_0, \dots, q_3$. However, the resulting solution has ϵ_{-2} in the denominator. We may remedy this situation as follows. Consider the construction of E_6^* as described in [Shioda 1995]: let $v_1, \dots, v_6$ be vectors in $\mathbb{R}^6$ with $\langle v_i, v_j\rangle = \delta_{ij} + 1/3$, and let $u = (\sum v_i)/3$. The $\mathbb{Z}$-span of $v_1, \dots, v_6, u$ is a lattice L isometric to E_6^*. It is clear that $v_1, \dots, v_5, u$

forms a basis of L. Here, we choose an isometry between the Mordell–Weil lattice and the lattice L, and let the specializations of $v_1, \ldots, v_6, u$ be $s_1, \ldots, s_6, r$, respectively. These satisfy $s_1 s_2 \ldots s_6 = r^3$. The 54 nonzero minimal vectors of E_6^* split up into two cosets (modulo E_6) of 27 each, of which we have chosen one. The specializations of these 27 special sections are given by

$$\{s_1, \ldots, s_{27}\} := \{s_i : 1 \le i \le 6\} \cup \{s_i/r : 1 \le i \le 6\} \cup \{r/(s_i s_j) : 1 \le i < j \le 6\}.$$

If

$$\delta_1 = r + \frac{1}{r} + \sum_{i \ne j} \frac{s_i}{s_j} + \sum_{i<j<k} \Big(\frac{r}{s_i s_j s_k} + \frac{s_i s_j s_k}{r} \Big),$$

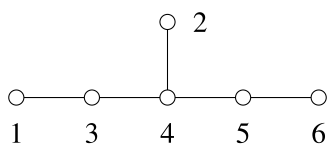
then δ_1 belongs to the $G = W(E_6)$-invariants of $\mathbb{Q}[s_1, \ldots, s_5, r, s_1^{-1}, \ldots, s_5^{-1}, r^{-1}]$, and explicit computations in [Shioda 2012] show that

$$\begin{aligned} \mathbb{Q}[s_1, \ldots, s_5, r, s_1^{-1}, \ldots, s_5^{-1}, r^{-1}]^G &= \mathbb{Q}[\delta_1, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_{-1}, \epsilon_{-2}] \\ &= \mathbb{Q}[p_0, p_1, p_2, q_0, q_1, q_2]. \end{aligned}$$

The explicit relation showing the second equality is as follows:

$$\begin{aligned} \delta_1 &= -2p_1, & \epsilon_{-1} &= p_2^2 - q_2, \\ \epsilon_2 &= 13p_2^2 + p_0 - q_2, & \epsilon_{-2} &= -2p_1 p_2 + 6p_2 + q_1, \\ \epsilon_1 &= 6p_2, & \epsilon_3 &= 8p_2^3 + 2p_0 p_2 + p_1^2 - 6p_1 - q_0 + 9. \end{aligned}$$

We make an additional observation. The six fundamental representations of the Lie algebra E_6 correspond to the fundamental weights in the following diagram, which displays the standard labeling of these representations.



The dimensions of these representations $V_1, \ldots, V_6$ are 27, 78, 351, 2925, 351, 27 respectively, and their characters are related to $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_{-1}, \epsilon_{-2}, \delta_1$ by the nice transformation

$$\begin{aligned} \chi_1 &= \epsilon_1, & \chi_2 &= \delta_1 + 6, & \chi_3 &= \epsilon_2, \\ \chi_4 &= \epsilon_3, & \chi_5 &= \epsilon_{-2}, & \chi_6 &= \epsilon_{-1}. \end{aligned}$$

This explains the reason for bringing in δ_1 into the picture, and also why there is a denominator when solving for $p_0, \ldots, q_2$ in terms of $\epsilon_1, \ldots, \epsilon_4, \epsilon_{-1}, \epsilon_{-2}$, as remarked in [Shioda 2012]. The coefficients ϵ_j are essentially the characters of $\bigwedge^j V$, where $V = V_1$ is the first fundamental representation, while ϵ_{-j} are those of $\bigwedge^j V^*$, where $V_6 = V^*$. Note that $\bigwedge^3 V \cong \bigwedge^3 V^*$. Therefore, from the expressions for $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_{-1}, \epsilon_{-2}$, we may obtain p_2, q_2, p_0, q_1, q_0, in terms of the remaining variable p_1, without introducing any denominators. However, representation V_2 cannot be obtained as a direct summand with multiplicity 1 from a tensor product of $\bigwedge^j V$ (for $1 \le j \le 3$) and $\bigwedge^k V^*$ (for $1 \le k \le 2$). On the other hand, we do have the isomorphism

$$(V_2 \otimes V_5) \oplus V_5 \oplus V_1 \cong \textstyle\bigwedge^4 V_1 \oplus (V_3 \otimes V_6) \oplus (V_6 \otimes V_6).$$

Therefore, we are able to solve for p_1 if we introduce a denominator of ϵ_{-2}, which is the character of V_5.

4. The E_7 case

4.1. *Results.* Next, we exhibit a multiplicative excellent family for the Weyl group of E_7. It is given by the Weierstrass equation

$$y^2 + txy = x^3 + (p_0 + p_1 t + p_2 t^2)\, x + q_0 + q_1 t + q_2 t^2 + q_3 t^3 - t^4.$$

For generic $\lambda = (p_0, \dots, p_2, q_0, \dots, q_3)$, this rational elliptic surface X_λ has a fiber of type I_2 at $t = \infty$, and no other reducible fibers. Hence, the Mordell–Weil group M_λ is E_7^*. We note that any elliptic surface with a fiber of type I_2 can be put into this Weierstrass form (in general over a small degree algebraic extension of the ground field), after a fractional linear transformation of the parameter t, and Weierstrass transformations of x and y.

Lemma 4. *The smooth part of the special fiber is isomorphic to the group scheme $\mathbb{G}_m \times \mathbb{Z}/2\mathbb{Z}$. The identity component is the nonsingular part of the curve*

$$y^2 + xy = x^3.$$

The x- and y-coordinates of a section of height 2 *are polynomials of degrees* 2 *and* 3 *respectively, and its specialization at $t = \infty$ is* $(\lim_{t\to\infty}(y + tx)/y, 0) \in k^* \times \{0, 1\}$. *A section of height* $3/2$ *has x and y coordinates of the form*

$$x = at + b \quad \textit{and} \quad y = ct^2 + dt + e.$$

and specializes at $t = \infty$ to $(c, 1)$.

Proof. First, to get a local chart for the elliptic surface near $t = \infty$, we set $x = \tilde{x}/u^2$, $y = \tilde{y}/u^3$ and $t = 1/u$, and look for u near 0. Therefore, the special fiber (before blow-up) is given by $\bar{y}^2 + \bar{x}\bar{y} = \bar{x}^3$, where $\bar{x} = \tilde{x}|_{u=0}$ and $\bar{y} = \tilde{y}|_{u=0}$ are the reductions of the coordinates at $u = 0$, respectively. It is an easy exercise to parametrize the smooth locus of this curve: It is given, for instance, by $\bar{x} = s/(s-1)^2$, $\bar{y} = s/(s-1)^3$. We then check that $s = (\bar{y} + \bar{x})/\bar{y}$ and the map from the smooth locus to $\mathbb{G}_m$ that takes the point $(\bar{x}, \bar{y})$ to s is a homomorphism from the secant group law to multiplication in k^*. This proves the first half of the lemma. Note that we could just as well have taken $1/s$ to be the parameter on $\mathbb{G}_m$; our choice is a matter of convention. To prove the specialization law for sections of height $3/2$, we may, for instance, take the sum of such a section Q with a section P of height 2 with specialization $(s, 0)$. A direct calculation shows that the y-coordinate of the sum has top (quadratic) coefficient cs. Therefore the specialization of Q must have the form κc, where κ is a constant not depending on Q. Finally, the sum of two sections Q_1 and Q_2 of height $3/2$ and having coefficients c_1 and c_2 for the t^2 term of their y-coordinates can be checked to specialize to $(c_1c_2, 0)$. It follows that $\kappa = \pm 1$, and we take the plus sign as a convention. (It is easy to see that both choices of sign are legitimate, since the sections of height 2 generate a copy of E_7, whereas the sections of height $3/2$ lie in the nontrivial coset of E_7 in E_7^*). $\square$

There are 56 sections of height $3/2$, with x and y coordinates in the form above. Substituting the formulas above for x and y into the Weierstrass equation, we get the following system of equations.

$$\begin{aligned} c^2 + ac + 1 &= 0, \\ q_3 + ap_2 + a^3 &= (2c+a)d + bc, \\ q_2 + bp_2 + 3a^2b &= (2c+a)e + (b+d)d, \\ q_1 + bp_1 + ap_0 + 3ab^2 &= (2d+b)e, \\ q_0 + bp_0 + b^3 &= e^2. \end{aligned}$$

We solve for a and b from the first and second equations, and then e from the third, assuming $c \neq 1$. Substituting these values back into the last two equations, we get two equations in the variables c and d. Taking the resultant of these two equations with respect to d, and dividing by $c^{30}(c^2-1)^4$, we obtain an equation of degree 56 in c, which is monic, reciprocal and has coefficients in $\mathbb{Z}[\lambda] = \mathbb{Z}[p_0, \ldots, q_3]$. We denote this polynomial by

$$\Phi_\lambda(X) = \prod_{i=1}^{56} (X - s(P)) = X^{56} + \epsilon_1 X^{55} + \epsilon_2 X^{54} + \cdots + \epsilon_1 X + \epsilon_0,$$

where P ranges over the 56 minimal sections of height $3/2$. It is clear that a, b, d, e are rational functions of c with coefficients in k_0.

We have a Galois representation on the Mordell–Weil lattice

$$\rho_\lambda : \mathrm{Gal}(k/k_0) \to \mathrm{Aut}(M_\lambda) \cong \mathrm{Aut}(E_7^*).$$

Here $\mathrm{Aut}(E_7^*) \cong \mathrm{Aut}(E_7) \cong W(E_7)$, the Weyl group of type E_7. The *splitting field* of M_λ is the fixed field k_λ of $\mathrm{Ker}(\rho_\lambda)$. By definition, $\mathrm{Gal}(k_\lambda/k_0) \cong \mathrm{Im}(\rho_\lambda)$. The splitting field k_λ is equal to the splitting field of the polynomial $\Phi_\lambda(X)$ over k_0, since the Mordell–Weil group is generated by the 56 sections of smallest height $P_i = (a_i t + b_i, c_i t^2 + d_i t + e_i)$. We also have $k_\lambda = k_0(P_1, \dots, P_{56}) = k_0(c_1, \dots, c_{56})$. We shall sometimes write e^α, (for $\alpha \in E_7^*$ a minimal vector) to refer to the specializations of these sections $c(P_i)$, for convenience.

Theorem 5. *Assume that λ is generic over $\mathbb{Q}$, i.e., the coordinates $p_0, \dots, q_3$ are algebraically independent over $\mathbb{Q}$.*

(1) *ρ_λ induces an isomorphism* $\mathrm{Gal}(k_\lambda/k_0) \cong W(E_7)$.

(2) *The splitting field k_λ is a purely transcendental extension of $\mathbb{Q}$, isomorphic to the function field $\mathbb{Q}(Y)$ of the toric hypersurface*

$$Y \subset \mathbb{G}_m^8 \quad \text{defined by } s_1 \dots s_7 = r^3.$$

There is an action of $W(E_7)$ on Y such that $\mathbb{Q}(Y)^{W(E_7)} = k_\lambda^{W(E_7)} = k_0$.

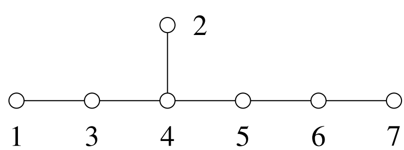
(3) *The ring of $W(E_7)$-invariants in the affine coordinate ring*

$$\mathbb{Q}[Y] = \frac{\mathbb{Q}[s_i, r, 1/s_i, 1/r]}{(s_1 \dots s_7 - r^3)} \cong \mathbb{Q}[s_1, \dots, s_6, r, s_1^{-1}, \dots, s_6^{-1}, r^{-1}]$$

is equal to the polynomial ring $\mathbb{Q}[\lambda]$:

$$\mathbb{Q}[Y]^{W(E_7)} = \mathbb{Q}[\lambda] = \mathbb{Q}[p_0, p_1, p_2, q_0, q_1, q_2, q_3].$$

In fact, we shall prove an explicit, invertible polynomial relation between the Weierstrass coefficients λ and the fundamental characters for E_7. Let $V_1, \dots, V_7$ be the fundamental representations of E_7, and $\chi_1, \dots, \chi_7$ their characters (on a maximal torus), as labeled below. For a description of the fundamental modules for the exceptional Lie groups see [Carter 2005, Section 13.8].



Note that since the weight lattice E_7^* has been equipped with a nice set of generators $(v_1, \dots, v_7, u)$ with $\sum v_i = 3u$ (as in [Shioda 1995]), the characters $\chi_1, \dots, \chi_7$ lie in the ring of Laurent polynomials $\mathbb{Q}[s_i, r, 1/s_i, 1/r]$, where s_i corresponds to e^{v_i} and r to e^u, and are obviously invariant under the (multiplicative) action of the Weyl group on this ring of Laurent polynomials. Explicit formulas for the χ_i are given in the auxiliary files.

We also note that the roots of Φ_λ are given by

$$s_i, \frac{1}{s_i} \quad \text{for } 1 \le i \le 7 \quad \text{and} \quad \frac{r}{s_i s_j}, \frac{s_i s_j}{r} \quad \text{for } 1 \le i < j \le 7.$$

Theorem 6. *For generic λ over $\mathbb{Q}$, we have*

$$\mathbb{Q}[\chi_1, \dots, \chi_7] = \mathbb{Q}[p_0, p_1, p_2, q_0, q_1, q_2, q_3].$$

The transformation between these sets of generators is

$$\begin{aligned}
\chi_1 &= 6p_2 + 25,\\
\chi_2 &= 6q_3 - 2p_1,\\
\chi_3 &= -q_2 + 13p_2^2 + 108p_2 + p_0 + 221,\\
\chi_4 &= 9q_3^2 - 6p_1q_3 - q_2 - q_0 + 8p_2^3 + 85p_2^2 + (2p_0 + 300)p_2 + p_1^2 + 10p_0 + 350,\\
\chi_5 &= (6p_2 + 26)q_3 + q_1 - 2p_1p_2 - 10p_1,\\
\chi_6 &= -q_2 + p_2^2 + 12p_2 + 27,\\
\chi_7 &= q_3,
\end{aligned}$$

with inverse

$$\begin{aligned}
p_2 &= (\chi_1 - 25)/6,\\
p_1 &= (6\chi_7 - \chi_2)/2,\\
p_0 &= -(3\chi_6 - 3\chi_3 + \chi_1^2 - 2\chi_1 + 7)/3,\\
q_3 &= \chi_7,\\
q_2 &= -(36\chi_6 - \chi_1^2 - 22\chi_1 + 203)/36,\\
q_1 &= (24\chi_7 + 6\chi_5 + (-\chi_1 - 5)\chi_2)/6,\\
q_0 &= (27\chi_2^2 - 8\chi_1^3 - 84\chi_1^2 + 120\chi_1 - 136)/108 - (\chi_1 + 2)\chi_6/3 - \chi_4 + (\chi_1 + 5)\chi_3/3.
\end{aligned}$$

Our formulas agree with those of Eguchi and Sakai [2003], who seem to derive these by using an ansatz.

Next, we describe two examples through specialization, one of "small Galois" (in which all sections are defined over $\mathbb{Q}[t]$) and one with "big Galois" (which has Galois group the full Weyl group).

Example 7. The values

$$\begin{aligned}
p_0 &= 244655370905444111/(3\mu^2),\\
p_1 &= -47883695294816415251 25/(16\mu^2),\\
q_3 &= 184185687325/(4\mu),\\
p_2 &= 19993710659027964447503892495507 6599/(12\mu^4),\\
q_2 &= 57918534120411335989995011407800421/(9\mu^3),\\
q_1 &= -17988091661721362494887555650280856 0625/(4\mu^4),\\
q_0 &= 353161437549197551159528020804697629366268908804692010 91/(1728\mu^6),
\end{aligned}$$

where $\mu = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 = 102102$, give rise to an elliptic surface for which we have $r = 2$, $s_1 = 3$, $s_2 = 5$, $s_3 = 7$, $s_4 = 11$, $s_5 = 13$, $s_6 = 17$, the simplest choice of multiplicatively independent elements. The Mordell–Weil group has a basis of sections for which $c \in \{3, 5, 7, 11, 13, 17, 15/2\}$. We write down their x-coordinates below:

$$\begin{aligned}
x(P_1) &= -(10/3)t - 707606695171055129/1563722760600,\\
x(P_2) &= -(26/5)t - 611410735289928023/1563722760600,\\
x(P_3) &= -(50/7)t - 513728975686763429/1563722760600,\\
x(P_4) &= -(122/11)t - 316023939417997169/1563722760600,\\
x(P_5) &= -(170/13)t - 216677827127591279/1563722760600,\\
x(P_6) &= -(290/17)t - 17562556436754779/1563722760600,\\
x(P_7) &= -(229/30)t - 140574879644393807/390930690150.
\end{aligned}$$

In the auxiliary files the x-and y-coordinates are listed, and it is verified that they satisfy the Weierstrass equation.

Example 8. The value $\lambda = \lambda_0 := (1, 1, 1, 1, 1, 1, 1)$ gives rise to an explicit polynomial $f(X) = \Phi_{\lambda_0}(X)$, given by

$$\begin{aligned}
f(X) = X^{56} &- X^{55} + 40X^{54} - 22X^{53} + 797X^{52} - 190X^{51} + 9878X^{50} - 1513X^{49}\\
&+ 82195X^{48} - 17689X^{47} + 496844X^{46} - 175584X^{45} + 2336237X^{44}\\
&- 1196652X^{43} + 8957717X^{42} - 5726683X^{41} + 28574146X^{40}\\
&- 20119954X^{39} + 75465618X^{38} - 53541106X^{37} + 163074206X^{36}\\
&- 110505921X^{35} + 287854250X^{34} - 181247607X^{33} + 420186200X^{32}\\
&- 243591901X^{31} + 518626022X^{30} - 278343633X^{29} + 554315411X^{28}\\
&- 278343633X^{27} + 518626022X^{26} - 243591901X^{25} + 420186200X^{24}\\
&- 181247607X^{23} + 287854250X^{22} - 110505921X^{21} + 163074206X^{20}\\
&- 53541106X^{19} + 75465618X^{18} - 20119954X^{17} + 28574146X^{16}\\
&- 5726683X^{15} + 8957717X^{14} - 1196652X^{13} + 2336237X^{12}\\
&- 175584X^{11} + 496844X^{10} - 17689X^{9} + 82195X^{8} - 1513X^{7}\\
&+ 9878X^{6} - 190X^{5} + 797X^{4} - 22X^{3} + 40X^{2} - X + 1,
\end{aligned}$$

for which we can show that the Galois group is the full group $W(E_7)$, as follows. The reduction of $f(X)$ modulo 7 shows that Frob_7 has cycle decomposition of type $(7)^8$, and similarly, Frob_{101} has cycle decomposition of type $(3)^2(5)^4(15)^2$.

This implies, as in [Shioda 1991b, Example 7.6], that the Galois group is the entire Weyl group.

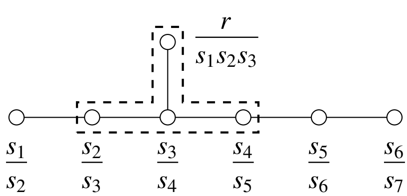
We can also describe degenerations of this family X_λ of rational elliptic surfaces by the method of "vanishing roots", where we drop the genericity assumption, and consider the situation where the elliptic fibration might have additional reducible fibers. Let $\psi : Y \to \mathbb{A}^7$ be the finite surjective morphism associated to

$$\mathbb{Q}[p_0, \dots, q_3] \hookrightarrow \mathbb{Q}[Y] \cong \mathbb{Q}[s_1, \dots, s_6, r, s_1^{-1}, \dots, s_6^{-1}, r^{-1}].$$

For $\xi = (s_1, \dots, s_7, r) \in Y$, let the multiset Π_ξ consist of the 126 elements s_i/r and r/s_i for $1 \le i \le 7$, s_i/s_j ((for $1 \le i \ne j \le 7$) and $s_i s_j s_k/r$ and $r/(s_i s_j s_k)$ for $1 \le i < j < k \le 7$, corresponding to the 126 roots of E_7. Let $2\nu(\xi)$ be the number of times 1 appears in Π_ξ, which is also the multiplicity of 1 as a root of $\Psi_\lambda(X)$ (to be defined in Section 4.2), where $\lambda = \psi(\xi)$. We call the associated roots of E_7 the *vanishing roots*, in analogy with vanishing cycles in the deformation of singularities. By abuse of notation we label the rational elliptic surface X_λ as X_ξ.

Theorem 9. *The surface X_ξ has new reducible fibers (necessarily at $t \ne \infty$) if and only if $\nu(\xi) > 0$. The number of roots in the root lattice T_{new} is equal to $2\nu(\xi)$, where $T_{new} := \bigoplus_{v \ne \infty} T_v$ is the new part of the trivial lattice.*

We may use this result to produce specializations with trivial lattice including A_1, corresponding to the entries in the table of [Oguiso and Shioda 1991, Section 1]. Note that in earlier work [Shioda 1991a; Shioda and Usui 1992], examples of rational elliptic surfaces were produced with a fiber of additive type, for instance, a fiber of type III (which contributes A_1 to the trivial lattice) or a fiber of type II. Using our excellent family, we can produce examples with the A_1 fiber being of multiplicative type I_2 and all other irreducible singular fibers being nodal (that is, I_1). We list below those types that are not already covered by [Shioda 2012]. To produce these examples, we use an embedding of the new part T_{new} of the fibral lattice into E_7, which gives us any extra conditions satisfied by $s_1, \dots, s_7, r$. The following multiplicative version of the labeling of simple roots of E_7 is useful (compare [Shioda 1995]).



For instance, to produce the example in line 18 of the table (that is, with $T_{\mathrm{new}} = D_4$), we may use the embedding into E_7 indicated by embedding the D_4 Dynkin diagram within the dashed lines in the figure above. Thus, we must force $s_2 = s_3 = s_4 = s_5$

Type	Fibral lattice	MW group	$\{s_1, \ldots, s_6, r\}$
2	A_1	E_7^*	3, 5, 7, 11, 13, 17, 2
4	A_1^2	D_6^*	3, 3, 5, 7, 11, 13, 2
7	A_1^3	$D_4^* \oplus A_1^*$	3, 3, 5, 5, 7, 11, 2
10	$A_1 \oplus A_3$	$A_1^* \oplus A_3^*$	3, 3, 3, 3, 5, 7, 2
13	A_1^4	$D_4^* \oplus \mathbb{Z}/2\mathbb{Z}$	−1, 2, 3, 5, 7, 49/30, 7
14	A_1^4	A_1^{*4}	3, 3, 5, 5, 7, 7, 2
17	$A_1 \oplus A_4$	$\frac{1}{10}\begin{pmatrix} 3 & 1 & -1 \\ 1 & 7 & 3 \\ -1 & 3 & 7 \end{pmatrix}$	3, 3, 3, 3, 3, 5, 2
18	$A_1 \oplus D_4$	A_1^{*3}	2, 3, 3, 3, 3, 5, 18
21	$A_1^2 \oplus A_3$	$A_3^* \oplus \mathbb{Z}/2\mathbb{Z}$	3, 5, 60, 30, 30, 30, 900
22	$A_1^2 \oplus A_3$	$A_1^{*2} \oplus \langle 1/4 \rangle$	3, 3, 5, 5, 5, 5, 2
24	A_1^5	$A_1^{*3} \oplus \mathbb{Z}/2\mathbb{Z}$	15/4, 2, 2, 3, 3, 5, 15
28	$A_1 \oplus A_5$	$A_2^* \oplus \mathbb{Z}/2\mathbb{Z}$	2, 3, 6, 6, 6, 6, 36
29	$A_1 \oplus A_5$	$A_1^* \oplus \langle 1/6 \rangle$	2, 2, 2, 2, 2, 2, 3
30	$A_1 \oplus D_5$	$A_1^* \oplus \langle 1/4 \rangle$	2, 2, 2, 2, 2, 3, 8
33	$A_1^2 \oplus A_4$	$\frac{1}{10}\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$	2, 2, 3, 3, 3, 3, 12
34	$A_1^2 \oplus D_4$	$A_1^{*2} \oplus \mathbb{Z}/2\mathbb{Z}$	2, 3, 3, 3, 3, 6, 18
38	$A_1^3 \oplus A_3$	$A_1^* \oplus \langle 1/4 \rangle \oplus \mathbb{Z}/2\mathbb{Z}$	2, 2, 3, 3, 3, 4, 12
42	A_1^6	$A_1^{*2} \oplus (\mathbb{Z}/2\mathbb{Z})^2$	6, −1, −1, 2, 2, 3, 6
47	$A_1 \oplus A_6$	$\langle 1/14 \rangle$	8, 8, 8, 8, 8, 8, 128
48	$A_1 \oplus D_6$	$A_1^* \oplus \mathbb{Z}/2\mathbb{Z}$	1, 2, 2, 2, 2, 2, 4
49	$A_1 \oplus E_6$	$\langle 1/6 \rangle$	2, 2, 2, 2, 2, 2, 8
52	$A_1^2 \oplus D_5$	$\langle 1/4 \rangle \oplus \mathbb{Z}/2\mathbb{Z}$	2, 2, 2, 2, 2, 4, 8
53	$A_1^2 \oplus A_5$	$\langle 1/6 \rangle \oplus \mathbb{Z}/2\mathbb{Z}$	2, 2, 4, 4, 4, 4, 16
57	$A_1^3 \oplus D_4$	$A_1^* \oplus (\mathbb{Z}/2\mathbb{Z})^2$	−1, 2, 2, 2, 2, −2, −4
58	$A_1 \oplus A_3^2$	$A_1^* \oplus \mathbb{Z}/4\mathbb{Z}$	I, I, I, I, 2, 2, 2
60	$A_1^4 \oplus A_3$	$\langle 1/4 \rangle \oplus (\mathbb{Z}/2\mathbb{Z})^2$	2, 2, 2, 2, −1, −1, 4
65	$A_1 \oplus E_7$	$\mathbb{Z}/2\mathbb{Z}$	1, 1, 1, 1, 1, 1, 1
70	$A_1 \oplus A_7$	$\mathbb{Z}/4\mathbb{Z}$	I, I, I, I, I, I, I
71	$A_1^2 \oplus D_6$	$(\mathbb{Z}/2\mathbb{Z})^2$	1, 1, 1, 1, 1, 1, −1
74	$A_1^2 \oplus A_3^2$	$(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/4\mathbb{Z})$	I, I, I, I, −1, −1, −1

Table 1. Examples of specializations of the E_7 family (types are from [Oguiso and Shioda 1991]).

and $r = s_1 s_2 s_3$, and a simple solution with no extra coincidences is given in the rightmost column (note that $s_7 = 18^3/(2 \cdot 3^4 \cdot 5) = 36/5$).

Here $I = \sqrt{-1}$.

Remark 10. For the examples in lines 58, 70 and 74 of the table, one can show that it is not possible to define a rational elliptic surface over $\mathbb{Q}$ in the form we have assumed, such that all the specializations s_i, r are rational. However, there do exist examples with all sections defined over $\mathbb{Q}$, not in the chosen Weierstrass form.

The surface with Weierstrass equation

$$y^2 + xy + \tfrac{1}{16}(c^2-1)(t^2-1)y = x^3 + \tfrac{1}{16}(c^2-1)(t^2-1)x^2$$

has a 4-torsion section $(0,0)$ and a nontorsion section

$$\big((c+1)(t^2-1)/8,\ (c+1)^2(t-1)^2(t+1)/32\big)$$

of height $1/2$, as well as two reducible fibers of type I_4 and a fiber of type I_2. It is an example of type 58.

The surface with Weierstrass equation

$$y^2 + xy + t^2 y = x^3 + t^2 x^2$$

has a 4-torsion section $(0,0)$, and reducible fibers of types I_8 and I_2. It is an example of type 70.

The surface with Weierstrass equation

$$y^2 + xy - \big(t^2 - \tfrac{1}{16}\big)y = x^3 - \big(t^2 - \tfrac{1}{16}\big)x^2$$

has two reducible fibers of type I_4 and two reducible fibers of type I_2. It also has a 4-torsion section $(0,0)$ and a 2-torsion section $\big((4t-1)/8,\ (4t-1)^2/32\big)$, which generate the Mordell–Weil group. It is an example of type 74. This last example is the universal elliptic curve with $\mathbb{Z}/4\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ torsion (compare [Kubert 1976]).

4.2. ***Proofs.*** We start by considering the coefficients ϵ_i of $\Phi_\lambda(X)$; we know that $(-1)^i\epsilon_i$ is simply the i-th elementary symmetric polynomial in the 56 specializations $s(P_i)$. One shows, either by explicit calculation with Laurent polynomials, or by calculating the decomposition of $\bigwedge^i V$ (where $V = V_7$ is the 56-dimensional representation of E_7), and expressing its character as polynomials in the fundamental characters, the following formulas. Some more details are in Section 7 and the auxiliary files in [Kumar and Shioda 2013].

$$\begin{aligned}
\epsilon_1 &= -\chi_7,\\
\epsilon_2 &= \chi_6 + 1,\\
\epsilon_3 &= -(\chi_7 + \chi_5),\\
\epsilon_4 &= \chi_6 + \chi_4 + 1,
\end{aligned}$$

$$\begin{aligned}
\epsilon_5 &= -(\chi_6+\chi_3-\chi_1^2+\chi_1+1)\chi_7+(\chi_1-1)\chi_5-\chi_2\chi_3,\\
\epsilon_6 &= \chi_1\chi_7^2+(\chi_5-(\chi_1+1)\chi_2)\chi_7+\chi_6^2+2(\chi_3-\chi_1^2+\chi_1+1)\chi_6\\
&\quad -\chi_2\chi_5-(2\chi_1+1)\chi_4+\chi_3^2+2(2\chi_1+1)\chi_3\\
&\quad +\chi_1\chi_2^2-2\chi_1^3+\chi_1^2+2\chi_1+1,\\
\epsilon_7 &= (-(\chi_1+1)\chi_6+2\chi_4-2(\chi_1+1)\chi_3+\chi_1^3-3\chi_1-1)\chi_7\\
&\quad -2(\chi_5-\chi_1\chi_2)\chi_6-(\chi_3-\chi_1^2+\chi_1+2)\chi_5+3\chi_2\chi_4\\
&\quad -(\chi_1+3)\chi_2\chi_3-\chi_2^3+(2\chi_1-1)\chi_1\chi_2.
\end{aligned}$$

On the other hand, we can explicitly calculate the first few coefficients ϵ_i of $\Phi_\lambda(X)$ in terms of the Weierstrass coefficients, obtaining the following equations. Details for the method are in Section 6.

$$\begin{aligned}
\epsilon_1 &= -q_3,\\
\epsilon_2 &= p_2^2+12p_2-q_2+28,\\
\epsilon_3 &= -3(2p_2+9)q_3-q_1+2p_1(p_2+5),\\
\epsilon_4 &= 9q_3^2-6p_1q_3-2q_2-q_0+8p_2^3+86p_2^2+2(p_0+156)p_2+p_1^2+10p_0+378,\\
\epsilon_5 &= (8q_2-20p_2^2-174p_2-7p_0-351)q_3-2p_1q_2+6(p_2+4)q_1\\
&\quad +14p_1p_2^2+108p_1p_2+2(p_0+101)p_1,\\
\epsilon_6 &= 12(4p_2+15)q_3^2-(5q_1+38p_1p_2+140p_1)q_3+4q_2^2\\
&\quad +(16p_2^2+96p_2-4p_0+155)q_2+2p_1q_1+3(4p_2+17)q_0+28p_2^4+360p_2^3\\
&\quad +(4p_0+1765)p_2^2+2(4p_1^2+21p_0+1950)p_2+29p_1^2+p_0^2+88p_0+3276,\\
\epsilon_7 &= -36q_3^3+42p_1q_3^2+(4q_2-20q_0-56p_2^3-628p_2^2-14(p_0+168)p_2-16p_1^2\\
&\quad -46p_0-2925)q_3+(3q_1+6p_1p_2+20p_1)q_2+(21p_2^2+162p_2-p_0+323)q_1\\
&\quad +6p_1q_0+42p_1p_2^3+448p_1p_2^2+2(p_0+799)p_1p_2+2p_1^3+6(p_0+316)p_1.
\end{aligned}$$

Equating the two expressions we have obtained for each ϵ_i, we get a system of seven equations, the first being

$$-\chi_7 = -q_3.$$

We label these equations (1) through (7). The last few of these polynomial equations are somewhat complicated, and so to obtain a few simpler ones, we may consider the 126 sections of height 2, which we analyze as follows. Substituting

$$x = at^2 + bt + c,$$
$$y = dt^3 + et^2 + ft + g$$

into the Weierstrass equation, we get another system of equations:

$$\begin{aligned}
a^3 &= d^2 + ad,\\
3a^2 b &= (2d+a)e + bd,\\
a(p_2 + 3ac + 3b^2) &= (2d+a)f + e^2 + be + cd + 1,\\
q_3 + bp_2 + ap_1 + 6abc + b^3 &= (2d+a)g + (2e+b)f + ce,\\
q_2 + cp_2 + bp_1 + ap_0 + 3ac^2 + 3b^2 c &= (2e+b)g + f^2 + cf,\\
q_1 + cp_1 + bp_0 + 3bc^2 &= (2f+c)g,\\
q_0 + cp_0 + c^3 &= g^2.
\end{aligned}$$

The specialization of such a section at $t = \infty$ is $1 + a/d$. Setting $d = ar$, we may as before eliminate other variables to obtain an equation of degree 126 for r. Substituting $r = 1/(u-1)$, we get a monic polynomial $\Psi_\lambda(X) = 0$ of degree 126 for u. Note that the roots are given by elements of the form

$$\begin{aligned}
&\frac{s_i}{r}, \frac{r}{s_i} && \text{for } 1 \le i \le 7,\\
&\frac{s_i}{s_j} && \text{for } 1 \le i \ne j \le 7, \quad \text{and}\\
&\frac{s_i s_j s_k}{r}, \frac{r}{s_i s_j s_k} && \text{for } 1 \le i < j < k \le 7.
\end{aligned}$$

As before, we can write the first few coefficients η_i of Ψ_λ in terms of $\lambda = (p_0, \ldots, q_3)$, as well as in terms of the characters χ_j, obtaining some more relations. We will only need the first two:

$$-\chi_1 + 7 = \eta_1 = -18 - 6p_2,$$
$$-6\chi_1 + \chi_3 + 28 = \eta_2 = p_0 + 72p_2 + 13p_2^2 - q_2 + 99$$

which we call (1′) and (2′), respectively.

Now we consider the system of six equations (1) through (4), (1′) and (2′). These may be solved for $(p_2, p_0, q_3, q_2, q_1, q_0)$ in terms of the χ_j and p_1. Substituting this solution into the other three relations (5), (6) and (7), we obtain three equations for p_1, of degrees 1, 2 and 3, respectively. These have a single common factor, linear in p_1, which we then solve. This gives us the proof of Theorem 6.

The proof of Theorem 5 is now straightforward. Part (1) asserts that the image of ρ_λ is surjective on to $W(E_7)$. This follows from a standard Galois-theoretic argument: Let F be the fixed field of $W(E_7)$ acting on $k_\lambda = \mathbb{Q}(\lambda)(s_1, \dots, s_6, r) = \mathbb{Q}(s_1, \dots, s_6, r)$, where the last equality follows from the explicit expression of $\lambda = (p_0, \dots, q_3)$ in terms of the χ_i, which are in $\mathbb{Q}(s_1, \dots, s_6, r)$. Then we have that $k_0 \subset F$ since $p_0, \dots, q_3$ are polynomials in the χ_i with rational coefficients, and the χ_i are manifestly invariant under the Weyl group. Therefore $[k_\lambda : k_0] \geq [k_\lambda : F] = |W(E_7)|$, where the latter equality is from Galois theory. Finally, $[k_\lambda : k_0] \leq |\mathrm{Gal}(k_\lambda/k_0)| \leq |W(E_7)|$, since $\mathrm{Gal}(k_\lambda/k_0) \hookrightarrow W(E_7)$. Therefore, equality is forced.

Another way to see that the Galois group is the full Weyl group is to show it for a specialization, such as Example 8, and use [Serre 1989, Section 9.2, Proposition 2].

Next, let Y be the toric hypersurface given by $s_1 \dots s_7 = r^3$. Its function field is the splitting field of $\Phi_\lambda(X)$, as we remarked above. We have seen that $\mathbb{Q}(Y)^{W(E_7)} = k_0 = \mathbb{Q}(\lambda)$. Since $\Phi_\lambda(X)$ is a monic polynomial with coefficients in $\mathbb{Q}[\lambda]$, we have that $\mathbb{Q}[Y]$ is integral over $\mathbb{Q}[\lambda]$. Therefore $\mathbb{Q}[Y]^{W(E_7)}$ is also integral over $\mathbb{Q}[\lambda]$, and contained in $\mathbb{Q}(Y)^{W(E_7)} = k_0 = \mathbb{Q}(\lambda)$. Since $\mathbb{Q}[\lambda]$ is a polynomial ring, it is integrally closed in its ring of fractions. Therefore $\mathbb{Q}[Y]^{W(E_7)} \subset \mathbb{Q}[\lambda]$.

We also know $\mathbb{Q}[\chi] = \mathbb{Q}[\chi_1, \dots, \chi_7] \subset \mathbb{Q}[Y]^{W(E_7)}$, since the χ_j are invariant under the Weyl group. Therefore, we have

$$\mathbb{Q}[\chi] \subset \mathbb{Q}[Y]^{W(E_7)} \subset \mathbb{Q}[\lambda]$$

and Theorem 6, which says $\mathbb{Q}[\chi] = \mathbb{Q}[\lambda]$, implies that all these three rings are equal. This completes the proof of Theorem 5.

Remark 11. This argument gives an independent proof of the fact that the ring of multiplicative invariants for $W(E_7)$ is a polynomial ring over $\chi_1, \dots, \chi_7$. See [Bourbaki 1968, Théorème VI.3.1 and Exemple 1] or [Lorenz 2005, Theorem 3.6.1] for the classical proof that the Weyl-orbit sums of a set of fundamental weights are a set of algebraically independent generators of the multiplicative invariant ring; from there to the fundamental characters is an easy exercise.

Remark 12. Now that we have found the explicit relation between the Weierstrass coefficients and the fundamental characters, we may go back and explore the "genericity condition" for this family to have Mordell–Weil lattice E_7^*. To do this, we compute the discriminant of the cubic in x, after completing the square in y, and take the discriminant with respect to t of the resulting polynomial of degree 10. A computation shows that this discriminant factors as the cube of a polynomial $A(\lambda)$ (which vanishes exactly when the family has a fiber of additive reduction, generically type II), times a polynomial $B(\lambda)$, whose zero locus corresponds to the occurrence of a reducible multiplicative fiber. In fact, we calculate (for instance, by evaluating the split case) that $B(\lambda)$ is the product of $(e^\alpha - 1)$, where α runs over the 126 roots of E_7. We deduce by further analyzing the type II case that the condition to have Mordell–Weil lattice E_7^* is that

$$\prod (e^\alpha - 1) = \Psi_\lambda(1) \neq 0.$$

Note that this is essentially the expression that occurs in Weyl's denominator formula. In addition, the condition for having only multiplicative fibers is that $\Psi_\lambda(1)$ and $A(\lambda)$ both be nonzero.

Finally, the proof of Theorem 9 follows immediately from the discussion in [Shioda 2010a; 2010b] — compare [Shioda 2010b, Section 2.3] for the additive reduction case.

5. The E_8 case

5.1. ***Results.*** Finally, we show a multiplicative excellent family for the Weyl group of E_8. It is given by the Weierstrass equation

$$y^2 = x^3 + t^2 x^2 + (p_0 + p_1 t + p_2 t^2)\, x + (q_0 + q_1 t + q_2 t^2 + q_3 t^3 + q_4 t^4 + t^5).$$

For generic $\lambda = (p_0, \ldots, p_2, q_0, \ldots, q_4)$, this rational elliptic surface X_λ has no reducible fibers, only nodal I_1 fibers at twelve points, including $t = \infty$. We will use the specialization map at ∞. The Mordell–Weil lattice M_λ is isomorphic to the lattice E_8. Any rational elliptic surface with a multiplicative fiber of type I_1 may be put in the form above (over a small degree algebraic extension of the base field), after a fractional linear transformation of t and Weierstrass transformations of x, y.

Lemma 13. *The smooth part of the special fiber is isomorphic to the group scheme $\mathbb{G}_m$. The identity component is the nonsingular part of the curve $y^2 = x^3 + x^2$. The x- and y-coordinates of a section of height* 2 *are polynomials of degrees* 2 *and* 3 *respectively, and its specialization at $t = \infty$ may be taken as*

$$\lim_{t \to \infty} (y + tx)/(y - tx) \in k^*.$$

The proof of the lemma is similar to that in the E_7 case (and simpler!), and we omit it.

There are 240 sections of minimal height 2, with x-and y-coordinates of the form

$$x = gt^2 + at + b,$$
$$y = ht^3 + ct^2 + dt + e.$$

Under the identification with $\mathbb{G}_m$ of the special fiber of the Néron model, this section goes to $(h+g)/(h-g)$. Substituting the formulas above for x and y into the Weierstrass equation, we get the following system of equations.

$$\begin{aligned}
h^2 &= g^3 + g^2, \\
2ch &= 3ag^2 + 2ag + 1, \\
2dh + c^2 &= q_4 + gp_2 + 3bg^2 + (2b + 3a^2)g + a^2, \\
2eh + 2cd &= q_3 + ap_2 + gp_1 + 6abg + 2ab + a^3, \\
2ce + d^2 &= q_2 + bp_2 + ap_1 + gp_0 + 3b^2g + b^2 + 3a^2b, \\
2de &= q_1 + bp_1 + ap_0 + 3ab^2, \\
e^2 &= q_0 + bp_0 + b^3.
\end{aligned}$$

Setting $h = gu$, we eliminate other variables to obtain an equation of degree 240 for u. Finally, substituting in $u = (v+1)/(v-1)$, we get a monic reciprocal equation $\Phi_\lambda(X) = 0$ satisfied by v, with coefficients in $\mathbb{Z}[\lambda] = \mathbb{Z}[p_0, \dots, p_2, q_0, \dots, q_4]$. We have

$$\Phi_\lambda(X) = \prod_{i=1}^{240}(X - s(P)) = X^{240} + \epsilon_1 X^{239} + \cdots + \epsilon_1 X + \epsilon_0,$$

where P ranges over the 240 minimal sections of height 2. It is clear that $a, \dots, h$ are rational functions of v, with coefficients in k_0.

We have a Galois representation on the Mordell–Weil lattice

$$\rho_\lambda : \mathrm{Gal}(k/k_0) \to \mathrm{Aut}(M_\lambda) \cong \mathrm{Aut}(E_8).$$

Here $\mathrm{Aut}(E_8) \cong W(E_8)$, the Weyl group of type E_8. The *splitting field* of M_λ is the fixed field k_λ of $\mathrm{Ker}(\rho_\lambda)$. By definition, $\mathrm{Gal}(k_\lambda/k_0) \cong \mathrm{Im}(\rho_\lambda)$. The splitting field k_λ is equal to the splitting field of the polynomial $\Phi_\lambda(X)$ over k_0, since the Mordell–Weil group is generated by the 240 sections of smallest height $P_i = (g_i t^2 + a_i t + b_i, h_i t^3 + c_i t^2 + d_i t + e_i)$. We also have

$$k_\lambda = k_0(P_1, \dots, P_{240}) = k_0(v_1, \dots, v_{240}).$$

Theorem 14. *Assume that λ is generic over $\mathbb{Q}$, that is, the coordinates $p_0, \dots, q_4$ are algebraically independent over $\mathbb{Q}$.*

(1) *ρ_λ induces an isomorphism* $\mathrm{Gal}(k_\lambda/k_0) \cong W(E_8)$.

(2) *The splitting field k_λ is a purely transcendental extension of $\mathbb{Q}$, and is isomorphic to the function field $\mathbb{Q}(Y)$ of the toric hypersurface $Y \subset \mathbb{G}_m^9$ defined by $s_1 \cdots s_8 = r^3$. There is an action of $W(E_8)$ on Y such that $\mathbb{Q}(Y)^{W(E_8)} = k_\lambda^{W(E_8)} = k_0$.*

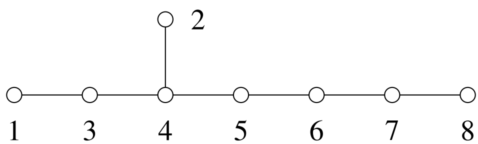
(3) *The ring of $W(E_8)$-invariants in the affine coordinate ring*

$$\mathbb{Q}[Y] = \mathbb{Q}[s_i, r, 1/s_i, 1/r]/(s_1 \ldots s_8 - r^3) \cong \mathbb{Q}[s_1, \ldots, s_7, r, s_1^{-1}, \ldots, s_7^{-1}, r^{-1}]$$

is equal to the polynomial ring $\mathbb{Q}[\lambda]$:

$$\mathbb{Q}[Y]^{W(E_8)} = \mathbb{Q}[\lambda] = \mathbb{Q}[p_0, p_1, p_2, q_0, q_1, q_2, q_3, q_4].$$

As in the E_7 case, we prove an explicit, invertible polynomial relation between the Weierstrass coefficients λ and the fundamental characters for E_8. Let $V_1, \ldots, V_8$ be the fundamental representations of E_8, and $\chi_1, \ldots, \chi_8$ their characters as labeled below.



Again, for the set of generators of E_8, we choose (as in [Shioda 1995]) vectors $v_1, \ldots, v_8, u$ with $\sum v_i = 3u$ and let s_i correspond to v_i and r to u, so that $\prod s_i = r^3$. The 240 roots of $\Phi_\lambda(X)$ are given by

$$s_i, \frac{1}{s_i} \quad \text{for } 1 \le i \le 8, \qquad \frac{s_i}{s_j} \quad \text{for } 1 \le i \ne j \le 8,$$
$$\frac{s_i s_j}{r}, \frac{r}{s_i s_j} \quad \text{for } 1 \le i < j \le 8, \quad \text{and} \quad \frac{s_i s_j s_k}{r}, \frac{r}{s_i s_j s_k} \quad \text{for } 1 \le i < j < k \le 8.$$

The characters $\chi_1, \ldots, \chi_7$ lie in the ring of Laurent polynomials $\mathbb{Q}[s_i, r, 1/s_i, 1/r]$, and are invariant under the multiplicative action of the Weyl group on this ring of Laurent polynomials. The χ_i may be explicitly computed using the software `LiE`, as indicated in Section 7 and the auxiliary files.

Theorem 15. *For generic λ over $\mathbb{Q}$, we have*

$$\mathbb{Q}[\chi_1, \ldots, \chi_8] = \mathbb{Q}[p_0, p_1, p_2, q_0, q_1, q_2, q_3, q_4].$$

The transformation between these sets of generators is

$$\chi_1 = -1600 q_4 + 1536 p_2 + 3875,$$
$$\chi_2 = 2(-45600 q_4 + 12288 q_3 + 40704 p_2 - 16384 p_1 + 73625),$$
$$\chi_3 = 64(14144 q_4^2 - 72(384 p_2 + 1225) q_4 + 11200 q_3 - 4096 q_2 + 13312 p_2^2$$
$$+ 87072 p_2 - 17920 p_1 + 16384 p_0 + 104625),$$

$$
\begin{aligned}
\chi_4 &= -91750400q_4^3 + 12288(25600p_2 + 222101)q_4^2 - 256(4530176q_3 - 65536q_2 \\
&\quad + 1392640p_2^2 + 21778944p_2 - 8159232p_1 + 2621440p_0 + 34773585)q_4 \\
&\quad + 32(4718592q_3^2 + 384(80896p_2 - 32768p_1 + 225379)q_3 - 29589504q_2 \\
&\quad + 30408704q_1 - 33554432q_0 + 4194304p_2^3 + 88129536p_2^2 \\
&\quad - 64(876544p_1 - 262144p_0 - 4399923)p_2 + 8388608p_1^2 - 133996544p_1 \\
&\quad + 65175552p_0 + 215596227), \\
\chi_5 &= 24760320q_4^2 - 64(106496q_3 + 738816p_2 - 163840p_1 + 2360085)q_4 \\
&\quad + 12288(512p_2 + 4797)q_3 - 30670848q_2 + 16777216q_1 + 20250624p_2^2 \\
&\quad - 512(16384p_1 - 235911)p_2 - 45154304p_1 + 13631488p_0 + 146325270, \\
\chi_6 &= 110592q_4^2 - 1536(128p_2 + 1235)q_4 + 724992q_3 - 262144q_2 + 65536p_2^2 \\
&\quad + 1062912p_2 - 229376p_1 + 2450240, \\
\chi_7 &= -4(3920q_4 - 1024q_3 - 1152p_2 - 7595), \\
\chi_8 &= -8(8q_4 - 31).
\end{aligned}
$$

Remark 16. We omit the inverse for conciseness here; it is easily computed in a computer algebra system and is available in the auxiliary files.

Remark 17. As before, our explicit formulas are compatible with those in [Eguchi and Sakai 2003]. Also, the proof of Theorem 14 gives another proof that the multiplicative invariants for $W(E_8)$ are freely generated by the fundamental characters (or by the orbit sums of the fundamental weights).

Example 18. Let $\mu = (2\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13\cdot 17\cdot 19) = 9699690$. Then

$$
\begin{aligned}
q_4 &= -224337445655593668343339/(2^5\cdot\mu^2), \\
q_3 &= 43080034312940338834622651824607856\,7/(2^{11}\cdot\mu^3), \\
q_2 &= 7255510194764901112739173303498415846257314640990576 9/(2^{22}\cdot 3^2\cdot\mu^4), \\
q_1 &= (-1288109930551729133820743237846836849158406377255698116491924 5 \\
&\qquad 30489)/(2^{29}\cdot 3\cdot\mu^5), \\
q_0 &= (88271767933236199294273033814854594019111918837196838709750423283 \\
&\qquad 443360357992650203)/(2^{42}\cdot 3^3\cdot\mu^6), \\
p_2 &= 146156773903879871001810589/(2^9\cdot 3\cdot\mu^2), \\
p_1 &= -24909805041567866985469379779685360019313/(2^{20}\cdot\mu^3), \\
p_0 &= 149210717611026376686431912157550398014717711388673 87/(2^{23}\cdot 3\cdot\mu^4).
\end{aligned}
$$

These values give an elliptic surface for which we have $r = 2$, $s_1 = 3$, $s_2 = 5$, $s_3 = 7$, $s_4 = 11$, $s_5 = 13$, $s_6 = 17$, $s_7 = 19$, the simplest choice of multiplicatively independent elements. Here, the specializations of a basis are given by $v \in \{3, 5, 7, 11, 13, 17, 19, 15/2\}$. Once again, we list the x-coordinates of the corresponding sections, and leave the rest of the verification to the auxiliary files.

$$\begin{aligned}
x(P_1) &= 3t^2 - \frac{99950606190359}{620780160}t + \frac{432532755764748812020964981 3}{2642523476911718400},\\
x(P_2) &= \frac{5}{4}t^2 - \frac{153332163637781}{1655413760}t + \frac{5414114237697608646836821}{5138596941004800},\\
x(P_3) &= \frac{7}{9}t^2 - \frac{203120672689603}{2793510720}t + \frac{6943164348569130636788638639}{7927570430735155200},\\
x(P_4) &= \frac{11}{25}t^2 - \frac{8564057914757}{147804800}t + \frac{115126372233675800396600989}{155442557465395200},\\
x(P_5) &= \frac{13}{36}t^2 - \frac{347479008951469}{6385167360}t + \frac{15713360768094961737403040541 7}{22197197206058434560 0},\\
x(P_6) &= \frac{17}{64}t^2 - \frac{1327421017414859}{26486620160}t + \frac{5942419292933021418457517303}{8901131711702630400},\\
x(P_7) &= \frac{19}{81}t^2 - \frac{489830985359431}{10056638592}t + \frac{4668513720174369644147745495 1}{7134813387661639680 0},\\
x(P_8) &= \frac{120}{169}t^2 - \frac{30706596009257}{440806080}t + \frac{7616444307482874366216546640 9}{5582330844976005120 0}.
\end{aligned}$$

Example 19. The value $\lambda = \lambda_0 := (1, 1, 1, 1, 1, 1, 1, 1)$ gives rise to an explicit polynomial $g(X) = \Phi_{\lambda_0}(X)$, for which we can show that the Galois group is $W(E_8)$, as follows. The reduction of $g(X)$ modulo 79 shows that Frob_{79} has cycle decomposition of type $(4)^2(8)^{29}$, and similarly, Frob_{179} has cycle decomposition of type $(15)^{16}$. We deduce, as in [Jouve et al. 2008, Section 3] or [Shioda 2009], that the Galois group is the entire Weyl group. Since the coefficients of $g(X)$ are large, we do not display it here, but it is included in the auxiliary files.

As in the case of E_7, we can also describe degenerations of this family of rational elliptic surfaces X_λ by the method of "vanishing roots", where we drop the genericity assumption, and consider the situation where the elliptic fibration might have additional reducible fibers. Let $\psi : Y \to \mathbb{A}^8$ be the finite surjective morphism associated to

$$\mathbb{Q}[p_0, \dots, q_4] \hookrightarrow \mathbb{Q}[Y] \cong \mathbb{Q}[s_1, \dots, s_7, r, s_1^{-1}, \dots, s_7^{-1}, r^{-1}].$$

For $\xi = (s_1, \dots, s_8, r) \in Y$, let the multiset Π_ξ consist of the 240 elements s_i and $1/s_i$ for $1 \le i \le 8$, s_i/s_j for $1 \le i \ne j \le 8$, s_is_j/r and $r/(s_is_j)$ for $1 \le i < j \le 8$, and $s_is_js_k/r$ and $r/(s_is_js_k)$ for $1 \le i < j < k \le 8$, corresponding to the 240 roots of E_8. Let $2\nu(\xi)$ be the number of times 1 appears in Π_ξ, which is also the

Type	Fibral lattice	MW group	$\{s_1, \dots, s_6, r\}$
1	0	E_8	3, 5, 7, 11, 13, 17, 19, 2
5	A_3	D_5^*	2, 2, 2, 2, 5, 7, 11, 3
8	A_4	A_4^*	2, 2, 2, 2, 2, 5, 7, 3
15	A_5	$A_2^* \oplus A_1^*$	2, 2, 2, 2, 2, 2, 5, 3
16	D_5	A_3^*	2, 3, 3, 3, 3, 3, 5, 18
25	A_6	$\frac{1}{7}\begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix}$	2, 2, 2, 2, 2, 2, 2, 3
26	D_6	A_1^{*2}	2, 3, 3, 3, 3, 3, 3, 18
35	A_3^2	$A_1^{*2} \oplus \mathbb{Z}/2\mathbb{Z}$	2, −1/2, 3, 3, 3, 1, 1, −3
36	A_3^2	$\langle 1/4 \rangle$	8, 8, 8, 8, 27, 27, 27, 1296
43	E_7	A_1^*	2, 2, 2, 2, 2, 2, 2, 8
44	A_7	$A_1^* \oplus \mathbb{Z}/2\mathbb{Z}$	2, 2, 2, 2, 2, 2, 2, −8
45	A_7	$\langle 1/8 \rangle$	8, 8, 8, 8, 8, 8, 8, 256
46	D_7	$\langle 1/4 \rangle$	2, 4, 4, 4, 4, 4, 4, 32
54	$A_3 \oplus D_4$	$\langle 1/4 \rangle \oplus \mathbb{Z}/2\mathbb{Z}$	2, −1, −1, −1, −1, 1, 1, 2
55	$A_3 \oplus A_4$	$\langle 1/20 \rangle$	16, 16, 16, 16, 32, 32, 32, 4096
62	E_8	0	1, 1, 1, 1, 1, 1, 1, 1
63	A_8	$\mathbb{Z}/3\mathbb{Z}$	1, 1, 1, 1, 1, 1, 1, ζ_3
64	D_8	$\mathbb{Z}/2\mathbb{Z}$	1, 1, 1, 1, 1, 1, 1, −1
67	A_4^2	$\mathbb{Z}/5\mathbb{Z}$	1, 1, 1, 1, ζ_5, ζ_5, ζ_5, ζ_5^3
72	$A_3 \oplus D_5$	$\mathbb{Z}/4\mathbb{Z}$	1, 1, 1, I, I, I, I, $-I$

Table 2. Examples of specializations of the E_8 family (types are from [Oguiso and Shioda 1991]).

multiplicity of 1 as a root of $\Phi_\lambda(X)$, with $\lambda = \psi(\xi)$. We call the associated roots of E_8 the *vanishing roots*, in analogy with vanishing cycles in the deformation of singularities. By abuse of notation we label the rational elliptic surface X_λ as X_ξ.

Theorem 20. *The surface X_ξ has new reducible fibers* (*necessarily at $t \neq \infty$*) *if and only if $\nu(\xi) > 0$. The number of roots in the root lattice T_{new} is equal to $2\nu(\xi)$, where $T_{\text{new}} := \bigoplus_{v \neq \infty} T_v$ is the new part of the trivial lattice.*

We may use this result to produce specializations with trivial lattice corresponding to most of the entries of [Oguiso and Shioda 1991], and a nodal fiber. We list below those types which are not already covered by [Shioda 1991a; 2012] or our examples for the E_7 case, which have an I_2 fiber.

Here ζ_3, I and ζ_5 are primitive third, fourth and fifth roots of unity.

Remark 21. As before, for the examples in lines 63, 67 and 72 of the table, one can show it is not possible to define a rational elliptic surface over $\mathbb{Q}$ in the form we have assumed, such that all the specializations s_i and r are rational. However, there do exist examples with all sections defined over $\mathbb{Q}$, not in the chosen Weierstrass form.

The surface with Weierstrass equation

$$y^2 + xy + t^3 y = x^3$$

has a 3-torsion point (0, 0) and a fiber of type I_9. It is an example of type 63.

The surface with Weierstrass equation

$$y^2 + (t+1)xy + ty = x^3 + tx^2$$

has a 5-torsion section (0, 0) and two fibers of type I_5. It is an example of type 67.

The surface with Weierstrass equation

$$y^2 + txy + \frac{t^2(t-1)}{16} y = x^3 + \frac{t(t-1)}{16} x^2$$

has a 4-torsion section (0, 0), and two fibers of types I_4 and I_1^*. It is an example of type 72.

Remark 22. Our tables and the one in [Shioda 2012] cover all the cases of [Oguiso and Shioda 1991], except lines 9, 27 and 73 of the table, with trivial lattice D_4, E_6 and D_4^2, respectively. Since these have fibers with additive reduction, examples for them may be directly constructed using the families in [Shioda 1991a]. For instance, the elliptic surface

$$y^2 = x^3 - xt^2$$

has two fibers of type I_0^* and Mordell–Weil group $(\mathbb{Z}/2\mathbb{Z})^2$. This covers line 73 of the table. For the other two cases, we refer the reader to the original examples of additive reduction in [Shioda 1991a, Section 3].

5.2. ***Proofs.*** The proof proceeds analogously to the E_7 case, with two differences: We only have one polynomial $\Phi_\lambda(X)$ to work with (as opposed to having $\Phi_\lambda(X)$ and $\Psi_\lambda(X)$), and the equations are a lot more complicated.

We first write down the relation between the coefficients ϵ_i for $1 \le i \le 9$, and the fundamental invariants χ_j; as before, we postpone the proofs to the auxiliary files and outline the idea in Section 7. Second, we write down the coefficients ϵ_i in terms of $\lambda = (p_0, \dots, p_2, q_0, \dots, q_4)$; see Section 6 for an explanation of how this is carried out. In the interest of brevity, we do not write out either of these sets of equations, but relegate them to the auxiliary computer files. Finally, setting the corresponding expressions equal to each other, we obtain a set of equations (1) through (9).

To solve these equations, proceed as follows: first use (1) through (5) to solve for $q_0, \dots, q_4$ in terms of χ_j and p_0, p_1, p_2. Substituting these in to the remaining equations, we obtain (6′) through (9′). These have low degree in p_0, which we eliminate, obtaining equations of relatively small degrees in p_1 and p_2. Finally, we take resultants with respect to p_1, obtaining two equations for p_2, of which the only common root is the one listed above. Working back, we solve for all the other variables, obtaining the system above and completing the proof of Theorem 15. The deduction of Theorem 14 now proceeds exactly as in the case of E_7.

Remark 23. As in the E_7 case, once we find the explicit relation between the Weierstrass coefficients and the fundamental characters, we may go back and explore the "genericity condition" for this family to have Mordell–Weil lattice isomorphic to E_8. To do this we compute the discriminant of the cubic in x, after completing the square in y, and take the discriminant with respect to t of the resulting polynomial of degree 11. A computation shows that this discriminant factors as the cube of a polynomial $A(\lambda)$ (which vanishes exactly when the family has a fiber of additive reduction, generically type II), and the product of $(e^\alpha - 1)$, where α runs over minimal vectors of E_8. Again, the genericity condition to have Mordell–Weil lattice exactly E_8 is just the nonvanishing of

$$\Phi_\lambda(1) = \prod (e^\alpha - 1),$$

the expression which occurs in the Weyl denominator formula. Furthermore, the condition to have only multiplicative fibers is that $\Phi_\lambda(1) A(\lambda) \neq 0$.

As before, the proof of Theorem 20 follows immediately from the results of [Shioda 2010a; 2010b], by degeneration from a flat family.

6. Resultants, interpolation and computations

We now explain a computational aid, used in obtaining the equations expressing the coefficients of Φ_λ (for E_8) or Ψ_λ (for E_7) in terms of the Weierstrass coefficients of the associated family of rational elliptic surfaces. We illustrate this using the system of equations obtained for sections of the E_8 family:

$$\begin{aligned}
h^2 &= g^3 + g^2, \\
2ch &= 3ag^2 + 2ag + 1, \\
c^2 + 2dh &= q_4 + gp_2 + 3bg^2 + (2b + 3a^2)g + a^2, \\
2eh + 2cd &= q_3 + ap_2 + gp_1 + 6abg + 2ab + a^3, \\
2ce + d^2 &= q_2 + bp_2 + ap_1 + gp_0 + 3b^2 g + b^2 + 3a^2 b, \\
2de &= q_1 + bp_1 + ap_0 + 3ab^2, \\
e^2 &= q_0 + bp_0 + b^3.
\end{aligned}$$

Setting $h = gu$ and solving the first equation for g, we have $g = u^2 - 1$. We solve the next three equations for c, d, e, respectively. This leaves us with three equations $R_1(a,b,u) = R_2(a,b,u) = R_3(a,b,u) = 0$. These have degrees 2, 2, 3 respectively in b. Taking the appropriate linear combination of R_1 and R_2 gives us an equation $S_1(a,b,u) = 0$ which is linear in b. Similarly, we may use R_1 and R_3 to obtain another equation $S_2(a,b,u) = 0$, linear in b. We write

$$\begin{aligned} S_1(a,b,u) &= s_{11}(a,u)b + s_{10}(a,u), \\ S_2(a,b,u) &= s_{21}(a,u)b + s_{20}(a,u), \\ R_1(a,b,u) &= r_2(a,u)b^2 + r_1(a,u)b + r_0(a,u). \end{aligned}$$

The resultant of the first two polynomials gives us an equation

$$T_1(a,u) = s_{11}s_{20} - s_{10}s_{21} = 0,$$

while the resultant of the first and third gives us

$$T_2(a,u) = r_2 s_{10}^2 - r_1 s_{10} s_{11} + r_0 s_{11}^2 = 0.$$

Finally, we substitute $u = (v+1)/(v-1)$ throughout, obtaining two equations $\tilde{T}_1(a,v) = 0$ and $\tilde{T}_2(a,v) = 0$.

Next, we would like to compute the resultant of $\tilde{T}_1(a,v)$ and $\tilde{T}_2(a,v)$, which have degrees 8 and 9 with respect to a, to obtain a single equation satisfied by v. However, the polynomials $\tilde{T}_1$ and $\tilde{T}_2$ are already fairly large (they take a few hundred kilobytes of memory), and since their degree in a is not too small, it is beyond the current reach of computer algebra systems such as `gp/PARI` or `Magma` to compute their resultant. It would take too long to compute their resultant, and another issue is that the resultant would take too much memory to store, certainly more than is available on the authors' computer systems (it would take more than 16GB of memory).

To circumvent this issue, what we shall do is to use several specializations of λ in $\mathbb{Q}^8$. Once we specialize, the polynomials take much less space to store, and the computations of the resultants becomes tremendously easier. Since the resultant can be computed via the Sylvester determinant

$$\begin{vmatrix} a_8 & \dots & a_2 & a_1 & a_0 & 0 & 0 & \dots & 0 \\ 0 & a_8 & \dots & a_2 & a_1 & a_0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_8 & \dots & a_2 & a_1 & a_0 & 0 \\ 0 & \dots & 0 & 0 & a_8 & \dots & a_2 & a_1 & a_0 \\ b_9 & b_8 & \dots & b_2 & b_1 & b_0 & 0 & \dots & 0 \\ 0 & b_9 & b_8 & \dots & b_2 & b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & b_9 & b_8 & \dots & b_2 & b_1 & b_0 \end{vmatrix},$$

where $\tilde{T}_1(a, v) = \sum a_i(v)a^i$ and $\tilde{T}_2(a, v) = \sum b_i(v)a^i$, we see that the resultant is a polynomial $Z(v) = \sum z_i v^i$ with coefficients z_i being polynomials in the coefficients of the a_i and the b_j, which happen to be elements of $\mathbb{Q}[\lambda]$ (recall that $\lambda = (p_0, \dots, p_2, q_0, \dots, q_4)$). Furthermore, we can bound the degrees $m_i(j)$ of $z_i(v)$ with respect to the j-th coordinate of λ, by using explicit bounds on the multidegrees of the a_i and b_i. Therefore, by using Lagrange interpolation (with respect to the eight variables λ_j) we can reconstruct $z_i(v)$ from its specializations for various values of λ. The same method lets us show that $Z(v)$ is divisible by v^{22} (for instance, by showing that z_0 through z_{21} are zero), and also by $(v+1)^{80}$ (by first shifting v by 1 and then computing the Sylvester determinant, and proceeding as before), as well as by $(v^2+v+1)^8$ (this time, using cube roots of unity). Finally, it is clear that $Z(v)$ is divisible by the square of the resultant $G(v)$ of s_{11} and s_{10} with respect to a. Removing these extraneous factors, we get a polynomial $\Phi_\lambda(v)$ that is monic and reciprocal of degree 240. We compute its top few coefficients by this interpolation method.

Finally, the interpolation method above is in fact completely rigorous. Namely, let $\epsilon_i(\lambda)$ be the coefficient of v^i in $\Phi_\lambda(v)$, with bounds $(m_1, \dots, m_8)$ for its multidegree, and $\epsilon'_i(\lambda)$ the putative polynomial we have computed using Lagrange interpolation on a set $L_1 \times \cdots \times L_8$, where $L_i = \{\ell_{i,0}, \dots, \ell_{i,m_i}\}$ for $1 \le i \le 8$ are sets of integers chosen generically enough to ensure that $G(v)$ has the correct degree and that $Z(v)$ is not divisible by any higher powers of v, $v+1$ or v^2+v+1 than in the generic case. Then since $\epsilon_j(\ell_{1,i_1}, \dots, \ell_{8,i_8}) = \epsilon'_j(\ell_{1,i_1}, \dots, \ell_{8,i_8})$ for all choices of $i_1, \dots, i_8$, we see that the difference of these polynomials must vanish.

7. Representation theory, and some identities in Laurent polynomials

Finally, we demonstrate how to deduce the identities relating the coefficients of $\Phi_{E_7,\lambda}(X)$ or $\Psi_{E_7,\lambda}(X)$ to the fundamental characters for E_7 (and similarly, the coefficients of $\Phi_{E_8,\lambda(X)}$ to the fundamental characters of E_8).

Conceptually, the simplest way to do this is to express the alternating powers of the 56-dimensional representation V_7 or the 133-dimensional representation V_1 in terms of the fundamental modules of E_7 and their tensor products. We know that the character χ_1 of V_1 is $7 + \sum e^\alpha$, where the sum is over the 126 roots of E_7. Therefore we have $(-1)\eta_1 = \chi_1 - 7$. For the next example, we consider $\bigwedge^2 V_1 = V_3 \oplus V_1$. This gives rise to the equation

$$\eta_2 + 7 \cdot (-1)\eta_1 + \binom{7}{2} = \chi_3 + \chi_1,$$

which gives the relation $\eta_2 = \chi_3 - 6\chi_1 + 28$.

A similar analysis can be carried out to obtain all the other identities used in our proofs, using the software `LiE` [LiE 2000].

A more explicit method is to compute the expressions for the χ_i as Laurent polynomials in $s_1, \ldots, s_6, r$ (note that $s_7 = r^3/(s_1 \ldots s_6)$), and then do the same for the ϵ_i or η_i. The latter calculation is simplified by computing the power sums $\sum(e^\alpha)^i$ (for α running over the smallest vectors of E_7^* or E_7), for $1 \le i \le 7$ and then using Newton's formulas to convert to the elementary symmetric polynomials, which are $(-1)^i \epsilon_i$ or $(-1)^i \eta_i$. Finally, we check the identities by direct computation in the Laurent polynomial ring (it may be helpful to clear out denominators). This method has the advantage that we obtain explicit expressions for the χ_i (and then for λ by Theorem 6) in terms of $s_1, \ldots, s_6, r$, which may then be used to generate examples such as Example 7.

8. Acknowledgements

We thank David Vogan for very helpful comments regarding computations in representation theory. The computer algebra systems `gp/PARI`, `Magma`, `Maxima` and the software `LiE` were used in the calculations for this paper. We thank Matthias Schütt and the referees for a careful reading of the paper and for many helpful comments.

The auxiliary computer files for checking our calculations are available from the source files in the preprint version of this paper, [Kumar and Shioda 2013].

References

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie, Chapitres IV, V, et VI*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. MR 39 #1590 Zbl 0483.22001

[Carter 2005] R. W. Carter, *Lie algebras of finite and affine type*, Cambridge Studies in Advanced Mathematics **96**, Cambridge University Press, Cambridge, 2005. MR 2006i:17001 Zbl 1110.17001

[Eguchi and Sakai 2003] T. Eguchi and K. Sakai, "Seiberg–Witten curve for E-string theory revisited", *Adv. Theor. Math. Phys.* **7**:3 (2003), 419–455. MR 2005g:81234 Zbl 1065.81105

[Jouve et al. 2008] F. Jouve, E. Kowalski, and D. Zywina, "An explicit integral polynomial whose splitting field has Galois group $W(E_8)$", *J. Théor. Nombres Bordeaux* **20**:3 (2008), 761–782. MR 2010e:12005 Zbl 1200.12003

[Kodaira 1963a] K. Kodaira, "On compact analytic surfaces, II", *Ann. of Math.* (2) **77**:3 (1963), 563–626. MR 32 #1730 Zbl 0118.15802

[Kodaira 1963b] K. Kodaira, "On compact analytic surfaces, III", *Ann. of Math.* (2) **78**:1 (1963), 1–40. MR 32 #1730 Zbl 0171.19601

[Kubert 1976] D. S. Kubert, "Universal bounds on the torsion of elliptic curves", *Proc. London Math. Soc.* (3) **33**:2 (1976), 193–237. MR 55 #7910 Zbl 0331.14010

[Kumar and Shioda 2013] A. Kumar and T. Shioda, "Multiplicative excellent families of elliptic surfaces of type E_7 or E_8", preprint, version 2, 2013. arXiv 1204.1531v2

[LiE 2000] M. A. A. van Leeuwen, A. M. Cohen, and B. Lisser, *LiE: A computer algebra package for Lie group computations, version 2.2.2*, 2000, available at http://www-math.univ-poitiers.fr/~maavl/LiE/.

[Lorenz 2005] M. Lorenz, *Multiplicative invariant theory*, Encyclopaedia of Mathematical Sciences **135**, Springer, Berlin, 2005. MR 2005m:13012 Zbl 1078.13003

[Oguiso and Shioda 1991] K. Oguiso and T. Shioda, "The Mordell–Weil lattice of a rational elliptic surface", *Comment. Math. Univ. St. Paul.* **40**:1 (1991), 83–99. MR 92g:14036 Zbl 0757.14011

[Serre 1989] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, edited by M. Brown, Aspects of Mathematics **E15**, Vieweg, Braunschweig, 1989. MR 90e:11086 Zbl 0676.14005

[Shioda 1990] T. Shioda, "On the Mordell–Weil lattices", *Comment. Math. Univ. St. Paul.* **39**:2 (1990), 211–240. MR 91m:14056 Zbl 0725.14017

[Shioda 1991a] T. Shioda, "Construction of elliptic curves with high rank via the invariants of the Weyl groups", *J. Math. Soc. Japan* **43**:4 (1991), 673–719. MR 92i:11059 Zbl 0751.14018

[Shioda 1991b] T. Shioda, "Theory of Mordell–Weil lattices", pp. 473–489 in *Proceedings of the International Congress of Mathematicians* (Kyoto, 1990), vol. 1, edited by I. Satake, Math. Soc. Japan, Tokyo, 1991. MR 93k:14046 Zbl 0746.14009

[Shioda 1995] T. Shioda, "A uniform construction of the root lattices E_6, E_7, E_8 and their dual lattices", *Proc. Japan Acad. Ser. A Math. Sci.* **71**:7 (1995), 140–143. MR 97e:17015 Zbl 0854.17008

[Shioda 2009] T. Shioda, "Some explicit integral polynomials with Galois group $W(E_8)$", *Proc. Japan Acad. Ser. A Math. Sci.* **85**:8 (2009), 118–121. MR 2011d:12003 Zbl 05651158

[Shioda 2010a] T. Shioda, "Gröbner basis, Mordell–Weil lattices and deformation of singularities, I", *Proc. Japan Acad. Ser. A Math. Sci.* **86**:2 (2010), 21–26. MR 2011e:14068 Zbl 1183.14050

[Shioda 2010b] T. Shioda, "Gröbner basis, Mordell–Weil lattices and deformation of singularities, II", *Proc. Japan Acad. Ser. A Math. Sci.* **86**:2 (2010), 27–32. MR 2011e:14069 Zbl 1186.14039

[Shioda 2012] T. Shioda, "Multiplicative excellent family of type E_6", *Proc. Japan Acad. Ser. A Math. Sci.* **88**:3 (2012), 46–51. MR 2908623 Zbl 06051465

[Shioda and Usui 1992] T. Shioda and H. Usui, "Fundamental invariants of Weyl groups and excellent families of elliptic curves", *Comment. Math. Univ. St. Paul.* **41**:2 (1992), 169–217. MR 93m:11047 Zbl 0815.14027

[Tate 1975] J. Tate, "Algorithm for determining the type of a singular fiber in an elliptic pencil", pp. 33–52 in *Modular functions of one variable, IV: Proceedings of the International Summer School* (Antwerp, 1972), edited by B. J. Birch and W. Kuyk, Lecture Notes in Math. **476**, Springer, Berlin, 1975. MR 52 #13850 Zbl 1214.14020

Communicated by János Kollár
Received 2012-04-08 Revised 2012-12-13 Accepted 2013-01-10

abhinav@math.mit.edu *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States*

shioda@rikkyo.ac.jp *Department of Mathematics, Rikkyo University, 3-34-1 Nishi-Ikebukuro Toshima-ku, Tokyo 171-8501, Japan*
http://www.rkmath.rikkyo.ac.jp/math/shioda/

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1643

Cohomological invariants of algebraic tori

Sam Blinstein and Alexander Merkurjev

Let G be an algebraic group over a field F. As defined by Serre, a cohomological invariant of G of degree n with values in $\mathbb{Q}/\mathbb{Z}(j)$ is a functorial-in-K collection of maps of sets $\mathrm{Tors}_G(K) \to H^n(K, \mathbb{Q}/\mathbb{Z}(j))$ for all field extensions K/F, where $\mathrm{Tors}_G(K)$ is the set of isomorphism classes of G-torsors over Spec K. We study the group of degree 3 invariants of an algebraic torus with values in $\mathbb{Q}/\mathbb{Z}(2)$. In particular, we compute the group $H^3_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(2))$ of unramified cohomology of an algebraic torus S.

1. Introduction

Let G be a linear algebraic group over a field F (of arbitrary characteristic). The notion of an *invariant* of G was defined in [Garibaldi et al. 2003] as follows. Consider the category Fields_F of field extensions of F and the functor

$$\mathrm{Tors}_G : \mathit{Fields}_F \to \mathit{Sets}$$

taking a field K to the set $\mathrm{Tors}_G(K)$ of isomorphism classes of (right) G-torsors over Spec K. Let

$$H : \mathit{Fields}_F \to \mathit{Abelian\ Groups}$$

be another functor. An H*-invariant* of G is then a morphism of functors

$$i : \mathrm{Tors}_G \to H,$$

viewing H with values in *Sets*, that is, a functorial in K collection of maps of sets $\mathrm{Tors}_G(K) \to H(K)$ for all field extensions K/F. We denote the group of H-invariants of G by $\mathrm{Inv}(G, H)$.

An invariant $i \in \mathrm{Inv}(G, H)$ is called *normalized* if $i(I) = 0$ for the trivial G-torsor I. The normalized invariants form a subgroup $\mathrm{Inv}(G, H)_{\mathrm{norm}}$ of $\mathrm{Inv}(G, H)$ and there is a natural isomorphism

The work of the second author has been supported by the NSF grant DMS #1160206.
MSC2010: primary 11E72; secondary 12G05.
Keywords: algebraic tori, cohomological invariants, Galois cohomology.

$$\mathrm{Inv}(G, H) \simeq H(F) \oplus \mathrm{Inv}(G, H)_{\mathrm{norm}},$$

so it is sufficient to study normalized invariants.

Typically, H is a cohomological functor given by Galois cohomology groups with values in a fixed Galois module. Of particular interest to us is the functor H which takes a field K/F to the Galois cohomology group $H^n(K, \mathbb{Q}/\mathbb{Z}(j))$, where the coefficients $\mathbb{Q}/\mathbb{Z}(j)$ are defined as follows. For a prime integer p different from the characteristic of F, the p-component $\mathbb{Q}_p/\mathbb{Z}_p(j)$ of $\mathbb{Q}/\mathbb{Z}(j)$ is the colimit over n of the étale sheaves $\mu_{p^n}^{\otimes j}$, where μ_m is the sheaf of m-th roots of unity. In the case $p = \mathrm{char}(F) > 0$, $\mathbb{Q}_p/\mathbb{Z}_p(j)$ is defined via logarithmic de Rham–Witt differentials; see Section 3b.

We write $\mathrm{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j))$ for the group of *cohomological invariants of G of degree n with values in $\mathbb{Q}/\mathbb{Z}(j)$*.

The second cohomology group $H^2(K, \mathbb{Q}/\mathbb{Z}(1))$ is canonically isomorphic to the Brauer group $\mathrm{Br}(K)$ of the field K. In Section 2c we prove (Theorem 2.4) that if G is a connected group (reductive if F is not perfect), then

$$\mathrm{Inv}(G, \mathrm{Br})_{\mathrm{norm}} \simeq \mathrm{Pic}(G).$$

The group $\mathrm{Inv}^3(G, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}}$ for a semisimple simply connected group G has been studied by Rost; see [Garibaldi et al. 2003].

An essential object in the study of cohomological invariants is the notion of a *classifying torsor*: a G-torsor $E \to X$ for a smooth variety X over F such that every G-torsor over an infinite field K/F is isomorphic to the pull-back of $E \to X$ along a K-point of X. If V is a generically free linear representation of G with a nonempty open subset $U \subset V$ such that there is a G-torsor $\pi : U \to X$, then π is classifying. Such representations exist (see Section 2b).

The generic fiber of π is the *generic torsor* over $\mathrm{Spec}\, F(X)$ attached to π. Evaluation at the generic torsor yields a homomorphism

$$\mathrm{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(F(X), \mathbb{Q}/\mathbb{Z}(j)), \tag{1-1}$$

and in Section 3 we show that the image of this map is contained in the subgroup $H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))$ of $H^n(F(X), \mathbb{Q}/\mathbb{Z}(j))$, where $\mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j))$ is the Zariski sheaf associated to the presheaf $W \mapsto H^n(W, \mathbb{Q}/\mathbb{Z}(j))$ of the étale cohomology groups. In fact, the image is contained in the subgroup $H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))_{\mathrm{bal}}$ of *balanced* elements, that is, elements that have the same images under the pull-back homomorphisms with respect to the two projections $(U \times U)/G \to X$. Moreover, the balanced elements precisely describe the image and we prove (Theorem 3.4):

Theorem A. *Let G be a smooth linear algebraic group over a field F. We assume that G is connected if F is a finite field. Let $E \to X$ be a classifying G-torsor with*

E a G-rational variety such that $E(F) \neq \varnothing$. Then (1-1) yields an isomorphism $\mathrm{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j)) \simeq H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))_{\mathrm{bal}}$.

At this point it is convenient to make use of a construction due to Totaro [1999]: because the Chow groups are homotopy invariant, the groups $\mathrm{CH}^n(X)$ do not depend on the choice of the representation V and the open set $U \subset V$ provided the codimension of $V \setminus U$ in V is large enough. This leads to the notation $\mathrm{CH}^n(BG)$, the Chow groups of the so-called *classifying space* BG, although BG itself is not defined in this paper.

Unfortunately, the étale cohomology groups with values in $\mathbb{Q}_p/\mathbb{Z}_p(j)$, where $p = \mathrm{char}(F) > 0$, are not homotopy invariant. In particular, we cannot use the theory of cycle modules of Rost [1996].

The main result of this paper is the exact sequence in Theorem 4.3 describing degree 3 cohomological invariants of an algebraic torus T. Writing $\widehat{T}_{\mathrm{sep}}$ for the character lattice of T over a separable closure of F and T° for the dual torus, we prove our main result:

Theorem B. *Let T be an algebraic torus over a field F. Then there is an exact sequence*

$$0 \to \mathrm{CH}^2(BT)_{\mathrm{tors}} \to H^1(F, T^0) \xrightarrow{\alpha} \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \\ \to H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\operatorname{Dec} \to H^2(F, T^0).$$

The homomorphism α is given by $\alpha(a)(b) = a_K \cup b$ for every $a \in H^1(F, T^0)$ and $b \in H^1(K, T)$ and every field extension K/F, where the cup-product is defined in (4-5), and Dec *is the subgroup of decomposable elements in the symmetric square $S^2(\widehat{T}_{\mathrm{sep}})$ defined in Section A-II.*

In the proof of the theorem we compute the group of balanced elements in the motivic cohomology group $H^4(BT, \mathbb{Z}(2))$ and relate it, using an exact sequence of B. Kahn and Theorem A, with the group of invariants $\mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}}$.

We also prove that the torsion group $\mathrm{CH}^2(BT)_{\mathrm{tors}}$ is finite of exponent 2 (Theorem 4.7) and the last homomorphism in the sequence is also of exponent 2 (see the discussion before Theorem 4.13).

Moreover, if p is an odd prime, the group $\mathrm{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{norm}}$, which is the p-primary component of $\mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}}$, splits canonically into the direct sum of *linear* invariants (those that induce group homomorphisms from Tors_T to H^3) and *quadratic* invariants, that is, the invariants i such that the function $h(a, b) := i(a + b) - i(a) - i(b)$ is bilinear and $h(a, a) = 2i(a)$ for all a and b. Furthermore, the groups of linear and quadratic invariants with values in $\mathbb{Q}_p/\mathbb{Z}_p(2)$ are canonically isomorphic to $H^1(F, T^\circ)\{p\}$ and $(H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\operatorname{Dec})\{p\}$, respectively.

We also prove (Theorem 4.10) that the degree 3 invariants have control over the structure of all invariants. Precisely, the group $\operatorname{Inv}^3(T_K, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}}$ is trivial for all K/F if and only if T is *special*, that is, T has no nontrivial torsors over any field K/F, which in particular means T has no nonconstant H-invariants for every functor H.

Our motivation for considering invariants of tori comes from their connection with unramified cohomology (defined in Section 5). Specifically, this work began as an investigation of a problem posed by Colliot-Thélène [1995, p. 39]: for n prime to $\operatorname{char}(F)$ and $i \geq 0$, determine the unramified cohomology group $H^i_{\mathrm{nr}}(F(S), \mu_n^{\otimes(i-1)})$, where $F(S)$ is the function field of a torus S over F. The connection is provided by Theorem 5.7 where we show that the unramified cohomology of a torus S is calculated by the invariants of an auxiliary torus:

Theorem C. *Let S be a torus over F and let $1 \to T \to P \to S \to 1$ be a flasque resolution of S, that is, T is flasque and P is quasisplit. Then there is a natural isomorphism*

$$H^n_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(j)) \simeq \operatorname{Inv}^n(T, \mathbb{Q}/\mathbb{Z}(j)).$$

By Theorem B and Theorem C, we have an exact sequence

$$\begin{aligned}0 \to \mathrm{CH}^2(BT)_{\mathrm{tors}} \to H^1(F, T^0) \xrightarrow{\alpha} \overline{H}^3_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(2))& \\ \to H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\operatorname{Dec} \to H^2(F, T^0)&\end{aligned}$$

describing the reduced third cohomology group

$$\overline{H}^3_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(2))) := H^3_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(2))/H^3(F, \mathbb{Q}/\mathbb{Z}(2)).$$

Moreover, for an odd prime p, we have a canonical direct sum decomposition of the p-primary components:

$$\overline{H}^3_{\mathrm{nr}}(F(S), \mathbb{Q}_p/\mathbb{Z}_p(2)) = H^1(F, T^0)\{p\} \oplus (H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\operatorname{Dec})\{p\}.$$

Note that the torus S determines T up to multiplication by a quasisplit torus. If X is a smooth compactification of S, one can take the torus T with $\widehat{T}_{\mathrm{sep}} = \operatorname{Pic}(X_{\mathrm{sep}})$; see [Colliot-Thélène and Sansuc 1977, §2].

In the present paper, F denotes a field of arbitrary characteristic, F_{sep} a separable closure of F, and Γ the absolute Galois group $\operatorname{Gal}(F_{\mathrm{sep}}/F)$ of F.

The word "scheme" over a field F means a separated scheme over F and, following [Fulton 1984], a "variety" over F is an integral scheme of finite type over F. If X is a scheme over F and L/F is a field extension then we write X_L for $X \times_F \operatorname{Spec} L$. When $L = F_{\mathrm{sep}}$ we write simply X_{sep}.

A "linear algebraic group over F" is an affine group scheme of finite type over F, not necessarily smooth.

2. Invariants of algebraic groups

2a. *Definitions and basic properties.* Let G be a linear algebraic group over a field F. Consider the functor

$$\mathrm{Tors}_G : \mathit{Fields}_F \to \mathit{Sets}$$

from the category of field extensions of F to the category of sets taking a field K to the set $\mathrm{Tors}_G(K)$ of isomorphism classes of (right) G-torsors over $\operatorname{Spec} K$. Note that if G is a smooth group, then there is a natural bijection

$$\mathrm{Tors}_G(K) \simeq H^1(K, G) := H^1(\mathrm{Gal}(K_{\mathrm{sep}}/K), G(K_{\mathrm{sep}})).$$

Let $H : \mathit{Fields}_F \to \mathit{Abelian\ Groups}$ be a functor. We also view H as a functor with values in Sets. Following [Garibaldi et al. 2003], we define an *H-invariant of G* as a morphism of functors $\mathrm{Tors}_G \to H$ from the category Fields_F to Sets. All the H-invariants of G form the abelian group of invariants $\mathrm{Inv}(G, H)$.

An invariant $i \in \mathrm{Inv}(G, H)$ is called *constant* if there is an element $h \in H(F)$ such that $i(I) = h_K$ for every G-torsor $I \to \operatorname{Spec} K$, where h_K is the image of h under natural map $H(F) \to H(K)$. The constant invariants form a subgroup $\mathrm{Inv}(G, H)_{\mathrm{const}}$ of $\mathrm{Inv}(G, H)$ isomorphic to $H(F)$. An invariant $i \in \mathrm{Inv}(G, H)$ is called *normalized* if $i(I) = 0$ for the trivial G-torsor I. The normalized invariants form a subgroup $\mathrm{Inv}(G, H)_{\mathrm{norm}}$ of $\mathrm{Inv}(G, H)$ and we have the decomposition

$$\mathrm{Inv}(G, H) = \mathrm{Inv}(G, H)_{\mathrm{const}} \oplus \mathrm{Inv}(G, H)_{\mathrm{norm}} \simeq H(F) \oplus \mathrm{Inv}(G, H)_{\mathrm{norm}},$$

so it suffices to determine the normalized invariants.

2b. *Classifying torsors.* Let G be a linear algebraic group over a field F. A G-torsor $E \to X$ over a smooth variety X over F is called *classifying* if for every field extension K/F, with K infinite, and for every G-torsor $I \to \operatorname{Spec} K$, there is a point $x : \operatorname{Spec} K \to X$ such that the torsor I is isomorphic to the fiber $E(x)$ of $E \to X$ over x, that is, $I \simeq E(x) := x^*(E) = \operatorname{Spec}(K) \times_X E$. The generic fiber $E_{\mathrm{gen}} \to \operatorname{Spec} F(X)$ of a classifying torsor is called a *generic G-torsor*; see [ibid., Part 1, §5.3].

If V is a generically free linear representation of G with a nonempty open subset $U \subset V$ such that there is a G-torsor $\pi : U \to X$, then π is classifying; see [ibid., Part 1, §5.4]. We will write U/G for X and call π a *standard classifying G-torsor*. Standard classifying G-torsors exist: we can embed G into $U := \mathbf{GL}_{n,F}$ for some n as a closed subgroup. Then U is an open subset in the affine space $M_n(F)$ on which G acts linearly and the canonical morphism $U \to X := U/G$ is a G-torsor. Note that $U(F) \neq \varnothing$.

We say that a G-variety Y is *G-rational* if there is an affine space V with a linear G-action such that Y and V have G-isomorphic nonempty open G-invariant subvarieties. Note that if $U \to U/G$ is a standard classifying G-torsor, then U is a G-rational variety.

Let $E \to X$ be a classifying G-torsor and let $H : \mathit{Fields}_F \to \mathit{Abelian\ Groups}$ be a functor. Define the map

$$\theta_G : \operatorname{Inv}(G, H) \to H(F(X)), \quad i \mapsto i(E_{\text{gen}}), \tag{2-1}$$

by sending an invariant to its value at the generic torsor E_{gen}.

Consider the following property of the functor H:

Property 2.1. The map $H(K) \to H(K((t)))$ is injective for any field extension K/F.

The following theorem, due to M. Rost, was proved in [Garibaldi et al. 2003, Part II, Theorem 3.3]. For completeness, we give a slightly modified proof in Section A-I.

Theorem 2.2. *Let G be a smooth linear algebraic group over F. If a functor $H : \mathit{Fields}_F \to \mathit{Abelian\ Groups}$ has Property 2.1, then the map θ_G is injective, that is, every H-invariant of G is determined by its value at the generic G-torsor.*

Let G' be a (closed) subgroup of G over F. The map of sets

$$H^1(K, G') \to H^1(K, G)$$

for every field extension K/F yields the *restriction map*

$$\operatorname{res} : \operatorname{Inv}(G, H) \to \operatorname{Inv}(G', H).$$

Choose standard torsors $\pi : U \to U/G$ and $\pi' : U \to U/G'$ (for example, with $U = \mathbf{GL}_{n,F}$ as above). The pull-back of π with respect to the natural morphism $\alpha : U/G' \to U/G$ is the push-forward of π' via the inclusion $G' \hookrightarrow G$. It follows that the diagram

$$\begin{tikzcd} \operatorname{Inv}(G, H) \arrow[r, "\operatorname{res}"] \arrow[d, "\theta_G"'] & \operatorname{Inv}(G', H) \arrow[d, "\theta_{G'}"'] \\ H(F(U/G)) \arrow[r, "\alpha^*"] & H(F(U/G')) \end{tikzcd}$$

is commutative.

2c. *The Brauer group invariants.* Let G be a smooth connected linear algebraic group over F. Every cohomological invariant of G of degree 1 is constant by [Knus et al. 1998, Proposition 31.15]. In this section we study (degree 2) Br-invariants for the Brauer group functor $K \mapsto \operatorname{Br}(K)$. We assume that G is reductive if $\operatorname{char}(F) > 0$.

Lemma 2.3. *For any field extension K/F such that F is algebraically closed in K, the natural map* $\operatorname{Pic}(G) \to \operatorname{Pic}(G_K)$ *is an isomorphism.*

Proof. We may assume that G is reductive by factoring out the unipotent radical in the case that F is perfect. There is an exact sequence (see [Colliot-Thélène 2004, Theorem 1.2])

$$1 \to C \to G' \to G \to 1$$

with C a torus and G' a reductive group with $\mathrm{Pic}(G'_L) = 0$ for any field extension L/F. Let T be the factor group of G' by the semisimple part. The result follows from the exact sequence [Sansuc 1981, Proposition 6.10] (note that G is reductive if L is not perfect)

$$\widehat{T}(L) \to \widehat{C}(L) \to \mathrm{Pic}(G_L) \to \mathrm{Pic}(G'_L) = 0$$

with $L = F$ and K since the groups $\widehat{T}(F)$ and $\widehat{C}(F)$ don't change when F is replaced by K. $\square$

Since for any G_K-torsor $E \to \mathrm{Spec}(K)$ over a field extension K/F one has [Sansuc 1981, Proposition 6.10] the exact sequence

$$\mathrm{Pic}(E) \to \mathrm{Pic}(G_K) \xrightarrow{\delta} \mathrm{Br}(K) \xrightarrow{\varepsilon} \mathrm{Br}(E), \tag{2-2}$$

we obtain the homomorphism

$$\nu : \mathrm{Pic}(G) \to \mathrm{Inv}(G, \mathrm{Br}),$$

which takes an element $\alpha \in \mathrm{Pic}(G)$ to the invariant that sends a G-torsor E over a field extension K/F to $\delta(\alpha_K)$. If E is a trivial torsor, that is, $E(K) \neq \varnothing$, then ε is injective and hence $\delta = 0$. It follows that the invariant $\nu(\alpha)$ is normalized.

Theorem 2.4. *Let G be a smooth connected linear algebraic group over F. Assume that G is reductive if* $\mathrm{char}(F) > 0$. *Then the map* $\nu : \mathrm{Pic}(G) \to \mathrm{Inv}(G, \mathrm{Br})_{\mathrm{norm}}$ *is an isomorphism.*

Proof. Choose a standard classifying G-torsor $U \to U/G$. Write K for the function field $F(U/G)$ and let U_{gen} be the generic G-torsor over K. Consider the commutative diagram

$$\begin{tikzcd}
 & \mathrm{Pic}(G) \arrow[r, "\nu"] \arrow[d, "j"'] & \mathrm{Inv}(G, \mathrm{Br})_{\mathrm{norm}} \arrow[d, "\theta_G"'] & \\
\mathrm{Pic}(U_{\mathrm{gen}}) \arrow[r] & \mathrm{Pic}(G_K) \arrow[r, "\delta"] & \mathrm{Br}(K) \arrow[r, "i"] & \mathrm{Br}(K(U_{\mathrm{gen}})),
\end{tikzcd}$$

where the bottom sequence is (2-2) for the G-torsor $U_{\mathrm{gen}} \to \mathrm{Spec}(K)$ followed by the injection $\mathrm{Br}(U_{\mathrm{gen}}) \to \mathrm{Br}(K(U_{\mathrm{gen}}))$ (see [Milne 1980, Chapter IV, Corollary 2.6]), and the map θ_G is evaluation at the generic torsor U_{gen} given in (2-1) and is injective by Theorem 2.2. Since the generic torsor is split over $K(U_{\mathrm{gen}})$, $\mathrm{Im}(\theta_G) \subset \mathrm{Ker}(i) = \mathrm{Im}(\delta)$. By Lemma 2.3, j is an isomorphism, hence ν is surjective.

Note that U_{gen} is a localization of U, hence $\operatorname{Pic}(U_{\mathrm{gen}})=0$ as $\operatorname{Pic}(U)=0$. It follows that ν is injective. □

An algebraic group G over a field F is called *special* if $H^1(K,G)=\{1\}$ for every field extension K/F, that is, all G-torsors over any field extension of F are trivial.

Corollary 2.5. *If the group G is special, then* $\operatorname{Pic}(G)=0$.

3. Invariants with values in $\mathbb{Q}/\mathbb{Z}(j)$

In this section we find a description for the group of cohomological invariants with values in $\mathbb{Q}/\mathbb{Z}(j)$ by identifying the image of the embedding θ_G in (2-1).

Let G be a linear algebraic group over a field F, let $H\subset G$ be a subgroup and let $E\to X$ be a G-torsor. Suppose that G/H is affine. Consider a G-action on $E\times(G/H)$ by $(e,g'H)g=(eg,g^{-1}g'H)$. By [Milne 1980, Theorem I.2.23], the affine G-equivariant projection $E\times(G/H)\to E$ descends to an affine morphism $Y\to X$. The (trivial right) H-torsor $E\times G\to E\times(G/H)$ descends to an H-torsor $E\to Y$. We will write E/H for Y.

Example 3.1. Let G be a linear algebraic group over a field F and let $E\to X$ be a G-torsor. Then for every $n>0$, $E^n:=E\times_F\cdots\times_F E$ (n times) is a G^n-torsor over X^n. Viewing G as the diagonal subgroup of G^n, we have the G-torsor $E^n\to E^n/G$.

3a. *Balanced elements.* Let G be a linear algebraic group over a field F. We assume that G is connected if F is finite. Let $E\to X$ be a G-torsor such that $E(F)\neq\varnothing$. We write p_1 and p_2 for the two projections $E^2/G=(E\times_F E)/G\to X$ (see Example 3.1).

Lemma 3.2. *Let K/F be a field extension and $x_1,x_2\in X(K)$. Then the G-torsors $E(x_1)$ and $E(x_2)$ over K are isomorphic if and only if there is a point $y\in(E^2/G)(K)$ such that $p_1(y)=x_1$ and $p_2(y)=x_2$.*

Proof. "⇒": By construction, we have G-equivariant morphisms $f_i:E(x_i)\to E$ for $i=1,2$. Choose an isomorphism $h:E(x_1)\xrightarrow{\sim}E(x_2)$ of G-torsors over K. The morphism $(f_1,f_2h):E(x_1)\to E^2$ yields the required point $\operatorname{Spec}K=E(x_1)/G\to E^2/G$.

"⇐": The pull-back of $E\to X$ with respect to any projection $E^2/G\to X$ coincides with the G-torsor $E^2\to E^2/G$, hence

$$E(x_1)=x_1^*(E)=y^*p_1^*(E)\simeq y^*(E^2)\simeq y^*p_2^*(E)=x_2^*(E)=E(x_2).\qquad\square$$

Let H be a (contravariant) functor from the category of schemes over F to the category of abelian groups. We have the two maps $p_i^*:H(X)\to H(E^2/G)$, $i=1,2$. An element $h\in H(X)$ is called *balanced* if $p_1^*(h)=p_2^*(h)$. We write $H(X)_{\mathrm{bal}}$ for the subgroup of balanced elements in $H(X)$. In other words, $H(X)_{\mathrm{bal}}=h_0(H(E^\bullet/G))$ in the notation of Section A-IV.

We can view H as a (covariant) functor $\mathit{Fields}_F \to \mathit{Sets}$ taking a field K to $H(K) := H(\operatorname{Spec} K)$.

Lemma 3.3. *Let $h \in H(X)_{\mathrm{bal}}$ be a balanced element, K/F a field extension and I a G-torsor over* $\operatorname{Spec}(K)$. *Let $x \in X(K)$ be a point such that $E(x) \simeq I$. Then the element $x^*(h)$ in $H(K)$ does not depend on the choice of x.*

Proof. Let $x_1, x_2 \in X(K)$ be two points such that $E(x_1) \simeq E(x_2)$. By Lemma 3.2, there is a point $y \in (E^2/G)(K)$ such that $p_1(y) = x_1$ and $p_2(y) = x_2$. Therefore

$$x_1^*(h) = y^*(p_1^*(h)) = y^*(p_2^*(h)) = x_2^*(h). \qquad \square$$

It follows from Lemma 3.3 that if the torsor $E \to X$ is classifying with $E(F) \neq \varnothing$, then every element $h \in H(X)_{\mathrm{bal}}$ determines an H-invariant i_h of G as follows. Let I be a G-torsor over a field extension K/F. We claim that there is a point $x \in X(K)$ such that $E(x) \simeq I$. If K is infinite, this follows from the definition of the classifying G-torsor. If K is finite then all G-torsors over K are trivial by [Lang 1956], as G is connected. Since $E(K) \neq \varnothing$, we can take for x the image in $X(K)$ of any point in $E(K)$. Defining $i_h(E) = x^*(h) \in H(K)$, we have a group homomorphism

$$H(X)_{\mathrm{bal}} \to \operatorname{Inv}(G, H), \quad h \mapsto i_h.$$

3b. ***Cohomology with values in $\mathbb{Q}/\mathbb{Z}(j)$.*** For every integer $j \geq 0$, the coefficients $\mathbb{Q}/\mathbb{Z}(j)$ are defined as the direct sum over all prime integers p of the objects $\mathbb{Q}_p/\mathbb{Z}_p(j)$ in the derived category of sheaves of abelian groups on the big étale site of $\operatorname{Spec} F$, where

$$\mathbb{Q}_p/\mathbb{Z}_p(j) = \operatorname*{colim}_n \mu_{p^n}^{\otimes j}$$

if $p \neq \operatorname{char} F$, with μ_{p^n} the sheaf of (p^n)-th roots of unity, and

$$\mathbb{Q}_p/\mathbb{Z}_p(j) = \operatorname*{colim}_n W_n\Omega^j_{\log}[-j]$$

if $p = \operatorname{char} F > 0$, with $W_n\Omega^j_{\log}$ the sheaf of *logarithmic de Rham–Witt differentials*; see [Illusie 1979, I.5.7; Kahn 1996].

We write $H^m(X, \mathbb{Q}/\mathbb{Z}(j))$ for the étale cohomology of a scheme X with values in $\mathbb{Q}/\mathbb{Z}(j)$. Then

$$H^m(X, \mathbb{Q}/\mathbb{Z}(j))\{p\} = \operatorname*{colim}_n H^m(X, \mu_{p^n}^{\otimes j})$$

if $p \neq \operatorname{char} F$ and

$$H^m(X, \mathbb{Q}/\mathbb{Z}(j))\{p\} = \operatorname*{colim}_n H^{m-j}(X, W_n\Omega^j_{\log})$$

if $p = \operatorname{char} F > 0$. In the latter case, the group $W_n\Omega^j_{\log}(F)$ is canonically isomorphic to $K_j^M(F)/p^n K_j^M(F)$, where $K_j^M(F)$ is Milnor's K-group of F (see

[Bloch and Kato 1986, Corollary 2.8]), hence by [Izhboldin 1991; Garibaldi et al. 2003, Part II, Appendix A], $H^s(F, W_n\Omega^j_{\log})$ is isomorphic to

$$H^s(F, K_j^M(F_{\mathrm{sep}})/p^n K_j^M(F_{\mathrm{sep}})) = \begin{cases} K_j^M(F)/p^n K_j^M(F) & \text{if } s = 0, \\ H^2(F, K_j^M(F_{\mathrm{sep}}))_{p^n} & \text{if } s = 1, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that in the case $p = \operatorname{char} F > 0$, we have

$$H^m(F, \mathbb{Q}/\mathbb{Z}(j))\{p\} = \begin{cases} K_j^M(F) \otimes (\mathbb{Q}_p/\mathbb{Z}_p) & \text{if } m = j, \\ H^2(F, K_j^M(F_{\mathrm{sep}}))\{p\} & \text{if } m = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

The *motivic complexes* $\mathbb{Z}(j)$, for $j = 0, 1, 2$, of étale sheaves on a smooth scheme X were defined by S. Lichtenbaum [1987; 1990]. We write $H^*(X, \mathbb{Z}(j))$ for the étale (hyper)cohomology groups of X with values in $\mathbb{Z}(j)$.

The complex $\mathbb{Z}(0)$ is equal to the constant sheaf $\mathbb{Z}$ and $\mathbb{Z}(1) = \mathbb{G}_{m,X}[-1]$, thus $H^n(X, \mathbb{Z}(1)) = H^{n-1}(X, \mathbb{G}_{m,X})$. In particular, $H^3(X, \mathbb{Z}(1)) = \operatorname{Br}(X)$, the cohomological Brauer group of X. The complex $\mathbb{Z}(2)$ is concentrated in degrees 1 and 2 and there is a product map $\mathbb{Z}(1) \otimes^L \mathbb{Z}(1) \to \mathbb{Z}(2)$; see [Lichtenbaum 1987, Proposition 2.5].

The exact triangle in the derived category of étale sheaves

$$\mathbb{Z}(j) \to \mathbb{Q} \otimes \mathbb{Z}(j) \to \mathbb{Q}/\mathbb{Z}(j) \to \mathbb{Z}(j)[1]$$

yields the connecting homomorphism

$$H^i(X, \mathbb{Q}/\mathbb{Z}(j)) \to H^{i+1}(X, \mathbb{Z}(j)),$$

which is an isomorphism if $X = \operatorname{Spec}(F)$ for a field F and $i > j$ [Kahn 1993, Lemme 1.1].

Write $\mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j))$ for the Zariski sheaf on a smooth scheme X associated to the presheaf $U \mapsto H^n(U, \mathbb{Q}/\mathbb{Z}(j))$ of étale cohomology groups.

Let G be a linear algebraic group over F. We assume that G is connected if F is a finite field and write $\operatorname{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j))$ for the group of *degree n invariants* of G for the functor $K \mapsto H^n(K, \mathbb{Q}/\mathbb{Z}(j))$. Note that Property 2.1 holds for this functor by [Garibaldi et al. 2003, Part 2, Proposition A.9].

Choose a classifying G-torsor $E \to X$ with E a G-rational variety such that $E(F) \neq \varnothing$. Applying the construction given in Section 3a to the functor $U \mapsto H^0_{\mathrm{Zar}}(U, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))$, we get a homomorphism

$$\varphi : H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))_{\mathrm{bal}} \to \operatorname{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j)).$$

Theorem 3.4. *Let G be a smooth linear algebraic group over a field F. We assume that G is connected if F is a finite field. Let $E \to X$ be a classifying G-torsor* *with E a G-rational variety such that $E(F) \neq \varnothing$. Then the homomorphism φ is an isomorphism.*

Proof. Let $E_{\mathrm{gen}} \to F(X)$ be the generic fiber of the classifying G-torsor $E \to X$. Note that since the pull-back of $E \to X$ with respect to any of the two projections $E^2/G \to X$ coincides with the G-torsor $E^2 \to E^2/G$, the pull-backs of the generic G-torsor $E_{\mathrm{gen}} \to \operatorname{Spec} F(X)$ with respect to the two morphisms $\operatorname{Spec} F(E^2/G) \to \operatorname{Spec} F(X)$ induced by the projections are isomorphic. It follows that for every invariant $i \in \operatorname{Inv}(G, H^*(\mathbb{Q}/\mathbb{Z}(j)))$ we have

$$p_1^*(i(E_{\mathrm{gen}})) = i(p_1^*(E_{\mathrm{gen}})) = i(p_2^*(E_{\mathrm{gen}})) = p_2^*(i(E_{\mathrm{gen}}))$$

in $H^*(F(E^2/G), \mathbb{Q}/\mathbb{Z}(j))$, that is, $i(E_{\mathrm{gen}}) \in H^*(F(X), \mathbb{Q}/\mathbb{Z}(j))_{\mathrm{bal}}$. By Proposition A.9, $\partial_x(h) = 0$ for every point $x \in X$ of codimension 1, hence

$$\theta_G(i) = i(E_{\mathrm{gen}}) \in H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))_{\mathrm{bal}}$$

by Proposition A.10. By Theorem 2.2, θ_G is injective and by construction, the composition $\theta_G \circ \varphi$ is the identity. It follows that φ is an isomorphism. □

Write $\overline{H}^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))$ for the factor group of $H^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))$ by the natural image of $H^n(F, \mathbb{Q}/\mathbb{Z}(j))$.

Corollary 3.5. *The isomorphism φ yields an isomorphism*

$$\overline{H}^0_{\mathrm{Zar}}(X, \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)))_{\mathrm{bal}} \xrightarrow{\sim} \operatorname{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j))_{\mathrm{norm}}.$$

4. Degree 3 invariants of algebraic tori

In this section we prove the main theorem that describes degree 3 invariants of an algebraic torus with values in $\mathbb{Q}/\mathbb{Z}(2)$.

4a. ***Algebraic tori.*** Let F be a field and $\Gamma = \operatorname{Gal}(F_{\mathrm{sep}}/F)$ the absolute Galois group of F. An *algebraic torus* of dimension n over F is an algebraic group T such that T_{sep} is isomorphic to the product of n copies of the multiplicative group $\mathbb{G}_m$; see [Colliot-Thélène and Sansuc 1977; Voskresenskiĭ 1998]. For an algebraic torus T over a field F, we write $\widehat{T}_{\mathrm{sep}}$ for the Γ-module of characters $\operatorname{Hom}(T_{\mathrm{sep}}, \mathbb{G}_m)$. The group $\widehat{T}_{\mathrm{sep}}$ is a Γ-*lattice*, that is, a free abelian group of finite rank with a continuous Γ-action. The contravariant functor $T \mapsto \widehat{T}_{\mathrm{sep}}$ is an antiequivalence between the category of algebraic tori and the category of Γ-lattices: the torus T and the group $T(F)$ can be reconstructed from the lattice $\widehat{T}_{\mathrm{sep}}$ by the formulas

$$T = \operatorname{Spec}(F_{\mathrm{sep}}[\widehat{T}_{\mathrm{sep}}]^{\Gamma}),$$
$$T(F) = \operatorname{Hom}_{\Gamma}(\widehat{T}_{\mathrm{sep}}, F_{\mathrm{sep}}^{\times}) = (\widehat{T}_{\mathrm{sep}}^{\circ} \otimes F_{\mathrm{sep}}^{\times})^{\Gamma},$$

where $\widehat{T}_{\mathrm{sep}}^{\circ} = \operatorname{Hom}(\widehat{T}_{\mathrm{sep}}, \mathbb{Z})$.

We write $\widehat{T}$ for the character group $\mathrm{Hom}_F(T, \mathbb{G}_m) = (\widehat{T}_{\mathrm{sep}})^\Gamma$ and T° for the *dual torus* having character lattice $\widehat{T}^\circ_{\mathrm{sep}}$.

A torus T is called *quasisplit* if T is isomorphic to the group of invertible elements of an étale F-algebra, or equivalently, the Γ-lattice $\widehat{T}_{\mathrm{sep}}$ is permutation, that is, $\widehat{T}_{\mathrm{sep}}$ has a Γ-invariant $\mathbb{Z}$-basis. An *invertible* torus is a direct factor of a quasisplit torus.

A torus T is called *flasque* or *coflasque* if $H^1(L, \widehat{T}^\circ_{\mathrm{sep}}) = 0$ or $H^1(L, \widehat{T}_{\mathrm{sep}}) = 0$, respectively, for every finite field extension L/F. A *flasque resolution* of a torus S is an exact sequence of tori $1 \to T \to P \to S \to 1$ with T flasque and P quasisplit. By [Colliot-Thélène and Sansuc 1977, §4; Voskresenskiĭ 1998, §4.7], the torus T in the flasque resolution is invertible if and only if S is a direct factor of a rational torus.

4b. ***Products.*** Let T be a torus over F and let $\widehat{T}(i)$ denote the complex $\widehat{T}_{\mathrm{sep}} \otimes \mathbb{Z}(i)$ of étale sheaves over F for $i = 0, 1, 2$. Thus, $\widehat{T}(0) = \widehat{T}_{\mathrm{sep}}$ and

$$\widehat{T}(1) = (\widehat{T}_{\mathrm{sep}} \otimes F^\times_{\mathrm{sep}})[-1] = T^\circ(F_{\mathrm{sep}})[-1].$$

Let S and T be algebraic tori over F and let i and j be nonnegative integers with $i + j \le 2$. For any smooth variety X over F, we have the product map

$$(\widehat{S}_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}})^\Gamma \otimes H^p(X, \widehat{S}^\circ(i)) \otimes H^q(X, \widehat{T}^\circ(j)) \to H^{p+q}(X, \mathbb{Z}(i + j)) \tag{4-1}$$

taking $a \otimes b \otimes c$ to $a \cup b \cup c$, via the canonical pairings between $\widehat{S}_{\mathrm{sep}}$ and $\widehat{S}^\circ_{\mathrm{sep}}$, $\widehat{T}_{\mathrm{sep}}$ and $\widehat{T}^\circ_{\mathrm{sep}}$, and the product map $\mathbb{Z}(i) \otimes^L \mathbb{Z}(j) \to \mathbb{Z}(i + j)$.

Recall that there is an isomorphism $H^n(F, \mathbb{Z}(k)) \simeq H^{n-1}(F, \mathbb{Q}/\mathbb{Z}(k))$ for $n > k$. In particular, we have the cup-product map

$$(\widehat{S}_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}})^\Gamma \otimes H^p(F, S) \otimes H^q(F, T) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \tag{4-2}$$

if $p + q = 2$.

If $S = T^\circ$ is the dual torus, then $(\widehat{S}_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}})^\Gamma = \mathrm{End}_\Gamma(\widehat{T}_{\mathrm{sep}})$ contains the canonical element 1_T. We then have the product map

$$H^p(X, \widehat{T}(i)) \otimes H^q(X, \widehat{T}^\circ(j)) \to H^{p+q}(X, \mathbb{Z}(i + j)) \tag{4-3}$$

and in particular, the product maps

$$H^1(F, \widehat{T}_{\mathrm{sep}}) \otimes H^1(F, T) \to H^2(F, \mathbb{Q}/\mathbb{Z}(1)) = \mathrm{Br}(F), \tag{4-4}$$

$$H^1(F, T^\circ) \otimes H^1(F, T) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2)), \tag{4-5}$$

$$H^2(F, T^\circ) \otimes H^0(F, T) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2)), \tag{4-6}$$

taking $a \otimes b$ to $1_T \cup a \cup b$ and applying (4-2).

As T is a commutative group, the set $H^1(K, T)$ is an abelian group. An invariant $i \in \mathrm{Inv}(T, H)$ for a functor H is called *linear* if $i_K : H^1(K, T) \to H(K)$ is a group homomorphism for every K/F. In the next section we will see that a normalized degree 3 invariant of a torus need not be linear.

4c. ***Main theorem.*** Let T be a torus over F and choose a standard classifying T-torsor $U \to U/T$ such that the codimension of $V \setminus U$ in V is at least 3. Such a torsor exists by [Edidin and Graham 1998, Lemma 9].

By [Sansuc 1981, Proposition 6.10], there is an exact sequence

$$F_{\mathrm{sep}}[U]^{\times}/F_{\mathrm{sep}}^{\times} \to \widehat{T}_{\mathrm{sep}} \to \mathrm{Pic}((U/T)_{\mathrm{sep}}) \to \mathrm{Pic}(U_{\mathrm{sep}}).$$

The codimension assumption implies that the side terms are trivial, hence the map $\widehat{T}_{\mathrm{sep}} \to \mathrm{Pic}((U/T)_{\mathrm{sep}})$ is an isomorphism. It follows that the classifying T-torsor $U \to U/T$ is universal in the sense of [Colliot-Thélène and Sansuc 1987a].

Write $K_*(F)$ for the (Quillen) K-groups of F and $\mathcal{K}_*$ for the Zariski sheaf associated to the presheaf $U \mapsto K_*(U)$. Then the groups $H^n_{\mathrm{Zar}}(U/T, \mathcal{K}_2)$ are independent of the choice of the classifying torsor; see [Edidin and Graham 1998]. So we write $H^n_{\mathrm{Zar}}(BT, \mathcal{K}_2)$ for this group (see Section A-IV). As T_{sep} is a split torus, by the Künneth formula (see Example A.5),

$$H^n_{\mathrm{Zar}}(BT_{\mathrm{sep}}, \mathcal{K}_2) = \begin{cases} K_2(F_{\mathrm{sep}}) & \text{if } n = 0, \\ \mathrm{Pic}((U/T)_{\mathrm{sep}}) \otimes F_{\mathrm{sep}}^{\times} = \widehat{T}_{\mathrm{sep}} \otimes F_{\mathrm{sep}}^{\times} = T^{\circ}(F_{\mathrm{sep}}) & \text{if } n = 1, \\ \mathrm{CH}^2((U/T)_{\mathrm{sep}}) = S^2(\widehat{T}_{\mathrm{sep}}) & \text{if } n = 2. \end{cases}$$

Applying the calculation of the $\mathcal{K}$-cohomology groups to the standard classifying T-torsor $U^i \to U^i/T$ for every $i > 0$ instead of $U \to U/T$, by Proposition B.3, we have the exact sequence

$$0 \to H^1(F, T^{\circ}) \xrightarrow{\alpha} \overline{H}^4(U^i/T, \mathbb{Z}(2)) \to \overline{H}^4((U^i/T)_{\mathrm{sep}}, \mathbb{Z}(2))^{\Gamma} \to H^2(F, T^{\circ}), \quad (4\text{-}7)$$

where $\overline{H}^4(U^i/T, \mathbb{Z}(2))$ is the factor group of $H^4(U^i/T, \mathbb{Z}(2))$ by $H^4(F, \mathbb{Z}(2))$, the map α is given by $\alpha(a) = q^*(a) \cup [U^i]$ with $q : U^i/T \to \operatorname{Spec} F$ the structure morphism, $[U^i]$ the class of the T-torsor $U^i \to U^i/T$ in $H^1(U^i/T, T)$, and the cup-product is taken for the pairing (B-6).

Taking the sequences (4-7) for all i (see Section A-IV), we get the exact sequence of cosimplicial groups

$$0 \to H^1(F, T^{\circ}) \xrightarrow{\alpha} \overline{H}^4(U^{\bullet}/T, \mathbb{Z}(2)) \to \overline{H}^4((U^{\bullet}/T)_{\mathrm{sep}}, \mathbb{Z}(2))^{\Gamma} \to H^2(F, T^{\circ}).$$

The first and the last cosimplicial groups in the sequence are constant, hence by Lemma A.2, the sequence

$$0 \to H^1(F, T^{\circ}) \xrightarrow{\alpha} \overline{H}^4(U/T, \mathbb{Z}(2))_{\mathrm{bal}} \to \overline{H}^4((U/T)_{\mathrm{sep}}, \mathbb{Z}(2))^{\Gamma}_{\mathrm{bal}} \to H^2(F, T^{\circ}) \quad (4\text{-}8)$$

is exact as $h_0(\overline{H}^4(U^\bullet/T,\mathbb{Z}(2)))=H^4(U/T,\mathbb{Z}(2))_{\mathrm{bal}}$.

The following theorem was proved by B. Kahn [1996, Theorem 1.1]:

Theorem 4.1. *Let X be a smooth variety over F. Then there is an exact sequence*

$$0\to \mathrm{CH}^2(X)\to H^4(X,\mathbb{Z}(2))\to H^0_{\mathrm{Zar}}(X,\mathcal{H}^3(\mathbb{Q}/\mathbb{Z}(2)))\to 0.$$

By Theorem 4.1, there is an exact sequence of cosimplicial groups

$$0\to \mathrm{CH}^2(U^\bullet/T)\to \overline{H}^4(U^\bullet/T,\mathbb{Z}(2))\to \overline{H}^0_{\mathrm{Zar}}(U^\bullet/T,\mathcal{H}^3(\mathbb{Q}/\mathbb{Z}(2)))\to 0.$$

As the functor CH^2 is homotopy invariant, by Lemma A.4, the first group in the sequence is constant. In view of Lemma A.2, and following the notation for the $\mathcal{H}$-cohomology, the sequence

$$0\to \mathrm{CH}^2(BT)\to \overline{H}^4(U/T,\mathbb{Z}(2))_{\mathrm{bal}}\to \overline{H}^0_{\mathrm{Zar}}(U/T,\mathcal{H}^3(\mathbb{Q}/\mathbb{Z}(2)))_{\mathrm{bal}}\to 0 \tag{4-9}$$

is exact. By Corollary 3.5, the last group in the sequence is canonically isomorphic to $\mathrm{Inv}(T,H^3(\mathbb{Q}/\mathbb{Z}(2)))_{\mathrm{norm}}$.

As the torus T_{sep} is split, all the invariants of T_{sep} are trivial hence the sequence (4-9) over F_{sep} yields an isomorphism

$$\overline{H}^4((U/T)_{\mathrm{sep}},\mathbb{Z}(2))_{\mathrm{bal}}\simeq \mathrm{CH}^2(BT_{\mathrm{sep}})\simeq S^2(\widehat{T}_{\mathrm{sep}}). \tag{4-10}$$

Combining (4-8), (4-9) and (4-10), we get the following diagram with an exact row and column:

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & H^1(F,T^0) \arrow[d, "\alpha"'] \arrow[dr] & & \\
0 \arrow[r] & \mathrm{CH}^2(BT) \arrow[r] \arrow[dr] & \overline{H}^4(U/T,\mathbb{Z}(2))_{\mathrm{bal}} \arrow[r] \arrow[d] & \mathrm{Inv}^3(T,\mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \arrow[r] & 0. \\
 & & S^2(\widehat{T}_{\mathrm{sep}})^\Gamma \arrow[d] & & \\
 & & H^2(F,T^0) & &
\end{tikzcd}$$

Write $\mathrm{Dec}=\mathrm{Dec}(\widehat{T}_{\mathrm{sep}})$ for the subgroup of *decomposable elements* in $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$ (see Section A-II).

Lemma 4.2. *The image of the homomorphism*

$$\mathrm{CH}^2(BT)\to \mathrm{CH}^2(BT_{\mathrm{sep}})^\Gamma=S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$$

in the diagram coincides with Dec.

Proof. Consider the Grothendieck ring $K_0(BT)$ of the category of T-equivariant vector bundles over $\mathrm{Spec}(F)$, or equivalently, of the category of finite dimensional linear representations of T. If T is split, every linear representation of T is a direct sum of one-dimensional representations. Therefore, there is an isomorphism between the group ring $\mathbb{Z}[\widehat{T}]$ of all formal finite sums $\sum_{x\in\widehat{T}} a_x e^x$ and $K_0(BT)$, taking e^x with $x \in \widehat{T}$ to the class of the 1-dimensional representation given by x. In general, for every torus T, we have $K_0(BT_{\mathrm{sep}}) = \mathbb{Z}[\widehat{T}_{\mathrm{sep}}]$ and $K_0(BT) = \mathbb{Z}[\widehat{T}_{\mathrm{sep}}]^\Gamma = K_0(BT_{\mathrm{sep}})^\Gamma$; see [Merkurjev and Panin 1997, page 136]. The group $\mathbb{Z}[\widehat{T}_{\mathrm{sep}}]^\Gamma$ is generated by the sums $\sum_{i=1}^n e^{\gamma_i x}$, where $\gamma_1, \gamma_2, \dots, \gamma_n$ are representatives of the left cosets of an arbitrary open subgroup Γ' in Γ and $x \in (\widehat{T}_{\mathrm{sep}})^{\Gamma'}$.

The equivariant Chern classes were defined in [Edidin and Graham 1998, §2.4]. The first Chern class $c_1 : K_0(BT_{\mathrm{sep}}) \to \mathrm{CH}^1(BT_{\mathrm{sep}}) = \widehat{T}_{\mathrm{sep}}$ takes e^x to x. In the diagram

$$\begin{tikzcd}
\mathbb{Z}[\widehat{T}_{\mathrm{sep}}]^\Gamma \arrow[r, equal] \arrow[d] & K_0(BT) \arrow[r, "c_2"] \arrow[d] & \mathrm{CH}^2(BT) \arrow[d] & \\
\mathbb{Z}[\widehat{T}_{\mathrm{sep}}] \arrow[r, equal] & K_0(BT_{\mathrm{sep}}) \arrow[r, "c_2"] & \mathrm{CH}^2(BT_{\mathrm{sep}}) \arrow[r, equal] & S^2(\widehat{T}_{\mathrm{sep}})
\end{tikzcd}$$

the second Chern class maps c_2 are surjective by [Esnault et al. 1998, Lemma C.3]. It follows from the formula $c_2(a+b) = c_2(a) + c_1(a)c_1(b) + c_2(b)$ that the composition

$$\mathbb{Z}[\widehat{T}_{\mathrm{sep}}]^\Gamma = K_0(BT) \to K_0(BT_{\mathrm{sep}}) \xrightarrow{c_2} \mathrm{CH}^2(BT_{\mathrm{sep}}) = S^2(\widehat{T}_{\mathrm{sep}}) \to S^2(\widehat{T}_{\mathrm{sep}})/(\widehat{T})^2$$

is a homomorphism and its image is generated by the elements (see Section A-II)

$$c_2\left(\sum_{i=1}^n e^{\gamma_i x}\right) = \sum_{i<j} (\gamma_i x)(\gamma_j x) = \mathrm{Qtr}(x). \qquad \square$$

By the restriction-corestriction argument, the kernel of the homomorphism

$$\mathrm{CH}^2(BT) \to \mathrm{CH}^2(BT_{\mathrm{sep}})^\Gamma = S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$$

coincides with the torsion subgroup $\mathrm{CH}^2(BT)_{\mathrm{tors}}$ in $\mathrm{CH}^2(BT)$.

The following theorem describes degree 3 invariants of an algebraic torus with values in $\mathbb{Q}/\mathbb{Z}(2)$:

Theorem 4.3. *Let T be an algebraic torus a field F. Then there is an exact sequence*

$$0 \to \mathrm{CH}^2(BT)_{\mathrm{tors}} \to H^1(F, T^0) \xrightarrow{\alpha} \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \\ \to S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \mathrm{Dec} \to H^2(F, T^0).$$

The homomorphism α is given by $\alpha(a)(b) = a_K \cup b$ for every $a \in H^1(F, T^0)$ and $b \in H^1(K, T)$ and every field extension K/F, where the cup-product is defined in (4-5).

Proof. The exactness of the sequence follows from the diagram before Lemma 4.2. It remains to describe the map α. Take an $a \in H^1(F, T^0)$ and consider the invariant i defined by $i(b) = a_K \cup b$, where the cup-product is given by (4-5). We need to prove that $i = \alpha(a)$. Choose a standard classifying T-torsor $U \to U/T$ and set $K = F(U/T)$. Let U_{gen} be the generic fiber of the classifying torsor. By Theorem 2.2, it suffices to show that $i(U_{\mathrm{gen}}) = \alpha(a)(U_{\mathrm{gen}})$ over K. This follows from the description of the map α in the exact sequence (4-7). □

Remark 4.4. In a similar (and much simpler) fashion one can describe degree 2 invariants of an algebraic torus T with values in $\mathbb{Q}/\mathbb{Z}(1)$, that is, invariants with values in the Brauer group by computing the étale motivic cohomology group $H^3(U/T, \mathbb{Z}(1)) = H^2(U/T, \mathbb{G}_m) = \mathrm{Br}(U/T)$ for a standard classifying T-torsor $U \to U/T$. One establishes canonical isomorphisms

$$H^1(F, \widehat{T}_{\mathrm{sep}}) \simeq \overline{H}^3(U/T, \mathbb{Z}(1))_{\mathrm{bal}} \simeq \mathrm{Inv}^2(T, \mathbb{Q}/\mathbb{Z}(1))_{\mathrm{norm}} = \mathrm{Inv}(T, \mathrm{Br})_{\mathrm{norm}}.$$

The composition takes an element $a \in H^1(F, \widehat{T}_{\mathrm{sep}})$ to the invariant $b \mapsto a_K \cup b$ for $b \in H^1(K, T)$ and a field extension K/F. This description shows that every normalized Br-invariant of T is linear.

4d. ***Torsion in*** $\mathbf{CH^2(BT)}$**.** We investigate the group $\mathrm{CH}^2(BT)_{\mathrm{tors}}$, the first term of the exact sequence in Theorem 4.3.

Let S be an algebraic torus over F. Using the Gersten resolution, [Quillen 1973, Proposition 5.8] we identify the group $H^0(S_{\mathrm{sep}}, \mathcal{K}_2)$ with a subgroup in $K_2(F_{\mathrm{sep}}(S))$. Set $\overline{H}^0(S_{\mathrm{sep}}, \mathcal{K}_2) := H^0(S_{\mathrm{sep}}, \mathcal{K}_2)/K_2(F_{\mathrm{sep}})$. By [Garibaldi et al. 2003, Part 2, §5.7], we have an exact sequence

$$0 \to \widehat{S}_{\mathrm{sep}} \otimes F_{\mathrm{sep}}^{\times} \to \overline{H}^0(S_{\mathrm{sep}}, \mathcal{K}_2) \xrightarrow{\lambda} \textstyle\bigwedge^2 \widehat{S}_{\mathrm{sep}} \to 0 \tag{4-11}$$

of Γ-modules, where $\lambda(\{e^x, e^y\}) = x \wedge y$ for $x, y \in \widehat{S}_{\mathrm{sep}}$.

Consider the Γ-homomorphism

$$\begin{aligned} \gamma : \textstyle\bigwedge^2 \widehat{S}_{\mathrm{sep}} &\to \overline{H}^0(S_{\mathrm{sep}}, \mathcal{K}_2) \\ x \wedge y &\mapsto \{e^x, e^y\} - \{e^y, e^x\}. \end{aligned}$$

We have $\lambda \circ \gamma = 2 \cdot \mathrm{Id}$, hence the connecting homomorphism

$$\partial : H^i(F, \textstyle\bigwedge^2 \widehat{S}_{\mathrm{sep}}) \to H^{i+1}(F, \widehat{S}_{\mathrm{sep}} \otimes F_{\mathrm{sep}}^{\times}) \tag{4-12}$$

satisfies $2\partial = 0$.

Lemma 4.5. *If S is an invertible torus, then the sequence of Γ-modules (4-11) is split.*

Proof. Suppose first that S is quasisplit. Let $\{x_1, x_2, \dots, x_m\}$ be a permutation basis for $\widehat{S}_{\text{sep}}$. Then the elements $x_i \wedge x_j$ for $i < j$ form a $\mathbb{Z}$-basis for $\bigwedge^2 \widehat{S}_{\text{sep}}$. The map $\bigwedge^2 \widehat{S}_{\text{sep}} \to \overline{H}^0(S_{\text{sep}}, \mathcal{K}_2)$, taking $x_i \wedge x_j$ to $\{e^{x_i}, e^{x_j}\}$ is a splitting for γ.

In general, find a torus S' such that $S \times S'$ is quasisplit. The desired splitting is the composition

$$\textstyle\bigwedge^2 \widehat{S}_{\text{sep}} \to \bigwedge^2 \widehat{S_{\text{sep}} \times S'_{\text{sep}}} \xrightarrow{\alpha} \overline{H}^0(S_{\text{sep}} \times S'_{\text{sep}}, \mathcal{K}_2) \xrightarrow{\beta} \overline{H}^0(S_{\text{sep}}, \mathcal{K}_2),$$

where α is a splitting for the torus $S \times S'$ and β is the pull-back map for the canonical inclusion $S \hookrightarrow S \times S'$. $\square$

Let

$$1 \to T \to P \to Q \to 1$$

be a coflasque resolution of T, that is, P is a quasisplit torus and Q is a coflasque torus; see [Colliot-Thélène and Sansuc 1977]. The torus P is an open set in the affine space of a separable F-algebra on which T acts linearly. Hence $P \to Q$ is a standard classifying T-torsor. By Theorem 2.2, the natural map

$$\theta_T : \operatorname{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2)) \to H^3(F(Q), \mathbb{Q}/\mathbb{Z}(2))$$

is injective.

Consider the spectral sequence (B-10) for the variety $X = Q$. By [Garibaldi et al. 2003, Part 2, Corollary 5.6], we have $H^1(Q_{\text{sep}}, \mathcal{K}_2) = 0$. In view of Proposition B.4, we have an injective homomorphism

$$\beta : H^2(F, \overline{H}^0(Q_{\text{sep}}, \mathcal{K}_2)) \to \overline{H}^4(Q, \mathbb{Z}(2)) \tag{4-13}$$

such that the composition of β with the homomorphism

$$H^2(F, Q^\circ) \to H^2(F, \overline{H}^0(Q_{\text{sep}}, \mathcal{K}_2))$$

is given by the cup-product with the class of the identity in $H^0(Q, Q)$.

Lemma 4.6. *For a coflasque torus Q, the group* $\mathrm{CH}^2(Q)$ *is trivial.*

Proof. By [Merkurjev and Panin 1997, Theorem 9.1], for every torus Q, the Grothendieck group $K_0(Q)$ is generated by the classes of the sheaves $i_*(P)$, where P is an invertible sheaf on Q_L, L/F a finite separable field extension and $i : Q_L \to Q$ is the natural morphism. By definition of a coflasque torus,

$$\operatorname{Pic}(Q_L) = H^1(L, \widehat{Q}_{\text{sep}}) = 0.$$

It follows that every invertible sheaf on Q_L is trivial, hence $K_0(Q) = \mathbb{Z} \cdot 1$. Since the group $\mathrm{CH}^2(Q)$ is generated by the second Chern classes of vector bundles on Q [Esnault et al. 1998, Lemma C.3], we have $\mathrm{CH}^2(Q) = 0$. $\square$

It follows from Proposition A.10, Theorem 4.1, and Lemma 4.6 that the homomorphism

$$\kappa : \overline{H}^4(Q, \mathbb{Z}(2)) \to \overline{H}^4(F(Q), \mathbb{Z}(2)) = \overline{H}^3(F(Q), \mathbb{Q}/\mathbb{Z}(2)) \tag{4-14}$$

is injective.

Consider the diagram

$$\begin{tikzcd}
0 \arrow[r] & Q^\circ(F_{\mathrm{sep}}) \arrow[r] \arrow[d, equal] & \overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2) \arrow[r] \arrow[d, "s"'] & \bigwedge^2 \widehat{Q}_{\mathrm{sep}} \arrow[r] \arrow[d, "t"'] & 0 \\
0 \arrow[r] & Q^\circ(F_{\mathrm{sep}}) \arrow[r] & P^\circ(F_{\mathrm{sep}}) \arrow[r] & T^\circ(F_{\mathrm{sep}}) \arrow[r] & 0
\end{tikzcd}$$

where s is the composition of the natural map $\overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2) \to \overline{H}^0(P_{\mathrm{sep}}, \mathcal{K}_2)$ and a splitting of $P^\circ(F_{\mathrm{sep}}) \to \overline{H}^0(P_{\mathrm{sep}}, \mathcal{K}_2)$ (see Lemma 4.5).

We have the following diagram

$$\begin{tikzcd}
 & & H^1(F, T^\circ) \arrow[r, "\alpha"] \arrow[d, hook, "\partial_1"] & \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \arrow[d, hook, "\theta_T"] \\
H^1(F, \overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2)) \arrow[r, "\varphi"] & H^1(F, \bigwedge^2 \widehat{Q}_{\mathrm{sep}}) \arrow[ur, "t^*"] \arrow[r, "\partial"] & H^2(F, Q^\circ) \arrow[r, "\sigma"] & \overline{H}^3(F(Q), \mathbb{Q}/\mathbb{Z}(2))
\end{tikzcd}$$

with the bottom sequence a complex, where σ is the composition of the maps in (4-13) and (4-14):

$$\begin{aligned}
H^2(F, Q^0) \xrightarrow{\psi} H^2(F, \overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2)) \xrightarrow{\beta} \overline{H}^4(Q, \mathbb{Z}(2)) \\
\xrightarrow{\kappa} \overline{H}^4(F(Q), \mathbb{Z}(2)) = \overline{H}^3(F(Q), \mathbb{Q}/\mathbb{Z}(2)),
\end{aligned}$$

with φ and ψ given by Galois cohomology applied to the exact sequence (4-11) for the torus Q. Note that the connecting map ∂_1 is injective as $H^1(F, P^\circ) = 0$ since P° is a quasisplit torus. As $2\partial = 0$ in (4-12), we have $2t^* = 0$.

The commutativity of the triangle follows from the definition of t^*. We claim that the square in the diagram is anticommutative. Note that $\partial_2(\xi) = [P_{\mathrm{gen}}]$, where $\partial_2 : H^0(F, Q) \to H^1(F, T)$ is the connecting homomorphism, P_{gen} is the generic fiber of the morphism $P \to Q$, and $\xi \in H^0(K, Q)$ is the generic point of Q with $K = F(Q)$. It follows from the description of the maps α and β in (4-7) and (4-13), respectively, and Lemma A.1 that

$$\sigma(\partial_1(a)) = \partial_1(a)_K \cup \xi = (-a_K) \cup \partial_2(\xi) = (-a_K) \cup [P_{\mathrm{gen}}] = -\theta_T(\alpha(a))$$

for every $a \in H^1(F, T^\circ)$.

The maps β and κ are injective, hence the bottom sequence in the diagram is exact. Thus, we have an exact sequence

$$H^1(F, \overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2)) \to H^1(F, \bigwedge^2 \widehat{Q}_{\mathrm{sep}}) \to \mathrm{Ker}(\alpha) \to 0$$

and $2\cdot\mathrm{Ker}(\alpha) = 2\cdot\mathrm{Im}(t^*) = 0$. Furthermore, $\mathrm{Ker}(\alpha) \simeq \mathrm{CH}^2(BT)_{\mathrm{tors}}$ by Theorem 4.3 and the group $H^1(F, \bigwedge^2 \widehat{Q}_{\mathrm{sep}})$ is finite as $\bigwedge^2 \widehat{Q}_{\mathrm{sep}}$ is a lattice.

We have proved:

Theorem 4.7. *Let* $1 \to T \to P \to Q \to 1$ *be a coflasque resolution of a torus* T. *Then there is an exact sequence*

$$H^1(F, \overline{H}^0(Q_{\mathrm{sep}}, \mathcal{K}_2)) \to H^1(F, \bigwedge^2 \widehat{Q}_{\mathrm{sep}}) \to \mathrm{CH}^2(BT)_{\mathrm{tors}} \to 0.$$

Moreover, $\mathrm{CH}^2(BT)_{\mathrm{tors}}$ *is a finite group satisfying* $2 \cdot \mathrm{CH}^2(BT)_{\mathrm{tors}} = 0$.

Corollary 4.8. *If* T° *is a birational direct factor of a rational torus*, *or if* T *is split over a cyclic field extension*, *then* $\mathrm{CH}^2(BT)_{\mathrm{tors}} = 0$, *that is*, *the map* α *in Theorem 4.3 is injective.*

Proof. The exact sequence $1 \to Q^\circ \to P^\circ \to T^\circ \to 1$ is a flasque resolution of T°. If T° is a birational direct factor of a rational torus, or if T is split over a cyclic field extension, the torus Q°, and hence Q, is invertible; see Section 4a and [Voskresenskiĭ 1998, §4, Theorem 3]. By Lemma 4.5, the sequence (4-11) for the torus Q is split, hence the first map in Theorem 4.7 is surjective. □

Question 4.9. Is $\mathrm{CH}^2(BT)_{\mathrm{tors}}$ trivial for every torus T?

4e. *Special tori.* Let T be an algebraic torus over a field F. The *tautological invariant* of the torus $T^\circ \times T$ is the normalized invariant taking a pair

$$(a, b) \in H^1(K, T^\circ) \times H^1(K, T)$$

to the cup-product $a \cup b \in H^3(K, \mathbb{Q}/\mathbb{Z}(2))$ defined in (4-5).

The following theorem shows that if a torus T has only trivial degree 3 normalized invariants with values in $\mathbb{Q}/\mathbb{Z}(2)$ universally, that is, over all field extensions of F, then T has no nonconstant invariants at all by the simple reason: every T-torsor over a field is trivial. Note that it follows from Theorem 2.4 that T has no degree 2 normalized invariants with values in $\mathbb{Q}/\mathbb{Z}(1)$ universally if and only if T is coflasque.

Theorem 4.10. *Let* T *be an algebraic torus over a field* F. *Then the following are equivalent*:

(1) $\mathrm{Inv}^3(T_K, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} = 0$ *for every field extension* K *of* F.

(2) *The tautological invariant of the torus* $T^\circ \times T$ *is trivial.*

(3) *The torus* T *is invertible.*

(4) *The torus* T *is special.*

Proof. (1) ⇒ (2): Let K/F be a field extension and $a \in H^1(K, T^\circ)$. By assumption, the degree 3 normalized invariant $i = \alpha(a)$ with values in $\mathbb{Q}/\mathbb{Z}(2)$, defined by $i(b) = a \cup b$ for every $b \in H^1(K, T)$, is trivial. In other words, the tautological invariant of the torus $T^\circ \times T$ is trivial.

(2) ⇒ (3): The image of the tautological invariant in the group

$$S^2(\widehat{T}^\circ_{\mathrm{sep}} \oplus \widehat{T}_{\mathrm{sep}})^\Gamma / \mathrm{Dec}$$

is represented by the identity $1_{\widehat{T}}$ in the direct factor $(\widehat{T}^\circ_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}})^\Gamma = \mathrm{End}_\Gamma(\widehat{T}_{\mathrm{sep}})$ of $S^2(\widehat{T}^\circ_{\mathrm{sep}} \oplus \widehat{T}_{\mathrm{sep}})^\Gamma$ (see Section A-II). The projection of Dec on the direct summand $(\widehat{T}^\circ_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}})^\Gamma$ is generated by the traces $\mathrm{Tr}(a \otimes b)$ for all open subgroups $\Gamma' \subset \Gamma$ and all $a \in (\widehat{T}^\circ_{\mathrm{sep}})^{\Gamma'}$ and $b \in (\widehat{T}_{\mathrm{sep}})^{\Gamma'}$. Hence $1_{\widehat{T}} = \sum_i \mathrm{Tr}(a_i \otimes b_i)$ for some open subgroups $\Gamma_i \subset \Gamma$, $a_i \in (\widehat{T}^\circ)^{\Gamma_i}$ and $b_i \in (\widehat{T})^{\Gamma_i}$. If $P_i = \mathbb{Z}[\Gamma/\Gamma_i]$, then a_i can be viewed as a Γ-homomorphism $\widehat{T} \to P_i$ and b_i can be viewed as a Γ-homomorphism $P_i \to \widehat{T}$ such that the composition

$$\widehat{T} \xrightarrow{(b_i)} P \xrightarrow{(a_i)} \widehat{T},$$

where $P = \coprod P_i$, is the identity. It follows that $\widehat{T}$ is a direct summand of a permutation Γ-module P and hence T is invertible.

(3) ⇒ (4): Obvious as every invertible torus is special.

(4) ⇒ (1): Obvious. □

Remark 4.11. The equivalence (3) ⇔ (4) was essentially proved in [Colliot-Thélène and Sansuc 1987b, Proposition 7.4].

4f. ***Linear and quadratic invariants.*** Let T be a torus over F. By Theorem 4.3, we have a natural homomorphism to the group of linear invariants:

$$\alpha : H^1(F, T^\circ) \to \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{lin}}.$$

Let S and T be algebraic tori over F. For every field extension K/F, the cup-product (4-2) yields a homomorphism

$$\varepsilon : (\widehat{T}^{\otimes 2}_{\mathrm{sep}})^\Gamma \to \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))$$

defined by $\varepsilon(a)(b) = a_K \cup b \cup b$ for $a \in (\widehat{T}^{\otimes 2}_{\mathrm{sep}})^\Gamma$ and $b \in H^1(K, T)$.

We say that an invariant $i \in \mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))$ is *quadratic* if the function

$$h(a, b) := i(a + b) - i(a) - i(b)$$

is bilinear and $h(a, a) = 2i(a)$ for all a and b. For example, the tautological invariant of the torus $T^\circ \times T$ in Section 4e is quadratic. We write $\mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{quad}}$ for the subgroup of all quadratic invariants in $\mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))$. The image of ε is contained in $\mathrm{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{quad}}$.

Lemma 4.12. *The composition of ε with the map*

$$\operatorname{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2)) \to S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec}$$

in Theorem 4.3 is induced by the natural homomorphism $\widehat{T}_{\mathrm{sep}}^{\otimes 2} \to S^2(\widehat{T}_{\mathrm{sep}})$.

Proof. Let $U \to U/T =: X$ be a standard classifying T-torsor as in Section 4c. Consider the commutative diagram

$$\begin{tikzcd}
(\widehat{T}_{\mathrm{sep}}^{\otimes 2})^\Gamma \otimes H^1(X, T)^{\otimes 2} \arrow[r, "\mathrm{prod}"] \arrow[d] & \overline{H}^4(X, \mathbb{Z}(2))_{\mathrm{bal}} \arrow[r] \arrow[d] & \operatorname{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \\
\widehat{T}_{\mathrm{sep}}^{\otimes 2} \otimes H^1(X_{\mathrm{sep}}, T)^{\otimes 2} \arrow[r, "\mathrm{prod}"] \arrow[d, "\eta"', "\wr"] & \overline{H}^4(X_{\mathrm{sep}}, \mathbb{Z}(2))_{\mathrm{bal}} \arrow[d, "\wr"] & \\
\widehat{T}_{\mathrm{sep}}^{\otimes 2} \otimes (\widehat{T}_{\mathrm{sep}}^{\circ})^{\otimes 2} \otimes \widehat{T}_{\mathrm{sep}}^{\otimes 2} \arrow[r, "\kappa"] & S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec} &
\end{tikzcd}$$

where the product maps are given by (4-1), η identifies

$$H^1(X_{\mathrm{sep}}, T) = \widehat{T}_{\mathrm{sep}}^{\circ} \otimes \operatorname{Pic}(X_{\mathrm{sep}})$$

with $\widehat{T}_{\mathrm{sep}}^{\circ} \otimes \widehat{T}_{\mathrm{sep}}$ and κ is given by the pairing between the first and second factors. Write $[U]$ for the class of the classifying torsor in $H^1(X, T)$. The image of $[U]$ in

$$H^1(X_{\mathrm{sep}}, T_{\mathrm{sep}}) = \widehat{T}_{\mathrm{sep}}^{\circ} \otimes \widehat{T}_{\mathrm{sep}} = \operatorname{End}(\widehat{T}_{\mathrm{sep}})$$

is the identity $1_{\widehat{T}_{\mathrm{sep}}}$. Hence for every $a \in (\widehat{T}_{\mathrm{sep}}^{\otimes 2})^\Gamma$, the image of $a \otimes [U] \otimes [U]$ under the diagonal map in the diagram coincides with the canonical image of a in $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec}$. □

The composition of the map $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma \to (\widehat{T}_{\mathrm{sep}}^{\otimes 2})^\Gamma$ given by $a \cdot b \mapsto a \otimes b + b \otimes a$ with the natural map $(\widehat{T}_{\mathrm{sep}}^{\otimes 2})^\Gamma \to S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$ is multiplication by 2. Then by Lemma 4.12, $2 \cdot S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec}$ is contained in the image of the map

$$\operatorname{Inv}^3(T, \mathbb{Q}/\mathbb{Z}(2)) \to S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec}.$$

Theorem 4.3 then yields:

Theorem 4.13. *Let T be an algebraic torus over F. Then 2 times the homomorphism $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec} \to H^2(F, T^0)$ from Theorem 4.3 is trivial. If p is an odd prime,*

$$\operatorname{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{norm}} = \operatorname{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{lin}} \oplus \operatorname{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{quad}}$$

and there are natural isomorphisms $\operatorname{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{lin}} \simeq H^1(F, T^\circ)\{p\}$ and

$$\operatorname{Inv}^3(T, \mathbb{Q}_p/\mathbb{Z}_p(2))_{\mathrm{quad}} \simeq (S^2(\widehat{T}_{\mathrm{sep}})^\Gamma / \operatorname{Dec})\{p\}.$$

Example 4.14. Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of n elements with the natural action of the symmetric group S_n. A continuous surjective group homomorphism $\Gamma \to S_n$ yields a separable field extension L/F of degree n. Consider the torus $T = R_{L/F}(\mathbb{G}_{m,L})/\mathbb{G}_m$, where $R_{L/F}$ is the Weil restriction; see [Voskresenskiĭ 1998, Chapter 1, §3.12]. Note that the generic maximal torus of the group $\mathbf{PGL}_n$ is of this form (see Section 5b). The character lattice $\widehat{T}_{\mathrm{sep}}$ is the kernel of the augmentation homomorphism $\mathbb{Z}[X] \to \mathbb{Z}$.

The dual torus T° is the norm one torus $R^{(1)}_{L/F}(\mathbb{G}_{m,L})$. For every field extension K/F, we have:

$$H^1(K, T) = \mathrm{Br}(KL/K), \qquad H^1(K, T^\circ) = K^\times/N(KL)^\times,$$

where $KL := K \otimes L$, N is the norm map for the extension KL/K and

$$\mathrm{Br}(KL/K) = \mathrm{Ker}(\mathrm{Br}(K) \to \mathrm{Br}(KL)).$$

The pairing

$$K^\times/N(KL)^\times \otimes \mathrm{Br}(KL/K) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2))$$

defines linear degree 3 invariants of both T and T°.

We claim that $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma/\mathrm{Dec} = 0$ and $S^2(\widehat{T}^\circ_{\mathrm{sep}})^\Gamma/\mathrm{Dec} = 0$, that is, T and T° have no nontrivial quadratic degree 3 invariants. We have $\widehat{T}^\circ_{\mathrm{sep}} = \mathbb{Z}[X]/\mathbb{Z}N_X$, where $N_X = \sum x_i$. The group $S^2(\widehat{T}^\circ_{\mathrm{sep}})^\Gamma$ is generated by $S := \sum_{i<j} x_i \cdot x_j$. As $S \in \mathrm{Dec}$, we have $S^2(\widehat{T}^\circ_{\mathrm{sep}})^\Gamma/\mathrm{Dec} = 0$.

Let $D = \sum x_i^2$ and $E := \mathrm{Qtr}(x_1 - x_2) = 2S - (n-1)D$, where the quadratic map Qtr is defined in Section A-II. The group $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$ is generated by E if n is even and by $E/2$ if n is odd. A computation shows that $nE/2 = \mathrm{Qtr}(nx_1 - N_X)$. It follows that the generator of $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma$ belongs to Dec, hence $S^2(\widehat{T}_{\mathrm{sep}})^\Gamma/\mathrm{Dec}$ is trivial.

Note that as the torus T is rational, it follows from Theorem 4.3 and Corollary 4.8 that $\mathrm{Inv}^3(T^\circ, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} \simeq \mathrm{Br}(L/F)$.

5. Unramified invariants

Let K/F be a field extension and v a discrete valuation of K over F with valuation ring O_v. We say that an element $a \in H^n(K, \mathbb{Q}/\mathbb{Z}(j))$ is *unramified with respect to* v if a belongs to the image of the map $H^n(O_v, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(K, \mathbb{Q}/\mathbb{Z}(j))$; see [Colliot-Thélène and Ojanguren 1989]. We write $H^n_{\mathrm{nr}}(K, \mathbb{Q}/\mathbb{Z}(j))$ for the subgroup of the elements in $H^n(K, \mathbb{Q}/\mathbb{Z}(j))$ that are unramified with respect to all discrete valuations of K over F. We have a natural homomorphism

$$H^n(F, \mathbb{Q}/\mathbb{Z}(j)) \to H^n_{\mathrm{nr}}(K, \mathbb{Q}/\mathbb{Z}(j)). \tag{5-1}$$

A dominant morphism of varieties $Y \to X$ yields a homomorphism

$$H^n_{\mathrm{nr}}(F(X), \mathbb{Q}/\mathbb{Z}(j)) \to H^n_{\mathrm{nr}}(F(Y), \mathbb{Q}/\mathbb{Z}(j)). \tag{5-2}$$

Proposition 5.1. *Let K/F be a purely transcendental field extension. Then the homomorphism (5-1) is an isomorphism.*

Proof. The statement is well known for the p-components if $p \neq \operatorname{char} F$; see, for example, [Colliot-Thélène and Ojanguren 1989, Proposition 1.2]. It suffices to consider the case $K = F(t)$ and prove the surjectivity of (5-1). The coniveau spectral sequence for the projective line $\mathbb{P}^1$ (see (A-1) in the Appendix) yields an exact sequence

$$H^n(\mathbb{P}^1, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(K, \mathbb{Q}/\mathbb{Z}(j)) \to \coprod_{x \in \mathbb{P}^1} H^{n+1}_x(\mathbb{P}^1, \mathbb{Q}/\mathbb{Z}(j))$$

and, therefore, a surjective homomorphism $H^n(\mathbb{P}^1, \mathbb{Q}/\mathbb{Z}(j)) \to H^n_{\mathrm{nr}}(K, \mathbb{Q}/\mathbb{Z}(j))$. By the projective bundle theorem (classical if $p \neq \operatorname{char}(F)$ and [Gros 1985, Theorem 2.1.11] if $p = \operatorname{char}(F) > 0$), we have

$$H^n(\mathbb{P}^1, \mathbb{Q}/\mathbb{Z}(j)) = H^n(F, \mathbb{Q}/\mathbb{Z}(j)) \oplus H^{n-2}(F, \mathbb{Q}/\mathbb{Z}(j-1))t,$$

where t is a generator of $H^2(\mathbb{P}^1, \mathbb{Z}(1)) = \operatorname{Pic}(\mathbb{P}^1) = \mathbb{Z}$. As t vanishes over the generic point of $\mathbb{P}^1$, the result follows. $\square$

Let G be a linear algebraic group over F. Choose a standard classifying G-torsor $U \to U/G$. An invariant $i \in \operatorname{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j))$ is called *unramified* if the image of i under $\theta_G : \operatorname{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(F(U/G), \mathbb{Q}/\mathbb{Z}(j))$ is unramified. This is independent of the choice of standard classifying torsor. Indeed, if $U' \to U'/G$ is another standard classifying G-torsor, then $(U \times V')/G \to U/G$ and $(V \times U')/G \to U'/G$ are vector bundles. Hence the field $F((U \times U')/G)$ is a purely transcendental extension of $F(U/G)$ and $F(U'/G)$ and by Proposition 5.1,

$$\begin{aligned} H^n_{\mathrm{nr}}(F(U/G), \mathbb{Q}/\mathbb{Z}(j)) &\simeq H^n_{\mathrm{nr}}(F((U \times U')/G), \mathbb{Q}/\mathbb{Z}(j)) \\ &\simeq H^n_{\mathrm{nr}}(F(U'/G), \mathbb{Q}/\mathbb{Z}(j)). \end{aligned}$$

We write $H^n_{\mathrm{nr}}(F(BG), \mathbb{Q}/\mathbb{Z}(j))$ for this common value and $\operatorname{Inv}^n_{\mathrm{nr}}(G, \mathbb{Q}/\mathbb{Z}(j))$ for the subgroup of unramified invariants. Similarly, we write $\operatorname{Br}_{\mathrm{nr}}(F(BG))$ for the *unramified Brauer group* $H^2_{\mathrm{nr}}(F(BG), \mathbb{Q}/\mathbb{Z}(1))$.

Proposition 5.2. *If G' be a subgroup of G and $i \in \operatorname{Inv}^n_{\mathrm{nr}}(G, \mathbb{Q}/\mathbb{Z}(j))$, then*

$$\operatorname{res}(i) \in \operatorname{Inv}^n_{\mathrm{nr}}(G', \mathbb{Q}/\mathbb{Z}(j)).$$

Proof. It is shown in Section 2b that there is a surjective morphism $X' \to X$ of the respective classifying varieties of G' and G, such that $\theta_G(i)_{F(X')} = \theta_{G'}(\mathrm{res}(i))$. Applying the homomorphism (5-2) we see that $\mathrm{res}(i)$ is unramified. $\square$

Proposition 5.3. *Let G be a smooth linear algebraic group over a field F. The map* $\mathrm{Inv}^n_{\mathrm{nr}}(G, \mathbb{Q}/\mathbb{Z}(j)) \to H^n_{\mathrm{nr}}(F(BG), \mathbb{Q}/\mathbb{Z}(j))$ *induced by θ_G is an isomorphism.*

Proof. By Theorem 3.4, it suffices to show that

$$H^n_{\mathrm{nr}}(F(U/G), \mathbb{Q}/\mathbb{Z}(j)) \subset H^n(F(U/G), \mathbb{Q}/\mathbb{Z}(j))_{\mathrm{bal}}.$$

We follow Totaro's approach; see [Garibaldi et al. 2003, p. 99]. Consider the open subscheme W of $(U^2/G)\times\mathbb{A}^1$ of all triples (u, u', t) such that $(2-t)u+(t-1)u' \in U$. We have the projection $q : W \to U^2/G$, the morphisms $f : W \to U/G$ defined by $f(u, u', t) = (2-t)u+(t-1)u'$, and $h_i : U^2/G \to W$ defined by $h_i(u, u') = (u, u', i)$ for $i = 1$ and 2. The composition $f \circ h_i$ is the projection $p_i : U^2/G \to U/G$ and $q \circ h_i$ is the identity of U^2/G.

Let w_i be the generic point of the preimage of i with respect to the projection $W \to \mathbb{A}^1$ and write O_i for the local ring of W at w_i. The morphisms q, f, and h_i yield F-algebra homomorphisms $F(U^2/G) \to O_i$, $F(U/G) \to O_i$ and $O_i \to F(U/G)$. Note that by Proposition A.11, we have

$$H^n_{\mathrm{nr}}(F(W), \mathbb{Q}/\mathbb{Z}(j)) \subset H^n(O_i, \mathbb{Q}/\mathbb{Z}(j)).$$

In the commutative diagram

$$\begin{tikzcd}
H^n_{\mathrm{nr}}(F(U/G), \mathbb{Q}/\mathbb{Z}(j)) \arrow[r, "f^*"] \arrow[d, hook] & H^n_{\mathrm{nr}}(F(W), \mathbb{Q}/\mathbb{Z}(j)) \arrow[d, hook] & H^n_{\mathrm{nr}}(F(U^2/G), \mathbb{Q}/\mathbb{Z}(j)) \arrow[l, "q^*"', "\sim"] \arrow[d, hook] \\
H^n(F(U/G), \mathbb{Q}/\mathbb{Z}(j)) \arrow[r, "f^*"] \arrow[dr, "p_i^*"'] & H^n(O_i, \mathbb{Q}/\mathbb{Z}(j)) \arrow[d, "h_i^*"] & H^n(F(U^2/G), \mathbb{Q}/\mathbb{Z}(j)) \arrow[l, "q^*"'] \arrow[dl, equal] \\
& H^n(F(U^2/G), \mathbb{Q}/\mathbb{Z}(j)) &
\end{tikzcd}$$

the top right map q^* is an isomorphism by Proposition 5.1 since the field extension $F(W)/F(U^2/G)$ is purely transcendental. It follows that the restriction of p_i^* on $H^n_{\mathrm{nr}}(F(U/G), \mathbb{Q}/\mathbb{Z}(j))$ coincides with $(q^*)^{-1} \circ f^*$ and hence is independent of i. $\square$

5a. *Unramified invariants of tori.*

Proposition 5.4. *If T is a flasque torus, then every invariant in* $\mathrm{Inv}^n(T, \mathbb{Q}/\mathbb{Z}(j))$ *is unramified.*

Proof. Consider an exact sequence of tori $1 \to T \to P \to Q \to 1$ with P quasisplit. Choose a smooth projective compactification X of Q; see [Colliot-Thélène et al. 2005]. As T is flasque, by [Colliot-Thélène and Sansuc 1977, Proposition 9], there is a T-torsor $E \to X$ extending the T-torsor $P \to Q$. The torsor E is classifying and T-rational. Choose an invariant in $\mathrm{Inv}^n(G, \mathbb{Q}/\mathbb{Z}(j))$ and consider its image a in $H^n(F(X), \mathbb{Q}/\mathbb{Z}(j))_{\mathrm{bal}}$ (see Theorem 3.4). We show that a is unramified with respect to every discrete valuation v on $F(X)$ over F; see [Colliot-Thélène 1995, Proposition 2.1.8]. By Proposition A.9, a is unramified with respect to the discrete valuation associated to every point $x \in X$ of codimension 1, that is, $\partial_x(a) = 0$.

As X is projective, the valuation ring O_v of the valuation v dominates a point $x \in X$. It follows from Proposition A.11 that a belongs to the image of

$$H^n(O_{X,x}, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(F(X), \mathbb{Q}/\mathbb{Z}(j)).$$

As the local ring $O_{X,x}$ is a subring of O_v, a belongs to the image of

$$H^n(O_v, \mathbb{Q}/\mathbb{Z}(j)) \to H^n(F(X), \mathbb{Q}/\mathbb{Z}(j))$$

and hence a is unramified with respect to v. □

Let T be a torus over F. By [Colliot-Thélène and Sansuc 1987b, Lemma 0.6], there is an exact sequence of tori $1 \to T \to T' \to P \to 1$ with T' flasque and P quasisplit. The following theorem computes the unramified invariants of T in terms of the invariants of T'.

Theorem 5.5. *There is a natural isomorphism*

$$\mathrm{Inv}^n_{\mathrm{nr}}(T, \mathbb{Q}/\mathbb{Z}(j)) \simeq \mathrm{Inv}^n(T', \mathbb{Q}/\mathbb{Z}(j)).$$

Proof. Choose an exact sequence $1 \to T' \to P' \to S \to 1$ with P' a quasisplit torus. Let S' be the cokernel of the composition $T \to T' \to P'$. We have an exact sequence $1 \to P \to S' \to S \to 1$. As P is quasisplit, the latter exact sequence splits at the generic point of S and therefore, $F(S')$ is a purely transcendental field extension of $F(S)$. It follows from Propositions 5.1, 5.3, and 5.4 that

$$\begin{aligned}\mathrm{Inv}^n_{\mathrm{nr}}(T, \mathbb{Q}/\mathbb{Z}(j)) \simeq H^n_{\mathrm{nr}}(F(S'), \mathbb{Q}/\mathbb{Z}(j)) \simeq H^n_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(j))\\ \simeq \mathrm{Inv}^n_{\mathrm{nr}}(T', \mathbb{Q}/\mathbb{Z}(j)) = \mathrm{Inv}^n(T', \mathbb{Q}/\mathbb{Z}(j)).\end{aligned}$$ □

The following corollary is essentially equivalent to [Colliot-Thélène and Sansuc 1987b, Proposition 9.5] in the case when F is of zero characteristic.

Corollary 5.6. *With notation as above, the isomorphism*

$$\mathrm{Inv}(T, \mathrm{Br}) \xrightarrow{\sim} \mathrm{Pic}(T) = H^1(F, \widehat{T})$$

identifies $\mathrm{Inv}_{\mathrm{nr}}(T, \mathrm{Br})$ *with the subgroup* $H^1(F, \widehat{T}')$ *of* $H^1(F, \widehat{T})$ *of all elements that are trivial when restricted to all cyclic subgroups of the decomposition group of* T.

Proof. The description of $H^1(F, \widehat{T}')$ as a subgroup of $H^1(F, \widehat{T})$ is given in [Colliot-Thélène and Sansuc 1987b, Proposition 9.5], and this part of the proof is characteristic free. □

In view of Propositions 5.1 and 5.3 we can calculate the group of unramified cohomology for the function field of an arbitrary torus in terms of the invariants of a flasque torus:

Theorem 5.7. *Let* S *be a torus over* F *and let* $1 \to T \to P \to S \to 1$ *be a flasque resolution of* S, *that is*, T *is flasque and* P *is quasisplit. Then there is a natural isomorphism*

$$H^n_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(j)) \simeq \mathrm{Inv}^n(T, \mathbb{Q}/\mathbb{Z}(j)).$$

Note that the torus S determines T up to multiplication by a quasisplit torus. If X is a smooth compactification of S, then one can take a torus T with $\widehat{T}_{\mathrm{sep}} \simeq \mathrm{Pic}(X_{\mathrm{sep}})$; see [Colliot-Thélène and Sansuc 1977, Proposition 6; Voskresenskiĭ 1998, §4.6].

Corollary 5.8. *A torus* S *has no nonconstant unramified degree* 3 *cohomology with values in* $\mathbb{Q}/\mathbb{Z}(2)$ *universally, that is*, $H^3_{\mathrm{nr}}(K(S), \mathbb{Q}/\mathbb{Z}(2)) = H^3(K, \mathbb{Q}/\mathbb{Z}(2))$ *for any field extension* K/F, *if and only if* S *is a direct factor of a rational torus.*

Proof. If S is a direct factor of a rational torus, then S has no nonconstant unramified cohomology by Proposition 5.1.

Conversely, let $1 \to T \to P \to S \to 1$ be a flasque resolution of S. By Theorem 5.7, $\mathrm{Inv}^3(T_K, \mathbb{Q}/\mathbb{Z}(2))_{\mathrm{norm}} = 0$ for every K/F. It follows from Theorem 4.10 that T is invertible and hence S is a factor of a rational torus (see Section 4a). □

Theorems 4.3, 5.7 and [Colliot-Thélène and Sansuc 1977, §2] yield the following proposition.

Proposition 5.9. *Let* S *be a torus over* F *and let* $1 \to T \to P \to S \to 1$ *be a flasque resolution of* S. *Then we have an exact sequence*

$$0 \to \mathrm{CH}^2(BT)_{\mathrm{tors}} \to H^1(F, T^0) \to \overline{H}^3_{\mathrm{nr}}(F(S), \mathbb{Q}/\mathbb{Z}(2)) \\ \to H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\mathrm{Dec} \to H^2(F, T^0).$$

For an odd prime p, *there is a canonical direct sum decomposition*

$$\overline{H}^3_{\mathrm{nr}}(F(S), \mathbb{Q}_p/\mathbb{Z}_p(2)) = H^1(F, T^0)\{p\} \oplus (H^0(F, S^2(\widehat{T}_{\mathrm{sep}}))/\mathrm{Dec})\{p\}.$$

If X *is a smooth compactification of* S, *one can take the torus* T *with* $\widehat{T}_{\mathrm{sep}} = \mathrm{Pic}(X_{\mathrm{sep}})$.

5b. ***The Brauer invariant for semisimple groups.*** The following theorem was proved by Bogomolov [1987, Lemma 5.7] in characteristic zero:

Theorem 5.10. *Let G be a (connected) semisimple group over a field F. Then* $\operatorname{Inv}_{\mathrm{nr}}(G, \mathrm{Br}) = \operatorname{Inv}(G, \mathrm{Br})_{\mathrm{const}} = \mathrm{Br}(F)$ *and* $\mathrm{Br}_{\mathrm{nr}}(F(BG)) = \mathrm{Br}(F)$.

Proof. Let $G' \to G$ be a simply connected cover of G and C the kernel of $G' \to G$. By Theorem 2.4 and [Sansuc 1981, Lemme 6.9(iii)], we have

$$\operatorname{Inv}(G, \mathrm{Br})_{\mathrm{norm}} = \operatorname{Pic}(G) = \widehat{C}(F).$$

As the map $\widehat{C}(F) \to \widehat{C}(F_{\mathrm{sep}})$ is injective, we can replace F by F_{sep} and assume that the group G is split.

Consider the variety $\mathcal{T}$ of maximal tori in G and the closed subscheme $\mathcal{X} \subset G \times \mathcal{T}$ of all pairs (g, T) with $g \in T$. The generic fiber of the projection $\mathcal{X} \to \mathcal{T}$ is the *generic torus* T_{gen} of G. Then T_{gen} is a maximal torus of G_K, where $K := F(\mathcal{T})$.

Every maximal torus in G is the factor torus of a maximal torus in G' by C. It follows that the variety $\mathcal{T}'$ of maximal tori in G' is naturally isomorphic to $\mathcal{T}$. Moreover, as the generic torus T'_{gen} of G' is a maximal torus of G'_K, we have $T_{\mathrm{gen}} \simeq T'_{\mathrm{gen}}/C_K$ and, therefore, an exact sequence of character groups

$$0 \to \widehat{T}_{\mathrm{gen}} \to \widehat{T}'_{\mathrm{gen}} \to \widehat{C}_K \to 0.$$

By Theorem 2.4, the composition of the natural homomorphism

$$\operatorname{Inv}(G, \mathrm{Br})_{\mathrm{norm}} \to \operatorname{Inv}(G_K, \mathrm{Br})_{\mathrm{norm}}$$

with the restriction $\operatorname{Inv}(G_K, \mathrm{Br})_{\mathrm{norm}} \to \operatorname{Inv}(T_{\mathrm{gen}}, \mathrm{Br})_{\mathrm{norm}}$ can be identified with the natural composition $\operatorname{Pic}(G) \to \operatorname{Pic}(G_K) \to \operatorname{Pic}(T_{\mathrm{gen}})$ and hence with the connecting homomorphism $\widehat{C}(F) = \widehat{C}(K) \to H^1(K, \widehat{T}_{\mathrm{gen}})$. Note that as $F = F_{\mathrm{sep}}$, the decomposition group of T_{gen} coincides with the Weyl group W of G by [Voskresenskiĭ 1988, Theorem 1], hence $H^1(K, \widehat{T}_{\mathrm{gen}}) \simeq H^1(W, \widehat{T}_{\mathrm{gen}})$.

Let w be a Coxeter element in W.[1] It is the product of reflections with respect to all simple roots (in some order). By [Humphreys 1990, Lemma, p. 76], 1 is not an eigenvalue of w on the space of weights $\widehat{T'}_{\mathrm{gen}} \otimes \mathbb{R}$. Let W_0 be the cyclic subgroup in W generated by w. It follows that the first term in the exact sequence

$$(\widehat{T'}_{\mathrm{gen}})^{W_0} \to \widehat{C}(K) \to H^1(W_0, \widehat{T}_{\mathrm{gen}})$$

is trivial, that is, the second map is injective. Hence every nonzero character χ in $\widehat{C}(K)$ restricts to a nonzero element in $H^1(W_0, \widehat{T}_{\mathrm{gen}})$. It follows that the image of χ in $H^1(W, \widehat{T}_{\mathrm{gen}})$ is ramified by Corollary 5.6, hence so is χ by Proposition 5.2. □

[1] We owe the idea to use the Coxeter element and the reference below to S. Garibaldi.

Appendix A: Generalities

A-I: *Proof of Theorem 2.2.* Suppose that $i(E_{\text{gen}}) = 0$ for an H-invariant i of G. Let K/F be a field extension and $I \to \operatorname{Spec} K$ a G-torsor. We need to show that $i(I) = 0$ in $H(K)$.

Suppose first that K is infinite. Find a point $x \in X(K)$ such that I is isomorphic to the pull-back of the classifying torsor with respect to x. Let x' be a rational point of X_K above x and write O for the local ring $O_{X_K,x'}$. The K-algebra O is a regular local ring with residue field K. Therefore, the completion $\widehat{O}$ is isomorphic to $K[[t_1, t_2, \dots, t_n]]$ over K. Let L be the quotient field of $\widehat{O}$, a field extension of $K(X)$. We have the following diagram of morphisms with a commutative square and three triangles:

$$\begin{tikzcd}
& \operatorname{Spec} K \arrow[dr, "x"] \arrow[d, shift left] & \\
\operatorname{Spec} L \arrow[ur] \arrow[r] \arrow[d] & \operatorname{Spec} \widehat{O} \arrow[u, shift left] \arrow[r] \arrow[d] & X \\
\operatorname{Spec} K(X) \arrow[r] & \operatorname{Spec} O \arrow[ur] &
\end{tikzcd}$$

The pull-back of the classifying torsor $E \to X$ via $\operatorname{Spec} K(X) \to X$ is $(E_{\text{gen}})_{K(X)}$. The G-torsor I is the pull-back of $E \to X$ with respect to x. Let $\widehat{E}$ be the pull-back of $E \to X$ via $\operatorname{Spec} \widehat{O} \to X$. Therefore, I is the pull-back of $\widehat{E}$. Since G is smooth, by a theorem of Grothendieck [Demazure and Grothendieck 1970, XXIV, Proposition 8.1], $\widehat{E}$ is the pull-back of I with respect to $\operatorname{Spec} \widehat{O} \to \operatorname{Spec}(K)$. It follows that $I_L \simeq (E_{\text{gen}})_L$ as torsors over L. Hence the images of $i(I)$ and $i(E_{\text{gen}})$ in $H(L)$ are equal and therefore, $i(I)_L = 0$. By Property 2.1, we have $i(I) = 0$.

If K is finite, we replace F by $F((t))$ and K by $K((t))$. By the first part of the proof, $i(I)$ belongs to the kernel of $H(K) \to H(K((t)))$ and hence is trivial by Property 2.1 again.

A-II: *Decomposable elements.* Let Γ be a profinite group and A a Γ-lattice. Write A^Γ for the subgroup of Γ-invariant elements in A. Let $\Gamma' \subset \Gamma$ be an open subgroup and choose representatives $\gamma_1, \gamma_2, \dots, \gamma_n$ for the left cosets of Γ' in Γ. We have the *trace map* $\operatorname{Tr} : A^{\Gamma'} \to A^\Gamma$ defined by $\operatorname{Tr}(a) = \sum_{i=1}^n \gamma_i a$.

Let $S^2(A)$ be the symmetric square of A. Consider the *quadratic trace map* $\operatorname{Qtr} : A^{\Gamma'} \to S^2(A)^\Gamma$ defined by $\operatorname{Qtr}(a) = \sum_{i<j} (\gamma_i a)(\gamma_j a)$. Write $\operatorname{Dec}(A)$ for the subgroup of *decomposable elements* in $S^2(A)^\Gamma$ generated by the square $(A^\Gamma)^2$ of A^Γ and the elements $\operatorname{Qtr}(a)$ for all open subgroups $\Gamma' \subset \Gamma$ and all $a \in A^{\Gamma'}$.

Let B be another Γ-lattice. We write $\mathrm{Dec}(A, B)$ for the subgroup of $(A \otimes B)^{\Gamma}$ generated by elements of the form $\mathrm{Tr}(a \otimes b)$ for all open subgroups $\Gamma' \subset \Gamma$ and all $a \in A^{\Gamma'}$, $b \in B^{\Gamma'}$.

There is a natural isomorphism $S^2(A \oplus B) \simeq S^2(A) \oplus (A \otimes B) \oplus S^2(B)$. Moreover, the equality $\mathrm{Qtr}(a+b) = \mathrm{Qtr}(a) + (\mathrm{Tr}(a) \otimes \mathrm{Tr}(b) - \mathrm{Tr}(a \otimes b)) + \mathrm{Qtr}(b)$ yields the decomposition

$$\mathrm{Dec}(A \oplus B) \simeq \mathrm{Dec}(A) \oplus \mathrm{Dec}(A, B) \oplus \mathrm{Dec}(B).$$

A-III: ***Cup-products.*** Let $1 \to T \to P \to Q \to 1$ be an exact sequence of tori. We consider the connecting maps

$$\partial_1 : H^p(F, \widehat{T}(i)) \to H^{p+1}(F, \widehat{Q}(i))$$

for the exact sequence $0 \to \widehat{Q}_{\mathrm{sep}} \to \widehat{P}_{\mathrm{sep}} \to \widehat{T}_{\mathrm{sep}} \to 0$ of character Γ-lattices and

$$\partial_2 : H^q(F, \widehat{Q}^{\circ}(j)) \to H^{q+1}(F, \widehat{T}^{\circ}(j))$$

for the dual sequence of lattices (see notation in Section 4b).

Lemma A.1. *Let $a \in H^p(F, \widehat{T}(i))$ and $b \in H^q(F, \widehat{Q}^{\circ}(j))$ with $i + j \leq 2$. Then $\partial_1(a) \cup b = (-1)^{p+1} a \cup \partial_2(b)$ in $H^{p+q+1}(F, \mathbb{Z}(i+j))$, where the cup-product is defined in (4-3).*

Proof. By [Cartan and Eilenberg 1999, Chapter V, Proposition 4.1], the elements $\partial_1(1_T)$ and $\partial_2(1_Q)$ in

$$H^1(F, \widehat{T}^{\circ}_{\mathrm{sep}} \otimes \widehat{Q}_{\mathrm{sep}}) = \mathrm{Ext}^1_{\Gamma}(\widehat{T}_{\mathrm{sep}}, \widehat{Q}_{\mathrm{sep}})$$

differ by a sign. Write τ for the isomorphism induced by permutation of the factors. By the standard properties of the cup-product, we have

$$\begin{aligned}
\partial_1(a) \cup b = 1_T \cup \partial_1(a) \cup b &= \partial_1(1_T) \cup a \cup b = (-1)^{pq} \tau(\partial_1(1_T) \cup b \cup a) \\
&= (-1)^{pq+1} \tau(\partial_2(1_Q) \cup b \cup a) = (-1)^{pq+1} \tau(1_Q \cup \partial_2(b) \cup a) \\
&= (-1)^{p+1} 1_Q \cup a \cup \partial_2(b) = (-1)^{p+1} a \cup \partial_2(b). \qquad \square
\end{aligned}$$

A-IV: ***Cosimplicial abelian groups.*** Let $A^{\bullet}$ be a cosimplicial abelian group

$$\begin{tikzcd} A^0 \arrow[r, shift left, "d^0"] \arrow[r, shift right, "d^1"'] & A^1 \arrow[r, shift left=2] \arrow[r] \arrow[r, shift right=2] & A^2 \arrow[r, shift left=3] \arrow[r, shift left=1] \arrow[r, shift right=1] \arrow[r, shift right=3] & \cdots \end{tikzcd}$$

and write $h_*(A^{\bullet})$ for the homology groups of the associated complex of abelian groups. In particular,

$$h_0(A^{\bullet}) = \mathrm{Ker}\left[(d^0 - d^1) : A^0 \to A^1\right].$$

We say that the cosimplicial abelian group $A^\bullet$ is *constant* if for every i, all the coface maps $d_j : A^{i-1} \to A^i$, $j = 0, 1, \dots, i$, are isomorphisms. In this case all the d_j are equal as $d_j = s_j^{-1} = d_{j+1}$, where the s_j are the codegeneracy maps. For a constant cosimplicial abelian group $A^\bullet$, we have $h_0(A^\bullet) = A^0$ and $h_i(A^\bullet) = 0$ for all $i > 0$. We will need the following straightforward statement.

Lemma A.2. *Let* $0 \to A^\bullet \to B^\bullet \to C^\bullet \to D^\bullet$ *be an exact sequence of cosimplicial abelian groups with* $A^\bullet$ *a constant cosimplicial group. Then the sequence of groups* $0 \to A^0 \to h_0(B^\bullet) \to h_0(C^\bullet) \to h_0(D^\bullet)$ *is exact.*

Let H be a contravariant functor from the category of schemes over F to the category of abelian groups. We say that H is *homotopy invariant* if for every vector bundle $E \to X$ over F, the induced map $H(X) \to H(E)$ is an isomorphism.

For an integer $d > 0$ consider the following property of the functor H:

Property A.3. For every closed subscheme Z of a scheme X with $\operatorname{codim}_X(Z) \geq d$, the natural homomorphism $H(X) \to H(X \setminus Z)$ is an isomorphism.

Let G be a linear algebraic group over a field F and choose a standard classifying G-torsor $U \to U/G$. Let U^i denote the product of i copies of U. We have the G-torsors $U^i \to U^i/G$.

Consider the cosimplicial abelian group $A^\bullet = H(U^\bullet/G)$ with $A^i = H(U^{i+1}/G)$ and coface maps $A^{i-1} \to A^i$ induced by the projections $U^{i+1}/G \to U^i/G$.

Lemma A.4. *Let* H *be a homotopy invariant functor satisfying* Property A.3 *for some* d. *Let* $U \to U/G$ *be a standard classifying* G*-torsor and* U' *an open subset of a* G*-representation* V'.

1. *If* $\operatorname{codim}_{V'}(V' \setminus U') \geq d$, *then the natural homomorphism*
$$H(U/G) \to H((U \times U')/G)$$
is an isomorphism.
2. *If* $\operatorname{codim}_V(V \setminus U) \geq d$, *then the cosimplicial group* $H(U^\bullet/G)$ *is constant.*

Proof. 1. The scheme $(U \times U')/G$ is an open subset of the vector bundle $(U \times V')/G$ over U/G with complement of codimension at least d. The map in question is the composition $H(U/G) \to H((U \times V')/G) \to H((U \times U')/G)$ and both maps in the composition are isomorphisms since H is homotopy invariant and satisfies Property A.3.

2. By the first part of the lemma applied to the G-torsor $U^i \to U^i/G$ and $U' = U$, the map $H(U^i/G) \to H(U^{i+1}/G)$ induced by a projection $U^{i+1}/G \to U^i/G$ is an isomorphism. $\square$

By Lemma A.4, if H is a homotopy invariant functor satisfying Property A.3 for some d, then the group $H(U/G)$ does not depend on the choice of the representation

V and the open set $U \subset V$ provided $\operatorname{codim}_V(V \setminus U) \geq d$. Following [Totaro 1999], we denote this group by $H(BG)$.

Example A.5. The split torus $T = (\mathbb{G}_m)^n$ over F acts freely on the product U of n copies of $\mathbb{A}^{r+1} \setminus \{0\}$ with $U/T \simeq (\mathbb{P}^r)^n$, that is, BT is "approximated" by the varieties $(\mathbb{P}^r)^n$ if "$r \gg 0$." We then have $\mathrm{CH}^*(BT) = S^*(\widehat{T})$, where S^* represents the symmetric algebra and $\widehat{T}$ is the character group of T; see [Edidin and Graham 1998, p. 607]. In particular, $\operatorname{Pic}(BT) = \mathrm{CH}^1(BT) = \widehat{T}$. More generally, by the Künneth formula [Esnault et al. 1998, Proposition 3.7],

$$H^*_{\mathrm{Zar}}(BT, \mathcal{K}_*) \simeq \mathrm{CH}^*(BT) \otimes K_*(F) \simeq S^*(\widehat{T}) \otimes K_*(F),$$

where $K_n(F)$ is the Quillen K-group of F and $\mathcal{K}_n$ is the Zariski sheaf associated to the presheaf $U \mapsto K_n(U)$.

A-V: *Étale cohomology.* For a scheme X and a closed subscheme $Z \subset X$ we write $H^*_Z(X, \mathbb{Q}/\mathbb{Z}(j))$ for the étale cohomology group of X with support in Z and values in $\mathbb{Q}/\mathbb{Z}(j)$ [Milne 1980, Chapter III, §1]. Write $X^{(i)}$ for the set of points in X of codimension i. For a point $x \in X^{(1)}$ set

$$H^*_x(X, \mathbb{Q}/\mathbb{Z}(j)) = \operatorname*{colim}_{x \in U} H^*_{\overline{\{x\}} \cap U}(U, \mathbb{Q}/\mathbb{Z}(j)),$$

where the colimit is taken over all open subsets $U \subset X$ containing x. If X is a variety, write

$$\partial_x : H^*(F(X), \mathbb{Q}/\mathbb{Z}(j)) \to H^{*+1}_x(X, \mathbb{Q}/\mathbb{Z}(j))$$

for the residue homomorphisms arising from the *coniveau spectral sequence* [Colliot-Thélène et al. 1997, 1.2]

$$E_1^{p,q} = \coprod_{x \in X^{(p)}} H^{p+q}_x(X, \mathbb{Q}/\mathbb{Z}(j)) \Rightarrow H^{p+q}(X, \mathbb{Q}/\mathbb{Z}(j)). \tag{A-1}$$

Let $f : Y \to X$ be a dominant morphism of varieties over F, $y \in Y^{(1)}$, and $x = f(y)$. If $x \in X^{(1)}$, there is a natural homomorphism

$$f^*_y : H^*_x(X, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_y(Y, \mathbb{Q}/\mathbb{Z}(j)).$$

The following lemma is straightforward.

Lemma A.6. *Let $f : Y \to X$ be a dominant morphism of varieties over F, $y \in Y^{(1)}$ and $x = f(y)$.*

(1) *If x is the generic point of X, then the composition*

$$H^*(F(X), \mathbb{Q}/\mathbb{Z}(j)) \xrightarrow{f^*} H^*(F(Y), \mathbb{Q}/\mathbb{Z}(j)) \xrightarrow{\partial_y} H^{*+1}_y(Y, \mathbb{Q}/\mathbb{Z}(j))$$

is trivial.

(2) *If $x \in X^{(1)}$, the diagram*

$$\begin{tikzcd} H^*(F(X), \mathbb{Q}/\mathbb{Z}(j)) \arrow[r, "\partial_x"] \arrow[d, "f^*"'] & H^{*+1}_x(X, \mathbb{Q}/\mathbb{Z}(j)) \arrow[d, "f^*_y"'] \\ H^*(F(Y), \mathbb{Q}/\mathbb{Z}(j)) \arrow[r, "\partial_y"] & H^{*+1}_y(Y, \mathbb{Q}/\mathbb{Z}(j)). \end{tikzcd}$$

is commutative.

Lemma A.7. *Let X be a geometrically irreducible variety, $Z \subset X$ a closed subvariety of codimension* 1, *and x the generic point of Z. Let P be a variety over F such that $P(K)$ is dense in P for every field extension K/F with K infinite, and let y be the generic point of $Z \times P$ in $Y := X \times P$. Then the homomorphism $f^*_y : H^*_x(X, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_y(Y, \mathbb{Q}/\mathbb{Z}(j))$ induced by the projection $f : Y \to X$ is injective.*

Proof. Assume first that the field F is infinite. An element $\alpha \in H^*_x(X, \mathbb{Q}/\mathbb{Z}(j))$ is represented by an element $h \in H^*_{Z\cap U}(U, \mathbb{Q}/\mathbb{Z}(j))$ for a nonempty open set $U \subset X$ containing x. If α belongs to the kernel of

$$f^*_y : H^*_x(X, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_y(Y, \mathbb{Q}/\mathbb{Z}(j)),$$

then there is an open subset $W \subset U \times P$ containing y such that h belongs to the kernel of the composition

$$g : H^*_{Z\cap U}(U, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_{(Z\cap U)\times P}(U \times P, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_{(Z\times P)\cap W}(W, \mathbb{Q}/\mathbb{Z}(j)).$$

As F is infinite, by the assumption on P, there is a rational point $t \in P$ in the image of the dominant composition $(Z \times P) \cap W \hookrightarrow Z \times P \to P$. We have $(U \times t) \cap W = U' \times t$ for an open subset $U' \subset U$ such that $x \in U'$. Composing g with the homomorphism $H^*_{(Z\times P)\cap W}(W, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_{Z\cap U'}(U', \mathbb{Q}/\mathbb{Z}(j))$ induced by the morphism $(U', Z\cap U') \to (W, (Z\times P)\cap W)$, $u \mapsto (u, t)$, we see that h belongs to the kernel of the restriction homomorphism $H^*_{Z\cap U}(U, \mathbb{Q}/\mathbb{Z}(j)) \to H^*_{Z\cap U'}(U', \mathbb{Q}/\mathbb{Z}(j))$, hence the image of α in $H^*_x(X, \mathbb{Q}/\mathbb{Z}(j))$ is trivial.

Suppose now that F is a finite field. Choose a prime integer p and an infinite algebraic pro-p-extension L/F. By the first part of the proof, the statement holds for the variety X_L over L. By the restriction-corestriction argument, $\operatorname{Ker}(f^*_y)$ is a p-primary torsion group. Since this holds for every prime p, we have $\operatorname{Ker}(f^*_y) = 0$. □

Corollary A.8. *Let G be a linear algebraic group over F, let $E \to X$ be a G-torsor over a geometrically irreducible variety X with E a G-rational variety and consider the first projection $p : E^2/G \to X$. Let $x \in X$ and $y \in E^2/G$ be points of* *codimension* 1 *such that* $p(y)=x$. *Then the homomorphism*

$$p_y^*: H_x^*(X, \mathbb{Q}/\mathbb{Z}(j)) \to H_y^*(E^2/G, \mathbb{Q}/\mathbb{Z}(j))$$

is injective.

Proof. Choose a linear G-space V and a nonempty G-variety U that is G-isomorphic to open subschemes of E and V. We can replace the variety E^2/G by $(E\times U)/G$, an open subscheme in the vector bundle $(E\times V)/G$ over X. Shrinking X around x, we may assume that the vector bundle is trivial, that is, $(E\times U)/G$ is isomorphic to an open subscheme in $X\times V$. The statement then follows from Lemma A.7. □

Proposition A.9. *In the conditions of Corollary A.8, let* $h\in H^*(F(X), \mathbb{Q}/\mathbb{Z}(j))_{\mathrm{bal}}$. *Then* $\partial_x(h)=0$ *for every point* $x\in X$ *of codimension* 1.

Proof. Let $y\in E^2/G$ be the point of codimension 1 such that $p_1(y)=x$. As $p_2(y)$ is the generic point of X, by Lemma A.6(1), $\partial_y(h')=0$, where $h'=p_1^*(h)=p_2^*(h)$ in $H^*(F(E^2/G), \mathbb{Q}/\mathbb{Z}(j))$. It follows from Lemma A.6(2) that $\partial_x(h)$ is in the kernel of $(p_1)_y^*: H_x^*(X, \mathbb{Q}/\mathbb{Z}(j)) \to H_y^*(E^2/G, \mathbb{Q}/\mathbb{Z}(j))$ and hence is trivial by Corollary A.8. □

The sheaf $\mathcal{H}^*(\mathbb{Q}/\mathbb{Z}(j))$ defined in Section 3 has a flasque resolution related to the Cousin complex by [Colliot-Thélène et al. 1997, §2] (for the p-components with $p\neq \operatorname{char} F$) and [Gros and Suwa 1988, Theorem 1.4] (for the p-component with $p=\operatorname{char} F>0$):

$$0\to \mathcal{H}^n(\mathbb{Q}/\mathbb{Z}(j)) \to \coprod_{x\in X^{(0)}} i_{x*}H_x^n(X, \mathbb{Q}/\mathbb{Z}(j)) \to \coprod_{x\in X^{(1)}} i_{x*}H_x^{n+1}(X, \mathbb{Q}/\mathbb{Z}(j)) \to \cdots,$$

where $i_x: \operatorname{Spec} F(x)\to X$ are the canonical morphisms. In particular, we have:

Proposition A.10. *Let* X *be a smooth variety over* F. *The sequence*

$$0\to H^0_{\mathrm{Zar}}(X, \mathcal{H}^*(\mathbb{Q}/\mathbb{Z}(j))) \to H^*(F(X), \mathbb{Q}/\mathbb{Z}(j)) \xrightarrow{\partial} \coprod_{x\in X^{(1)}} H_x^{*+1}(X, \mathbb{Q}/\mathbb{Z}(j)),$$

where $\partial=\coprod \partial_x$, *is exact.*

Proposition A.11. *Let* X *be a smooth variety over* F *and* $x\in X$. *The sequence*

$$0\to H^*(O_{X,x}, \mathbb{Q}/\mathbb{Z}(j))) \to H^*(F(X), \mathbb{Q}/\mathbb{Z}(j)) \xrightarrow{\partial} \coprod_{\substack{x'\in X^{(1)}\\ x'\in\overline{\{x\}}}} H_{x'}^{*+1}(X, \mathbb{Q}/\mathbb{Z}(j))$$

is exact.

Appendix B: Spectral sequences

B-I: ***Hochschild–Serre spectral sequence.*** Let

$$\mathcal{A} \xrightarrow{W} \mathcal{B} \xrightarrow{V} \mathcal{C}$$

be additive left exact functors between abelian categories with enough injective objects. If W takes injective objects to V-acyclic ones, there is a spectral sequence

$$E_2^{p,q} = R^p V(R^q W(A)) \Rightarrow R^{p+q}(VW)(A)$$

for every complex A in $\mathcal{A}$ bounded from below.

We have exact triangles in the derived category of $\mathcal{B}$:

$$\tau_{\le n} RW(A) \to RW(A) \to \tau_{\ge n+1} RW(A) \to \tau_{\le n} RW(A)[1], \tag{B-1}$$

$$\tau_{\le n-1} RW(A) \to \tau_{\le n} RW(A) \to R^n W(A)[-n] \to \tau_{\le n-1} RW(A)[1]. \tag{B-2}$$

The filtration on $R^n(VW)(A)$ is defined by

$$F^j R^n(VW)(A) = \operatorname{Im}(R^n V(\tau_{\le(n-j)} RW(A)) \to R^n V(RW(A)) = R^n(VW)(A)).$$

As $\tau_{\ge n+1} RW(A)$ is acyclic in degrees less than or equal to n, the morphism

$$R^n V(\tau_{\le n} RW(A)) \to R^n V(RW(A)) = R^n(VW)(A)$$

is an isomorphism, in particular, $F^0 R^n(VW)(A) = R^n(VW)(A)$.

The *edge* homomorphism is defined as the composition

$$R^n(VW)(A) \xrightarrow{\sim} R^n V(\tau_{\le n} RW(A)) \to R^n V(R^n W(A)[-n]) = V(R^n W(A)).$$

Moreover, the kernel $F^1 R^n(VW)(A)$ of the edge homomorphism is isomorphic to $R^n V(\tau_{\le n-1} RW(A))$. We define the morphism d_n as the composition

$$d_n : F^1 R^n(VW)(A) \to R^n V(R^{n-1} W(A)[-n+1]) = R^1 V(R^{n-1} W(A)) = E_2^{1,n-1}.$$

B-II: ***First spectral sequence.*** Let X be a smooth variety over a field F. We have the functors

$$\mathrm{Sheaves}_{\text{ét}}(X) \xrightarrow{q_*} \mathrm{Sheaves}_{\text{ét}}(F) \xrightarrow{V} Ab,$$

where q_* is the push-forward map for the structure morphism $q : X \to \operatorname{Spec}(F)$ and $V(M) = H^0(F, M)$.

Consider the Hochschild–Serre spectral sequence

$$E_2^{p,q} = H^p(F, H^q(X_{\mathrm{sep}}, \mathbb{Z}(2)) \Rightarrow H^{p+q}(X, \mathbb{Z}(2)). \tag{B-3}$$

Set $\Delta(i) := Rq_*(\mathbb{Z}(i))$ for $i = 1$ or 2. Then $\Delta(i)$ is the complex of étale sheaves on F concentrated in degrees ≥ 1. The j-th term $F^j H^n(X, \mathbb{Z}(i))$ of the filtration on $H^n(X, \mathbb{Z}(i))$ coincides with the image of the canonical homomorphism

$$H^n(F, \tau_{\le(n-j)}\Delta(i)) \to H^n(F, \Delta(i)) = H^n(X, \mathbb{Z}(i)).$$

Let M be a Γ-lattice viewed as an étale sheaf over F. Note that there are canonical isomorphisms

$$H^*(F, M^\circ \otimes \Delta(i)) = \mathrm{Ext}^*_F(M, \Delta(i)) = \mathrm{Ext}^*_X(q^*M, \mathbb{Z}(i)), \tag{B-4}$$

where $M^\circ := \mathrm{Hom}(M, \mathbb{Z})$ is the dual lattice.

Consider also the following product map:

$$\mathbb{Z}(1) \otimes^L \Delta(1) \to Rq_*(q^*\mathbb{Z}(1) \otimes^L \mathbb{Z}(1)) \to Rq_*(\mathbb{Z}(1) \otimes^L \mathbb{Z}(1)) \to Rq_*(\mathbb{Z}(2)).$$

The complex $\mathbb{Z}(1) \otimes^L \tau_{\le 2}\Delta(1)$ is trivial in degrees greater than 3, hence we have a commutative diagram

$$\begin{tikzcd}
\mathbb{Z}(1) \otimes^L \tau_{\le 2}\Delta(1) \arrow[r, "\mathrm{prod}"] \arrow[d] & \tau_{\le 3} Rq_*(\mathbb{Z}(2)) \arrow[d] \arrow[r, phantom, "="] & \tau_{\le 3}\Delta(2) \\
\mathbb{Z}(1) \otimes^L \Delta(1) \arrow[r, "\mathrm{prod}"] & Rq_*(\mathbb{Z}(2)) \arrow[r, phantom, "="] & \Delta(2).
\end{tikzcd}$$

There are canonical morphisms from (B-2):

$$h_2 : \tau_{\le 2}\Delta(1)[2] \to H^2(X_{\mathrm{sep}}, \mathbb{Z}(1)) \quad \text{and} \quad h_3 : \tau_{\le 3}\Delta(2)[3] \to H^3(X_{\mathrm{sep}}, \mathbb{Z}(2)).$$

Consider an element

$$\delta \in H^1(F, M \otimes F^\times_{\mathrm{sep}}) = \mathrm{Ext}^1_F(M^\circ, \mathbb{G}_{m,F}) = \mathrm{Ext}^2_F(M^\circ, \mathbb{Z}(1)),$$

and view δ as a morphism $\delta : M^\circ \to \mathbb{Z}(1)[2]$ in $D^+(\mathrm{Sheaves}_{\text{ét}}(F))$.

The following diagram

$$\begin{tikzcd}
M^\circ \otimes \Delta(1)[2] \arrow[r, "\delta\otimes 1"] & \mathbb{Z}(1) \otimes^L \Delta(1)[4] \arrow[r, "\mathrm{prod}"] & \Delta(2)[4] \\
M^\circ \otimes \tau_{\le 2}\Delta(1)[2] \arrow[r, "\delta\otimes 1"] \arrow[u, "(1\otimes i_2)[2]"] \arrow[d, "1\otimes h_2"'] & \mathbb{Z}(1) \otimes^L \tau_{\le 2}\Delta(1)[4] \arrow[r, "\mathrm{prod}"] \arrow[u, "(1\otimes i_2)[4]"] \arrow[d, "1\otimes h_2"'] & \tau_{\le 3}\Delta(2)[4] \arrow[u, "(i_3)[4]"] \arrow[d, "h_3"'] \\
M^\circ \otimes H^2(X_{\mathrm{sep}}, \mathbb{Z}(1)) \arrow[r, "\delta\otimes 1"] & \mathbb{Z}(1) \otimes^L H^2(X_{\mathrm{sep}}, \mathbb{Z}(1))[2] \arrow[r, "\mathrm{prod}"] & H^3(X_{\mathrm{sep}}, \mathbb{Z}(2))[1]
\end{tikzcd}$$

where $i_2 : \tau_{\le 2}\Delta(1) \to \Delta(1)$ and $i_3 : \tau_{\le 3}\Delta(2) \to \Delta(2)$ are natural morphisms, is commutative.

By (B-4), we have

$$H^0(F, M^\circ \otimes \Delta(1)[2]) = \mathrm{Ext}^2_F(M, \Delta(1)) = \mathrm{Ext}^2_X(q^*M, \mathbb{Z}(1)).$$

Furthermore, the diagram above yields a commutative square

$$\begin{tikzcd} \operatorname{Ext}^2_X(q^*M, \mathbb{Z}(1)) \arrow[r, "q^*(\delta)\,\cup\,-"] \arrow[d, "d_2"'] & F^1H^4(X, \mathbb{Z}(2)) \arrow[d, "d_4"'] \\ \operatorname{Hom}_\Gamma(M, H^2(X_{\mathrm{sep}}, \mathbb{Z}(1)) \arrow[r, "j"] & H^1(F, H^3(X_{\mathrm{sep}}, \mathbb{Z}(2))), \end{tikzcd}$$

where d_2 is the edge map coming from the spectral sequence

$$\operatorname{Ext}^p_F(M, H^q(X_{\mathrm{sep}}, \mathbb{Z}(1))) \Rightarrow \operatorname{Ext}^{p+q}_X(q^*M, \mathbb{Z}(1)) \tag{B-5}$$

and j coincides with the composition

$$\begin{aligned}\operatorname{Hom}_\Gamma(M, H^2(X_{\mathrm{sep}}, \mathbb{Z}(1)) &= H^0(F, M^\circ \otimes H^2(X_{\mathrm{sep}}, \mathbb{Z}(1)))\\ &\xrightarrow{\delta\,\cup\,-} H^1(F, F^\times_{\mathrm{sep}} \otimes H^2(X_{\mathrm{sep}}, \mathbb{Z}(1))) \xrightarrow{\rho} H^1(F, H^3(X_{\mathrm{sep}}, \mathbb{Z}(2))),\end{aligned}$$

with ρ given by the product map.

Now suppose the group $H^2(X_{\mathrm{sep}}, \mathbb{Z}(1))$, which is canonically isomorphic to $\operatorname{Pic}(X_{\mathrm{sep}})$, is a lattice. Let $M = \operatorname{Pic}(X_{\mathrm{sep}})$ and consider the torus T over F with $\widehat{T}_{\mathrm{sep}} = M$. It follows that

$$\delta \in H^1(F, T^\circ) = H^1(F, \widehat{T}_{\mathrm{sep}} \otimes F^\times_{\mathrm{sep}}) = H^2(F, \widehat{T}_{\mathrm{sep}} \otimes \mathbb{Z}(1)),$$

where T° is the dual torus. Note that $\delta \cup 1_M = \delta$, where

$$1_M \in H^0(F, M^\circ \otimes H^2(X_{\mathrm{sep}}, \mathbb{Z}(1))) = \operatorname{End}_\Gamma(M)$$

is the identity.

The top map in the last diagram is given by the pairing

$$\begin{gathered} H^1(X, T^0) \otimes H^1(X, T) \to F^1H^4(X, \mathbb{Z}(2)),\\ a \otimes b \mapsto a \cup b, \end{gathered} \tag{B-6}$$

defined as the cup-product in (4-3),

$$H^2(X, \widehat{T}(1)) \otimes H^2(X, \widehat{T}^\circ(1)) \to F^1H^4(X, \mathbb{Z}(2)),$$

if we identify $\operatorname{Ext}^2_X(q^*M, \mathbb{Z}(1))$ with $H^2(X, \widehat{T}^\circ(1)) = H^1(X, T)$.

In this case, the homomorphism

$$\rho : H^1(F, T^\circ) \to H^1(F, H^3(X_{\mathrm{sep}}, \mathbb{Z}(2))) \tag{B-7}$$

is given by the product homomorphism

$$T^\circ(F_{\mathrm{sep}}) = F^\times_{\mathrm{sep}} \otimes \widehat{T}_{\mathrm{sep}} = F^\times_{\mathrm{sep}} \otimes \operatorname{Pic}(X_{\mathrm{sep}}) \to H^3(X_{\mathrm{sep}}, \mathbb{Z}(2)).$$

A T-torsor $E \to X$ is called *universal* if the class of E in

$$H^1(X,T) = \mathrm{Ext}^2_X(q^*M, \mathbb{Z}(1))$$

satisfies $d_2([E]) = 1_M$; see [Colliot-Thélène and Sansuc 1987a].

Commutativity of the previous diagram gives:

Proposition B.1. *Let X be a smooth variety over F such that* $\mathrm{Pic}(X_{\mathrm{sep}})$ *is a lattice. Let T be the torus over F satisfying $\widehat{T}_{\mathrm{sep}} = \mathrm{Pic}(X_{\mathrm{sep}})$ and let E be a universal T-torsor over X with the class $[E] \in H^1(X,T)$. Then for every $\delta \in H^1(F, T^\circ)$, we have*

$$d_4(q^*(\delta) \cup [E]) = \rho(\delta),$$

where $d_4 : F^1H^4(X, \mathbb{Z}(2)) \to H^1(F, H^3(X_{\mathrm{sep}}, \mathbb{Z}(2)))$ is the map induced by the Hochschild–Serre spectral sequence (B-3) and the cup-product is taken for the pairing (B-6).

B-III: ***Second spectral sequence.*** We assume that $H^3(X_{\mathrm{sep}}, \mathbb{Z}(2)) = 0$, hence in particular $E_2^{0,3} = 0$ in the spectral sequence (B-3) and so $E_\infty^{2,2} \subset E_2^{2,2}$. Therefore, we have a canonical map

$$e_4 : F^2H^4(X, \mathbb{Z}(2)) \to E_\infty^{2,2} \hookrightarrow E_2^{2,2} = H^2(F, H^2(X_{\mathrm{sep}}, \mathbb{Z}(2)).$$

Let N be a Γ-lattice. Consider an element

$$\gamma \in H^2(F, N \otimes F_{\mathrm{sep}}^\times) = \mathrm{Ext}^2_F(N^\circ, \mathbb{G}_{m,F}) = \mathrm{Ext}^3_F(N^\circ, \mathbb{Z}(1)),$$

and view γ as a morphism $\gamma : N^\circ \to \mathbb{Z}(1)[3]$ in $D^+(\mathrm{Sheaves}_{\text{ét}}(F))$.

As above, the commutative diagram

$$\begin{tikzcd}
N^\circ \otimes \Delta(1)[1] \arrow[r, "\gamma\otimes 1"] & \mathbb{Z}(1) \otimes^L \Delta(1)[4] \arrow[r, "\mathrm{prod}"] & \Delta(2)[4] \\
N^\circ \otimes \tau_{\le 1}\Delta(1)[1] \arrow[r, "\gamma\otimes 1"] \arrow[u, "(1\otimes i_1)[1]"] \arrow[d, "1\otimes h_1"'] & \mathbb{Z}(1) \otimes^L \tau_{\le 1}\Delta(1)[4] \arrow[r, "\mathrm{prod}"] \arrow[u, "(1\otimes i_1)[4]"] \arrow[d, "1\otimes h_1"'] & \tau_{\le 2}\Delta(2)[4] \arrow[u, "(i_2)[4]"] \arrow[d, "h_2"'] \\
N^\circ \otimes H^1(X_{\mathrm{sep}}, \mathbb{Z}(1)) \arrow[r, "\gamma\otimes 1"] & \mathbb{Z}(1) \otimes^L H^1(X_{\mathrm{sep}}, \mathbb{Z}(1))[3] \arrow[r, "\mathrm{prod}"] & H^2(X_{\mathrm{sep}}, \mathbb{Z}(2))[2]
\end{tikzcd}$$

where i_1, i_2, h_1 and h_2 are defined in a similar fashion as in Section B-II, yields a commutative square

$$\begin{tikzcd}
\mathrm{Ext}^1_X(q^*N, \mathbb{Z}(1)) \arrow[r, "q^*(\gamma) \cup -"] \arrow[d, "d_1"'] & F^2H^4(X, \mathbb{Z}(2)) \arrow[d, "e_4"] \\
\mathrm{Hom}_\Gamma(N, H^1(X_{\mathrm{sep}}, \mathbb{Z}(1)) \arrow[r, "k"] & H^2(F, H^2(X_{\mathrm{sep}}, \mathbb{Z}(2)))
\end{tikzcd}$$

where d_1 is the edge map coming from the spectral sequence

$$\mathrm{Ext}_F^p(N, H^q(X_{\mathrm{sep}}, \mathbb{Z}(1))) \Rightarrow \mathrm{Ext}_X^{p+q}(q^*N, \mathbb{Z}(1))$$

and k coincides with the composition

$$\begin{aligned}\mathrm{Hom}_\Gamma(N, H^1(X_{\mathrm{sep}}, \mathbb{Z}(1)) = H^0(F, N^\circ \otimes H^1(X_{\mathrm{sep}}, \mathbb{Z}(1)))\\ \xrightarrow{\gamma\,\cup\,-} H^2(F, F_{\mathrm{sep}}^\times \otimes H^1(X_{\mathrm{sep}}, \mathbb{Z}(1))) \to H^2(F, H^2(X_{\mathrm{sep}}, \mathbb{Z}(2)))\end{aligned}$$

with the last homomorphism given by the product map.

Suppose N is a Γ-lattice in $F_{\mathrm{sep}}[X]^\times$ such that the composition

$$N \hookrightarrow F_{\mathrm{sep}}[X]^\times \to F_{\mathrm{sep}}[X]^\times / F_{\mathrm{sep}}^\times$$

is an isomorphism. Consider the torus Q with $\widehat{Q}_{\mathrm{sep}} = N$, so that $\gamma \in H^2(F, Q^\circ)$.

Note that $\gamma \cup i_N = \gamma$, where

$$i_N \in H^0(F, N^\circ \otimes H^1(X_{\mathrm{sep}}, \mathbb{Z}(1))) = \mathrm{Hom}_\Gamma(N, F_{\mathrm{sep}}[X]^\times)$$

is the embedding.

The top map in the previous diagram is given by the pairing

$$\begin{aligned}H^2(X, Q^0) \otimes H^0(X, Q) \to F^2H^4(X, \mathbb{Z}(2)),\\ a \otimes b \mapsto a \cup b,\end{aligned} \tag{B-8}$$

defined as the cup-product in (4-3),

$$H^3(X, \widehat{Q}(1)) \otimes H^1(X, \widehat{Q}^\circ(1))) \to H^4(X, \mathbb{Z}(2)),$$

if we identify $\mathrm{Ext}_X^1(q^*N, \mathbb{Z}(1))$ with $H^1(X, \widehat{Q}^\circ(1)) = H^0(X, Q)$.

The inclusion of $\widehat{Q}_{\mathrm{sep}}$ into $F_{\mathrm{sep}}[X]^\times$ yields a morphism $\varepsilon : X \to Q$ that can be viewed as an element of $H^0(X, Q)$. Consider the map

$$\mu : H^2(F, Q^\circ) \to H^2(F, H^2(X_{\mathrm{sep}}, \mathbb{Z}(2))) \tag{B-9}$$

given by composition with the product homomorphism

$$Q^\circ(F_{\mathrm{sep}}) = F_{\mathrm{sep}}^\times \otimes \widehat{Q}_{\mathrm{sep}} \to F_{\mathrm{sep}}^\times \otimes H^1(X_{\mathrm{sep}}, \mathbb{Z}(1)) \to H^2(X_{\mathrm{sep}}, \mathbb{Z}(2)).$$

We have proved:

Proposition B.2. *Let X be a smooth variety over F such that $H^3(X_{\mathrm{sep}}, \mathbb{Z}(2)) = 0$. Let N be a Γ-lattice in $F_{\mathrm{sep}}[X]^\times$ such that the composition*

$$N \hookrightarrow F_{\mathrm{sep}}[X]^\times \to F_{\mathrm{sep}}[X]^\times / F_{\mathrm{sep}}^\times$$

is an isomorphism. Let Q be the torus over F satisfying $\widehat{Q}_{\mathrm{sep}} = N$. Then for every $\gamma \in H^2(F, Q^\circ)$, we have

$$e_4(q^*(\gamma) \cup \varepsilon) = \mu(\gamma),$$

where $e_4 : F^2 H^4(X, \mathbb{Z}(2)) \to H^2(F, H^2(X_{\mathrm{sep}}, \mathbb{Z}(2)))$ is the map induced by the Hochschild–Serre spectral sequence (B-3) and the cup-product is taken for the pairing (B-8).

B-IV: ***Relative étale cohomology.*** Let X be a smooth variety over F. Following B. Kahn [1996, §3], we define the relative étale cohomology groups as follows. Recall that $\Delta(i) = Rq_*(\mathbb{Z}(i))$ for $i = 1$ and 2, where $q : X \to \mathrm{Spec}(F)$ is the structure morphism, and let $\Delta'(i)$ be the cone of the natural morphism $\mathbb{Z}(i) \to \Delta(i)$ in $D_+(\mathrm{Sheaves}_{\text{ét}}(F))$. Define

$$H^*(X/F, \mathbb{Z}(2)) := H^*(F, \Delta'(2)).$$

(Note that our indexing is different from that in [Kahn 1996, §3].)

There is an infinite exact sequence

$$\cdots \to H^i(F, \mathbb{Z}(2)) \to H^i(X, \mathbb{Z}(2)) \to H^i(X/F, \mathbb{Z}(2)) \to H^{i+1}(F, \mathbb{Z}(2)) \to \cdots$$

If X has a rational point, we have

$$H^i(X/F, \mathbb{Z}(2)) = \overline{H}^i(X, \mathbb{Z}(2)) := H^i(X, \mathbb{Z}(2))/H^i(F, \mathbb{Z}(2)).$$

There is a Hochschild–Serre type spectral sequence [Kahn 1996, §3]

$$E_2^{p,q} = H^p(F, H^q(X_{\mathrm{sep}}/F_{\mathrm{sep}}, \mathbb{Z}(2))) \Rightarrow H^{p+q}(X/F, \mathbb{Z}(2)), \tag{B-10}$$

and we have by [Kahn 1996, Lemma 3.1] that

$$H^q(X_{\mathrm{sep}}/F_{\mathrm{sep}}, \mathbb{Z}(2)) = \begin{cases} 0 & \text{if } q \le 0, \\ \text{uniquely divisible group} & \text{if } q = 1, \\ \overline{H}^0_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2) & \text{if } q = 2, \\ H^1_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2) & \text{if } q = 3. \end{cases}$$

It follows that $E_2^{p,q} = 0$ if $q \le 1$ and $p > 0$. Comparing the spectral sequences (B-3) and (B-10), by Proposition B.1 we have:

Proposition B.3. *Let X be a smooth variety over F such that $X(F) \neq \varnothing$. If $H^0_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2) = K_2(F_{\mathrm{sep}})$, then the spectral sequence (B-10) yields an exact sequence*

$$0 \to H^1(F, H^1_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2)) \xrightarrow{\alpha} \overline{H}^4(X, \mathbb{Z}(2)) \to \overline{H}^4(X_{\mathrm{sep}}, \mathbb{Z}(2))^\Gamma \to H^2(F, H^1_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2)).$$

If, moreover, the group $\mathrm{Pic}(X_{\mathrm{sep}})$ *is a lattice and* T *is the torus over* F *such that* $\widehat{T}_{\mathrm{sep}} = \mathrm{Pic}(X_{\mathrm{sep}})$, *then* $\alpha(\rho(\delta)) = q^*(\delta) \cup [E]$ *for every* $\delta \in H^1(F, T^\circ)$, *where* ρ *is defined in* (B-7) *and* E *is a universal* T*-torsor over* X.

Comparing the spectral sequences (B-3) and (B-10), by Proposition B.2 we have:

Proposition B.4. *Let* X *be a smooth variety over* F *such that* $X(F) \neq \varnothing$. *If* $H^1_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2) = 0$, *then the spectral sequence* (B-10) *yields an exact sequence*

$$0 \to H^2(F, \overline{H}^0_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2)) \xrightarrow{\beta} \overline{H}^4(X, \mathbb{Z}(2)) \to \overline{H}^4(X_{\mathrm{sep}}, \mathbb{Z}(2))^\Gamma \to H^3(F, \overline{H}^0_{\mathrm{Zar}}(X_{\mathrm{sep}}, \mathcal{K}_2)).$$

If N *is a* Γ*-lattice in* $F_{\mathrm{sep}}[X]^\times$ *such that the composition*

$$N \hookrightarrow F_{\mathrm{sep}}[X]^\times \to F_{\mathrm{sep}}[X]^\times / F_{\mathrm{sep}}^\times$$

is an isomorphism and Q *is the torus over* F *satisfying* $\widehat{Q}_{\mathrm{sep}} = N$, *then* $\beta(\mu(\gamma)) = q^*(\gamma) \cup \varepsilon$ *for every* $\gamma \in H^2(F, Q^\circ)$, *where* μ *is defined in* (B-9) *and* $\varepsilon \in H^0(X, Q)$ *is given by the inclusion of* $\widehat{Q}_{\mathrm{sep}}$ *into* $F_{\mathrm{sep}}[X]^\times$.

Acknowledgements

The authors would like to thank Jean-Louis Colliot-Thélène, Skip Garibaldi, Thomas Geisser, David Harari, Bruno Kahn, Kazuya Kato, Alena Pirutka and the referee for helpful discussions and comments.

References

[Bloch and Kato 1986] S. Bloch and K. Kato, "p-adic étale cohomology", *Inst. Hautes Études Sci. Publ. Math.* 63 (1986), 107–152. MR 87k:14018 Zbl 0613.14017

[Bogomolov 1987] F. A. Bogomolov, "The Brauer group of quotient spaces of linear representations", *Izv. Akad. Nauk SSSR Ser. Mat.* **51**:3 (1987), 485–516, 688. In Russian; translated in *Math. USSR-Izv.* **30**:3 (1988), 455–485. MR 88m:16006 Zbl 0641.14005

[Cartan and Eilenberg 1999] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton Landmarks in Mathematics, Princeton University Press, 1999. MR 2000h:18022 Zbl 0933.18001

[Colliot-Thélène 1995] J.-L. Colliot-Thélène, "Birational invariants, purity and the Gersten conjecture", pp. 1–64 in *K-theory and algebraic geometry: connections with quadratic forms and division algebras* (Santa Barbara, CA, 1992), edited by B. Jacob and A. Rosenberg, Proc. Sympos. Pure Math. **58**, Amer. Math. Soc., Providence, RI, 1995. MR 96c:14016 Zbl 0834.14009

[Colliot-Thélène 2004] J.-L. Colliot-Thélène, "Résolutions flasques des groupes réductifs connexes", *C. R. Math. Acad. Sci. Paris* **339**:5 (2004), 331–334. MR 2005f:20081 Zbl 1076.20034

[Colliot-Thélène and Ojanguren 1989] J.-L. Colliot-Thélène and M. Ojanguren, "Variétés unirationnelles non rationnelles: Au-delà de l'exemple d'Artin et Mumford", *Invent. Math.* **97**:1 (1989), 141–158. MR 90m:14012 Zbl 0686.14050

[Colliot-Thélène and Sansuc 1977] J.-L. Colliot-Thélène and J.-J. Sansuc, "La R-équivalence sur les tores", *Ann. Sci. École Norm. Sup.* (4) **10**:2 (1977), 175–229. MR 56 #8576 Zbl 0356.14007

[Colliot-Thélène and Sansuc 1987a] J.-L. Colliot-Thélène and J.-J. Sansuc, "La descente sur les variétés rationnelles, II", *Duke Math. J.* **54**:2 (1987), 375–492. MR 89f:11082 Zbl 0659.14028

[Colliot-Thélène and Sansuc 1987b] J.-L. Colliot-Thélène and J.-J. Sansuc, "Principal homogeneous spaces under flasque tori: Applications", *J. Algebra* **106**:1 (1987), 148–205. MR 88j:14059 Zbl 0597.14014

[Colliot-Thélène et al. 1997] J.-L. Colliot-Thélène, R. T. Hoobler, and B. Kahn, "The Bloch–Ogus–Gabber theorem", pp. 31–94 in *Algebraic K-theory* (Toronto, ON, 1996), edited by V. P. Snaith, Fields Inst. Commun. **16**, Amer. Math. Soc., Providence, RI, 1997. MR 98j:14021 Zbl 0911.14004

[Colliot-Thélène et al. 2005] J.-L. Colliot-Thélène, D. Harari, and A. N. Skorobogatov, "Compactification équivariante d'un tore (d'après Brylinski et Künnemann)", *Expo. Math.* **23**:2 (2005), 161–170. MR 2006c:14076 Zbl 1078.14076

[Demazure and Grothendieck 1970] M. Demazure and A. Grothendieck (editors), *Schémas en groupes, III: Structure des schémas en groupes réductifs. Exposés XIX à XXVI.* (Séminaire de Géométrie Algébrique du Bois Marie 1962/64 (SGA 3)), Lecture Notes in Mathematics **153**, Springer, Berlin, 1970. MR 43 #223c Zbl 0212.52810

[Edidin and Graham 1998] D. Edidin and W. Graham, "Equivariant intersection theory", *Invent. Math.* **131**:3 (1998), 595–634. MR 99j:14003a Zbl 0940.14003

[Esnault et al. 1998] H. Esnault, B. Kahn, M. Levine, and E. Viehweg, "The Arason invariant and mod 2 algebraic cycles", *J. Amer. Math. Soc.* **11**:1 (1998), 73–118. MR 98d:14010 Zbl 1025.11009

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] **2**, Springer, Berlin, 1984. MR 85k:14004 Zbl 0541.14005

[Garibaldi et al. 2003] S. Garibaldi, A. Merkurjev, and J.-P. Serre, *Cohomological invariants in Galois cohomology*, University Lecture Series **28**, American Mathematical Society, Providence, RI, 2003. MR 2004f:11034 Zbl 1159.12311

[Gros 1985] M. Gros, *Classes de Chern et classes de cycles en cohomologie de Hodge–Witt logarithmique*, Mém. Soc. Math. France (N.S.) **21**, Société Mathématique de France, Paris, 1985. MR 87m:14021 Zbl 0615.14011

[Gros and Suwa 1988] M. Gros and N. Suwa, "La conjecture de Gersten pour les faisceaux de Hodge–Witt logarithmique", *Duke Math. J.* **57**:2 (1988), 615–628. MR 89h:14006b Zbl 0715.14011

[Humphreys 1990] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge University Press, 1990. MR 92h:20002 Zbl 0725.20028

[Illusie 1979] L. Illusie, "Complexe de de Rham–Witt et cohomologie cristalline", *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR 82d:14013 Zbl 0436.14007

[Izhboldin 1991] O. Izhboldin, "On p-torsion in K_*^M for fields of characteristic p", pp. 129–144 in *Algebraic K-theory*, edited by A. A. Suslin, Adv. Soviet Math. **4**, Amer. Math. Soc., Providence, RI, 1991. MR 92f:11165 Zbl 0746.19002

[Kahn 1993] B. Kahn, "Descente galoisienne et K_2 des corps de nombres", *K-Theory* **7**:1 (1993), 55–100. MR 94i:11094 Zbl 0780.12007

[Kahn 1996] B. Kahn, "Applications of weight-two motivic cohomology", *Doc. Math.* **1**:17 (1996), 395–416. MR 98b:14007 Zbl 0883.19002

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications **44**, American Mathematical Society, Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Lang 1956] S. Lang, "Algebraic groups over finite fields", *Amer. J. Math.* **78** (1956), 555–563. MR 19,174a Zbl 0073.37901

[Lichtenbaum 1987] S. Lichtenbaum, "The construction of weight-two arithmetic cohomology", *Invent. Math.* **88**:1 (1987), 183–215. MR 88d:14011 Zbl 0615.14004

[Lichtenbaum 1990] S. Lichtenbaum, "New results on weight-two motivic cohomology", pp. 35–55 in *The Grothendieck Festschrift, III*, edited by P. Cartier et al., Progr. Math. **88**, Birkhäuser, Boston, MA, 1990. MR 92m:14030 Zbl 0809.14004

[Merkurjev and Panin 1997] A. S. Merkurjev and I. A. Panin, "K-theory of algebraic tori and toric varieties", *K-Theory* **12**:2 (1997), 101–143. MR 98j:19005 Zbl 0882.19002

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, 1980. MR 81j:14002 Zbl 0433.14012

[Quillen 1973] D. Quillen, "Higher algebraic K-theory, I", pp. 85–147 in *Algebraic K-theory, I: Higher K-theories* (Seattle, Wash., 1972), edited by H. Bass, Lecture Notes in Math. **341**, Springer, Berlin, 1973. MR 49 #2895 Zbl 0292.18004

[Rost 1996] M. Rost, "Chow groups with coefficients", *Doc. Math.* **1**:16 (1996), 319–393. MR 98a:14006 Zbl 0864.14002

[Sansuc 1981] J.-J. Sansuc, "Groupe de Brauer et arithmétique des groupes algébriques linéaires sur un corps de nombres", *J. Reine Angew. Math.* **327** (1981), 12–80. MR 83d:12010 Zbl 0468.14007

[Totaro 1999] B. Totaro, "The Chow ring of a classifying space", pp. 249–281 in *Algebraic K-theory* (Seattle, WA, 1997), edited by W. Raskind and C. Weibel, Proc. Sympos. Pure Math. **67**, Amer. Math. Soc., Providence, RI, 1999. MR 2001f:14011 Zbl 0967.14005

[Voskresenskiĭ 1988] V. E. Voskresenskiĭ, "Maximal tori without affect in semisimple algebraic groups", *Mat. Zametki* **44**:3 (1988), 309–318, 410. MR 90a:11046

[Voskresenskiĭ 1998] V. E. Voskresenskiĭ, *Algebraic groups and their birational invariants*, Translations of Mathematical Monographs **179**, American Mathematical Society, Providence, RI, 1998. MR 99g:20090

Communicated by Jean-Louis Colliot-Thélène
Received 2012-04-23 Revised 2012-10-08 Accepted 2012-11-09

blinstein@math.ucla.edu *Department of Mathematics, University of California, Los Angeles, Los Angeles CA 90095-1555, United States*

merkurev@math.ucla.edu *Department of Mathematics, University of California, Los Angeles, Los Angeles CA 90095-1555, United States*

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1685

On abstract representations of the groups of rational points of algebraic groups and their deformations

Igor A. Rapinchuk

In this paper, we continue our study, begun in an earlier paper, of abstract representations of elementary subgroups of Chevalley groups of rank ≥ 2. First, we extend the methods to analyze representations of elementary groups over arbitrary associative rings and, as a consequence, prove the conjecture of Borel and Tits on abstract homomorphisms of the groups of rational points of algebraic groups for groups of the form $\mathbf{SL}_{n,D}$, where D is a finite-dimensional central division algebra over a field of characteristic 0. Second, we apply the previous results to study deformations of representations of elementary subgroups of universal Chevalley groups of rank ≥ 2 over finitely generated commutative rings.

1. Introduction and statement of the main results

The goal of this paper is twofold. First, we extend the methods and results developed in our paper [Rapinchuk 2011] to analyze abstract representations of Chevalley groups over commutative rings to elementary groups over arbitrary associative rings. As a consequence of this analysis, we prove the conjecture of Borel and Tits [1973, 8.19] on abstract homomorphisms of the groups of rational points of algebraic groups for groups of the form $\mathbf{SL}_{n,D}$, where D is a finite-dimensional central division algebra over a field of characteristic 0. Second, we apply the results of [Rapinchuk 2011] to study deformations of representations of the elementary subgroup $\Gamma = E(\Phi, R)$ of a universal Chevalley group associated to a root system Φ of rank ≥ 2 over a finitely generated commutative ring R. This relies on the description, obtained in [Rapinchuk 2011], of representations with *nonreductive image*, which are at the heart of the Borel–Tits conjecture (recall that representations with reductive image were completely described in [Borel and Tits 1973]). We also use techniques of representation and character varieties (see [Lubotzky and Magid 1985]) in conjunction with the fact that such Γ satisfies Kazhdan's property (T), which was recently established in [Ershov et al. 2011].

MSC2010: primary 20G15; secondary 14L15.
Keywords: abstract homomorphisms, algebraic groups, rigidity, character varieties.

Before formulating of our first result, let us recall the statement of the Borel–Tits conjecture. For an algebraic G defined over a field k, let G^+ denote the subgroup of $G(k)$ generated by the k-points of split (smooth) connected unipotent k-subgroups.

Conjecture (BT). *Let G and G' be algebraic groups defined over infinite fields k and k', respectively. If $\rho\colon G(k) \to G'(k')$ is any abstract homomorphism such that $\rho(G^+)$ is Zariski-dense in $G'(k')$, then there exist a commutative finite-dimensional k'-algebra C and a ring homomorphism $f_C\colon k \to C$ such that $\rho = \sigma \circ r_{C/k'} \circ F$, where $F\colon G(k) \to {}_CG(C)$ is induced by f_C (${}_CG$ is the group obtained by change of scalars), $r_{C/k'}\colon {}_CG(C) \to R_{C/k'}({}_CG)(k')$ is the canonical isomorphism (here $R_{C/k'}$ denotes the functor of restriction of scalars), and σ is a rational k'-morphism of $R_{C/k'}({}_CG)$ to G'.*

If an abstract homomorphism $\rho\colon G(k) \to G'(k')$ admits a factorization as in (BT), we will say that ρ has a *standard description*.

Remarks. (1) Another frequently used definition of G^+, which appears in the introduction of [Borel and Tits 1973], is that it is the subgroup of $G(k)$ generated by the k-points of the unipotent radicals of the parabolic k-subgroups of G. Recall that if G is reductive, then the k-split smooth connected unipotent k-subgroups all lie in the unipotent radicals of minimal parabolic k-subgroups, so in this case, the two definitions coincide. However, they may differ for general smooth connected affine k-groups. Now, it follows from [Conrad et al. 2010, Proposition C.3.11, Theorem C.3.12] that in the case of a general smooth connected affine k-group G, one can also describe G^+ as the subgroup of $G(k)$ generated by the k-points of the k-split unipotent radicals of the minimal pseudoparabolic k-subgroups.

(2) It was pointed out to us by B. Conrad and G. Prasad that, using techniques from the theory of pseudoreductive groups (developed in [Conrad et al. 2010, Chapter 9]), one can construct counterexamples to (BT) over all local and global function fields of characteristic 2 (or, more generally, over any field k of characteristic 2 such that $[k : k^2] = 2$). The groups that arise in these counterexamples are perfect and k-simple. So one should exclude fields of characteristic 2 (and possibly also characteristic 3) in the statement of (BT).

Our result concerning (BT) is as follows. Given a finite-dimensional central division algebra D over a field k, we let $G = \mathbf{SL}_{n,D}$ denote the algebraic k-group such that $G(k) = \mathrm{SL}_n(D)$, the group of elements of $\mathrm{GL}_n(D)$ having reduced norm 1; recall that G is an inner form of type A_l (see [Knus et al. 1998; Platonov and Rapinchuk 1994] for details).

Theorem 1. *Let D be a finite-dimensional central division algebra over a field k of characteristic* 0, *and let $G = \mathbf{SL}_{n,D}$, where $n \geq 3$. Let $\rho\colon G(k) \to \mathrm{GL}_m(K)$ be a finite-dimensional linear representation of $G(k)$ over an algebraically closed* *field K of characteristic* 0, *and set $H = \overline{\rho(G(k))}$ (Zariski-closure). Then the abstract homomorphism $\rho\colon G(k) \to H(K)$ has a standard description.*

In fact, we will see in Section 3 that a similar, but somewhat weaker, statement can be established for representations of elementary groups over arbitrary associative rings, not just division algebras (see Theorem 3.2 for a precise statement). It should be observed that while the overall structure of the proof of Theorem 1 resembles that of the Main Theorem of [Rapinchuk 2011], the analogs of the K-theoretic results of Stein [1973], which played a crucial role in [Rapinchuk 2011], were not available in the noncommutative setting. So part of our argument is dedicated to developing the required K-theoretic results, which is done in Section 2 using the computations of relative K_2 groups given by Bak and Rehmann [1982].

As we have already mentioned, results describing representations of a given group Γ with nonreductive image can be used to analyze *deformations* of representations of Γ, which is the second major theme of this paper. Formally, over a field of characteristic 0, deformations of (completely reducible) n-dimensional representations of a finitely generated group Γ can be understood in terms of the corresponding character variety $X_n(\Gamma)$. For $\Gamma = E(\Phi, R)$, the elementary subgroup of $G(R)$, where G is a universal Chevalley–Demazure group scheme corresponding to a reduced irreducible root system of rank > 1 and R is a finitely generated commutative ring, we use the results of [Rapinchuk 2011] to estimate the dimension of $X_n(\Gamma)$ as a function of n. (We note that it was recently shown in [Ershov et al. 2011] that such Γ possesses Kazhdan's property (T) and hence is finitely generated, so the representation variety $R_n(\Gamma)$ and the associated character variety $X_n(\Gamma)$ are defined. See Section 4 for a brief review of these notions and [Lubotzky and Magid 1985] for complete details.) To put our result into perspective, we recall that for $\Gamma = F_d$, the free group on $d > 1$ generators, the dimension $\varkappa_n(\Gamma) := \dim X_n(\Gamma)$ is given by

$$\varkappa_n(\Gamma) = (d-1)n^2 + 1,$$

i.e., the growth of $\varkappa_n(\Gamma)$ is *quadratic* in n. It follows that the rate of growth cannot be more than quadratic for *any* finitely generated group (and it is indeed quadratic in some important situations such as $\Gamma = \pi_g$, the fundamental group of a compact orientable surface of genus $g > 1$ [Rapinchuk et al. 1996]). At the other end of the spectrum are the groups Γ, called *SS-rigid*, for which $\varkappa_n(\Gamma) = 0$ for all $n \geq 1$. For example, according to the superrigidity theorem of Margulis [1991, Chapter VII, Theorems 5.6, 5.25, and A], all irreducible higher-rank lattices are *SS*-rigid (see Section 5 regarding the superrigidity of groups like $E(\Phi, \mathcal{O})$, where $\mathcal{O}$ is a ring of algebraic integers). Now, in [Rapinchuk 2013], we show that if Γ is not *SS*-rigid, then the rate of growth of $\varkappa_n(\Gamma)$ is at least linear. It follows that unless Γ is *SS*-rigid, the growth rate of $\varkappa_n(\Gamma)$ is between linear and quadratic. Our result shows that for $\Gamma = E(\Phi, R)$ as above, this rate is the minimal possible, namely linear.

To formulate our result, we recall that a pair (Φ, R) consisting of a reduced irreducible root system of rank > 1 and a commutative ring R was called *nice* in [Rapinchuk 2011] if $2 \in R^{\times}$ whenever Φ contains a subsystem of type B_2 and $2, 3 \in R^{\times}$ if Φ is of type G_2.

Theorem 2. *Let Φ be a reduced irreducible root system of rank ≥ 2, R a finitely generated commutative ring such that (Φ, R) is a nice pair, and G the universal Chevalley–Demazure group scheme of type Φ. Let $\Gamma = E(\Phi, R)$ denote the elementary subgroup of $G(R)$, and consider the variety $X_n(\Gamma)$ of characters of n-dimensional representations of Γ over an algebraically closed field K of characteristic* 0. *Then there exists a constant $c = c(R)$ (depending only on R) such that $\varkappa_n(\Gamma) := \dim X_n(\Gamma)$ satisfies*

$$\varkappa_n(\Gamma) \leq c \cdot n$$

for all $n \geq 1$.

The proof is based on a suitable variation of the approach, going back to A. Weil, of bounding the dimension of the tangent space to $X_n(\Gamma)$ at a point $[\rho]$ corresponding to a representation $\rho \colon \Gamma \to \mathrm{GL}_n(K)$ by the dimension of the cohomology group $H^1(\Gamma, \mathrm{Ad}_{\mathrm{GL}_n} \circ \rho)$. Using the results of [Rapinchuk 2011], we describe the latter space in terms of certain spaces of derivations of R. This leads to the conclusion that the constant c in Theorem 2 does not exceed the minimal number of generators d of R (i.e., the smallest integer such that there exists a surjection $\mathbb{Z}[X_1, \dots, X_d] \twoheadrightarrow R$). In fact, if R is the ring of integers or S-integers in a number field L, then $c = 0$ (see Lemma 4.7), so we obtain that $\varkappa_n(\Gamma) = 0$ for all n, i.e., Γ is *SS*-rigid. We then show in Section 5 that the results of [Rapinchuk 2011] actually imply that $\Gamma = E(\Phi, R)$ is in fact superrigid in this case. The proof of Theorem 2 uses the validity of property (T) for $\Gamma = E(\Phi, R)$. On the other hand, groups of this form account for most of the known examples of linear Kazhdan groups, so it is natural to ask if the conclusion of Theorem 2 can be extended to all discrete linear Kazhdan groups.

Conjecture. *Let Γ be a discrete linear group having Kazhdan's property* (*T*). *Then there exists a constant $c = c(\Gamma)$ such that*

$$\varkappa_n(\Gamma) \leq c \cdot n$$

for all $n \geq 1$.

The paper is organized as follows. In Section 2, we begin by summarizing some well known facts from K-theory and then use the results of [Bak and Rehmann 1982] to obtain a description of the group K_2 of certain associative rings similar to the one given by Stein in the commutative case. This is then used in the proof of Theorem 1, which is given in Section 3, along with similar results for arbitrary associative rings. Next, we begin Section 4 with a brief review of representation and character varieties and some related cohomological machinery, after which we turn to the proof of Theorem 2. Finally, in Section 5, we show how the techniques of [Rapinchuk 2011], along with some considerations involving derivations, can be used to establish various rigidity results for the elementary groups $E(\Phi, \mathcal{O})$, where $\mathcal{O}$ is a ring of algebraic integers.

Notations and conventions. Throughout the paper, Φ will denote a reduced irreducible root system of rank ≥ 2. All of our rings are assumed to be associative and unital. As noted earlier, if R is a commutative ring, we say that the pair (Φ, R) is *nice* if $2 \in R^{\times}$ whenever Φ contains a subsystem of type B_2 and $2, 3 \in R^{\times}$ if Φ is of type G_2. Finally, given an algebraic group H, we let H° denote the connected component of the identity.

2. K-theoretic preliminaries

In this section, we develop the K-theoretic results that will be needed in the proof of Theorem 1. Even though the statements in this section are consequences of some well known results, to the best of our knowledge, they have never appeared explicitly in the literature in the form that we require. The main objective will be to use the computations of Bak and Rehmann [1982] to establish certain analogs in the noncommutative setting of Stein's [1973] description of the group K_2 of a semilocal commutative ring (see Propositions 2.3 and 2.4 below).

We begin by recalling some standard definitions. Let R be an associative unital ring. For $1 \leq i, j \leq n$, $i \neq j$, and $r \in R$, let $e_{ij}(r) \in \mathrm{GL}_n(R)$ be the elementary matrix with r in the (i, j)-th place, and let $E_n(R)$ denote the subgroup of $\mathrm{GL}_n(R)$, called the *elementary group*, generated by all the $e_{ij}(r)$. If $n \geq 3$, it is well known that the elementary matrices in $\mathrm{GL}_n(R)$ satisfy the following relations:

(R1) $e_{ij}(r)e_{ij}(s) = e_{ij}(r+s)$.

(R2) $[e_{ij}(r), e_{kl}(s)] = 1$ if $i \neq l$ and $j \neq k$.

(R3) $[e_{ij}(r), e_{jl}(s)] = e_{il}(rs)$ if $i \neq l$.

The *Steinberg group* over R, denoted $\mathrm{St}_n(R)$, is defined to be the group generated by all symbols $x_{ij}(r)$ with $1 \leq i, j \leq n$, $i \neq j$, and $r \in R$ subject to the natural analogs of the relations (R1)–(R3) written in terms of the $x_{ij}(r)$. From the definition, it is clear that there exists a canonical surjective group homomorphism

$$\pi_R\colon\ \mathrm{St}_n(R) \to E_n(R), \quad x_{ij}(r) \mapsto e_{ij}(r),$$

and we set

$$K_2(n, R) = \ker(\mathrm{St}_n(R) \xrightarrow{\pi_R} E_n(R)).$$

It is easy to see that there exist natural homomorphisms $\mathrm{St}_n(R) \to \mathrm{St}_{n+1}(R)$ and $E_n(R) \hookrightarrow E_{n+1}(R)$, which induce homomorphisms $K_2(n, R) \to K_2(n+1, R)$ [Hahn and O'Meara 1989, §1.4]. Also notice that the pair $(\mathrm{St}_n(R), \pi_R)$ is functorial in the following sense: given a homomorphism of rings $f\colon R \to S$, there is a commutative diagram of group homomorphisms

$$\begin{array}{ccc} \mathrm{St}_n(R) & \xrightarrow{\tilde{F}} & \mathrm{St}_n(S) \\ {\scriptstyle \pi_R}\big\downarrow & & \big\downarrow{\scriptstyle \pi_S} \\ E_n(R) & \xrightarrow{F} & E_n(S) \end{array}$$

where F and $\tilde{F}$ are the homomorphisms induced by f defined on generators by

$$F\colon e_{ij}(t) \mapsto e_{ij}(f(t)) \quad \text{and} \quad \tilde{F}\colon x_{ij}(t) \mapsto x_{ij}(f(t)).$$

It follows from the commutativity of the above diagram that $\tilde{F}$ induces a homomorphism $K_2(n, R) \to K_2(n, S)$. In the following proposition, we derive some general properties of $K_2(n, R)$ that will be needed later in this section:

Proposition 2.1. (a) *Suppose R is an associative unital ring such that $R/J(R)$ is artinian, where $J(R)$ is the Jacobson radical of R. Then the natural map $K_2(3, R) \to K_2(4, R)$ is an isomorphism. If, moreover, R is finitely generated as a module over its center, then $K_2(n, R)$ is a central subgroup of $\mathrm{St}_n(R)$ for $n \geq 3$.*

(b) *Suppose C is a commutative finite dimensional algebra over a field K, and let $A = M_m(C)$ be the ring of $m \times m$ matrices over C. For $a \in C$ and $1 \leq k, l \leq m$, let $\tilde{y}_{kl}(a) \in A$ be the matrix with a as the (k, l) entry and 0 for all other entries. Then for $n \geq 3$, the maps*

$$\tilde{\psi}(x_{ij}^A(\tilde{y}_{kl}(a))) = x^C_{(i-1)m+k,(j-1)m+l}(a),$$
$$\psi(e_{ij}^A(\tilde{y}_{kl}(a))) = e^C_{(i-1)m+k,(j-1)m+l}(a),$$

where the $x_{ij}^A(a)$ and $e_{ij}^A(a)$ are the generators of $\mathrm{St}_n(A)$ and $E_n(A)$ and the $x_{ij}^C(c)$ and $e_{ij}^C(c)$ are the generators of $\mathrm{St}_{nm}(C)$ and $E_{nm}(C)$, respectively, define isomorphisms $\tilde{\psi}\colon \mathrm{St}_n(A) \to \mathrm{St}_{nm}(C)$ and $\psi\colon E_n(A) \to E_{nm}(C)$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathrm{St}_n(A) & \xrightarrow{\tilde{\psi}} & \mathrm{St}_{nm}(C) \\ {\scriptstyle \pi_A}\big\downarrow & & \big\downarrow{\scriptstyle \pi_C} \\ E_n(A) & \xrightarrow{\psi} & E_{nm}(C) \end{array} \tag{1}$$

In particular, $K_2(n, M_m(C)) \simeq K_2(nm, C)$.

Proof. (a) By Theorem 7 of [van der Kallen 1976], the fact that $R/\mathrm{Rad}(R)$ is artinian implies that it has property SR_2^*, and then Theorem 6 of the same work yields the required isomorphism. Now, if R is finitely generated as a module over its center, then according to [Hahn and O'Meara 1989, Theorem 1.4.15], $\pi_R\colon \mathrm{St}_n(R) \to E_n(R)$ is a central extension for $n \geq 4$ (in fact, a *universal* central extension for $n \geq 5$). So in view of the canonical isomorphism $K_2(3, R) \simeq K_2(4, R)$, we obtain that $K_2(n, R)$ is a central subgroup of $\mathrm{St}_n(R)$ for $n \geq 3$, as claimed.

(b) First notice that the natural group isomorphism $\mathrm{GL}_n(A) \xrightarrow{\sim} \mathrm{GL}_{nm}(C)$ restricts to a group homomorphism $\psi\colon E_n(A) \to E_{nm}(C)$. By direct computation with commutator relations, one sees that ψ is surjective for $n \geq 3$ and hence an isomorphism. Moreover, on generators it is given by the second formula in the statement. Now, since A is generated additively by the $\tilde{y}_{kl}(a)$, with $1 \leq k, l \leq m$, it follows that the $\tilde{x}_{ij}^A(\tilde{y}_{kl}(a))$ generate $\mathrm{St}_n(A)$, so it suffices to define $\tilde{\psi}$ on these elements and check the defining relations. This is done by direct computation using the definition of $\tilde{\psi}$ given above.

Next, since without loss of generality $m \geq 2$, we have $nm \geq 6$, so as noted in the proof of (a), $\pi_C\colon \mathrm{St}_{nm}(C) \to E_{nm}(C)$ is a *universal* central extension and $\pi_A\colon \mathrm{St}_n(A) \to E_n(A)$ is a central extension. Hence, there exists a unique group homomorphism $\tilde{\varphi}\colon St_{nm}(C) \to \mathrm{St}_n(A)$ making the diagram

$$\begin{tikzcd} \mathrm{St}_n(A) \arrow[d, "\pi_A"'] & \mathrm{St}_{nm}(C) \arrow[l, "\tilde{\varphi}"'] \arrow[d, "\pi_C"] \\ E_n(A) \arrow[r, "\psi"] & E_{nm}(C) \end{tikzcd} \tag{2}$$

commute, and by universality, we conclude that $\tilde{\psi} \circ \tilde{\varphi} = \mathrm{id}_{\mathrm{St}_{nm}(C)}$. On the other hand, by the commutativity of the diagrams (1) and (2), we have that for any $x \in \mathrm{St}_n(A)$,

$$(\psi \circ \pi_A \circ \tilde{\varphi} \circ \tilde{\psi})(x) = (\pi_C \circ \tilde{\psi})(x) = (\psi \circ \pi_A)(x).$$

Since ψ is an isomorphism, we conclude that $(\tilde{\varphi} \circ \tilde{\psi})(x) = x z_x$, where $z_x \in K_2(n, A)$. The centrality of $K_2(n, A)$ then implies that the map $x \mapsto z_x$ is a homomorphism $\mathrm{St}_n(A) \to K_2(n, A)$, which must be trivial as $\mathrm{St}_n(A)$ is a perfect group and $K_2(n, A)$ is commutative. Thus, $\tilde{\varphi} \circ \tilde{\psi} = \mathrm{id}_{\mathrm{St}_n(A)}$, as required. It then follows that $K_2(n, A) \simeq K_2(nm, C)$. □

Next, let us summarize the results of [Bak and Rehmann 1982] dealing with relative K_2 groups of associative rings (see Theorem 2.2 below). From now on, we will always assume that $n \geq 3$. First, we need to introduce some additional notation. As above, let R be an associative unital ring. Given $u \in R^\times$, we define, for $i \neq j$, the following standard elements of $\mathrm{St}_n(R)$:

$$w_{ij}(u) = x_{ij}(u)x_{ji}(-u^{-1})x_{ij}(u) \quad \text{and} \quad h_{ij}(u) = w_{ij}(u)w_{ij}(-1).$$

Notice that the image $\pi_R(h_{ij}(u))$ in $E_n(R)$ is the diagonal matrix with u as the i-th diagonal entry, u^{-1} as the j-th diagonal entry, and 1s everywhere else on the diagonal. We will also need the following noncommutative version of the usual Steinberg symbols: for $u, v \in R^\times$, let

$$c(u, v) = h_{12}(u)h_{12}(v)h_{12}(vu)^{-1}.$$

One easily sees that $\pi_R(c(u, v))$ is the diagonal matrix with $uvu^{-1}v^{-1}$ as its first diagonal entry and 1s everywhere else on the diagonal. Let $U_n(R)$ be the subgroup of $\mathrm{St}_n(R)$ generated by all the $c(u, v)$ with $u, v \in R^\times$.

As in the commutative case, one can also consider relative versions of these constructions. Let $\mathfrak{a}$ be a two-sided ideal of R and

$$\mathrm{GL}_n(R, \mathfrak{a}) = \ker(\mathrm{GL}_n(R) \to \mathrm{GL}_n(R/\mathfrak{a}))$$

be the congruence subgroup of level $\mathfrak{a}$. Define $E_n(R, \mathfrak{a})$ to be the normal subgroup of $E_n(R)$ generated by all elementary matrices $e_{ij}(a)$ with $a \in \mathfrak{a}$. Now letting

$$\mathrm{St}_n(R, \mathfrak{a}) = \ker(\mathrm{St}_n(R) \to \mathrm{St}_n(R/\mathfrak{a})),$$

we have a natural homomorphism $\mathrm{St}_n(R, \mathfrak{a}) \to E_n(R, \mathfrak{a})$, and we set

$$K_2(n, R, \mathfrak{a}) = \ker(\mathrm{St}_n(R, \mathfrak{a}) \to E_n(R, \mathfrak{a})).$$

Finally, let

$$U_n(R, \mathfrak{a}) := \langle c(u, 1+a) \mid u \in R^\times,\ 1+a \in (1+\mathfrak{a}) \cap R^\times \rangle$$

(notice this is contained in $\mathrm{St}_n(R, \mathfrak{a})$). We should point out that even though for a noncommutative ring, the groups $U_n(R)$ and $U_n(R, \mathfrak{a})$ may not lie in $K_2(n, R)$, it is well known that any element of $K_2(n, R) \cap U_n(R)$ is automatically contained in the center of $\mathrm{St}_n(R)$ [Milnor 1971, Corollary 9.3]. This will be needed in Proposition 2.3 below.

Theorem 2.2 [Bak and Rehmann 1982, Theorem 2.9, Corollary 2.11]. *Let R be an associative unital ring. Suppose that $\mathfrak{a}$ is a two-sided ideal contained in the Jacobson radical $J(R)$ of R and that R is additively generated by $R^\times$. Assume $n \geq 3$. Then the following are true*:

(1) $K_2(n, R, \mathfrak{a}) \subset U_n(R, \mathfrak{a})$, *and the canonical sequence below is exact*:

$$1 \to U_n(R, \mathfrak{a}) \to U_n(R) \to U_n(R/\mathfrak{a}) \to 1.$$

(2) *If, moreover,* $K_2(n, R/\mathfrak{a}) \subset U_n(R/\mathfrak{a})$, *then* $K_2(n, R) \subset U_n(R)$ *and the natural sequence*

$$1 \to K_2(n, R, \mathfrak{a}) \to K_2(n, R) \to K_2(n, R/\mathfrak{a}) \to 1$$

is exact.

The theorem yields the following:

Proposition 2.3. *Suppose that* R *is either a finite-dimensional algebra over an algebraically closed field* K *or a finite ring with* $2 \in R^\times$. *Then* $K_2(n, R) \subset U_n(R)$, *and consequently,* $K_2(n, R)$ *is a central subgroup of* $\mathrm{St}_n(R)$.

Proof. Let $J = J(R)$ be the Jacobson radical of R. To apply Theorem 2.2, we need to verify that in both cases, R is additively generated by its units and that $K_2(n, R/J) \subset U_n(R/J)$.

If R is a finite-dimensional algebra over K, then we can view R as a connected algebraic ring over K, and it follows from [Rapinchuk 2011, Corollary 2.5] that R is generated by $R^\times$.[1] Now suppose that R is a finite ring. Since R is obviously artinian, R/J is semisimple [Lam 2001, Theorem 4.14], so by the Artin–Wedderburn theorem [Lam 2001, Theorem 3.5] and the fact that finite division rings are commutative [Lam 2001, Theorem 13.1], we have

$$R/J \simeq M_{n_1}(F_1) \oplus \cdots \oplus M_{n_r}(F_r),$$

where $F_1, \ldots, F_r$ are finite fields with $F_i \neq \mathbb{F}_2$, the field of two elements, for all i as $2 \in R^\times$. It follows that R/J is additively generated by its units. On the other hand, the canonical map $R \to R/J$ induces a surjective homomorphism $R^\times \to (R/J)^\times$, which, combined with the fact that J lies in the linear span of $R^\times$ [Lam 2001, Lemma 4.3], yields that R is additively generated by $R^\times$.

Next, let us show that $K_2(n, R/J) \subset U_n(R/J)$ in both cases. If R is a finite-dimensional K-algebra, then as above R/J is semisimple. So since there are no nontrivial division algebras over algebraically closed fields, the Artin–Wedderburn theorem implies that

$$R/J \simeq M_{n_1}(K) \oplus \cdots \oplus M_{n_s}(K).$$

Thus, in both cases, R/J is a direct sum of matrix algebras over fields. Since K_2 commutes with finite direct sums, we may assume without loss of generality that $A := R/J \simeq M_m(F)$ with F a field. By Proposition 2.1, we have isomorphisms $\tilde{\psi}\colon \mathrm{St}_n(A) \to \mathrm{St}_{nm}(F)$ and $\psi\colon E_n(A) \to E_{nm}(F)$ that induce an isomorphism

[1] All the background on algebraic rings needed in this paper can be found in [Rapinchuk 2011, §2]. M. Kassabov has also informed us that the notion of an algebraic ring actually goes back to [Greenberg 1964], where one can find proofs of some basic properties.

$K_2(n, A) \simeq K_2(nm, F)$. Now let $u \in F^\times$ and $t_u = \mathrm{diag}(u, 1, \ldots, 1) \in M_m(F)$. By direct computation, one checks that

$$\tilde{\psi}(h_{12}^A(t_u)) = h_{1,m+1}^F(u),$$

and therefore, for $u, v \in F^\times$, we have

$$\tilde{\psi}(c(t_u, t_v)) = c_{1,m+1}(u, v),$$

where $c_{1,m+1}(u, v) = h_{1,m+1}^F(u)h_{1,m+1}^F(v)h_{1,m+1}^F(vu)^{-1}$. On the other hand, by Matsumoto's theorem, the group $K_2(nm, F)$ is generated by the Steinberg symbols $c_{1,m+1}(u, v)$ [Steinberg 1968]; consequently, we see $K_2(n, R/J) \subset U_n(R/J)$, as claimed. Hence, $K_2(n, R) \subset U_n(R)$ by Theorem 2.2. As noted above, it now follows from [Milnor 1971, Corollary 9.3] that $K_2(n, R)$ lies in the center of $\mathrm{St}_n(R)$. □

An important ingredient in the proof of Theorem 1 will be the following:

Proposition 2.4. *Let k and K be fields of characteristic 0 with K algebraically closed. Suppose that D is a finite-dimensional central division algebra over k, A a finite-dimensional algebra over K, and $f : D \to A$ a ring homomorphism with Zariski-dense image. Then for $n \geq 3$, $K_2(n, A)$ coincides with the subgroup*

$$U_n'(A) = \big\langle c(u, v) \mid u, v \in \overline{f(L^\times)} \big\rangle,$$

where L is an arbitrary maximal subfield of D.

We begin with the following:

Lemma 2.5. *Let A, D, and f be as above, and set $C = \overline{f(k)}$ (Zariski closure). Then*

$$A \simeq D \otimes_k C \simeq M_s(C) \tag{3}$$

as K-algebras, where $s^2 = \dim_k D$. Moreover, if L is any maximal subfield of D, then the second isomorphism can be chosen so that $L \otimes_k C \simeq D_s(C)$, where $D_s(C) \subset M_s(C)$ is the subring of diagonal matrices.

Proof. We start with the proof of the first isomorphism in (3). To begin, we note that since k and K are both fields of characteristic 0, C is a finite-dimensional algebra over K by [Rapinchuk 2011, Lemma 2.13, Proposition 2.14]. Moreover, by [Greenberg 1964, Proposition 5.1], the natural inclusion $C \hookrightarrow A$ is a homomorphism of K-algebras (this also follows from the proof of [Rapinchuk 2011, Proposition 2.14]). Now consider the map

$$\theta : D \otimes_k C \to A, \quad (x, c) \mapsto f(x)c.$$

We claim that θ is an isomorphism. From the above remark, it is clear that θ is a homomorphism of K-algebras (where $D \otimes_k C$ is endowed with the natural K-algebra structure coming from C). For surjectivity, first note that since $\mathrm{im}\,\theta$

contains $f(D)$, it is Zariski-dense in A. On the other hand, let $x_1, \dots, x_{s^2}$ be a basis of D over k. Then

$$\operatorname{im}\theta = f(x_1)C + \cdots + f(x_{s^2})C$$

and therefore is closed. Hence, θ is surjective. To prove injectivity, notice that since D is a central simple algebra, $\ker\theta = D \otimes_k \mathfrak{c}$ for some ideal $\mathfrak{c} \subset C$ [Farb and Dennis 1993, Theorem 3.5]. On the other hand, since by construction the restriction $\theta|_{\mathfrak{c}}$ is an embedding, we have $\mathfrak{c} = 0$, and θ is injective.

Now let us consider the second isomorphism. First, since C is a commutative artinian algebraic ring, by [Rapinchuk 2011, Proposition 2.20], we can write

$$C = C_1 \times \cdots \times C_r,$$

where each C_i is a local commutative algebraic ring. Moreover, since tensor products commute with finite products and $M_s(C_1 \times \cdots \times C_r) = M_s(C_1) \times \cdots \times M_s(C_r)$, it suffices to establish the isomorphism when C is a local algebraic ring. So suppose that is the case, and let $J(C)$ be the Jacobson radical of C. Then it follows from [Rapinchuk 2011, Corollary 2.6, Proposition 2.19] that $C/J(C) \simeq K$, so composing f with the canonical map $C \to C/J(C)$, we obtain an embedding $k \hookrightarrow K$. Consequently, as K is algebraically closed, the division algebra D splits over K, i.e., there exists an isomorphism

$$\tau\colon D \otimes_k K \xrightarrow{\sim} M_s(K). \tag{4}$$

Notice also if L is a maximal subfield of D, we can choose τ so that $L \otimes_k K \simeq D_s(K)$. Indeed, since L is separable over k (as $\operatorname{char} k = 0$) and $[L : k] = s$, we can write $L = k[X]/(f)$, where f is a separable polynomial of degree s. Then by the Chinese remainder theorem, $L \otimes_k K \simeq K^s$. But any subalgebra of $M_s(K)$ that is isomorphic to K^s is conjugate to $D_s(K)$ [Gille and Szamuely 2006, Lemma 2.2.9], so it follows that τ can be composed with an inner automorphism of $M_s(K)$ to have the required form.

Now consider the natural (surjective) map

$$D \otimes_k C \to D \otimes_k (C/J(C)) = D \otimes_k K.$$

Since D is a central simple algebra, the same argument as above shows that the kernel of this map is contained in the Jacobson radical $J(D \otimes_k C)$, and the fact that $D \otimes_k K \simeq M_s(K)$ is semisimple implies that it actually coincides with $J(D \otimes_k C)$. So by the Wedderburn–Malcev theorem [Pierce 1982, Corollary 11.6], there exists a section

$$\alpha\colon M_s(K) \simeq D \otimes_k K \hookrightarrow D \otimes_k C.$$

We claim that the following map gives the required isomorphism:

$$\beta\colon M_s(K)\otimes_K C \to D\otimes_k C, \quad m\otimes c \mapsto \alpha(m)\cdot(1\otimes c).$$

Indeed, injectivity is proved by the same argument as above, and surjectivity follows by dimension count. Thus, $M_s(C) \simeq M_s(K)\otimes_K C \simeq D\otimes_k C$, and it follows immediately from the above remarks that $D_s(C) \simeq L\otimes_k C$. □

Proof of Proposition 2.4. By Lemma 2.5, we have $L\otimes_k C \simeq D_s(C)$. Moreover, $L\otimes_k C \simeq \overline{f(L)}$. Indeed, since $k\subset L$, we have

$$f(L)\subset \theta(L\otimes_k C)\subset \overline{f(L)}.$$

On the other hand, the same argument as in the proof of Lemma 2.5 shows that $\theta(L\otimes_k C)$ is closed.

Next, since $A\simeq M_s(C)$ and C is a finite-dimensional K-algebra, there exists by Proposition 2.1 an isomorphism $\tilde{\psi}\colon \mathrm{St}_n(A)\to \mathrm{St}_{ns}(C)$ that induces an isomorphism $K_2(n, A)\simeq K_2(ns, C)$. Now, C is a semilocal commutative ring that is additively generated by its units, so $K_2(ns, C)$ coincides with the subgroup $U_{ns}(C)$ of $\mathrm{St}_{ns}(C)$ generated by the Steinberg symbols $c_{1,s+1}(u, v)$ taken with respect to the root $\alpha_{1,s+1}$ (i.e., $c_{1,s+1}(u, v) = h_{1,s+1}(u)h_{1,s+1}(v)h_{1,s+1}(vu)^{-1}$) by [Stein 1973, Theorem 2.13]. As we noted in the proof of Proposition 2.3, we have

$$\tilde{\psi}(c(t_u, t_v)) = c_{1,s+1}(u, v),$$

where for $u\in C^\times$, we set $t_u = \mathrm{diag}(u, 1, \dots, 1)\in M_s(C)$. Thus, $K_2(n, A)$ is contained in the group generated by the symbols $c(t_u, t_v)$. On the other hand, since all of the t_u are diagonal matrices, they lie in the image of $L\otimes_k C$; hence, $K_2(n, A)\subset U_n'(A)$. Since clearly $U_n'(A)\subset K_2(n, A)$, this concludes the proof. □

3. Abstract homomorphisms over noncommutative rings

The main goal of this section is to give the proof of Theorem 1. Before beginning the argument, we would like to give an alternative statement of Theorem 1, which can be generalized (in a somewhat weaker form) to (essentially) arbitrary associative rings. First, we need to observe that if B is a finite-dimensional algebra over an algebraically closed field K, then the elementary group $E_n(B)$ has the structure of a connected algebraic K-group. Indeed, using the regular representation of B over K, it is easy to see that $\mathrm{GL}_n(B)$ is a Zariski-open subset of $M_n(B)$ and hence an algebraic group over K. Now let us view B as a connected algebraic ring over K, and for $i, j\in\{1, \dots, n\}$, $i\neq j$, consider the regular maps

$$\varphi_{ij}\colon B\to \mathrm{GL}_n(B), \quad b\mapsto e_{ij}(b).$$

Set $W_{ij} = \operatorname{im} \varphi_{ij}$. Then each W_{ij} contains the identity matrix $I_n \in \mathrm{GL}_n(B)$, and by definition, $E_n(B)$ is generated by the W_{ij}. So $E_n(B)$ is a connected algebraic group by [Borel 1991, Proposition 2.2].

Theorem 3.1. *Suppose k and K are fields of characteristic 0 with K algebraically closed, D is a finite-dimensional central division algebra over k, and n is an integer ≥ 3. Let $\rho\colon E_n(D) \to \mathrm{GL}_m(K)$ be a finite-dimensional linear representation, and set $H = \overline{\rho(E_n(D))}$ (Zariski closure). Then there exist a finite-dimensional associative K-algebra $\mathcal{B}$, a ring homomorphism $f\colon D \to \mathcal{B}$ with Zariski-dense image, and a morphism $\sigma : E_n(\mathcal{B}) \to H$ of algebraic K-groups such that*

$$\rho = \sigma \circ F,$$

where $F\colon E_n(D) \to E_n(\mathcal{B})$ is the group homomorphism induced by f.

We also have the following result for general associative rings:

Theorem 3.2. *Suppose R is an associative ring with $2 \in R^{\times}$, K is an algebraically closed field of characteristic 0, and n is an integer ≥ 3. Let $\rho\colon E_n(R) \to \mathrm{GL}_m(K)$ be a finite-dimensional linear representation, set $H = \overline{\rho(E_n(R))}$, and let H° denote the connected component of H. If the unipotent radical of H° is commutative, there exist a finite-dimensional associative K-algebra $\mathcal{B}$, a ring homomorphism $f\colon R \to \mathcal{B}$ with Zariski-dense image, and a morphism $\sigma : E_n(\mathcal{B}) \to H$ of algebraic K-groups such that for a suitable finite-index subgroup $\Delta \subset E_n(R)$, we have*

$$\rho|_\Delta = (\sigma \circ F)|_\Delta,$$

where $F\colon E_n(R) \to E_n(\mathcal{B})$ is the group homomorphism induced by f.

As we indicated in the introduction, the proofs of Theorems 3.1 and 3.2 are based on a natural extension of the approach developed in our earlier paper [Rapinchuk 2011]. More precisely, we will first associate to ρ an algebraic ring A, then show that ρ can be lifted to a representation $\tilde{\tau}\colon \mathrm{St}_n(A) \to H$ of the Steinberg group, and finally use the results of Section 2 to verify that $\tilde{\sigma}$ descends to an abstract representation of $E_n(A)$. Then, to conclude the argument, we will prove that this abstract representation is actually a morphism of algebraic groups.

We begin with the construction of the algebraic ring A attached to a given representation ρ.

Proposition 3.3. *Suppose R is an associative ring, K is an algebraically closed field, and $n \geq 3$. Given a representation $\rho\colon E_n(R) \to \mathrm{GL}_m(K)$, there exists an associative algebraic ring A together with a homomorphism of abstract rings $f\colon R \to A$ having Zariski-dense image such that for all $i, j \in \{1, \ldots, n\}$, $i \neq j$, there is an injective regular map $\psi_{ij}\colon A \to H$ into $H := \overline{\rho(E_n(R))}$ satisfying*

$$\rho(e_{ij}(t)) = \psi_{ij}(f(t)) \tag{5}$$

for all $t \in R$.

Proof. This statement goes back to [Kassabov and Sapir 2009] (see also [Rapinchuk 2011, Theorem 3.1]). For the sake of completeness, we indicate the main points of the construction. Let $A = \overline{\rho(e_{13}(R))}$. If $\boldsymbol{\alpha}\colon A \times A \to A$ denotes the restriction of the matrix product in H to A, it is clear $(A, \boldsymbol{\alpha})$ is a commutative algebraic subgroup of H. We let $f\colon R \to A$ be the map defined by $t \mapsto \rho(e_{13}(t))$. From the definition, it follows that

$$\boldsymbol{\alpha}(f(t_1), f(t_2)) = f(t_1 + t_2)$$

for all $t_1, t_2 \in R$. To define the multiplication operation $\boldsymbol{\mu}\colon A \times A \to A$, we will need the elements

$$\overline{w}_{12} = e_{12}(1)e_{21}(-1)e_{12}(1) \quad \text{and} \quad \overline{w}_{23} = e_{23}(1)e_{32}(-1)e_{23}(1)$$

(notice that these are simply the images under π_R of the elements $w_{ij}(1)$ considered in Section 2). By direct computation, one sees that

$$\overline{w}_{12}^{-1} e_{13}(r)\overline{w}_{12} = e_{23}(r), \quad \overline{w}_{23} e_{13}(r)\overline{w}_{23}^{-1} = e_{12}(r),$$

and

$$[e_{12}(r), e_{23}(s)] = e_{13}(rs)$$

for all $r, s \in R$, where $[g, h] = ghg^{-1}h^{-1}$. Now let $\boldsymbol{\mu}\colon A \times A \to H$ be the regular map defined by

$$\boldsymbol{\mu}(a_1, a_2) = [\rho(\overline{w}_{23})a_1\rho(\overline{w}_{23})^{-1}, \rho(\overline{w}_{12})^{-1}a_2\rho(\overline{w}_{12})].$$

Then the above relations yield

$$\boldsymbol{\mu}(f(t_1), f(t_2)) = f(t_1 t_2),$$

so, in particular, $\boldsymbol{\mu}(f(R) \times f(R)) \subset f(R)$, which implies that $\boldsymbol{\mu}(A \times A) \subset A$ and allows us to view $\boldsymbol{\mu}$ as a regular map $\boldsymbol{\mu}\colon A \times A \to A$. Since by our assumption R is a (unital) associative ring and f has Zariski-dense image, it follows that $(A, \boldsymbol{\alpha}, \boldsymbol{\mu})$ is a (unital) associative algebraic ring as defined in [Rapinchuk 2011, §2]. Furthermore, by our construction, (5) obviously holds for the inclusion map $\psi_{13}\colon A \to H$. Finally, using an appropriate element $\overline{w}_{ij}$, we can conjugate any root subgroup $e_{ij}(R)$ into $e_{13}(R)$, from which the existence of all the other maps ψ_{ij} follows. □

Remark 3.4. Observe that if R is an infinite division ring, then the algebraic ring A constructed in Proposition 3.3 is automatically connected. Indeed, the connected component A° is easily seen to be a two-sided ideal of A. So if $A \neq A^\circ$, then $f^{-1}(A^\circ)$ would be a proper two-sided ideal of finite index in R, which is impossible. In particular, we see that in the situation of Theorem 3.1, the algebraic ring associated to ρ is connected.

Next, we show that the representation ρ can be lifted to a representation of the Steinberg group $\mathrm{St}_n(A)$. The precise statement is given by the following proposition:

Proposition 3.5. *Suppose R is an associative ring, K is an algebraically closed field, and $n \geq 3$, and let $\rho \colon E_n(R) \to \mathrm{GL}_m(K)$ be a representation. Let A and $f \colon R \to A$ be the algebraic ring and ring homomorphism constructed in Proposition 3.3. Then there exists a group homomorphism $\tilde{\tau} \colon \mathrm{St}_n(A) \to H \subset \mathrm{GL}_m(K)$ such that $\tilde{\tau} \colon x_{ij}(a) \mapsto \psi_{ij}(a)$ for all $a \in A$ and all $i, j \in \{1, \ldots, n\}$, $i \neq j$. Consequently, $\tilde{\tau} \circ \tilde{F} = \rho \circ \pi_R$, where $\tilde{F} \colon \mathrm{St}_n(R) \to \mathrm{St}_n(A)$ is the homomorphism induced by f.*

Proof. This proposition is proved in exactly the same way as [Rapinchuk 2011, Proposition 4.2]. We simply note that since $\mathrm{St}_n(A)$ is generated by the symbols $x_{ij}(a)$ subject to the relations (R1)–(R3) given in Section 2, to establish the existence of $\tilde{\tau}$, it suffices to verify that relations (R1)–(R3) are satisfied if the $x_{ij}(a)$ are replaced by $\psi_{ij}(a)$, which follows from (5) and the fact that f has Zariski-dense image. For the second statement, we observe that the maps $\tilde{\tau} \circ \tilde{F}$ and $\rho \circ \pi_R$ both send the symbol $x_{ij}(s)$ to $\psi_{ij}(f(s)) = \rho(e_{ij}(s)) = (\rho \circ \pi_R)(x_{ij}(s))$, so they must coincide on $\mathrm{St}_n(R)$. □

To analyze the representation $\tilde{\sigma}$ that we have just constructed, we will need some additional information on the structure of the group $\mathrm{St}_n(A)$.

Proposition 3.6. *Let K be an algebraically closed field of characteristic 0 and n an integer ≥ 3. Suppose A is an associative algebraic ring over K such that $2 \in A^{\times}$, and let A° denote the connected component of 0_A. Then*

(i) $\mathrm{St}_n(A) = \mathrm{St}_n(A^{\circ}) \times P$, *where P is a finite group and*

(ii) $K_2(n, A^{\circ})$ *is a central subgroup of* $\mathrm{St}_n(A^{\circ})$.

Proof. (i) First, since $\operatorname{char} K = 0$, by [Rapinchuk 2011, Proposition 2.14], we have $A = A^{\circ} \oplus S$ with S a finite ring. So

$$\mathrm{St}_n(A) = \mathrm{St}_n(A^{\circ}) \times \mathrm{St}_n(S),$$

and we need to show that $\mathrm{St}_n(S)$ is a finite group. Now, since $E_n(S)$ is obviously a finite group and $K_2(n, S)$ is by definition the kernel of the canonical map $\pi_S \colon \mathrm{St}_n(S) \to E_n(S)$, we see that the finiteness of $\mathrm{St}_n(S)$ is equivalent to that of $K_2(n, S)$. On the other hand, since $2 \in S^{\times}$, Proposition 2.3 implies that $K_2(n, S)$ is a central subgroup of $\mathrm{St}_n(S)$. So we can use the argument given in the proof of [Rapinchuk 2011, Proposition 4.5] and consider the Hochschild–Serre spectral sequence

$$H^1(\mathrm{St}_n(S), \mathbb{Q}/\mathbb{Z}) \to H^1(K_2(\Phi, S), \mathbb{Q}/\mathbb{Z})^{\mathrm{St}_n(S)} \to H^2(E_n(S), \mathbb{Q}/\mathbb{Z})$$

(where all groups act trivially on $\mathbb{Q}/\mathbb{Z}$) corresponding to the short exact sequence

$$1 \to K_2(n, S) \to \mathrm{St}_n(S) \xrightarrow{\pi_S} E_n(S) \to 1$$

to conclude that $K_2(n, S)$ is finite.

(ii) By [Rapinchuk 2011, Proposition 2.14], A° is a finite-dimensional K-algebra, so the assertion follows from Proposition 2.3. $\square$

Remark 3.7. We would like to point out that the assumption that $2 \in A^\times$ is needed to guarantee that the finite ring S that appears in the proof of Proposition 3.6(i) above is additively generated by its units, which then enables us to use Proposition 2.3 to conclude that $K_2(n, S)$ is a central subgroup of $\mathrm{St}_n(S)$. If S is a finite commutative ring, then, as we show in [Rapinchuk 2011, Proposition 4.5], this assumption is not needed since in that case S can be written as a finite product of commutative local rings, which are automatically generated by their units.

To complete the proofs of Theorems 3.1 and 3.2, the basic idea will be to show that the homomorphism $\tilde{\tau}$ constructed in Proposition 3.5 descends to a (rational) representation of $E_n(A)$. Let us make this more precise. Given a representation $\rho\colon E_n(R) \to \mathrm{GL}_m(K)$, let $f\colon R \to A$ be the ring homomorphism associated to ρ (Proposition 3.3), and let $\tilde{F}\colon \mathrm{St}_n(R) \to \mathrm{St}_n(A)$ and $F\colon E_n(R) \to E_n(A)$ denote the group homomorphisms induced by f. Then under the hypotheses of Theorems 3.1 and 3.2, we have $\mathrm{St}_n(A) = \mathrm{St}_n(A^\circ)$ (Remark 3.4) and $\mathrm{St}_n(A) = \mathrm{St}_n(A^\circ) \times P$ (Proposition 3.6), respectively, so in both cases, $\tilde{\Delta} := \tilde{F}^{-1}(\mathrm{St}_n(A^\circ))$ and $\Delta := \pi_R(\tilde{\Delta})$ are finite-index subgroups of $\mathrm{St}_n(R)$ and $E_n(R)$. Moreover, $F(\Delta) \subset E_n(A^\circ)$ clearly. Thus, letting $\tilde{\sigma}$ denote the restriction of $\tilde{\tau}$ to $\mathrm{St}_n(A^\circ)$, we see that the solid arrows in

$$\begin{tikzcd}
\tilde{\Delta} \arrow[r, "\tilde{F}"] \arrow[d, "\pi_R"'] & \mathrm{St}_n(A^\circ) \arrow[d, "\pi_{A^\circ}"] \arrow[ddr, "\tilde{\sigma}"] & \\
\Delta \arrow[r, "F"] \arrow[drr, "\rho"'] & E_n(A^\circ) \arrow[dr, dotted, "\sigma"] & \\
& & H^\circ
\end{tikzcd} \tag{6}$$

form a commutative diagram. In the remainder of this section, we will show that under our assumptions, there exists a group homomorphism $\sigma\colon E_n(A^\circ) \to H^\circ$ (in fact, a morphism of algebraic groups) making the full diagram commute. In the situation of Theorem 3.1, the existence of the required *abstract* homomorphism σ will be shown in Proposition 3.8 below. For Theorem 3.2, we will first need to establish the somewhat weaker result that there exists a homomorphism $\bar{\sigma}\colon E_n(A^\circ) \to \overline{H}$ such that $\bar{\sigma} \circ \pi_{A^\circ} = \nu \circ \tilde{\sigma}$, where $Z(H^\circ)$ is the center of H°, $\overline{H} = H^\circ/Z(H^\circ)$, and $\nu\colon H^\circ \to \overline{H}$ is the canonical map (see Proposition 3.10).

Proposition 3.8. *Suppose k and K are fields of characteristic 0 with K algebraically closed, D is a finite-dimensional central division algebra over k, and n is an integer ≥ 3. Let $\rho\colon E_n(D) \to \mathrm{GL}_m(K)$ be a representation, and let A denote the algebraic ring associated to ρ (Proposition 3.3). Then $A = A^\circ$ is a finite-dimensional K-algebra and there exists a homomorphism of abstract groups $\sigma : E_n(A^\circ) \to H^\circ$ making the diagram (6) commute.*

Proof. We have $A = A^\circ$ by Remark 3.4, and A° is a finite-dimensional K-algebra by [Rapinchuk 2011, Proposition 2.14]. Next, by Proposition 2.4, $K_2(n, A)$ coincides with the subgroup

$$U_n'(A) = \langle c(u, v) \mid u, v \in \overline{f(L^\times)} \rangle$$

of $\mathrm{St}_n(A)$, where L is an arbitrary maximal subfield of D and $f : D \to A$ is the ring homomorphism associated to ρ. Now, from the construction of $\tilde{\sigma}$ and the definition of $c(u, v)$, we have

$$\tilde{\sigma}(c(u, v)) = H_{12}(u)H_{12}(v)H_{12}(vu)^{-1},$$

where for $r \in A^\times$, we set

$$H_{12}(r) = W_{12}(r)W_{12}(-1) \quad \text{and} \quad W_{12}(r) = \psi_{12}(r)\psi_{21}(-r^{-1})\psi_{12}(r).$$

By [Rapinchuk 2011, Proposition 2.4], the map $A^\times \to A^\times,\ t \mapsto t^{-1}$ is regular, which implies that the map

$$\Theta\colon A^\times \times A^\times \to H, \quad (u, v) \mapsto \tilde{\tau}(c(u, v))$$

is also regular. On the other hand, as we observed earlier, $\pi_D(h_{ij}(u)) \in E_n(D)$ is a diagonal matrix with u as the i-th diagonal entry, u^{-1} as the j-th diagonal entry, and 1s everywhere else on the diagonal. In particular, for $u, v \in L^\times$, it follows that

$$\pi_D(h_{12}(u)h_{12}(v)h_{12}(vu)^{-1}) = 1.$$

So by Proposition 3.5,

$$\tilde{\sigma}(c(f(u), f(v))) = \rho(\pi_D(h_{12}(u)h_{12}(v)h_{12}(vu)^{-1})) = 1$$

for all $u, v \in L^\times$. By the regularity of Θ, we obtain that $\tilde{\sigma}(c(a, b)) = 1$ for all $a, b \in \overline{f(L^\times)}$, and consequently, $\tilde{\sigma}$ vanishes on $K_2(n, A)$. Since the canonical homomorphism $\pi_A\colon \mathrm{St}_n(A) \to E_n(A)$ is surjective and $K_2(n, A) = \ker \pi_A$ by definition, the existence of σ now follows. □

The proof of Theorem 3.2 will require the following proposition, which contains analogs of results established in [Rapinchuk 2011, §5]:

Proposition 3.9. *Suppose R is an associative ring with $2 \in R^\times$, K is an algebraically closed field of characteristic 0, and $n \geq 3$. Let $\rho\colon E_n(R) \to \mathrm{GL}_m(K)$ be a* *representation, set* $H = \overline{\rho(E_n(R))}$, *and let* A *denote the algebraic ring associated to* ρ. *Then the following hold*:

(i) *The group* H° *coincides with* $\tilde{\sigma}(\mathrm{St}_n(A^\circ))$ *and is its own commutator.*

(ii) *Let* U *and* $Z(H^\circ)$ *be the unipotent radical and center of* H°, *respectively. If* U *is commutative*, *then* $Z(H^\circ) \cap U = \{e\}$, *and consequently*, $Z(H^\circ)$ *is finite and is contained in any Levi subgroup of* H°.

Proof. (i) It follows from Proposition 3.5 that $\tilde{\sigma}(\mathrm{St}_n(A^\circ))$ coincides with the (abstract) group $\mathcal{H} \subset H$ generated by all the $\psi_{ij}(A^\circ)$ with $i, j \in \{1, \ldots, n\}, i \neq j$. Since $\psi_\alpha(A^\circ)$ is clearly a connected subgroup of H, by [Borel 1991, Proposition 2.2], $\mathcal{H}$ is Zariski-closed and connected; hence, $\mathcal{H} \subset H^\circ$. On the other hand, by Proposition 3.6, $\mathrm{St}_n(A^\circ)$ is a finite-index subgroup of $\mathrm{St}_n(A)$, from which it follows that $\tilde{\sigma}(\mathrm{St}_n(A))$ is Zariski-closed. Since $\tilde{\sigma}(\mathrm{St}_n(A))$ contains $\rho(E_n(R))$, it is Zariski-dense in H and therefore coincides with H. So $\mathcal{H}$ is a closed subgroup of finite index in H; hence, $\mathcal{H} \supset H^\circ$, and consequently, $\mathcal{H} = H^\circ$. Furthermore, from the definition of the Steinberg group, one easily sees that $\mathrm{St}_n(A^\circ)$ coincides with its commutator subgroup, so the same is true for H°.

(ii) Using the fact that H° coincides with its commutator subgroup, one can now apply the argument given in the proof of [Rapinchuk 2011, Proposition 5.5]. □

Now set $\overline{H} = H^\circ / Z(H^\circ)$. Since $Z(H^\circ)$ is a closed normal subgroup of H°, $\overline{H}$ is an (affine) algebraic group and the canonical map $\nu \colon H^\circ \to \overline{H}$ is a morphism of algebraic groups [Borel 1991, Theorem 6.8].

Proposition 3.10. *Suppose* R *is an associative ring with* $2 \in R^\times$, K *is an algebraically closed field of characteristic* 0, *and* $n \geq 3$. *Let* $\rho \colon E_n(R) \to \mathrm{GL}_m(K)$ *be a representation*, *set* $H = \overline{\rho(E_n(R))}$, *and let* A *denote the algebraic ring associated to* ρ. *Then* A° *is a finite-dimensional* K*-algebra and there exists a homomorphism* $\bar{\sigma} \colon E_n(A^\circ) \to \overline{H}$ *such that* $\bar{\sigma} \circ \pi_{A^\circ} = \nu \circ \tilde{\sigma}$.

Proof. Since $\operatorname{char} K = 0$, by [Rapinchuk 2011, Proposition 2.14], A° is a finite-dimensional K-algebra. Furthermore, $H^\circ = \tilde{\sigma}(\mathrm{St}_n(A^\circ))$ by Proposition 3.9 and $K_2(n, A^\circ) = \ker \pi_{A^\circ}$ is a central subgroup of $\mathrm{St}_n(A^\circ)$ by Proposition 2.3, from which the existence of $\bar{\sigma}$ follows. □

The remaining step in the proof is to show that the (abstract) homomorphisms $\sigma \colon E_n(A^\circ) \to H^\circ$ and $\bar{\sigma} \colon E_n(A^\circ) \to \overline{H}$ constructed in Propositions 3.8 and 3.10, respectively, are actually morphisms of algebraic groups (see Proposition 3.12 below). In the latter case, this will allow us to lift $\bar{\sigma}$ to a morphism of algebraic groups $\sigma \colon E_n(A^\circ) \to H^\circ$ making the diagram (6) commute. Our proof of rationality here will deviate from the approach of [Rapinchuk 2011] as rather than using results about the "big cell" of $E_n(A^\circ)$, we will instead apply the following geometric lemma:

Lemma 3.11. *Let X, Y, and Z be irreducible varieties over an algebraically closed field K of characteristic* 0*. Suppose $s\colon X \to Y$ and $t\colon X \to Z$ are regular maps with s dominant such that for any $x_1, x_2 \in X$ with $s(x_1) = s(x_2)$, we have $t(x_1) = t(x_2)$. Then there exists a rational map $h\colon Y \dashrightarrow Z$ such that $h \circ s = t$ on a suitable open subset of X.*

Proof. Let $W \subset X \times Y \times Z$ be the subset

$$W = \{(x, y, z) \mid s(x) = y,\ t(x) = z\}.$$

Notice that W is the graph of the morphism

$$\varphi\colon X \to Y \times Z, \quad x \mapsto (s(x), t(x)),$$

so W is an irreducible variety isomorphic to X. Now consider the projection $\mathrm{pr}_{Y\times Z}\colon X \times Y \times Z \to Y \times Z$, and let $U = \mathrm{pr}_{Y\times Z}(W)$ and $V = \overline{U}$, where the bar denotes the Zariski closure. Then V is an irreducible variety. Moreover, U is constructible by [Humphreys 1975, Theorem 4.4] so in particular contains a dense open subset P of V, which is itself an irreducible variety. Let now $p\colon P \to Y$ be the projection to the first component. We claim that p gives a birational isomorphism between P and Y. From our assumptions, we see that p is dominant, and since $\operatorname{char} K = 0$, p is also separable. So it follows from [Humphreys 1975, Theorem 4.6] that to show that p is birational, we only need to verify that it is injective. Consider $u_1 = (y_1, z_1)$ and $u_2 = (y_2, z_2)$ in P, with $y_1 = y_2$. By our construction, there exist $x_1, x_2 \in X$ such that $s(x_1) = y_1$, $t(x_1) = z_1$, $s(x_2) = y_2$, and $t(x_2) = z_2$. Since $s(x_1) = s(x_2)$, we have $t(x_1) = t(x_2)$, so $u_1 = u_2$, as needed.

Since p is birational, we can now take $h = \pi_Z \circ p^{-1}\colon Y \dashrightarrow Z$, where we let $\pi_Z\colon Y \times Z \to Z$ be the projection. □

Now let $\rho\colon E_n(R) \to \mathrm{GL}_m(K)$ be a representation as in Theorem 3.1 or 3.2, and let A denote the algebraic ring associated to ρ. Also let Q be the set of all pairs (i, j) with $1 \le i, j \le n$, $i \ne j$. Then, as we already observed at the beginning of this section, $E_n(A^\circ)$ is the connected algebraic group generated by the images $W_q = \operatorname{im}\varphi_q$ of the regular maps

$$\varphi_q\colon A^\circ \to \mathrm{GL}_n(A^\circ), \quad a \mapsto e_q(a)$$

for all $q \in Q$. In particular, [Borel 1991, Proposition 2.2] implies that there exists a finite sequence $(\alpha(1), \dots, \alpha(v))$ in Q such that

$$E_n(A^\circ) = W_{\alpha(1)}^{\varepsilon_1} \cdots W_{\alpha(v)}^{\varepsilon_v},$$

where each $\varepsilon_i = \pm 1$. Let

$$X = \prod_{i=1}^{v} (A^\circ)_{\alpha(i)}$$

be the product of v copies of A° indexed by the $\alpha(i)$, and define a regular map $s\colon X \to E_n(A^\circ)$ by

$$s(a_{\alpha(1)}, \ldots, a_{\alpha(v)}) = \varphi_{\alpha(1)}(a_{\alpha(1)})^{\varepsilon_1} \cdots \varphi_{\alpha(v)}(a_{\alpha(v)})^{\varepsilon_v}. \tag{7}$$

Also let

$$t\colon X \to H^\circ, \quad t(a_{\alpha(1)}, \ldots, a_{\alpha(v)}) = \psi_{\alpha(1)}(a_{\alpha(1)})^{\varepsilon_1} \cdots \psi_{\alpha(v)}(a_{\alpha(v)})^{\varepsilon_v}, \tag{8}$$

where the $\psi_{\alpha(i)}$ are the regular maps from Proposition 3.3. With this setup, we can now prove:

Proposition 3.12. *The homomorphisms $\sigma\colon E_n(A^\circ) \to H^\circ$ and $\bar{\sigma}\colon E_n(A^\circ) \to \bar{H}$ constructed in Propositions 3.8 and 3.10, respectively, are morphisms of algebraic groups.*

Proof. We will only consider σ as the argument for $\bar{\sigma}$ is completely analogous. Set $Y = E_n(A^\circ)$ and $Z = H^\circ$, and let $s\colon X \to Y$ and $t\colon X \to Z$ be the regular maps defined in (7) and (8). From the construction of σ, it is clear that $(\sigma \circ s)(x) = t(x)$, so in particular, $s(x_1) = s(x_2)$ for $x_1, x_2 \in X$ implies that $t(x_1) = t(x_2)$. Hence, by Lemma 3.11, σ is a rational map. Therefore, there exists an open subset $V \subset E_n(A^\circ)$ such that $\sigma|_V$ is regular. So it follows from the next lemma that $\sigma\colon E_n(A^\circ) \to H^\circ$ is a morphism. □

Lemma 3.13 [Rapinchuk 2011, Lemma 6.4]. *Let K be an algebraically closed field, and let $\mathscr{G}$ and $\mathscr{G}'$ be affine algebraic groups over K with $\mathscr{G}$ connected. Suppose $f\colon \mathscr{G} \to \mathscr{G}'$ is an abstract group homomorphism,[2] and assume there exists a Zariski-open set $V \subset \mathscr{G}$ such that $\varphi := f|_V$ is a regular map. Then f is a morphism of algebraic groups.*

Theorem 3.1 now follows from Propositions 3.8 and 3.12 with $\mathscr{B} = A^\circ$ $(= A)$. For Theorem 3.2, we again take $\mathscr{B} = A^\circ$, and it remains to show that one can lift the morphism $\bar{\sigma}\colon E_n(A^\circ) \to \bar{H}$ to a morphism $\sigma\colon E_n(A^\circ) \to H^\circ$ making the diagram (6) commute. This is accomplished through a suitable modification of the argument used in the proof of [Rapinchuk 2011, Proposition 6.6]. For this, we need some analogs of results contained in [Rapinchuk 2011, §6] regarding the structure of $E_n(B)$ as an algebraic K-group, where B is an arbitrary finite-dimensional algebra over an algebraically closed field K. Let $J = J(B)$ be the Jacobson radical of B. Then by the Wedderburn–Malcev theorem [Pierce 1982, Corollary 11.6], there exists a semisimple subalgebra $\bar{B} \subset B$ such that $B = \bar{B} \oplus J$ as K-vector spaces and $\bar{B} \simeq B/J$ as K-algebras. Furthermore, since K is algebraically closed, the Artin–Wedderburn theorem implies that

$$\bar{B} = M_{n_1}(K) \times \cdots \times M_{n_r}(K).$$

[2] Here we tacitly identify $\mathscr{G}$ and $\mathscr{G}'$ with the corresponding groups $\mathscr{G}(K)$ and $\mathscr{G}'(K)$ of K-points.

Now consider the group homomorphism $E_n(B) \to E_n(\bar{B})$ induced by the canonical map $B \to B/J$ (notice that this is a morphism of algebraic groups as $B \to B/J$ is a homomorphism of algebraic rings: see [Rapinchuk 2011, Lemma 2.9]), and let $E_n(J)$ be its kernel. It is clear that $E_n(J)$ is a closed normal subgroup of $E_n(B)$. Note that

$$E_n(M_{n_i}(K)) \simeq E_{nn_i}(K) \simeq \mathrm{SL}_{nn_i}(K),$$

so $E_n(\bar{B})$ is a semisimple simply connected algebraic group. It is also easy to see that for any $a, b \geq 1$, we have

$$[\mathrm{GL}_n(B, J^a), \mathrm{GL}_n(B, J^b)] \subset \mathrm{GL}_n(B, J^{a+b}),$$

where $\mathrm{GL}_n(B, J^s) = \ker(\mathrm{GL}_n(B) \to \mathrm{GL}_n(B/J^s))$. Since J is a nilpotent ideal, it follows that $E_n(J)$ is a nilpotent group. In particular, we obtain that

$$E_n(B) = E_n(J) \rtimes E(\bar{B}) \tag{9}$$

is a Levi decomposition of $E_n(B)$ [Rapinchuk 2011, Proposition 6.5].

Now, using the Levi decomposition (9) for $B = \mathcal{B}$ as well as the fact that the center $Z(H^\circ)$ is finite (Proposition 3.9), one can directly imitate the argument of [Rapinchuk 2011, Proposition 6.6] to conclude the proof of Theorem 3.2.

Finally, to derive Theorem 1 from Theorem 3.1, we first note that by Lemma 2.5, we have K-algebra isomorphisms

$$\mathcal{B} \simeq D \otimes_k C \simeq M_s(C),$$

where $s^2 = \dim_k D$ and $C = \overline{f(k)}$ (as above, $f : D \to \mathcal{B}$ is the ring homomorphism associated to ρ). Consequently, $E_n(\mathcal{B}) \simeq E_n(M_s(C)) \simeq E_{ns}(C)$. Moreover, since C is a finite-dimensional K-algebra, in particular a semilocal commutative ring, $E_{ns}(C) \simeq \mathrm{SL}_{ns}(C)$ [Matsumoto 1966, Corollary 2]. So since $G = \mathbf{SL}_{n,D}$ is K-isomorphic to SL_{ns} [Platonov and Rapinchuk 1994, 2.3.1], we see $E_n(\mathcal{B}) \simeq G(C)$. Letting $f_C : k \to C$ be the restriction of f to k, we now obtain Theorem 1.

4. Applications to representation varieties and deformations of representations

In this section, we will prove Theorem 2. To estimate the dimension of the character variety $X_n(\Gamma)$ for an elementary subgroup Γ as in the statement of Theorem 2, we will exploit the well known connection, going back to A. Weil, between the tangent space of $X_n(\Gamma)$ at a given point and the 1-cohomology of Γ with coefficients in the space of a naturally associated representation. We then use the results of [Rapinchuk 2011] on standard descriptions of representations of Γ to relate the latter space to a certain space of derivations of the finitely generated ring R used to define Γ (see Proposition 4.4). Since the dimensions of spaces of derivations are finite and are bounded by a constant depending only on R, we obtain the required bound on $\dim X_n(\Gamma)$. Throughout this section, we will work over a fixed algebraically closed field K of characteristic 0.

We begin by summarizing some key definitions and basic properties related to representation and character varieties, mostly following the first two chapters of [Lubotzky and Magid 1985]. Let Γ be a finitely generated group, and fix an integer $n \geq 1$. Recall that the n-th *representation scheme* of Γ is the functor $\mathfrak{R}_n(\Gamma)$ from the category of commutative K-algebras to the category of sets defined by

$$\mathfrak{R}_n(\Gamma)(A) = \mathrm{Hom}(\Gamma, \mathrm{GL}_n(A)).$$

More generally, if $\mathcal{G}$ is a linear algebraic group over K, we let the representation scheme of Γ with values in $\mathcal{G}$ be the functor $\mathfrak{R}(\Gamma, \mathcal{G})$ defined by

$$\mathfrak{R}(\Gamma, \mathcal{G})(A) = \mathrm{Hom}(\Gamma, \mathcal{G}(A)).$$

Because for any commutative K-algebra A, a homomorphism $\rho\colon \Gamma \to \mathrm{GL}_n(A)$ is determined by the images of the generators, subject to the defining relations of Γ, one shows that $\mathfrak{R}_n(\Gamma)$ is an affine K-scheme represented by a finitely generated K-algebra $\mathfrak{A}_n(\Gamma)$. Similarly, $\mathfrak{R}(\Gamma, \mathcal{G})$ is an affine K-scheme represented by a finitely generated K-algebra $\mathfrak{A}(\Gamma, \mathcal{G})$ [Lubotzky and Magid 1985, Proposition 1.2]. The set $\mathfrak{R}_n(\Gamma)(K)$ of K-points of $\mathfrak{R}_n(\Gamma)$ is then denoted $R_n(\Gamma)$ and is called the n-th *representation variety* of Γ. It is an affine variety over K with coordinate ring $A_n(\Gamma) = \mathfrak{A}_n(\Gamma)_{\mathrm{red}}$, the quotient of $\mathfrak{A}_n(\Gamma)$ by its nilradical. The representation variety $R(\Gamma, \mathcal{G})$ is defined analogously.

Now let $\rho_0 \in R(\Gamma, \mathcal{G})$. To describe the Zariski tangent space of $\mathfrak{R}(\Gamma, \mathcal{G})$ at ρ_0, denoted $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$, we will use the algebra of dual numbers $K[\varepsilon]$ (where $\varepsilon^2 = 0$). More specifically, it is well known that $\mathfrak{R}(\Gamma, \mathcal{G})(K[\varepsilon])$ is the tangent bundle of $\mathfrak{R}(\Gamma, \mathcal{G})$, and therefore, $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$ can be identified with the fiber over ρ_0 of the map $\mu\colon \mathfrak{R}(\Gamma, \mathcal{G})(K[\varepsilon]) \to \mathfrak{R}(\Gamma, \mathcal{G})(K)$ induced by the augmentation homomorphism $K[\varepsilon] \to K$, $\varepsilon \mapsto 0$ [Borel 1991, AG 16.2]. In other words, we have

$$T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G})) = \{\rho \in \mathrm{Hom}(\Gamma, \mathcal{G}(K[\varepsilon])) \mid \mu \circ \rho = \rho_0\}.$$

For us, it will be useful to have the following alternative description of $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$. Let $\tilde{\mathfrak{g}}$ be the Lie algebra of $\mathcal{G}$. Notice that $\tilde{\mathfrak{g}}$ has a natural Γ-action given by

$$\gamma \cdot x = \mathrm{Ad}(\rho_0(\gamma))x$$

for $\gamma \in \Gamma$ and $x \in \tilde{\mathfrak{g}}$, where $\mathrm{Ad}\colon \mathcal{G}(K) \to \mathrm{GL}(\tilde{\mathfrak{g}})$ is the adjoint representation. Now $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$ can be identified with the space $Z^1(\Gamma, \tilde{\mathfrak{g}})$ of 1-cocycles [Lubotzky and Magid 1985, Proposition 2.2]. Indeed, an element $c \in Z^1(\Gamma, \tilde{\mathfrak{g}})$ is by definition a map $c\colon \Gamma \to \tilde{\mathfrak{g}}$ such that

$$c(\gamma_1\gamma_2) = c(\gamma_1) + \mathrm{Ad}(\rho_0(\gamma_1))c(\gamma_2).$$

On the other hand, we have an isomorphism $\mathcal{G}(K[\varepsilon]) \simeq \tilde{\mathfrak{g}} \rtimes \mathcal{G}$ given by

$$B + C\varepsilon \mapsto (CB^{-1}, B).$$

Hence, an element $\rho \in T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$ is a homomorphism $\rho\colon \Gamma \to \tilde{\mathfrak{g}} \rtimes \mathcal{G}$ whose projection to the second factor is ρ_0. In other words, it arises from a map $c\colon \Gamma \to \tilde{\mathfrak{g}}$ such that the map

$$\Gamma \to \tilde{\mathfrak{g}} \rtimes \mathcal{G}, \quad \gamma \mapsto (c(\gamma), \rho_0(\gamma))$$

is a group homomorphism. With the above identification, this translates into the condition

$$c(\gamma_1\gamma_2) = c(\gamma_1) + \mathrm{Ad}(\rho_0(\gamma_1))c(\gamma_2),$$

giving the required isomorphism of $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$ with $Z^1(\Gamma, \tilde{\mathfrak{g}})$. Also notice that for any finite-index subgroup $\Delta \subset \Gamma$ (which is automatically finitely generated), the natural restriction maps $\mathfrak{R}(\Gamma, \mathcal{G}) \to \mathfrak{R}(\Delta, \mathcal{G})$ and $Z^1(\Gamma, \tilde{\mathfrak{g}}) \to Z^1(\Delta, \tilde{\mathfrak{g}})$ induce a commutative diagram

$$\begin{tikzcd}
T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G})) \arrow[r] \arrow[d] & Z^1(\Gamma, \tilde{\mathfrak{g}}) \arrow[d] \\
T_{\rho_0}(\mathfrak{R}(\Delta, \mathcal{G})) \arrow[r] & Z^1(\Delta, \tilde{\mathfrak{g}})
\end{tikzcd}$$

where the horizontal maps are the isomorphisms described above.

Next, let us recall a characterization of the space $B^1(\Gamma, \tilde{\mathfrak{g}})$ of 1-coboundaries that will be used later; for this, we need to consider the action of $\mathcal{G}(K)$ on $R(\Gamma, \mathcal{G})$. Given $\rho_0 \in R(\Gamma, \mathcal{G})$, let $\psi_{\rho_0}\colon \mathcal{G}(K) \to R(\Gamma, \mathcal{G})$ be the orbit map, i.e., the map defined by

$$\psi_{\rho_0}(T) = T\rho_0 T^{-1}, \quad T \in \mathcal{G}(K).$$

Direct computation shows that under the isomorphism $T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G})) \simeq Z^1(\Gamma, \tilde{\mathfrak{g}})$, the image of the differential $(d\psi_{\rho_0})_e\colon T_e(\mathcal{G}) \to T_{\rho_0}(R(\Gamma, \mathcal{G})) \subset T_{\rho_0}(\mathfrak{R}(\Gamma, \mathcal{G}))$ consists of maps $\tau\colon \Gamma \to \tilde{\mathfrak{g}}$ such that there exists $A \in \tilde{\mathfrak{g}}$ with

$$\tau(\gamma) = A - \mathrm{Ad}(\rho_0(\gamma))A$$

for all $\gamma \in \Gamma$, i.e., the image coincides with $B^1(\Gamma, \tilde{\mathfrak{g}})$ [Lubotzky and Magid 1985, Proposition 2.3]. In fact, if $O(\rho_0)$ is the orbit of ρ_0 in $R(\Gamma, \mathcal{G})$ under the action of $\mathcal{G}(K)$, then $B^1(\Gamma, \tilde{\mathfrak{g}})$ can be identified with $T_{\rho_0}(O(\rho_0)) \subset T_{\rho_0}(R(\Gamma, \mathcal{G}))$ [Lubotzky and Magid 1985, Corollary 2.4].

As a special case of the preceding constructions, we can consider the action of $\mathrm{GL}_n(K)$ on $R_n(\Gamma)$. The n-th *character variety* of Γ, denoted $X_n(\Gamma)$, is by definition the (categorical) quotient of $R_n(\Gamma)$ by $\mathrm{GL}_n(K)$; i.e., it is the affine K-variety with coordinate ring $A_n(\Gamma)^{\mathrm{GL}_n(K)}$. Let $\pi : R_n(\Gamma) \to X_n(\Gamma)$ be the canonical map. Then each fiber $\pi^{-1}(x)$ contains a semisimple representation, and moreover, if $\rho_1, \rho_2 \in R_n(\Gamma)$ are semisimple with $\pi(\rho_1) = \pi(\rho_2)$, then $\rho_1 = T\rho_2 T^{-1}$ for some $T \in \mathrm{GL}_n(K)$. In particular, we see that π induces a bijection between the isomorphism classes of semisimple representations and the points of $X_n(\Gamma)$ [Lubotzky and Magid 1985, Theorem 1.28].

We turn to the proof of Theorem 2. In the remainder of this section, Γ will be the elementary subgroup $E(\Phi, R) \subset G(R)$, where Φ is a reduced irreducible root system of rank ≥ 2, G a universal Chevalley–Demazure group scheme of type Φ, and R a finitely generated commutative ring such that (Φ, R) is a nice pair. By recent work of Ershov, Jaikin-Zapirain, and Kassabov [2011], it is known that Γ has Kazhdan's property (T). In particular, Γ is finitely generated and satisfies the condition

for any finite-index subgroup $\Delta \subset \Gamma$, the abelianization $\Delta^{\mathrm{ab}} = \Delta/[\Delta, \Delta]$ is finite (FAb)

[de la Harpe and Valette 1989]. This has the following consequence:

Proposition 4.1 [Rapinchuk 1999, Proposition 2]. *Let Γ be a group satisfying (FAb). For any $n \geq 1$, there exists a finite collection $G_1, \ldots, G_d$ of algebraic subgroups of $\mathrm{GL}_n(K)$ such that for any completely reducible representation $\rho\colon \Gamma \to \mathrm{GL}_n(K)$, the Zariski closure $\overline{\rho(\Gamma)}$ is conjugate to one of the G_i. Moreover, for each i, the connected component G_i° is a semisimple group.*

Thus, if we denote by $R_n(\Gamma)_{ss}$ the set of completely reducible representations $\rho\colon \Gamma \to \mathrm{GL}_n(K)$, we have[3]

$$R_n(\Gamma)_{ss} = \bigcup_{\substack{i\in\{1,\ldots,d\},\\ g\in \mathrm{GL}_n(K)}} gR'(\Gamma, G_i)g^{-1},$$

where for an algebraic subgroup $\mathcal{G} \subset \mathrm{GL}_n(K)$, we set

$$R'(\Gamma, \mathcal{G}) = \{\rho\colon \Gamma \to \mathcal{G} \mid \overline{\rho(\Gamma)} = \mathcal{G}\}.$$

Therefore, letting $\pi : R_n(\Gamma) \to X_n(\Gamma)$ be the canonical map, we obtain that

$$X_n(\Gamma) = \bigcup_{i=1}^{d} \pi(R'(\Gamma, G_i)). \tag{10}$$

[3] Observe that if $\mathcal{G} \subset \mathrm{GL}_n(K)$ is an algebraic subgroup such that $\mathcal{G}^\circ$ is semisimple, then $\mathcal{G}$ is completely reducible; hence, any representation $\rho\colon \Gamma \to \mathrm{GL}_n(K)$ with $\overline{\rho(\Gamma)} = \mathcal{G}$ is completely reducible.

Notice that if $\mathscr{G} \subset \mathrm{GL}_n(K)$ is an algebraic group such that $\mathscr{G}^\circ$ is semisimple, then $R'(\Gamma, \mathscr{G})$ is an open subvariety of $R(\Gamma, \mathscr{G})$. Indeed, let

$$R''(\Gamma, \mathscr{G}) = \{\rho \colon \Gamma \to \mathscr{G} \mid \overline{\rho(\Gamma)} \supset \mathscr{G}^\circ\}.$$

Since $\mathscr{G}^\circ$ is semisimple, $R''(\Gamma, \mathscr{G})$ is easily seen to be an open subvariety in $R(\Gamma, \mathscr{G})$ [Rapinchuk 1998, Lemma 4]. On the other hand, we obviously have

$$R'(\Gamma, \mathscr{G}) = R''(\Gamma, \mathscr{G}) \cap \left(R(\Gamma, \mathscr{G}) \setminus \bigcup_{i=1}^{l} R(\Gamma, \mathscr{H}_i)\right),$$

where $\mathscr{H}_1, \ldots, \mathscr{H}_l$ are the algebraic subgroups of $\mathscr{G}$ such that

$$\mathscr{G} \supsetneq \mathscr{H}_i \supset \mathscr{G}^\circ.$$

Now let $W \subset X_n(\Gamma)$ be an irreducible component of maximal dimension so that $\dim X_n(\Gamma) = \dim W$. Then it follows from (10) that we can find an irreducible component V of some $R'(\Gamma, G_i)$ such that $\overline{\pi(V)} = W$. Since $\pi|_V$ is dominant and separable (as $\operatorname{char} K = 0$), it follows from [Borel 1991, AG 17.3] that there exists $\rho_0 \in V$ that is a simple point (of $R'(\Gamma, G_i)$) such that $\pi(\rho_0)$ is simple and the differential

$$(d\pi)_{\rho_0} \colon T_{\rho_0}(V) \to T_{\pi(\rho_0)}(W) \tag{11}$$

is surjective. Next, let $\psi_{\rho_0} \colon G_i \to R(\Gamma, G_i)$ be the orbit map. By the construction of π, we have $(\pi \circ \psi_{\rho_0})(T) = \pi(\rho_0)$ for any $T \in G_i$, so $d(\pi \circ \psi_{\rho_0})_e = 0$. On the other hand, as we noted above, the image of the differential $(d\psi_{\rho_0})_e$ is the space $B = B^1(\Gamma, \tilde{\mathfrak{g}}_i)$, where $\tilde{\mathfrak{g}}_i$ is the Lie algebra of G_i with Γ-action given by $\mathrm{Ad} \circ \rho_0$. Since ρ_0 is a simple point, it lies on a unique irreducible component of $R'(\Gamma, G_i)$, so it follows that the image of ψ_{ρ_0} (i.e., the orbit of ρ_0) is contained in V. Consequently, (11) factors through

$$T_{\rho_0}(V)/B \to T_{\pi(\rho_0)}(W).$$

Since obviously $\dim_K T_{\rho_0}(V) \le \dim_K T_{\rho_0}(\mathfrak{R}(\Gamma, G_i))$ and

$$T_{\rho_0}(\mathfrak{R}(\Gamma, G_i)) \simeq Z^1(\Gamma, \tilde{\mathfrak{g}}),$$

we therefore obtain that

$$\dim X_n(\Gamma) = \dim W \le \dim_K H^1(\Gamma, \tilde{\mathfrak{g}}_i). \tag{12}$$

Thus, the proof of Theorem 2 is now reduced to considering the following situation. Suppose $\rho_0 \colon \Gamma \to \mathrm{GL}_n(K)$ is a completely reducible representation, set $\mathscr{G} = \overline{\rho_0(\Gamma)}$ (note that the connected component $\mathscr{G}^\circ$ is semisimple), and let $\tilde{\mathfrak{g}}$ be the Lie algebra of $\mathscr{G}$, considered as a Γ-module via $\mathrm{Ad} \circ \rho_0$. We need to give an upper bound on $\dim_K H^1(\Gamma, \tilde{\mathfrak{g}})$. This will be made more precise in Proposition 4.4 below after some preparatory remarks.

First, notice that for the purpose of estimating $\dim_K H^1(\Gamma, \tilde{\mathfrak{g}})$, we may compose ρ_0 with the adjoint representation and assume without loss of generality that the group $\mathscr{G}$ is adjoint. Now, since $\mathscr{G}^\circ$ is semisimple, ρ_0 has a standard description by [Rapinchuk 2011, Theorem 6.7], i.e., there exist a commutative finite-dimensional K-algebra A_0, a ring homomorphism

$$f_0 \colon R \to A_0 \tag{13}$$

with Zariski-dense image, and a morphism of algebraic groups

$$\theta \colon G(A_0) \to \mathscr{G} \tag{14}$$

such that on a suitable finite-index subgroup $\Delta \subset \Gamma$, we have

$$\rho_0|_\Delta = (\theta \circ F_0)|_\Delta, \tag{15}$$

where $F_0 \colon \Gamma \to G(A_0)$ is the group homomorphism induced by f_0. Moreover, it follows from [Rapinchuk 2011, Proposition 5.3] that $\theta(G(A_0)) = \mathscr{G}^\circ$.

Next, let $\mathscr{G}_1, \dots, \mathscr{G}_r$ be the (almost) simple components of $\mathscr{G}^\circ$ [Borel 1991, Proposition 14.10]. Since $\mathscr{G}^\circ$ is adjoint, the product map

$$\mathscr{G}_1 \times \cdots \times \mathscr{G}_r \to \mathscr{G}^\circ$$

is an isomorphism. The following lemma gives a more detailed description of A_0:

Lemma 4.2. *The algebraic ring A_0 is isomorphic to the product $\underbrace{K \times \cdots \times K}_{r \text{ copies}}$.*

Proof. Let J_0 be the Jacobson radical of A_0. Since $\mathscr{G}^\circ$ is semisimple (in particular, reductive), $J_0 = \{0\}$ by [Rapinchuk 2011, Lemma 5.7], and consequently by [Rapinchuk 2011, Proposition 2.20], we have

$$A_0 \simeq K^{(1)} \times \cdots \times K^{(s)},$$

where $K^{(i)} \simeq K$ for all i. Thus, $G(A_0) = G(K^{(1)}) \times \cdots \times G(K^{(s)})$. As we observed above, the map θ is surjective, so since $G(K)$ is an almost simple group, it follows that $s \geq r$. On the other hand, by [Rapinchuk 2011, Theorem 3.1], for each root $\alpha \in \Phi$, there exists an injective map $\psi_\alpha \colon A_0 \to \mathscr{G}$ such that

$$\theta(e_\alpha(a)) = \psi_\alpha(a), \tag{16}$$

where $e_\alpha(A_0)$ is the 1-parameter root subgroup of $G(A_0)$ corresponding to the root α [Rapinchuk 2011, Proposition 4.2]. Now if $s > r$, then θ would kill some simple component $G(K^{(i)})$ of $G(A_0)$. Since $G(K^{(i)})$ intersects each root subgroup $e_\alpha(A_0)$, the maps ψ_α would not be injective, a contradiction. So $s = r$, as claimed. □

Thus, we can write $f_0\colon R \to A_0$ as

$$f_0(t) = (f_0^{(1)}(t), \dots, f_0^{(r)}(t)) \tag{17}$$

for some ring homomorphisms $f_0^{(i)}\colon R \to K$.

Remark 4.3. Notice that for each i, the image $\theta(G(K^{(i)}))$ intersects a unique simple factor of $\mathscr{G}^\circ$, say $\theta(G(K^{(i)})) \cap \mathscr{G}_i \neq \{e\}$, and then $\theta(G(K^{(i)})) = \mathscr{G}_i$. Furthermore, it follows from the proof of Lemma 4.2 that θ is an isogeny, so since char $K = 0$, the differential $(d\theta)_e\colon \mathfrak{g} \to \tilde{\mathfrak{g}}_i$ gives an isomorphism of Lie algebras. In particular, we see that the Lie algebras of all the simple factors $\mathscr{G}_i$ are isomorphic (in fact, they are isomorphic as $G(K)$-modules with $G(K)$ acting via $\mathrm{Ad} \circ \theta$).

To formulate the next result, we need to introduce some notation. Suppose $g\colon R \to K$ is a ring homomorphism. Then we will let $\mathrm{Der}^g(R, K)$ denote the space of K-valued derivations of R with respect to g, i.e., an element $\delta \in \mathrm{Der}^g(R, K)$ is a map $\delta\colon R \to K$ such that for any $r_1, r_2, \in R$,

$$\delta(r_1 + r_2) = \delta(r_1) + \delta(r_2) \quad \text{and} \quad \delta(r_1 r_2) = \delta(r_1) g(r_2) + g(r_1)\delta(r_2).$$

Proposition 4.4. *Suppose $\rho_0\colon \Gamma \to \mathrm{GL}_n(K)$ is a linear representation, and set $\mathscr{G} = \overline{\rho_0(\Gamma)}$. Let $\tilde{\mathfrak{g}}$ denote the Lie algebra of $\mathscr{G}$, and assume $\mathscr{G}^\circ$ is semisimple. Then*

$$\dim_K H^1(\Gamma, \tilde{\mathfrak{g}}) \leq \sum_{i=1}^{r} \dim_K \mathrm{Der}^{f_0^{(i)}}(R, K),$$

where the $f_0^{(i)}$ are the ring homomorphisms appearing in (17).

We first note two facts that will be needed in the proof. Let $\Lambda \subset \Gamma$ be any finite-index subgroup. Then, as we have already seen, the space of 1-cocycles $Z^1(\Lambda, \tilde{\mathfrak{g}})$ can be naturally identified with the tangent space

$$T_{\rho_0}(\mathfrak{R}(\Lambda, \mathscr{G})) = \{\rho \in \mathrm{Hom}(\Lambda, \mathscr{G}(K[\varepsilon])) \mid \mu \circ \rho = \rho_0\}. \tag{18}$$

Also observe that the restriction map

$$\mathrm{res}_{\Gamma/\Lambda}\colon H^1(\Gamma, \tilde{\mathfrak{g}}) \to H^1(\Lambda, \tilde{\mathfrak{g}})$$

is injective. Indeed, since $[\Gamma : \Lambda] < \infty$, the corestriction map

$$\mathrm{cor}_{\Gamma/\Lambda}\colon H^1(\Lambda, \tilde{\mathfrak{g}}) \to H^1(\Gamma, \tilde{\mathfrak{g}})$$

is defined and the composition $\mathrm{cor}_{\Gamma/\Lambda} \circ \mathrm{res}_{\Gamma/\Lambda}$ coincides with multiplication by $[\Gamma : \Lambda]$. Since char $K = 0$, the injectivity of $\mathrm{res}_{\Gamma/\Lambda}$ follows.

Proof of Proposition 4.4. Set

$$X = \mathrm{Der}^{f_0^{(1)}}(R, K) \oplus \cdots \oplus \mathrm{Der}^{f_0^{(r)}}(R, K),$$

and let $\Delta \subset \Gamma$ be the finite-index subgroup appearing in (15). We will show that there exists a linear map $\psi \colon X \to H^1(\Delta, \tilde{\mathfrak{g}})$ such that

$$\operatorname{res}(H^1(\Gamma, \tilde{\mathfrak{g}})) \subset \operatorname{im}(\psi). \tag{19}$$

The proposition then follows from the injectivity of the restriction map.

The map ψ is constructed as follows. Choose derivations $\delta_i \in \operatorname{Der}^{f_0^{(i)}}(R, K)$ for $i = 1, \dots, r$, and let

$$B = \underbrace{K[\varepsilon] \times \cdots \times K[\varepsilon]}_{r \text{ copies}}$$

(with $\varepsilon^2 = 0$). Then

$$f_{\delta_1,\dots,\delta_r} \colon R \to B, \quad s \mapsto (f_0^{(1)}(s) + \delta_1(s)\varepsilon, \dots, f_0^{(r)}(s) + \delta_r(s)\varepsilon)$$

is a ring homomorphism and hence induces a group homomorphism

$$F_{\delta_1,\dots,\delta_r} \colon \Gamma \to G(B)$$

(recall that $\Gamma = E(R) \subset G(R)$). On the other hand, we have

$$G(B) \simeq (\mathfrak{g} \oplus \cdots \oplus \mathfrak{g}) \rtimes (G(K) \times \cdots \times G(K)) \simeq \operatorname{Lie}(G(A_0)) \rtimes G(A_0)$$

and

$$\mathcal{G}(K[\varepsilon]) \simeq \tilde{\mathfrak{g}} \rtimes \mathcal{G},$$

so we can define a group homomorphism $\tilde{\theta} \colon G(B) \to \mathcal{G}(K[\varepsilon])$ by the formula

$$(x, g) \mapsto ((d\theta)_e(x), \theta(g)),$$

where $\theta \colon G(A_0) \to \mathcal{G}$ is the morphism appearing in (14). Notice that since by Remark 4.3, the differential of θ gives a homomorphism

$$(d\theta)_e \colon \mathfrak{g} \to \tilde{\mathfrak{g}}_i$$

for each factor $\mathfrak{g}$ of $\operatorname{Lie}(G(A_0))$, the map $\tilde{\theta}$ can also be described as follows. Let $x_1, \dots, x_r \in \mathfrak{g}$ and $g \in G(A_0)$. Then

$$\tilde{\theta}(x_1, \dots, x_r, g) = \left(\sum_{i=1}^{r} (d\theta)_e(x_i), \theta(g) \right).$$

Now, $\tilde{\theta} \circ F_{\delta_1,\dots,\delta_r}$ is a homomorphism $\Gamma \to \mathcal{G}(K[\varepsilon])$, and in view of (15), we have

$$\mu \circ (\tilde{\theta} \circ F_{\delta_1,\dots,\delta_r}|_\Delta) = \rho_0.$$

It follows from (18) that

$$c_{\delta_1,\dots,\delta_r} := \tilde{\theta} \circ \operatorname{pr} \circ F_{\delta_1,\dots,\delta_r}|_\Delta,$$

where $\mathrm{pr}\colon G(B) \to \mathrm{Lie}(G(A_0))$ is the projection, is an element of $Z^1(\Delta, \tilde{\mathfrak{g}})$. Now put

$$\psi((\delta_1, \ldots, \delta_r)) = [c_{\delta_1,\ldots,\delta_r}],$$

where $[c_{\delta_1,\ldots,\delta_r}]$ denotes the class of $c_{\delta_1,\ldots,\delta_r}$ in $H^1(\Delta, \tilde{\mathfrak{g}})$.

Let us now turn to the proof of the inclusion (19). Suppose $\rho\colon \Gamma \to \mathscr{G}(K[\varepsilon])$ is a homomorphism with $\mu \circ \rho = \rho_0$. By [Rapinchuk 2011, Proposition 2.14, Theorem 3.1], we can associative to ρ a commutative finite-dimensional K-algebra A together with a ring homomorphism $f\colon R \to A$ with Zariski-dense image.

Lemma 4.5. *Let A be the finite-dimensional commutative K-algebra associated to ρ. Then*

$$A \simeq \tilde{K}^{(1)} \times \cdots \times \tilde{K}^{(r)},$$

where, as above, r is the number of simple components of $\mathscr{G}^\circ$ and, for each i, $\tilde{K}^{(i)}$ is isomorphic to either K or $K[\varepsilon]$ (with $\varepsilon^2 = 0$).

Proof. Let J be the Jacobson radical of A. Since the unipotent radical U of $\overline{\rho(\Gamma)}^\circ$ is commutative (which follows from the fact that $\tilde{\mathfrak{g}}$ is the unipotent radical of $\mathscr{G}(K[\varepsilon])$), we have $J^2 = \{0\}$ by [Rapinchuk 2011, Lemma 5.7]. Now by our assumption, $\mu \circ \rho = \rho_0$, where $\mu\colon \mathscr{G}(K[\varepsilon]) \to \mathscr{G}(K)$ is the homomorphism induced by ring homomorphism $K[\varepsilon] \to K,\ \varepsilon \mapsto 0$. In particular, for any root $\alpha \in \Phi$, we have

$$\mu(\rho(e_\alpha(r))) = \rho_0(e_\alpha(r)) \tag{20}$$

for all $r \in R$. Since μ is a morphism of algebraic groups and the algebraic rings A and A_0 associated to ρ and ρ_0, respectively, are by construction the connected components of $\overline{\rho(e_\alpha(R))}$ and $\overline{\rho_0(e_\alpha(R))}$ for any root α [Rapinchuk 2011, Theorem 3.1], it follows that μ induces a surjective map $\nu\colon A \to A_0$. Moreover, since (20) holds for all roots $\alpha \in \Phi$, the construction of the ring operations on A and A_0 given in [Rapinchuk 2011, Theorem 3.1] implies that ν is actually a ring homomorphism. Also notice that since J is commutative and nilpotent, we have $J \subset \ker \nu$ by the definition of μ. On the other hand, the ring A_0 is semisimple by Lemma 4.2, so $J = \ker \nu$. Thus, $A_0 \simeq A/J \simeq K \times \cdots \times K$.

Next, by the Wedderburn–Malcev theorem, we can find a semisimple subalgebra $\tilde{B} \subset A$ such that $A = \tilde{B} \oplus J$ as K-vector spaces and $\tilde{B} \simeq A/J \simeq K \times \cdots \times K$ as K-algebras [Pierce 1982, Corollary 11.6]. Let $e_i \in \tilde{B}$ be the i-th standard basis vector. Since $e_1, \ldots, e_r$ are idempotent and we have $e_1 + \cdots + e_r = 1$ and $e_i e_j = 0$ for $i \neq j$, it follows that we can write $A = \bigoplus_{i=1}^r A_i$, where $A_i = e_i A$. Clearly, $A_i = \tilde{B}_i \oplus J_i$ with $\tilde{B}_i = e_i \tilde{B} \simeq K$ and $J_i = e_i J$; in particular, A_i is a local K-algebra with maximal ideal J_i such that $J_i^2 = \{0\}$. To complete the proof, it obviously suffices to show that $s_i := \dim_K J_i \leq 1$ for all i.

Now, by [Rapinchuk 2011, Proposition 6.5], for each $i = 1, \dots, r$, we have a Levi decomposition

$$G(A_i) = \underbrace{(\mathfrak{g} \oplus \cdots \oplus \mathfrak{g})}_{s_i \text{ copies}} \rtimes G(K),$$

where $\mathfrak{g}$ is the Lie algebra of $G(K)$. Also, by [Rapinchuk 2011, Theorem 6.7], there exists a morphism

$$\sigma\colon G(A) \to \mathcal{G}(K[\varepsilon]) \tag{21}$$

of algebraic groups such that for a suitable subgroup of finite index $\Delta' \subset \Gamma$, we have

$$\rho|_{\Delta'} = \sigma \circ F|_{\Delta'}, \tag{22}$$

where $F\colon \Gamma \to G(A)$ denotes the group homomorphism induced by f. Since $\mu \circ \rho = \rho_0$ and for $\tilde{\Delta} = \Delta \cap \Delta'$, we have

$$\rho_0|_{\tilde{\Delta}} = (\theta \circ F_0)|_{\tilde{\Delta}} \quad \text{and} \quad \rho|_{\tilde{\Delta}} = \sigma \circ F|_{\tilde{\Delta}}$$

by (15) and (22), it follows that the diagram

$$\begin{tikzcd} G(A) \arrow[r, "\sigma"] \arrow[d, "\tilde{\nu}"'] & \mathcal{G}(K[\varepsilon]) \arrow[d, "\mu"] \\ G(A_0) \arrow[r, "\theta"] & \mathcal{G} \end{tikzcd} \tag{23}$$

commutes (where $\tilde{\nu}$ is the homomorphism induced by ν). Now Remark 4.3, together with the definition of ν, implies that $(\theta \circ \tilde{\nu})(G(A_i)) = \mathcal{G}_i$, where $\mathcal{G}_i$ is a simple factor of $\mathcal{G}$. Since $G(A_i)$ coincides with its commutator subgroup [Stein 1971, Corollary 4.4], we obtain that $\sigma(G(A_i))$ is a subgroup of $\mathcal{G}(K[\varepsilon])$ that maps to $\mathcal{G}_i$ under μ and coincides with its commutator, so the fact that the simple factors $\mathcal{G}_1, \dots, \mathcal{G}_r$ of $\mathcal{G}$ commute elementwise implies that $\sigma(G(A_i)) \subset \tilde{\mathfrak{g}}_i \rtimes \mathcal{G}_i$, where $\mathfrak{g}_i$ is the Lie algebra of $\mathcal{G}_i$. On the other hand, by [Rapinchuk 2011, Theorem 3.1], for each root $\alpha \in \Phi$, there exists an injective map $\tilde{\psi}_\alpha\colon A \to \mathcal{G}(K[\varepsilon])$ such that

$$\sigma(e_\alpha(a)) = \tilde{\psi}_\alpha(a), \tag{24}$$

where $e_\alpha(A)$ is the 1-parameter root subgroup of $G(A)$ corresponding to the root α. So since $\tilde{\mathfrak{g}}_i \simeq \mathfrak{g}$ by Remark 4.3, the same argument as in the proof of Lemma 4.2 shows that $s_i \leq 1$. $\square$

For ease of notation, we will view A as a subalgebra of

$$\tilde{A} := \underbrace{K[\varepsilon] \times \cdots \times K[\varepsilon]}_{r \text{ copies}}. \tag{25}$$

Then, using the lemma and the assumption that $\mu \circ \rho = \rho_0$, we can write the homomorphism $f\colon R \to A$ in the form

$$f(t) = (f_0^{(1)}(t) + \delta_1(t)\varepsilon, \ldots, f_0^{(r)}(t) + \delta_r(t)\varepsilon) \tag{26}$$

with $(\delta_1, \ldots, \delta_r) \in X$ and $\delta_i = 0$ for $i = r_2 + 1, \ldots, r$.

To describe the cohomology class corresponding to ρ, we will now need to analyze more closely the morphism σ introduced in (21). First, we note that if $\bar{A} = A/J$ and $G(A, J)$ is the congruence subgroup

$$G(A, J) = \ker(G(A) \to G(\bar{A})),$$

then by [Rapinchuk 2011, Proposition 6.5],

$$G(A) = G(A, J) \rtimes G(\bar{A})$$

is a Levi decomposition of $G(A)$. Now by [Borel 1991, Proposition 11.23], any two Levi subgroups of $(\mathcal{G}(K[\varepsilon]))^\circ$ are conjugate under an element of the unipotent radical $R_u(\mathcal{G}(K[\varepsilon]))^\circ$, which can be identified with $\mathcal{G}^\circ(K[\varepsilon], (\varepsilon)) \simeq \tilde{\mathfrak{g}}$. In our case, we can apply this to the groups $\sigma(G(\bar{A}))$ and $\theta(G(A_0)) = G^\circ$ (where θ is the morphism from (14)) to conclude that $B\theta(G(A_0))B^{-1} = \sigma(G(\bar{A}))$ for some $B \in \mathcal{G}(K[\varepsilon], (\varepsilon)) \simeq \tilde{\mathfrak{g}}$. By direct computation, one sees that for any $X \in \mathcal{G}$ and $B = I + \varepsilon Y \in \mathcal{G}(K[\varepsilon], (\varepsilon))$,

$$BXB^{-1} = (I + \varepsilon(Y - XYX^{-1}))X,$$

which shows that

$$\rho(\gamma) = \sigma(F(\gamma)) = \big((\sigma \circ \mathrm{pr} \circ F)(\gamma) + Y - \mathrm{Ad}(\theta(F_0(\gamma))(Y), \theta(F_0(\gamma)))\big)$$

for all $\gamma \in \tilde{\Delta} = \Delta \cap \Delta'$ (where Δ and Δ' are the finite-index subgroups of Γ appearing in (15) and (22), respectively). Since $\theta(F_0(\gamma)) = \rho_0(\gamma)$ for $\gamma \in \tilde{\Delta}$, we can rewrite this as

$$\rho(\gamma) = (c(\gamma), \rho_0(\gamma)),$$

where

$$c(\gamma) = (\sigma \circ \mathrm{pr} \circ F)(\gamma) + Y - \mathrm{Ad}(\rho_0(\gamma))(Y).$$

Using (18), we obtain $c \in Z^1(\tilde{\Delta}, \tilde{\mathfrak{g}})$. Now let $b_Y \in B^1(\tilde{\Delta}, \tilde{\mathfrak{g}})$ be the 1-coboundary defined by $b_Y(\gamma) = Y - \mathrm{Ad}(\rho_0(\gamma))Y$, and put $\tilde{c} = c - b_Y$ (thus, $\tilde{c}$ and c define the same element of $H^1(\tilde{\Delta}, \tilde{g})$). Then

$$\tilde{c}(\gamma) = (\sigma \circ \mathrm{pr} \circ F)(\gamma)$$

for all $\gamma \in \tilde{\Delta}$. To complete the proof of the proposition, we will need the following:

Lemma 4.6. *Assume that K is an algebraically closed field of characteristic* 0. *Let $\pi : \mathscr{G} \to \mathscr{G}'$ be an isogeny of absolutely almost simple algebraic groups. Let $\mathfrak{g}$ and $\mathfrak{g}'$ denote the Lie algebras of $\mathscr{G}$ and $\mathscr{G}'$, respectively. Set*

$$\mathscr{H} = \mathfrak{g} \rtimes \mathscr{G} \quad \textit{and} \quad \mathscr{H}' = \mathfrak{g}' \rtimes \mathscr{G}',$$

where $\mathscr{G}$ and $\mathscr{G}'$ act on $\mathfrak{g}$ and $\mathfrak{g}'$, respectively, via the adjoint representation. Then for any morphism $\varphi \colon \mathscr{H} \to \mathscr{H}'$ such that $\varphi|_{\mathscr{G}} = \pi$, there exists $a \in K$ such that

$$\varphi(X, g) = (a(d\pi)_e(X), \pi(g)).$$

Proof. Since $\operatorname{char} K = 0$ and $\mathfrak{g}$ and $\mathfrak{g}'$ are simple Lie algebras, the adjoint representations $\mathrm{Ad}\colon \mathscr{G} \to \mathrm{GL}(\mathfrak{g})$ and $\mathrm{Ad}\colon \mathscr{G}' \to \mathrm{GL}(\mathfrak{g}')$ are both irreducible. Let us now view $\mathfrak{g}'$ as a $\mathscr{G}$-module with $\mathscr{G}$ acting via π. Then both $\varphi|_{\mathfrak{g}}$ and $(d\pi)_e$ are $\mathscr{G}$-equivariant homomorphisms of irreducible $\mathscr{G}$-modules. So by Schur's lemma, $\varphi|_{\mathfrak{g}} = a(d\pi)_e$ for some $a \in K$ [Artin 1991, Theorem 9.6]. □

Now, as above, we consider A as a subalgebra of the algebra $\tilde{A}$ appearing in (25); after possible renumbering, we may assume that, in the notation of Lemma 4.5, we have $\tilde{K}^{(i)} \simeq K[\varepsilon]$ for $i = 1, \dots, s$, where $s = \dim_K J(A)$, and $\tilde{K}^{(i)} \simeq K$ for $i = s+1, \dots, r$. We will view $G(A)$ as a subgroup of

$$G(\tilde{A}) \simeq \mathrm{Lie}(G(A_0)) \rtimes G(A_0)$$

and write $G(A_0) = G(K^{(1)}) \times \cdots \times G(K^{(r)})$ and $\mathrm{Lie}(G(A_0)) = \mathfrak{g}_1 \oplus \cdots \oplus g_r$, where $G(K^{(i)}) = G(K)$ and $\mathfrak{g}_i = \mathfrak{g}$ for all i. We will also regard $\sigma : G(A) \to G(K[\varepsilon])$ as a morphism $\sigma\colon G(\tilde{A}) \to \mathscr{G}(K[\varepsilon])$ with $\sigma|_{\mathfrak{g}_i} = 0$ for all $i > s$. Now since by our construction, the cocycles c and $\tilde{c}$ lie in the same cohomology class, we may assume without loss of generality that σ has the form

$$\sigma(x_1, \dots, x_r, g) = (\sigma|_{\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r}(x_1, \dots, x_r), \theta(g))$$

for $(x_1, \dots, x_r, g) \in (\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r) \rtimes G(\bar{A})$. By Remark 4.3, for each factor $G(K^{(i)})$ of $G(A_0)$, the differential $(d\theta)_e \colon \mathfrak{g}_i \to \tilde{\mathfrak{g}}_i$ yields an isomorphism of $G(K)$-modules (with $G(K)$ acting on $\tilde{\mathfrak{g}}_i$ via $\mathrm{Ad} \circ \theta$). Furthermore, since $\sigma|_{G(\bar{A})} = \theta$, the same argument as used in the proof of Lemma 4.5 shows that $\sigma(\mathfrak{g}_i) = \tilde{\mathfrak{g}}_i$ for $i = 1, \dots, s$. Now applying Lemma 4.6 to the restrictions $\sigma|_{\mathfrak{g}_i \rtimes G(K^{(i)})}$ and $((d\theta)_e, \theta)|_{\mathfrak{g}_i \rtimes G(K^{(i)})}$, we get

$$\sigma|_{\mathfrak{g}_i} = a(d\theta)_e|_{\mathfrak{g}_i}$$

for some $a \in K$ (possibly 0). Repeating for all factors shows that we have

$$\sigma|_{\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r}(x_1, \dots x_r) = \sum_{i=1}^{r} a_i (d\theta)_e(x_i)$$

for $(x_1, \dots, x_r) \in \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$. So replacing the element $(\delta_1, \dots, \delta_r)$ in (26) by $(a_1\delta_1, \dots, a_r\delta_r)$, we have

$$\tilde{c}(\gamma) = c_{\delta_1,\dots,\delta_r}(\gamma)$$

for all $\gamma \in \tilde{\Delta}$. Now let $\psi((\delta_1, \dots, \delta_r)) = d_{\delta_1,\dots,\delta_r} \in Z^1(\Delta, \tilde{\mathfrak{g}})$, and let c_ρ be the element of $Z^1(\Gamma, \tilde{\mathfrak{g}})$ corresponding to ρ. It follows that

$$\mathrm{res}_{\Delta/\tilde{\Delta}}(\mathrm{res}_{\Gamma/\Delta}([c_\rho])) = \mathrm{res}_{\Delta/\tilde{\Delta}}([d_{\delta_1,\dots,\delta_r}]),$$

where

$$\mathrm{res}_{\Gamma/\Delta}\colon H^1(\Gamma, \tilde{\mathfrak{g}}) \to H^1(\Delta, \tilde{\mathfrak{g}}) \quad \text{and} \quad \mathrm{res}_{\Delta/\tilde{\Delta}}\colon H^1(\Delta, \tilde{\mathfrak{g}}) \to H^1(\tilde{\Delta}, \tilde{\mathfrak{g}})$$

are the restriction maps. So the injectivity of the restriction maps yields

$$\mathrm{res}_{\Gamma/\Delta}([c_\rho]) = [d_{\delta_1,\dots,\delta_r}],$$

which shows that

$$\mathrm{res}(H^1(\Gamma, \tilde{\mathfrak{g}})) \subset \mathrm{im}(\psi).$$

This completes the proof of the proposition. □

Proof of Theorem 2. In view of (12) and Proposition 4.4, it remains to show that $r \le n$ and to give a bound on the dimension of the space $\mathrm{Der}^g(R, K)$, for any ring homomorphism $g\colon R \to K$, which is independent of g. Notice that $\mathscr{G}^\circ \subset \mathrm{GL}_n(K)$ and $\mathscr{G}^\circ = \mathscr{G}_1 \times \cdots \times \mathscr{G}_r$, so we have

$$n \ge \mathrm{rk}\,\mathscr{G}^\circ = \sum_{i=1}^{r} \mathrm{rk}\,\mathscr{G}_i \ge r$$

as needed. For the second task, we have the following (elementary) lemma:

Lemma 4.7. *Let R be a finitely generated commutative ring, and let d denote the minimal number of generators of R (i.e., the smallest integer such that there exists a surjection $\mathbb{Z}[x_1, \dots, x_d] \twoheadrightarrow R$). Then for any field K and ring homomorphism $g\colon R \to K$, $\dim_K \mathrm{Der}^g(R, K) \le d$. If, moreover, K is a field of characteristic 0, R is an integral domain with field of fractions L, and g is injective, then $\dim_K \mathrm{Der}^g(R, K) \le l$, where l is the transcendence degree of L over its prime subfield.*

Proof. Let $S = \{r_1, \dots, r_d\}$ be a minimal set of generators of R. Since any element $\delta \in \mathrm{Der}^g(R, K)$ is completely determined by its values on the elements of S, the map

$$\delta \mapsto (\delta(r_1), \dots, \delta(r_d))$$

defines an injection $\mathrm{Der}^g(R, K) \to K^d$, so $\dim_K \mathrm{Der}^g(R, K) \le d$ as claimed.

Now suppose that R is a finitely generated integral domain and g is injective. Since $\mathrm{char}\, K = 0$, after possibly localizing R with respect to the multiplicative set $\mathbb{Z} \setminus \{0\}$ (which does not affect the dimension of the space $\mathrm{Der}^g(R, K)$), we can use Noether's normalization lemma to write R as an integral extension of $S = \mathbb{Q}[x_1, \ldots, x_l]$ so that the field of fractions of R is a separable extension of that of S. Combining this with the assumption that g is injective, one easily sees that any derivation δ of R is uniquely determined by its restriction to S [Lang 2002, Chapter VII, Theorem 5.1], so in particular,

$$\dim_K \mathrm{Der}^g(R, K) \leq \dim_K \mathrm{Der}^g(S, K) =: s.$$

On the other hand, the argument given in the previous paragraph shows that $s \leq l$, which completes the proof. □

Remark 4.8. Notice that the estimate $\dim_K \mathrm{Der}^g(R, K) \leq l$ may not be true if g is not injective. Indeed, take $K = \overline{\mathbb{Q}}$, and let $R_0 = \mathbb{Z}[X, Y]$ and $R = \mathbb{Z}[X, Y]/(X^3 - Y^2)$. Furthermore, let

$$f : \mathbb{Z}[X, Y] \to \overline{\mathbb{Q}}, \quad \varphi(X, Y) \mapsto \varphi(0, 0)$$

and $g : R \to \overline{\mathbb{Q}}$ denote the induced homomorphism. The space $\mathrm{Der}^f(R_0, \overline{\mathbb{Q}})$ is spanned by the linearly independent derivations δ_x and δ_y defined by

$$\delta_x(\varphi(X, Y)) = \frac{\partial \varphi}{\partial X}(0, 0) \quad \text{and} \quad \delta_y(\varphi(X, Y)) = \frac{\partial \varphi}{\partial Y}(0, 0),$$

so $\dim_{\overline{\mathbb{Q}}} \mathrm{Der}^f(R_0, \overline{\mathbb{Q}}) = 2$. Now notice that the natural map

$$\mathrm{Der}^g(R, \overline{\mathbb{Q}}) \to \mathrm{Der}^f(R_0, \overline{\mathbb{Q}})$$

is bijective. Indeed, it is obviously injective, and since any $\delta \in \mathrm{Der}^f(R_0, \overline{\mathbb{Q}})$ vanishes on the elements of the ideal $(X^3 - Y^2)R_0$, it is also surjective. Thus, $\dim_{\overline{\mathbb{Q}}} \mathrm{Der}^g(R, \overline{\mathbb{Q}}) = 2$. On the other hand, if L is the fraction field of R, then $l := \mathrm{tr}\deg_{\mathbb{Q}} L$ is 1.

5. Applications to rigidity

In this section, we will show how our results from [Rapinchuk 2011] imply various forms of classical rigidity for the elementary groups $E(\Phi, \mathcal{O})$, where Φ is a reduced irreducible root system of rank > 1 and $\mathcal{O}$ is a ring of algebraic integers (or S-integers) in a number field. It is worth mentioning that all forms of rigidity ultimately boil down to the fact that $\mathcal{O}$ does not admit nontrivial derivations.

To fix notations, let Φ be a reduced irreducible root system of rank > 1, G the universal Chevalley–Demazure group scheme of type Φ, and $\mathcal{O}$ a ring of algebraic S-integers in a number field L such that $(\Phi, \mathcal{O})$ is a nice pair. Furthermore, let $\Gamma = E(\Phi, \mathcal{O})$ be the elementary subgroup of $G(\mathcal{O})$.

Proposition 5.1. *Let $\rho\colon \Gamma \to \mathrm{GL}_m(K)$ be an abstract linear representation over an algebraically closed field K of characteristic* 0. *Then there exist*

(i) *a finite-dimensional commutative K-algebra*

$$A \simeq K^{(1)} \times \cdots \times K^{(r)}$$

with $K^{(i)} \simeq K$ for all i,

(ii) *a ring homomorphism $f = (f^{(1)}, \ldots, f^{(r)})\colon \mathbb{O} \to A$ with Zariski-dense image, where each $f^{(i)}\colon \mathbb{O} \to K^{(i)}$ is the restriction to $\mathbb{O}$ of an embedding $\varphi_i : L \hookrightarrow K$ and $\varphi_1, \ldots, \varphi_r$ are all* distinct, *and*

(iii) *a morphism of algebraic groups $\sigma : G(A) \to \mathrm{GL}_m(K)$*

such that for a suitable subgroup of finite index $\Delta \subset \Gamma$, we have

$$\rho|_\Delta = \sigma|_\Delta.$$

Proof. Let $H = \overline{\rho(\Gamma)}$, where, as before, the bar denotes Zariski closure. We begin by showing that the connected component H° is automatically reductive. Suppose this is not the case, and let U be the unipotent radical of H°. Since the commutator subgroup $U' = [U, U]$ is a closed normal subgroup of H, the quotient $\check{H} = H/U'$ is affine, so we have a closed embedding $\iota\colon \check{H} \to \mathrm{GL}_{m'}(K)$ for some m'. Then $\check{\rho} = \iota \circ \pi \circ \rho$, where $\pi : H \to \check{H}$ is the quotient map, is a linear representation of Γ such that $\overline{\check{\rho}(\Gamma)}^\circ = \check{H}^\circ$ has commutative unipotent radical. So we can now apply [Rapinchuk 2011, Theorem 6.7] to obtain a finite-dimensional commutative K-algebra $\check{A}$, a ring homomorphism $\check{f} : \mathbb{O} \to \check{A}$ (which is injective as any nonzero ideal in $\mathbb{O}$ has finite index) with Zariski-dense image, and a morphism $\check{\sigma} : G(\check{A}) \to \check{H}$ of algebraic groups such that for a suitable finite-index subgroup $\check{\Delta} \subset \Gamma$, we have

$$\check{\rho}|_{\check{\Delta}} = (\check{\sigma} \circ \check{F})|_{\check{\Delta}},$$

where $\check{F}\colon \Gamma \to G(\check{A})$ is the group homomorphism induced by $\check{f}$.

Now let J be the Jacobson radical of $\check{A}$. Since $\check{H}^\circ$ has commutative unipotent radical, $J^2 = \{0\}$ by [Rapinchuk 2011, Lemma 5.7]. We claim that in fact $J = \{0\}$. Indeed, using the Wedderburn–Malcev theorem as in the proof of Lemma 4.5, we can write $\check{A} = \bigoplus_{i=1}^r \check{A}_i$, where for each i, $\check{A}_i = K \oplus J_i$ is a finite-dimensional local K-algebra with maximal ideal J_i such that $J_i^2 = \{0\}$. Then it suffices to show that $J_i = \{0\}$ for all i. So we may assume that $\check{A}$ is itself a local K-algebra of this form. Then, fixing a K-basis $\{\varepsilon_1, \ldots, \varepsilon_s\}$ of J, we have

$$\check{f}(x) = f_0(x) + \delta_1(x)\varepsilon_1 + \cdots + \delta_s(x)\varepsilon_s,$$

where $f_0\colon \mathbb{O} \to K$ is an injective ring homomorphism and $\delta_1, \ldots, \delta_s \in \mathrm{Der}^{f_0}(\mathbb{O}, K)$. On the other hand, since the fraction field of $\mathbb{O}$ is a number field, it follows from

Lemma 4.7 that the derivations $\delta_1, \dots, \delta_s$ are identically zero. So the fact that $\check{f}$ has Zariski-dense image forces $J = \{0\}$. Consequently, $\check{A} \simeq K \times \cdots \times K$.

Now by [Rapinchuk 2011, Proposition 5.3], $\check{\sigma} \colon G(\check{A}) \to \check{H}^\circ$ is surjective, so $\check{H}^\circ$ is semisimple and in particular reductive [Borel 1991, Proposition 14.10]. It follows that $U = [U, U]$ [Borel 1991, Corollary 14.11], and hence, being a nilpotent group, $U = \{e\}$, which contradicts our original assumption. Thus, H° must be reductive, as claimed.

We can now apply [Rapinchuk 2011, Theorem 6.7] to ρ to obtain a finite-dimensional commutative K-algebra A, a ring homomorphism $f \colon \mathbb{O} \to A$ with Zariski-dense image, and a morphism $\sigma \colon G(A) \to H$ of algebraic groups such that for a suitable subgroup of finite index $\Delta \subset \Gamma$, we have

$$\rho|_\Delta = (\sigma \circ F)|_\Delta.$$

Moreover, the fact that H° is reductive implies that $A = K \times \cdots \times K$ [Rapinchuk 2011, Proposition 2.20, Lemma 5.7]. So we can write $f = (f^{(1)}, \dots, f^{(r)})$ for some ring homomorphisms $f^{(1)}, \dots, f^{(r)} \colon \mathbb{O} \to K$. It is easy to see that all of the $f^{(i)}$ are injective, and since L is the fraction field of $\mathbb{O}$, it follows that each homomorphism $f^{(i)}$ is a restriction to $\mathbb{O}$ of an embedding $\varphi_i \colon L \hookrightarrow K$. Finally, since f has Zariski-dense image, all of the φ_i must be *distinct*, completing the proof. $\square$

Keeping the notations of the proposition, we have the following:

Corollary 5.2. *Any representation $\rho \colon \Gamma \to \mathrm{GL}_m(K)$ is completely reducible.*

Proof. By Proposition 5.1, we have $\rho|_\Delta = \sigma|_\Delta$, so since $G(B)$ is a semisimple group and char $K = 0$, $\rho|_\Delta$ is completely reducible. Since Δ is a finite-index subgroup of Γ, it follows that ρ is also completely reducible. $\square$

SS-rigidity and local rigidity. Notice that since by Lemma 4.7 there are no nonzero derivations $\delta \colon \mathbb{O} \to K$, Proposition 4.4 and the estimate given in (12) yield that for $\Gamma = E(\Phi, \mathbb{O})$, we have $\dim X_n(\Gamma) = 0$ for all $n \geq 1$, i.e., Γ is *SS*-rigid. In fact, Corollary 5.2 implies that Γ is *locally rigid*, that is, $H^1(\Gamma, \mathrm{Ad} \circ \rho) = 0$ for *any* representation $\rho \colon \Gamma \to \mathrm{GL}_m(K)$. This is shown in [Lubotzky and Magid 1985], and we recall the argument for the reader's convenience. Let $V = K^m$. It is well known that

$$H^1(\Gamma, \mathrm{End}_K(V, V)) = \mathrm{Ext}^1_\Gamma(V, V)$$

[Lubotzky and Magid 1985, page 37], and $\mathrm{Ext}^1_\Gamma(V, V) = 0$ by Corollary 5.2. But $\mathrm{Ad} \circ \rho$, whose underlying vector space is $M_m(K)$, can be naturally identified as a Γ-module with $\mathrm{End}_K(V, V)$, so $H^1(\Gamma, \mathrm{Ad} \circ \rho) = 0$, as claimed.

Superrigidity (compare [Bass et al. 1967, §16; Margulis 1991, Chapter VII]). Let $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ ($n \geq 3$) and consider an abstract representation $\rho \colon \Gamma \to \mathrm{GL}_m(K)$. There exists a rational representation $\sigma \colon \mathrm{SL}_n(K) \to \mathrm{GL}_m(K)$ such that $\rho|_\Delta = \sigma|_\Delta$ for a suitable finite-index subgroup $\Delta \subset \Gamma$. Indeed, let $f : \mathbb{Z} \to A$ be the homomorphism associated to ρ. Since $A \simeq K^{(1)} \times \cdots \times K^{(r)}$ by Proposition 5.1, we see that f is simply a diagonal embedding of $\mathbb{Z}$ into $K \times \cdots \times K$. But f has Zariski-dense image, so $r = 1$, and the rest follows.

Notice that for a general ring of S-integers $\mathbb{O}$, the algebraic group $G(A)$ that arises in Proposition 5.1 can be described as follows. Let $\mathcal{G} = R_{L/\mathbb{Q}}({}_L G)$, where ${}_L G$ is the algebraic group obtained from G by extending scalars from $\mathbb{Q}$ to L and $R_{L/\mathbb{Q}}$ is the functor of restriction of scalars. Then $\mathcal{G}(K) \simeq G(K) \times \cdots \times G(K)$ with the factors corresponding to all of the distinct embeddings of L into K [Platonov and Rapinchuk 1994, §2.1.2]. The group $G(A)$ is then obtained from $\mathcal{G}(K)$ by simply projecting to the factors corresponding to the embeddings $\varphi_1, \dots, \varphi_r$, so any representation of $E(\Phi, \mathbb{O})$ factors through $\mathcal{G}$.

Remark 5.3. Let us point out that another situation in which $\mathrm{Der}^f(R, K) = 0$ occurs is if K is a field of characteristic $p > 0$ and R is a commutative ring of characteristic p such that $R^p = R$. This allows one to use arguments similar to the ones presented in this section to recover results of Seitz [1997]. Details will be published elsewhere.

Acknowledgments

I would like to thank my advisor Professor G. A. Margulis for suggesting the problems considered in this paper and for a number of helpful discussions. I am indebted to G. Prasad for insightful conversations about pseudoreductive groups during my visit to the University of Michigan. I would also like to thank B. Conrad for detailed remarks on the connections between pseudoreductive groups and Conjecture (BT). Finally, I thank M. Kassabov for useful communications regarding algebraic rings.

I was supported by an NSF Postdoctoral Fellowship in the final stages of the preparation of this paper.

References

[Artin 1991] M. Artin, *Algebra*, Prentice Hall, Englewood Cliffs, NJ, 1991. MR 92g:00001 MR 0788:00001

[Bak and Rehmann 1982] A. Bak and U. Rehmann, "The congruence subgroup and metaplectic problems for $\mathrm{SL}_{n\geq 2}$ of division algebras", *J. Algebra* **78**:2 (1982), 475–547. MR 85j:11041 Zbl 0495.20022

[Bass et al. 1967] H. Bass, J. Milnor, and J.-P. Serre, "Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$)", *Inst. Hautes Études Sci. Publ. Math.* **33** (1967), 59–137. MR 39 #5574 Zbl 0174.05203

[Borel 1991] A. Borel, *Linear algebraic groups*, 2nd ed., Graduate Texts in Mathematics **126**, Springer, New York, 1991. MR 92d:20001 Zbl 0726.20030

[Borel and Tits 1973] A. Borel and J. Tits, "Homomorphismes "abstraits" de groupes algébriques simples", *Ann. of Math.* (2) **97** (1973), 499–571. MR 47 #5134 Zbl 0272.14013

[Conrad et al. 2010] B. Conrad, O. Gabber, and G. Prasad, *Pseudo-reductive groups*, New Mathematical Monographs **17**, Cambridge University Press, 2010. MR 2011k:20093 Zbl 1216.20038

[Ershov et al. 2011] M. Ershov, A. Jaikin-Zapirain, and M. Kassabov, "Property (T) for groups graded by root systems", preprint, 2011. arXiv 1102.0031

[Farb and Dennis 1993] B. Farb and R. K. Dennis, *Noncommutative algebra*, Graduate Texts in Mathematics **144**, Springer, New York, 1993. MR 94j:16001 Zbl 0797.16001

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics **101**, Cambridge University Press, 2006. MR 2007k:16033 Zbl 1137.12001

[Greenberg 1964] M. J. Greenberg, "Algebraic rings", *Trans. Amer. Math. Soc.* **111** (1964), 472–481. MR 28 #3040 Zbl 0135.21503

[Hahn and O'Meara 1989] A. J. Hahn and O. T. O'Meara, *The classical groups and K-theory*, Grundlehren Math. Wiss. **291**, Springer, Berlin, 1989. MR 90i:20002 Zbl 0683.20033

[de la Harpe and Valette 1989] P. de la Harpe and A. Valette, *La propriété (T) de Kazhdan pour les groupes localement compacts*, Astérisque **175**, Société Mathématique de France, Paris, 1989. MR 90m:22001 Zbl 0759.22001

[Humphreys 1975] J. E. Humphreys, *Linear algebraic groups*, Graduate Texts in Mathematics **21**, Springer, New York, 1975. MR 53 #633 Zbl 0325.20039

[van der Kallen 1976] W. van der Kallen, "Injective stability for K_2", pp. 77–154 in *Algebraic K-theory* (Evanston, IL, 1976), edited by M. R. Stein, Lecture Notes in Mathematics **551**, Springer, Berlin, 1976. MR 58 #22243 Zbl 0349.18009

[Kassabov and Sapir 2009] M. Kassabov and M. V. Sapir, "Nonlinearity of matrix groups", *J. Topol. Anal.* **1**:3 (2009), 251–260. MR 2010m:20078 Zbl 1189.20042

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, Amer. Math. Soc. Colloq. Publ. **44**, American Mathematical Society, Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Lam 2001] T. Y. Lam, *A first course in noncommutative rings*, 2nd ed., Graduate Texts in Mathematics **131**, Springer, New York, 2001. MR 2002c:16001 Zbl 0980.16001

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics **211**, Springer, New York, 2002. MR 2003e:00003 Zbl 0984.00001

[Lubotzky and Magid 1985] A. Lubotzky and A. R. Magid, *Varieties of representations of finitely generated groups*, Mem. Amer. Math. Soc. **58**, 1985. MR 87c:20021 Zbl 0598.14042

[Margulis 1991] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergeb. Math. Grenzgeb. (3) **17**, Springer, Berlin, 1991. MR 92h:22021 Zbl 0732.22008

[Matsumoto 1966] H. Matsumoto, "Subgroups of finite index in certain arithmetic groups", pp. 99–103 in *Algebraic groups and discontinuous subgroups* (Boulder, CO, 1965), American Mathematical Society, Providence, RI, 1966. MR 34 #4373 Zbl 0178.35302

[Milnor 1971] J. Milnor, *Introduction to algebraic K-theory*, Annals of Mathematics **72**, Princeton University Press, 1971. MR 50 #2304 Zbl 0237.18005

[Pierce 1982] R. S. Pierce, *Associative algebras*, Graduate Texts in Mathematics **88**, Springer, New York, 1982. MR 84c:16001 Zbl 0497.16001

[Platonov and Rapinchuk 1994] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics **139**, Academic Press, Boston, MA, 1994. MR 95b:11039 Zbl 0841.20046

[Rapinchuk 1998] A. S. Rapinchuk, "On SS-rigid groups and A. Weil's criterion for local rigidity, I", *Manuscripta Math.* **97**:4 (1998), 529–543. MR 99m:20019 Zbl 0920.20004

[Rapinchuk 1999] A. Rapinchuk, "On the finite-dimensional unitary representations of Kazhdan groups", *Proc. Amer. Math. Soc.* **127**:5 (1999), 1557–1562. MR 99h:22004 Zbl 0926.22001

[Rapinchuk 2011] I. A. Rapinchuk, "On linear representations of Chevalley groups over commutative rings", *Proc. Lond. Math. Soc.* (3) **102**:5 (2011), 951–983. MR 2012e:20108 Zbl 1232.20049

[Rapinchuk 2013] I. A. Rapinchuk, "On the character varieties of finitely generated groups", preprint, 2013. arXiv 1308.2692

[Rapinchuk et al. 1996] A. S. Rapinchuk, V. V. Benyash-Krivetz, and V. I. Chernousov, "Representation varieties of the fundamental groups of compact orientable surfaces", *Israel J. Math.* **93** (1996), 29–71. MR 98a:57002 Zbl 0857.14012

[Seitz 1997] G. M. Seitz, "Abstract homomorphisms of algebraic groups", *J. London Math. Soc.* (2) **56**:1 (1997), 104–124. MR 99b:20077 Zbl 0904.20038

[Stein 1971] M. R. Stein, "Generators, relations and coverings of Chevalley groups over commutative rings", *Amer. J. Math.* **93** (1971), 965–1004. MR 48 #437 Zbl 0246.20034

[Stein 1973] M. R. Stein, "Surjective stability in dimension 0 for K_2 and related functors", *Trans. Amer. Math. Soc.* **178** (1973), 165–191. MR 48 #6267 Zbl 0267.18015

[Steinberg 1968] R. Steinberg, *Lectures on Chevalley groups*, Yale University, New Haven, CT, 1968. MR 57 #6215 Zbl 1196.22001

Communicated by Brian Conrad
Received 2012-06-09 Revised 2012-06-15 Accepted 2012-09-07

rapinch@math.harvard.edu *Department of Mathematics, Harvard University, 1 Oxford Street, Cambridge, MA, 02138, United States*

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1725

Betti diagrams from graphs

Alexander Engström and Matthew T. Stamps

The emergence of Boij–Söderberg theory has given rise to new connections between combinatorics and commutative algebra. Herzog, Sharifan, and Varbaro recently showed that every Betti diagram of an ideal with a k-linear minimal resolution arises from that of the Stanley–Reisner ideal of a simplicial complex. In this paper, we extend their result for the special case of 2-linear resolutions using purely combinatorial methods. Specifically, we show bijective correspondences between Betti diagrams of ideals with 2-linear resolutions, threshold graphs, and anti-lecture-hall compositions. Moreover, we prove that any Betti diagram of a module with a 2-linear resolution is realized by a direct sum of Stanley–Reisner rings associated to threshold graphs. Our key observation is that these objects are the lattice points in a normal reflexive lattice polytope.

1. Introduction

A fundamental problem in commutative algebra is to characterize the coarsely graded Betti numbers of the finitely generated graded modules over a fixed polynomial ring. Originating with Hilbert in the 1890s, this task largely eluded mathematicians until 2006, when Boij and Söderberg introduced the following relaxation: Instead of trying to determine whether or not a table of nonnegative integers is the Betti diagram of a module, one should try to determine if some rational scalar of the table is the Betti diagram of a module. This shifted the viewpoint to studying rays in a rational cone and with this new geometric picture, the subject has seen a great deal of progress over the last six years. In particular, the idea led Boij and Söderberg [2008] to conjecture that every Betti diagram of a module can be decomposed in a specific and predictable way. Eisenbud and Schreyer [2009] proved this for Cohen–Macaulay modules, and Boij and Söderberg [2012] later extended that proof to the general setting.

A natural question that arises from Boij–Söderberg theory is the following: If a module is constructed from a combinatorial object, such as the edge ideal of a graph or the Stanley–Reisner ideal of a simplicial complex, can any of the combinatorial properties of that object be seen in the Boij–Söderberg decomposition of the module? Herzog, Sharifan, and Varbaro [Herzog et al. 2012] recently gave an elegant partial

MSC2010: primary 13D02; secondary 05C25.
Keywords: linear resolutions, Boij–Söderberg theory, threshold graphs.

answer to this question for the special case of ideals with k-linear resolutions by showing that every Betti diagram of an ideal with a k-linear minimal resolution can be realized by the Stanley–Reisner ideal of a certain simplicial complex. More specifically, they prove that from the coefficients of a Boij–Söderberg decomposition of a k-linear Betti diagram, one obtains an O-sequence which, by a famous result of Eagon and Reiner along with Macaulay's theorem, yields a simplicial complex with the desired properties. Nagel and Sturgeon [2013] employ a similar approach to show that the k-linear Betti diagrams can be realized with hyperedge ideals of k-uniform Ferrers hypergraphs.

In this paper, we restrict our attention to the case of 2-linear resolutions and give an alternate characterization of the Betti diagrams of ideals with 2-linear minimal resolutions using purely combinatorial means. We show that every Betti diagram from an ideal with a 2-linear resolution is realized by a Stanley–Reisner ring constructed from a threshold graph and that this correspondence is a bijection.

Theorem 4.12. *For every* 2*-linear ideal* I *in* S, *there is a unique threshold graph* T *on* $n+1$ *vertices with* $\beta(S/I) = \beta(\Bbbk[T])$.

Moreover, for any such ideal, we give an efficient algorithm for constructing its corresponding threshold graph that avoids expensive computations like Hochster's formula; rather, we can generate all such Betti diagrams recursively with affine transformations, avoiding operators such as Ext and Tor. Even more interesting, we find that these diagrams are the lattice points of a normal reflexive lattice simplex that is combinatorially equivalent to a simplex of anti-lecture-hall compositions and, from this geometric picture, we prove that any Betti diagram of a *module* with a 2-linear resolution arises from a direct sum of Stanley–Reisner rings constructed from threshold graphs.

Theorem 4.16. *For every finitely generated*, *graded* S*-module* M *with* 2*-linear minimal free resolution and* $\beta_{0,0}(M) = m$, *there is a collection of* m *threshold graphs* $\{T_1, \ldots, T_m\}$, *not necessarily distinct*, *such that* $\beta(M) = \beta(\Bbbk[T_1] \oplus \cdots \oplus \Bbbk[T_m])$.

The paper is organized as follows: In Section 2, we give a quick review of the necessary concepts from commutative algebra and Boij–Söderberg theory. In Section 3, we interpret the main theorem of Boij–Söderberg theory in terms of linear algebra for the special case of modules with k-linear minimal resolutions. We prove our main theorems in Section 4 and conclude with some interesting connections to discrete geometry in Section 5.

2. Preliminaries

We begin with a review of the basic definitions and theorems from Boij–Söderberg theory. For a more detailed introduction, we recommend [Fløystad 2012].

Commutative algebra. Let $\Bbbk$ be a field and $S = \Bbbk[x_1, \dots, x_n]$. For any finitely generated graded S-module M, let M_i denote its graded piece of degree i and let $M(d)$ denote the *twisting* of M by d, that is, the module such that $M(d)_i \cong M_{i+d}$. A *minimal graded free resolution* of M is an exact complex

$$0 \leftarrow M \leftarrow F_0 \leftarrow F_1 \leftarrow \cdots \leftarrow F_l,$$

where each F_i is a graded free S-module of the form

$$\bigoplus_{j \in \mathbb{Z}} S(-j)^{\beta_{i,j}}$$

such that the number of basis elements is minimal and each map is graded.

The value $\beta_{i,j}$ is called the i-th *graded Betti number* of degree j. These numbers are a refinement of the ordinary Betti numbers $\beta_i = \sum_j \beta_{i,j}$ and are independent of the choice of resolution of M, thus yielding an important numerical invariant of M. We often express the graded Betti numbers in a two-dimensional array called the *Betti diagram* of M, denoted by $\beta(M)$. Since $\beta_{i,j} = 0$ whenever $i > j$, it is customary to write $\beta(M)$ such that $\beta_{i,j}$ is in position $(j - i,\ i)$. That is,

$$\beta(M) = \begin{bmatrix} \beta_{0,0} & \beta_{1,1} & \dots & \beta_{l,l} \\ \beta_{0,1} & \beta_{1,2} & \dots & \beta_{l,l+1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{0,r} & \beta_{1,r+1} & \dots & \beta_{l,l+r} \end{bmatrix}.$$

A Betti diagram is called *pure* if every column has at most one nonzero entry, that is, for each $i \in \{0, \dots, l\}$, $\beta_{i,j} \neq 0$ for at most one $j \in \mathbb{Z}$.

Boij–Söderberg theory. Let $\mathbb{Z}^{n+1}_{\text{deg}}$ denote the set of strictly increasing nonnegative integer sequences $\boldsymbol{d} = (d_0, \dots, d_s)$ with $s \leq n$, called *degree sequences*, along with the partial order given by

$$(d_0, \dots, d_s) \geq (e_0, \dots, e_t)$$

whenever $s \leq t$ and $d_i \geq e_i$ for all $i \in \{0, \dots, s\}$. To every $\boldsymbol{d} \in \mathbb{Z}^{n+1}_{\text{deg}}$, we associate a pure Betti diagram $\pi(\boldsymbol{d})$ with entries defined as follows:

$$\pi_{i,j}(\boldsymbol{d}) = \begin{cases} \displaystyle\prod_{k \neq 0,i} \left| \frac{d_k - d_0}{d_k - d_i} \right| & i \geq 0,\ j = d_i, \\ 0 & \text{otherwise.} \end{cases}$$

The main theorem of Boij–Söderberg theory states that the Betti diagram of any graded S-module can be written as a positive rational combination of $\pi(\boldsymbol{d})$'s. It was originally conjectured by Boij and Söderberg [2008], proven for Cohen–Macaulay modules by Eisenbud and Schreyer [2009], and then generalized thus:

Theorem 2.1 [Boij and Söderberg 2012]. *For every graded S-module M, there exists a vector $c \in \mathbb{Q}^p_{\geq 0}$ and a chain of degree sequences $\boldsymbol{d}^1 < \boldsymbol{d}^2 < \cdots < \boldsymbol{d}^p$ in $\mathbb{Z}^{n+1}_{\text{deg}}$ such that*

$$\beta(M) = c_1\pi(\boldsymbol{d}^1) + \cdots + c_p\pi(\boldsymbol{d}^p).$$

The combination in Theorem 2.1 is called a *Boij–Söderberg decomposition* of M and the entries of c are called *Boij–Söderberg coefficients*. This decomposition is not unique in general, but there is a simple algorithm for computing a set of coefficients that satisfy the theorem, see [Fløystad 2012].

3. Betti diagrams of 2-linear resolutions

An ideal I in S is called k-*linear* if $\beta_{i,j}(I) = 0$ whenever $j - i \neq k - 1$. If I is 2-linear, then the Betti diagram of $M = S/I$ looks like

$$\beta(M) = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_s \end{bmatrix}$$

for some $s \leq n$. Our aim is to translate the statement of Theorem 2.1, for S-modules with 2-linear resolutions, into linear algebraic terms. For this, it will be convenient to consider the *reduced Betti vector* $\omega(M) = [\beta_1, \ldots, \beta_s]$ in place of $\beta(M)$.

If M is a 2-linear S-module, then every $\boldsymbol{d}^l$ in Theorem 2.1 is of the form $(0, 2, \ldots, l+1)$. So, let $\pi^l = \pi(\boldsymbol{d}^l)$, ω^l be the reduced Betti vector corresponding to π^l, and Ω be the lower-diagonal $n \times n$ matrix whose l-th row is ω^l. We leave it to the reader to verify the following:

Lemma 3.1. *The matrix Ω is invertible and has ij-entry $\omega^i_j = j\binom{i+1}{j+1}$. Moreover, the ij-entry of Ω^{-1} is $(-1)^{i-j}\frac{1}{i}\binom{i+1}{j+1}$.*

Since any subset of row vectors in Ω forms a chain in $\mathbb{Z}^{n+1}_{\text{deg}}$, we can replace the vector $c \in \mathbb{Q}^p_{>0}$ in Theorem 2.1 with a vector $c \in \mathbb{Q}^n_{\geq 0}$ such that $\sum_i c_i = \beta_{0,0}(M)$.

Theorem 3.2. *For every 2-linear (graded) S-module M with $\beta_{0,0}(M) = m$,*

$$\beta(M) = c_1\pi^1 + \cdots + c_n\pi^n,$$

where $c = \omega(M)\Omega^{-1} \in \mathbb{Q}^n_{\geq 0}$ and $\sum_i c_i = m$.

Remark 3.3. When $\beta_{0,0}(M) = 1$, Theorem 3.2 asserts that $\omega(M)$ is a lattice point in the $(n-1)$-dimensional simplex spanned by row vectors of Ω.

We conclude this section with some classic examples of 2-linear ideals that arise from graph theory. A *graph* G consists of a finite set $V(G)$, called the *vertex set*, and a subset $E(G)$ of $\binom{V(G)}{2}$, called the *edge set*. To simplify notation, we write uv instead of $\{u, v\}$ for each edge in G. For any subset of vertices $W \subset V(G)$, the *induced subgraph* $G[W]$ is the graph with vertex set W and edge set $E(G) \cap \binom{W}{2}$.

If $W = V(G) \setminus S$ for some $S \subseteq V(G)$, we write $G \setminus S$ instead of $G[W]$. A subgraph C of the form $V(C) = \{v_1, \dots, v_l\}$ and $E(C) = \{v_i v_{i+1} \mid 1 \le i < l\} \cup \{v_1 v_l\}$ is called a *cycle* of length l. We say G is *chordal* if it has no induced cycles of length greater than three or, equivalently, if $E(C) \subsetneq E(G[C])$ for every cycle of length greater than three. The elements of $E(G[C]) \setminus E(C)$ are called *chords*. Chordal graphs have many interesting properties that are actively studied in graph theory. For a thorough introduction to graph theory, we recommend [Diestel 2010].

Given a graph G with vertex set $[n+1] = \{1, \dots, n+1\}$, where n is the number of indeterminates in S, let $R = \Bbbk[x_1, \dots, x_{n+1}]$, let

$$I^c(G) = \langle x_i x_j \mid ij \notin E(G) \rangle \subseteq R$$

be the ideal generated by the monomials corresponding to nonedges in G, and let $\Bbbk[G]$ be the quotient $R/I^c(G)$. The knowledgeable reader may observe that $I^c(G)$ is the edge ideal of the complement of G and $\Bbbk[G]$ is the Stanley–Reisner ring of the clique complex of G. The following theorem was first proved by Fröberg [1990] and then by Dochtermann and Engström [2009], using topological combinatorics.

Theorem 3.4. *A graph G is chordal if and only if $I^c(G)$ is 2-linear. Whenever this is the case,*

$$\beta_{i,j}(\Bbbk[G]) = \sum_{W \in \binom{V(G)}{j}} (-1 + \#\,\textit{components of}\ G[W])$$

for $i = j - 1 \ge 1$.

Example 3.5. If G consists of $n+1$ isolated vertices, then the $\binom{n+1}{i+1}$ induced subgraphs of G with $i+1$ vertices each have $i+1$ connected components. Thus, $\beta_{i,i+1}(\Bbbk[G]) = i\binom{n+1}{i+1}$ for each $i \ge 1$.

Example 3.6. If G consists of a complete graph on n vertices plus an isolated vertex v, then the $\binom{n}{i}$ induced subgraphs of G with $i+1$ vertices that contain v each have two connected components and the remaining induced subgraphs of G (with $i+1$ vertices) are connected. Thus, $\beta_{i,i+1}(\Bbbk[G]) = \binom{n}{i}$ for each $i \ge 1$.

Remark 3.7. If we apply Theorems 3.2 and 3.4 to $\Bbbk[G]$ for some chordal graph G, we get a formula that takes the number of connected components of induced subgraphs of G as input and yields a vector $c \in \mathbb{Q}^n_{\ge 0}$, namely $\omega(\Bbbk[G])\Omega^{-1}$, whose entries sum to 1. It is natural to ask what this formula says if G is not chordal. If the entries of c fail to be nonnegative or sum to 1, then we get a certificate that G is not chordal. Since measuring how far a graph is from being chordal is nontrivial from the viewpoint of complexity, one is inclined to ask if this procedure characterizes chordal graphs.

Alas, this turns out to not be the case — there are nonchordal graphs that yield admissible c's — but these *false chordal graphs* seem to be few. Examples of false

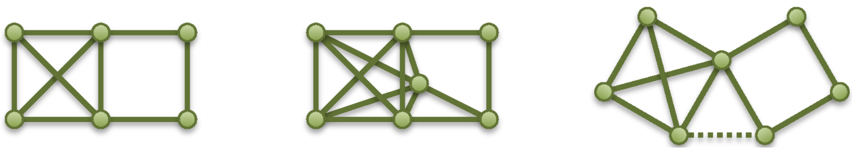


Figure 1. The single false chordal graph on six vertices along with two examples on seven vertices.

chordal graphs on six and seven vertices are illustrated in Figure 1. All other false chordal graphs on seven vertices arise from expanding a (possibly empty) clique of the six-vertex graph or coning over the whole six-vertex graph. We offer some computer-generated statistics on the size of each class of graphs for a given number of vertices:

	1	2	3	4	5	6	7
Chordal	1	2	4	10	27	94	393
False chordal	0	0	0	0	0	1	15
Not chordal	0	0	0	1	7	62	651

4. Betti diagrams from graphs

In this section, we study the Betti diagrams corresponding to a special class of chordal graphs called threshold graphs. We show that threshold graphs on a fixed vertex set have distinct Betti diagrams, that every Betti diagram of a chordal graph is that of a threshold graph on the same number of vertices, that every Betti diagram of an S-algebra with a 2-linear resolution is that of a threshold graph on $n+1$ vertices, and that every Betti diagram of an S-module with a 2-linear resolution is that of a direct sum of Stanley–Reisner rings constructed from threshold graphs on $n+1$ vertices, where n is the number of indeterminates in S.

Betti diagrams from threshold graphs. In a graph G, two vertices are said to be *adjacent* if they are contained in an edge of G. A vertex adjacent to no others is called *isolated* and a vertex adjacent to all others is called *dominating*. For every graph G on n vertices, let G_* be the graph on $n+1$ vertices obtained by adding an isolated vertex to G and, similarly, let G^* be the graph obtained by adding a dominating vertex to G. A graph G is called *threshold* if it can be constructed from a single vertex and a sequence of the operations * and $_*$. It is well known that if G is chordal, then so are G_* and G^*, and thus, all threshold graphs are chordal. We refer to Mahadev and Peled [1995] for a survey that includes the following lemma:

Lemma 4.1. *There are 2^n threshold graphs on $n+1$ vertices. Moreover, every threshold graph is determined by a unique sequence of * and $_*$ operations.*

The Betti diagram of a threshold graph can be constructed recursively in a similar manner to the graph itself. As such, we can quickly calculate the Betti diagram of a threshold graph without the computations in Theorem 3.4.

Proposition 4.2. *If G is a chordal graph on n vertices, then*

$$\omega(\Bbbk[G^*]) = [\omega(\Bbbk[G]) \mid 0] \quad \textit{and} \quad \omega(\Bbbk[G_*]) = \omega(\Bbbk[G])\Lambda + \eta_n, \tag{1}$$

where Λ is the $(n-1)\times n$-matrix whose (i, j) position is 1 if $i = j$ or $j-1$ and 0 otherwise, and η_n is the vector whose i-th entry is $\binom{n}{i}$.

Proof. This is a simple application of Theorem 3.4. For the first part, any subset of vertices containing the dominating vertex in G^* spans a connected graph and therefore, the only nonzero parts of $\omega(\Bbbk[G^*])$ come from $\omega(\Bbbk[G])$. For the second part, we consider whether or not a subset of vertices in G_* contains the isolated vertex v: The induced subgraphs that *do not* contain v contribute $[\omega(\Bbbk[G]) \mid 0]$ to $\omega(\Bbbk[G_*])$ while those that *do* contain v contribute $[0 \mid \omega(\Bbbk[G])] + \eta_n$. □

As a corollary, we find that distinct threshold graphs on a fixed number of vertices have distinct Betti diagrams.

Corollary 4.3. *If T and T' are threshold graphs on the same number of vertices and $\omega(\Bbbk[T]) = \omega(\Bbbk[T'])$, then $T \cong T'$.*

Proof. For any chordal graph G on k vertices, $\omega_{k+1}(\Bbbk[G_*]) \neq \omega_{k+1}(\Bbbk[G^*]) = 0$ by Proposition 4.2. Therefore, since distinct threshold graphs have distinct sequences of $_*$ and * (Lemma 4.1), they must also have distinct Betti diagrams. □

Betti diagrams from chordal graphs. Next, we show that every Betti diagram from a chordal graph arises as the Betti diagram of a threshold graph on the same number of vertices. Moreover, for a given chordal graph, we present an efficient algorithm for constructing its "threshold representative".

Let $\sim_\beta$ be the equivalence relation for graphs on $[n+1]$ defined by

$$G \sim_\beta H \text{ if and only if } \beta(\Bbbk[G]) = \beta(\Bbbk[H])$$

and let $[G]_\beta$ denote the equivalence class of G with respect to $\sim_\beta$. For a chordal graph G on $n+1$ vertices, a threshold graph T (on $n+1$ vertices) is called a *threshold representative* of G if $T \in [G]_\beta$. The next theorem follows from the notion of *algebraic shifting* and can be pieced together from results in [Goodarzi and Yassemi 2012; Klivans 2007; Woodroofe 2011], but we offer a purely graph-theoretic proof instead.

Theorem 4.4. *Every chordal graph G has a unique threshold representative T.*

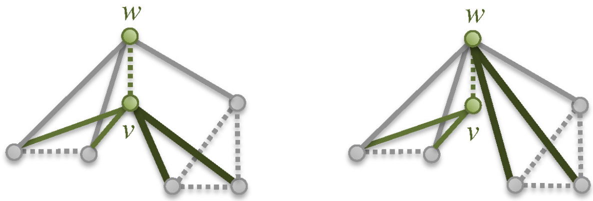


Figure 2. A comparison of a graph G (left) with $G_{v\to w}$ (right).

We proceed with some new machinery. For a graph G with $v, w \in V(G)$, we define a new graph $G_{v\to w}$ on $V(G)$ with

$$E(G_{v\to w}) := \big(E(G) \setminus \{uv \mid u \in N(v; w)\}\big) \cup \big\{uw \mid u \in N(v; w)\big\},$$

where $N(x) = \{y \in V(G) \mid xy \in E(G)\}$ is the *neighborhood* of a vertex x and $N(v; w) = N(v) \setminus \big(\{w\} \cup N(w)\big)$. See Figure 2.

Lemma 4.5. *Let G be a chordal graph.*

(1) *If G is connected with $vw \in E(G)$, then $G_{v\to w}$ is chordal.*

(2) *If G is disconnected with $v, w \in V(G)$ in separate components, then $G_{v\to w}$ is chordal.*

Proof. For each part, we suppose C is a cycle with length $l \geq 4$ in $G' = G_{v\to w}$ and show that C has a chord in G'.

In (1), if $w \notin V(C)$, then $C \subseteq G$ since the only new edges of G' contain w and therefore C has at least one chord in G. If every chord of C in G is removed in G', then they must each contain v and thus $G[V(C \setminus v) \cup w]$ is an induced cycle, which is a contradiction. If $w \in V(C)$, $v \notin V(C)$, and C does not have a chord in G', then $G[V(C) \cup v]$ is an induced cycle since $N(v) \subseteq N(w)$ in G', another contradiction. If $v, w \in V(C)$, then $vw \in E(C)$ and xw is a chord of C in G', where x is the other neighbor of v in C, since $N(v) \subseteq N(w)$ in G'.

In (2), if $w \notin V(C)$, then C contains a chord in $G \setminus w = G' \setminus w \subseteq G'$. So suppose $w \in V(C)$ and C has no chord in G'. Then $G[V(C \setminus w)]$ is contained in the connected component of either v or w in G. If the former is true, then $G[V(C \setminus w) \cup v]$ is an induced cycle and if the latter is true, then C itself is an induced cycle in G, both of which are contradictions. □

For a graph H with $W \subseteq V(H)$, let $\kappa_H(W)$ denote the number of connected components in $H[W]$.

Lemma 4.6. *Let G be a chordal graph.*

(1) *If G is connected with $vw \in E(G)$, then $G_{v\to w} \in [G]_\beta$.*

(2) *If G is disconnected with $v, w \in V(G)$ in separate components, then $G_{v\to w}$ is in $[G]_\beta$.*

Proof. This is a straightforward application of Theorem 3.4 after we make the following calculations. For each part, let $G' = G_{v \to w}$ and $W \subseteq V(G)$.

In (1), if $v, w \notin W$, then $\kappa_G(W) = \kappa_{G'}(W)$ since $G \setminus \{v, w\} = G' \setminus \{v, w\}$ and if $v, w \in W$, then $\kappa_G(W) = \kappa_{G'}(W)$ because the component in $G[W]$ containing v and w spans the same set of vertices as that of $G'[W]$. For the remaining subsets of $V(G)$, we prove that $\kappa_G(W \cup v) + \kappa_G(W \cup w) = \kappa_{G'}(W \cup v) + \kappa_{G'}(W \cup w)$ for every $W \subseteq V(G) \setminus \{v, w\}$. Let $m_\circ(W)$, $m_w(W)$, and $m_v(W)$ denote the number of connected components of $G[W]$ that do not contain any elements of $N(v) \cup N(w)$, $N(v) \setminus N(w)$, and $N(w) \setminus N(v)$, respectively. It is straightforward to check that $\kappa_G(W \cup v) = 1 + m_\circ(W) + m_w(W)$, $\kappa_G(W \cup w) = 1 + m_\circ(W) + m_v(W)$, $\kappa_{G'}(W \cup v) = 1 + m_\circ(W) + m_v(W) + m_w(W)$, and $\kappa_{G'}(W \cup w) = 1 + m_\circ(W)$.

In (2), we record the difference between $\kappa_G(W)$ and $\kappa_{G'}(W)$. If $v, w \notin W$, then $\kappa_G(W) = \kappa_{G'}(W)$ since $G \setminus \{v, w\} = G' \setminus \{v, w\}$. If $v, w \in W$, then $\kappa_G(W) = \kappa_{G'}(W)$ because every vertex in the component of v in $G[W]$ gets moved to the component of w in $G'[W]$. If $v \in W$ and $w \notin W$, then $\kappa_G(W) = \kappa_{G'}(W) - 1$. If $w \in W$ and $v \notin W$, then $\kappa_G(W) = \kappa_{G'}(W) + 1$. $\square$

Proof of Theorem 4.4. We induct on $|V(G)|$. Let G be a chordal graph on n vertices and fix a vertex $v \in V(G)$. We will apply the operations ${}_{v \to w}$ or ${}_{w \to v}$ to G to a get a graph where v is either dominating or isolated.

If G is connected and v is not dominating, then for any vertex $u \in G$ with $d(u, v) = 2$, let $w \in N(v) \cap N(u)$ and replace G with $G_{w \to v}$. Repeat this until v is a dominating vertex, that is, there are no more elements u with $d(v, u) = 2$. The process terminates since G is finite and connected. By Lemma 4.5, the graph G is chordal at every step and by Lemma 4.6, its Betti diagram stays fixed. Since v is dominating and $G \setminus v$ is chordal (being an induced subgraph of a chordal graph), $\beta(\Bbbk[G]) = \beta(\Bbbk[G \setminus v])$. So, by induction, there is a unique (up to isomorphism) threshold graph T such that $\beta(\Bbbk[T^*]) = \beta(\Bbbk[T]) = \beta(\Bbbk[G \setminus v]) = \beta(\Bbbk[G])$.

If G is disconnected, let $w \in V(G)$ be in a separate component in G from v. By Lemmas 4.5 and 4.6, $G_{v \to w}$ is chordal and $\beta(\Bbbk[G]) = \beta(\Bbbk[G_{v \to w}])$; by induction, there exists a unique (up to isomorphism) threshold graph $T \in [G_{v \to w} \setminus v]_\beta$. Thus, $T_* = T \cup \{\alpha\} \in [G]_\beta$ and $\beta(\Bbbk[T_*]) = \beta(\Bbbk[G])$. $\square$

Remark 4.7. The algorithm presented in the proof of Theorem 4.4 is fast. A crude analysis of the complexity is as follows: For each vertex of G, we decompose G into its connected components which takes $O(|V(G)| + |E(G)|)$ and then we repeatedly apply the operations ${}_{v \to w}$ or ${}_{w \to v}$; by amortized analysis, this takes only $O(|E(G)|)$ since each edge is moved at most once. Thus, the total complexity is $O(|V(G)|(|V(G)| + |E(G)|)) \approx O(|V(G)|^3)$. The authors suspect that a more thorough analysis would yield a complexity of $O(|V(G)|^2)$, which is the best one could hope for with this problem.

As simple corollaries of Theorem 4.4, we recover two special classes of graphs that are invariant under β.

Corollary 4.8. *If G is a tree on $n+1$ vertices, then $\beta_{i,i+1}(\Bbbk[G]) = i\binom{n}{i+1}$.*

Proof. Since G has exactly n edges and ${}_{v\to w}$ preserves the number of edges in G, the procedure outlined in the proof of Theorem 4.4 yields a threshold representative T of G that is a star on $n+1$ vertices, that is, a single dominating vertex v and no other edges. Therefore, $T \setminus v$ consists of n isolated points and, by Proposition 4.2 and Example 3.5, $\beta_{i,i+1}(\Bbbk[G]) = \beta_{i,i+1}(\Bbbk[T]) = \beta_{i,i+1}(\Bbbk[T \setminus v]) = (i)\binom{n}{i+1}$. □

The graph from a triangulation of a polygon is called *maximally outerplanar*.

Corollary 4.9. *If G is a maximal outerplanar graph on $n+1$ vertices, then $\beta_{i,i+1}(\Bbbk[G]) = i\binom{n-1}{i+1}$.*

Proof. By Theorem 4.4, the threshold representative T of G consists of a dominating vertex v and a path on $V(T)\setminus v$. In particular, $T \setminus v$ is a tree on n vertices. The result now follows from Proposition 4.2 and Corollary 4.8. □

Betti diagrams of algebras and modules. Here we present the main results of the paper — that every Betti diagram from a 2-linear ideal in S arises from a Stanley–Reisner ring of a threshold graph on $n+1$ vertices and that every Betti diagram from an S-module with a 2-linear resolution arises from a direct sum of Stanley–Reisner rings constructed from threshold graphs on $n+1$ vertices.

To begin, we establish bijections between the set of threshold graphs on $n+1$ vertices, the set of Betti diagrams from 2-linear ideals in S, and the set of anti-lecture-hall compositions of length n bounded above by 1. An integer sequence $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ of the form

$$t \geq \frac{\lambda_1}{1} \geq \frac{\lambda_2}{2} \geq \cdots \geq \frac{\lambda_n}{n} \geq 0$$

is called *anti-lecture-hall composition of length n* bounded above by t. These sequences were introduced in [Corteel and Savage 2003] and are a well-studied variation of the *lecture hall partitions* in [Bousquet-Mélou and Eriksson 1997a; 1997b]. For our purposes, we only need this result of Corteel, Lee, and Savage:

Theorem 4.10 ([Corteel et al. 2005]). *There are $(t+1)^n$ anti-lecture-hall compositions of length n bounded above by t.*

We remark that $\Bbbk[G] = R$ if G is the complete graph on $n+1$ vertices, so we shall ignore that graph for the rest of the paper.

Proposition 4.11. *The set of noncomplete threshold graphs on $n+1$ vertices, the set of Betti diagrams of quotients of S by 2-linear ideals, and the set of anti-lecture-hall compositions of length n with $\lambda_1 = 1$ are in bijective correspondence.*

Proof. By Lemma 4.1 and Corollary 4.3, there are $2^n - 1$ noncomplete threshold graphs on $n+1$ vertices, each of which corresponds to a distinct Betti diagram. It suffices to show that the Betti diagrams of quotients of S by 2-linear ideals inject into the anti-lecture-hall compositions of length n with $\lambda_1 = 1$, since by Theorem 4.10, there are exactly $2^n - 1$ of them.

Let I be a 2-linear ideal in S and let Ψ be the unimodular matrix with ij-entry equal to $\binom{i-1}{j-1}$. Then there exists $\lambda = [\lambda_1, \ldots, \lambda_n] \in \mathbb{Z}^n$ such that $\omega(S/I) = \lambda\Psi$. By Theorem 3.2, we have

$$\lambda\Psi\Omega^{-1} = [c_1, \ldots, c_n] \in \mathbb{Q}^n_{\geq 0}$$

with $\sum_{i=1}^n c_i = 1$. We leave it to the reader to verify that $\Psi \cdot \Omega^{-1}$ has ij-entry $1/i$ if $i = j$, $-1/i$ if $i = j+1$, and 0 otherwise. Thus, $c_i = \lambda_i/i - \lambda_{i+1}/(i+1)$ for all $i \in [n-1]$ and $c_n = \lambda_n/n$. In particular, we get

$$1 = \sum_{i=1}^{n} c_i = \frac{\lambda_1}{1} \geq \frac{\lambda_2}{2} \geq \cdots \geq \frac{\lambda_n}{n} = c_n \geq 0$$

and hence, λ is an anti-lecture-hall composition with $\lambda_1 = 1$. □

The first part of our main theorem is a simple corollary of Proposition 4.11. In particular, it asserts that the injection in Proposition 4.2 is in fact a bijection.

Theorem 4.12 (Main Theorem, Part 1). *For every* 2*-linear ideal I in S, there is a unique threshold graph T on $n+1$ vertices with $\beta(S/I) = \beta(\Bbbk[T])$.*

Remark 4.13. For a given 2-linear ideal I in S, it is easy to construct the graph T realizing its Betti diagram.

Example 4.14. To illustrate Theorem 4.12 at work, consider the ideal

$$I = \langle x_1^2, x_1x_2, x_1x_3, x_1x_4, x_2^2, x_1x_5 + x_2x_4, x_4^2 \rangle \subseteq S = \Bbbk[x_1, \ldots, x_5].$$

Then

$$\beta(S/I) = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 7 & 11 & 6 & 1 & 0 \end{bmatrix}.$$

In order to find a threshold graph T on six vertices whose Betti diagram is $\beta(S/I)$, we sequentially apply the inverses of the affine transformations in Proposition 4.2 depending on whether or not the sequences end in 0. (We leave it to the reader to verify that the inverse of Λ in Proposition 4.2 is the $n \times (n-1)$-matrix whose (i, j) position is $(-1)^{i+j}$ if $i \leq j$ and 0 otherwise.)

$$[7, 11, 6, 1, 0] \xrightarrow{-^*} [7, 11, 6, 1] \xrightarrow{\Lambda^{-1}} [3, 2, 0] \xrightarrow{-^*} [3, 2] \xrightarrow{\Lambda^{-1}} [1] \xrightarrow{\Lambda^{-1}} [0]$$

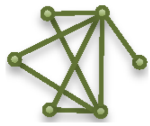


Figure 3. The threshold graph T on six vertices with $\omega(\Bbbk[T]) = [7, 11, 6, 1, 0]$.

From this, we see that $\beta(S/I) = \beta(\Bbbk[T])$, where T is the threshold graph with sequence ${}_{**}{}^{*}{}_{*}{}^{*}$ drawn in Figure 3.

For the rest of the paper, we take a more geometric approach. Specifically, we make use of the fact (Remark 3.3) that the reduced Betti vectors of these diagrams are lattice points in the $(n-1)$-dimensional lattice simplex P_n spanned by the row vectors of Ω. Illustrations of P_1 through P_4, labeled by reduced Betti vectors, Boij–Söderberg coefficients, truncated coordinates (see Section 5), and corresponding chordal graphs are shown in Figures 4 and 5, with the threshold graphs colored dark green. Notice that each P_n contains two copies of P_{n-1}, colored blue and red, corresponding to the first and second equations, respectively, in (1) (see Proposition 4.2).

We continue with some standard definitions from discrete geometry. The integer points $\mathbb{Z}^d \subseteq \mathbb{R}^d$ form a *lattice*. The integer points of a polytope are its *lattice points*

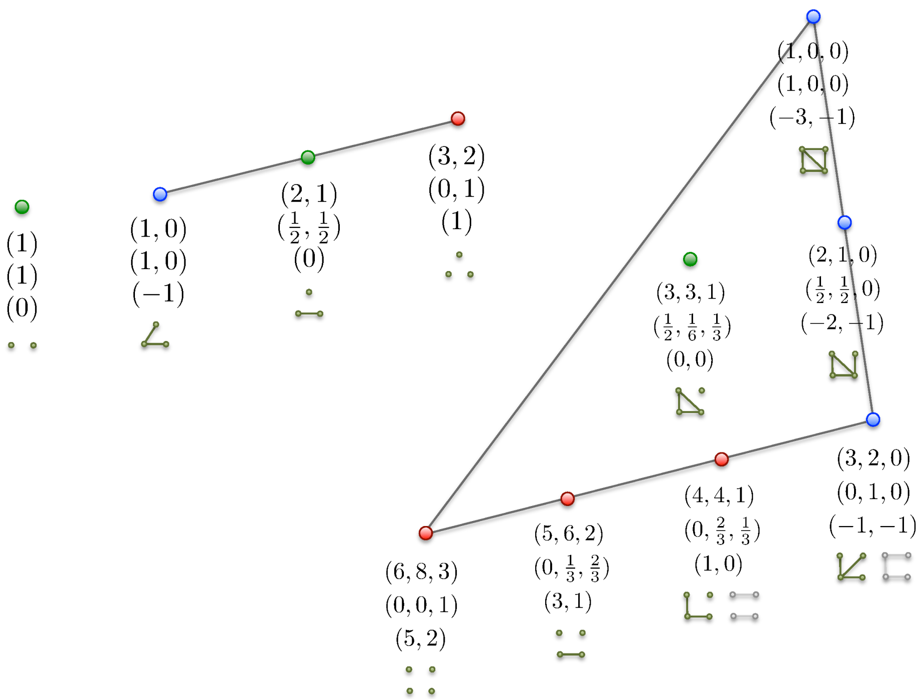


Figure 4. The lattice polytopes P_1 (left), P_2 (middle), and P_3 (right).

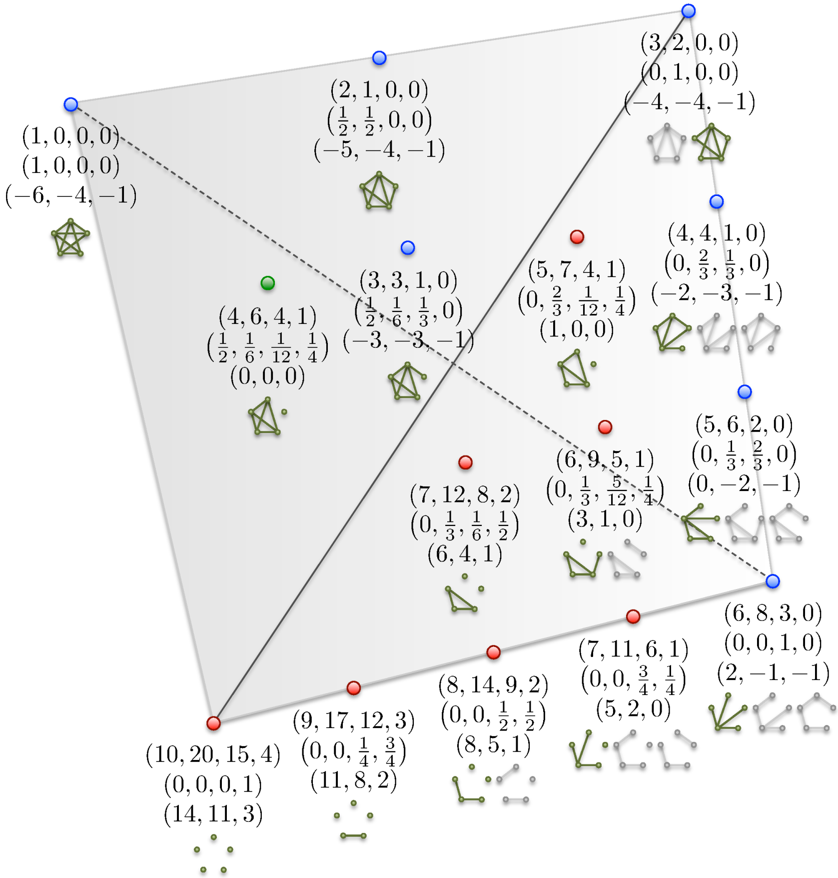


Figure 5. The lattice polytope P_4.

and a polytope is called a *lattice polytope* if all its vertices are lattice points. For a polytope P with vertices $\{v_1, \dots, v_s\}$ and $t \in \mathbb{N}$, let tP denote the t-th *dilation* of P, that is, the polytope attained by taking the convex hull of the points $\{t \cdot v_1, \dots, t \cdot v_s\}$, let $S_P \subseteq \mathbb{Z}^{d+1}$ denote the semigroup generated by $\big\{[1, p_1, \dots, p_d] : (p_1, \dots, p_d) \in P \cap \mathbb{Z}^d\big\}$, and let $\mathrm{gp}(S_P)$ be the smallest group containing S_P, that is the group of differences in S_P. We say P is *normal* if $x \in \mathrm{gp}(S_P)$ such that $s \cdot x \in S_P$ for some $s \in \mathbb{N}$ implies that $x \in S_P$. We refer to [Barvinok 2002; Bruns et al. 1997] for questions on lattice polytopes.

Proposition 4.15. *The lattice simplex P_n is normal for each $n \in \mathbb{N}$.*

Proof. It is straightforward to check that the anti-lecture-hall compositions of length n bounded above by 1 are the lattice points of the n-dimensional lattice simplex spanned by $(0, \dots, 0)$ and the compositions $\lambda^l = (1, 2, \dots, l, 0, \dots, 0)$ for $l \in [n]$. Let Q_n be the facet spanned by the λ^l. Since normality is preserved under unimodular transformations, we prove that Q_n is normal and apply Ψ from the proof of Proposition 4.11.

To begin, we must truncate the coordinates of Q_n since it is an $(n-1)$-dimensional simplex. Removing the first coordinate yields the simplex with vertices $(0, \dots, 0)$ and $(2, 3, \dots, l, 0, \dots, 0)$ for $l \in [n]$. Then S_{Q_n} is the set of all anti-lecture-hall compositions and $\mathrm{gp}(S_{Q_n}) = \mathbb{Z}^n$. From here it is clear that if $\lambda \in \mathbb{Z}^n$ and $s \cdot \lambda \in S_{Q_n}$ for some $s \in \mathbb{N}$, then $\lambda \in S_{Q_n}$. Hence, Q_n is normal. □

A convenient consequence of normality is that every lattice point in the t-th dilation of a normal polytope P can be written as a sum of t, not necessarily distinct, lattice points in P. With that, we can prove the second part of our main theorem.

Theorem 4.16 (Main Theorem, Part 2). *For every finitely generated, graded S-module M with a 2-linear minimal free resolution and $\beta_{0,0}(M) = m$, there is a collection of m threshold graphs $\{T_1, \dots, T_m\}$, not necessarily distinct, such that $\beta(M) = \beta(\Bbbk[T_1] \oplus \cdots \oplus \Bbbk[T_m])$.*

Proof. By Theorem 3.2, $\omega(M)$ is a lattice point in mP_n and is a sum of m lattice points $p_1, \dots, p_m$ in P_n, by Proposition 4.15. Applying Theorem 4.12 yields a threshold graph T_i such that $p_i = \omega(\Bbbk[T_i])$ for each $i \in [m]$, and thus,

$$\beta(M) = \beta(T_1) + \cdots + \beta(T_m) = \beta(\Bbbk[T_1] \oplus \cdots \oplus \Bbbk[T_m]). \qquad \square$$

Remark 4.17. The decomposition in Theorem 4.16 is often not unique. So in the more general setting of modules, we do not know how to construct the family of trees representing a given Betti diagram as we do in the special case of algebras, see Theorem 4.12 and Example 4.14.

5. The geometry of P_n and Q_n

In the previous section, we used the geometry of the lattice simplex P_n of reduced Betti vectors of 2-linear ideals in S (or equivalently, the lattice simplex Q_n of nonzero anti-lecture-hall compositions of length n) to prove algebraic statements about Betti diagrams of algebras and modules with 2-linear resolutions, but these polytopes have many other beautiful geometric properties which make them interesting on their own. In this section, we take the opportunity to showcase a few of these properties. Specifically, we remark that P_n has a simple Ehrhart polynomial, by a result from [Corteel et al. 2005], and we prove that P_n is reflexive.

Given a d-dimensional polytope P, let $\mathrm{Ehr}_P(t)$ denote the number of lattice points in tP. It is well known that $\mathrm{Ehr}_P(t)$ is a degree d polynomial in t, called the *Ehrhart polynomial* of P, with constant term 1 and leading coefficient equal to the volume of P, and that Ehrhart polynomials are preserved under unimodular transformations. For an introduction to Ehrhart theory, see [Beck and Robins 2007].

Theorem 5.1. *For every $n, t \in \mathbb{N}$, $\operatorname{Ehr}_{P_n}(t) = \operatorname{Ehr}_{Q_n}(t) = (t+1)^n - t^n$.*

Proof. Since the matrix Ψ^{-1} in the proof of Proposition 4.11 is unimodular, we know that $\operatorname{Ehr}_{P_n}(t) = \operatorname{Ehr}_{Q_n}(t)$. So, let $A_n(t)$ denote the number of anti-lecture-hall compositions of length n with $\lambda_1 \leq t$. Theorem 4.10 gives us $A_n(t) = (t+1)^n$. Since every point in the tQ_n satisfies $\lambda_1 = t$, it follows immediately that

$$\operatorname{Ehr}_{P_n}(t) = \operatorname{Ehr}_{Q_n}(t) = A_n(t) - A_n(t-1) = (t+1)^n - t^n. \qquad \square$$

Next, we prove that P_n is reflexive. For this, we need the concept of a dual (or polar) of a polytope, but restrict to the case of simplices, since those are the only polytopes we consider.

Definition 5.2. Let the vertices of a d-simplex P be recorded as the rows of the $d \times (d-1)$ matrix M and let M^* be the $(d-1) \times d$ matrix such that MM^* has value -1 everywhere outside the diagonal. The d-simplex whose vertices are the columns of M^* is the *dual* P^* of P.

If P is a lattice polytope containing 0 as an interior point such that P^* is also a lattice polytope, then P and P^* are called *reflexive*. These polytopes have several interesting properties and characterizations, for instance, a lattice polytope P is reflexive if and only if its only interior lattice point is 0 and if u and v are two lattice points on the boundary of P, then either u and v are on the same facet, or $u+v$ is in P. This is an important concept with interesting connections to geometry and theoretical physics. For an exposition suitable for researchers with a background in discrete mathematics, we refer to Batyrev and Nill [2008].

Because P_n is an $(n-1)$-dimensional simplex with coordinates in $\mathbb{Z}^n$, for each lattice point $p \in P_n$, we define

$$p_t = [p_1, \ldots, p_{n-1}] := [p - \eta_n]_{2 \leq i \leq n} = \left[p_2 - \tbinom{n}{2}, \ldots, p_n - \tbinom{n}{n}\right]$$

to be the *truncated coordinates* of p in P_n.

Theorem 5.3. *The simplex P_n realized in the truncated coordinates is a reflexive lattice polytope.*

Proof. We begin by removing the left-most column of Ω to get the $n \times (n-1)$ matrix Ω'_n. Then the truncated coordinates of P_n are the rows of $\Omega_n = \Omega'_n - \eta_n \mathbf{1}_n$. More explicitly, the ij-entry of Ω'_n is $(j+1)\binom{i+1}{j+2}$ and the j entry of η_n is $\binom{n}{j+1}$.

The dual of P_n, in truncated coordinates, is the simplex whose vertices are the columns of the $(n-1) \times (n)$ matrix Ξ_n satisfying that all values of $\Omega_n \Xi_n$ outside the diagonal are -1. If all entries of Ξ_n are integers, then the dual of P_n is a lattice polytope and hence, P_n is reflexive. To show this, we construct Ξ_n explicitly with three $(n-1) \times n$ matrices, Ξ'_n, Ξ''_n, and Ξ'''_n. The ij-entries of Ξ'_n are $-(i+2)(-1)^{i+j}\binom{i}{j-1}$ and the matrices Ξ''_n and Ξ'''_n are all zero, with the exceptions that the first column of Ξ''_n is $-2(-1)^i$, and the bottom right-most entry of Ξ'''_n is $1-n$. We consider $\Xi_n = \Xi'_n + \Xi''_n + \Xi'''_n$.

To calculate the product $\Omega_n \Xi_n$, we separate both Ω_n and Ξ_n into the sums above and then multiply them. The matrix multiplications are straightforward applications of elementary combinatorics, so we only record the results:

(1) The matrix $\Omega'_n \Xi'_n$ is the sum of two matrices. The only nonzero elements of the first are the diagonal ii-entries $i(i+1)$ and the only nonzero elements of the second are the first column $i1$-entries $-i(i+1)$.

(2) The matrix $\eta_n \mathbf{1}_n \Xi'_n$ has 1s everywhere, except that the first column is constant with $-2n+1$ and the last column is $n+1$.

(3) The matrix $\Omega'_n \Xi''_n$ has 0s everywhere, except that the first column's $i1$-entry is $i(i+1)-2$.

(4) The matrix $\eta_n \mathbf{1}_n \Xi''_n$ has 0s everywhere, except that the first column is constant with $2n-2$.

(5) The matrix $\Omega'_n \Xi'''_n$ has 0s everywhere, except that the rightmost bottom corner is $-n^2$.

(6) The matrix $\eta_n \mathbf{1}_n \Xi'''_n$ has 0s everywhere, except that the rightmost column is constant $-n$.

Summing up, we conclude that the ij-entry of $\Omega_n \Xi_n = (\Omega'_n - \eta_n \mathbf{1}_n)(\Xi'_n + \Xi''_n + \Xi'''_n)$ is

$$-1 \text{ if } i \neq j, \quad i^2+i-1 \text{ if } i = j < n, \quad n \text{ if } i = j = n. \qquad \square$$

Acknowledgements

The authors thank Mats Boij, for several insightful conversations regarding this work, Benjamin Braun, for pointing out the connection to anti-lecture-hall compositions, and the anonymous referee, for the helpful suggestions to improve this paper. Engström thanks the Miller Institute for Basic Research at UC Berkeley for funding. Stamps thanks the Mathematical Sciences Research Institute for support to attend the 2011 Summer Graduate Workshop on Commutative Algebra.

References

[Barvinok 2002] A. Barvinok, *A course in convexity*, Graduate Studies in Mathematics **54**, American Mathematical Society, Providence, RI, 2002. MR 2003j:52001 Zbl 1014.52001

[Batyrev and Nill 2008] V. Batyrev and B. Nill, "Combinatorial aspects of mirror symmetry", pp. 35–66 in *Integer points in polyhedra—geometry, number theory, representation theory, algebra, optimization, statistics* (Snowbird, UT, 2006), edited by M. Beck et al., Contemp. Math. **452**, Amer. Math. Soc., Providence, RI, 2008. MR 2009m:14059 Zbl 1161.14037

[Beck and Robins 2007] M. Beck and S. Robins, *Computing the continuous discretely: Integer-point enumeration in polyhedra*, Springer, New York, 2007. MR 2007h:11119 Zbl 1114.52013

[Boij and Söderberg 2008] M. Boij and J. Söderberg, "Graded Betti numbers of Cohen–Macaulay modules and the multiplicity conjecture", *J. Lond. Math. Soc.* (2) **78**:1 (2008), 85–106. MR 2009g:13018 Zbl 1189.13008

[Boij and Söderberg 2012] M. Boij and J. Söderberg, "Betti numbers of graded modules and the multiplicity conjecture in the non-Cohen–Macaulay case", *Algebra Number Theory* **6**:3 (2012), 437–454. MR 2966705 Zbl 1259.13009

[Bousquet-Mélou and Eriksson 1997a] M. Bousquet-Mélou and K. Eriksson, "Lecture hall partitions", *Ramanujan J.* **1**:1 (1997), 101–111. MR 99c:05015 Zbl 0909.05008

[Bousquet-Mélou and Eriksson 1997b] M. Bousquet-Mélou and K. Eriksson, "Lecture hall partitions, II", *Ramanujan J.* **1**:2 (1997), 165–185. MR 99c:05016 Zbl 0909.05009

[Bruns et al. 1997] W. Bruns, J. Gubeladze, and N. V. Trung, "Normal polytopes, triangulations, and Koszul algebras", *J. Reine Angew. Math.* **485** (1997), 123–160. MR 99c:52016 Zbl 0866.20050

[Corteel and Savage 2003] S. Corteel and C. D. Savage, "Anti-lecture hall compositions", *Discrete Math.* **263**:1-3 (2003), 275–280. MR 2003m:05007 Zbl 1019.05004

[Corteel et al. 2005] S. Corteel, S. Lee, and C. D. Savage, "Enumeration of sequences constrained by the ratio of consecutive parts", *Sém. Lothar. Combin.* **54A** (2005), [article] B54Aa. MR 2006f:05011 Zbl 1086.05010

[Diestel 2010] R. Diestel, *Graph theory*, 4th ed., Graduate Texts in Mathematics **173**, Springer, Heidelberg, 2010. MR 2011m:05002 Zbl 1204.05001

[Dochtermann and Engström 2009] A. Dochtermann and A. Engström, "Algebraic properties of edge ideals via combinatorial topology", *Electron. J. Combin.* **16**:2 (2009), 1–24. MR 2010f:13027 Zbl 1161.13013

[Eisenbud and Schreyer 2009] D. Eisenbud and F.-O. Schreyer, "Betti numbers of graded modules and cohomology of vector bundles", *J. Amer. Math. Soc.* **22**:3 (2009), 859–888. MR 2011a:13024 Zbl 1213.13032

[Fløystad 2012] G. Fløystad, "Boij–Söderberg theory: Introduction and survey", pp. 1–54 in *Progress in commutative algebra, I*, edited by C. Francisco et al., de Gruyter, Berlin, 2012. MR 2932580 Zbl 1260.13020

[Fröberg 1990] R. Fröberg, "On Stanley–Reisner rings", pp. 57–70 in *Topics in algebra, II* (Warsaw, 1988), edited by S. Balcerzyk et al., Banach Center Publ. **26**, PWN, Warsaw, 1990. MR 93f:13009 Zbl 0741.13006

[Goodarzi and Yassemi 2012] A. Goodarzi and S. Yassemi, "Shellable quasi-forests and their h–triangles", *Manuscripta Math.* **137**:3-4 (2012), 475–481. MR 2012m:13041 Zbl 1246.13029

[Herzog et al. 2012] J. Herzog, L. Sharifan, and M. Varbaro, "Graded Betti numbers of componentwise linear ideals", preprint, 2012. To appear in *Proc. Amer. Math. Soc.* arXiv 1111.0442

[Klivans 2007] C. J. Klivans, "Threshold graphs, shifted complexes, and graphical complexes", *Discrete Math.* **307**:21 (2007), 2591–2597. MR 2008j:05373 Zbl 1127.05086

[Mahadev and Peled 1995] N. V. R. Mahadev and U. N. Peled, *Threshold graphs and related topics*, Annals of Discrete Mathematics **56**, North-Holland, Amsterdam, 1995. MR 97h:05001 Zbl 0852.05001

[Nagel and Sturgeon 2013] U. Nagel and S. Sturgeon, "Combinatorial interpretations of some Boij–Söderberg decompositions", *J. Algebra* **381** (2013), 54–72. MR 3030509

[Woodroofe 2011] R. Woodroofe, "Erdős–Ko–Rado theorems for simplicial complexes", *J. Combin. Theory Ser. A* **118**:4 (2011), 1218–1227. MR 2012a:05351 Zbl 1231.05308

Communicated by David Eisenbud
Received 2012-11-06 Revised 2013-01-25 Accepted 2013-03-12

alexander.engstrom@aalto.fi *Department of Mathematics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland*

matthew.stamps@aalto.fi *Department of Mathematics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland*

ALGEBRA AND NUMBER THEORY 7:7 (2013)
dx.doi.org/10.2140/ant.2013.7.1743

Hopf monoids from class functions on unitriangular matrices

Marcelo Aguiar, Nantel Bergeron and Nathaniel Thiem

We build, from the collection of all groups of unitriangular matrices, Hopf monoids in Joyal's category of species. Such structure is carried by the collection of class function spaces on those groups and also by the collection of superclass function spaces in the sense of Diaconis and Isaacs. Superclasses of unitriangular matrices admit a simple description from which we deduce a combinatorial model for the Hopf monoid of superclass functions in terms of the Hadamard product of the Hopf monoids of linear orders and of set partitions. This implies a recent result relating the Hopf algebra of superclass functions on unitriangular matrices to symmetric functions in noncommuting variables. We determine the algebraic structure of the Hopf monoid: it is a free monoid in species with the canonical Hopf structure. As an application, we derive certain estimates on the number of conjugacy classes of unitriangular matrices.

Introduction

A Hopf monoid (in Joyal's category of species) is an algebraic structure akin to that of a Hopf algebra. Combinatorial structures that compose and decompose give rise to Hopf monoids. These objects are the subject of [Aguiar and Mahajan 2010, Part II]. The few basic notions and examples needed for our purposes are reviewed in Section 1, including the Hopf monoids of linear orders, set partitions, and simple graphs and the Hadamard product of Hopf monoids.

The main goal of this paper is to construct a Hopf monoid out of the groups of unitriangular matrices with entries in a finite field and to do this in a transparent manner. The structure exists on the collection of function spaces on these groups and also on the collections of class function and superclass function spaces. It is induced by two simple operations on this collection of groups: the passage from a

Aguiar is supported in part by NSF grant DMS-1001935. Bergeron is supported in part by CRC and NSERC. Thiem is supported in part by NSF FRG DMS-0854893.
MSC2010: primary 05E10; secondary 05E05, 05E15, 16T05, 16T30, 18D35, 20C33.
Keywords: unitriangular matrix, class function, superclass function, Hopf monoid, Hopf algebra.

matrix to its principal minors gives rise to the product, and direct sum of matrices gives rise to the coproduct.

Class functions are defined for arbitrary groups. An abstract notion and theory of superclass functions (and supercharacters) for arbitrary groups exists [Diaconis and Isaacs 2008]. While a given group may admit several such theories, there is a canonical choice of superclasses for a special class of groups known as algebra groups. These notions are briefly discussed in Section 4.1. Unitriangular groups are the prototype of such groups, and we employ the corresponding notion of superclasses in Section 4.2. The study of unitriangular superclasses and supercharacters was initiated in [André 1995a; 1995b], making use of the method of Kirillov [1995], and by more elementary means in [Yan 2001].

Preliminaries on unitriangular matrices are discussed in Section 2. The Hopf monoids $\mathbf{f}(\mathrm{U})$ of functions and $\mathbf{cf}(\mathrm{U})$ of class functions are constructed in Section 3. The nature of the construction is fairly general; in particular, the same procedure yields the Hopf monoid $\mathbf{scf}(\mathrm{U})$ of superclass functions in Section 4.2.

Unitriangular matrices over $\mathbb{F}_2$ may be identified with simple graphs, and direct sums and the passage to principal minors correspond to simple operations on graphs. This yields a combinatorial model for $\mathbf{f}(\mathrm{U})$ in terms of the Hadamard product of the Hopf monoids of linear orders and of graphs, as discussed in Section 3.6. The conjugacy classes on the unitriangular groups exhibit great complexity and considerable attention has been devoted to their study [Goodwin 2006; Higman 1960; Kirillov 1995; Vera-López et al. 2008]. We do not attempt an explicit combinatorial description of the Hopf monoid $\mathbf{cf}(\mathrm{U})$. On the other hand, superclasses are well-understood (Section 4.3), and such a combinatorial description exists for $\mathbf{scf}(\mathrm{U})$. In Section 4.5, we obtain a combinatorial model in terms of the Hadamard product of the Hopf monoids of linear orders and of set partitions. This has as a consequence the main result of [Aguiar et al. 2012], as we explain in Section 6.2.

Employing the combinatorial models, we derive structure theorems for the Hopf monoids $\mathbf{f}(\mathrm{U})$ and $\mathbf{scf}(\mathrm{U})$ in Section 5. Our main results state that both are free monoids with the canonical Hopf structure (in which the generators are primitive).

Applications are presented in Section 6. With the aid of Lagrange's theorem for Hopf monoids, one may derive estimates on the number of conjugacy classes of unitriangular matrices in the form of certain recursive inequalities. We obtain this application in Section 6.1, where we also formulate a refinement of Higman's conjecture on the polynomiality of these numbers. Other applications involving the Hopf algebra of superclass functions of [Aguiar et al. 2012] are given in Section 6.2.

We employ two fields: the *base* field $\Bbbk$ and the field of *matrix entries* $\mathbb{F}$. We consider algebras and groups of matrices with entries in $\mathbb{F}$; all other vector spaces are over $\Bbbk$. The field of matrix entries is often assumed to be finite and sometimes to be $\mathbb{F}_2$.

1. Hopf monoids

We review the basics on Hopf monoids and recall three examples built from linear orders, set partitions, and simple graphs, respectively. We also consider the Hadamard product of Hopf monoids. In later sections, Hopf monoids are built from functions on unitriangular matrices. The constructions of this section will allow us to provide combinatorial models for them.

1.1. ***Species and Hopf monoids.*** For the precise definitions of *vector species* and *Hopf monoid*, we refer to [Aguiar and Mahajan 2010, Chapter 8]. The main ingredients are reviewed below.

A *vector species* $\boldsymbol{p}$ is a collection of vector spaces $\boldsymbol{p}[I]$, one for each finite set I, equivariant with respect to bijections $I \cong J$. A morphism of species $f : \boldsymbol{p} \to \boldsymbol{q}$ is a collection of linear maps $f_I : \boldsymbol{p}[I] \to \boldsymbol{q}[I]$ that commute with bijections.

A *decomposition* of a finite set I is a finite sequence $(S_1, \ldots, S_k)$ of disjoint subsets of I whose union is I. In this situation, we write

$$I = S_1 \sqcup \cdots \sqcup S_k.$$

A *Hopf monoid* consists of a vector species $\boldsymbol{h}$ equipped with two collections μ and Δ of linear maps

$$\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \xrightarrow{\mu_{S_1,S_2}} \boldsymbol{h}[I] \quad \text{and} \quad \boldsymbol{h}[I] \xrightarrow{\Delta_{S_1,S_2}} \boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2].$$

There is one map in each collection for each finite set I and each decomposition $I = S_1 \sqcup S_2$. This data is subject to a number of axioms, of which the main ones follow.

Associativity. For each decomposition $I = S_1 \sqcup S_2 \sqcup S_3$, the diagrams

$$\begin{tikzcd}
\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \otimes \boldsymbol{h}[S_3] \arrow[r, "\mathrm{id}\otimes\mu_{S_2,S_3}"] \arrow[d, "\mu_{S_1,S_2}\otimes\mathrm{id}"'] & \boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2 \sqcup S_3] \arrow[d, "\mu_{S_1,S_2\sqcup S_3}"] \\
\boldsymbol{h}[S_1 \sqcup S_2] \otimes \boldsymbol{h}[S_3] \arrow[r, "\mu_{S_1\sqcup S_2,S_3}"'] & \boldsymbol{h}[I]
\end{tikzcd} \tag{1}$$

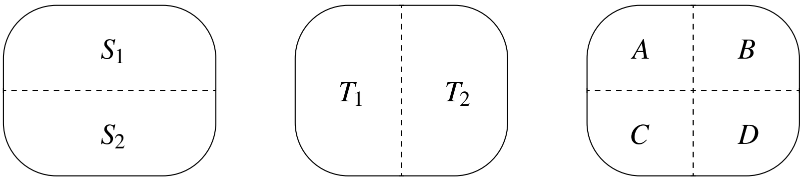
$$\begin{tikzcd}
\boldsymbol{h}[I] \arrow[r, "\Delta_{S_1\sqcup S_2,S_3}"] \arrow[d, "\Delta_{S_1,S_2\sqcup S_3}"'] & \boldsymbol{h}[S_1 \sqcup S_2] \otimes \boldsymbol{h}[S_3] \arrow[d, "\Delta_{S_1,S_2}\otimes\mathrm{id}"] \\
\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2 \sqcup S_3] \arrow[r, "\mathrm{id}\otimes\Delta_{S_2,S_3}"'] & \boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \otimes \boldsymbol{h}[S_3]
\end{tikzcd} \tag{2}$$

commute.

Compatibility. Fix decompositions $S_1 \sqcup S_2 = I = T_1 \sqcup T_2$, and consider the resulting pairwise intersections

$$A := S_1 \cap T_1, \quad B := S_1 \cap T_2, \quad C := S_2 \cap T_1, \quad \text{and} \quad D := S_2 \cap T_2,$$

as illustrated below:



(3)

For any such pair of decompositions, the diagram

$$\begin{tikzcd}
\boldsymbol{h}[A] \otimes \boldsymbol{h}[B] \otimes \boldsymbol{h}[C] \otimes \boldsymbol{h}[D] \arrow[rr, "\cong"] & & \boldsymbol{h}[A] \otimes \boldsymbol{h}[C] \otimes \boldsymbol{h}[B] \otimes \boldsymbol{h}[D] \arrow[d, "\mu_{A,C} \otimes \mu_{B,D}"] \\
\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \arrow[u, "\Delta_{A,B} \otimes \Delta_{C,D}"] \arrow[r, "\mu_{S_1,S_2}"'] & \boldsymbol{h}[I] \arrow[r, "\Delta_{T_1,T_2}"'] & \boldsymbol{h}[T_1] \otimes \boldsymbol{h}[T_2]
\end{tikzcd} \tag{4}$$

must commute. The top arrow stands for the map that interchanges the middle factors.

In addition, the Hopf monoid $\boldsymbol{h}$ is *connected* if $\boldsymbol{h}[\varnothing] = \Bbbk$ and the maps

$$\boldsymbol{h}[I] \otimes \boldsymbol{h}[\varnothing] \underset{\Delta_{I,\varnothing}}{\overset{\mu_{I,\varnothing}}{\rightleftarrows}} \boldsymbol{h}[I] \quad \text{and} \quad \boldsymbol{h}[\varnothing] \otimes \boldsymbol{h}[I] \underset{\Delta_{\varnothing,I}}{\overset{\mu_{\varnothing,I}}{\rightleftarrows}} \boldsymbol{h}[I]$$

are the canonical identifications.

The collection μ is the *product*, and the collection Δ is the *coproduct* of the Hopf monoid $\boldsymbol{h}$.

A Hopf monoid is (co)commutative if the left (right) diagram below commutes for all decompositions $I = S_1 \sqcup S_2$:

$$\begin{tikzcd}
\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \arrow[rr, "\cong"] \arrow[dr, "\mu_{S_1,S_2}"'] & & \boldsymbol{h}[S_2] \otimes \boldsymbol{h}[S_1] \arrow[dl, "\mu_{S_2,S_1}"] \\
& \boldsymbol{h}[I] &
\end{tikzcd} \qquad
\begin{tikzcd}
\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \arrow[rr, "\cong"] & & \boldsymbol{h}[S_2] \otimes \boldsymbol{h}[S_1] \\
& \boldsymbol{h}[I] \arrow[ul, "\Delta_{S_1,S_2}"] \arrow[ur, "\Delta_{S_2,S_1}"'] &
\end{tikzcd} \tag{5}$$

The top arrows stand for the map that interchanges the factors.

A morphism of Hopf monoids $f : \boldsymbol{h} \to \boldsymbol{k}$ is a morphism of species that commutes with μ and Δ.

1.2. ***The Hopf monoid of linear orders.*** For any finite set I, $\mathrm{L}[I]$ is the set of all linear orders on I. For instance, if $I = \{a, b, c\}$,

$$\mathrm{L}[I] = \{abc, bac, acb, bca, cab, cba\}.$$

Let $\boldsymbol{L}[I]$ denote the vector space with basis $\mathrm{L}[I]$. The collection $\boldsymbol{L}$ is a vector species.

Let $I = S_1 \sqcup S_2$. Given linear orders ℓ_i on S_i, $i = 1, 2$, their concatenation $\ell_1 \cdot \ell_2$ is a linear order on I. This is the list consisting of the elements of S_1 as ordered by ℓ_1 followed by those of S_2 as ordered by ℓ_2. Given a linear order ℓ on I and $S \subseteq I$, the restriction $\ell|_S$ (the list consisting of the elements of S written in the order in which they appear in ℓ) is a linear order on S. These operations give rise to maps

$$\begin{aligned} \mathrm{L}[S_1] \times \mathrm{L}[S_2] &\to \mathrm{L}[I], \\ (\ell_1, \ell_2) &\mapsto \ell_1 \cdot \ell_2 \end{aligned} \quad \text{and} \quad \begin{aligned} \mathrm{L}[I] &\to \mathrm{L}[S_1] \times \mathrm{L}[S_2], \\ \ell &\mapsto (\ell|_{S_1}, \ell|_{S_2}). \end{aligned} \tag{6}$$

Extending by linearity, we obtain linear maps

$$\mu_{S_1,S_2} : \boldsymbol{L}[S_1] \otimes \boldsymbol{L}[S_2] \to \boldsymbol{L}[I] \quad \text{and} \quad \Delta_{S_1,S_2} : \boldsymbol{L}[I] \to \boldsymbol{L}[S_1] \otimes \boldsymbol{L}[S_2]$$

that turn $\boldsymbol{L}$ into a Hopf monoid. For instance, given linear orders ℓ_i on S_i, $i = 1, 2$, the commutativity of (4) boils down to the fact that the concatenation of $\ell_1|_A$ and $\ell_2|_C$ agrees with the restriction to T_1 of $\ell_1 \cdot \ell_2$. The Hopf monoid $\boldsymbol{L}$ is cocommutative but not commutative. For more details, see [Aguiar and Mahajan 2010, Section 8.5].

1.3. ***The Hopf monoid of set partitions.*** A *partition* of a finite set I is a collection X of disjoint nonempty subsets whose union is I. The subsets are the *blocks* of X.

Given a partition X of I and $S \subseteq I$, the restriction $X|_S$ is the partition of S whose blocks are the nonempty intersections of the blocks of X with S. Let $I = S_1 \sqcup S_2$. Given partitions X_i of S_i, $i = 1, 2$, their union is the partition $X_1 \sqcup X_2$ of I whose blocks are the blocks of X_1 and the blocks of X_2. A quasishuffle of X_1 and X_2 is any partition of I whose restriction to S_i is X_i, $i = 1, 2$.

Let $\Pi[I]$ denote the set of partitions of I and $\boldsymbol{\Pi}[I]$ the vector space with basis $\Pi[I]$. A Hopf monoid structure on $\boldsymbol{\Pi}$ is defined and studied in [Aguiar and Mahajan 2010, Section 12.6]. Among its various linear bases, we are interested in the basis $\{m_X\}$ on which the operations are as follows. The product

$$\mu_{S_1,S_2} : \boldsymbol{\Pi}[S_1] \otimes \boldsymbol{\Pi}[S_2] \to \boldsymbol{\Pi}[I]$$

is given by

$$\mu_{S_1,S_2}(m_{X_1} \otimes m_{X_2}) = \sum_{\substack{X:\ X|_{S_1}=X_1 \\ X|_{S_2}=X_2}} m_X. \tag{7}$$

The coproduct

$$\Delta_{S_1,S_2} : \mathbf{\Pi}[I] \to \mathbf{\Pi}[S_1] \otimes \mathbf{\Pi}[S_2]$$

is given by

$$\Delta_{S_1,S_2}(m_X) = \begin{cases} m_{X|_{S_1}} \otimes m_{X|_{S_2}} & \text{if } S_1 \text{ is the union of some blocks of } X, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Note that the following conditions are equivalent for a partition X of I:

- S_1 is the union of some blocks of X.
- S_2 is the union of some blocks of X.
- $X = X|_{S_1} \sqcup X|_{S_2}$.

These operations turn the species $\mathbf{\Pi}$ into a Hopf monoid that is both commutative and cocommutative.

1.4. ***The Hopf monoid of simple graphs.*** A (*simple*) *graph* g on a finite set I is a relation on I that is symmetric and irreflexive. The elements of I are the *vertices* of g There is an *edge* between two vertices when they are related by g.

Given a graph g on I and $S \subseteq I$, the restriction $g|_S$ is the graph on S whose edges are the edges of g between elements of S. Let $I = S_1 \sqcup S_2$. Given graphs g_i of S_i, $i = 1, 2$, their union is the graph $g_1 \sqcup g_2$ of I whose edges are those of g_1 and those of g_2. A quasishuffle of g_1 and g_2 is any graph on I whose restriction to S_i is g_i, $i = 1, 2$.

Let $\mathrm{G}[I]$ denote the set of graphs on I and $\boldsymbol{G}[I]$ the vector space with basis $\mathrm{G}[I]$. A Hopf monoid structure on $\boldsymbol{G}$ is defined and studied in [Aguiar and Mahajan 2010, Section 13.2]. We are interested in the basis $\{m_g\}$ on which the operations are as follows. The product

$$\mu_{S_1,S_2} : \boldsymbol{G}[S_1] \otimes \boldsymbol{G}[S_2] \to \boldsymbol{G}[I]$$

is given by

$$\mu_{S_1,S_2}(m_{g_1} \otimes m_{g_2}) = \sum_{\substack{g:\ g|_{S_1}=g_1 \\ g|_{S_2}=g_2}} m_g. \tag{9}$$

The coproduct

$$\Delta_{S_1,S_2} : \boldsymbol{G}[I] \to \boldsymbol{G}[S_1] \otimes \boldsymbol{G}[S_2]$$

is given by

$$\Delta_{S_1,S_2}(m_g) = \begin{cases} m_{g|_{S_1}} \otimes m_{g|_{S_2}} & \text{if no edge of } g \text{ connects } S_1 \text{ to } S_2, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Note that no edge of g connects S_1 to S_2 if and only if $g = g|_{S_1} \sqcup g|_{S_2}$.

These operations turn the species $\boldsymbol{G}$ into a Hopf monoid that is both commutative and cocommutative.

Remark 1. The dual of a species $\boldsymbol{p}$ is the collection $\boldsymbol{p}^*$ of dual vector spaces: $\boldsymbol{p}^*[I] = \boldsymbol{p}[I]^*$. A species $\boldsymbol{p}$ is said to be finite-dimensional if each space $\boldsymbol{p}[I]$ is finite-dimensional. Dualizing the operations of a finite-dimensional Hopf monoid $\boldsymbol{h}$, one obtains a Hopf monoid $\boldsymbol{h}^*$. The Hopf monoid $\boldsymbol{h}$ is called self-dual if $\boldsymbol{h} \cong \boldsymbol{h}^*$. In general, such isomorphism is not unique.

Over a field of characteristic 0, a Hopf monoid that is connected, commutative, and cocommutative is always self-dual. This is a consequence of the Cartier–Milnor–Moore theorem. (The isomorphism with the dual is not canonical.)

In particular, the Hopf monoids $\boldsymbol{\Pi}$ and $\boldsymbol{G}$ are self-dual. In [Aguiar and Mahajan 2010], the preceding descriptions of these Hopf monoids are stated in terms of their duals $\boldsymbol{\Pi}^*$ and $\boldsymbol{G}^*$. A different description of $\boldsymbol{\Pi}$ is given in [Aguiar and Mahajan 2010, Section 12.6.2]. To reconcile the two, one should use the explicit isomorphism $\boldsymbol{\Pi} \cong \boldsymbol{\Pi}^*$ given in [Aguiar and Mahajan 2010, Proposition 12.48].

1.5. ***The Hadamard product.*** Given species $\boldsymbol{p}$ and $\boldsymbol{q}$, their *Hadamard product* is the species $\boldsymbol{p} \times \boldsymbol{q}$ defined by

$$(\boldsymbol{p} \times \boldsymbol{q})[I] = \boldsymbol{p}[I] \otimes \boldsymbol{q}[I].$$

If $\boldsymbol{h}$ and $\boldsymbol{k}$ are Hopf monoids, then so is $\boldsymbol{h} \times \boldsymbol{k}$ with the following operations. Let $I = S_1 \sqcup S_2$. The product is

$$\begin{tikzcd}
(\boldsymbol{h} \times \boldsymbol{k})[S_1] \otimes (\boldsymbol{h} \times \boldsymbol{k})[S_2] \arrow[rr, dashed] \arrow[d, equal] & & (\boldsymbol{h} \times \boldsymbol{k})[I] \arrow[d, equal] \\
\boldsymbol{h}[S_1] \otimes \boldsymbol{k}[S_1] \otimes \boldsymbol{h}[S_2] \otimes \boldsymbol{k}[S_2] \arrow[r, "\cong"'] & \begin{array}{c}\boldsymbol{h}[S_1] \otimes \boldsymbol{h}[S_2] \\ \otimes\, \boldsymbol{k}[S_1] \otimes \boldsymbol{k}[S_2]\end{array} \arrow[r, "\mu_{S,T} \otimes \mu_{S,T}"'] & \boldsymbol{h}[I] \otimes \boldsymbol{k}[I]
\end{tikzcd}$$

and the coproduct is defined dually. If $\boldsymbol{h}$ and $\boldsymbol{k}$ are (co)commutative, then so is $\boldsymbol{h} \times \boldsymbol{k}$. For more details, see [Aguiar and Mahajan 2010, Section 8.13].

2. Unitriangular matrices

This section sets up the basic notation pertaining to unitriangular matrices and discusses two simple but important constructions: direct sum of matrices and the passage from a matrix to its principal minors. The Hopf monoid constructions of later sections are based on them. The key results are Lemmas 2 and 3. The former is the reason why we must use unitriangular matrices: for arbitrary matrices, the passage to principal minors is not multiplicative. The latter will be responsible (in later sections) for the necessary compatibility between the product and coproduct of the Hopf monoids.

Let $\mathbb{F}$ be a field, I a finite set, and ℓ a linear order on I. Let $\mathrm{M}(I)$ denote the algebra of matrices

$$A = (a_{ij})_{i,j\in I}, \quad a_{ij} \in \mathbb{F} \text{ for all } i, j \in I.$$

The general linear group $\mathrm{GL}(I)$ consists of the invertible matrices in $\mathrm{M}(I)$, and the subgroup $\mathrm{U}(I, \ell)$ consists of the *upper ℓ-unitriangular* matrices

$$U = (u_{ij})_{i,j\in I}, \quad u_{ii} = 1 \text{ for all } i \in I \text{ and } u_{ij} = 0 \text{ whenever } i >_\ell j.$$

If ℓ' is another linear order on I, then $\mathrm{U}(I, \ell)$ and $\mathrm{U}(I, \ell')$ are conjugate subgroups of $\mathrm{GL}(I)$. However, we want to keep track of all groups in this collection and of the manner in which they interact.

2.1. ***Direct sum of matrices.*** Suppose $I = S_1 \sqcup S_2$ is a decomposition. Given $A = (a_{ij}) \in \mathrm{M}(S_1)$ and $B = (b_{ij}) \in \mathrm{M}(S_2)$, their direct sum is

$$A \oplus B = (c_{ij}) \in \mathrm{M}(I),$$

the matrix with entries

$$c_{ij} = \begin{cases} a_{ij} & \text{if both } i, j \in S_1, \\ b_{ij} & \text{if both } i, j \in S_2, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\ell \in \mathrm{L}[I]$. The direct sum of an $\ell|_{S_1}$-unitriangular and an $\ell|_{S_2}$-unitriangular matrix is ℓ-unitriangular. The morphism of algebras

$$\mathrm{M}(S_1) \times \mathrm{M}(S_2) \to \mathrm{M}(I), \quad (A, B) \mapsto A \oplus B$$

thus restricts to a morphism of groups

$$\sigma_{S_1,S_2} : \mathrm{U}(S_1, \ell|_{S_1}) \times \mathrm{U}(S_2, \ell|_{S_2}) \to \mathrm{U}(I, \ell). \tag{11}$$

(The dependence of σ_{S_1,S_2} on ℓ is left implicit.)

Direct sum of matrices is associative; thus, for any decomposition $I = S_1 \sqcup S_2 \sqcup S_3$, the diagram

$$\begin{tikzcd} \mathrm{U}(S_1, \ell|_{S_1}) \times \mathrm{U}(S_2, \ell|_{S_2}) \times \mathrm{U}(S_3, \ell|_{S_3}) \arrow[r, "\sigma_{S_1,S_2}\times\mathrm{id}"] \arrow[d, "\mathrm{id}\times\sigma_{S_2,S_3}"'] & \mathrm{U}(S_1 \sqcup S_2, \ell|_{S_1\sqcup S_2}) \times \mathrm{U}(S_3, \ell|_{S_3}) \arrow[d, "\sigma_{S_1\sqcup S_2,S_3}"] \\ \mathrm{U}(S_1, \ell|_{S_1}) \times \mathrm{U}(S_2 \sqcup S_3, \ell|_{S_2\sqcup S_3}) \arrow[r, "\sigma_{S_1,S_2\sqcup S_3}"'] & \mathrm{U}(I, \ell) \end{tikzcd} \tag{12}$$

commutes. Note also that, with these definitions, $A \oplus B$ and $B \oplus A$ are the same matrix. Thus, the following diagram commutes:

$$\begin{tikzcd} \mathrm{U}(S_1, \ell|_{S_1}) \times \mathrm{U}(S_2, \ell|_{S_2}) \arrow[rr, "\cong"] \arrow[dr, "\sigma_{S_1,S_2}"'] & & \mathrm{U}(S_2, \ell|_{S_2}) \times \mathrm{U}(S_1, \ell|_{S_1}) \arrow[dl, "\sigma_{S_2,S_1}"] \\ & \mathrm{U}(S_1 \sqcup S_2, \ell) & \end{tikzcd} \tag{13}$$

2.2. *Principal minors.* Given $A = (a_{ij}) \in \mathrm{M}(I)$, the *principal minor* indexed by $S \subseteq I$ is the matrix

$$A_S = (a_{ij})_{i,j \in S}.$$

In general, A_S is not invertible even if A is. In addition, the assignment $A \mapsto A_S$ does not preserve multiplications. On the other hand, if U is ℓ-unitriangular, then U_S is $\ell|_S$-unitriangular. In regard to multiplicativity, we have the following basic fact.

We say that S is an ℓ-segment if $i, k \in S$ and $i <_\ell j <_\ell k$ imply that also $j \in S$.

Let $E_{ij} \in \mathrm{M}(I)$ denote the *elementary* matrix in which the (i, j) entry is 1 and all other entries are 0.

Lemma 2. *Let $\ell \in \mathrm{L}[I]$ and $S \subseteq I$. The map*

$$\mathrm{U}(I, \ell) \to \mathrm{U}(S, \ell|_S), \quad U \mapsto U_S$$

is a morphism of groups if and only if S is an ℓ-segment.

Proof. Suppose the map is a morphism of groups. Let $i, j, k \in I$ be such that $i, k \in S$ and $i <_\ell j <_\ell k$. The matrices

$$\mathrm{Id} + E_{ij} \quad \text{and} \quad \mathrm{Id} + E_{jk}$$

are in $\mathrm{U}(I, \ell)$, and

$$(\mathrm{Id} + E_{ij}) \cdot (\mathrm{Id} + E_{jk}) = \mathrm{Id} + E_{ij} + E_{jk} + E_{ik}.$$

If $j \notin S$, then the two matrices are in the kernel of the map while their product is mapped to $\mathrm{Id} + E_{ik} \neq \mathrm{Id}$. Thus, $j \in S$ and S is an ℓ-segment.

The converse implication follows from the fact that if U and V are ℓ-unitriangular, then the (i, k) entry of UV is

$$\sum_{i \leq_\ell j \leq_\ell k} u_{ij} v_{jk}. \qquad \square$$

Let $I = S_1 \sqcup S_2$ be a decomposition with $\ell_i \in \mathrm{L}[S_i]$, $i = 1, 2$. We define a map

$$\pi_{S_1,S_2} : \mathrm{U}(I, \ell_1 \cdot \ell_2) \to \mathrm{U}(S_1, \ell_1) \times \mathrm{U}(S_2, \ell_2) \tag{14}$$

by

$$U \mapsto (U_{S_1}, U_{S_2}).$$

Note that S_1 is an initial segment for $\ell_1 \cdot \ell_2$ and S_2 is a final segment for $\ell_1 \cdot \ell_2$. Thus, π_{S_1,S_2} is a morphism of groups by Lemma 2.

If $R \subseteq S \subseteq I$, then $(A_S)_R = A_R$. This implies the following commutativity for any decomposition $I = S_1 \sqcup S_2 \sqcup S_3$ and $\ell_i \in \mathrm{L}[S_i]$, $i = 1, 2, 3$:

$$\begin{CD} \mathrm{U}(I, \ell_1 \cdot \ell_2 \cdot \ell_3) @>{\pi_{S_1 \sqcup S_2, S_3}}>> \mathrm{U}(S_1 \sqcup S_2, \ell_1 \cdot \ell_2) \times \mathrm{U}(S_3, \ell_3) \\ @V{\pi_{S_1, S_2 \sqcup S_3}}VV @VV{\pi_{S_1,S_2} \times \mathrm{id}}V \\ \mathrm{U}(S_1, \ell_1) \times \mathrm{U}(S_2 \sqcup S_3, \ell_2 \cdot \ell_3) @>>{\mathrm{id} \times \pi_{S_2,S_3}}> \mathrm{U}(S_1, \ell_1) \times \mathrm{U}(S_2, \ell_2) \times \mathrm{U}(S_3, \ell_3) \end{CD} \tag{15}$$

2.3. ***Direct sums and principal minors.*** The following key result relates the collection of morphisms σ to the collection π:

Lemma 3. *Fix two decompositions $I = S_1 \sqcup S_2 = T_1 \sqcup T_2$, and let A, B, C, and D be the resulting intersections, as in (3). Let ℓ_i be a linear order on S_i, $i = 1, 2$, and $\ell = \ell_1 \cdot \ell_2$. Then the following diagram commutes*:

$$\begin{CD} \mathrm{U}(T_1, \ell|_{T_1}) \times \mathrm{U}(T_2, \ell|_{T_2}) @>{\pi_{A,C} \times \pi_{B,D}}>> \begin{array}{c}\mathrm{U}(A, \ell_1|_A) \times \mathrm{U}(C, \ell_2|_C) \\ \times\, \mathrm{U}(B, \ell_1|_B) \times \mathrm{U}(D, \ell_2|_D)\end{array} \\ @V{\sigma_{T_1,T_2}}VV @. \\ \mathrm{U}(I, \ell) @. @VV{\cong}V \\ @V{\pi_{S_1,S_2}}VV @. \\ \mathrm{U}(S_1, \ell_1) \times \mathrm{U}(S_2, \ell_2) @<<{\sigma_{A,B} \times \sigma_{C,D}}< \begin{array}{c}\mathrm{U}(A, \ell_1|_A) \times \mathrm{U}(B, \ell_1|_B) \\ \times\, \mathrm{U}(C, \ell_2|_C) \times \mathrm{U}(D, \ell_2|_D)\end{array} \end{CD} \tag{16}$$

Proof. First note that since $\ell|_{T_1} = (\ell_1|_A) \cdot (\ell_2|_C)$, $\pi_{A,C}$ does map as stated in the diagram and similarly for $\pi_{B,D}$. The commutativity of the diagram boils down to the simple fact that

$$(U \oplus V)_{S_1} = U_A \oplus V_B$$

(and a similar statement for S_2, C, and D). This holds for any $U \in \mathrm{M}(T_1)$ and $V \in \mathrm{M}(T_2)$. $\square$

3. A Hopf monoid of (class) functions

We employ the operations of Section 2 (direct sum of matrices and the passage from a matrix to its principal minors) to build a Hopf monoid structure on the collection of function spaces on unitriangular matrices. The collection of class function spaces gives rise to a Hopf submonoid. With matrix entries in $\mathbb{F}_2$, the Hopf monoid of functions may be identified with the Hadamard product of the Hopf monoids of linear orders and of simple graphs.

3.1. ***Functions.*** Given a set X, let $\mathsf{f}(X)$ denote the vector space of functions on X with values on the base field $\Bbbk$. The functor

$$\mathsf{f} : \{\text{sets}\} \to \{\text{vector spaces}\}$$

is contravariant. If at least one of two sets X_1 and X_2 is finite, then there is a canonical isomorphism

$$\mathsf{f}(X_1 \times X_2) \cong \mathsf{f}(X_1) \otimes \mathsf{f}(X_2). \tag{17}$$

A function $f \in \mathsf{f}(X_1 \times X_2)$ corresponds to $\sum_i f_i^1 \otimes f_i^2 \in \mathsf{f}(X_1) \otimes \mathsf{f}(X_2)$ if and only if

$$f(x_1, x_2) = \sum_i f_i^1(x_1) f_i^2(x_2) \quad \text{for all } x_1 \in X_1 \text{ and } x_2 \in X_2.$$

Given an element $x \in X$, let $\kappa_x : X \to \Bbbk$ denote its characteristic function:

$$\kappa_x(y) = \begin{cases} 1 & \text{if } y = x, \\ 0 & \text{if not.} \end{cases} \tag{18}$$

Suppose now that X is finite. As x runs over the elements of X, the functions κ_x form a linear basis of $\mathsf{f}(X)$. If $\varphi : X \to X'$ is a function and x' is an element of X', then

$$\kappa_{x'} \circ \varphi = \sum_{\varphi(x) = x'} \kappa_x. \tag{19}$$

Under (17),

$$\kappa_{(x_1, x_2)} \leftrightarrow \kappa_{x_1} \otimes \kappa_{x_2}. \tag{20}$$

3.2. ***Class functions on groups.*** Given a group G, let $\mathsf{cf}(G)$ denote the vector space of *class functions* on G. These are the functions $f : G \to \Bbbk$ that are constant on conjugacy classes of G. If $\varphi : G \to G'$ is a morphism of groups and f is a class function on G', then $f \circ \varphi$ is a class function on G. In this manner,

$$\mathsf{cf} : \{\text{groups}\} \to \{\text{vector spaces}\}$$

is a contravariant functor. If at least one of two groups G_1 and G_2 is finite, then there is a canonical isomorphism

$$\mathsf{cf}(G_1 \times G_2) \cong \mathsf{cf}(G_1) \otimes \mathsf{cf}(G_2) \tag{21}$$

obtained by restriction from the isomorphism (17).

Given a conjugacy class C of G, let $\kappa_C : G \to \Bbbk$ denote its characteristic function:

$$\kappa_C(x) = \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{if not.} \end{cases} \tag{22}$$

Suppose G has finitely many conjugacy classes. As C runs over the conjugacy classes of G, the functions κ_C form a linear basis of $\mathsf{cf}(G)$. If C' is a conjugacy class of G', then

$$\kappa_{C'} \circ \varphi = \sum_{\varphi(C) \subseteq C'} \kappa_C. \tag{23}$$

The conjugacy classes of $G_1 \times G_2$ are of the form $C_1 \times C_2$, where C_i is a conjugacy class of G_i, $i = 1, 2$. Under (21),

$$\kappa_{C_1 \times C_2} \leftrightarrow \kappa_{C_1} \otimes \kappa_{C_2}. \tag{24}$$

3.3. ***Functions on unitriangular matrices.*** We assume for the rest of this section that the field $\mathbb{F}$ of matrix entries is finite. Thus, all groups $\mathrm{U}(I, \ell)$ of unitriangular matrices are finite.

We define a vector species $\mathsf{f}(\mathrm{U})$ as follows. On a finite set I,

$$\mathsf{f}(\mathrm{U})[I] = \bigoplus_{\ell \in \mathrm{L}[I]} \mathsf{f}(\mathrm{U}(I, \ell)).$$

In other words, $\mathsf{f}(\mathrm{U})[I]$ is the direct sum of the spaces of functions on all unitriangular groups on I. A bijection $\sigma : I \cong J$ induces an isomorphism $\mathrm{U}(I, \ell) \cong \mathrm{U}(J, \sigma \cdot \ell)$ and therefore an isomorphism $\mathsf{f}(\mathrm{U})[I] \cong \mathsf{f}(\mathrm{U})[J]$. Thus, $\mathsf{f}(\mathrm{U})$ is a species.

Let $I = S_1 \sqcup S_2$ and $\ell_i \in \mathrm{L}[S_i]$, $i = 1, 2$. Applying the functor f to the morphism π_{S_1,S_2} in (14) and composing with the isomorphism in (17), we obtain a linear map

$$\mathsf{f}(\mathrm{U}(S_1, \ell_1)) \otimes \mathsf{f}(\mathrm{U}(S_2, \ell_2)) \to \mathsf{f}(\mathrm{U}(I, \ell_1 \cdot \ell_2)).$$

Adding over all $\ell_1 \in \mathrm{L}[S_1]$ and $\ell_2 \in \mathrm{L}[S_2]$, we obtain a linear map

$$\mu_{S_1,S_2} : \mathsf{f}(\mathrm{U})[S_1] \otimes \mathsf{f}(\mathrm{U})[S_2] \to \mathsf{f}(\mathrm{U})[I]. \tag{25}$$

Explicitly, given functions $f : \mathrm{U}(S_1, \ell_1) \to \Bbbk$ and $g : \mathrm{U}(S_2, \ell_2) \to \Bbbk$,

$$\mu_{S_1,S_2}(f \otimes g) : \mathrm{U}(I, \ell_1 \cdot \ell_2) \to \Bbbk$$

is the function given by

$$U \mapsto f(U_{S_1}) g(U_{S_2}). \tag{26}$$

Similarly, from the map σ_{S_1,S_2} in (11), we obtain the components

$$\mathsf{f}(\mathrm{U}(I, \ell)) \to \mathsf{f}(\mathrm{U}(S_1, \ell|_{S_1})) \otimes \mathsf{f}(\mathrm{U}(S_2, \ell|_{S_2}))$$

(one for each $\ell \in \mathrm{L}[I]$) of a linear map

$$\Delta_{S_1,S_2} : \mathsf{f}(\mathrm{U})[I] \to \mathsf{f}(\mathrm{U})[S_1] \otimes \mathsf{f}(\mathrm{U})[S_2]. \tag{27}$$

Explicitly, given a function $f : \mathrm{U}(I, \ell) \to \Bbbk$, we have $\Delta_{S_1,S_2}(f) = \sum_i f_i^1 \otimes f_i^2$, where

$$f_i^1 : \mathrm{U}(S_1, \ell|_{S_1}) \to \Bbbk \quad \text{and} \quad f_i^2 : \mathrm{U}(S_2, \ell|_{S_2}) \to \Bbbk$$

are functions such that

$$f(U_1 \oplus U_2) = \sum_i f_i^1(U_1) f_i^2(U_2) \qquad \text{for all } U_1 \in \mathrm{U}(S_1, \ell|_{S_1}) \text{ and } U_2 \in \mathrm{U}(S_2, \ell|_{S_2}). \tag{28}$$

Proposition 4. *With the operations (25) and (27), the species* $\mathsf{f}(\mathrm{U})$ *is a connected Hopf monoid. It is cocommutative.*

Proof. Axioms (1), (2), and (4) follow from (12), (15), and (16) by functoriality. In the same manner, cocommutativity (5) follows from (13). □

We describe the operations on the basis of characteristic functions (18). Let $U_i \in \mathrm{U}(S_i, \ell_i)$, $i = 1, 2$. It follows from (19) and (20), or from (26), that the product is

$$\mu_{S_1,S_2}(\kappa_{U_1} \otimes \kappa_{U_2}) = \sum_{\pi_{S_1,S_2}(U)=(U_1,U_2)} \kappa_U = \sum_{\substack{U_{S_1}=U_1 \\ U_{S_2}=U_2}} \kappa_U. \tag{29}$$

Similarly, the coproduct is

$$\Delta_{S_1,S_2}(\kappa_U) = \sum_{\sigma_{S_1,S_2}(U_1,U_2)=U} \kappa_{U_1} \otimes \kappa_{U_2} = \begin{cases} \kappa_{U_{S_1}} \otimes \kappa_{U_{S_2}} & \text{if } U = U_{S_1} \oplus U_{S_2}, \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

3.4. *Constant functions.* Let $\mathbf{1}_\ell$ denote the constant function on $\mathrm{U}(I, \ell)$ with value 1. Let $I = S_1 \sqcup S_2$. It follows from (26) that

$$\mu_{S_1,S_2}(\mathbf{1}_{\ell_1} \otimes \mathbf{1}_{\ell_2}) = \mathbf{1}_{\ell_1 \cdot \ell_2}$$

for any $\ell_1 \in \mathrm{L}[S_1]$ and $\ell_2 \in \mathrm{L}[S_2]$. Similarly, we see from (28) that

$$\Delta_{S_1,S_2}(\mathbf{1}_\ell) = \mathbf{1}_{\ell|_{S_1}} \otimes \mathbf{1}_{\ell|_{S_2}}$$

for any $\ell \in \mathrm{L}[I]$. We thus have:

Corollary 5. *The collection of maps*

$$\boldsymbol{L}[I] \to \mathsf{f}(\mathrm{U})[I], \quad \ell \mapsto \mathbf{1}_\ell$$

is an injective morphism of Hopf monoids.

3.5. ***Class functions on unitriangular matrices.*** Let $\mathsf{cf}(\mathrm{U})[I]$ be the direct sum of the spaces of class functions on all unitriangular groups on I:

$$\mathsf{cf}(\mathrm{U})[I] = \bigoplus_{\ell \in \mathrm{L}[I]} \mathsf{cf}(\mathrm{U}(I, \ell)).$$

This defines a subspecies $\mathsf{cf}(\mathrm{U})$ of $\mathsf{f}(\mathrm{U})$.

Proceeding in the same manner as in Section 3.3, we obtain linear maps

$$\mathsf{cf}(\mathrm{U})[S_1] \otimes \mathsf{cf}(\mathrm{U})[S_2] \underset{\Delta_{S_1,S_2}}{\overset{\mu_{S_1,S_2}}{\rightleftarrows}} \mathsf{cf}(\mathrm{U})[I]$$

by applying the functor cf to the morphisms π_{S_1,S_2} and σ_{S_1,S_2}. This is meaningful since the latter are morphisms of groups (in the case of π_{S_1,S_2}, by Lemma 2).

Proposition 6. *With these operations, the species* $\mathsf{cf}(\mathrm{U})$ *is a connected cocommutative Hopf monoid. It is a Hopf submonoid of* $\mathsf{f}(\mathrm{U})$.

Proof. As in the proof of Proposition 4, the first statement follows by functoriality. The second follows from the naturality of the inclusion of class functions and its compatibility with the isomorphisms in (17) and (21). □

We describe the operations on the basis of characteristic functions (22). Let C_i be a conjugacy class of $\mathrm{U}(S_i, \ell_i)$, $i = 1, 2$. It follows from (23) and (24) that the product is

$$\mu_{S_1,S_2}(\kappa_{C_1} \otimes \kappa_{C_2}) = \sum_{\pi_{S_1,S_2}(C) \subseteq C_1 \times C_2} \kappa_C, \tag{31}$$

where the sum is over conjugacy classes C in $\mathrm{U}(I, \ell_1 \cdot \ell_2)$. Similarly, the coproduct is

$$\Delta_{S_1,S_2}(\kappa_C) = \sum_{\sigma_{S_1,S_2}(C_1 \times C_2) \subseteq C} \kappa_{C_1} \otimes \kappa_{C_2}. \tag{32}$$

Here C is a conjugacy class of $\mathrm{U}(I, \ell)$, and the sum is over pairs of conjugacy classes C_i of $\mathrm{U}(S_i, \ell|_{S_i})$.

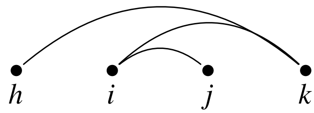
Remark 7. Let

$$\mathscr{F} : \{\text{groups}\} \to \{\text{vector spaces}\}$$

be a functor that is contravariant and *bilax monoidal* in the sense of [Aguiar and Mahajan 2010, Section 3.1]. The construction of the Hopf monoids $\mathsf{f}(\mathrm{U})$ and $\mathsf{cf}(\mathrm{U})$ can be carried out for any such functor $\mathscr{F}$ in place of cf in exactly the same manner. It can also be carried out for a covariant bilax monoidal functor $\mathscr{F}$ in a similar manner.

3.6. ***A combinatorial model.*** To a unitriangular matrix $U \in \mathrm{U}(I, \ell)$, we associate a graph $g(U)$ on I as follows: there is an edge between i and j if $i < j$ in ℓ and $u_{ij} \neq 0$. For example, given nonzero entries $a, b, c \in \mathbb{F}$,

$$\ell = hijk, \quad U = \begin{pmatrix} 1 & 0 & 0 & a \\ & 1 & b & c \\ & & 1 & 0 \\ & & & 1 \end{pmatrix} \Longrightarrow g(U) =$$



$. \qquad (33)$

Recall the Hopf monoids $\boldsymbol{L}$ and $\boldsymbol{G}$ and the notion of Hadamard product from Section 1. Let

$$\phi : \boldsymbol{L} \times \boldsymbol{G} \to \mathsf{f}(\mathrm{U})$$

be the map with components

$$(\boldsymbol{L} \times \boldsymbol{G})[I] \to \mathsf{f}(\mathrm{U})[I]$$

given as follows. On a basis element $\ell \otimes m_g \in \boldsymbol{L}[I] \otimes \boldsymbol{G}[I] = (\boldsymbol{L} \times \boldsymbol{G})[I]$, we set

$$\phi(\ell \otimes m_g) = \sum_{\substack{U \in \mathrm{U}(I,\ell) \\ g(U) = g}} \kappa_U \in \mathsf{f}(\mathrm{U}(I, \ell)) \subseteq \mathsf{f}(\mathrm{U})[I] \tag{34}$$

and extend by linearity. The map relates the m-basis of $\boldsymbol{G}$ to the basis of characteristic functions (18) of $\mathsf{f}(\mathrm{U})$.

Proposition 8. *Let $\mathbb{F}$ be an arbitrary finite field. The map $\phi : \boldsymbol{L} \times \boldsymbol{G} \to \mathsf{f}(\mathrm{U})$ is an injective morphism of Hopf monoids.*

Proof. From the definition of the Hopf monoid operations on a Hadamard product and formulas (6) and (9), it follows that

$$\mu_{S_1,S_2}\big((\ell_1 \otimes m_{g_1}) \otimes (\ell_2 \otimes m_{g_2})\big) = \sum_{\substack{g|_{S_1} = g_1 \\ g|_{S_2} = g_2}} \ell_1 \cdot \ell_2 \otimes m_g.$$

Comparing with (29), we see that products are preserved since given $U \in \mathrm{U}(I, \ell)$, we have

$$g(U_{S_i}) = g(U)|_{S_i}.$$

The verification for coproducts is similar, employing (6), (10), and (30) and the fact that given $I = S_1 \sqcup S_2$ and $U_i \in \mathrm{U}(S_i, \ell|_{S_i})$, we have

$$g(U_1 \oplus U_2) = g(U_1) \sqcup g(U_2).$$

Consider the map $\psi : \mathsf{f}(\mathrm{U}) \to \boldsymbol{L} \times \boldsymbol{G}$ given by

$$\psi(\kappa_U) = \ell \otimes m_{g(U)} \tag{35}$$

for any $U \in \mathrm{U}(I, \ell)$. Then

$$\psi\phi(\ell \otimes m_g) = (q-1)^{e(g)} \ell \otimes m_g,$$

where q is the cardinality of $\mathbb{F}$ and $e(g)$ is the number of edges in g. Thus, ϕ is injective. □

We mention that the map ψ in (35) is a morphism of comonoids but not of monoids in general.

Assume now that the matrix entries are from $\mathbb{F}_2$, the field with two elements. In this case, the matrix U is uniquely determined by the linear order ℓ and the graph $g(U)$. Therefore, the map ϕ is invertible with inverse ψ.

Corollary 9. *There is an isomorphism of Hopf monoids*

$$\mathsf{f}(\mathrm{U}) \cong \boldsymbol{L} \times \boldsymbol{G}$$

between the Hopf monoid of functions on unitriangular matrices with entries in $\mathbb{F}_2$ and the Hadamard product of the Hopf monoids of linear orders and simple graphs.

On an arbitrary function $f : \mathrm{U}(I, \ell) \to \mathbb{k}$, the isomorphism is given by

$$\psi(f) = \ell \otimes \sum_{U \in \mathrm{U}(I,\ell)} f(U) m_{g(U)}.$$

The coefficients of the m-basis elements are the values of f.

4. A Hopf monoid of superclass functions

An abstract notion of *superclass* (and *supercharacter*) has been introduced by Diaconis and Isaacs [2008]. We only need a minimal amount of related concepts that we review in Sections 4.1 and 4.2. For this purpose, we first place ourselves in the setting of algebra groups. In Section 4.2, we construct a Hopf monoid structure on the collection of spaces of superclass functions on the unitriangular groups by the same procedure as that in Section 3. The combinatorics of these superclasses is understood from the thesis of Yan [2001] (reviewed in slightly different terms in Section 4.3), and this allows us to obtain an explicit description for the Hopf monoid operations in Section 4.4. This leads to a theorem in Section 4.5 identifying the Hopf monoid of superclass functions with matrix entries in $\mathbb{F}_2$ to the Hadamard product of the Hopf monoids of linear orders and set partitions. The combinatorial models for functions and for superclass functions are related in Section 4.6.

4.1. ***Superclass functions on algebra groups.*** Let $\mathfrak{n}$ be a *nilpotent* algebra: an associative, nonunital algebra in which every element is nilpotent. Let $\bar{\mathfrak{n}} = \mathbb{F} \oplus \mathfrak{n}$ denote the result of adjoining a unit to $\mathfrak{n}$. The set

$$G(\mathfrak{n}) = \{1 + n \mid n \in \mathfrak{n}\}$$

is a subgroup of the group of invertible elements of $\bar{\mathfrak{n}}$. A group of this form is called an *algebra group*. (This is the terminology employed in [Diaconis and Isaacs 2008] and, in a slightly different context, [André 1999; Isaacs 1995].)

A morphism of nilpotent algebras $\varphi : \mathfrak{m} \to \mathfrak{n}$ has a unique unital extension $\overline{\mathfrak{m}} \to \bar{\mathfrak{n}}$, and this sends $G(\mathfrak{m})$ to $G(\mathfrak{n})$. A *morphism of algebra groups* is a map of this form.

Warning. When we refer to the algebra group $G(\mathfrak{n})$, it is implicitly assumed that the algebra $\mathfrak{n}$ is given as well.

Following Yan [2001], we define an equivalence relation on $G(\mathfrak{n})$ as follows. Given $x, y \in G(\mathfrak{n})$, we write $x \sim y$ if there exist $g, h \in G(\mathfrak{n})$ such that

$$y - 1 = g(x - 1)h. \tag{36}$$

Following now Diaconis and Isaacs [2008], we refer to the equivalence classes of this relation as *superclasses* and to the functions $G(\mathfrak{n}) \to \Bbbk$ constant on these classes as *superclass functions*. The set of such functions is denoted $\mathsf{scf}(G(\mathfrak{n}))$.

Since

$$gxg^{-1} - 1 = g(x - 1)g^{-1},$$

we have that $x \sim gxg^{-1}$ for any $x, g \in G(\mathfrak{n})$. Thus, each superclass is a union of conjugacy classes, and hence, every superclass function is a class function:

$$\mathsf{scf}(G(\mathfrak{n})) \subseteq \mathsf{cf}(G(\mathfrak{n})). \tag{37}$$

A morphism $\varphi : G(\mathfrak{m}) \to G(\mathfrak{n})$ of algebra groups preserves the relation $\sim$. Therefore, if $f : G(\mathfrak{n}) \to \Bbbk$ is a superclass function on $G(\mathfrak{n})$, then $f \circ \varphi$ is a superclass function on $G(\mathfrak{m})$. In this manner,

$$\mathsf{scf} : \{\text{algebra groups}\} \to \{\text{vector spaces}\}$$

is a contravariant functor. In addition, the inclusion (37) is natural with respect to morphisms of algebra groups.

The direct product of two algebra groups is another algebra group. Indeed,

$$G(\mathfrak{n}_1) \times G(\mathfrak{n}_2) \cong G(\mathfrak{n}_1 \oplus \mathfrak{n}_2)$$

and $\mathfrak{n}_1 \oplus \mathfrak{n}_2$ is nilpotent. Moreover,

$$(x_1, x_2) \sim (y_1, y_2) \iff (x_1 \sim y_1 \text{ and } x_2 \sim y_2).$$

Therefore, a superclass of the product is a pair of superclasses from the factors, and if at least one of the two groups is finite, there is a canonical isomorphism

$$\mathsf{scf}(G(\mathfrak{n}_1) \times G(\mathfrak{n}_2)) \cong \mathsf{scf}(G(\mathfrak{n}_1)) \otimes \mathsf{scf}(G(\mathfrak{n}_2)).$$

4.2. ***Superclass functions on unitriangular matrices.*** Given a finite set I and a linear order ℓ on I, let $\mathfrak{n}(I,\ell)$ denote the subalgebra of $\mathrm{M}(I)$ consisting of strictly upper-triangular matrices

$$N=(n_{ij})_{i,j\in I},\quad n_{ij}=0 \text{ whenever } i\geq_\ell j.$$

Then $\mathfrak{n}(I,\ell)$ is nilpotent and $G(\mathfrak{n}(I,\ell))=\mathrm{U}(I,\ell)$. Thus, the unitriangular groups are algebra groups.

We assume from now on that the field $\mathbb{F}$ is finite.

We define, for each finite set I,

$$\mathsf{scf}(\mathrm{U})[I]=\bigoplus_{\ell\in\mathrm{L}[I]}\mathsf{scf}(\mathrm{U}(I,\ell)).$$

This defines a species $\mathsf{scf}(\mathrm{U})$. Proceeding in the same manner as in Sections 3.3 and 3.5, we obtain linear maps

$$\mathsf{scf}(\mathrm{U})[S_1]\otimes\mathsf{scf}(\mathrm{U})[S_2]\underset{\Delta_{S_1,S_2}}{\overset{\mu_{S_1,S_2}}{\rightleftarrows}}\mathsf{scf}(\mathrm{U})[I]$$

by applying the functor scf to the morphisms π_{S_1,S_2} and σ_{S_1,S_2}. This is meaningful since the latter are morphisms of algebra groups: it was noted in Section 2.1 that σ_{S_1,S_2} is the restriction of a morphism defined on the full matrix algebras while the considerations of Lemma 2 show that π_{S_1,S_2} is the restriction of a morphism defined on the algebra of upper-triangular matrices.

Proposition 10. *With these operations, the species* $\mathsf{scf}(\mathrm{U})$ *is a connected cocommutative Hopf monoid. It is a Hopf submonoid of* $\mathsf{cf}(\mathrm{U})$.

Proof. As in the proof of Proposition 4, the first statement follows by functoriality. The second follows from the naturality of the inclusion (37). □

(31) and (32) continue to hold for the (co)product of superclass functions.

The constant function $\mathbf{1}_\ell$ is a superclass function. Thus, the morphism of Hopf monoids of Corollary 5 factors through $\mathsf{scf}(\mathrm{U})$ and $\mathsf{cf}(\mathrm{U})$:

$$\boldsymbol{L}\hookrightarrow\mathsf{scf}(\mathrm{U})\hookrightarrow\mathsf{cf}(\mathrm{U})\hookrightarrow\mathsf{f}(\mathrm{U}).$$

4.3. ***Combinatorics of the superclasses.*** Yan [2001] showed superclasses are parametrized by certain combinatorial data essentially along the lines presented below.

According to (36), two unitriangular matrices U_1 and U_2 are in the same superclass if and only if $U_2-\mathrm{Id}$ is obtained from $U_1-\mathrm{Id}$ by a sequence of elementary row and column operations. The available operations are from the unitriangular group itself, so the pivot entries cannot be normalized. Thus, each superclass contains a unique matrix U such that $U-\mathrm{Id}$ has at most one nonzero entry in each row and each column. We refer to this matrix U as the *canonical representative* of the superclass.

We proceed to encode such representatives in terms of combinatorial data.

We first discuss the combinatorial data. Let ℓ be a linear order on a finite set I and X a partition of I. Let us say that $i, j \in I$ *bound an arc* if

- i precedes j in ℓ,
- i and j are in the same block of X, say S, and
- no other element of S lies between i and j in the order ℓ.

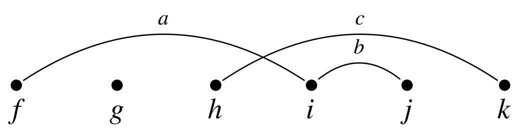
The set of *arcs* is

$$A(X,\ell) := \{(i,j) \mid i \text{ and } j \text{ bound an arc}\}.$$

Consider also a function

$$\alpha : A(X,\ell) \to \mathbb{F}^{\times}$$

from the set of arcs to the nonzero elements of $\mathbb{F}$. We say that the pair (X,α) is an *arc diagram* on the linearly ordered set (I,ℓ). We may visualize an arc diagram:



Here the combinatorial data is

$$\ell = fghijk, \quad X = \{\{f,i,j\},\{g\},\{h,k\}\}, \quad \alpha(f,i) = a,\ \alpha(i,j) = b,\ \alpha(h,k) = c.$$

Fix the linear order ℓ. To an arc diagram (X,α) on (I,ℓ), we associate a matrix $U_{X,\alpha}$ with entries

$$u_{ij} = \begin{cases} \alpha(i,j) & \text{if } (i,j) \in A(X,\ell), \\ 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, the matrix $U_{X,\alpha}$ is ℓ-unitriangular and $U_{X,\alpha} - \mathrm{Id}$ has at most one nonzero entry in each row and each column. In the above example,

$$U_{X,\alpha} = \begin{pmatrix} 1 & 0 & 0 & a & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & c \\ & & & 1 & d & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{pmatrix}.$$

Conversely, any canonical representative matrix $U \in \mathrm{U}(I,\ell)$ is of the form $U_{X,\alpha}$ for a unique arc diagram (X,α) on (I,ℓ): the location of the nonzero entries determines the set of arcs, and the values of the entries determine the function α. The smallest equivalence relation on I containing the set of arcs determines the partition X.

In conclusion, the canonical representatives, and hence the superclasses, are in bijection with the set of arc diagrams. We let $C_{X,\alpha}$ denote the superclass of $\mathrm{U}(I, \ell)$ containing $U_{X,\alpha}$, and we write $\kappa_{X,\alpha}$ for the characteristic function of this class. As (X, α) runs over all arc diagrams on (I, ℓ), these functions form a basis of the space $\mathsf{scf}(\mathrm{U}(I, \ell))$.

We describe principal minors and direct sums of the canonical representatives. To this end, fix $\ell \in \mathrm{L}[I]$ and recall the notions of union and restriction of set partitions discussed in Section 1.3.

Let $S \subseteq I$ be an arbitrary subset. Given a partition X of I, let $A(X, \ell)|_S$ denote the subset of $A(X, \ell)$ consisting of those arcs (i, j) where both i and j belong to S. We let $\alpha|_S$ denote the restriction of α to $A(X, \ell)|_S$. We have $A(X, \ell)|_S \subseteq A(X|_S, \ell|_S)$, and if S is an ℓ-segment, then

$$A(X, \ell)|_S = A(X|_S, \ell|_S). \tag{38}$$

In this case, we obtain an arc diagram $(X|_S, \alpha|_S)$ on $(S, \ell|_S)$, and we have

$$(U_{X,\alpha})_S = U_{X|_S,\alpha|_S}. \tag{39}$$

Suppose now that $I = S_1 \sqcup S_2$ and (X_i, α_i) is an arc diagram on $(S_i, \ell|_{S_i})$, $i = 1, 2$. Then

$$A(X_1 \sqcup X_2, \ell) = A(X_1, \ell|_{S_1}) \sqcup A(X_2, \ell|_{S_2}). \tag{40}$$

Let $\alpha_1 \sqcup \alpha_2$ denote the common extension of α_1 and α_2 to this set. Then the pair $(X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2)$ is then an arc diagram on (I, ℓ) and

$$U_{X_1,\alpha_1} \oplus U_{X_2,\alpha_2} = U_{X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2}. \tag{41}$$

4.4. ***Combinatorics of the (co)product.*** We now describe the product and coproduct of the Hopf monoid $\mathsf{scf}(\mathrm{U})$ on the basis $\{\kappa_{X,\alpha}\}$ of Section 4.3. We employ (31) and (32), which, as discussed in Section 4.2, hold for superclass functions.

Let $I = S_1 \sqcup S_2$ and $\ell_i \in \mathrm{L}[S_i]$, $i = 1, 2$, and consider the product

$$\mathsf{scf}(\mathrm{U}(S_1, \ell_1)) \times \mathsf{scf}(\mathrm{U}(S_2, \ell_2)) \to \mathsf{scf}(\mathrm{U}(I, \ell_1 \cdot \ell_2)).$$

Let (X_i, α_i) be an arc diagram on (I, ℓ_i), $i = 1, 2$. According to (31), we have

$$\mu_{S_1,S_2}(\kappa_{X_1,\alpha_1} \otimes \kappa_{X_2,\alpha_2}) = \sum_{\pi_{S_1,S_2}(C_{X,\alpha}) \subseteq C_{X_1,\alpha_1} \times C_{X_2,\alpha_2}} \kappa_{X,\alpha},$$

a sum over arc diagrams (X, α) on $(I, \ell_1 \cdot \ell_2)$. Since $\pi_{S,T}$ preserves superclasses,

$$\begin{aligned} \pi_{S_1,S_2}(C_{X,\alpha}) \subseteq C_{X_1,\alpha_1} \times C_{X_2,\alpha_2} &\iff \pi_{S_1,S_2}(U_{X,\alpha}) \in C_{X_1,\alpha_1} \times C_{X_2,\alpha_2} \\ &\iff (U_{X,\alpha})_{S_i} \in C_{X_i,\alpha_i}, \quad i = 1, 2. \end{aligned}$$

In view of (39), this is in turn equivalent to

$$X|_{S_i} = X_i \quad \text{and} \quad \alpha|_{S_i}, = \alpha_i, \quad i = 1, 2.$$

In conclusion,

$$\mu_{S_1,S_2}(\kappa_{X_1,\alpha_1} \otimes \kappa_{X_2,\alpha_2}) = \sum_{\substack{X|_{S_i}=X_i \\ \alpha|_{S_i}=\alpha_i}} \kappa_{X,\alpha}. \tag{42}$$

The sum is over all arc diagrams (X, α) on $(I, \ell_1 \cdot \ell_2)$ whose restriction to S_i is (X_i, α_i) for $i = 1, 2$.

Take now $\ell \in \mathrm{L}[I]$, $I = S_1 \sqcup S_2$, and consider the coproduct

$$\mathsf{scf}(\mathrm{U}(I, \ell)) \to \mathsf{scf}(\mathrm{U}(S_1, \ell|_{S_1})) \times \mathsf{scf}(\mathrm{U}(S_2, \ell|_{S_2})).$$

Let (X, α) be an arc diagram on (I, ℓ). According to (32), we have

$$\Delta_{S_1,S_2}(\kappa_{X,\alpha}) = \sum_{\sigma_{S_1,S_2}(C_{X_1,\alpha_1} \times C_{X_2,\alpha_2}) \subseteq C_{X,\alpha}} \kappa_{X_1,\alpha_1} \otimes \kappa_{X_2,\alpha_2},$$

a sum over arc diagrams (X_i, α_i) on $(S_i, \ell|_{S_i})$. The superclass $C_{X_1,\alpha_1} \times C_{X_2,\alpha_2}$ contains $(U_{X_1,\alpha_1}, U_{X_2,\alpha_2})$, and hence, its image under σ_{S_1,S_2} contains

$$U_{X_1,\alpha_1} \oplus U_{X_2,\alpha_2} = U_{X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2}$$

by (41). Therefore,

$$\sigma_{S_1,S_2}(C_{X_1,\alpha_1} \times C_{X_2,\alpha_2}) \subseteq C_{X,\alpha} \iff X_1 \sqcup X_2 = X \quad \text{and} \quad \alpha_1 \sqcup \alpha_2 = \alpha.$$

Note that $X_1 \sqcup X_2 = X$ if and only if S_1 (or equivalently, S_2) is a union of blocks of X. In this case, $X_i = X|_{S_i}$ and $\alpha_i = \alpha|_{S_i}$. In conclusion,

$$\Delta_{S_1,S_2}(\kappa_{X,\alpha}) = \begin{cases} \kappa_{X|_{S_1},\alpha|_{S_1}} \otimes \kappa_{X|_{S_2},\alpha|_{S_2}} & \text{if } S_1 \text{ is the union of some blocks of } X, \\ 0 & \text{otherwise.} \end{cases} \tag{43}$$

4.5. ***Decomposition as a Hadamard product.*** The apparent similarity between the combinatorial description of the Hopf monoid operations of $\mathsf{scf}(\mathrm{U})$ in Section 4.4 and those of the Hopf monoids $\boldsymbol{L}$ and $\boldsymbol{\Pi}$ in Sections 1.2 and 1.3 can be formalized. Recall the Hadamard product of Hopf monoids from Section 1.5.

Let

$$\phi : \boldsymbol{L} \times \boldsymbol{\Pi} \to \mathsf{scf}(\mathrm{U})$$

be the map with components

$$(\boldsymbol{L} \times \boldsymbol{\Pi})[I] \to \mathsf{scf}(\mathrm{U})[I]$$

given as follows. On a basis element $\ell \otimes m_X \in \boldsymbol{L}[I] \otimes \boldsymbol{\Pi}[I] = (\boldsymbol{L} \times \boldsymbol{\Pi})[I]$, we set

$$\phi(\ell \otimes m_X) = \sum_{\alpha : A(X,\ell) \to \mathbb{F}^\times} \kappa_{X,\alpha} \in \mathsf{scf}(\mathrm{U}(I,\ell)) \subseteq \mathsf{scf}(\mathrm{U})[I] \tag{44}$$

and extend by linearity. The morphism ϕ adds labels to the arcs in all possible ways.

Proposition 11. *Let $\mathbb{F}$ be an arbitrary finite field. The map $\phi : \boldsymbol{L} \times \boldsymbol{\Pi} \to \mathsf{scf}(\mathrm{U})$ is an injective morphism of Hopf monoids.*

Proof. This follows by comparing definitions, as in the proof of Proposition 8. The relevant equations are (6), (7), and (8) for the operations of $\boldsymbol{L} \times \boldsymbol{\Pi}$ and (42) and (43) for the operations of $\mathsf{scf}(\mathrm{U})$. □

When the field of matrix entries is $\mathbb{F}_2$, the arc labels are uniquely determined. The map ϕ is then invertible with inverse ψ given by

$$\psi(\kappa_{X,\alpha}) = \ell \otimes m_X$$

for any arc diagram (X, α) on a linearly ordered set (I, ℓ). We thus have:

Corollary 12. *There is an isomorphism of Hopf monoids*

$$\mathsf{scf}(\mathrm{U}) \cong \boldsymbol{L} \times \boldsymbol{\Pi}$$

between the Hopf monoid of superclass functions on unitriangular matrices with entries in $\mathbb{F}_2$ and the Hadamard product of the Hopf monoids of linear orders and set partitions.

4.6. ***Relating the combinatorial models.*** The results of Section 4.5 provide a combinatorial model for the Hopf monoid $\mathsf{scf}(\mathrm{U})$. They parallel those of Section 3.6 that do the same for $\mathsf{f}(\mathrm{U})$. We now interpret the inclusion $\mathsf{scf}(\mathrm{U}) \hookrightarrow \mathsf{f}(\mathrm{U})$ in these terms.

Let X be a partition on a linearly ordered set (I, ℓ). We may regard the set of arcs $A(X, \ell)$ as a simple graph on I. Let $G(X, \ell)$ denote the set of simple graphs g on I such that

- g contains the graph $A(X, \ell)$ and
- if $i < j$ in ℓ and $g \setminus A(X, \ell)$ contains an edge between i and j, then there exists k such that

$$i < k < j \text{ in } \ell \text{ and either } (i,k) \in A(X,\ell) \text{ or } (k,j) \in A(X,\ell).$$

The following illustrates the extra edges (dotted) that may be present in g when an arc (solid) is present in $A(X, \ell)$:

Define a map

$$\boldsymbol{L} \times \boldsymbol{\Pi} \to \boldsymbol{L} \times \boldsymbol{G}$$

with components

$$(\boldsymbol{L} \times \boldsymbol{\Pi})[I] \to (\boldsymbol{L} \times \boldsymbol{G})[I], \quad \ell \otimes m_X \mapsto \ell \otimes \sum_{g \in G(X,\ell)} m_g.$$

Proposition 13. *The map $\boldsymbol{L} \times \boldsymbol{\Pi} \to \boldsymbol{L} \times \boldsymbol{G}$ is an injective morphism of Hopf monoids. Moreover, the following diagram commutes*:

$$\begin{tikzcd}
\boldsymbol{L} \times \boldsymbol{\Pi} \arrow[r, hook] \arrow[d, "\phi"'] & \boldsymbol{L} \times \boldsymbol{G} \arrow[d, "\phi"] \\
\mathsf{scf}(\mathrm{U}) \arrow[r, hook] & \mathsf{f}(\mathrm{U})
\end{tikzcd}$$

Proof. It is enough to prove the commutativity of the diagram since all other maps in the diagram are injective morphisms. The commutativity boils down to the following fact. Given $X \in \Pi[I]$, $\ell \in \mathrm{L}[I]$, and $U \in \mathrm{U}(I, \ell)$,

$$U \in C_{X,\alpha} \text{ for some } \alpha : A(X, \ell) \to \mathbb{F}^{\times} \iff g(U) \in G(X, \ell).$$

This expresses the fact that a matrix U belongs to the superclass $C_{X,\alpha}$ if and only if the nonzero entries of $U - \mathrm{Id}$ are located either above or to the right of the nonzero entries of the representative $U_{X,\alpha}$. $\square$

5. Freeness

We prove that the Hopf monoids $\mathsf{f}(\mathrm{U})$ and $\mathsf{scf}(\mathrm{U})$ are free and the Hopf structure is isomorphic to the canonical one on a free monoid. We assume that the base field $\Bbbk$ is of characteristic 0, which enables us to apply the results of the Appendix.

5.1. ***A partial order on arc diagrams.*** Let (I, ℓ) be a linearly ordered set. Given arc diagrams (X, α) and (Y, β) on (I, ℓ), we write

$$(X, \alpha) \leq (Y, \beta)$$

if

$$A(X, \ell) \subseteq A(Y, \ell) \quad \text{and} \quad \beta|_{A(X,\ell)} = \alpha.$$

In other words, every arc of X is an arc of Y and with the same label. In particular, the partition Y is coarser than X. On the other hand, the following arc diagrams are incomparable (regardless of the labels) even though the partition on the right is the coarsest one:

The poset of arc diagrams has a unique minimum (the partition into singletons, for which there are no arcs) but several maximal elements. The arc diagrams above are the two maximal elements when $\ell = ijk$ (up to a choice of labels).

A partition X of the linearly ordered set (I, ℓ) is *atomic* if no proper initial ℓ-segment of I is a union of blocks of X. Equivalently, there is no decomposition $I = S_1 \sqcup S_2$ into proper ℓ-segments such that $X = X|_{S_1} \sqcup X|_{S_2}$.

atomic nonatomic

If (X, α) is a maximal element of the poset of arc diagrams, then X is an atomic partition. But if X is atomic, (X, α) need not be maximal (regardless of α).

maximal ($\Longrightarrow$atomic) atomic not maximal

5.2. ***A second basis for*** **scf(U).** We employ the partial order of Section 5.1 to define a new basis $\{\lambda_{X,\alpha}\}$ of $\mathbf{scf}(\mathrm{U}(I, \ell))$ by

$$\lambda_{X,\alpha} = \sum_{(X,\alpha)\le(Y,\beta)} \kappa_{Y,\beta}. \tag{45}$$

The product of the Hopf monoid $\mathbf{scf}(\mathrm{U})$ takes a simple form on the λ-basis.

Proposition 14. *Let* $I = S_1 \sqcup S_2$ *and* $\ell_i \in \mathrm{L}[S_i]$, $i = 1, 2$. *Then*

$$\mu_{S_1,S_2}(\lambda_{X_1,\alpha_1} \otimes \lambda_{X_2,\alpha_2}) = \lambda_{X_1\sqcup X_2,\alpha_1\sqcup\alpha_2} \tag{46}$$

for any arc diagrams (X_i, α_i) *on* (S_i, ℓ_i), $i = 1, 2$.

Proof. We calculate using (42) and (45):

$$\mu_{S_1,S_2}(\lambda_{X_1,\alpha_1}\otimes\lambda_{X_2,\alpha_2}) = \sum_{(X_i,\alpha_i)\le(Y_i,\beta_i)} \mu_{S_1,S_2}(\kappa_{Y_1,\beta_1}\otimes\kappa_{Y_2,\beta_2}) = \sum_{(X_i,\alpha_i)\le(Y|_{S_i},\beta|_{S_i})} \kappa_{Y,\beta}.$$

Now by (38) and (40), we have

$$(X_i, \alpha_i) \le (Y|_{S_i}, \beta|_{S_i}),\ i = 1, 2 \iff (X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2) \le (Y, \beta).$$

Therefore,

$$\mu_{S_1,S_2}(\lambda_{X_1,\alpha_1} \otimes \lambda_{X_2,\alpha_2}) = \sum_{(X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2) \leq (Y,\beta)} \kappa_{Y,\beta} = \lambda_{X_1 \sqcup X_2, \alpha_1 \sqcup \alpha_2}. \qquad \square$$

The coproduct of the Hopf monoid $\mathbf{scf}(\mathrm{U})$ takes the same form on the λ-basis as on the κ-basis.

Proposition 15. *Let $I = S_1 \sqcup S_2$ and $\ell \in \mathrm{L}[I]$. Then*

$$\Delta_{S_1,S_2}(\lambda_{X,\alpha}) = \begin{cases} \lambda_{X|_{S_1},\alpha|_{S_1}} \otimes \lambda_{X|_{S_2},\alpha|_{S_2}} & \text{if } S_1 \text{ is the union of some blocks of } X, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

Proof. Suppose first that S_1 is not the union of blocks of X. Then the same is true for any partition coarser than X, in particular for any partition Y entering in (45). In view of (43), we then have $\Delta_{S_1,S_2}(\lambda_{X,\alpha}) = 0$.

Otherwise, $X = X|_{S_1} \sqcup X|_{S_2}$ and $\alpha = \alpha|_{S_1} \sqcup \alpha|_{S_2}$. Among the arc diagrams (Y, β) entering in (45), only those for which $Y = Y|_{S_1} \sqcup Y|_{S_2}$ contribute to the coproduct, in view of (43). These arc diagrams are of the form $Y = Y_1 \sqcup Y_2$ and $\beta = \beta_1 \sqcup \beta_2$, and by (40), we must have

$$A(X|_{S_i}, \ell|_{S_i}) \subseteq A(Y_i, \ell|_{S_i}) \quad \text{and} \quad \beta_i|_{A(X|_{S_i},\ell|_{S_i})} = \alpha|_{S_i}, \quad i = 1, 2.$$

We then have

$$\begin{aligned} \Delta_{S_1,S_2}(\lambda_{X,\alpha}) &= \sum_{(X,\alpha) \leq (Y,\beta)} \Delta(\kappa_{Y,\beta}) \\ &= \sum_{(X|_{S_i},\alpha|_{S_i}) \leq (Y_i,\beta_i)} \kappa_{Y_1,\beta_1} \otimes \kappa_{Y_2,\beta_2} = \lambda_{X|_{S_1},\alpha|_{S_1}} \otimes \lambda_{X|_{S_2},\alpha|_{S_2}}. \qquad \square \end{aligned}$$

Remark 16. The relationship between the λ- and κ-bases of $\mathbf{scf}(\mathrm{U})$ is somewhat reminiscent of that between the p- and m-bases of $\mathbf{\Pi}$ [Aguiar and Mahajan 2010, Equation (12.5)]. However, the latter involves a sum over all partitions coarser than a given one. For this reason, the morphism ϕ in (44), which relates the m-basis to the κ-basis, does not relate the p-basis to the λ-basis in the same manner.

5.3. ***Freeness of* $\mathbf{scf}(\mathrm{U})$.** Let $\boldsymbol{q}$ be a species such that $\boldsymbol{q}[\varnothing] = 0$. A new species $\mathcal{T}(\boldsymbol{q})$ is defined by $\mathcal{T}(\boldsymbol{q})[\varnothing] = \Bbbk$ and, on a finite nonempty set I,

$$\mathcal{T}(\boldsymbol{q})[I] = \bigoplus_{\substack{I = I_1 \sqcup \cdots \sqcup I_k \\ k \geq 1,\ I_j \neq \varnothing}} \boldsymbol{q}[I_1] \otimes \cdots \otimes \boldsymbol{q}[I_k].$$

The sum is over all decompositions of I into nonempty subsets. The number k of subsets is therefore bounded above by $|I|$.

The species $\mathcal{T}(\boldsymbol{q})$ is a connected monoid with product given by concatenation. To describe this in detail, let $I = S \sqcup T$ and choose decompositions $S = S_1 \sqcup \cdots \sqcup S_k$ and $T = T_1 \sqcup \cdots \sqcup T_l$ and elements $x_i \in \boldsymbol{q}[S_i]$, $i = 1, \ldots, k$, and $y_j \in \boldsymbol{q}[T_j]$, $j = 1, \ldots, l$. Write

$$x = x_1 \otimes \cdots \otimes x_i \in \mathcal{T}(\boldsymbol{q})[S] \quad \text{and} \quad y = y_1 \otimes \cdots \otimes y_j \in \mathcal{T}(\boldsymbol{q})[T].$$

The product is

$$\begin{aligned} \mu_{S,T}(x \otimes y) &= x_1 \otimes \cdots \otimes x_i \otimes y_1 \otimes \cdots \otimes y_j \\ &\qquad \in \boldsymbol{q}[S_1] \otimes \cdots \otimes \boldsymbol{q}[S_k] \otimes \boldsymbol{q}[T_1] \otimes \cdots \otimes \boldsymbol{q}[T_l] \subseteq \mathcal{T}(\boldsymbol{q})[I]. \end{aligned}$$

The monoid $\mathcal{T}(\boldsymbol{q})$ is *free* on the species $\boldsymbol{q}$: a map of species $\boldsymbol{q} \to \boldsymbol{m}$ from $\boldsymbol{q}$ to a monoid $\boldsymbol{m}$ has a unique extension to a morphism of monoids $\mathcal{T}(\boldsymbol{q}) \to \boldsymbol{m}$.

The monoid $\mathcal{T}(\boldsymbol{q})$ may carry several coproducts that turn it into a connected Hopf monoid. The *canonical* structure is the one for which the elements of $\boldsymbol{q}$ are *primitive*. This means that

$$\Delta_{S,T}(x) = 0$$

for every $x \in \boldsymbol{q}[I]$ and every decomposition $I = S \sqcup T$ into nonempty subsets.

More details can be found in [Aguiar and Mahajan 2010, Sections 11.2.1–11.2.2].

Let $\mathrm{D}(I, \ell)$ denote the set of arc diagrams (X, α) for which X is an atomic set partition of the linearly ordered set (I, ℓ). Let $\boldsymbol{d}(I, \ell)$ be the vector space with basis $\mathrm{D}(I, \ell)$. Define a species $\boldsymbol{d}$ by

$$\boldsymbol{d}[I] = \bigoplus_{\ell \in \mathrm{L}[I]} \boldsymbol{d}(I, \ell).$$

Consider the map of species $\boldsymbol{d} \to \mathsf{scf}$ with components

$$\boldsymbol{d}[I] \to \mathsf{scf}(\mathrm{U})[I], \qquad (X, \alpha) \mapsto \lambda_{X,\alpha}.$$

The map sends the summand $\boldsymbol{d}(I, \ell)$ of $\boldsymbol{d}[I]$ to the summand $\mathsf{scf}(\mathrm{U}(I, \ell))$ of $\mathsf{scf}(\mathrm{U})[I]$. By freeness, it extends to a morphism of monoids

$$\mathcal{T}(\boldsymbol{d}) \to \mathsf{scf}(\mathrm{U}).$$

Proposition 17. *The map* $\mathcal{T}(\boldsymbol{d}) \to \mathsf{scf}(\mathrm{U})$ *is an isomorphism of monoids. In particular, the monoid* $\mathsf{scf}(\mathrm{U})$ *is free.*

Proof. Let (X, α) be an arbitrary arc diagram on (I, ℓ). Let $I_1, \ldots, I_k$ be the minimal ℓ-segments of I, numbered from left to right, such that each I_j is a union of blocks of X. Let $\ell_j = \ell|_{I_j}$, $X_j = X|_{I_j}$, and $\alpha_j = \alpha|_{I_j}$. Then X_j is an atomic partition of (I_j, ℓ_j),

$$X_1 \sqcup \cdots \sqcup X_j = X \quad \text{and} \quad \alpha_1 \sqcup \cdots \sqcup \alpha_j = \alpha.$$

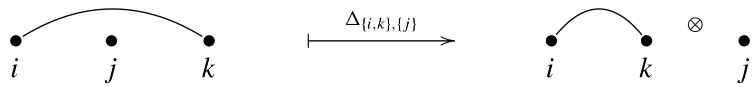
By (46),

$$\mu_{I_1,\dots,I_k}(\lambda_{X_1,\alpha_1}\otimes\cdots\otimes\lambda_{X_k,\alpha_k})=\lambda_{X,\alpha}.$$

Thus, the morphism $\mathcal{T}(\boldsymbol{d})\to\mathbf{scf}(\mathrm{U})$ sends the basis element $(X_1,\alpha_1)\otimes\cdots\otimes(X_k,\alpha_k)$ of $\boldsymbol{d}(I_1,\ell_1)\otimes\cdots\otimes\boldsymbol{d}(I_k,\ell_k)$ to the basis element $\lambda_{X,\alpha}$ of $\mathbf{scf}(\mathrm{U}(I,\ell))$ and is therefore an isomorphism. □

We may state Proposition 17 by saying that the superclass functions $\lambda_{X,\alpha}$ freely generate the monoid $\mathbf{scf}(\mathrm{U})$ as (X,α) runs over all arc diagrams for which X is an atomic set partition.

The generators, however, need not be primitive. For instance,



which is not 0. Nevertheless, Proposition 23 allows us to conclude the following:

Corollary 18. *Let $\Bbbk$ be a field of characteristic* 0. *There exists an isomorphism of Hopf monoids* $\mathbf{scf}(\mathrm{U})\cong\mathcal{T}(\boldsymbol{d})$, *where the latter is endowed with its canonical Hopf structure.*

As discussed in the Appendix, an isomorphism may be constructed with the aid of the first Eulerian idempotent.

Let $\boldsymbol{\Pi}_a(I,\ell)$ be the vector space with basis the set of atomic partitions on (I,ℓ). When the field of matrix entries is $\mathbb{F}_2$, arc diagrams reduce to atomic set partitions and $\boldsymbol{d}(I,\ell)$ identifies with $\boldsymbol{\Pi}_a(I,\ell)$. Combining Corollaries 12 and 18, we obtain an isomorphism of Hopf monoids

$$\boldsymbol{L}\times\boldsymbol{\Pi}\cong\mathcal{T}(\boldsymbol{\Pi}_a), \tag{48}$$

where

$$\boldsymbol{\Pi}_a[I]=\bigoplus_{\ell\in\mathrm{L}[I]}\boldsymbol{\Pi}_a(I,\ell).$$

5.4. ***A second basis for G and for* f(U).** Given two unitriangular matrices U and $V\in\mathrm{U}(I,\ell)$, let us write $U\le V$ if

$$u_{ij}=v_{ij}\quad\text{whenever } u_{ij}\neq 0.$$

In other words, some zero entries in U may be nonzero in V; the other entries are the same for both matrices.

We define a new basis $\{\lambda_U\}$ of $\mathbf{f}(\mathrm{U}(I,\ell))$ by

$$\lambda_U=\sum_{U\le V}\kappa_V.$$

Let $I = S_1 \sqcup S_2$, $U \in \mathrm{U}(I,\ell)$, and $g_i \in \mathrm{U}(S_i,\ell_i)$, $i = 1, 2$. It is easy to derive the following formulas from (29) and (30):

$$\mu_{S_1,S_2}(\lambda_{U_1} \otimes \lambda_{U_2}) = \lambda_{U_1 \oplus U_2}, \tag{49}$$

$$\Delta_{S_1,S_2}(\lambda_U) = \begin{cases} \lambda_{U_{S_1}} \otimes \lambda_{U_{S_2}} & \text{if } U = U_{S_1} \oplus U_{S_2}, \\ 0 & \text{otherwise.} \end{cases} \tag{50}$$

(49) implies that $\mathsf{f}(\mathrm{U})$ is a free monoid with generators λ_U indexed by unitriangular matrices U for which the graph $g(U)$ is connected.

For completeness, one may define a new basis $\{p_g\}$ of $\boldsymbol{G}[I]$ by

$$p_g = \sum_{g \subseteq h} m_h. \tag{51}$$

The sum is over all simple graphs h with vertex set I and with the same or more edges than g. Let $I = S_1 \sqcup S_2$, $g \in \mathrm{G}[I]$, and $g_i \in \mathrm{G}[S_i]$, $i = 1, 2$. From (9) and (10), one obtains

$$\mu_{S_1,S_2}(p_{g_1} \otimes p_{g_2}) = p_{g_1 \sqcup g_2}, \tag{52}$$

$$\Delta_{S_1,S_2}(p_g) = \begin{cases} p_{g|_{S_1}} \otimes p_{g|_{S_2}} & \text{if no edge of } g \text{ connects } S_1 \text{ to } S_2, \\ 0 & \text{otherwise.} \end{cases} \tag{53}$$

Equation (52) implies that $\boldsymbol{G}$ is the free commutative monoid on the species of connected graphs. From (44), we deduce that the morphism ϕ of Proposition 8 takes the following form on these bases:

$$\phi(\ell \otimes p_g) = \sum_{\substack{U \in \mathrm{U}(I,\ell) \\ g(U) = g}} \lambda_U.$$

6. Applications

We conclude with some applications and remarks regarding past and future work.

6.1. ***Counting conjugacy classes.*** Let $k_n(q)$ be the number of conjugacy classes of the group of unitriangular matrices of size n with entries in the field with q elements. Higman's conjecture states that, for fixed n, k_n is a polynomial function of q. Much effort has been devoted to the precise determination of these numbers or their asymptotic behavior [Goodwin 2006; Goodwin and Röhrle 2009; Higman 1960; Robinson 1998; Vera-López and Arregi 2003; Vera-López et al. 2008].

We fix q and let n vary. It turns out that the existence of a Hopf monoid structure on class functions imposes certain linear conditions on the sequence $k_n(q)$, as we explain next.

Given a finite-dimensional Hopf monoid $\boldsymbol{h}$, consider the generating function

$$\mathsf{T}_{\boldsymbol{h}}(x) = \sum_{n\geq 0} \dim_{\Bbbk}(\boldsymbol{h}[n]_{\mathrm{S}_n})x^n. \tag{54}$$

Here $[n]$ denotes the set $\{1, 2, \dots, n\}$ and $\boldsymbol{h}[n]_{\mathrm{S}_n}$ is the (quotient) space of coinvariants for the action of the symmetric group (afforded by the species structure of $\boldsymbol{h}$).

For example, since

$$(\boldsymbol{L} \times \boldsymbol{\Pi})[n]_{\mathrm{S}_n} = (\boldsymbol{L}[n] \otimes \boldsymbol{\Pi}[n])_{\mathrm{S}_n} \cong \boldsymbol{\Pi}[n],$$

we have

$$\mathsf{T}_{\boldsymbol{L}\times\boldsymbol{\Pi}}(x) = \sum_{n\geq 0} B_n x^n, \tag{55}$$

where B_n is the *n-th Bell number*, the number of partitions of the set $[n]$.

On the other hand, from (48),

$$\mathsf{T}_{\boldsymbol{L}\times\boldsymbol{\Pi}}(x) = \mathsf{T}_{\mathcal{T}(\boldsymbol{\Pi}_a)}(x).$$

It is a general fact that, for a species $\boldsymbol{q}$ with $\boldsymbol{q}[\varnothing] = 0$,

$$\mathsf{T}_{\mathcal{T}(\boldsymbol{q})}(x) = \frac{1}{1 - \mathsf{T}_{\boldsymbol{q}}(x)}.$$

(This follows from [Bergeron_F et al. 1998, Theorem 2.b, Section 1.4] for instance). Therefore,

$$\mathsf{T}_{\boldsymbol{L}\times\boldsymbol{\Pi}}(x) = \frac{1}{1 - \sum_{n\geq 1} A_n x^n}, \tag{56}$$

where A_n is the number of atomic partitions of the linearly ordered set $[n]$.

From (55) and (56), we deduce that

$$\sum_{n\geq 0} B_n x^n = \frac{1}{1 - \sum_{n\geq 1} A_n x^n},$$

a fact known from [Bergeron and Zabrocki 2009].

Consider now the injections

$$\mathsf{scf}(\mathrm{U}) \hookrightarrow \mathsf{cf}(\mathrm{U}) \quad \text{and} \quad \boldsymbol{L} \times \boldsymbol{\Pi} \hookrightarrow \mathsf{scf}(\mathrm{U}).$$

Both are morphisms of Hopf monoids (Propositions 10 and 11). *Lagrange's theorem* for Hopf monoids implies in this situation that both quotients

$$\frac{\mathsf{T}_{\mathsf{cf}(\mathrm{U})}(x)}{\mathsf{T}_{\mathsf{scf}(\mathrm{U})}(x)} \quad \text{and} \quad \frac{\mathsf{T}_{\mathsf{scf}(\mathrm{U})}(x)}{\mathsf{T}_{\boldsymbol{L}\times\boldsymbol{\Pi}}(x)}$$

belong to $\mathbb{N}[[x]]$, that is, have nonnegative (integer) coefficients [Aguiar and Lauve 2012, Corollary 13]. In particular,

$$\frac{\mathsf{T}_{\mathbf{cf}(\mathrm{U})}(x)}{\mathsf{T}_{\boldsymbol{L}\times\boldsymbol{\Pi}}(x)} \in \mathbb{N}[[x]]$$

as well.

We have

$$\mathbf{cf}(\mathrm{U})[n]_{\mathrm{S}_n} = \Bigl(\bigoplus_{\ell\in \mathrm{L}[n]} \mathbf{cf}(\mathrm{U}([n],\ell))\Bigr)_{\mathrm{S}_n} \cong \mathbf{cf}(\mathrm{U}([n])).$$

Therefore,

$$\mathsf{T}_{\mathbf{cf}(\mathrm{U})}(x) = \sum_{n\geq 0} k_n(q)x^n.$$

By combining the above, we deduce

$$\Bigl(\sum_{n\geq 0} k_n(q)x^n\Bigr)\Bigl(1-\sum_{n\geq 1} A_n x^n\Bigr) \in \mathbb{N}[[x]],$$

whence the following result:

Corollary 19. *The following linear inequalities are satisfied for every $n\in\mathbb{N}$ and every prime power q:*

$$k_n(q) \geq \sum_{i=0}^{n-1} A_{n-i}k_i(q). \tag{57}$$

For instance, for $n=8$, the inequality is

$$k_6(q) \geq 92 + 22k_1(q) + 6k_2(q) + 2k_3(q) + k_4(q) + k_5(q).$$

Inequality (57) is stronger than merely stating that there are more conjugacy classes than superclasses. For instance, for $q=2$ and $n=6$, the right-hand side of the inequality is 213 (provided we use the correct values for $k_i(2)$ for $i\leq 5$) while there are only $B_6=203$ superclasses. The first few values of the sequence $k_n(2)$ appear in [OEIS Foundation 2010] as A007976; in particular, $k_6(2)=275$.

The numbers $k_n(q)$ are known for $n\leq 13$ from work of Vera-López and Arregi [1992; 1995; 2003]; see also [Vera-López et al. 2008]. (There is an incorrect sign in the value given for $k_7(q)$ in [Vera-López and Arregi 1995, page 923]: the lowest term should be $-7q$.)

We may derive additional information on these numbers from the injective morphism of Hopf monoids (Proposition 6)

$$\mathbf{cf}(\mathrm{U}) \hookrightarrow \mathbf{f}(\mathrm{U}).$$

Define a sequence of integers $c_n(q)$, $n \geq 1$, by means of

$$\sum_{n\geq 0} k_n(q)x^n = \frac{1}{1-\sum_{n\geq 1} c_n(q)x^n}. \tag{58}$$

Arguing as above, we obtain the following result:

Corollary 20. *The following linear inequalities are satisfied for every $n \in \mathbb{N}$ and every prime power q*:

$$q^{\binom{n}{2}} \geq \sum_{i=1}^{n} q^{\binom{n-i}{2}} c_i(q). \tag{59}$$

Through (58), these inequalities impose further constraints on the numbers $k_n(q)$. The first few values of the sequence $c_n(q)$ are as follows with $t = q-1$:

$$\begin{aligned}
c_1(q) &= 1,\\
c_2(q) &= t,\\
c_3(q) &= t^2+t,\\
c_4(q) &= 2t^3+4t^2+t,\\
c_5(q) &= 5t^4+14t^3+9t^2+t,\\
c_6(q) &= t^6+18t^5+55t^4+54t^3+16t^2+t.
\end{aligned}$$

Conjecture 21. *There exist polynomials $p_n(t) \in \mathbb{N}[t]$ such that $c_n(q) = p_n(q-1)$ for every prime power q and every $n \geq 1$.*

Using the formulas given by Vera-López et al. [2008, Corollaries 10–11] for computing $k_n(q)$, we have verified the conjecture for $n \leq 13$.

Polynomiality of $k_n(q)$ is equivalent to that of $c_n(q)$. On the other hand, the nonnegativity of c_n as a polynomial of t implies that of k_n but not conversely. Thus, Conjecture 21 is a strong form of Higman's.

It is possible to show, using the methods of [Aguiar and Mahajan 2012], that the monoid $\mathbf{cf}(\mathrm{U})$ is free. This implies that the integers $c_n(q)$ are nonnegative for every $n \geq 1$ and prime power q.

6.2. ***From Hopf monoids to Hopf algebras.*** It is possible to associate a number of graded Hopf algebras to a given Hopf monoid $\boldsymbol{h}$. This is the subject of [Aguiar and Mahajan 2010, Part III]. In particular, there are two graded Hopf algebras $\mathcal{K}(\boldsymbol{h})$ and $\overline{\mathcal{K}}(\boldsymbol{h})$ related by a canonical surjective morphism

$$\mathcal{K}(\boldsymbol{h}) \twoheadrightarrow \overline{\mathcal{K}}(\boldsymbol{h}).$$

The underlying spaces of these Hopf algebras are

$$\mathcal{K}(\boldsymbol{h}) = \bigoplus_{n\geq 0} \boldsymbol{h}[n] \quad \text{and} \quad \overline{\mathcal{K}}(\boldsymbol{h}) = \bigoplus_{n\geq 0} \boldsymbol{h}[n]_{\mathrm{S}_n},$$

where $\boldsymbol{h}[n]_{\mathrm{S}_n}$ is as in (54). The product and coproduct of these Hopf algebras is built from those of the Hopf monoid $\boldsymbol{h}$ together with certain canonical transformations. The latter involve certain combinatorial procedures known as *shifting* and *standardization*. For more details, we refer to [Aguiar and Mahajan 2010, Chapter 15].

For example, one has that

$$\overline{\mathcal{K}}(\boldsymbol{L}) = \Bbbk[x]$$

is the polynomial algebra on one primitive generator while $\mathcal{K}(\boldsymbol{L})$ is the Hopf algebra introduced by Patras and Reutenauer [2004].

According to [Aguiar and Mahajan 2010, Section 17.4], $\overline{\mathcal{K}}(\boldsymbol{\Pi})$ is the ubiquitous Hopf algebra of *symmetric functions* while $\mathcal{K}(\boldsymbol{\Pi})$ is the Hopf algebra of *symmetric functions in noncommuting variables*, an object studied in various references including [Aguiar and Mahajan 2006, Section 6.2; Bergeron et al. 2006; Bergeron and Zabrocki 2009; Rosas and Sagan 2006].

For any Hopf monoid $\boldsymbol{h}$, one has [Aguiar and Mahajan 2010, Theorem 15.13]

$$\overline{\mathcal{K}}(\boldsymbol{L} \times \boldsymbol{h}) \cong \mathcal{K}(\boldsymbol{h}).$$

Combining with Corollary 12, we obtain that, when the field of coefficients is $\mathbb{F}_2$,

$$\overline{\mathcal{K}}(\mathbf{scf}(\mathrm{U})) \cong \overline{\mathcal{K}}(\boldsymbol{L} \times \boldsymbol{\Pi}) \cong \mathcal{K}(\boldsymbol{\Pi}).$$

In other words, the Hopf algebra constructed from superclass functions on unitriangular matrices (with entries in $\mathbb{F}_2$) via the functor $\overline{\mathcal{K}}$ is isomorphic to the Hopf algebra of symmetric functions in noncommuting variables. This is the main result of [Aguiar et al. 2012].

The freeness of the Hopf algebra $\mathcal{K}(\boldsymbol{\Pi})$, a fact known from [Harčenko 1978; Wolf 1936], is a consequence of Proposition 17.

We mention that one may arrive at Corollary 19 by employing the Hopf algebra $\overline{\mathcal{K}}(\mathbf{cf}(\mathrm{U}))$ (rather than the Hopf monoid $\mathbf{cf}(\mathrm{U})$) and appealing to Lagrange's theorem for graded connected Hopf algebras.

6.3. ***Supercharacters and beyond.*** The notion of superclass on a unitriangular group comes with a companion notion of *supercharacter* and a full-fledged theory relating them. This is due to the pioneering work of André [1995a; 1995b] and later Yan [2001]. Much of this theory extends to algebra groups [André 1999; Diaconis and Isaacs 2008; Diaconis and Thiem 2009]. More recently, a connection with classical work on Schur rings has been understood [Hendrickson 2010].

In regards to the object of present interest, the Hopf monoid $\mathbf{scf}(\mathrm{U})$, this implies the existence of a second canonical linear basis consisting of supercharacters. The work of André and Yan provides a *character formula*, which yields the change of basis between superclass functions and supercharacters. We plan to study the Hopf monoid structure of $\mathbf{scf}(\mathrm{U})$ on the supercharacter basis in future work.

Appendix: On free Hopf algebras and Hopf monoids

A free algebra may carry several Hopf algebra structures. It always carries a canonical one in which the generators are primitive. It turns out that under certain conditions, any Hopf structure on a free algebra is isomorphic to the canonical one. We provide such a result below. An analogous result holds for Hopf monoids in vector species. This is applied in the paper in Section 5.

We assume that the base field $\Bbbk$ is of characteristic 0.

We employ the *first Eulerian idempotent* [Gerstenhaber and Schack 1991; Loday 1998, Section 4.5.2; Reutenauer 1993, Section 8.4]. For any connected Hopf algebra H, the identity map $\mathrm{id} : H \to H$ is locally unipotent with respect to the convolution product of $\mathrm{End}(H)$. Therefore,

$$\boldsymbol{e} := \log(\mathrm{id}) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k} (\mathrm{id} - \iota\epsilon)^{*k} \tag{60}$$

is a well-defined linear endomorphism of H. Here

$$\iota : \Bbbk \to H \quad \text{and} \quad \epsilon : H \to \Bbbk$$

denote the unit and counit maps of H, respectively, and the powers are with respect to the convolution product. It is an important fact that if H is in addition cocommutative, then $\boldsymbol{e}(x)$ is a primitive element of H for any $x \in H$. In fact, the operator $\boldsymbol{e}$ is in this case a projection onto the space of primitive elements [Patras 1994; Schmitt 1994, pages 314–318].

Let $T(V)$ denote the free algebra on a vector space V:

$$T(V) = \bigoplus_{n\geq 0} V^{\otimes n}.$$

The product is concatenation of tensors. We say in this case that V freely generates.

The unique morphisms of algebras

$$\Delta : T(V) \to T(V) \otimes T(V) \quad \text{and} \quad \epsilon : T(V) \to \Bbbk$$

given for all $v \in V$ by

$$\Delta(v) = 1 \otimes v + v \otimes 1 \quad \text{and} \quad \epsilon(v) = 0$$

turn $T(V)$ into a connected, cocommutative Hopf algebra. This is the *canonical* Hopf structure on $T(V)$.

Proposition 22. *Let $\Bbbk$ be a field of characteristic* 0*. Let H be a connected cocommutative Hopf algebra over $\Bbbk$. Suppose $H \cong T(W)$ as algebras in such a way that the image of W lies in the kernel of the ϵ. Then there exists a* (*possibly different*) *isomorphism of Hopf algebras $H \cong T(W)$, where the latter is endowed with its canonical Hopf structure.*

Proof. We may assume $H = T(W)$ as algebras for some subspace W of $\ker(\epsilon)$. Since H is connected and $\Bbbk$ is of characteristic 0, the Eulerian idempotent $\boldsymbol{e}$ is defined. Let $V = \boldsymbol{e}(W)$. We show below that $V \cong W$ and that V freely generates H. Since H is cocommutative, V consists of primitive elements, and therefore, $H \cong T(V)$ as Hopf algebras. This completes the proof.

Let

$$H_+ = \bigoplus_{n\geq 1} W^{\otimes n}.$$

Since ϵ is a morphism of algebras, $H_+ \subseteq \ker(\epsilon)$, and since both spaces are of codimension 1, they must agree: $H_+ = \ker(\epsilon)$.

Define $\Delta_+ : H_+ \to H_+ \otimes H_+$ by

$$\Delta_+(x) = \Delta(x) - 1 \otimes x - x \otimes 1.$$

By counitality,

$$(\epsilon \otimes \mathrm{id})\Delta_+ = 0 = (\mathrm{id} \otimes \epsilon)\Delta_+.$$

Therefore, $\Delta_+(H_+) \subseteq \ker(\epsilon) \otimes \ker(\epsilon) = H_+ \otimes H_+$, and hence,

$$\Delta_+^{(k-1)}(H_+) \subseteq H_+^{\otimes k}$$

for all $k \geq 1$. In addition, since $H = T(W)$ as algebras,

$$\mu^{(k-1)}(H_+^{\otimes k}) \subseteq \sum_{n\geq 2} W^{\otimes n}$$

for all $k \geq 2$.

Take $w \in W$. Then

$$\begin{aligned}\boldsymbol{e}(w) &= \sum_{k\geq 1} \frac{(-1)^{k+1}}{k} (\mathrm{id} - \iota\epsilon)^{*k}(w) = w + \sum_{k\geq 2} \frac{(-1)^{k+1}}{k} \mu^{(k-1)} \Delta_+^{(k-1)}(w) \\ &\equiv w + \sum_{n\geq 2} W^{\otimes n}.\end{aligned}$$

By triangularity, $\boldsymbol{e} : W \to V$ is invertible, and hence, V generates H.

Now take $w_1, w_2 \in W$. It follows from the above that

$$\boldsymbol{e}(w_1)\boldsymbol{e}(w_2) \equiv w_1 w_2 + \sum_{n \geq 3} W^{\otimes n},$$

and a similar triangular relation holds for higher products. Hence, V generates H freely. □

The Eulerian idempotent is defined for connected Hopf monoids in species by the same formula as (60). Let $\boldsymbol{p}$ be a species such that $\boldsymbol{p}[\varnothing] = 0$. The free monoid $\mathcal{T}(\boldsymbol{p})$ and its canonical Hopf structure is discussed in [Aguiar and Mahajan 2010, Section 11.2]. The arguments in Proposition 22 may easily be adapted to this setting to yield the following result:

Proposition 23. *Let $\Bbbk$ be a field of characteristic 0. Let $\boldsymbol{h}$ be a connected cocommutative Hopf monoid in vector species over $\Bbbk$. Suppose $\boldsymbol{h} \cong \mathcal{T}(\boldsymbol{p})$ as monoids for some species $\boldsymbol{p}$ such that $\boldsymbol{p}[\varnothing] = 0$. Then there exists a (possibly different) isomorphism of Hopf monoids $\boldsymbol{h} \cong \mathcal{T}(\boldsymbol{p})$, where the latter is endowed with its canonical Hopf structure.*

Acknowledgements

We thank the referee and Franco Saliola for useful comments and suggestions.

References

[Aguiar and Lauve 2012] M. Aguiar and A. Lauve, "Lagrange's theorem for Hopf monoids in species", preprint, 2012. arXiv 1105.5572v3

[Aguiar and Mahajan 2006] M. Aguiar and S. Mahajan, *Coxeter groups and Hopf algebras*, Fields Institute Monographs **23**, American Mathematical Society, Providence, RI, 2006. MR 2008d:20072 Zbl 1106.16039

[Aguiar and Mahajan 2010] M. Aguiar and S. Mahajan, *Monoidal functors, species and Hopf algebras*, CRM Monograph Series **29**, American Mathematical Society, Providence, RI, 2010. MR 2012g:18009 Zbl 1209.18002

[Aguiar and Mahajan 2012] M. Aguiar and S. Mahajan, "On the Hadamard product of Hopf monoids", preprint, 2012. arXiv 1209.1363

[Aguiar et al. 2012] M. Aguiar, C. André, C. Benedetti, N. Bergeron, Z. Chen, P. Diaconis, A. Hendrickson, S. Hsiao, I. M. Isaacs, A. Jedwab, K. Johnson, G. Karaali, A. Lauve, T. Le, S. Lewis, H. Li, K. Magaard, E. Marberg, J.-C. Novelli, A. Pang, F. Saliola, L. Tevlin, J.-Y. Thibon, N. Thiem, V. Venkateswaran, C. R. Vinroot, N. Yan, and M. Zabrocki, "Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras", *Adv. Math.* **229**:4 (2012), 2310–2337. MR 2880223 Zbl 1237.05208

[André 1995a] C. A. M. André, "Basic characters of the unitriangular group", *J. Algebra* **175**:1 (1995), 287–319. MR 96h:20081a Zbl 0835.20052

[André 1995b] C. A. M. André, "Basic sums of coadjoint orbits of the unitriangular group", *J. Algebra* **176**:3 (1995), 959–1000. MR 96h:20081b Zbl 0837.20050

[André 1999] C. A. M. André, "Irreducible characters of finite algebra groups", pp. 65–80 in *Matrices and group representations* (Coimbra, 1998), Textos Mat. Sér. B **19**, Universidade de Coimbra, Coimbra, 1999. MR 2001g:20009 Zbl 0972.20005

[Bergeron_F et al. 1998] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial species and tree-like structures*, Encyclopedia of Mathematics and its Applications **67**, Cambridge University Press, 1998. MR 2000a:05008 Zbl 0888.05001

[Bergeron and Zabrocki 2009] N. Bergeron and M. Zabrocki, "The Hopf algebras of symmetric functions and quasi-symmetric functions in non-commutative variables are free and co-free", *J. Algebra Appl.* **8**:4 (2009), 581–600. MR 2011a:05372 Zbl 1188.16030

[Bergeron et al. 2006] N. Bergeron, C. Hohlweg, M. Rosas, and M. Zabrocki, "Grothendieck bialgebras, partition lattices, and symmetric functions in noncommutative variables", *Electron. J. Combin.* **13**:1 (2006), R75. MR 2007e:05176 Zbl 1098.05079

[Diaconis and Isaacs 2008] P. Diaconis and I. M. Isaacs, "Supercharacters and superclasses for algebra groups", *Trans. Amer. Math. Soc.* **360**:5 (2008), 2359–2392. MR 2009c:20012 Zbl 1137.20008

[Diaconis and Thiem 2009] P. Diaconis and N. Thiem, "Supercharacter formulas for pattern groups", *Trans. Amer. Math. Soc.* **361**:7 (2009), 3501–3533. MR 2010g:20013 Zbl 1205.20006

[Gerstenhaber and Schack 1991] M. Gerstenhaber and S. D. Schack, "The shuffle bialgebra and the cohomology of commutative algebras", *J. Pure Appl. Algebra* **70**:3 (1991), 263–272. MR 92e:13008 Zbl 0728.13003

[Goodwin 2006] S. M. Goodwin, "On the conjugacy classes in maximal unipotent subgroups of simple algebraic groups", *Transform. Groups* **11**:1 (2006), 51–76. MR 2006k:20096 Zbl 1118.20041

[Goodwin and Röhrle 2009] S. M. Goodwin and G. Röhrle, "Calculating conjugacy classes in Sylow p-subgroups of finite Chevalley groups", *J. Algebra* **321**:11 (2009), 3321–3334. MR 2010g:20083 Zbl 1210.20044

[Harčenko 1978] V. K. Harčenko, "Algebras of invariants of free algebras", *Algebra i Logika* **17**:4 (1978), 478–487. In Russian; translated in *Algebra and Logic* **17**:4 (1978), 316–321. MR 80e:16003

[Hendrickson 2010] A. O. F. Hendrickson, "Supercharacter theories and Schur rings", preprint, 2010. arXiv 1006.1363v1

[Higman 1960] G. Higman, "Enumerating p-groups, I: Inequalities", *Proc. London Math. Soc.* (3) **10** (1960), 24–30. MR 22 #4779 Zbl 093.02603

[Isaacs 1995] I. M. Isaacs, "Characters of groups associated with finite algebras", *J. Algebra* **177**:3 (1995), 708–730. MR 96k:20011 Zbl 0839.20010

[Kirillov 1995] A. A. Kirillov, "Variations on the triangular theme", pp. 43–73 in *Lie groups and Lie algebras: E. B. Dynkin's seminar*, edited by S. G. Gindikin and E. B. Vinberg, Amer. Math. Soc. Transl. Ser. 2 **169**, American Mathematical Society, Providence, RI, 1995. MR 97a:20072 Zbl 0840.22015

[Loday 1998] J.-L. Loday, *Cyclic homology*, 2nd ed., Grundlehren Math. Wiss. **301**, Springer, Berlin, 1998. MR 98h:16014 Zbl 0885.18007

[OEIS Foundation 2010] OEIS Foundation, "The on-line encyclopedia of integer sequences", 2010, http://oeis.org.

[Patras 1994] F. Patras, "L'algèbre des descentes d'une bigèbre graduée", *J. Algebra* **170**:2 (1994), 547–566. MR 96a:16043 Zbl 0819.16033

[Patras and Reutenauer 2004] F. Patras and C. Reutenauer, "On descent algebras and twisted bialgebras", *Mosc. Math. J.* **4**:1 (2004), 199–216. MR 2005e:16067 Zbl 1103.16026

[Reutenauer 1993] C. Reutenauer, *Free Lie algebras*, London Math. Soc. Monogr. (N.S.) **7**, Oxford University Press, New York, 1993. MR 94j:17002 Zbl 0798.17001

[Robinson 1998] G. R. Robinson, "Counting conjugacy classes of unitriangular groups associated to finite-dimensional algebras", *J. Group Theory* **1**:3 (1998), 271–274. MR 99h:14025 Zbl 0926.20031

[Rosas and Sagan 2006] M. H. Rosas and B. E. Sagan, "Symmetric functions in noncommuting variables", *Trans. Amer. Math. Soc.* **358**:1 (2006), 215–232. MR 2006f:05184 Zbl 1071.05073

[Schmitt 1994] W. R. Schmitt, "Incidence Hopf algebras", *J. Pure Appl. Algebra* **96**:3 (1994), 299–330. MR 95m:16033 Zbl 0808.05101

[Vera-López and Arregi 1992] A. Vera-López and J. M. Arregi, "Conjugacy classes in Sylow p-subgroups of $\mathrm{GL}(n, q)$", *J. Algebra* **152**:1 (1992), 1–19. MR 94b:20048 Zbl 0777.20015

[Vera-López and Arregi 1995] A. Vera-López and J. M. Arregi, "Some algorithms for the calculation of conjugacy classes in the Sylow p-subgroups of $\mathrm{GL}(n, q)$", *J. Algebra* **177**:3 (1995), 899–925. MR 96j:20029 Zbl 0839.20061

[Vera-López and Arregi 2003] A. Vera-López and J. M. Arregi, "Conjugacy classes in unitriangular matrices", *Linear Algebra Appl.* **370** (2003), 85–124. MR 2004i:20091 Zbl 1045.20045

[Vera-López et al. 2008] A. Vera-López, J. M. Arregi, L. Ormaetxea, and F. J. Vera-López, "The exact number of conjugacy classes of the Sylow p-subgroups of $\mathrm{GL}(n, q)$ modulo $(q-1)^{13}$", *Linear Algebra Appl.* **429**:2-3 (2008), 617–624. MR 2009d:20117 Zbl 1141.20008

[Wolf 1936] M. C. Wolf, "Symmetric functions of non-commutative elements", *Duke Math. J.* **2**:4 (1936), 626–637. MR 1545953 Zbl 0016.00501

[Yan 2001] N. Yan, *Representation theory of the finite unipotent linear groups*, Ph.D. thesis, University of Pennsylvania, 2001, http://repository.upenn.edu/dissertations/AAI3015396/. MR 2702153

Communicated by Ravi Vakil
Received 2012-04-10 Revised 2012-08-08 Accepted 2012-10-07

maguiar@math.tamu.edu *Department of Mathematics, Texas A&M University, College Station, TX, 77843, United States*
http://www.math.tamu.edu/~maguiar

bergeron@mathstat.yorku.ca *Department of Mathematics and Statistics, York University, Toronto, ON, M3J 1P3, Canada*
http://www.math.yorku.ca/bergeron

thiem@colorado.edu *Department of Mathematics, University of Colorado, Boulder, CO, 80309, United States*
http://math.colorado.edu/~thiemn

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.

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 7 No. 7 2013

Weil representation and transfer factor 1535
TERUJI THOMAS

Analytic families of finite-slope Selmer groups 1571
JONATHAN POTTHARST

Multiplicative excellent families of elliptic surfaces of type E_7 or E_8 1613
ABHINAV KUMAR and TETSUJI SHIODA

Cohomological invariants of algebraic tori 1643
SAM BLINSTEIN and ALEXANDER MERKURJEV

On abstract representations of the groups of rational points of algebraic groups and their deformations 1685
IGOR A. RAPINCHUK

Betti diagrams from graphs 1725
ALEXANDER ENGSTRÖM and MATTHEW T. STAMPS

Hopf monoids from class functions on unitriangular matrices 1743
MARCELO AGUIAR, NANTEL BERGERON and NATHANIEL THIEM