# Algebra & Number Theory

Volume 5
2011
No. 1

mathematical sciences publishers

# Algebra & Number Theory

www.jant.org



See inside back cover or www.jant.org for submission instructions.

The subscription price for 2011 is US $150/year for the electronic version, and $210/year (+$35 shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Algebra & Number Theory (ISSN 1937-0652) at Mathematical Sciences Publishers, Department of Mathematics, 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

http://msp.org/

A NON-PROFIT CORPORATION

Typeset in LaTeX

ALGEBRA AND NUMBER THEORY 5:1(2011)

msp

# Formules pour l'invariant de Rost

Philippe Gille et Anne Quéguiner-Mathieu

On donne une formule exacte pour l'invariant de Rost $H^1(k, G) \to H^3(k)$ des groupes linéaires spéciaux.

We provide an exact formula for the Rost invariant $H^1(k, G) \to H^3(k)$ of special linear groups.



## 1. Introduction et notations

Soient $k$ un corps, $k_s$ une clôture séparable de $k$ et $\Gamma_k = \mathrm{Gal}(k_s/k)$ le groupe de Galois absolu de $k$. Si $G/k$ est un groupe semi-simple simplement connexe absolument presque $k$-simple, on dispose de l'invariant de Rost [Esnault et al. 1998; Garibaldi et al. 2003]

$$r_G : H^1(k, G) \to H^3(k) := H^3(k, (\mathbb{Q}/\mathbb{Z})(2))$$

qui associe à la classe d'un $G$-torseur une classe de cohomologie galoisienne de degré 3, où pour $d \geq 0$, $H^{d+1}(k, (\mathbb{Q}/\mathbb{Z})(d))$ désigne le groupe de cohomologie galoisienne modifié à la Kato [1982] sur la composante $p$-primaire si $k$ est de caractéristique $p$ positive [Garibaldi et al. 2003, p. 151].

*MSC2000:* 11E72.

*Mots-clefs:* cohomologie galoisienne, Galois cohomology, linear algebraic groups, groupes algébriques linéaires.

Le premier but de cet article est d'établir une formule exacte pour l'invariant de Rost dans le cas du groupe $G = \mathrm{SL}_1(A)$, où $A$ est une $k$-algèbre centrale simple. On sait alors que la suite exacte

$$1 \to \mathrm{SL}_1(A) \to \mathrm{GL}_1(A) \xrightarrow{\mathrm{Nrd}_A} \mathbb{G}_m \to 1$$

induit un isomorphisme $H^1(k, G) \cong k^\times/\mathrm{Nrd}(A^\times)$, et que le cup-produit avec la classe de Brauer $[A] \in \mathrm{Br}(k) = H^2(k, (\mathbb{Q}/\mathbb{Z})(1))$ engendre le groupe des invariants de degré 3 de $\mathrm{SL}_1(A)$ [Garibaldi et al. 2003, p. 107]. Pour comparer ce générateur et l'invariant de Rost, on convient d'identifier le groupe de Brauer de $k$ et $H^2(k)$ par le cobord, comme dans [Gille et Szamuely 2006, §4.4]. Notons que cette convention est opposée à celle de [Garibaldi et al. 2003, p. 151], donnée par le produit croisé (voir [Knus et al. 1998, p. 397]).

**Théorème 1.1.** *Soit $A$ une $k$-algèbre simple centrale de degré $n$. On note $[A] \in {}_n\mathrm{Br}(k) = {}_nH^2(k)$ sa classe dans le groupe de Brauer de $k$. Soit*

$$[v] \in H^1(k, \mathrm{SL}_1(A)) = k^\times/\mathrm{Nrd}(A^\times).$$

(1) *Si l'indice $\mathrm{ind}_k(A)$ est inversible dans $k$, on a*

$$r_{\mathrm{SL}_1(A)}\big([v]\big) = (v) \cup [A] \in H^3(k),$$

*où $(v)$ désigne la classe dans ${}_nH^1(k) = k^\times/k^{\times n}$ d'un représentant quelconque de $[v]$.*

(2) *Si $k$ est de caractéristique $p > 0$ et $A$ est d'indice $p^h$, alors*

$$r_{\mathrm{SL}_1(A)}\big([v]\big) = -(v) \cup [A] \in H^3(k).$$

Pour la définition du cup-produit dans le second cas, voir le § 4D.

La méthode employée consiste à utiliser des cocycles explicites pour des algèbres cycliques, l'ingrédient fondamental étant les extensions centrales de [Brylinski et Deligne 2001] qui permettent une approche galoisienne des extensions centrales de [Matsumoto 1969].

On propose ensuite deux applications du résultat principal. Le § 5 précise les résultats obtenus dans [Merkurjev et al. 2002] et [Garibaldi et Quéguiner-Mathieu 2007] concernant la restriction de l'invariant de Rost aux torseurs issus du centre du groupe. La seconde application concerne les groupes de type $G_2$, $F_4$ et $E_8$ dont le centre est trivial. Ces groupes possèdent un sous-groupe de la forme $A = \mu_l \times \mu_l \times \mathbb{Z}/l\mathbb{Z}$ de centralisateur fini, et l'on donne au § 6 une description de l'invariant de Rost pour les torseurs issus de $A$. On en déduit (voir corollaire 6.7) que la partie modulo 3 de l'invariant de Rost coïncide avec l'invariant des algèbres d'Albert construit dans [Rost 1991], et décrit également dans [Petersson et Racine 1996].

## 2. Extensions de groupes

Dans toute cette partie, $A$ désigne un groupe abélien, et

$$0 \to A \to E \xrightarrow{p} G \to 1$$

est une extension d'un certain groupe $G$ par $A$. On rappelle que l'action de $E$ sur $A$ par automorphisme intérieur se factorise en une action $\theta : G \to \mathrm{Aut}(A)$, ce qui munit $A$ d'une structure de $G$-module. De plus, si l'on fixe une action de $G$ sur $A$, les extensions de $G$ par $A$ qui induisent cette action sont classifiées par le groupe $H^2(G, A)$ (voir par exemple [Weibel 1994, §6.6]).

**2A.** ***Extensions de* Γ*-groupes.*** Soit $\Gamma$ un groupe. On suppose ici que $A$, $E$ et $G$ sont des $\Gamma$-groupes et que l'extension ci-dessus est centrale et compatible à l'action de $\Gamma$. Elle donne lieu à une suite exacte longue d'ensembles pointés [Serre 1994, §I.5.4]

$$1 \to H^0(\Gamma, A) \to H^0(\Gamma, E) \to H^0(\Gamma, G) \to H^1(\Gamma, A) \to H^1(\Gamma, E) \to H^1(\Gamma, G) \xrightarrow{\Delta} H^2(\Gamma, A).$$

Nous nous proposons de donner une description du bord $\Delta$ en termes d'extensions de groupes. On rappelle que l'ensemble des 1-cocycles $Z^1(\Gamma, G)$ n'est pas autre chose que l'ensemble des sections de $G \rtimes \Gamma \to \Gamma$. En effet, si $z \in Z^1(\Gamma, G)$, on lui associe la section $u_z : \Gamma \to G \rtimes \Gamma$, $\sigma \mapsto z_\sigma\, \sigma$ (*ibid*, §5.1, exercice 1). Étant donné un 1-cocycle $z : \Gamma \to G$, on peut retirer en arrière par $u_z$ l'extension

$$\begin{array}{ccccccccc} 1 & \longrightarrow & A & \longrightarrow & E \rtimes \Gamma & \longrightarrow & G \rtimes \Gamma & \longrightarrow & 1 \\ & & \| & & \uparrow & & \uparrow{\scriptstyle u_z} & & \\ 1 & \longrightarrow & A & \longrightarrow & E(z) & \longrightarrow & \Gamma & \longrightarrow & 1. \end{array}$$

**Lemme 2.1.** $\Delta([z]) = \big[E(z)\big] \in H^2(\Gamma, A)$.

*Démonstration.* On vérifie par calcul. Pour chaque $\sigma \in \Gamma$, on choisit un relevé par $e_\sigma \in E$ de $z_\sigma$ de sorte que $z_e$ est l'élément neutre de $E$. Alors $\Delta([z])$ est par définition la classe du 2-cocycle

$$a_{\sigma,\tau} = e_\sigma\, {}^\sigma e_\tau\, e_{\sigma\tau}^{-1} \in A.$$

De l'autre coté, $f_\sigma := e_\sigma\, \sigma \in E(z)$ définit un relevé de $\sigma$ pour le morphisme $E(z) \to \Gamma$. La classe de $E(z)$ dans $H^2(\Gamma, A)$ est celle du 2-cocycle (voir [Weibel 1994, th. 6.6.3])

$$f_\sigma\, f_\tau\, f_{\sigma\tau}^{-1} = e_\sigma\, \sigma\, e_\tau\, \tau\, \tau^{-1}\sigma^{-1}\, e_{\sigma\tau}^{-1} = e_\sigma\, (\sigma e_\tau \sigma^{-1})\, e_{\sigma\tau}^{-1} = e_\sigma\, {}^\sigma e_\tau\, e_{\sigma\tau}^{-1} = a_{\sigma,\tau}. \quad \square$$

On va maintenant établir une version tordue. Soit $z \in Z^1(\Gamma, G)$. On peut tordre l'extension $1 \to A \to E \to G \to 1$ par $z$

$$1 \to A \to {}_zE \to {}_zG \to 1$$

et étudier le bord $\Delta_z : H^1(\Gamma, {}_zG) \to H^2(\Gamma, A)$ [Serre 1994, §I.5.7]. Un 1-cocycle $w \in Z^1(\Gamma, {}_zG)$ est simplement une section $u_w$ de la projection $G \rtimes^z \Gamma = {}_zG \rtimes \Gamma \to \Gamma$ via $\sigma \mapsto w_\sigma \sigma$. En appliquant la formule ci-dessus, on obtient que le bord $\Delta_z([w])$ est la classe de l'extension retirée en arrière

$$\begin{array}{ccccccccc} 1 & \longrightarrow & A & \longrightarrow & E \rtimes^z \Gamma & \to & G \rtimes^z \Gamma & \longrightarrow & 1 \\ & & \| & & \uparrow & & \uparrow{\scriptstyle u_w} & & \\ 1 & \longrightarrow & A & \longrightarrow & E(w) & \longrightarrow & \Gamma & \longrightarrow & 1. \end{array}$$

**2B.** ***Lemme préliminaire.*** On s'intéresse aux sous-groupes suivants du groupe $\mathrm{Aut}(E)$ des automorphismes de $E$ :

$$\mathrm{Aut}(E, A) = \big\{\phi \in \mathrm{Aut}(E) \mid \phi(A) = A\big\},$$
$$\mathrm{Aut}_A(E) = \big\{\phi \in \mathrm{Aut}(E) \mid \phi_{|A} = \mathrm{Id}_{|A}\big\}.$$

**Lemme 2.2.** *Le noyau de l'application naturelle* $\mathrm{Aut}_A(E) \to \mathrm{Aut}(G)$ *est en bijection avec le groupe* $Z^1(G, A)$ *des* 1*-cocycles de* $G$ *à valeurs dans* $A$.

*Démonstration.* Soit $\alpha = (a_g) \in Z^1(G, A)$ un 1-cocycle. On lui associe l'application $\phi_\alpha : E \to E$, définie par $\phi_\alpha(x) = a_{p(x)}x$, pour tout $x \in E$. Quels que soient $x_1, x_2 \in E$, on a

$$\phi_\alpha(x_1x_2) = a_{p(x_1x_2)}\, x_1\, x_2 = a_{p(x_1)}\, x_1\, a_{p(x_2)}\, x_1^{-1}x_1\, x_2 = \phi_\alpha(x_1)\, \phi_\alpha(x_2).$$

Ainsi, $\phi_\alpha$ est un morphisme, dont la restriction à $A$ est l'identité. Son noyau, qui est contenu dans $A$, est trivial. Enfin, si $x \in E$, on a $x = \phi_\alpha(a_{p(x)}^{-1}x)$, de sorte que $\phi_\alpha \in \mathrm{Aut}_A(E)$. Comme $p(\phi_\alpha(x)) = p(x)$, l'automorphisme induit sur $G$ est trivial.

Ainsi, l'application $\alpha \mapsto \phi_\alpha$ est un morphisme de $Z^1(G, A)$ dans le noyau $\ker(\mathrm{Aut}_A(E) \to \mathrm{Aut}(G))$. Pour montrer qu'elle est bijective, on va exhiber sa réciproque. Soit donc $\phi : E \xrightarrow{\sim} E$ un automorphisme induisant l'identité sur $A$ et sur $G$. L'application $x \in E \to \phi(x)x^{-1}$ induit une application $\psi : G = E/A \to A$. De plus, si $g_1$ et $g_2 \in G$ ont pour antécédents respectifs $x_1$ et $x_2 \in E$, alors

$$\psi(g_1g_2) = \phi(x_1x_2)(x_1x_2)^{-1} = \psi(g_1)x_1\, \psi(g_2)x_1^{-1} = \psi(g_1)\, \big(g_1.\psi(g_2)\big).$$

On a donc bien défini ainsi une application de $\ker(\mathrm{Aut}_A(E) \to \mathrm{Aut}(G))$ dans $Z^1(G, A)$, qui par un calcul direct est la réciproque de la précédente. □

**2C.** ***Extensions centrales d'un groupe parfait.*** On suppose dans ce paragraphe que l'extension de groupes

$$0 \to A \to E \xrightarrow{p} G \to 1,$$

est centrale. Autrement dit, $A$ est contenu dans le centre $Z(E)$ de $E$, de sorte que l'action de $G$ sur $A$ est triviale. Si de plus le groupe $G$ est parfait, alors $Z^1(G, A) = \mathrm{Hom}(G, A) = 0$. Par le lemme précédent, le morphisme

$$\mathrm{Aut}_A(E) \hookrightarrow \mathrm{Aut}(G)$$

est donc injectif. On en déduit que deux automorphismes de $\mathrm{Aut}(E, A)$ qui induisent les mêmes morphismes sur $A$ et sur $G$ sont égaux, d'où un morphisme injectif

$$\mathrm{Aut}(E, A) \hookrightarrow \mathrm{Aut}(A) \times \mathrm{Aut}(G).$$

**Lemme 2.3.** *On considère une extension centrale*

$$0 \to A \to E \xrightarrow{p} G \to 1,$$

*d'un groupe parfait $G$, et une action $f = (f_1, f_2) : \Gamma \to \mathrm{Aut}(A) \times \mathrm{Aut}(G)$ d'un groupe $\Gamma$ sur $A$ et sur $G$. Les deux assertions suivantes sont équivalentes* :

(1) *L'image de $f$ est incluse dans le sous-groupe*

$$\mathrm{Aut}(E, A) \hookrightarrow \mathrm{Aut}(A) \times \mathrm{Aut}(G),$$

*c'est-à-dire l'action $f$ s'étend à $E$, et ce de manière unique* ;

(2) *Pour tout $\gamma \in \Gamma$, l'extension $(f_2(\gamma)^{-1})^* f_1(\gamma)_*(E)$ de $G$ par $A$ est équivalente à $E$, où $(f_2(\gamma)^{-1})^*$ et $f_1(\gamma)_*$ désignent respectivement le pull-back et le push-out associés aux morphismes $f_2(\gamma)^{-1}$ et $f_1(\gamma)$.*

*Démonstration.* Pour tout $\gamma \in \Gamma$, on considère les extensions $E'_\gamma = f_1(\gamma)_*(E)$ et $E''_\gamma = (f_2(\gamma)^{-1})^* f_1(\gamma)_*(E)$. Par définition, on a un diagramme commutatif

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A & \longrightarrow & E & \longrightarrow & G & \longrightarrow & 1 \\
 & & \downarrow{\scriptstyle f_1(\gamma)} & & \downarrow & & \| & & \\
0 & \longrightarrow & A & \longrightarrow & E'_\gamma & \longrightarrow & G & \longrightarrow & 1 \\
 & & \| & & \uparrow & & \uparrow{\scriptstyle f_2(\gamma)^{-1}} & & \\
0 & \longrightarrow & A & \longrightarrow & E''_\gamma & \longrightarrow & G & \longrightarrow & 1.
\end{array} \tag{2-1}$$

Supposons que l'action $f$ s'étend à $E$, c'est-à-dire que $f(\gamma) \in \mathrm{Aut}(E, A)$. On a alors

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & E & \longrightarrow & G & \longrightarrow & 1 \\ & & {\scriptstyle f_1(\gamma)^{-1}}\downarrow\wr & & \downarrow\wr & & {\scriptstyle f_2(\gamma)^{-1}}\downarrow\wr & & \\ 0 & \longrightarrow & A & \longrightarrow & E & \longrightarrow & G & \longrightarrow & 1 \end{array}$$

En combinant avec (2-1), on en déduit que les extensions $E$ et $E''_\gamma$ sont équivalentes. Réciproquement, si les extensions $E$ et $E''_\gamma$ sont équivalentes, on a un diagramme commutatif

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & E''_\gamma & \longrightarrow & G & \longrightarrow & 1 \\ & & \| & & \downarrow\wr & & \| & & \\ 0 & \longrightarrow & A & \longrightarrow & E & \longrightarrow & G & \longrightarrow & 1 \end{array}$$

qui, combiné avec (2-1), nous donne un automorphisme de $E$ dont l'image dans $\mathrm{Aut}(A) \times \mathrm{Aut}(G)$ est $(f_1(\gamma), f_2(\gamma))$. □

**2D.** ***Extension centrale d'un groupe abélien.*** Si le $G$-module $A$ est trivial, par le théorème des coefficients universels, on a une suite exacte scindée

$$0 \to \mathrm{Ext}^1_{\mathbb{Z}}(G_{ab}, A) \to H^2(G, A) \xrightarrow{\delta} \mathrm{Hom}(H_2(G, \mathbb{Z}), A) \to 0,$$

où $G_{ab} = G/[G, G] = H_1(G, \mathbb{Z})$ est l'abélianisé de $G$ [Weibel 1994, exercice 6.1.5(3) et théorème 6.1.11]. Supposons de plus que $G = B$ est abélien. Alors $H_2(B, \mathbb{Z}) = \Lambda^2 B$ [Brown 1982, théorème V.6.4(iii)] et la suite exacte devient

$$0 \to \mathrm{Ext}^1_{\mathbb{Z}}(B, A) \to H^2(B, A) \xrightarrow{\delta} \mathrm{Hom}(\Lambda^2 B, A) \to 0.$$

De plus, par [Brown 1982, §V.6, exercice 5], l'image sous $\delta$ de la classe d'un cocycle $f \in Z^2(B, A)$ est donnée par $b_1 \wedge b_2 \mapsto f(b_1, b_2) - f(b_2, b_1)$. En combinant avec [Gille et Szamuely 2006, exemple 3.2.6], on en déduit aisément

**Lemme 2.4.** *Soit*

$$0 \to A \to E \xrightarrow{p} B \to 1$$

*une extension centrale d'un groupe abélien $B$. L'image sous $\delta$ de sa classe dans $H^2(B, A)$ est le "relevé des commutateurs"*

$$\begin{array}{rcl} \Lambda^2 B & \to & A \\ b_1 \wedge b_2 & \mapsto & [e_1, e_2] = e_1 e_2 e_1^{-1} e_2^{-1}, \end{array}$$

*où $e_i \in E$ vérifie $p(e_i) = b_i$.*

**Remarque 2.5.** Soit $\phi : \Lambda^2 B \to A$ un morphisme ; on peut le voir comme une application bilinéaire alternée $B \times B \to A$, donc comme un 2-cocycle $\phi \in Z^2(B, A)$.

Il faut prendre garde que l'image sous $\delta$ de sa classe dans $H^2(B, A)$ est $2\phi$. En effet, puisque $\phi$ est alternée, la formule rappelée ci-dessus donne

$$\delta(\phi)(b_1 \wedge b_2) = \phi(b_1, b_2) - \phi(b_2, b_1) = 2\,\phi(b_1, b_2).$$

**2E.** ***Extensions d'un groupe cyclique.*** On suppose dans ce paragraphe que le groupe $G$ est cyclique d'ordre $n$. Le choix d'un générateur $\sigma \in G$ permet d'identifier $H^2(G, A)$ avec le groupe $A^G/N(A)$ où $N$ est l'application du $G$-module $A$ dans lui même définie par $N(a) = \Sigma_{i=0}^{n-1}\sigma^i \cdot a$ ; voir par exemple [Gille et Szamuely 2006, exemple 3.2.9].

**Lemme 2.6.** *Soit $G$ un groupe cyclique d'ordre $n$, de générateur $\sigma$, et soit*

$$0 \to A \to E \xrightarrow{p} G \to 1,$$

*une extension de $G$ par $A$. Choisissons $e \in E$ tel que $p(e) = \sigma$. Alors, $e^n \in A^G$ et sa classe dans $A^G/N(A) \simeq H^2(G, A)$ est la classe de l'extension $E$.*

*Démonstration.* Clairement, $e^n$ est dans le noyau de $p$ ; de plus, $\sigma$ agissant sur $A$ par conjugaison par $e$, l'élément $e^n$ est invariant sous l'action de $G$. Montrons maintenant que sa classe dans $A^G/N(A)$ est celle de l'extension $E$.

Pour cela, on considère le caractère $\chi : G \to \mathbb{Z}/n\mathbb{Z}$ associé au choix de $\sigma$, c'est-à-dire défini par $\chi(\sigma) = 1$, et le bord $\partial\chi \in H^2(G, \mathbb{Z})$ provenant de la suite exacte de $G$-modules

$$0 \to \mathbb{Z} \xrightarrow{n} \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \to 0.$$

D'après [Gille et Szamuely 2006, 3.4.11(3)], l'isomorphisme entre $A^G/N(A)$ et $H^2(G, A)$ est le cup-produit par $\partial\chi$. En particulier, il est fonctoriel en $A$. Il suffit donc de montrer le théorème pour l'extension $0 \to \mathbb{Z} \xrightarrow{n} \mathbb{Z} \xrightarrow{p} G \to 1$, où $p$ est définie par $p(1) = \sigma$. On vérifie facilement (voir exemple 3.2.6 du même ouvrage) que sa classe dans $H^2(G, \mathbb{Z})$ est représentée par le 2-cocycle

$$(\sigma^i, \sigma^j) \mapsto \begin{cases} 0 & \text{si } i + j < n, \\ 1 & \text{si } i + j \geq n, \end{cases}$$

qui n'est rien d'autre que $\partial\chi$. Elle correspond donc, dans $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}^G/N(\mathbb{Z})$, à $\bar{1}$. Or, si l'on choisit $e = 1 \in \mathbb{Z}$ comme relevé de $\sigma$, on a $e^n = n$ qui a pour antécédent $1 \in A^G = \mathbb{Z}$. Le théorème est donc prouvé dans ce cas.

Le cas général s'en déduit de la manière suivante. Étant donnée l'extension

$$0 \to A \to E \xrightarrow{p} G \to 1,$$

le choix d'un élément $e \in E$ tel que $p(e) = \sigma$ induit un diagramme commutatif

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z} & \xrightarrow{n} & \mathbb{Z} & \longrightarrow & G & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \phi} & & \| & & \\ 0 & \longrightarrow & A & \longrightarrow & E & \xrightarrow{p} & G & \longrightarrow & 0, \end{array}$$

où $\phi$ est définie par $\phi(1) = e$, de sorte que $\psi$ vérifie $\psi(1) = e^n$. Ainsi, l'extension qui nous intéresse est le push-out suivant $\psi$ de la précédente et l'élément correspondant dans $A^G/N(A)$ est la classe de $\psi(1) = e^n$. □

**2F.** ***Extensions de groupes et 2-extensions au sens de Yoneda.*** Fixons une action $\theta : G \to \mathrm{Aut}(A)$, de sorte que $A$ est muni d'une structure de $G$-module. Le groupe $H^2(G, A)$ est isomorphe au groupe $\mathrm{Ext}^2_{\mathbb{Z}[G]}(\mathbb{Z}, A)$ des 2-extensions de $\mathbb{Z}[G]$-modules au sens de Yoneda [Mac Lane 1963, IV, corollaire 5.2]. On va ici décrire explicitement cet isomorphisme en terme d'extensions de groupes et de $G$-modules.

**Lemme 2.7.** *Considérons une suite exacte de $G$-modules* :

$$0 \longrightarrow A \longrightarrow B \xrightarrow{d} C \longrightarrow \mathbb{Z} \longrightarrow 0.$$

(1) *Pour tout $c \in C$, l'ensemble*

$$E_c = \big\{(g, b) \in G \times B \mid d(b) = g.c - c\big\},$$

*est muni de la loi de groupe suivante $(g_1, b_1)(g_2, b_2) = (g_1 g_2, b_1 + g_1.b_2)$. De plus, $E_c$ est une extension de $G$ par $A$ d'action $\theta$.*

(2) *Si $c$ et $c' \in C$ ont la même image dans $\mathbb{Z}$, alors les extensions $E_c$ et $E_{c'}$ sont canoniquement isomorphes.*

*Démonstration.* (1) Si $d(b_i) = g_i.c - c$ pour $i = 1, 2$, on a

$$d(b_1 + g_1.b_2) = g_1.c - c + g_1.(g_2.c - c) = (g_1 g_2).c - c.$$

Ainsi, la loi définie ci-dessus est bien une loi interne sur $E_c$ ; on vérifie facilement qu'elle munit $E_c$ d'une structure de groupe. En particulier, l'inverse de $(g, b) \in E_c$ est $(g, b)^{-1} = (g^{-1}, -g^{-1}.b)$.

Comme l'image dans $\mathbb{Z}$ de $g \cdot c - c$ est triviale, la projection $E_c \to G$ est surjective ; son noyau $\{(1_G, a),\ a \in A\}$ est isomorphe à $A$. Ceci prouve que $E_c$ est une extension de $G$ par $A$. De plus, l'action induite de $G$ sur $A$ est donnée par $(g, 0)(1_G, a)(g^{-1}, 0) = (g, g.a)(g^{-1}, 0) = (1_G, g.a)$, qui coïncide bien avec l'action initiale.

(2) Si $c$ et $c'$ ont la même image dans $\mathbb{Z}$, il existe $b_0 \in B$ tel que $c' = c + d(b_0)$. L'application

$$\begin{array}{ccc} E_c & \to & E_{c'} \\ (g, b) & \mapsto & (g, b + g \cdot b_0 - b_0), \end{array}$$

qui agit comme l'identité sur $\{(1_G, a),\ a \in A\}$, est un isomorphisme entre ces deux extensions. □

**Proposition 2.8.** *Soit* $0 \to A \to B \to C \to \mathbb{Z} \to 0$ *une suite exacte de* $G$*-modules. Sa classe dans* $\mathrm{Ext}^2_{\mathbb{Z}[G]}(\mathbb{Z}, A) \simeq H^2(G, A)$ *est la classe de l'extension de groupes* $0 \to A \to E_c \to G \to 1$*, où* $c$ *est un élément de* $C$ *d'image* 1 *dans* $\mathbb{Z}$*.*

*Démonstration.* Notons tout d'abord que, en vertu du lemme 2.7(2), la classe de l'extension $E_c$ ne dépend pas du choix de $c \in C$ d'image 1 dans $\mathbb{Z}$.

Pour montrer la proposition, on utilise la description de l'opération inverse donnée dans [Mac Lane 1963, §IV.6], qui à une extension de groupe d'action $\theta$

$$0 \to A \xrightarrow{i} E \xrightarrow{p} G \to 1$$

associe sa classe caractéristique dans $\mathrm{Ext}^2_{\mathbb{Z}[G]}(\mathbb{Z}, A)$. Notant $\mathbb{Z}[G]^{(E)}$ le $G$-module libre de base $E$, cette classe est représentée par

$$\chi(E): \qquad 0 \to A \xrightarrow{\alpha} \mathbb{Z}[G]^{(E)}/L \xrightarrow{\beta} \mathbb{Z}[G] \xrightarrow{\epsilon} \mathbb{Z} \to 0$$

où $L$ est engendré par $[1_G]$ et les $[e_1 e_2] - p(e_1).[e_2] - [e_1]$ pour $e_1, e_2$ parcourant $E$. Les morphismes $\alpha$ et $\beta$ sont définis respectivement par $\alpha(a) = [i(a)] \in \mathbb{Z}[G]^{(E)}/L$ et $\beta([e]) = [p(e)] - [1_G]$, pour tout $e \in E$. Enfin, $\epsilon$ est le morphisme d'augmentation.

L'élément $1_G \in \mathbb{Z}[G]$ a pour image 1 dans $\mathbb{Z}$. Pour montrer la proposition, il suffit donc de remarquer que l'application $e \in E \mapsto (p(e), [e])$ est un isomorphisme entre l'extension $E_{1_G}$ associée à $\chi(E)$ suivant le lemme 2.7 et l'extension initiale $E$. □

Ce formalisme est bien commode quand on a affaire à une situation équivariante, c'est-à-dire quand on a une action

$$f = (f_1, f_2) : \Gamma \to \mathrm{Aut}(A) \times \mathrm{Aut}(G)$$

d'un groupe $\Gamma$ sur $A$ et sur $G$ qui est compatible avec l'action de $G$ sur $A$, c'est-à-dire telle que $f_1(\gamma)(g.a) = f_2(\gamma)(g) \cdot f_1(\gamma)(a)$ pour tous $\gamma \in \Gamma$, $g \in G$ et $a \in A$. En d'autres mots, on a une structure de $G \rtimes \Gamma$-module sur $A$ définie par

$$(g\gamma).a = g.f_1(\gamma)(a).$$

Supposons que la suite exacte de $G$-modules

$$0 \to A \to B \to C \to \mathbb{Z} \to 0$$

est en réalité une suite de $G \rtimes \Gamma$-modules. Pour tout $c \in C$, on peut alors définir par le lemme 2.7 une extension $E_c$ de $G$ par $A$ et une extension $\tilde{E}_c$ de $G \rtimes \Gamma$ par $A$. De plus, le diagramme suivant est commutatif :

$$\begin{array}{ccccccccc}
 & & & & 1 & & 1 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & A & \longrightarrow & E_c & \longrightarrow & G & \longrightarrow & 1 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & A & \longrightarrow & \widetilde{E}_c & \longrightarrow & G \rtimes \Gamma & \longrightarrow & 1. \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & \Gamma & = & \Gamma & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 1 & & 1 & &
\end{array}$$

**Lemme 2.9.** *Si $c \in C^\Gamma$, alors le groupe $\Gamma$ agit sur $E_c$ ; de plus, $\widetilde{E}_c$ est isomorphe à $E_c \rtimes \Gamma$.*

*Démonstration.* Si $c$ est invariant sous l'action de $\Gamma$, alors pour tout $\gamma \in \Gamma$ et $(g, b) \in E_c$, le couple $(\gamma \cdot g, \gamma \cdot b)$ appartient à $E_c$. Ceci définit une action de $\Gamma$ sur $E_c$. De plus, on peut vérifier que l'application $\tilde{E}_c \to E_c \times \Gamma$, qui à $(g\gamma, b)$ associe $\big((g, b), \gamma\big)$ induit un isomorphisme de groupes entre $\tilde{E}_c$ et $E_c \rtimes \Gamma$. □

## 3. Extensions de Brylinski–Deligne

**3A.** ***Rappels.*** Soit $G/k$ un groupe semi-simple simplement connexe absolument presque $k$-simple. On note $H^i(G, \mathcal{K}_j)$ les groupes de $K^M$-cohomologie de la variété $G$ définis par les complexes de Gersten [Gille et Szamuely 2006, §8.1]. On a alors $H^0(G, \mathcal{K}_2) = K_2(k)$ et $H^1(G, \mathcal{K}_2) = \mathbb{Z}$ (voir [Brylinski et Deligne 2001, proposition 4.6] si $G$ est déployé et [Garibaldi et al. 2003, (6.7), (6.10) et (6.12), p. 116–118] pour le cas général). Ainsi, en revenant à définition de la cohomologie en termes de complexes de Gersten, le complexe de $G(k)$-modules

$$K_2(k(G)) \to \bigoplus_{x \in G^{(1)}} k(x)^\times \xrightarrow{\partial} \bigoplus_{x \in G^{(2)}} \mathbb{Z}$$

induit une 2-extension

$$0 \to K_2(k) \to K_2(k(G)) \to \mathcal{Y}(G) \to \mathbb{Z} \to 0,$$

où

$$\mathcal{Y}(G) = \ker\Big( \bigoplus_{x \in G^{(1)}} k(x)^\times \xrightarrow{\partial} \bigoplus_{x \in G^{(2)}} \mathbb{Z} \Big).$$

Cette extension étant équivariante pour l'action de $G(k)$ par translation à gauche, elle définit une classe de $\mathrm{Ext}^2_{\mathbb{Z}[G(k)]}(\mathbb{Z}, K_2(k)) = H^2(G(k), K_2(k))$. D'après la proposition 2.8, cette classe est représentée par l'extension centrale de Brylinski–Deligne

$$(\mathscr{E}(G,c,k)) \qquad 0 \to K_2(k) \to \mathscr{E}(G,c,k) \to G(k) \to 1,$$

où $c \in \mathscr{X}(G)$ est un élément d'image 1 dans $\mathbb{Z}$, et $\mathscr{E}(G,c,k)$ est définie comme dans le lemme 2.7(1). Notons que l'extension $\mathscr{E}(G,c,k)$, contrairement à sa classe, dépend explicitement du choix de $c$, dorénavant fixé.

**Remarque 3.1.** Dans l'article [Brylinski et Deligne 2001], cette extension n'est pas définie de cette façon alors qu'elle l'était dans une version préliminaire. Les deux constructions coïncident et c'est d'ailleurs celle présentée ici qui est utilisée dans [Gille 2000].

Notons $\mathrm{Aut}(G)$ le $k$-groupe algébrique des automorphismes de $G$. Le groupe $\mathrm{Aut}(G)(k)$ agit naturellement sur le complexe ci-dessus, d'où une action sur la 2-extension de $G(k)$-modules associée, qui est triviale sur $K_2(k)$. De plus, on a :

**Lemme 3.2.** (1) *L'action de* $\mathrm{Aut}(G)(k)$ *sur* $H^1(G, \mathscr{K}_2) = \mathbb{Z}$ *est triviale.*

(2) *Si* $G(k)$ *est parfait, alors l'extension de Brylinski–Deligne* $\mathscr{E}(G,c,k)$ *est canoniquement* $\mathrm{Aut}(G)(k)$*-équivariante.*

*Démonstration.*

(1) Considérons une représentation fidèle $\rho : G \rtimes \mathrm{Aut}(G) \to \mathrm{SL}_N$, et notons $\rho_0 : G \to \mathrm{SL}_N$ sa restriction à $G$. Le morphisme induit

$$\mathbb{Z} = H^1(\mathrm{SL}_N, \mathscr{K}_2) \xrightarrow{\rho_0^*} H^1(G, \mathscr{K}_2) = \mathbb{Z}$$

est la multiplication par un entier strictement positif $d_{\rho_0}$, l'indice de Dynkin de $\rho_0$ [Garibaldi et al. 2003, p. 122]. Soit maintenant $f \in \mathrm{Aut}(G)(k)$. Du diagramme commutatif

$$\begin{array}{ccc} G & \xrightarrow{\rho_0} & \mathrm{SL}_N \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathrm{Int}(\rho(1,f))} \\ G & \xrightarrow{\rho_0} & \mathrm{SL}_N, \end{array}$$

on déduit le diagramme

$$\begin{array}{ccc} H^1(G, \mathscr{K}_2) & \xleftarrow{\rho_0^\star} & H^1(\mathrm{SL}_N, \mathscr{K}_2) \\ {\scriptstyle f^\star}\uparrow & & \uparrow{\scriptstyle (\mathrm{Int}(\rho(1,f)))^\star} \\ H^1(G, \mathscr{K}_2) & \xleftarrow{\rho_0^\star} & H^1(\mathrm{SL}_N, \mathscr{K}_2). \end{array}$$

Or la flèche verticale de droite est l'identité de $\mathbb{Z}$ puisque l'action de $\mathrm{SL}_N$ sur lui-même par automorphismes intérieurs induit une action triviale de $\mathrm{SL}_N(k)$ sur

$H^1(\mathrm{SL}_N, \mathscr{K}_2)$ [Garibaldi et al. 2003, lemme 6.9]. Par commutativité du diagramme, il en est de même de celle de gauche, ce qui prouve la première assertion.

(2) La description explicite de $\mathscr{E}(G, c, k) \subset G(k) \times K_2(k(G))$ donnée par le lemme 2.7(1) montre que l'action de $f \in \mathrm{Aut}(G)(k)$ sur $G(k) \times K_2(k(G))$, induit une application

$$\mathscr{E}(G, c, k) \to \mathscr{E}(G, f.c, k),$$

qui vaut l'identité sur $K_2(k)$. Ainsi, $\mathscr{E}(G, f.c, k)$ n'est rien d'autre que le pull-back $(f^{-1})^\star(\mathscr{E}(G, c, k))$. Or, comme l'action de $\mathrm{Aut}(G)(k)$ sur $H^1(G, \mathscr{K}_2) = \mathbb{Z}$ est triviale, les éléments $c$ et $f.c$ ont la même image dans $\mathbb{Z}$. Par le lemme 2.7(2), on en déduit que les extensions de Brylinski–Deligne correspondante $\mathscr{E}(G, c, k)$ et $(f^{-1})^\star(\mathscr{E}(G, c, k))$ sont équivalentes. Enfin, $G(k)$ étant supposé parfait, ceci prouve que l'action de $\mathrm{Aut}(G)(k)$ s'étend à $\mathscr{E}(G, c, k)$ par le lemme 2.3. □

**3B.** ***Action galoisienne.*** Soit maintenant $L/k$ une extension galoisienne finie de groupe $\Gamma$. En poussant $c$ dans $\mathscr{Z}(G_L)$, on obtient une extension centrale

$$\mathscr{E}(G, c, L): \qquad 0 \to K_2(L) \to \mathscr{E}(G, c, L) \to G(L) \to 1.$$

**Lemme 3.3.** *Si $G(L)$ est parfait, $\mathscr{E}(G, c, L)$ est une extension canoniquement $(\mathrm{Aut}(G)(L) \rtimes \Gamma)$-équivariante.*

*Démonstration.* Les groupes $\Gamma$ et $\mathrm{Aut}(G)(L)$ agissent tous deux sur l'extension $\mathscr{E}(G, c, L)$ en vertu des lemmes 2.9 et 3.2(2). Il est clair que sur $K_2(L)$ et sur $G(L)$, ces deux actions sont compatibles, c'est-à-dire induisent un morphisme

$$\mathrm{Aut}(G)(L) \rtimes \Gamma \to \mathrm{Aut}(K_2(L)) \times \mathrm{Aut}(G(L)).$$

Or on vient de voir que les images de $\mathrm{Aut}(G)(L)$ et de $\Gamma$ sont toutes deux incluses dans

$$\mathrm{Aut}(\mathscr{E}(G, c, L), K_2(L)) \subset \mathrm{Aut}(K_2(L)) \times \mathrm{Aut}(G(L)).$$

Le morphisme ci-dessus se factorise donc bien par $\mathrm{Aut}(\mathscr{E}(G, c, L), K_2(L))$. □

**Lemme 3.4.** *On suppose $G(L)$ parfait. Soient $z \in Z^1(\Gamma, \mathrm{Aut}(G)(L))$ un 1-cocycle et $\phi : G \times_k L \cong {}_zG \times_k L$ une trivialisation satisfaisant $z_\gamma = \phi^{-1}\gamma(\phi)$ pour tout $\gamma \in \Gamma$. On pose*

$$d := \phi^*(c_L) \in \mathscr{Z}({}_zG \times_k L).$$

*Alors l'extension tordue par $z$*

$$0 \to K_2(L) \to {}_z\mathscr{E}(G, c, L) \to {}_zG(L) \to 1$$

*est $(\mathrm{Aut}({}_zG)(L) \rtimes \Gamma)$-isomorphe à l'extension $\mathscr{E}({}_zG, d, L)$.*

Notons que la propriété galoisienne $\mathscr{Z}({}_zG) = \mathscr{Z}({}_zG \times_k L)^\Gamma$ est cruciale ici [Colliot-Thélène et Raskind 1985, proposition 3.6] pour savoir que $d \in \mathscr{Z}({}_zG)$.

*Démonstration.* La description explicite de ${}_z\mathscr{E}(G, c, L)$ et $\mathscr{E}({}_zG, d, L)$ donnée par le lemme 2.7(1) montre que $\phi$ induit un isomorphisme entre ces deux extensions de ${}_zG(L)$ par $K_2(L)$. Toutes deux sont $\mathrm{Aut}({}_zG)(L) \rtimes \Gamma$ équivariantes, mais il n'est pas clair à priori que l'isomorphisme entre les deux soit également équivariant. Le lemme 2.3 permet de le montrer sans aucun calcul ; il suffit en effet d'observer que l'action de $\mathrm{Aut}({}_zG)(L) \rtimes \Gamma$ sur $K_2(L)$ et ${}_zG(L)$ est la même pour chacune des deux extensions. □

La proposition qui suit est l'ingrédient clef pour effectuer des calculs. Elle repose sur le lien, établi dans [Gille 2000, Lemme 5], entre le bord associé à l'extension de Brylinski–Deligne, et l'invariant de Rost. Précisément, étant donné un 1-cocycle $z \in Z^1(\Gamma, \mathrm{Aut}(G)(L))$, on y montre l'anti-commutativité du diagramme

$$\begin{array}{ccc} H^1(\Gamma, {}_zG(L)) & \xrightarrow{r_{zG}} & \mathrm{Ker}\big(H^3(k) \to H^3(L)\big) \\ \downarrow{\scriptstyle \rho} & & \| \\ H^2(\Gamma, K_2(L)) & \xrightarrow{a_k^L} & \mathrm{Ker}\big(H^3(k) \to H^3(L)\big), \end{array}$$

où $r_{zG}$ désigne l'invariant de Rost, $a_k^L$ est la flèche construite par B. Kahn [1993], et $\rho : H^1(\Gamma, {}_zG(L)) \to H^2(\Gamma, K_2(L))$ est le bord associé à l'extension de Brylinski–Deligne $\mathscr{E}({}_zG, d_0, L)$ pour un certain $d_0 \in \mathfrak{L}({}_zG)$ d'image 1 dans $\mathbb{Z} = H^1({}_zG, \mathscr{K}_2)$. On a alors :

**Proposition 3.5.** *On suppose $G(L)$ parfait. On se donne $u \in Z^1(\Gamma, {}_zG(L))$, vu comme une section du morphisme $G(L) \rtimes^z \Gamma \to \Gamma$, et on considère l'extension retirée en arrière*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_2(L) & \longrightarrow & \mathscr{E}(G, c, L) \rtimes^z \Gamma & \longrightarrow & G(L) \rtimes^z \Gamma & \longrightarrow & 1 \\ & & \| & & \uparrow & & \uparrow{\scriptstyle u} & & \\ 0 & \longrightarrow & K_2(L) & \longrightarrow & E(z, u) & \longrightarrow & \Gamma & \longrightarrow & 1. \end{array}$$

*Alors on a*

$$a_k^L\big(r_{zG}([u])\big) = -\big[E(z, u)\big] \in H^2(\Gamma, K_2(L)).$$

*Démonstration.* Comme on vient de l'expliquer, il suffit de calculer l'image de

$$[u] \in H^1(\Gamma, {}_zG(L))$$

sous le bord associé l'extension de Brylinski–Deligne $\mathscr{E}({}_zG, d_0, L)$, où $d_0 \in \mathfrak{L}({}_zG)$ est un élément d'image 1 dans $\mathbb{Z} = H^1({}_zG, \mathscr{K}_2)$. Si l'on prend $d_0 = \phi^*(c_L)$, le lemme 3.4 indique que l'on peut remplacer $\mathscr{E}({}_zG, d_0, L)$ par l'extension tordue ${}_z\mathscr{E}(G, c, L)$. La version tordue du lemme 2.1 montre que l'extension de groupes $E(z, u)$ ci-dessus représente la classe $\rho([u])$. □

## 4. Calculs explicites pour $\mathrm{SL}_1(A)$

### 4A. *Algèbres cycliques.*

**Proposition 4.1.** *Soient $L/k$ une extension galoisienne cyclique de degré $n$ et de groupe $\Gamma = \langle\sigma\rangle$, et $b \in k^\times$. On considère la $k$-algèbre cyclique*

$$A = (L/k, \sigma, b) = L \oplus Ly \oplus \cdots \oplus Ly^{n-1}$$

*définie par les relations $y^n = b$ et $\lambda y = y\sigma(\lambda)$, pour $\lambda \in L$. Quel que soit $[v] \in k^\times/\mathrm{Nrd}(A^\times) \cong H^1(\Gamma, \mathrm{SL}_1(A)(L))$, on a*

$$a_k^L\big(r_{\mathrm{SL}_1(A)}([v])\big) = \big[\{v, b\}\big] \in K_2(k)/N(K_2(L)) \cong H^2(\Gamma, K_2(L)).$$

*Démonstration.* Notons

$$f_b = \begin{bmatrix} 0 & 0 & \cdots & 0 & b \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \in \mathrm{GL}_n(k).$$

On rappelle [Gille et Szamuely 2006, §2.5] que l'algèbre $A$ est la tordue $A = {}_zM_n(k)$ de $M_n(k)$ par le cocycle $z : \Gamma \to \mathrm{PGL}_n(L)$, défini par $\sigma^i \mapsto [f_b^i]$. Un point remarquable est que $z$ est à valeurs dans $\mathrm{PGL}_n(k)$, c'est-à-dire définit un morphisme de groupes $\Gamma \to \mathrm{PGL}_n(k)$. On se donne également une trivialisation $\phi : M_n(L) \cong A \otimes_k L$ telle que $z_\tau = \phi^{-1}\tau(\phi)$ pour tout $\tau \in \Gamma$. Pour utiliser la proposition 3.5, on doit réaliser la classe $[u]$ par un cocycle $u \in Z^1(\Gamma, {}_z\,\mathrm{SL}_n(L))$. Le $\mathrm{SL}_1(A)$-torseur correspondant à $v$ est donné par l'équation $v = \mathrm{Nrd}_A(y)$.

Notons $(\lambda_i)_{i=1,\ldots,n}$ la base canonique des cocaractères du tore diagonal $\mathbb{G}_m^n$ de $\mathrm{GL}_n$. L'élément $\phi(\lambda_1(v)) \in \mathrm{GL}_1(A)(L)$ satisfait

$$\mathrm{Nrd}_{A_L}(\phi(\lambda_1(v))) = \mathrm{Nrd}_{M_n(k)}(\lambda_1(v)) = \det(\lambda_1(v)) = v.$$

Un 1-cocycle représentant ce torseur est donc donné par

$$u_\sigma = \phi(\lambda_1(v))^{-1}\,\sigma(\phi(\lambda_1(v))) \in \mathrm{SL}_1(A)(L).$$

Vu comme section du morphisme $\mathrm{SL}_n(L) \rtimes^z \Gamma \to \Gamma$, ce cocycle est défini par l'homomorphisme $u : \Gamma \to \mathrm{SL}_n(L) \rtimes^z \Gamma$ donné par $\sigma \to t\,\sigma$, où

$$t = \lambda_1(v)^{-1}\lambda_2(v) = \begin{bmatrix} v^{-1} & 0 & & \cdots & 0 & 0 \\ 0 & v & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 & 1 \end{bmatrix} \in \mathrm{SL}_n(L).$$

On se donne $c \in \mathscr{Z}(\mathrm{SL}_n)$ d'image 1 dans $H^1(\mathrm{SL}_n, \mathscr{K}_2)$. On considère alors l'extension retirée en arrière

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_2(L) & \rightarrowtail & \mathscr{E}(\mathrm{SL}_n, c, L) \rtimes^z \Gamma & \twoheadrightarrow & \mathrm{SL}_n(L) \rtimes^z \Gamma & \longrightarrow & 1 \\ & & \| & & \uparrow & & \uparrow u & & \\ 0 & \longrightarrow & K_2(L) & \longrightarrow & E(z,u) & \longrightarrow & \Gamma & \longrightarrow & 1 \end{array}$$

dont on va calculer la classe. Suivant le lemme 2.6, celle-ci est donnée par l'élément

$$(\tilde{t}\,\sigma)^n \in K_2(L) \subset \mathscr{E}(\mathrm{SL}_n, c, L) \rtimes^z \Gamma$$

où $\tilde{t}$ désigne un relevé de $t$ dans $\mathscr{E}(\mathrm{SL}_n, c, L)$ que l'on choisit bien sûr provenant de $\mathscr{E}(\mathrm{SL}_n, c, k)$. En utilisant la forme précise du cocycle $z$, on écrit alors

$$\begin{aligned} (\tilde{t}\,\sigma)^n &= \tilde{t}\,(\sigma\tilde{t}\sigma^{-1}) \cdots (\sigma^{n-1}\,\tilde{t}\,\sigma^{1-n}) \\ &= \tilde{t}\,(f_b.\sigma(\tilde{t})) \cdots (f_b^{n-1}.\sigma^{n-1}(\tilde{t})) \\ &= \tilde{t}\,(f_b.\tilde{t}) \cdots (f_b^{n-1}.\tilde{t}) \in \mathscr{E}(\mathrm{SL}_n, c, L). \end{aligned} \tag{4-1}$$

Notant $T = \mathrm{Ker}(\mathbb{G}_m^n \to \mathbb{G}_m)$ le tore maximal standard de $\mathrm{SL}_n$ et $T_{ad} = \mathbb{G}_m^n/\mathbb{G}_m$ celui de $\mathrm{PGL}_n$, on a alors besoin de l'accouplement

$$h : T_{ad}(L) \times T(L) \to K_2(L)$$

défini par $h(x, y) = (x.\tilde{y})\tilde{y}^{-1}$ où $\tilde{y}$ est un relevé (arbitraire) de $y$ dans $\mathscr{E}(\mathrm{SL}_n, c, L)$ [Brylinski et Deligne 2001, §4.13]. En pratique, $h$ se calcule avec le diagramme commutatif

$$\begin{array}{ccccc} T_{ad}(L) & \times & T(L) & \longrightarrow & K_2(L) \\ \uparrow & & \downarrow & & \| \\ (L^\times)^n & \times & (L^\times)^n & \longrightarrow & K_2(L) \\ (x_i) & & (y_i) & \mapsto & \sum_{i=1}^n \{x_i, y_i\}. \end{array}$$

Écrivons $f_b = \lambda_1(b)\, f_1$ et $f_b^i = (\lambda_1(b) \cdots \lambda_i(b))\, f_1^i$ $(i = 1, \ldots, n-1)$. Vu que $f_1$ normalise le tore $T$, on a $f_1^i.\tilde{t} \in T(L)$. Par suite,

$$f_b^i.\tilde{t} = (\lambda_1(b) \cdots \lambda_i(b)).(f_1^i.\tilde{t}) = h\big(\lambda_1(b) \cdots \lambda_i(b),\, f_1^i.t\big) f_1^i.\tilde{t}.$$

Reportant ceci dans (4-1), on obtient

$$(\tilde{t}\,\sigma)^n = \alpha \times \tilde{t}\,(f_1.\tilde{t}) \cdots (f_1^{n-1}.\tilde{t}),$$

où $\alpha = \sum_{i=1}^{n-1} h\big(\lambda_1(b) \cdots \lambda_i(b),\, f_1^i.t\big) \in K_2(L)$. Le terme de droite,

$$\tilde{t}\,(f_1.\sigma(\tilde{t})) \cdots (f_1^{n-1}.\sigma^{d-1}(\tilde{t})),$$

étant celui du cas $b=1$, est inessentiel ; il appartient donc à l'image de $N : K_2(L) \to K_2(L)$. Il reste donc à calculer le premier terme $\alpha$. Vu que $t = \lambda_1(v)^{-1}\lambda_2(v)$, on a

$$\alpha = \sum_{i=1}^{n-1} h\big(\lambda_1(b)\cdots\lambda_i(b),\ \lambda_i(v)^{-1}\lambda_{i+1}(v)\big) = \sum_{i=1}^{n-1}\{b, v^{-1}\} = -(n-1)\{b, v\}.$$

On conclut que $[E(z,u)] = \{b, v\} \in K_2(L)^\Gamma / N.K_2(L)$, d'où la formule voulue par application de la proposition 3.5. □

**4B.** ***Calcul dans un cas particulier.***

**Lemme 4.2.** *Soit $n$ un entier $\geq 1$. On suppose que $k$ admet une racine primitive $n$-ième de l'unité $\zeta_n$. On pose $K = k((x))$ et on considère l'extension de Kummer $L = K(x^{1/n})$. On note $\sigma$ le générateur de* $\mathrm{Gal}(L/K)$ *défini par $\sigma(x^{1/n}) = \zeta_n x^{1/n}$ et $\chi \in H^1(K, \mathbb{Z}/n\mathbb{Z})$ le caractère de* $\mathrm{Gal}(L/K) \to \mathbb{Z}/n\mathbb{Z}$ *appliquant $\sigma$ sur* 1.

(1) *On a un isomorphisme*

$$K_2(k)/nK_2(k) \xrightarrow{\chi \cup h_n}{}^{\sim} \mathrm{Ker}\big(H^3(K) \to H^3(L)\big)$$

*où $h_n : K_2(k)/nK_2(k) \to H^2(k, \mu_n^{\otimes 2})$ désigne le symbole galoisien.*

(2) *Le composé*

$$K_2(k)/nK_2(k) \xrightarrow{\chi.h_n}{}^{\sim} \mathrm{Ker}\big(H^3(K) \to H^3(L)\big) \xrightarrow{a_K^L} H^2(\mathrm{Gal}(L/K), K_2(L))$$

*est injectif et applique le symbole $\{x, y\} \in K_2(k)$ sur*

$$\big[\{y, x\}\big] \in K_2(L)^{\mathrm{Gal}(L/K)}/N.K_2(L) \cong H^2(\mathrm{Gal}(L/K), K_2(L)).$$

(3) *Soient $b, v \in k^\times$. On considère l'algèbre $A = (L/K, \sigma, b)$. Alors*

$$r_{\mathrm{SL}_1(A)}([v]) = (v) \cup [A] \in H^3(K).$$

*Démonstration.* (1) On observe tout d'abord que

$$\mathrm{Ker}\big(H^3(K, \mu_n^{\otimes 2}) \to H^3(L, \mu_n^{\otimes 2})\big) \cong \mathrm{Ker}\big(H^3(K) \to H^3(L)\big).$$

On a un diagramme commutatif exact de suites exactes de résidus [Garibaldi et al. 2003, (8.4), p. 20]

$$\begin{array}{ccccccccc}
0 & \longrightarrow & H^3(k, \mu_n^{\otimes 2}) & \to & H^3(K, \mu_n^{\otimes 2}) & \xrightarrow{\partial_K} & H^2(k, \mu_n) & \longrightarrow & 0 \\
 & & \| & & \mathrm{res}\downarrow & & \downarrow{\scriptstyle \times n} & & \\
0 & \longrightarrow & H^3(k, \mu_n^{\otimes 2}) & \to & H^3(L, \mu_n^{\otimes 2}) & \xrightarrow{\partial_K} & H^2(k, \mu_n) & \longrightarrow & 0
\end{array}$$

Ceci produit un isomorphisme

$$\mathrm{Ker}\big(H^3(K, \mu_n^{\otimes 2}) \to H^3(L, \mu_n^{\otimes 2})\big) \xrightarrow{\sim} H^2(k, \mu_n) \xrightarrow{\sim} H^2(k, \mu_n^{\otimes 2}),$$

avec l'identification $\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} \mu_n$ envoyant 1 sur $\zeta_n$. La réciproque de cette application est donnée par le cup-produit par $\chi$. En le combinant avec l'isomorphisme de Merkurjev–Suslin [1982]

$$h_n : K_2(k)/nK_2(k) \xrightarrow{\sim} H^2(k, \mu_n^{\otimes 2}),$$

on obtient un isomorphisme

$$K_2(k)/nK_2(k) \xrightarrow{\sim} \mathrm{Ker}\big(H^3(K, \mu_n^{\otimes 2}) \to H^3(L, \mu_n^{\otimes 2}))\big), \{u, v\} \mapsto \chi \cup h_n(\{u, v\}).$$

(2) Suivant [Gille 2000, lemme 2], on a

$$a_K^L\big(\chi \cup h_n(\{u, v\})\big) = -\{u, v\} \in K_2(L)^{\mathrm{Gal}(L/K)}/N.K_2(L).$$

Il reste à vérifier que le morphisme $K_2(k)/nK_2(k) \to K_2(L)^{\mathrm{Gal}(L/K)}/N.K_2(L)$ est injectif. Cela découle du fait qu'il est scindé par la spécialisation

$$s : K_2(L) \to K_2(k), \quad \alpha \mapsto \partial_L\big((-x^{1/n}).\alpha\big),$$

où l'on voit $(-x^{1/n}).\alpha \in K_3^M(L)$.

(3) La proposition 4.1 indique que

$$a_K^L\big(r_{\mathrm{SL}_1(A)}([v])\big) = \big[\{v, b\}\big] \in K_2(L)^\Gamma/N.K_2(L) \cong H^2(\Gamma, K_2(L)).$$

Suivant (2), ceci entraîne

$$r_{\mathrm{SL}_1(A)}([v]) = \chi \cup h_n\big(\{b, v\}\big) = \chi \cup (b) \cup (v) = (v) \cup [A] \in H^3(K). \qquad \square$$

**4C.** ***Démonstration du théorème 1.1.***

**Lemme 4.3.** *Soit $n$ un entier $\geq 1$ inversible dans $k$. Il existe un entier positif $m$ tel que pour tout corps $F/k$ et toute algèbre simple centrale $A/F$ de degré $n$, l'invariant de Rost $F^\times/\mathrm{Nrd}(A^\times) \to H^3(F)$ est donné par $[v] \mapsto m\,(v) \cup [A]$.*

La démonstration passe par les torseurs versels [Garibaldi et al. 2003, p. 11].

*Démonstration.* Pour tout corps $F/k$, on définit

$$a_F : H^1(F, \mu_n \times \mathrm{PGL}_n) = F^\times/(F^\times)^n \times H^1(F, \mathrm{PGL}_n) \to H^3(F)$$

par

$$a_F\big((v), [A]\big) := \text{Invariant de Rost du } \mathrm{SL}_1(A)\text{-torseur } v = \mathrm{Nrd}_A(y).$$

Les $a_F$ définissent un invariant cohomologique du groupe $\mu_n \times \mathrm{PGL}_n$. Notant $\mathscr{A}/k(X)$ une algèbre simple centrale "verselle" de degré $n$, on sait, selon [Garibaldi et al. 2003, théorème 11.5, p. 137], qu'il existe un entier $m$ tel que

$$r\big([P_{k(X)(t)}])\big) = m\,(t) \cup [\mathscr{A}_{k(X)}] \in H^3(k(X)(t)).$$

Le principe de spécialisation [Garibaldi et al. 2003, (3.3), p. 109] nous permet de conclure que $a_F\big((v),[A]\big)=m\,(v)\cup[A]$ pour tout corps $F/k$, tout $c\in F^\times$ et toute $F$-algèbre simple centrale $A$. □

Pour établir le théorème 1.1(1), on commence par le cas où $n$ est inversible dans $k$. L'idée est de tester l'entier $m$ sur un exemple. Plus précisément, il suffit d'exhiber une extension $F/k$, une algèbre $A/F$ de degré $n$, et un élément $v\in F^\times$ tel que $r_{\mathrm{SL}_1(A)}([v])=(v)\cup[A]$ et tel que $r_{\mathrm{SL}_1(A)}([v])$ soit d'ordre exactement $n$.

Posons $F=\bar{k}((t))((y))((x))$. On note $A$ l'algèbre cyclique sur $F$ présentée par $X^n=x$, $Y^n=y$ et $XY=\zeta_n YX$. Alors $(t)\cup[A]=(t)\cup(x)\cup(y)$ est d'ordre $n$ dans $H^3(F,\mu_n^{\otimes 2})\cong\mathbb{Z}/n\mathbb{Z}$. On conclut que $m=1$ par le lemme 4.2(3).

**4D.** ***Cas d'une p-algèbre.*** On suppose ici que $k$ est de caractéristique $p>0$ et que $A$ est d'indice $p^h$. Si $d\geq 0$, rappelons la définition du groupe

$$H^{d+1}(k,(\mathbb{Q}/\mathbb{Z})(d))=\bigcup_{m\geq 1}H^{d+1}(k,(\mathbb{Z}/p^m\mathbb{Z})(d)).$$

au moyen des différentielles logarithmiques. Pour tout schéma $X$ de caractéristique $p$, on note $W_m\Omega^d_X$ le faisceau de de Rham–Witt sur $X_{ét}$ de degré $m$ et de poids $d$ [Illusie 1979, I.1] et on note $W_m\Omega^q_{X,\log}$ le sous-faisceau de $W_m\Omega^d_X$ engendré localement pour la topologie étale par les différentielles logarithmiques

$$\mathrm{dlog}(x_1)\wedge\cdots\wedge\mathrm{dlog}(x_q).$$

On note $\nu_m(d)/X=W_m\Omega^d_{X,\log}$ et $\nu(d)/X=W_1\Omega^q_{X,\log}$.

Par définition, on a $H^{d+1}(k,(\mathbb{Z}/p^n\mathbb{Z})(d))=H^1(k,\nu_m(d)(k_s))$. Par ailleurs, le théorème de Bloch–Gabber–Kato [Bloch et Kato 1986] établit un isomorphisme canonique

$$K^M_d(k)/p^mK^M_d(k)\xrightarrow{\sim}\nu_m(d)(k),\quad \{x_1,\dots,x_d\}\mapsto\frac{dx_1}{x_1}\wedge\cdots\wedge\frac{dx_d}{x_d}.$$

On a donc un isomorphisme

$$H^{d+1}(k,(\mathbb{Z}/p^n\mathbb{Z})(d))\xrightarrow{\sim}H^1\big(k,K^M_d(k_s)/p^mK^M_d(k_s)\big)$$

Pour tout entier $r\geq 0$, le produit

$$K^M_d(k_s)/p^mK^M_d(k_s)\times K^M_r(k_s)/p^mK^M_r(k_s)\to K^M_{r+d}(k_s)/p^mK^M_{r+d}(k_s)$$

induit le cup-produit

$$H^1\big(k,K^M_d(k_s)/p^mK^M_d(k_s)\big)\times H^0\big(k,K^M_r(k_s)/p^mK^M_r(k_s)\big)\to H^1\big(k,K^M_{d+r}(k_s)/p^mK^M_{d+r}(k_s)\big).$$

Vu que $K_r^M(k)/p^m K_r^M(k) = H^0\big(k, K_r^M(k_s)/p^m K_r^M(k_s)\big)$, on a donc un cup-produit

$$H^{d+1}(k, (\mathbb{Z}/p^m\mathbb{Z})(d)) \times K_r^M(k)/p^m K_r^M(k) \to H^{d+r+1}(k, (\mathbb{Z}/p^m\mathbb{Z})(d+r)).$$

Rappelons que l'on a un isomorphisme $H^2(k, (\mathbb{Z}/p^m\mathbb{Z})(1)) \xrightarrow{\sim} {}_{p^m}\mathrm{Br}(k)$ (Kato, voir [Gille et Szamuely 2006, th. 9.2.4] dans le cas $m = 1$). En particulier, cela permet de voir la classe de $[A]$ dans $H^2(k, (\mathbb{Q}_p/\mathbb{Z}_p)(1))$.

Nous allons maintenant montrer que l'argument de relèvement en caractéristique nulle de l'invariant de Rost [Gille 2000, §5.1] fonctionne ici. Soit $K$ un corps complet pour une valuation discrète, de caractéristique nulle et de corps résiduel $k$. On note $O$ son anneau des entiers et $\mathscr{A}$ une $R$-algèbre d'Azumaya relevant la $k$-algèbre simple centrale $A$. On a le diagramme commutatif

$$\begin{array}{ccccc}
K^\times/\mathrm{Nrd}(\mathscr{A}_K^\times) & \cong & H^1(K, \mathrm{SL}_1(\mathscr{A}_K)) & \longrightarrow & H^3(K, \mathbb{Q}/\mathbb{Z}(2)) \\
\uparrow & & \uparrow & & \uparrow \\
O^\times/\mathrm{Nrd}(\mathscr{A}^\times) & \cong & H^1(O, \mathrm{SL}_1(\mathscr{A})) & & -i_k^K \\
\downarrow & & \downarrow & & \uparrow \\
k^\times/\mathrm{Nrd}(A^\times) & \cong & H^1(k, \mathrm{SL}_1(A)) & \longrightarrow & H^3(k, (\mathbb{Q}_p/\mathbb{Z}_p)(2))
\end{array} \tag{4-2}$$

où $i_k^K$ désigne le morphisme (injectif) de relèvement de Kato qui satisfait aux compatibilités suivantes.

**Lemme 4.4.** (1) *On a*

$$i_k^K(\alpha \cup \{\bar{u}_1, \dots, \bar{u}_r\}) = i_k^K(\alpha) \cup \{u_1, \dots, u_r\}$$

*pour tous* $u_1, \dots, u_r \in O^\times$ *et* $\alpha \in H^{d+1}(k, (\mathbb{Z}/p^m\mathbb{Z})(d))$.

(2) $i_k([A]) = [\mathscr{A}_K] \in H^2(K, \mathbb{Q}_p/\mathbb{Z}_p)(1))$.

*Démonstration.* (1) Cela suit de la définition de $i_k^K$ ; voir la preuve de [Kato 1982, proposition 2].

(2) Suivant le théorème d'Albert (voir [Gille et Szamuely 2006, §9.1]), $A$ est semblable à une algèbre cyclique $B = (L/k, \sigma, \bar{b})$ (conventions du proposition 4.1) où $k'/k$ est une extension cyclique de corps de groupe $\langle\sigma\rangle$ d'ordre $p^s$ et $\bar{b} \in k^\times$. Sa classe dans le groupe de Brauer est $\chi \cup (\bar{b})$ où $\chi : \mathrm{Gal}(k'/k) \to \mathbb{Z}/p^s\mathbb{Z}$ est le caractère appliquant $\sigma$ sur 1 et $\bar{b} \in k^\times$ [Gille et Szamuely 2006, 4.7.3]. Soit $O'/O$ une extension galoisienne cyclique relevant $k'/k$ et $b \in O^\times$ un relevé de $\bar{b}$. Alors la $O$-algèbre d'Azumaya cyclique $\mathscr{B} = (O'/O, \sigma, b)$ relève $B$. On a $\mathscr{B}_K = (K'/K, \sigma, b)$ où $K' = \mathrm{Frac}(O')$ et sa classe dans $\mathrm{Br}(K)$ est $\Theta \cup (b)$ où $\Theta : \mathrm{Gal}(K'/K) \to \mathbb{Z}/p^s\mathbb{Z}$ est le caractère appliquant $\sigma$ sur 1. Mais par définition $i_k^K : H^1(k, \mathbb{Z}/p^s\mathbb{Z}) \to H^1(K, \mathbb{Z}/p^s\mathbb{Z})$ applique $\chi$ sur $\Theta$, donc (1) permet de conclure que $i_k([B]) = [\mathscr{B}_K]$. □

Nous montrons le théorème 1.1(2). Étant donné $u \in O^\times$ de réduction $\bar{u} \in k$, on a

$$\begin{aligned} r_K([u]) &= (u) \cup [\mathscr{A}_K] && \text{[Cas de car. nulle]} \\ &= (u) \cup i_k^K[A] && \text{[Lemme 4.4(2)]} \\ &= i_k^K\big((\bar{u}) \cup [A]\big) && \text{[Lemme 4.4(1)].} \end{aligned}$$

Le diagramme (4-2) ci-dessus permet de conclure que $r_k([\bar{u}]) = -(\bar{u}) \cup [A]$.

## 5. Restriction au centre de l'invariant de Rost

Soit $G/k$ un groupe semi-simple simplement connexe absolument presque $k$-simple. On suppose dans cette partie que la caractéristique du corps $k$ ne divise pas l'exposant du centre $Z$ de $G$, de sorte que $Z$ est lisse. On s'intéresse à la restriction $\rho_G$ de l'invariant de Rost aux torseurs issus de $Z$, c'est-à-dire à la composée

$$H^1(k, Z) \to H^1(k, G) \to H^3(k),$$

où la première application est induite par l'inclusion de $Z$ dans $G$ et la seconde est l'invariant de Rost. L'ingrédient majeur pour faire ce calcul est la description des invariants des tores quasi-triviaux donnée dans [Merkurjev et al. 2002, théorème 1.1], qui montre qu'il existe une classe de cohomologie $t_{R,G} \in H^2(k, Z)$ telle que $\rho_G$ est donné par un certain cup-produit avec $t_{R,G}$. On sait de plus que cette classe $t_{R,G}$ est, suivant le type du groupe $G$, soit la classe nulle, soit un multiple non trivial de la classe de Tits $t_G \in H^2(k, Z)$ (voir [Merkurjev et al. 2002] pour $G$ de type classique ou [Garibaldi et Quéguiner-Mathieu 2007] pour $G$ de type exceptionnel). Le théorème 1.1 permet d'énoncer le résultat plus précis suivant :

**Corollaire 5.1.** *Si $G$ est de type $A$, $C_\ell$ avec $\ell$ impair, $D$, $E_6$ ou $E_7$, la restriction $\rho_G$ de l'invariant de Rost aux torseurs issus du centre est le cup-produit avec la classe de Tits du groupe $G$, $t_G \in H^2(k, Z)$, le cup-produit étant induit par l'application bilinéaire $Z(k_s) \times Z(k_s) \to \mu_n^{\otimes 2}$ spécifiée pour chaque type de groupe dans* [Garibaldi et Quéguiner-Mathieu 2007, §2]. *Pour les autres groupes, $\rho_G$ est nulle.*

*Démonstration.* Si le groupe $G$ est de type ${}^1A_{\ell-1}$, il est de la forme $G = \mathrm{SL}_1(A)$ pour une certaine $k$-algèbre centrale simple $A$ de degré $\ell$, et il a pour centre $Z = \mu_\ell$. Par [Knus et al. 1998, (31.7)], avec l'identification que l'on a choisie entre le groupe de Brauer de $k$ et $H^2(k)$, la classe de Tits de $G$ est alors la classe de Brauer de l'algèbre $A$, $t_{\mathrm{SL}_1(A)} = [A] \in H^2(k, \mu_n)$, et le corollaire découle donc dans ce cas de la description de l'invariant de Rost donnée dans le théorème 1.1.

Quand la composée $\rho_G$ est nulle, le résultat est prouvé dans [Merkurjev et al. 2002] ou [Garibaldi et Quéguiner-Mathieu 2007]. Pour les autres types de groupes, on sait (*loc. cit.*) que la restriction à $H^1(k, Z)$ de l'un des générateurs du groupe des invariants de degré 3 de $G$ est le cup-produit avec la classe de Tits. De sorte que $t_{R,G}$

est de la forme $m\,t_G$, où $m$ est un entier premier à l'indice de Dynkin $n_G$ du groupe $G$. Si $n_G$ est pair et $Z$ d'exposant 2, alors $m$ est impair et $m\,t_G = t_G \in H^2(k, Z)$. Le corollaire est ainsi prouvé pour $G$ de type $C_\ell$, $D_\ell$ avec $\ell$ pair et $E_7$.

Si $G$ est de type $E_6$, la preuve de [Garibaldi et Quéguiner-Mathieu 2007, §11] montre que si l'invariant de Rost pour le groupe $\mathrm{SL}_1(D)$ est le cup-produit avec $m[D]$, alors $t_{R,G} = m\,t_G$. En combinant avec le théorème 1.1, on obtient donc le résultat annoncé.

Il reste à prouver le corollaire pour les groupes de type ${}^2A_{\ell-1}$, et pour les groupes de type $D_\ell$ avec $\ell$ impair. On commence par les formes extérieures de $A_{\ell-1}$.

*Groupes de type* ${}^2A_{\ell-1}$. L'argument présenté ici est essentiellement tiré de [Merkurjev et al. 2002]. Le groupe $G$ est de la forme $G = \mathrm{SU}(B, \tau)$, où $B$ est une algèbre centrale simple de degré $\ell$ sur une extension quadratique $K$ de $k$, et $\tau$ est une involution $K/k$-semi-linéaire de $B$. Son centre est une forme tordue $Z = \mu_{\ell[K]}$ du groupe des racines $\ell$-ièmes de l'unité.

Supposons tout d'abord que $\ell$ est impair. Par [Merkurjev et al. 2002, (6)], le groupe $\mu_{\ell[K]}$ s'insère dans une suite exacte impliquant des tores quasi-triviaux, et en considérant la suite induite en cohomologie, on observe que $H^2(k, Z)$ s'injecte dans $H^2(k, R_{K/k}(\mathbb{G}_m)) = \mathrm{Br}(K)$. On en déduit que la restriction

$$\mathrm{res}_{K/k} : H^2(k, Z) \to H^2(K, Z)$$

est injective. Or, par le cas intérieur, la différence $t_{R,G} - t_G \in H^2(k, Z)$ est nulle sur $K$. On a donc bien $t_{R,G} = t_G$.

Supposons maintenant que $\ell = 2m$ est pair. Dans ce cas, par [Merkurjev et al. 2002, proposition 5.2], l'application

$$(\lambda_\star, \mathrm{res}_{K/k}) : H^2(k, Z) \to H^2(k, \mu_2) \times H^2(K, \mu_n)$$

est injective, où $\lambda_\star$ est induite par l'élévation à la puissance $m$. A nouveau par le cas intérieur, $\mathrm{res}_{K/k}(t_{R,G}) = \mathrm{res}_{K/k}(t_G)$. Il reste donc à montrer que $\lambda_\star(t_{R,G}) = \lambda_\star(t_G)$, qui par [Knus et al. 1998, (31.8)] est la classe de Brauer de l'algèbre discriminante $\mathcal{D}(B, \tau)$.

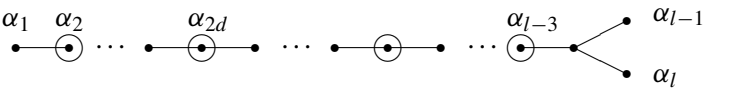
Pour cela rappelons que, par [Knus et al. 1998, (30.13)], le groupe $H^1(k, Z)$ est un quotient de $\{(x, y) \in k^\times \times K^\times \mid x^n = N_{K/k}(y)\}$. De plus, si on note $(x, y)_n$ la classe du couple $(x, y) \in k^\times \times K^\times$, alors par les calculs de [Merkurjev et al. 2002, p. 819], pour toute classe $\theta \in H^2(k, Z)$, le cup produit avec $(x, y)_n$ est donné par

$$(x, y)_n \cdot \theta = x \cdot \lambda_\star(\theta) + N_{K/k}(y \cdot \mathrm{res}_{K/k}(\theta)).$$

Ainsi, par [Merkurjev et al. 2002, §4.1], si l'algèbre $B$ est déployée, alors $\lambda_\star(t_{R,G})$ est la classe de Brauer de l'algèbre discriminante $\mathcal{D}(B, \tau)$, ce qui prouve que $\lambda_\star(t_{R,G}) = \lambda_\star(t_G)$ dans le cas déployé. Le cas général s'en déduit par extension des

scalaires aux corps des fonctions $E$ du transfert de $K$ à $k$ de la variété de Severi–Brauer de $B$ (*loc. cit.*). En effet, $E$ déploie $B$ et la restriction $\mathrm{Br}(k) \to \mathrm{Br}(E)$ est injective.

*Groupes de type $D_\ell$ avec $\ell$ impair.* Supposons maintenant que $G$ est de type $D_\ell$ avec $\ell$ impair. Son centre est $\mu_4$ dans le cas intérieur, et une forme tordue $\mu_{4[K]}$ dans le cas extérieur. On sait alors que $t_{R,G}$ est égal à $t_G$ ou $3t_G$, mais les arguments donnés dans [Merkurjev et al. 2002, §4.2.1, §4.3.1] ne permettent pas de lever l'ambiguïté. On va donc utiliser la méthode développée dans [Garibaldi et Quéguiner-Mathieu 2007], qui s'applique aussi aux groupes classiques. Nous présentons ici une esquisse de la preuve, comprenant les calculs qui ne figurent pas dans cet article-là, où le lecteur trouvera néanmoins certains détails des arguments. Rappelons d'abord que l'on peut supposer, pour calculer $t_{R,G}$, que le groupe $G$ a un indice de Tits de la forme suivante [Garibaldi et Quéguiner-Mathieu 2007, §4] :




Le groupe $G$ contient donc un $k$-tore déployé $S'$ de rang $(\ell-3)/2$. Considérons le sous-groupe dérivé $G'$ du centralisateur dans $G$ de $S'$. C'est un groupe semi-simple simplement connexe, dont le diagramme de Dynkin est obtenu à partir du diagramme ci-dessus en supprimant les sommets entourés. C'est donc un produit de groupes de type $A$. De plus, par [Garibaldi et Quéguiner-Mathieu 2007, proposition 5.5], le centre $Z$ de $G$ est contenu dans $G'$. On va donc calculer $\rho_G$ en passant par ce sous-groupe, et en utilisant le théorème 1.1.

Pour cela, il nous faut décrire précisément le groupe $G'$ et l'inclusion de $Z$ dans $G'$. Pour $i = 1, 3, \ldots, \ell-4$, on note $G'_i$ la composante de $G'$ correspondant au sommet $\alpha_i$ du diagramme ; elle est de la forme $G'_i = \mathrm{SL}_1(Q_i)$, où $Q_i$ est l'algèbre de Tits associée au poids fondamental $\omega_i$. Or, en consultant les tables de [Bourbaki 1982], on observe que les poids $\omega_1, \omega_3 \ldots, \omega_{\ell-4}$ sont égaux modulo le réseau des racines. Il découle donc de [Tits 1971, p. 211] (cf. également [Knus et al. 1998, (27.7)]) que les algèbres $Q_i$ sont toutes isomorphes. Ainsi, on a $G'_i = \mathrm{SL}_1(Q)$ pour une certaine algèbre de quaternions $Q$. Notons maintenant $G'_\ell$ la composante associée au sous-diagramme de sommets $\alpha_\ell$, $\alpha_{\ell-2}$ et $\alpha_{\ell-1}$. L'algèbre de Tits correspondant au poids $\omega_\ell$ est une algèbre de degré 4 sur $k$ dans le cas intérieur et sur $K$ dans le cas extérieur. Le groupe $G'_\ell$ est $G'_\ell = \mathrm{SL}_1(D)$ dans le premier cas et $G'_\ell = \mathrm{SU}(D, \tau)$ dans le second, où $\tau$ est une involution $K/k$ semi-linéaire de $D$. De plus, à nouveau par [Knus et al. 1998, (27.7)], on a $[Q] = 2[D]$ dans le cas intérieur et $\mathrm{res}_{K/k}([Q]) = 2[D]$ dans le cas extérieur.

La description de l'inclusion de $Z$ dans $G'$ peut se faire au niveau de la clôture séparable $k_s$ de $k$. Comme dans [Garibaldi et Quéguiner-Mathieu 2007, §8], le poids $\omega_\ell$ étant d'ordre 4 dans le quotient $\Lambda/\Lambda_r$, l'application $z_{\omega_\ell}$ associée par la

proposition 6.2 du même article induit un isomorphisme entre $\mu_4$ et $Z$. Avec les notations de [Garibaldi et Quéguiner-Mathieu 2007, §5.1], elle est donnée par

$$z_{\omega_\ell}(\zeta) = h_{4\omega_\ell}(\zeta) = h_1(\zeta^2)\, h_3(\zeta^2) \ldots h_{\ell-4}(\zeta^2) h_{\ell-2}(\zeta^2) h_{\ell-1}(\zeta^{\ell-2})\, h_\ell(\zeta^\ell).$$

On peut décrire de manière analogue le centre de $\mathrm{SL}_1(D)$. En comparant les deux formules, on obtient que l'inclusion de $Z$ dans le produit $Z'_1 \times Z'_3 \times \cdots \times Z'_{\ell-4} \times Z'_\ell$, où $Z'_i$ désigne le centre de $G'_i$, est donnée par

$$\zeta \mapsto (\zeta^2, \ldots, \zeta^2, \zeta^{\ell-2}).$$

Rappelons que le centre $Z'_i$ est isomorphe à $\mu_2$ pour $i = 1, 3, \ldots, \ell-4$, tandis que $Z'_l$ est isomorphe à $Z$. L'application induite au niveau des $H^1$ est donc

$$\begin{array}{ccc} H^1(k, Z) & \mapsto & \prod H^1(k, Z'_i) \\ a & \mapsto & (\lambda_\star(a), \ldots, \lambda_\star(a), (\ell-2)a), \end{array}$$

où $\lambda_\star$ désigne comme précédemment l'application induite par l'élévation au carré. Pour conclure, il ne reste plus qu'à appliquer la formule [Garibaldi et Quéguiner-Mathieu 2007, (5.8)].

Plaçons-nous tout d'abord dans le cas intérieur. Si $\ell \equiv 3[4]$, on obtient

$$\rho_G(a) = (\ell-2)(a) \cdot [D] = (a) \cdot [D].$$

Si maintenant $\ell \equiv 1[4]$, on obtient

$$\rho_G(a) = \lambda_\star(a) \cdot [Q] + (3a) \cdot [D].$$

Or $(2a) \cdot [D] = (a) \cdot [2D] = \lambda_\star(a) \cdot [Q]$, qui est d'ordre 2 dans $H^3(k)$.

On obtient donc dans les deux cas $\rho_G(a) = (a) \cdot [D]$. Or, pour faire ce calcul, on a identifié $\mu_4$ et $Z$ par l'intermédiaire du poids $\omega_\ell$. Par [Knus et al. 1998, (31.7)], cette identification induit une application $H^2(k, Z) \to H^2(k, \mu_4)$ qui envoie la classe de Tits $t_G$ sur la classe de Brauer de l'algèbre de Tits associée à $\omega_\ell$ qui est justement $[D]$. Le corollaire est donc prouvé dans le cas intérieur.

Notons que l'on aurait aussi pu identifier $Z$ et $\mu_4$ en utilisant le poids $\omega_{\ell-1}$. Un calcul analogue montre qu'il faudrait alors remplacer $\ell-2$ par $\ell$ dans la formule ci-dessus. Mais il faudrait également remplacer $D$ par l'algèbre de Tits associée au poids $\omega_{\ell-1}$, qui est l'algèbre opposée $D^{\mathrm{op}}$. Comme prévu, la formule ne dépend donc pas de l'identification choisie.

Dans le cas extérieur, un calcul analogue, conduit à $\rho_G(a) = a \cdot t_{G'_\ell}$. Si l'on identifie $Z$ et $Z'_\ell$ à $\mu_{4[K]}$ par le poids $\omega_\ell$, en appliquant à nouveau [Knus et al. 1998, (31.7)], on observe que $\mathrm{res}_{K/k}(t_G) = \mathrm{res}_{K/k}(t_{G'_\ell}) = [D]$ et $\lambda_\star(t_G) = \lambda_\star(t_{G_\ell}) = [Q]$. Par [Merkurjev et al. 2002, proposition 5.2], ceci permet d'identifier $t_G = t_{G'_\ell}$, et termine la preuve. □

## 6. Groupes exceptionnels de type $G_2$, $F_4$ et $E_8$

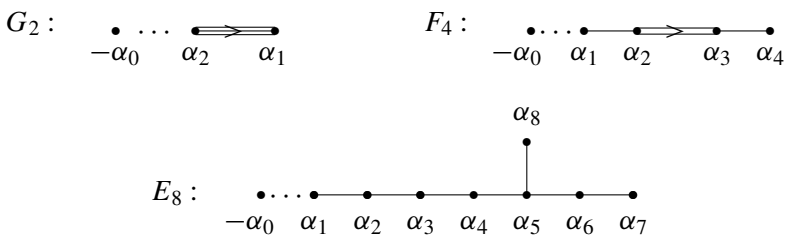
On note $G$ le groupe déployé de type $G_2$ (resp. $F_4$, $E_8$). Il admet un unique (à $G(k_s)$-conjugaison près) $k$-sous-groupe $A = \mu_l \times \mu_l \times \mathbb{Z}/l\mathbb{Z}$ avec $l = 2$ (resp. $l = 3$, $l = 5$) tel que $Z_G(A)$ est fini dont on va donner une description précise ci-dessous. On se propose de calculer le composé

$$H^1_{\mathrm{fppf}}(k, A) \to H^1_{\mathrm{fppf}}(k, G) \to H^3(k),$$

le cas de $G_2$ étant bien connu des spécialistes. La théorie de la cohomologie plate [Berhuy et al. 2007, appendice B] n'est nécessaire que dans le cas de caractéristique $l$. Notons que ce type de torseurs intervient dans l'étude de la dimension essentielle de $G$ [Reichstein et Youssin 2000; Chernousov et Serre 2006; Gille et Reichstein 2009].

Si $n$ est un entier strictement positif, on note $C_n$ le groupe cyclique d'ordre $n$ et $A_n$ le groupe alterné en $n$ lettres.

**6A.** ***Sous-groupes finis.*** Soit $T$ un tore déployé maximal de $G$ et $W = N_G(T)/T$ son groupe de Weyl. On considère les diagrammes de Dynkin étendus respectifs




où $\alpha_0$ désigne la plus grande racine. D'après [Borel et De Siebenthal 1949], ceci indique que $G$ admet un sous-groupe $H$ de type $A_1 \times A_1$ (resp. $A_2 \times A_2$, resp. $A_4 \times A_4$) dont le premier facteur contient le sommet $-\alpha_0$. On a $H = (\mathrm{SL}_l \times \mathrm{SL}_l)/\mu$ où $\mu \subset \mu_l \times \mu_l$ désigne le groupe fondamental de $H$.

**Lemme 6.1.** *Si $l$ est inversible dans $k$, alors $H = Z_G(Z(H))$.*

*Démonstration.* Le fait que $H = Z_G(Z(H))^0$ fait partie du théorème de Borel – de Siebenthal ; voir [Gille 2010]. Il faut montrer la connexité de $Z_G(Z(H))$ et on peut supposer $k$ algébriquement clos. Vu que $Z(H) = (\mu_l \times \mu_l)/\mu$ est cyclique et que $G$ est simplement connexe, on sait que $Z_G(Z(H))$ est connexe d'après [Springer et Steinberg 1970, §3.9]. □

De façon plus précise, on va décrire $H = H'.H''$ au moyen d'un couple de Killing $(T, B)$ de $G$. Celui-ci définit une base $\Delta$ du système de racines $\Phi(G, T)$ ; pour chaque $\alpha \in \Phi(G, T)$, on note $U_\alpha \subset G$ le sous-groupe radiciel associé. On

pose $H' = \langle U'_+, U'_- \rangle$, $H'' = \langle U''_+, U''_- \rangle$, où $U'_\pm$ et $U''_\pm$ sont définis selon le cas par

| | $G_2$ | $F_4$ | $E_8$ |
|---|---|---|---|
| $U'_\pm$ | $U_{\pm\alpha_0}$ | $\langle U_{\pm\alpha_0}, U_{\pm\alpha_1} \rangle$ | $\langle U_{\pm\alpha_0}, U_{\pm\alpha_1}, U_{\pm\alpha_2}, U_{\pm\alpha_3} \rangle$ |
| $U''_\pm$ | $U_{\pm\alpha_1}$ | $\langle U_{\pm\alpha_3}, U_{\pm\alpha_4} \rangle$ | $\langle U_{\pm\alpha_5}, U_{\pm\alpha_6}, U_{\pm\alpha_7}, U_{\pm\alpha_8} \rangle$ |

Ainsi le groupe $H$ est muni du couple de Killing $(T, T.U'_+.U''_+) = (T, H \cap B)$ définissant les diagrammes de Dynkin $\Delta'$ et $\Delta''$. On a $\Delta' = \{-\alpha_0\}$, resp. $\{-\alpha_0, \alpha_1\}$, resp. $\{-\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$ ; on note $\omega'_0$, etc. les poids fondamentaux de $H'$ et on identifie $H'$ à $\mathrm{SL}_2$ (resp. $\mathrm{SL}_3$, $\mathrm{SL}_5$) par la représentation de plus haut poids $\omega'_0$. De même, $\Delta'' = \{\alpha_1\}$, resp. $\{\alpha_3, \alpha_4\}$, resp. $\{\alpha_5, \alpha_6, \alpha_7, \alpha_8\}$ ; on note $\omega''_1$ (resp. $\omega''_3$, $\omega''_4$, etc.) les poids fondamentaux de $H''$ que l'on identifie à $\mathrm{SL}_2$ (resp. $\mathrm{SL}_3$, $\mathrm{SL}_5$) par la représentation de plus haut poids $\omega''_1$ (resp. $\omega''_3$, resp. $\omega''_8$). Nous allons déterminer le groupe $\mu$ en suivant la méthode de [Tits 1990, §1.7.1]. (Il est aussi possible d'utiliser la théorie des représentations, voir [Garibaldi 2009, p. 40] dans le cas $E_8$.)

**Lemme 6.2.** *Dans le cas de $G_2$ (resp. $F_4$), $\mu$ est le $\mu_2$ (resp. $\mu_3$) diagonal de $\mathrm{SL}_2 \times \mathrm{SL}_2$ (resp. $\mathrm{SL}_3 \times \mathrm{SL}_3$). Dans le cas de $E_8$, $\mu$ est le sous-groupe $\mu_5$ plongé de $\mu_5 \times \mu_5$ plongé par $x \mapsto (x, x^2)$.*

*Démonstration.* On note $T' = T \cap H'$, $T'' = T \cap H''$ et $\alpha_* = \alpha_2$ (resp. $\alpha_2$, resp. $\alpha_4$). Suivant [Tits 1990, §1.7.1], le dual $\hat{\mu}$ est le quotient de $\hat{T}' \oplus \hat{T}''$ par $\hat{T}$. De façon plus précise, on a le diagramme commutatif

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \hat{T} & \longrightarrow & \hat{T}' & \oplus & \hat{T}'' & \longrightarrow & \hat{\mu} & \longrightarrow & 0 \\ & & & & \downarrow & & \downarrow & & \| & & \\ & & & & \mathbb{Z}/l\mathbb{Z} & \oplus & \mathbb{Z}/l\mathbb{Z} & \longrightarrow & \hat{\mu} & \longrightarrow & 0 \end{array}$$

et $\hat{T}$ est le sous-réseau de $\hat{T}' \oplus \hat{T}''$ engendré par les $\alpha_0^\vee(\omega_\alpha)\omega'_0 - \omega'_\alpha$ pour $\alpha \in \Delta' \setminus \{\alpha_0\}$, $\alpha_0^\vee(\omega_{\alpha_*}) - \omega'_0$ et les $\alpha_0^\vee(\omega_\alpha) - \omega''_\alpha$ pour $\alpha \in \Delta''$. Le fait que le premier facteur $\mathbb{Z}/l\mathbb{Z}$ soit engendré par $\omega'_0$ et le second facteur $\mathbb{Z}/l\mathbb{Z}$ soit engendré par $\omega''_1$ (resp. $\omega''_3$, resp. $\omega''_8$) simplifie la présentation du groupe $\hat{\mu}$. Pour l'étude cas par cas, on se refère aux tables de [Bourbaki 1982] renumérotées dans le cas de $E_8$ par Tits [1990].

*Cas de $G_2$ :* $\hat{\mu}$ est le quotient du sous-groupe de $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ par la relation

$$\omega'_0 + \alpha_0^\vee(\omega_2)\omega''_2 = \omega'_0 + \frac{(\alpha_0, \omega_2)}{(\alpha_0, \alpha_0)}\omega''_2 = \omega'_0 - \omega''_2$$

puisque $\alpha_0 = -\omega_2$. Ainsi $\mu$ est le noyau de $\mu_2 \times \mu_2 \to \mu_2$, $(x, y) \mapsto xy^{-1}$.

*Cas de $F_4$ :* $\hat{\mu}$ est le quotient du sous-groupe de $\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ par la relation

$$\omega'_0 + \alpha_0^\vee(\omega_3)\omega''_3 = \omega'_0 + \frac{(\alpha_0, \omega_3)}{(\alpha_0, \alpha_0)}\omega''_3 = \omega'_0 - 2\omega''_3$$

puisque $\alpha_0 = -\omega_1$. Ainsi $\mu$ est le noyau de $\mu_3 \times \mu_3 \to \mu_3$, $(x, y) \mapsto xy^{-2}$.

*Cas de $E_8$ :* $\hat{\mu}$ est le quotient du sous-groupe de $\mathbb{Z}/5\mathbb{Z} \oplus \mathbb{Z}/5\mathbb{Z}$ par la relation $\omega_0' + \alpha_0^\vee(\omega_8)\omega_5'' = \omega_0' - 3\omega_8''$. Ainsi $\mu$ est le noyau de $\mu_5 \times \mu_5 \to \mu_5$, $(x, y) \to xy^{-3}$ (voir aussi [Garibaldi 2009, p. 40]). □

On note $h$ l'image respective de

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathrm{SL}_2(k) \times \mathrm{SL}_2(k),$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \in \mathrm{SL}_3(k) \times \mathrm{SL}_3(k),$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \in \mathrm{SL}_5(k) \times \mathrm{SL}_5(k)$$

dans $H(k)$. L'élément $h$ est d'ordre $l = 2$ (resp. 3, 5) et agit de façon anisotrope à travers son image $w \in N_H(T)/T$ (c'est-à-dire $T^w$ est fini). On pose $A = T^w \times \langle h \rangle$, il contient l'image $\mu_l^{(1)} = Z(H')$ du centre du premier facteur ainsi qu'un sous-groupe $\mu_l^{(2)}$ que l'on définit de la façon suivante. Pour $G_2$, c'est le sous-groupe de $H$

$$\mu_2 = \mu_4/\mu_2, \quad [x] \mapsto \begin{bmatrix} x & 0 \\ 0 & -x \end{bmatrix} \times \begin{bmatrix} x & 0 \\ 0 & -x \end{bmatrix}.$$

Pour $F_4$ et $E_6$, c'est l'image du plongement $\mu_l \to H$ donné respectivement par

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{bmatrix},$$

$$x \mapsto \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x^2 & 0 & 0 \\ 0 & 0 & 0 & x^3 & 0 \\ 0 & 0 & 0 & 0 & x^4 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x^2 & 0 & 0 & 0 \\ 0 & 0 & x^4 & 0 & 0 \\ 0 & 0 & 0 & x^6 & 0 \\ 0 & 0 & 0 & 0 & x^8 \end{bmatrix}.$$

**Lemme 6.3.** *Dans les trois cas précédents, si $l$ est inversible dans $k$, on a*

(1) $T^w = \mu_l^{(1)} \times \mu_l^{(2)}$ *et* $A = \mu_l^{(1)} \times \mu_l^{(2)} \times \mathbb{Z}/l\mathbb{Z}$.

(2) $Z_G(T^w) = T \times \mathbb{Z}/l\mathbb{Z}$ *et* $Z_G(A) = A$.

(3) $N_G(T^w) = N_G(T, T^w)$.

*Démonstration.* La première assertion est évidente puisque $T^w$ se calcule dans $H$.

(2) On a $H = Z_G(Z(H))$ suivant le lemme 6.1 et $Z(H) \subset T^w$, d'où $Z_G(T^w) = Z_H(T^w)$. Il suit que $Z_G(T^w) = T \times \mathbb{Z}/l\mathbb{Z}$ puisque $\mathbb{Z}/l\mathbb{Z} = Z_{S_l}(\mathbb{Z}/l\mathbb{Z})$. Vu que $T^w = T^A \subset A$, on en déduit $Z_G(A) = A$.

(3) Le $k$-groupe $N_G(T^w)$ normalise $T = Z_G(T^w)^0$. D'où on déduit $N_G(T^w) = N_G(T, T^w)$. □

On considère le morphisme

$$\phi : N_G(T, T^w) \to \mathrm{Aut}_{k-gp}(T^w) = \mathrm{Aut}_{k-gp}(\mu_l^{(1)} \times \mu_l^{(2)}) = \mathrm{GL}_2(\mathbb{F}_l).$$

Par ailleurs, on note $N_{\langle w\rangle}$ (resp. $N_w$) la préimage de $N_W(\langle w\rangle)$ (resp. $Z_W(w)$) dans $N_G(T)$.

**Lemme 6.4.** *On suppose que $k$ est de caractéristique $\neq l$.*

(1) *On a des suites exactes de $k$-groupes*

$$1 \to T \times \mathbb{Z}/l\mathbb{Z} \to N_G(T, T^w) \xrightarrow{\phi} \mathrm{Aut}(T^w) \to 1$$

*où* $\mathrm{Aut}(T^w) \cong \mathrm{GL}_2(\mathbb{Z}/l\mathbb{Z})$ *et*

$$1 \to A \to N_G(T, T^w) \cap Z_G(\mathbb{Z}/l\mathbb{Z}) \xrightarrow{\phi'} \mathrm{SL}_2(\mathbb{F}_l) \to 1,$$

*où $\phi'$ est la restriction de $\phi$ à $N_G(T, T^w) \cap Z_G(\mathbb{Z}/l\mathbb{Z})$.*

(2) *L'inclusion $N_G(T, T^w) \cap Z_G(\mathbb{Z}/l\mathbb{Z}) \subset N_w$ induit un isomorphisme*

$$\mathrm{SL}_2(\mathbb{F}_l) \xrightarrow{\sim} Z_W(w)/\langle w\rangle.$$

(3) *L'inclusion $N_G(T, T^w) \to N_{\langle w\rangle}$ est un isomorphisme et induit un isomorphisme* $\mathrm{GL}_2(\mathbb{F}_l) \xrightarrow{\sim} N_{\langle w\rangle}/\langle w\rangle$.

(4) *Le morphisme*

$$N_G(T, T^w)(k) \xrightarrow{\phi} \mathrm{GL}_2(\mathbb{F}_l).$$

*est surjectif.*

(5) *Le morphisme*

$$N_G(T, T^w)(k) \cap Z_G(\mathbb{Z}/l\mathbb{Z})(k) \xrightarrow{\phi'} \mathrm{SL}_2(\mathbb{F}_l).$$

*est surjectif.*

Pour établir (2), on va procéder cas par cas en utilisant la forme explicite du groupe fini $Z_W(w)$. Dans le cas de $G_2$, $W$ est le groupe diédral d'ordre 12 et $\langle w\rangle$ est central, ainsi $Z_W(w) = W$. Dans les deux autres cas, la forme explicite de $Z_W(w)$ peut être relevée des tables de [Carter 1972] ou plus rapidement du fait que $Z_W(w)$

est un groupe de réflexion complexe [Springer 1974]. On a $Z_W(w) = C_6.A_4$ dans le cas $F_4$ et $Z_W(w) = C_{10}.A_5$ dans le cas $E_8$.

*Démonstration du lemme 6.4.* (1) On a

$$\ker(\phi) = N_G(T, T^w) \cap Z_G(T^w) = N_G(T, T^w) \cap \left(T \times \mathbb{Z}/l\mathbb{Z}\right) = T \times \mathbb{Z}/l\mathbb{Z}$$

suivant le lemme 6.3(2). Pour la surjectivité de $\phi$, on peut supposer $k$ algébriquement clos. Suivant les tables de Griess [1991], puisque $A$ est non toral, on a $N_G(A) = A \rtimes \mathrm{SL}_3(\mathbb{F}_l)$. L'image de $N_G(T^w) \cap N_G(\mathbb{Z}/l\mathbb{Z}) = N_G(A, T^w, \mathbb{Z}/l\mathbb{Z})$ dans $\mathrm{SL}_3(\mathbb{F}_l)$ est le sous-groupe des éléments de la forme

$$\begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{bmatrix}$$

qui est isomorphe à $\mathrm{GL}_2(\mathbb{F}_l)$. Ainsi $\mathrm{Im}(\phi) \cong \mathrm{GL}_2(\mathbb{F}_l)$. La surjectivité de $\phi'$ résulte du même fait et le noyau de $\phi'$ est donné par $\ker(\phi') = \ker(\phi) \cap Z_G(\mathbb{Z}/l\mathbb{Z}) = (T \times \mathbb{Z}/l\mathbb{Z}) \cap Z_G(\mathbb{Z}/l\mathbb{Z}) = A$ suivant le lemme 6.3(2).

(2) Le morphisme $N_G(T, T^w) \cap Z_G(\mathbb{Z}/l\mathbb{Z}) \to N_w$ produit en effet un plongement $\mathrm{SL}_2(\mathbb{F}_l) \hookrightarrow Z_W(w)/\langle w \rangle$, qui est un isomorphisme puisque ces deux groupes ont même ordre.

(3) De même, le morphisme $N_G(T, T^w) \to N_{\langle w \rangle}$ produit un plongement

$$\mathrm{GL}_2(\mathbb{F}_l) \hookrightarrow N_{\langle w \rangle}/\langle w \rangle.$$

Comme l'on a une suite exacte

$$1 \to Z_w \to N_{\langle w \rangle} \to Aut(\langle w \rangle) = (\mathbb{Z}/l\mathbb{Z})^\times,$$

on déduit en comptant que $\mathrm{GL}_2(\mathbb{F}_l) \xrightarrow{\sim} N_{\langle w \rangle}/\langle w \rangle$. Ainsi

$$N_G(T, T^w)/T \xrightarrow{\sim} N_{\langle w \rangle}/T,$$

d'où l'on conclut que $N_G(T, T^w) \xrightarrow{\sim} N_{\langle w \rangle}$.

(4) Le morphisme $N_G(T)(k) \to W$ étant surjectif, le morphisme

$$N_{\langle w \rangle}(k) \to N_W(\langle w \rangle)$$

l'est également. En utilisant (3), on conclut que $N_G(T, T^w)(k) \longrightarrow \mathrm{GL}_2(\mathbb{F}_l)$ est surjectif.

(5) Si $r \in \mathrm{SL}_2(\mathbb{F}_l)$, le (4) produit un relevé $g \in N_G(T, T^w)(k)$ qui agit trivialement sur $\mathbb{Z}/l\mathbb{Z}$. □

**Remarque 6.5.** L'hypothèse sur la caractéristique est technique et peut probablement être levée dans les lemmes 6.1, 6.3 et 6.4.

**6B.** ***Calcul explicite.*** On a $H^1_{\text{fppf}}(k, A) = k^\times/(k^\times)^l \times k^\times/(k^\times)^l \times H^1(k, \mathbb{Z}/l\mathbb{Z})$, un élément est donc le produit de deux éléments $(a), (b) \in k^\times/(k^\times)^l$ et d'un caractère $\chi \in H^1(k, \mathbb{Z}/l\mathbb{Z})$.

**Théorème 6.6.** *Le composé*

$$r_G \circ i_k : H^1_{\text{fppf}}(k, A) \to H^1_{\text{fppf}}(k, G) \to H^3(k),$$

*applique* $(a) \times (b) \times \chi$ *sur* $-\chi \cup (a) \cup (b)$ *si* $l \in k^\times$ *et sur* $\chi \cup (a) \cup (b)$ *si* $k$ *est de caractéristique* $l$.

*Démonstration* (s'appuyant principalement sur [Chernousov 1994]). On note $f = r_G \circ i_k$. On suppose que $G = E_8$, les deux autres cas étant similaires. Comme au Section 4D, l'argument de relèvement en caractéristique nulle [Gille 2000, théorème 2] permet le cas échéant de supposer que $\text{car}(k) \neq l$. En outre, par restriction-corestriction, il est loisible de supposer que $k$ contient une racine cinquième de l'unité $\zeta$. On regarde d'abord la restriction au facteur

$$H^1(k, \mu_5^{(2)} \times \mathbb{Z}/5\mathbb{Z}) = k^\times/(k^\times)^5 \times H^1(k, \mathbb{Z}/5\mathbb{Z}).$$

Celle-ci définit un invariant cohomologique $f'$ de $\mu_5^{(2)} \times \mathbb{Z}/5\mathbb{Z}$ dans $H^3(k, \mu_5^{\otimes 2})$. La description de ces invariants est bien connue [Garibaldi 2009, proposition §2.1 et §6.7]. On sait alors qu'il existe (de façon unique !) des éléments $(c_0) \in k^\times/(k^\times)^5$, $\alpha_0 \in H^2(k, \mu_5^{(2)})$ et $\beta_0 \in H^2(k, \mu_5)$ tels que

$$f'\big((b) \times \chi\big) = (c_0) \cup (b) \cup \chi + \alpha_0 \cup \chi + (b) \cup \beta_0.$$

Vu que $\mu_5^{(2)} \subset T$ et que $H^1(k, T) = 0$, cet invariant est nul si $\chi = 0$, d'où $\beta_0 = 0$. De même, $\mathbb{Z}/l\mathbb{Z}$ se plonge dans un $k$-tore déployé, donc $\alpha_0 = 0$ et

$$f'\big((b) \times \chi\big) = (c_0) \cup (b) \cup \chi.$$

On note $z$ l'image dans $Z^1(k, H)$ du cocycle $(b) \times \chi$. Suivant [Gille et Szamuely 2006, §2.5, 4.7], on a

$${}_zH = \big(\mathrm{SL}_1(B) \times \mathrm{SL}_1(C)\big)/\mu$$

où $B = (\chi, b)$ et $C = (\chi, b^2)$ sont des $k$-algèbres cycliques de degré 5 (dans le cas de $F_4$, on a ${}_zH = \big(\mathrm{SL}_1(B) \times \mathrm{SL}_1(B)\big)/\mu$ avec $B = (\chi, b)$). On considère le diagramme commutatif [Gille 2000, lemme 7]

$$\begin{array}{ccccc} H^1(k, A) & \twoheadrightarrow & H^1(k, {}_zG) & \xrightarrow{r_{zG}} & H^3(k) \\ \downarrow ?+(b)\cup\chi & & \downarrow \tau_z & & \downarrow ?+f((b)\times\chi) \\ H^1(k, A) & \twoheadrightarrow & H^1(k, G) & \xrightarrow{r_G} & H^3(k) \end{array} \qquad (6\text{-}1)$$

où $\tau_z$ désigne la bijection de torsion. Vu que le facteur $\mu_5^{(1)}$ s'applique sur le centre de $\mathrm{SL}_1(B)$, il vient le diagramme commutatif

$$\begin{array}{ccccc} k^\times/(k^\times)^5 & \longrightarrow & k^\times/\mathrm{Nrd}(B^\times) & \xrightarrow{\sim} & H^1(k, \mathrm{SL}_1(B)) \\ \| & & & & \downarrow \\ k^\times/(k^\times)^5 = H^1(k, \mu_5^{(1)}) & \longrightarrow & H^1(k, A) & \longrightarrow & H^1(k, {}_zG). \end{array}$$

Nous affirmons que le composé $H^1(k, \mathrm{SL}_1(B)) \to H^1(k, {}_zG) \to H^3(k)$, la seconde flèche étant $r_{{}_zG}$, est l'invariant de Rost de $\mathrm{SL}_1(B)$. En effet, on sait par [Garibaldi et al. 2003, (9.11), p. 129] que ce composé est $d \times r_{\mathrm{SL}_1(B)}$ où $d$ est l'indice de Dynkin (ou multiplicateur de Rost) du morphisme $\mathrm{SL}_1(B) \to {}_zG$. Cet indice correspond au morphisme $\mathbb{Z} = H^1({}_zG, \mathscr{K}_2) \to H^1(\mathrm{SL}_1(B), \mathscr{K}_2) = \mathbb{Z}$ et est donc le même que celui du morphisme $\mathrm{SL}_5 \to G$. Pour voir que ce multiplicateur est 1, on utilise le morphisme

$$\alpha_0^\vee : \mathrm{SL}_2 \to \mathrm{SL}_5 \to G$$

attaché à la coracine (courte) $\alpha_0^\vee$ commune à $\mathrm{SL}_5$ et $G$. On considère le diagramme commutatif

$$\begin{array}{ccc} \mathbb{Z} = H^1_{\mathrm{Zar}}(G, \mathscr{K}_2) & \xrightarrow{(\alpha_0^\vee)^*} & H^1_{\mathrm{Zar}}(\mathrm{SL}_2, \mathscr{K}_2) = \mathbb{Z} \\ \times d \downarrow & & \| \\ \mathbb{Z} = H^1_{\mathrm{Zar}}(\mathrm{SL}_5, \mathscr{K}_2) & \xrightarrow{(\alpha_0^\vee)^*} & H^1_{\mathrm{Zar}}(\mathrm{SL}_2, \mathscr{K}_2) = \mathbb{Z}. \end{array}$$

Suivant [Garibaldi et al. 2003, (7.5), p. 121], les flèches horizontales sont $id_{\mathbb{Z}}$, d'où $d = 1$. Ainsi, l'image de $(a)$ par le composé

$$k^\times/(k^\times)^5 = H^1(k, \mu_5^{(1)}) \to H^1(k, A) \to H^1(k, {}_zG) \to H^3(k)$$

est $(a) \cup \chi \cup (b)$ suivant le théorème 1.1. En remontant avec le diagramme (6-1), il vient

$$f\big((a) \times (b) \times \chi\big) = (a) \cup \chi \cup (b) - (c_0) \cup (b) \cup \chi.$$

On utilise maintenant l'action du groupe $N_G(T^w)(k) \cap Z_G(\mathbb{Z}/l\mathbb{Z})(k)$ sur $G$ et $A$. Suivant le lemme 6.4(5), il existe $g \in N_G(T^w)(k) \cap Z_G(\mathbb{Z}/l\mathbb{Z})(k)$ qui agit sur $T^w$ suivant

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{F}_5).$$

Vu que $G(k)$ agit trivialement sur $H^1(k, G)$, $(a)$ et $(b)$ jouent donc des rôles équivalents dans la formule cherchée, d'où la trivialité de $(c_0)$ et de $\beta_0$. On conclut que

$$f\big((a) \times (b) \times \chi\big) = -(a) \cup (b) \cup \chi = -\chi \cup (a) \cup (b). \qquad \square$$

**6C.** ***Algèbres d'Albert.*** La partie modulo 3 de l'invariant de Rost des groupes de type $F_4$ définit un invariant des algèbres d'Albert. En application du théorème 6.6, nous montrons que c'est exactement l'invariant construit dans [Rost 1991], et décrit également dans [Petersson et Racine 1996].

**Corollaire 6.7.** *On suppose que* $\mathrm{car}(k) \neq 3$. *Soient* $A/k$ *une algèbre simple centrale de degré* 3 *et* $\lambda \in k^\times$. *On note* $J(A,\lambda)$ *l'algèbre d'Albert de "première construction de Tits" et* $[J(A,\lambda)] \in H^1(k, F_4)$ *sa classe d'isomorphisme. Alors la partie modulo* 3 *de son invariant de Rost est*

$$g_3([J(A,\lambda)]) = [A] \cup (\lambda) \in H^3(k, \mu_3^{\otimes 2}).$$

La preuve passe par le lemme suivant :

**Lemme 6.8.** *Soit* $G/k$ *un groupe semi-simple déployé. Soit* $T/k$ *un tore maximal déployé de* $G$.

(1) *Soient* $S_1, S_2$ *des* $k$*-sous-groupes de* $T$. *Alors* $S_1$ *et* $S_2$ *sont* $G(k_s)$*-conjugués si et seulement si* $S_1$ *et* $S_2$ *sont* $G(k)$*-conjugués.*

(2) *Soient* $H_1$, $H_2$ *des sous-groupes réductifs déployés de* $G$ *de rang maximal. Alors* $H_1$ *et* $H_2$ *sont* $G(k_s)$*-conjugués si et seulement si* $H_1$ *et* $H_2$ *sont* $G(k)$*-conjugués.*

Le (1) est bien connu dans le cas des sous-groupes finis constants d'ordre premier à la caractéristique [Serre 2010, 1.1.1] qui est d'ailleurs le cas d'application.

*Démonstration.* (1) Les $k$-groupes $G_1 = Z(S_1)^0$ et $G_2 = Z(S_2)^0$ sont réductifs et admettent $T$ comme tore maximal. Si $S_1$ et $S_2$ sont $G(k_s)$-conjugués, il en est de même de $G_1$ et $G_2$ et ils sont alors $N_G(T)(k_s)$-conjugués. On pose $W = N_G(T)/T$, c'est un $k$-groupe fini constant et $N_G(T)(k) \to W$ est surjectif. Ainsi quitte à conjuguer par un élément de $G(k)$, on est ramené au cas où $G_1$ et $G_2$ sont $T(k_s)$-conjugués, d'où $S_1 = S_2$.

(2) On suppose donc que $H_1$ et $H_2$ sont $G(k_s)$-conjugués. On peut supposer que $T \subset H_i$ pour $i = 1, 2$ suivant le théorème de conjugaison de Borel–Tits. On pose $S_i = Z(H_i) \subset T$. A conjugaison de $G(k)$ près, le (1) montre que $Z(H_1) = Z(H_2)$. Par application du théorème de Borel de Siebenthal, il vient $H_1 = H_2$. □

Ainsi les sous-groupes semi-simples de $F_4$ qui sont déployés de type $A_2^2$ sont conjugués sous $F_4(k)$. On rappelle la présentation de $F_4$ donnée dans [Garibaldi 2009, §7.2]. On note $M = M_3(k)$ et $V = M \times M \times M$ qui est muni de la forme cubique

$$f(X, Y, Z) = \det(X) + \det(Y) + \det(Z) - \mathrm{Tr}(XYZ).$$

Le sous-groupe de $\mathrm{GL}(V)$ qui fixe $(I_3, 0, 0)$ et la forme $f$ est le groupe $F_4$. L'action de $\mathrm{SL}_3 \times \mathrm{SL}_3$ sur $V$ suivant

$$(A, C).(X, Y, Z) := (AXA^{-1}, AYC^{-1}, CYA^{-1})$$

fixe $(I_3, 0, 0)$ et la forme $f$ ; son noyau est le $\mu_3$ diagonal. En d'autres mots, on dispose d'un plongement

$$(\mathrm{SL}_3 \times \mathrm{SL}_3)/\mu_3 \to F_4.$$

Ce groupe est $F_4(k)$-conjugué au groupe $H$ considéré au § 6A, on peut donc abusivement le noter $H$ aussi.

On va identifier le sous-groupe $\mathrm{PGL}_3 \times \mu_3$ de $F_4$ considéré dans [Garibaldi 2009, §7.2]. Le sous-groupe $\mathrm{PGL}_3$ est le groupe diagonal $\mathrm{SL}_3/\mu_3 \subset H = (\mathrm{SL}_3 \times \mathrm{SL}_3)/\mu_3$ et le $\mu_3$ noté ici $\mu_3^{(4)}$ agit selon

$$\zeta.(X, Y, Z) = (X, \zeta Y, \zeta^2 Z).$$

Ainsi $\mu_3^{(4)}$ est le sous-groupe $(\mathrm{SL}_3 \times \mathrm{SL}_3)/\mu_3$ donné par

$$\zeta \mapsto \begin{bmatrix} \zeta & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

c'est-à-dire le sous-groupe $\mu_3^{(1)}$. On dispose par ailleurs du sous-groupe $\mu_3^{(5)} \times \mathbb{Z}/3\mathbb{Z}$ de $\mathrm{PGL}_3$ donné par les générateurs

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

On observe que son image dans $H$ est le sous-groupe $\mu_3^{(2)} \times \mathbb{Z}/3\mathbb{Z}$. Ainsi

$$\mu_3^{(1)} \times \mu_3^{(2)} \times \mathbb{Z}/3\mathbb{Z} = \mu_3^{(4)} \times \mu_3^{(5)} \times \mathbb{Z}/3\mathbb{Z}.$$

*Démonstration de la corollaire 6.7.* Si $A$ est déployée, $J(A, \lambda)$ est triviale et la formule également. Si $A$ est un corps gauche on sait d'après Wedderburn que $A$ est cyclique, c'est-à-dire $A = (L/k, \sigma, b)$ avec les conventions de la proposition 4.1. Sa classe dans $H^1(k, \mathrm{PGL}_3)$ est l'image de $(\chi, (b))$ par la flèche

$$H^1(k, \mu_3) \times H^1(k, \mathbb{Z}/3\mathbb{Z}) \longrightarrow H^1(k, \mathrm{PGL}_3).$$

Or l'image de

$$\begin{array}{ccc} H^1(k, \mu_3^{(1)}) \times H^1(k, \mu_3^{(2)}) \times H^1(k, \mathbb{Z}/3\mathbb{Z}) & \longrightarrow & H^1(k, \mu_3^{(1)}) \times H^1(k, \mathrm{PGL}_3) \\ & & \downarrow \\ & & H^1(k, F_4) \end{array}$$

applique $(a) \times (b) \times \chi$ sur $J(A, a)$ d'après [Knus et al. 1998, corollaire 39.9]. Suivant le théorème 6.6, le composé

$$H^1(k, \mu_3^{(1)}) \times H^1(k, \mu_3^{(2)}) \times H^1(k, \mathbb{Z}/3\mathbb{Z}) \longrightarrow H^1(k, F_4) \xrightarrow{r} H^3(k)$$

applique $(a) \times (b) \times \chi$ sur $-\chi \cup (a) \cup (b)$, ce qui permet de conclure que

$$g_3(J(A, a)) = -\chi \cup (a) \cup (b) = [A] \cup (a). \qquad \square$$

## Remerciements

Nous remercions Skip Garibaldi et le rapporteur pour les remarques ayant permis d'améliorer cet article, ainsi que R. Parimala qui a suggéré le corollaire 6.7. Le second auteur remercie L. Breen et B. Oliver pour les échanges enrichissants qu'elle a eu avec eux.

## Bibliographie

[Berhuy et al. 2007] G. Berhuy, C. Frings et J.-P. Tignol, "Galois cohomology of the classical groups over imperfect fields", *J. Pure Appl. Algebra* **211**:2 (2007), 307–341. MR 2009f:12004 Zbl 1121.11035

[Bloch et Kato 1986] S. Bloch et K. Kato, "$p$-adic étale cohomology", *Inst. Hautes Études Sci. Publ. Math.* 63 (1986), 107–152. MR 87k:14018 Zbl 0613.14017

[Borel et De Siebenthal 1949] A. Borel et J. De Siebenthal, "Les sous-groupes fermés de rang maximum des groupes de Lie clos", *Comment. Math. Helv.* **23** (1949), 200–221. MR 11,326d Zbl 0034.30701

[Bourbaki 1982] N. Bourbaki, *Groupes et algèbres de Lie, chapitre 9: groupes de Lie réels compacts*, Masson, Paris, 1982. MR 84i:22001 Zbl 0505.22006

[Brown 1982] K. S. Brown, *Cohomology of groups*, Graduate Texts in Mathematics **87**, Springer, New York, 1982. MR 83k:20002 Zbl 0584.20036

[Brylinski et Deligne 2001] J.-L. Brylinski et P. Deligne, "Central extensions of reductive groups by $\mathbf{K}_2$", *Publ. Math. Inst. Hautes Études Sci.* **94** (2001), 5–85. MR 2004a:20049 Zbl 1093.20027

[Carter 1972] R. W. Carter, "Conjugacy classes in the Weyl group", *Compositio Math.* **25** (1972), 1–59. MR 47 #6884 Zbl 0254.17005

[Chernousov 1994] V. I. Chernousov, "A remark on the (mod 5)-invariant of Serre for groups of type $E_8$", *Mat. Zametki* **56**:1 (1994), 116–121, 157. MR 95i:20072 Zbl 0835.20059

[Chernousov et Serre 2006] V. Chernousov et J.-P. Serre, "Lower bounds for essential dimensions via orthogonal representations", *J. Algebra* **305**:2 (2006), 1055–1070. MR 2007i:20070 Zbl 1181.20042

[Colliot-Thélène et Raskind 1985] J.-L. Colliot-Thélène et W. Raskind, "$\mathcal{K}_2$-cohomology and the second Chow group", *Math. Ann.* **270**:2 (1985), 165–199. MR 86m:14005 Zbl 0536.14004

[Esnault et al. 1998] H. Esnault, B. Kahn, M. Levine et 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

[Garibaldi 2009] S. Garibaldi, *Cohomological invariants: exceptional groups and spin groups*, Mem. Amer. Math. Soc. **937**, American Mathematical Society, Providence, 2009. MR 2010g:20079 Zbl 1191.11009

[Garibaldi et Quéguiner-Mathieu 2007] S. Garibaldi et A. Quéguiner-Mathieu, "Restricting the Rost invariant to the center", *Algebra i Analiz* **19**:2 (2007), 52–73. Republié dans *St. Petersburg Math. J.* **19**:2 (2008), 197–213. MR 2008f:12009 Zbl 05342956

[Garibaldi et al. 2003] S. Garibaldi, A. Merkurjev et J.-P. Serre, *Cohomological invariants in Galois cohomology*, Univ. Lecture Ser. **28**, Amer. Math. Soc., Providence, RI, 2003. MR 1999383 Zbl 1159.12311

[Gille 2000] P. Gille, "Invariants cohomologiques de Rost en caractéristique positive", *K-Theory* **21**:1 (2000), 57–100. MR 2001k:11064 Zbl 0993.20031

[Gille 2010] P. Gille, "The Borel-de Siebenthal's theorem", 2010, http://www.math.ens.fr/~gille/prenotes/bds.pdf.

[Gille et Reichstein 2009] P. Gille et Z. Reichstein, "A lower bound on the essential dimension of a connected linear group", *Comment. Math. Helv.* **84**:1 (2009), 189–212. MR 2009j:11066 Zbl 1173.11022

[Gille et Szamuely 2006] P. Gille et T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics **101**, Cambridge University Press, Cambridge, 2006. MR 2007k:16033 Zbl 1137.12001

[Griess 1991] R. L. Griess, Jr., "Elementary abelian $p$-subgroups of algebraic groups", *Geom. Dedicata* **39**:3 (1991), 253–305. MR 92i:20047 Zbl 0733.20023

[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

[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

[Kato 1982] K. Kato, "Galois cohomology of complete discrete valuation fields", pp. 215–238 dans *Algebraic K-theory* (Oberwolfach, 1980), vol. 2, édité par R. K. Dennis, Lecture Notes in Math. **967**, Springer, Berlin, 1982. MR 84k:12006 Zbl 0506.12022

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost et J.-P. Tignol, *The book of involutions*, Colloquium Publications **44**, American Mathematical Society, Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Mac Lane 1963] S. Mac Lane, *Homology*, Grundlehren der math. Wiss. **114**, Springer, Berlin, 1963.

[Matsumoto 1969] H. Matsumoto, "Sur les sous-groupes arithmétiques des groupes semi-simples déployés", *Ann. Sci. École Norm. Sup.* (4) **2** (1969), 1–62. MR 39 #1566 Zbl 0261.20025

[Merkurjev et Suslin 1982] A. S. Merkurjev et A. A. Suslin, "$\mathcal{K}$-cohomologie des variétés de Severi-Brauer et l'homomorphisme de norme résiduelle", *Izv. Akad. Nauk SSSR Ser. Mat.* **46**:5 (1982), 1011–1046. En russe; traduction anglaise: *Math. USSR Izv.* **21** (1983), 307–340. MR 84i:12007

[Merkurjev et al. 2002] A. S. Merkurjev, R. Parimala et J.-P. Tignol, "Invariants of quasitrivial tori and the Rost invariant", *Algebra i Analiz* **14**:5 (2002), 110–151. Republié dans *St. Petersburg Math. J.* **14**:5 (2003), 791–821. MR 2004c:11045 Zbl 1041.11023

[Petersson et Racine 1996] H. P. Petersson et M. L. Racine, "An elementary approach to the Serre–Rost invariant of Albert algebras", *Indag. Math.* (*N.S.*) **7**:3 (1996), 343–365. MR 99j:17045 Zbl 0872.17029

[Reichstein et Youssin 2000] Z. Reichstein et B. Youssin, "Essential dimensions of algebraic groups and a resolution theorem for $G$-varieties", *Canad. J. Math.* **52**:5 (2000), 1018–1056. MR2001k:2001k:14088 Zbl 1044.14023

[Rost 1991] M. Rost, "A (mod 3) invariant for exceptional Jordan algebras", *C. R. Acad. Sci. Paris Sér. I Math.* **313**:12 (1991), 823–827. MR 92j:19002 Zbl 0756.17014

[Serre 1994] J.-P. Serre, *Cohomologie galoisienne*, 5ème éd., Lecture Notes in Mathematics **5**, Springer, Berlin, 1994. MR 96b:12010 Zbl 0812.12002

[Serre 2010] J.-P. Serre, "Le groupe de Cremona et ses sous-groupes finis", pp. 75–100 = exposé 1000 dans *Séminaire Bourbaki* 2008/2009, Astérisque **332**, Soc. Math. de France, Paris, 2010. MR 2648675

[Springer 1974] T. A. Springer, "Regular elements of finite reflection groups", *Invent. Math.* **25** (1974), 159–198. MR 50 #7371 Zbl 0287.20043

[Springer et Steinberg 1970] T. A. Springer et R. Steinberg, "Conjugacy classes", pp. 167–266 dans *Seminar on Algebraic Groups and Related Finite Groups* (Princeton, 1968/69), édité par A. Borel et al., Lecture Notes in Mathematics **131**, Springer, Berlin, 1970. MR 42 #3091

[Tits 1971] J. Tits, "Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque", *J. Reine Angew. Math.* **247** (1971), 196–220. MR 43 #3269 Zbl 0227.20015

[Tits 1990] J. Tits, "Strongly inner anisotropic forms of simple algebraic groups", *J. Algebra* **131**:2 (1990), 648–677. MR 91g:20069 Zbl 0697.20029

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Adv. Math. **38**, Cambridge University Press, Cambridge, 1994. MR 95f:18001 Zbl 0797.18001

Communicated by Michel van den Bergh
Received 2009-07-22 Revised 2010-05-10 Accepted 2010-05-29

Philippe.Gille@ens.fr *UMR 8552 du CNRS, Département de Mathématiques et Applications, École Normale Supérieure, F-75005 Paris, France*
http://www.math.ens.fr/~gille/

queguin@math.univ-paris13.fr *Université Paris 13, LAGA, UMR CNRS 7539, et Université Paris 12, Institut Galilée, Avenue J.-B. Clément, 93430 Villetaneuse, France*
http://www-math.univ-paris13.fr/~queguin/

ALGEBRA AND NUMBER THEORY 5:1(2011)

msp

# Modular abelian varieties of odd modular degree

Soroosh Yazdani

We study modular abelian varieties with odd congruence number by examining the cuspidal subgroup of $J_0(N)$. We show that the conductor of such abelian varieties must be of a special type. For example, if $N$ is the conductor of an absolutely simple modular abelian variety with odd congruence number, then $N$ has at most two prime divisors, and if $N$ is odd, then $N = p^\alpha$ or $N = pq$ for some primes $p$ and $q$. In the second half of the paper, we focus on modular elliptic curves with odd modular degree. Our results, combined with the work of Agashe, Ribet, and Stein for elliptic curves to have odd modular degree. In the process we prove Watkins' conjecture for elliptic curves with odd modular degree and a nontrivial rational torsion point.

Let $E/\mathbb{Q}$ be an elliptic curve over the rational numbers. From [Wiles 1995; Taylor and Wiles 1995], we know that $E$ is modular (see [Breuil et al. 2001]), which implies that there is a surjective map $\pi : X_0(N) \to E$ defined over the rationals. As such, we have a new invariant attached to the elliptic curve, namely the minimal degree of $\pi$, which we call the *modular degree* of $E$. This invariant is related to many other invariants of the elliptic curve. For instance, this number is closely related to the congruences between $E$ and other modular forms (see Section 1A and [Agashe et al. 2008]). Also, we know that finding a good bound on the degree of $\pi$ in terms of $N$ is equivalent to the $ABC$ conjecture (see [Murty 1999; Frey 1997]).

After calculating the modular degree of various elliptic curves, Watkins conjectured that $2^r$ divides the modular degree of the elliptic curve $E$, where $r$ is the rank of $E(\mathbb{Q})$ (see [Watkins 2002]). In the particular case when the modular degree of $E$ is odd, Watkins' conjecture implies $E(\mathbb{Q})$ is finite. Searching through Cremona, Stein, and Watkins' databases [Stein and Watkins 2002; Cremona] for elliptic curves of odd modular degree, Calegari and Emerton [2009] observed that all such elliptic curves have bad reduction at no more than two primes. By studying the Atkin–Lehner involution on elliptic curves $E$ having odd modular degree, they

This research was partially supported by NSERC..

*MSC2000:* primary 11F33; secondary 11G05.

*Keywords:* modular form, modular curve, elliptic curve, congruence number.

demonstrated that such curves have an even analytic rank and that there are at most two odd primes dividing their conductor (see Section 2A). Dummigan [2006] has provided a heuristic explanation for Watkins' conjecture. His method uses the Selmer group of the symmetric square of $E$ and its relationship to congruences between modular forms.

The goal of this paper is to extend the results of Calegari and Emerton to modular abelian varieties having odd modular exponents and odd congruence number (see Section 1A for definition). We find the necessary conditions for a modular abelian variety to have odd congruence number. Specifically in Theorem 2.15 we show that if a modular abelian variety with conductor $N$ has odd congruence number, then $N = 2p, 4p^a, 8p^a, pq$ where $p$ and $q$ are odd primes, or $N$ is a power of a prime. In Section 3 we study elliptic curves having odd congruence number. Recall the result of Agashe, Ribet, and Stein [Agashe et al. 2008] that elliptic curves with semistable reduction at 2 have odd congruence number if and only if they have odd modular degree (see Theorem 1.1).[1] We find more stringent conditions that elliptic curves with odd congruence number need to satisfy. Specifically in Theorem 3.8 we show that if an elliptic curve $E$ with conductor $N$ has odd congruence number, then if it has a trivial torsion structure then $N$ is prime and $E$ has an even analytic rank, otherwise $N$ has at most two prime divisors and has rank 0. Furthermore, we find families of elliptic curves that any elliptic curve with odd congruence number and a nontrivial torsion point must belong to one of these families (see Theorem 3.8). We expect that the elliptic curves in these families have odd modular degrees, although to prove this we need a better understanding of the rational torsion points of $J_0(N)$.

We now give a quick overview. In Section 1, we review some definitions and results needed in the rest of the paper. Specifically, in Section 1B we recall how to calculate the rational cuspidal subgroup of $J_0(N)$, and in Section 1C we study the action of the Hecke algebra and Atkin–Lehner involutions on this subgroup. In Section 2, we study modular abelian varieties with odd congruence number, and show that all such abelian varieties have at most two primes of bad reduction. A key component of this argument is that if $A$ is a modular abelian variety having odd congruence number and non-prime-power conductor, $A$ must have a rational 2-torsion point (Theorem 2.1). We also show that if $A$ has odd congruence number and a rational 2-torsion point, then all the new rational 2-torsion points of $J_0(N)$ map injectively to $A$ (see Section 2C). We use this fact and our analysis of the cuspidal subgroup to show that if $A$ has odd congruence number and is semistable away from 2, then it has at most two primes of bad reduction (Theorem 2.12) and the primes dividing the conductor must satisfy certain congruences (Theorem 2.15). The other useful result is that if $p^2|N$ for some odd prime $N$, then $A$ must have a

[1] In fact, by searching through Cremona's table of elliptic curves, it seems that an elliptic curve has odd congruence number if and only if it has odd modular degree.

complex multiplication or an inner twist (Proposition 2.10). In Section 3 we apply our results to elliptic curves. Theorem 2.15 gives us conditions that the conductor of an elliptic curve with odd congruence number must satisfy. In each subsection of Section 3 we study one of these cases, get more stringent conditions on the conductor, and show that in almost all cases the rank of such elliptic curves is zero (Theorem 3.8).

## 1. Preliminaries

Let $N$ be a positive integer and $X_0(N)$ be the moduli space of generalized elliptic curves with a cyclic subgroup of order $N$. Let $\mathscr{C}_N \subset X_0(N)$ be the set of cusps of $X_0(N)$, that is, $\mathscr{C}_N = \pi^{-1}(\infty)$, where $\pi : X_0(N) \to X_0(1)$ is the natural degeneracy map, and $\infty$ is the unique cusp on $X_0(1)$. All such cusps can be represented as rational numbers $a/b \in \mathbb{H}$, with $a$ and $b$ positive coprime integers and $b|N$. Furthermore, there is a unique representative for any cusp with $a \le (b, N/b)$. Under this representation, $\infty = 1/N$. For any divisor $r$ of $N$ with $\gcd(r, N/r) = 1$, we can define the Atkin–Lehner involution $w_r : X_0(N) \to X_0(N)$ by sending $(E, D)$ to $(E/D[r], (E[r] + D)/D[r])$[2]. We usually abuse notation by letting $w_{\bar{r}} = w_r$ whenever $\bar{r} = \prod_{l|r} l$ (for example, $w_4 = w_2$ on $X_0(4N)$).

Let $S(N)$ be the space of weight two cuspforms on $\Gamma_0(N)$. Let $\mathbb{T}$ denote the $\mathbb{Z}$-algebra of the Hecke operators acting on $S(N)$. As usual, we denote $J_0(N) = \mathrm{Jac}(X_0(N))$. Then $\mathbb{T}$ acts faithfully on $J_0(N)$ by Picard functoriality. We also have the standard Albanese embedding $i : X_0(N) \to J_0(N)$ via $i(z) = (z) - (\infty)$. Note that for any map $w : X_0(N) \to X_0(N)$ we have the induced map

$$\begin{aligned} w_* : J_0(N) &\to J_0(N), \\ \sum (z) &\mapsto \sum (w(z)). \end{aligned}$$

**1A.** ***Congruence numbers.*** Recall that attached to any newform $f \in S(N)$ we have a modular abelian variety $A_f$. Specifically, let $I_f$ be the kernel of $\mathbb{T} \to \mathbb{C}$ induced by $f$. Then we have $A_f = J_0(N)/I_f$, which we refer to as the *optimal quotient* attached to $f$. Conversely, if $A$ is a simple quotient of $J_0(N)$ that is stable under the action of $\mathbb{T}$ and the Atkin–Lehner involutions, then we can find a modular eigenform $f \in S(N)$ such that $A$ is isogenous to $A_f$. In this case, we say that $f$ is attached to $A$. Furthermore all modular forms attached to $A$ are Galois conjugate to $f$. Let $\phi : J_0(N) \to A$ be a surjective morphism. Then the dual morphism is $\phi^\vee : A^\vee \to J_0(N)^\vee$. Since $J_0(N)$ is self-dual, we can compose these two morphisms to get

$$\psi : A^\vee \to A.$$

[2] As usual, $G[r]$ is the set of $r$-torsion points of the group $G$.

Following [Agashe et al. 2008], we define the *modular number* to be the order of $\ker(\psi)$, and the *modular exponent* to be its exponent, denoted by $\widetilde{n}_A$. If $A$ is an elliptic curve, then $\widetilde{n}_A$ equals the modular degree of $A$. In fact, in the case of elliptic curves we get that $\ker(\psi) = A[\deg(\pi)]$, where $\pi : X_0(N) \to A$ (see Lemma 2.2).

Now let $\phi : J_0(N) \to A$ be any optimal modular abelian quotient. Let $B = \ker(\phi)$, which is an abelian variety since $A$ is an optimal quotient. Let $\mathbb{T}_A$ be the $\mathbb{Z}$-algebra of the Hecke operators acting on $A$. Similarly, let $\mathbb{T}_B$ be the $\mathbb{Z}$-algebra of the Hecke operators acting on $B$. There is an injective map $\mathbb{T} \to \mathbb{T}_A \oplus \mathbb{T}_B$ with a finite index, given by the restriction map. The order of the cokernel of $\mathbb{T} \to \mathbb{T}_A \oplus \mathbb{T}_B$ is the *congruence number* of $A$. The exponent of this cokernel is the *congruence exponent* of $A$, which is denoted by $\widetilde{r}_A$ (see [Agashe et al. 2008, Lemma 4.3]). Let $\mathfrak{m} \subset \mathbb{T}$ be a maximal ideal of $\mathbb{T}$. Then $A[\mathfrak{m}] \neq 0$ (resp. $B[\mathfrak{m}] \neq 0$) if and only if the image of $\mathfrak{m}$ in $\mathbb{T}_A$ (resp. $\mathbb{T}_B$) is a proper maximal ideal. If $A[\mathfrak{m}]$ and $B[\mathfrak{m}]$ are both nontrivial, then by tensoring $\mathbb{T} \to \mathbb{T}_A \oplus \mathbb{T}_B$ by $\mathbb{T}/\mathfrak{m}$, we see that the cokernel is a nontrivial vector space over $\mathbb{T}/\mathfrak{m}$, which means that the characteristic of $\mathbb{T}/\mathfrak{m}$ divides the congruence exponent of $A$. On the other hand, if $A[\mathfrak{m}] \neq 0$, then $A^\vee[\mathfrak{m}] \neq 0$, and if $A^\vee[\mathfrak{m}] \cap B[\mathfrak{m}] \neq 0$, the characteristic of $\mathbb{T}/\mathfrak{m}$ divides the modular exponent.

In [Agashe et al. 2008], the relationship between the modular exponent and the congruence exponent was studied, and the following was proved.

**Theorem 1.1.** *If $f \in S(N)$ is a newform, then $\widetilde{n}_{A_f} \mid \widetilde{r}_{A_f}$. If, moreover, $p^2 \nmid N$, then* $\operatorname{ord}_p(\widetilde{n}_{A_f}) = \operatorname{ord}_p(\widetilde{r}_{A_f})$.

In particular, if $f$ is a newform of level $N$ and $4 \nmid N$, then the modular exponent of $A_f$ is odd if and only if its congruence exponent is odd.

**1B.** ***Cuspidal subgroup.*** The cuspidal subgroup of $J_0(N)$ is the subgroup generated by the cusps of $X_0(N)$. The goal of this section is to understand the rational points of the cuspidal subgroup of $J_0(N)$, denoted by $C_N$. This problem was studied for $N$ a power of a prime in [Ling 1997] and for $N$ the product of the two primes in [Chua and Ling 1997]. Set

$$P_d = \frac{1}{\gcd(d, N/d)} \sum_{i=1}^{\gcd(d,N/d)} (id/N)^3.$$

**Proposition 1.2** [Ling 1997]. *The rational cuspidal subgroup $C_N \subset J_0(N)$ is generated by the elements $\phi(\gcd(d, N/d))(P_d - P_1)$.*

Here and in Proposition 1.4, $\phi$ denotes the Euler totient function. However,

[3] Our notation is slightly different from [Ling 1997; Chua and Ling 1997]. Specifically their $P_d$ is $\gcd(d, N/d)$ times our $P_d$.

*outside this section, $\phi$ is reserved for the map $\phi : J_0(N) \to A$*, and whenever results from this section are used, we will have $\phi(\gcd(d, N/d)) = 1$.

We will calculate the order of certain elements in $C_N$. Recall the Dedekind eta function, defined as

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n), \quad \text{with } q = e^{2\pi i \tau}.$$

Let $\eta(M\tau) = \eta_M(\tau)$. We use $\eta_M$ to construct functions with divisors supported on the cusps. In particular, for $M|N$, $\eta_M$ has a zero of order

$$\frac{1}{24} \frac{N d'^2}{d t M}, \tag{1}$$

at the cusp of $X_0(N)$ corresponding to $x/d \in \mathbb{H}$, where $d' = \gcd(d, M)$ and $t = \gcd(d, N/d)$; see, for example, [Ogg 1974].

**Proposition 1.3** [Ligozat 1975]. *Let $\boldsymbol{r} = (r_\delta)$ be a collection of rational numbers $r_\delta \in \mathbb{Q}$ indexed by all the positive divisors of $\delta|N$. Then the function $g_{\boldsymbol{r}} = \prod_{\delta|N} \eta_\delta^{r_\delta}$ is a modular function on $X_0(N)$ if and only if the following conditions are satisfied*:

(1) *All the rational numbers $r_\delta$ are rational integers.*

(2) $\sum_{\delta|N} r_\delta \delta \equiv 0 \pmod{24}$.

(3) $\sum_{\delta|N} r_\delta N/\delta \equiv 0 \pmod{24}$.

(4) $\sum_{\delta|N} r_\delta = 0$.

(5) $\prod_{\delta|N} \delta^{r_\delta}$ *is a square of a rational number.*

We also know that the lattice of divisors linearly equivalent to zero supported on the cusps is generated by the divisors of $g_{\mathbf{r}}$ that are modular functions. Let $N = \prod_{i=1}^{k} p_i^{s_i}$ be the prime factorization of $N$, and let $V$ be the rational vector space spanned by $P_d$ for $d|N$. We can represent this vector space as the tensor product of the vector spaces $V_{p_i}$, where $V_{p_i}$ is the $(s_i + 1)$-dimensional space generated by $P_1, P_{p_i}, \ldots, P_{p_i^{s_i}}$. (The isomorphism between $V$ and the tensor product $\bigotimes_i V_{p_i}$ is the natural one sending $P_{\prod p_i^{\alpha_i}}$ to $P_{p_1^{\alpha_1}} \otimes \cdots \otimes P_{p_k^{\alpha_k}}$.) Similarly, let $W$ be the rational vector space of functions $g_{\boldsymbol{r}}$ (as defined in Proposition 1.3) under multiplication. Then we have $W \simeq \bigotimes W_{p_i}$ where $W_{p_i}$ is the $(s_i + 1)$-dimensional vector space generated by $\eta_1, \eta_{p_i}, \ldots, \eta_{p_i^{s_i}}$. We have an isomorphism $\Lambda : V \to W$ where $\Lambda^{-1}(g)$ is the divisor attached to $g$. We can verify that this isomorphism can be written very explicitly as

$$24 \Lambda_{p_1} \otimes \cdots \otimes \Lambda_{p_k},$$

where $\Lambda_{p_i} : V_{p_i} \to W_{p_i}$ is expressed in the basis above as the tridiagonal matrix

$$\Lambda_{p_i} = \frac{1}{(p_i^2-1)\phi(p_i^{s_i})} \begin{pmatrix} p_i(p_i-1) & -p_i & & & & & \\ -(p_i-1) & p_i^2+1 & -p_i & & & & \\ & -p_i & p_i^2+1 & -p_i & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -p_i & p_i^2+1 & -(p_i-1) & \\ & & & & -p_i & p_i(p_i-1) \end{pmatrix}.$$

When $f \in W$ is a modular function, $\Lambda^{-1}(f)$ is linearly equivalent to zero. Therefore, by combining Proposition 1.3 and the isomorphism above we get:

**Proposition 1.4.** *An element $v \in \bigotimes V_{p_i} = V$ is linearly equivalent to zero if the following conditions are satisfied*:

(1) *All the coefficients in $\Lambda v$ are integral.*

(2) *$v$ has degree 0.*

(3) *$v$ is integral and the coefficient of $P_d$ divides $\phi(d, N/d)$.*

(4) *For each $i$, the number $(e_1 \otimes \cdots \otimes f_i \otimes \cdots \otimes e_k)\Lambda v$ is even, where $e_i = (1, 1, 1, \ldots, 1) \in W_{p_i}^\vee$ and $f_i = (0, 1, 0, 1, \ldots) \in W_{p_i}^\vee$.*

We use Proposition 1.4 to calculate the order of the elements in $C_N$. Specifically, for an integral element $v \in V$ of degree zero, the order of $v$ in $C_N$ is the smallest positive integer $n$ such that $nv$ satisfies all the conditions in Proposition 1.4. Notice that if $N = 2^{s_2}M$ where $M$ is a square-free odd integer and $s_2 < 4$ (the case we deal with in this paper), then condition (3) is reduced to the coefficients of $v$ being integral. Therefore, the denominator of $\Lambda(v)$ gives the order of $v$ or half of it.

We can use Proposition 1.4 to calculate the order of elements of $C_N$ considered in Table 1. For example, consider an element $v = v_{p_1} \otimes \cdots \otimes v_{p_k}$, where $v_{p_i} = P_1 + b_i P_{p_i}$, representing an element in $C_N$ when $N$ is square-free and not a prime. This is a generalization of [Ogg 1974] in the case where $N = pq$. Note that

$$\begin{aligned} \Lambda v &= \frac{24}{\prod_i (p_i^2-1)(p_i-1)} (\Lambda_{p_1} v_{p_1} \otimes \cdots \otimes \Lambda_{p_k} v_{p_k}) \\ &= \frac{24}{\prod_i (p_i^2-1)} \left( \begin{pmatrix} p_1 - b_1 \\ (p_1-b_1)b_1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} p_k - b_k \\ (p_k-b_k)b_k \end{pmatrix} \right) \\ &= \frac{24}{\prod_i (p_i+b_i)} \left( \begin{pmatrix} 1 \\ b_1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 \\ b_k \end{pmatrix} \right). \end{aligned}$$

Considering the coefficient of the first coordinate, the order is divisible by $n$, the numerator of the (reduced) fraction $\prod_i (p_i + b_i)/24$. On the other hand, $\Lambda(nv)$

| $N$ | Cuspidal element | Order |
|---|---|---|
| $p$ | $P_1 - P_p$ | $\mathrm{Num}\left(\frac{p-1}{12}\right)$ |
| $\prod_{i=1}^{t} p_i$ | $(P_1 + b_1 P_{p_1}) \otimes \cdots \otimes (P_1 + b_k P_{p_k})$ | $\mathrm{Num}\left(\frac{\prod_i (p_i + b_i)}{24}\right)$ |
| Conditions: $t > 1$; $b_i = \pm 1$ for $i = 1, 2, \ldots, t$; $b_j = -1$ for at least one $j$. | | |
| $4p$ | $P_2 - P_{2p}$ | $\frac{p-1}{2}$ |
| Conditions: $p$ is odd. | | |
| $4\prod_{i=1}^{t} p_i$ | $P_2 \otimes (P_1 + b_1 P_{p_1}) \otimes \cdots \otimes (P_1 + b_k P_{p_k})$ | $\left(\frac{\prod_i (p_i + b_i)}{4}\right)$ |
| Conditions: $t > 1$; all the $p_i$ are odd; $b_i = \pm 1$ for $i = 1, \ldots, t$; $b_j = -1$ for some $j$. | | |
| $8\prod_{i=1}^{t} p_i$ | $(P_1 - P_8) \otimes (P_1 + b_1 P_{p_1}) \otimes \cdots \otimes (P_1 + b_k P_{p_k})$ | $\frac{\prod_i p_i + b_i}{2}$ |
| Conditions: all the $p_i$ are odd. | | |

**Table 1.** Order of elements in $C_N$. "Num" is the numerator of a reduced fraction.

equals $\binom{1}{b_1} \otimes \cdots \otimes \binom{1}{b_k}$. Therefore

$$(e_1 \otimes \cdots \otimes f_i \otimes \cdots \otimes e_t)\Lambda(nv)$$

is even, implying that $nv$ is linearly equivalent to zero. Hence $v$ has order $n$. The other entries in Table 1 are calculated the same way.

**1C.** ***Hecke action.*** In this section we recall the explicit action of the Hecke operators $T_l$ on the rational cuspidal divisors of $X_0(N)$. This is fairly standard, although the representation of these actions as the tensor product of matrices is not that common. The following is the main result of this section.

**Proposition 1.5.** (1) *Let $p \nmid N$. Then $T_p : V \to V$ acts as multiplication by $p+1$.*

(2) *Let $p|N$ and $V = \bigotimes V_{p_i}$. Then $T_p$ acts trivially on $V_{p_i}$ for $p_i \neq p$, and as*

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ p-1 & 0 & \cdots & 0 & 0 \\ 0 & p & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & p & p \end{pmatrix}$$

*on $V_p$ with the standard basis, (the diagonal elements are all 0 except for the first and last, and the subdiagonal elements are all $p$ except for the first).*

(3) *For $p|N$ we have $w_p$ acting trivially on $V_{p_i}$ for $p_i \neq p$, and as*

$$\begin{pmatrix} 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix} : V_p \to V_p.$$

We will omit the proof of this proposition.

**Remark 1.6.** Applying $w_2$ to $P_2$ when $N = 4M$ with $M$ odd, we see that $w_2$ has a fixed point on $X_0(4M)$.

We can use this explicit formula to calculate the action of $T_p$ on various elements in the cuspidal subgroup.

**Proposition 1.7.** *Let $M = \prod p_i$ be an odd square-free integer and $N = 2^a M$ for some $a < 4$. Let $p|N$ and assume that an element $v \in V$ can be written as $v_p \otimes v'$, where $V \simeq V_p \otimes V'$ in an obvious way, $v_p \in V_p$, and $v' \in V'$. Then*

(1) *If $p \| N$ and $v_p = P_1 - P_p$ then $T_p v = v$.*

(2) *If $p \| N$ and $v_p = P_1 + P_p$ then $T_p v = v + u$ where $u = 2(p-1)P_p \otimes v'$.*

(3) *If $p = 2$, $N = 4M$, and $v_2 = P_2$ then $T_2 v = u$ with $u = 2P_4 \otimes v'$.*

(4) *If $p = 2$, $N = 8M$, and $v_2 = P_1 - P_8$ then $T_2 v = u$ where $u = (P_1 + P_2 - 2P_4) \otimes v'$.*

*Furthermore if $v$ represents an element in the cuspidal group of order 2, then $u$ is linearly equivalent to 0.*

*Proof.* Calculating the action of various Hecke operators on these elements is a straightforward matrix multiplication. As for proving that $u$ is linearly equivalent to 0, one can check directly that $\Lambda u$ satisfies conditions (1) and (4) of Proposition 1.3, as long as $\Lambda(2v)$ does. □

**Corollary 1.8.** *Let $v \in V$ be a cuspidal element considered in Table 1. Assume that the order of $v$ in $C_N$ is even and let $\lambda$ be such that $\lambda v \in C_N[2]$. Then for $p \| N$ (resp. $p = 2$ and $4|N$) we have $T_p(\lambda v) = \lambda v$ (resp. $T_2(\lambda v) = 0$).*

*Proof.* This follows immediately from Proposition 1.7. □

**Remark 1.9.** We can also write the elements in Table 1 using the Atkin–Lehner involution. For example, the element $(P_1 - b_1 P_{p_1}) \otimes \cdots \otimes (P_1 - b_1 P_{p_k})$ can be written as $(1 - b_1 w_{p_1}) \cdots (1 - b_k w_{p_k}) P_1$.

Recall that if $A$ is a simple new modular form, then for $p \| N$, $T_p|_A$ is acting as either 1 or $-1$, and when $p^2|N$ then $T_p|_A = 0$. Hence the corollary tells us that if one of the elements in Table 1 has an even order, then we have a nontrivial element in $C_N[2]$ that is new. This will be used to create congruences between modular forms in later sections.

## 2. Modular abelian varieties with odd congruence number

In this section we will study simple modular abelian varieties with odd congruence number. By examining the twists of modular abelian varieties, the action of the Atkin–Lehner involutions, and the order of the cuspidal subgroup, we demonstrate that if we have an absolutely simple modular abelian variety with odd congruence number, then its conductor $N$ has at most two prime divisors. We also show that the odd part of $N$ is either square-free or a power of a prime, and if $16|N$, then $N$ is a power of 2. Furthermore, we find some congruences that prime divisors of $N$ must satisfy.

Throughout this section we let $A$ be an optimal modular abelian variety with conductor $N$ and we fix a surjective map $\phi : J_0(N) \to A$ defined over $\mathbb{Z}[1/N]$. Furthermore, let $\pi : X_0(N) \to A$ be the composition of the Albanese embedding and $\phi$. As usual, let $\mathbb{T}$ be the Hecke algebra acting on $J_0(N)$ and $S(N)$.

**2A.** ***Atkin–Lehner involution.*** The goal of this section is to prove the following:

**Theorem 2.1.** *Let $A$ be a new simple modular abelian variety with odd modular exponent. If $A(\mathbb{Q})$ has no 2-torsion points, the conductor of $A$ is a power of a prime. Furthermore if $A$ has good reduction at 2 and $A(\mathbb{F}_2)$ has no 2-torsion points, the conductor of $A$ is a power of a prime.*

This theorem was proved by Calegari and Emerton [2009, Theorem 2.1] in the case where $A$ is an elliptic curve. Here, we apply their techniques to higher-dimensional modular abelian varieties.

**Lemma 2.2.** *Let $k$ be a field and $f : X/k \to Y/k$ be a degree $m$ map between curves. Then the composition*

$$\mathrm{Jac}(Y) \simeq \mathrm{Jac}(Y)^\vee \xrightarrow{f^*} \mathrm{Jac}(X)^\vee \simeq \mathrm{Jac}(X) \xrightarrow{f_*} \mathrm{Jac}(Y)$$

*is multiplication by $m$.*

*Proof.* It suffices to verify the lemma for the points $(z_1) - (z_2) \in \mathrm{Jac}(Y)$, since these points generate $\mathrm{Jac}(Y)$. Unraveling the definitions we get

$$\begin{aligned} f_*\big(f^*((z_1) - (z_2))\big) &= f_*\Big(\sum_{f(y_1)=z_1} (y_1) - \sum_{f(y_2)=z_2} (y_2)\Big) \\ &= \Big(\sum_{y_1 \in f^{-1}(z_1)} (z_1) - \sum_{y_2 = f^{-1}(z_2)} (z_2)\Big) = m((z_1) - (z_2)), \end{aligned}$$

where the summations are understood to account for multiplicities. □

**Lemma 2.3.** *Let $w$ be an involution on $X_0(N)$. Assume that*

$$\begin{array}{ccc} X_0(N) & \xrightarrow{\pi} & \\ {\scriptstyle w}\downarrow & & A \\ X_0(N) & \xrightarrow{\pi} & \end{array}$$

*commutes. Then the modular exponent of $A$ is even.*

*Proof.* The assumptions imply that $\pi$ factors through

$$X_0(N) \longrightarrow X_0(N)/w \longrightarrow A.$$

Therefore $\phi$ factors through

$$\mathrm{Jac}(X_0(N)) \longrightarrow \mathrm{Jac}(X_0(N)/w) \longrightarrow A.$$

Dualizing this diagram and using the autoduality of $J_0(N)$, we get

$$\begin{array}{ccccc} A^\vee & \longrightarrow & \mathrm{Jac}(X_0(N)/w)^\vee & \longrightarrow & J_0(N)^\vee \\ {\scriptstyle\delta}\downarrow & & \downarrow & & \downarrow \\ A & \longleftarrow & \mathrm{Jac}(X_0(N)/w) & \longleftarrow & J_0(N). \end{array}$$

By Lemma 2.2, the middle arrow is multiplication by 2, since the degree of $X_0(N) \to X_0(N)/w$ is 2. Using the commutativity of this diagram, we can see that $A^\vee[2] \subset \ker(\delta)$. Recalling that the modular exponent is the exponent of the kernel of $\delta$, we conclude that the modular exponent of $A$ is even. □

Recall that for an involution map $w : X_0(N) \to X_0(N)$, we get the induced map $w_* : J_0(N) \to J_0(N)$. Let $A$ be an optimal modular abelian variety, and $\phi : J_0(N) \to A$ the associated surjective map. Then if $w_*$ keeps $\ker(\phi)$ invariant, then $w_*$ acts on $A$ as well (this happens when, for example, $w$ is an Atkin–Lehner involution and $A$ is new). The following lemma deals with the case when $w_*$ is trivial on $A$.

**Lemma 2.4.** *Let $k$ be either $\mathbb{Q}$ or $\mathbb{F}_p$ with $p \nmid N$. Let $A$ be an optimal modular abelian variety with odd modular exponent. As before let $\pi : X_0(N) \to A$ be the composition of Albanese embedding $X_0(N) \to J_0(N)$ and $\phi$. Assume that for some involution $w$, $w_* : J_0(N) \to J_0(N)$ descends down to a trivial action on $A$. Then $\pi(w(z)) - \pi(z)$ is a nontrivial $k$-rational 2-torsion point for all $z \in X_0(N)(\bar{k})$.*

*Proof.* Recall that $P_1$ is the cusp at infinity and $\pi(z) = \phi(z - P_1)$. Then we get

$$\begin{aligned} \pi(w(z)) - \pi(z) &= \phi(w(z) - P_1) - \phi(z - P_1) \\ &= \phi(w(z) - w(P_1)) - \phi(z - P_1) + \phi(w(P_1) - P_1) \\ &= w_*(\phi(z - P_1)) - \phi(z - P_1) + \phi(w(P_1) - P_1) = \pi(w(P_1)). \end{aligned}$$

Therefore $\pi(w(z)) = \pi(z) + \pi(w(P_1))$ for all $z \in X_0(N)$. Applying this equation to $w(z)$ we get $\pi(w(w(z))) = \pi(w(z)) + \pi(w(P_1)) = \pi(z) + 2\pi(w(P_1))$. Therefore, $2\pi(w(P_1)) = 0$. By Lemma 2.3, if $A$ has odd modular exponent, then $\pi(w(z)) - \pi(z)$ is nontrivial. Thus, $\pi(w(P_1))$ is a nontrivial 2-torsion point of $A$. It is $k$-rational because $w(P_1)$ is also $k$-rational. □

*Proof of Theorem 2.1.* Let $W$ be the group of Atkin–Lehner involutions on $X_0(N)$, and let $k = \mathbb{Q}$ or $\mathbb{F}_2$ when $N$ is odd. Since we are assuming that $A$ is new and simple, for any Atkin–Lehner involution $w \in W$, we have $w_*(z)$ is either $z$ or $-z$ for all $z \in A(\bar{k})$. This gives us a group homomorphism $W \to \{\pm 1\}$. Let $W_0$ be the kernel of this map. Note that $W_0$ has index at most 2 in $W$. Assume that $N$ is not a power of a prime; hence $W$ will have more than 2 elements. Therefore, we can find a nontrivial element $w \in W_0$; then $w_*(z) = z$ for all $z \in A(\bar{k})$. Applying Lemma 2.4, we find that $0 \neq \pi(w(P_1)) \in A[2](k)$. Therefore, if $A[2](k) = 0$ then $N$ must be a power of a prime. □

Lemma 2.4 can also be used to find the signs of the Atkin–Lehner involutions on $A$ in certain cases.

**Lemma 2.5.** *Let $A$ be a new modular simple abelian variety with conductor $N$ and odd modular exponent. If the Atkin–Lehner involution $w_r : X_0(N) \to X_0(N)$ has a fixed point then $(w_r)_*$ acts as $-1$ on $A$. Specifically, $(w_N)_*$ acts as $-1$ on $A$. When $N = 2M$ (resp. $N = 4M$), $(w_2)_*$ acts as 1 (resp. $(w_2)_*$ acts as $-1$) on $A$.*

*Proof.* Let $P \in X_0(N)(\bar{\mathbb{Q}})$ be the fixed point of $w_r$. Then $\pi(P) = \pi(w_r(P))$, which implies that $\pi(w_r(P)) - \pi(P) = 0$. However, we know that if $(w_r)_* = 1$ then $\pi(w_r(z)) - \pi(z) = \pi(w_r(P_1))$ for any $z \in X_0(N)(\bar{\mathbb{Q}})$. Specifically, we get $\pi(w_r(z)) = \pi(z)$, which by Lemma 2.3 implies that $A$ has an even congruence number. Therefore $(w_r)_* = -1$ when $w_r$ has a fixed point in $X_0(N)$.

Finally, the point $\sqrt{-N}$ is fixed by $w_N$. When $N = 2M$, we can check that $1/(M - i\sqrt{M})$ is fixed under $(w_M)_*$. Similarly, when $N = 4M$, $P_2$ is fixed under $(w_2)_*$. Therefore, we have the desired result. □

Since $(w_N)_*$ is the sign of the functional equation, we get the following:

**Corollary 2.6.** *If $A$ is a simple modular abelian variety with odd congruence number, then the analytic rank of $A$ is even.*

**Remark 2.7.** Calegari and Emerton used Theorem 2.1 for modular elliptic curves $E$ with odd modular degree and conductor $N$ to show that $N$ has at most two odd prime divisors. Specifically, since $E[2](\mathbb{Q})$ has at most 4 elements, an immediate corollary of Theorem 2.1 is that if $N$ has more than 3 prime divisors, then $E$ has even modular degree. Similarly, if $E$ has good reduction at 2, then since $E[2](\mathbb{F}_2)$ has at most two elements, they conclude that if $N$ has more than two prime divisors then $E$ has even modular degree.

**2B.** ***Nonsemistable case.*** The goal of this section is to prove the following:

**Theorem 2.8.** *Let $A$ be an absolutely simple modular abelian variety $A$ of level $N$ with odd congruence number. Let $\delta_p = 0$ for the odd primes $p$ and $\delta_2 = 2$. Assume that $p^{2+\delta_p} | N$. Then $A$ has good reduction away from $p$ and 2, and has potentially good reduction everywhere. Specifically, if $p$ is odd and $p^2 | N$, then $N = p^s$, $N = 4p^s$, or $N = 8p^s$ for $s \geq 2$, and if $16 | N$ then $N = 2^s$.*

We expect this theorem to be true without assuming $A$ to be absolutely simple; however, at this moment we do not know how to overcome the difficulty with the inner forms in that case. To prove this theorem, we use the technique of [Calegari and Emerton 2009] to show that such modular abelian varieties have inner twists or complex multiplication by a character of conductor $p$. Using the results of [Ribet 1981] on inner twists, we will prove that $A$ must have potentially good reduction everywhere if $A$ is absolutely simple, and that $A$ has good reduction away from $p$, and possibly 2. We have the following lemma.

**Lemma 2.9.** *If $\mathrm{End}_{\overline{\mathbb{Q}}}(A) \otimes \mathbb{Q}$ is a matrix algebra, then $A$ is not absolutely simple.*

*Proof.* Assume that $R = \mathrm{End}_{\overline{\mathbb{Q}}}(A) \otimes \mathbb{Q}$ is a matrix algebra. We can find the projections $e_1, e_2 \in R$ such that $e_1 + e_2 = \mathrm{Id}$, $e_1 e_2 = 0$, and $e_1, e_2 \notin \{0, \mathrm{Id}\}$. For some integer $n$, $ne_i \in \mathrm{End}_{\overline{\mathbb{Q}}}(A)$. If we assume that $A$ is absolutely simple, the image of $ne_i A$ must be $A$ or 0. However, since the product of $ne_1$ and $ne_2$ equals $n^2 e_1 e_2 = 0$, one of them must be 0. Assume without loss of generality that $ne_2 = 0$ in $\mathrm{End}_{\overline{\mathbb{Q}}}(A)$. This implies that $e_2 = 0$, which contradicts the assumption $e_2 \notin \{0, \mathrm{Id}\}$. Therefore, $A$ is not absolutely simple. □

This lemma is used in conjunction with Ribet's result on the endomorphism algebra of modular abelian varieties with inner twists. Specifically, let $A$ be a $d$-dimensional simple modular abelian variety. There are $d$ modular eigenforms of weight 2 and level $N$ associated with $A$, which are Galois conjugate to each other. Let $f = \sum a_n q^n$ be one such eigenform and let $E = \mathbb{Q}(\ldots, a_n, \ldots)$ be the field of definition of $f$. We know that $\mathrm{End}_{\mathbb{Q}}(A) \otimes \mathbb{Q} = E$. Let

$$D = \mathrm{End}_{\overline{\mathbb{Q}}}(A) \otimes \mathbb{Q}$$

be the algebra of endomorphisms of $A$. From [Ribet 1980] we know that $E$ is its own commutant in $D$, and therefore $D$ is a central simple algebra over some subfield $F$ of $E$. If we assume that $A$ is absolutely simple, then $D$ must be some division algebra with centre $E$. Furthermore, $D$ must be either $E$ (which forces $E = F$) or a quaternion division algebra over $F$ (which forces $E$ to be a quadratic extension of $F$).

**Proposition 2.10.** *Let $A$ be an absolutely simple modular abelian variety $A$ of level $N$ with odd congruence number. Let $\delta_p = 0$ for odd primes and $\delta_2 = 2$. If*

*$p^{2+\delta_p}|N$, then $A$ has a complex multiplication or an inner twist and $A$ has potentially good reduction everywhere. Specifically, for any other prime number $q$ if $q|N$ then $q^2|N$.*

*Proof.* Assume that $A$ is of dimension $d$, and let

$$f_A = \sum a_n q^n \in \mathbb{C}((q))$$

be a normalized eigenform associated with $A$. Let $E = \mathbb{Q}(\dots, a_i, \dots) \subset \mathbb{C}$. Let $\chi$ be the quadratic character with conductor $p$. Since $p^{2+\delta_p}|N$, $\chi \otimes f_A$ is another modular eigenform in $S_2(\Gamma_0(N))$ (see [Shimura 1971]). Since $\chi$ is a quadratic character, $\chi$ takes values in $\pm 1$, and as a result $\chi \otimes f_A \equiv f_A \pmod{\lambda}$ for any $\lambda|2$. If $A$ has odd congruence number, then $\chi \otimes f_A$ must be in the same conjugacy class as $f_A$. If $\chi \otimes f_A = f_A$, then $A$ has complex multiplication by $\chi$, and therefore $A$ has potentially good reduction everywhere. In this case, $A$ must be an elliptic curve, because if $A$ has complex multiplication and has a dimension greater than 1, then the ring of endomorphisms of $A$ is a matrix algebra, which contradicts the absolute simplicity assumption. In general, $A$ might have an inner twist, and $\chi \otimes f_A = \gamma(f_A)$ for some $\gamma \in \mathrm{Hom}(E, \mathbb{C})$. Let $\Gamma \subset \mathrm{Hom}(E, \mathbb{C})$ such that for any $\gamma \in \Gamma$ we can find a character $\chi_\gamma$ such that $\chi_\gamma \otimes f_A = \gamma(f_A)$. By [Ribet 1981], $F = E^\Gamma$ and (as discussed above) $D = \mathrm{End}_{\bar{\mathbb{Q}}} A \otimes \mathbb{Q}$ must be a quaternion algebra. However, using [Ribet 1980, Theorem 3], $A$ has potentially good reduction everywhere, as desired.

The final claim of the lemma follows by noting that if $q|N$ but $q^2 \nmid N$, then $A$ has multiplicative reduction over any field extension. □

*Proof of Theorem 2.8.* Assume that $p^{2+\delta_p}|N$ and $q^{2+\delta_q}|N$ for distinct primes $p$ and $q$. In this case, assuming that $A$ has no complex multiplication, $A$ has more inner twists, and the subset $\Gamma \subset \mathrm{Hom}(E, \mathbb{C})$ will have at least four elements, $\gamma_1$, $\gamma_p$, $\gamma_q$, and $\gamma_{pq}$. But that means that $|E : F| \geq 4$, which shows that $D$ must be a matrix algebra. However, Lemma 2.9 forces $A$ not to be absolutely simple, which contradicts our assumption. Since we are assuming $A$ is absolutely simple if $A$ has complex multiplication, then $A$ is an elliptic curve. Therefore it will have complex multiplication by $\chi_p$ and $\chi_q$, which is impossible. □

**2C.** ***Number of primes of bad reduction.*** We now show that a modular abelian variety with odd congruence number has bad reduction at no more than two primes. Let $A$ be an absolutely simple optimal abelian variety of conductor $N$. Let $B = \ker(\phi)$ where $\phi$ is the modular uniformization map $\phi : J_0(N) \to A$. Assume that $N$ is a not a power of a prime. Then Theorem 2.1 says that $A[2](\mathbb{Q})$ has a nontrivial element. Let $z \in A[2](\mathbb{Q})$ be a nontrivial rational 2-torsion point of $A$, and let $\mathfrak{m} \subset \mathbb{T}$ be the annihilator of $z$. Since $z \in A[\mathfrak{m}] \neq 0$, we get that $A^\vee[\mathfrak{m}] \neq 0$. Therefore, if $B[\mathfrak{m}] \neq 0$ as well, then $A$ will have an even congruence number. We will show that when $N$ has more than two prime divisors, then $B[\mathfrak{m}] \neq 0$.

**Lemma 2.11.** *Let $A$ be a new simple modular abelian variety, $0 \neq z = A[2](\mathbb{Q})$, and let $\mathfrak{m}$ be the annihilator of $z$ in $\mathbb{T}$. Then $\mathfrak{m}$ is generated by 2, $T_l - (l+1)$ for $l \nmid N$, $T_p - 1$ for $p|N$ but $p^2 \nmid N$, and $T_p$ for $p^2|N$.*

*Proof.* Clearly $z$ is killed by 2, and since $A$ is a new modular abelian variety, if $p \| N$, we have $T_p(z) = \pm z = z$, and if $p^2|N$ then $T_p(z) = 0$. Let $\rho$ be the Galois representation

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{T}/\mathfrak{m}).$$

Since $A$ has a rational 2 torsion point,

$$\rho \simeq \begin{pmatrix} \chi & * \\ 0 & 1 \end{pmatrix},$$

where $\chi$ is the cyclotomic character. Therefore $\mathrm{trace}(\rho(\mathrm{Frob}(l))) = 1 + \chi(l) = l + 1$ for $l \nmid 2N$, and hence by the Eichler–Shimura relationship we get that $T_l - (l+1) \in \mathfrak{m}$ for $l \nmid 2N$. When $l = 2$ and $N$ is odd, by Theorem 2.1 we get that $A$ has ordinary reduction at 2, and hence $T_2 \notin \mathfrak{m}$, or equivalently $T_2 - (2+1) \in \mathfrak{m}$. □

Recall that $C_N \subset J_0(N)$ is the rational cuspidal subgroup of $J_0(N)$. Let $\mathfrak{m} \subset \mathbb{T}$ be the annihilator of $z \in A[2]$. We can use the elements considered in Table 1 to show that $B[\mathfrak{m}] \neq 0$ when $N$ has more than two prime divisors. Specifically, if $v \in C_N$ of even order is such that $\phi(v) = 0$, then $v \in B \cap C_N$. If $v$ is also a cusp of the type considered in Table 1 and of even order, then by Corollary 1.8 we get that for some integer $\lambda$ we have that $\lambda v \in C_N[\mathfrak{m}]$. Therefore, to show that $A$ has an even congruence number, we only need to check that such $v$'s have even order and that $\phi(v) = 0$.

**Theorem 2.12.** *Let $A$ be a new absolutely simple optimal modular abelian variety with odd congruence number. Then $N$ has at most two prime factors.*

*Proof.* If $A$ has an inner twist or complex multiplication, then the result follows by Theorem 2.8. Assume that $A$ has odd congruence number with no inner twist or complex multiplication. Assume to the contrary that $N$ has more than two prime factors. Then $N = 2^\alpha M$ with $M$ a square-free odd integer, and $\alpha < 4$. Furthermore, by Theorem 2.1, we can find a nontrivial $z \in A[2](\mathbb{Q})$. Let $\mathfrak{m}$ be the annihilator of $z$. We will consider three main cases, based on the valuation of $N$ at 2.

Assume that $4 \nmid N$. Since $w_N = \prod_{l|N} w_l$, and $(w_N)_* = -1$, there is an odd number of primes such that $(w_l)_*$ act as $-1$ on $A$. Therefore, we can select three distinct prime divisors of $N$ — call them $p$, $q$, and $r$ — such that $(w_p)_*$ acts as $-1$, while $(w_r)_* = (w_q)_*$. If $2 \| N$, by Lemma 2.5 $(w_2)_*$ acts as $+1$. Therefore, without loss of generality assume that $2 \nmid pq$.

Let $s_p$, $s_q = \pm 1$ and let

$$v = (1 - w_{qr})(1 + s_p w_p)(1 + s_q w_q) P_1 = (1 + s_p w_p)(1 + s_q w_q)(1 - s_q w_r) P_1.$$

Consulting Table 1 we get that $v$ has order $\mathrm{Num}\big((1+s_p p)(1+s_q q)(1-s_q r)/24\big)$. If we select $s_p \equiv -p \pmod 4$ and $s_q \equiv -q \pmod 4$, then this order is even. Next we show that $\phi(v)=0$. Note that $\pi(w_{qr}(\tau)) = \pi(\tau)+a$ for any $\tau \in X_0(N)$, where $a$ is some 2-torsion point. Let $P = (1+s_p w_p)(1+s_q w_q)P_1 = P_1 \pm P_p \pm P_q \pm P_{pq}$. Then

$$\phi(v) = \phi(w_{qr}(P) - P) = \sum_{m|pq} \pi(w_{qr}(P_m)) - \pi(P_m) = 4a = 0,$$

which shows that $A$ has an even congruence number.

Assume that $4\|N$. By Lemma 2.5 we know that $(w_2)_*$ acts as $-1$. Let $p, q|N$ and let $v = (1-w_p)(1+s_q w_q)P_2$ with $s_q = \pm 1$. From Table 1 we get that the order of $v$ is $\mathrm{Num}\big((1-p)(1+s_q q)/4\big)$. If we select $s_q \equiv -q \pmod 4$, then $v$ will have an even order. Since $(w_2)_*$ is acting as $-1$, either $(w_p)_*$ or $(w_{2p})_*$ is acting trivially on $A$. Let $w$ be the corresponding Atkin–Lehner involution. Note that because $w_2(P_2) = P_2$, $v = (1-w)(1+s_q w_q)P_2$. Furthermore, $\pi(w(\tau)) - \pi(\tau) = a \in A[2]$ for any $\tau \in X_0(N)$. As a result,

$$\phi(v) = \pi(P_2) - \pi(w(P_2)) + s_q(\pi(P_{2q}) - \pi(w(P_{2q})) = a + s_q a = 0.$$

Therefore $\phi(v) = 0$, which proves that in this case $A$ has an even congruence number.

Finally assume that $8\|N$, and let $p, q|N$ be two distinct odd divisors of $N$. Let $(w_p)_*$ and $(w_q)_*$ act as $s_p$ and $s_q$ on $A$. Let

$$v = (1-w_2)(1+s_p w_p)(1+s_q w_q)P_1 = (1-w_2)(1+s_p s_q w_{pq})(1+s_p w_p)P_1.$$

Then, again from Table 1, $v$ has order $\mathrm{Num}\big((1+s_p p)(1+s_q q)/2\big)$ that is even. Note that, similar to the case when $N$ is odd, we can write $v = (1-w)P$ for some Atkin–Lehner involution $w$ such that $w_* = 1$ and some $P = (1-w_2)(1 \pm w')P_1$. That shows $\phi(v) = 0$. Therefore $A$ in this case will have an even congruence number again. □

Combining this result with Theorem 2.8, we get:

**Corollary 2.13.** *Let $A$ be an absolutely simple modular abelian variety with odd congruence number and conductor $N$. Then $N$ has at most two prime divisors. Furthermore, if $N$ is not square-free, then $N = 2^a$, $p^b$, $4p^b$, or $8p^b$, where $p$ is an odd prime.*

**2D. *Congruence classes of primes.*** Let $A$ be a simple modular abelian variety of conductor $N$ with odd congruence number, and without complex multiplication or an inner twist. As usual let $\pi : X_0(N) \to A$ to be the composition of the Albanese embedding with the modular uniformization $\phi$. Assume that $N$ is not a power of a prime, which by Theorem 2.1 implies that $A[2](\mathbb{Q})$ is nontrivial. From the previous

sections we know that $N$ has at most two prime factors, say $p$ and $q$. In this section we find congruences that $p$ and $q$ must satisfy. As in the proof of Theorem 2.12, we use different techniques depending on the valuation of $N$ at 2.

If $N$ is odd, then $N = pq$ with both $p$ and $q$ being odd. By Lemma 2.5, we know that $(w_{pq})_*$ is acting as $-1$ on $A$. Therefore, assume without loss of generality that $(w_q)_*$ is acting trivially on $A$ and $(w_p)_*$ is acting as $-1$. Let $v = (1 \pm w_p)(1 - w_q)P_1$. Again, $\pi(\tau) - \pi(w_q(\tau)) = a \in A[2]$ for all $\tau \in X_0(N)$. As a result,

$$\phi(v) = \pi(P_1) - \pi(w_q(P_1)) \pm (\pi(P_p) - \pi(w_q(P_p))) = a \pm a = 0.$$

Note that the order of $v$ is $\mathrm{Num}\big((p \pm 1)(q-1)/24\big)$. Since we are assuming that $A$ has odd congruence number, we get that $p \equiv \pm 3 \pmod 8$ and $q \equiv 3 \pmod 4$.

**Corollary 2.14.** *Let $A$ be a modular abelian variety with odd congruence number and conductor $pq$, where $p$ and $q$ are odd. Then $A[2](\mathbb{Q})$ is at least 2-dimensional over $\mathbb{F}_2$.*

*Proof.* We prove this by finding two distinct points in $C_N[\mathfrak{m}]$. First note that $P_1 - P_p$ has order $(p-1)(q^2-1)/24$ and $P_1 - P_q$ has order $(p^2-1)(q-1)/24$. Therefore, both

$$u = \frac{(p-1)(q^2-1)}{48}(P_1 - P_p) \quad \text{and} \quad u' = \frac{(p^2-1)(q-1)}{48}(P_1 - P_q)$$

are of order 2. We can easily check that $T_p u = u$ and $T_q u' = u'$. On the other hand

$$u + T_q u = \frac{(p-1)(q^2-1)}{48}(P_1 - P_p + P_q - P_{pq}),$$

which is zero. Similarly, we get $u' + T_p u' = 0$. Therefore, $u, u' \in C_N[\mathfrak{m}]$. Furthermore, we know that $\Lambda(u + u')$ has integral coefficients, but

$$(1, 0) \otimes (1, 1)\Lambda(u + u') = (q-1)/2,$$

which is not even since $q \equiv 3 \pmod 4$. Therefore, $u + u' \neq 0$, which implies that $C_N[\mathfrak{m}]$ is at least 2-dimensional over $\mathbb{F}_2$. Since we are assuming that $A$ has odd congruence number, $C_N[\mathfrak{m}]$ injects in $A$, which is the desired result. □

If $N = 2p$, we know by Lemma 2.5 that $(w_2)_*$ acts trivially and $(w_p)_*$ acts as $-1$ on $A$. Therefore, $\pi(P_2) = \pi(w_2(P_1)) \in A[2]$, and $P_2 - P_1$ (which has order $(p^2-1)/8$) must have an even order. Let $v = ((p^2-1)/16)(P_2 - P_1) \in C_N[2]$. By Corollary 1.8 we get $v \in C_N[\mathfrak{m}]$. Note that

$$\phi(v) = \pi\Big(\frac{p^2-1}{16}(P_2 - P_1)\Big) = \frac{p^2-1}{16}\pi(P_2),$$

so if $(p^2-1)/16$ is even, then $\pi(v) = 0$. This implies that $v \in C_N[\mathfrak{m}] \cap B$, and, in turn, that the congruence number is even. Since we are assuming that

the congruence number of $A$ is odd, we get that $(p^2-1)/16$ is odd; that is, $p^2-1 \equiv 16 \pmod{32}$. That implies that $p \equiv \pm 7 \pmod{16}$. However, we also know that $w_2$ cannot have any fixed points. This implies that $-2$ is not a quadratic residue mod $p$, which means that $p \equiv 5, 7, 13,$ or $15 \pmod{16}$. Therefore $p \equiv 7 \pmod{16}$.

If $N = 4p$, then we know that $(w_2)_*$ acts as $-1$ on $A$, while $(w_p)_*$ acts trivially. Therefore, $\pi(P_2) - \pi(P_{2p}) = \pi(P_2) - \pi(w_p(P_2)) \in A[2]$. The order of $P_2 - P_{2p}$ is $(p-1)/2$. Therefore, if $A$ has odd congruence number, $(p-1)/4$ must be odd, hence $p \equiv 5 \pmod{8}$.

If $N = 8p$, we can check that $(1-w_2)(1-w_p)P_1$ vanishes in $A$, and that it has order $(p-1)/2$. Therefore, $4 \nmid p-1$, otherwise $A$ will have an even congruence number. Therefore $p \equiv 3 \pmod{4}$. (We can probably say more, if we figure out the sign of $(w_p)_*$.)

We combine the above results in the following theorem.

**Theorem 2.15.** *Let $A$ be a new modular abelian variety with odd congruence number and conductor $N$. Assume that $A$ has no inner twists or complex multiplications. Then one of the following must be true.*

(1) *$N$ is a prime number $p$.*

(2) *$N = pq$ and $p \equiv \pm 3 \pmod{8}$ and $q \equiv 3 \pmod{4}$.*

(3) *$N = 2p$ and $p \equiv 7 \pmod{16}$.*

(4) *$N = 4p$ and $p \equiv 5 \pmod{8}$.*

(5) *$N = 8p$ and $p \equiv 3 \pmod{4}$.*

## 3. Elliptic curves with odd congruence number

In this section, we apply the results of the previous section to the case of elliptic curves. We show that the conductors of all such elliptic curves are of the form $p$, $pq$, $2p$, $4p$, or one of the finitely many exceptions. We study each class to demonstrate that all such elliptic curves have finite Mordell–Weil group, except possibly when the conductor is prime. Furthermore, we know from the result of [Agashe et al. 2008] that when $4 \nmid N$, then having odd congruence number is the same as having odd modular degree. As a result, we can state many of our results in terms of modular degrees. We conjecture that in fact having odd congruence number is equivalent to having odd modular degree in all cases.

***Complex multiplication.*** Let $E$ be an elliptic curve of conductor $N$. If $p^2|N$ for an odd prime $p$, then by Section 2B we know that $E$ has complex multiplication. We also showed that if $16|N$ then $E$ must have complex multiplication. There are only finitely many elliptic curves over rationals with complex multiplication and the conductor $2^m p^n$ for some prime number $p$. The following is the list of all such

elliptic curves that have odd modular degree: $E = 27A, 32A, 36A, 49A, 243B$. We also verify that all such elliptic curves have rank 0, as predicted by Watkins' conjecture.

We will now focus our attention on elliptic curves without complex multiplication, that is, elliptic curves with conductor $N = p, 2p, 4p, 8p$, or $pq$ for some odd primes $p$ and $q$. Each of the remaining sections deals with one of these remaining cases.

***Prime level.*** Let $E$ be an elliptic curve with odd congruence number and a prime conductor $N$. Mestre and Oesterlé [1989] have studied the elliptic curves of prime conductors, and they have demonstrated that aside from elliptic curves $11A$, $17A$, $19A$, and $37B$, all such elliptic curves have either a trivial torsion subgroup or a $\mathbb{Z}/2\mathbb{Z}$ torsion subgroup. The above cases have the torsion structures $\mathbb{Z}/5\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, and $\mathbb{Z}/3\mathbb{Z}$, respectively. Mestre and Oesterlè also showed that if $E_{\text{tors}}$ is $\mathbb{Z}/2\mathbb{Z}$, then $E$ is a Neumann–Setzer curve and $N = u^2 + 64$. Stein and Watkins [2004] have studied the parity of congruence numbers of Neumann–Setzer curves and they show that $E$ has odd congruence number if and only if $u \equiv 3 \pmod 8$. Furthermore one can show that Neumann–Setzer curves have rank 0 using descent. We will give another proof of this fact using $L$-functions.

**Proposition 3.1.** *Let $E$ be an elliptic curve over $\mathbb{Q}$ with a prime conductor $N$. Assume that $E_{\text{tors}}$ is nontrivial. Then $L(E, 1) \neq 0$, hence $E(\mathbb{Q})$ has rank 0.*

*Proof.* Recall that

$$L(E, 1) = 2\pi i \int_0^{i\infty} f_E(z)\, dz \equiv \pi(P_N) \pmod{\Lambda_E},$$

where $\mathbb{C}/\Lambda_E \simeq E(\mathbb{C})$. Therefore, if $L(E, 1) = 0$, then $\pi(P_N) = 0$, or alternatively $\phi(P_1 - P_N) = 0$. By [Mazur 1977; Mestre and Oesterlé 1989] (see also [Emerton 2003]) we know that $J_0(N)_{\text{tors}}$ is generated by the cusp $P_1 - P_N$, and for any elliptic curve quotient of $J_0(N) \to E$, $E_{\text{tors}}$ is generated by the image of $\pi(P_1) - \pi(P_N)$. Since we are assuming that $E$ has a nontrivial torsion structure, $\pi(P_1) - \pi(P_N) \neq 0$, which implies that $L(E, 1) \neq 0$. Therefore the rank of $E(\mathbb{Q})$ is zero by [Kolyvagin 1988; Gross and Zagier 1986]. □

The case when $E$ has trivial torsion structure and odd congruence number was studied in [2009], where it is shown that $E$ has an even analytic rank (since $(w_N)_* = -1$), supersingular reductions at 2 and $E(\mathbb{R})$ is connected. From a search in Cremona's database, it appears that if an elliptic curve $E$ has supersingular reduction at 2, Mordell–Weil rank 0, and a connected real component, then $E$ will have odd congruence number.

***Level $N = pq$.*** We consider elliptic curves of odd modular degree and conductor $N = pq$, where $p$ and $q$ are both odd primes. Let $E$ be such an elliptic curve. Assume throughout this section that $(w_p)_* = -1$ on $E$. By Theorem 2.15, we know that $p \equiv \pm 3 \pmod 8$ and $q \equiv 3 \pmod 4$. We will show that with a few exceptions, $p, q \equiv 3 \pmod 8$, and that all such elliptic curves have finite Mordell–Weil group over $\mathbb{Q}$.

Recall that by Corollary 2.14 we know that $E[2](\mathbb{Q}) = (\mathbb{Z}/2\mathbb{Z})^2$. First, we show that if $E_{\text{tors}}$ is $\mathbb{Z}/2 \times \mathbb{Z}/4$, then $E$ has conductor 15 or 21. We can prove a general result about semistable elliptic curves with $E_{\text{tors}} = \mathbb{Z}/2 \times \mathbb{Z}/4$ and good reduction at 2:

**Lemma 3.2.** *Let $E$ be a semistable elliptic curve with good reduction at 2. $E_{\text{tors}} = \mathbb{Z}/2 \times \mathbb{Z}/4$, and let $Q \in E(\mathbb{Z}[1/N])$ be a point of order 4. Let $\overline{Q}$ be the reduction of $Q$ mod 2. Then $\overline{Q}$ has order 4 in $E(\mathbb{F}_2)$.*

*Proof.* We can check that an elliptic curve $E$ with good reduction at 2 and a rational 2-torsion point has a minimal model

$$E : y^2 + xy = x^3 + a_2x^2 + a_4x.$$

Since $E[2] = \mathbb{Z}/2 \times \mathbb{Z}/2$, $(4a_2+1)^2 - 64a_4$ is a perfect square. The $x$ coordinates of the 2-torsion points are $0$, $4\alpha$, and $\beta/4$, where $\alpha$ and $\beta$ are both (odd) integers since we are assuming that $E$ is in minimal model. Since $E$ is assumed to be semistable, $\alpha$ and $\beta$ are coprime. Note that the point $(\beta/4, -\beta/8) \in E(\mathbb{Q})$ maps to the identity under the reduction mod 2 map. Using the notation of [Silverman 1992], we have

$$b_2 = 16\alpha + \beta, \quad b_4 = 2\alpha\beta, \quad b_6 = 0, \quad b_8 = -\alpha^2\beta^2, \quad \Delta = \alpha^2\beta^2(16\alpha - \beta)^2.$$

Let $Q \in E(\mathbb{Q})$ be a point of order 4, and let $x(Q) = x_0$. Recall that we want to show $\overline{Q} \in E(\mathbb{F}_2)$ is a point of order 4. We have that $x([2]Q) = 0$, $4\alpha$, or $\beta/4$. If $\overline{Q}$ has order less than 4, then $2\overline{Q}$ must be the identity element, that implies that $x([2]Q) = \beta/4$. In that case

$$\frac{\beta}{4} = \frac{x_0^4 - b_4x_0^2 - b_8}{4x_0^3 + b_2x_0^2 + 2b_4x_0} = \frac{x_0^4 - 2\alpha\beta x_0^2 + \alpha^2\beta^2}{4x_0^3 + (16\alpha + \beta)x_0^2 + 4\alpha\beta x_0},$$

so

$$0 = x_0^4 - \beta x_0^3 - \left(6\alpha\beta + \frac{\beta^2}{4}\right)x_0^2 - \alpha\beta^2 x_0 + \alpha^2\beta^2 = \left(x_0^2 - \frac{\beta}{2}x_0 + \alpha\beta\right)^2 - \left(4\alpha\beta + \frac{\beta^2}{2}\right)x_0^2.$$

Therefore, $16\alpha\beta + 2\beta^2 = 2\beta(8\alpha + \beta)$ must be a perfect square; however that is not possible because $\alpha$ and $\beta$ are odd. As a result, $x([2]Q) = 0$ or $4\alpha$. Therefore, $[2]\overline{Q}$ has order 2 in $E(\mathbb{F}_2)$. This shows that $\overline{Q}$ has order 4, which is the desired result. □

**Proposition 3.3.** *Let $E$ be an elliptic curve with conductor $pq$ and torsion group $\mathbb{Z}/2 \times \mathbb{Z}/4$. Then $pq = 15$ or $21$.*

*Proof.* Using the same notation as in Lemma 3.2, let 0, $4\alpha$, and $\beta/4$ be the $x$-coordinates of the 2-torsion points of $E$. Let $Q$ be a point in $E_{\text{tors}}$ of order 4. By Lemma 3.2, $x([2]Q) = 0$ or $4\alpha$. Without loss of generality, assume that $x([2]Q) = 0$, since if $x([2]Q) = 4\alpha$, then we can change the coordinates to find another model with $x([2]Q') = 0$. Let $x_0 = x(Q)$. Then $x_0^4 - 2\alpha\beta x_0^2 + \alpha^2\beta^2 = 0$, which implies that $x_0^2 = \alpha\beta$. Since $\alpha$ and $\beta$ are coprime, they are both perfect squares, or negatives of perfect squares (both of the same sign). Since $E$ is of conductor $pq$, $\Delta = \alpha^2\beta^2(16\alpha - \beta)^2$ is a product of the powers of $p$ and $q$. Let $a^2 = \pm\alpha$ and $b^2 = \pm\beta$. Then, $a^4b^4(4a - b)(4a + b)$ is a product of the powers of $p$ and $q$. Note that $(4a - b, 4a + b) = 1$, which implies that all factors are pairwise coprime. Note that if $|4a + b| = |4a - b| = 1$, then either $a = 0$ or $b = 0$ contrary to our assumptions. Therefore we will assume without loss of generality that $4a + b > 1$.

If $4a - b \neq \pm 1$, then $a^2 = b^2 = 1$, which means $E$ is the elliptic curve $15A$. If $4a - b = \pm 1$ then $|b| > 1$; therefore $|a| = 1$. Since we are assuming that $4a + b > 1$ we get that $a = 1$, and $4a - b = 1$ leads to elliptic curve $21A$ and $4a - b = -1$ leads to elliptic curve $15A$. This completes our proof. □

**Remark 3.4.** Note that the previous proposition seems a bit tedious. It is straightforward to show that 3 must divide the conductor by the Hasse–Weil bound. Unfortunately, it is not clear how this observation can simplify the argument.

An immediate corollary of the above is that for an elliptic curve $E$ of conductor $pq$ and ordinary reduction at 2, we have $E_{\text{tors}} = (\mathbb{Z}/2\mathbb{Z})^2$, since the only other option is $E_{\text{tors}} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$. However the Hasse–Weil bounds for elliptic curves rule this case out.

**Theorem 3.5.** *Assume that $E$ is an elliptic curve with odd modular degree. Furthermore, assume that the conductor of $E$ is $pq$ with $pq \neq 21$ or $15$. Then $p, q \equiv 3 \pmod 8$.*

*Proof.* By Corollary 2.14 we know that $E[2](\mathbb{F}_2)$ is nontrivial, hence $E$ has good ordinary reduction at 2. Therefore, for $pq \neq 21$ and 15 we have $E(\mathbb{Q})_{\text{tors}} = (\mathbb{Z}/2\mathbb{Z})^2$. Recall that we are assuming $(w_p)_* = -1$ and $(w_q)_* = 1$ on $E$. Note that

$$\begin{aligned}\pi(\tau) - \pi(w_p(\tau)) &= \phi(\tau - P_1) - \phi(w_p(P_1) - P_1) - \phi(w_p(\tau) - w_p(P_1)) \\ &= \pi(\tau) - (w_p)_*(\pi(\tau)) - \pi(w_p(P_1)) = 2\pi(\tau) - \pi(P_p),\end{aligned}$$

for any $\tau \in X_0(N)$. When $\tau$ is a cusp of $X_0(N)$, $\pi(\tau)$ is a torsion point, and since $E_{\text{tors}} = E[2]$ we get $2\pi(\tau) = 0$. Therefore

$$\pi(\tau) - \pi(w_p(\tau)) = \pi(P_p).$$

Let $v = (1 + w_q)(1 - w_p)P_1$. Then

$$\phi(v) = \big(\pi(P_1) - \pi(w_p(P_1))\big) + \big(\pi(P_q) - \pi(w_q(P_q))\big) = 2\pi(P_p) = 0.$$

As a result, $v \in B \cap C_N$. Therefore, by Corollary 1.8, if $v$ has even order then $E$ will have an even congruence number. Since we are assuming that $E$ has odd congruence number, $v$ must have odd order. The order of this point is $\mathrm{Num}((q+1)(p-1)/24)$. Since $q \equiv 3 \pmod 4$, $4|q+1$. If $p \equiv -3 \pmod 8$, then $v$ will have an even order, and $E$ will have an even congruence number. Therefore $p \equiv 3 \pmod 8$, and $2 \| p+1$. If $q \equiv -1 \pmod 8$, again $v$ will have an even order. Therefore, $q \equiv 3 \pmod 8$, which is the desired result. □

We also get the following corollary.

**Corollary 3.6.** *Assume that $E$ is an elliptic curve with odd congruence number and conductor $pq$ with $pq \neq 15$ or 21. Then there exist odd integers $r$ and $s$ such that $|p^r - q^s| = 16$.*

*Proof.* Following the notation of Lemma 3.2, we have $\Delta = \alpha^2\beta^2(16\alpha - \beta)^2$ for some odd integers $\alpha$ and $\beta$, coprime to each other. Assume that $\alpha^2 \neq 1$; then $|\alpha| = p^r, q^s$, or $p^r q^s$. In the last case, $\beta^2 = (16\alpha - \beta)^2 = 1$, which is not possible. Therefore assume without loss of generality that $\alpha = \pm p^r$. If $\beta = \pm q^s$, $16\alpha - \beta = \pm 1$, which leads to the Diophantine equation $\pm 16p^r - \pm q^s = \pm 1$. We get the same Diophantine equation if $\beta = \pm 1$. Therefore, we need to solve the Diophantine equation

$$q^s - 16p^r = \pm 1.$$

Since $q^s \equiv 3 \pmod 8$ for all odd $s$ and $q^s \equiv 1 \pmod{16}$ for all even $s$, we conclude that $s$ is even and

$$q^s - 16p^r = 1.$$

This leads to $(q^{s/2} - 1)(q^{s/2} + 1) = 16p^r$, and since $(q^{s/2} - 1, q^{s/2} + 1) = 2$, $q^{s/2} = 7$ or 9. Therefore $q^s = 81$, which forces $p = 5$. This is not congruent to 3 (mod 8), so we get that $\alpha = \pm 1$.

If $\beta^2 = 1$, then $|\pm 16 - \beta|$ is 15 or 17, which again contradicts $p, q \equiv 3 \pmod 8$. We get the same result if $(\pm 16 - \beta)^2 = 1$. Therefore, $\beta = \pm p^r$ and $\pm 16 - \beta = \pm q^s$. This leads to the Diophantine equation $|p^r - q^s| = 16$. Since $p, q \equiv 3 \pmod 8$, $r \equiv s \pmod 2$. If they are both even, then the difference of the two squares equals 16, which forces $N = 15$. Therefore, $r$ and $s$ are odd, which is the desired result. Finally note that in this case the elliptic curve has the model

$$E : y^2 + xy = x^3 + \frac{15 + p^r}{4}x^2 + p^r x.$$ □

**Theorem 3.7.** *Let $E$ be an elliptic curve with conductor $pq$ and odd congruence number. Then $L(E, 1) \neq 0$; hence $E$ has rank 0.*

*Proof.* For $pq = 15$ or $21$ we can check that $E$ has Mordell–Weil rank 0. Therefore assume that $pq \neq 15$ or $21$. Recall that in Corollary 2.14 we showed that

$$u = \frac{(p-1)(q^2-1)}{48}(P_1 - P_p) \quad \text{and} \quad u' = \frac{(q-1)(p^2-1)}{48}(P_1 - P_q)$$

have order two, and $\phi(u)$ and $\phi(u')$ are linearly independent, hence they generate $E[2]$. However, since $p, q \equiv 3 \pmod 8$ we get that $u$ and $u'$ are odd multiples of $P_1 - P_p$ and $P_1 - P_q$, respectively. So $\pi(P_p)$ and $\pi(P_q)$ also generate $E[2]$. Therefore, $\phi(P_p - P_q)$ is nontrivial. Applying the Atkin–Lehner involution $w_p$ to $P_p - P_q$, we get that $\phi(P_1 - P_{pq})$ is nontrivial. Therefore, $\pi(P_{pq}) \neq 0$, which implies that $L(E, 1) \neq 0$. □

***Level $N = 2p$.*** Now we take the case when $N = 2p$, for $p$ an odd prime. Specifically, we want to show that $L(E, 1) \neq 0$. In this case it seems more straightforward to prove this using analytic tools.

Specifically, let $f_E(q) = \sum a_n q^n$ be the modular form attached to the elliptic curve $E$, and let $\Omega_E$ be the real period of $E$. Note that $L(f_E, 1) \in \mathbb{R}$ since the Fourier coefficients of $f_E$ are rational integers. Therefore, the order of $\pi(P_{2p})$ is the order of $L(f_E, 1) \in \mathbb{R}/\Omega_E\mathbb{Z}$. We know that $L(f_E, s)$ has an Euler product expansion

$$L(f_E, s) = \prod_p L_p(f_E, s),$$

and $L_2(f_E, s) = \dfrac{1}{1 - a_2 2^{-s}}$. Similarly

$$\pi(P_p) = 2\pi i \int_{\frac{1}{2}}^{i\infty} f_E(z)\,dz = 2\pi i \int_0^{i\infty} f_E(z+1/2)\,dz = 2\pi i \int_0^{i\infty} \sum (-1)^n a_n q^n\,dz,$$

which implies that $\pi(P_p)$ can be written as $L(g, 1)$ where $L(g, s)$ has an Euler product expansion

$$L(g, s) = \left(-1 + \frac{a_2}{2^s} + \frac{a_4}{4^s} + \ldots\right) \prod_{p>2} L_p(f_E, s) = -\frac{1 - a_2 2^{1-s}}{1 - a_2 2^{-s}} \prod_{p>2} L_p(f_E, s).$$

Therefore $L(g, 1) = L(f_E, 1)(a_2 - 1)$, and more appropriately for us

$$\pi(P_p) \equiv (a_2 - 1)\pi(P_{2p}) \pmod{\Omega_E\mathbb{Z}}.$$

We know that if $E$ has odd congruence number, then $(w_2)_*$ is acting trivially, which implies that $a_2 = -1$. Therefore

$$\pi(P_p) \equiv -2\pi(P_{2p}) \pmod{\Omega_E\mathbb{Z}}.$$

However, we also know that $P_{2p} = w_2(P_p)$, and $\pi(w_2(P_p)) = \pi(P_p) + \alpha$ where $\alpha$ is a 2-torsion point in $E$. Since both $\pi(P_p)$ and $\pi(P_{2p})$ are equivalent to real numbers, $\alpha$ is also equivalent to a real number, which implies that $\alpha \equiv \Omega_E/2 \pmod{\Omega_E\mathbb{Z}}$. As a result we obtain successively

$$\begin{aligned}
\pi(P_p) &\equiv \pi(P_{2p}) + \Omega_E/2 \pmod{\Omega_E\mathbb{Z}},\\
-2\pi(P_{2p}) &\equiv \pi(P_{2p}) + \Omega_E/2 \pmod{\Omega_E\mathbb{Z}},\\
-3\pi(P_{2p}) &\equiv \Omega_E/2 \pmod{\Omega_E\mathbb{Z}},\\
\pi(P_{2p}) &\equiv \Omega_E\left(\tfrac{1}{3}k - \tfrac{1}{6}\right) \pmod{\Omega_E\mathbb{Z}},
\end{aligned}$$

for some integer $k$. Therefore, $\pi(P_{2p}) \neq 0$ and $L(f_E, 1) \neq 0$. We also observe that $\pi(P_{2p})$ will either be a 6-torsion point (for $k \equiv 0$ or $1 \pmod 3$), or a 2-torsion point (for $k \equiv 2 \pmod 3$).

In either case, we have an elliptic curve with conductor $2p$ and a rational 2-torsion point. Such elliptic curves were studied in [Ivorra 2004], whose results allow us to put stringent conditions on the values for $p$. Ivorra's Theorem 1 says that if $p \geq 29$, there is an integer $k \geq 4$ such that one of $p+2^k$, $p-2^k$, or $2^k - p$ is a perfect square. However, we already know from Theorem 2.15 that $p \equiv 7 \pmod{16}$. Putting these two facts together, we get that $p = 2^k - m^2$. In fact, in this case, Ivorra's result says that $7 \leq k$ is odd and our elliptic curve is isogenous to

$$y^2 + xy = x^3 + \frac{m-1}{4}x^2 + 2^{k-6}x.$$

Searching through Cremona's database, we find out that the only elliptic curves with odd modular degrees and conductors $2p$ with $p \leq 29$ are $E = 14A$ and $E = 46A$, and both of these are of the form above.

***Level $N = 4p$.*** As with the case of $N = 2p$, we can use Theorem 2 of [Ivorra 2004] to parametrize all elliptic curves with conductor $4p$ and a rational 2-torsion point. Specifically, for $p > 29$, $p = a^2 + 4$ for some integer $a \equiv 1 \pmod 4$, and $E$ is isomorphic to one of the following two isogenous elliptic curves:

$$E : y^2 = x^3 + ax^2 - x, \quad E' : y^2 = x^3 - 2ax^2 + px.$$

We can calculate the rank of such elliptic curves using a standard 2-descent. In fact, if we let $\phi : E \to E'$ and $\phi'$ be the dual isogeny, using notation from [Silverman 1992] we get

$$|S^{\phi}(E, \mathbb{Q})| = |S^{\phi'}(E, \mathbb{Q})| = 2,$$

which implies that

$$|E(\mathbb{Q})/\phi'(E'(\mathbb{Q}))| = |E'(\mathbb{Q})/\phi(E(\mathbb{Q}))| = 2,$$

which, by the exact sequence

$$0 \to E'(\mathbb{Q})[\phi']/\phi(E(\mathbb{Q}))[2] \to E(\mathbb{Q})/\phi'(E'(\mathbb{Q})) \to E(\mathbb{Q})/2E(\mathbb{Q}) \to E'(\mathbb{Q})/\phi(E(\mathbb{Q})) \to 0,$$

gives us $|E(\mathbb{Q})/2E(\mathbb{Q})| \leq 4$. This forces the rank of $E(\mathbb{Q})$ to be 0.

For $p \leq 29$, we can consult Cremona's table to get the elliptic curves $20A$, $52C$, and $116C$. In fact all these elliptic curves are of the model constructed above.

*Level* $N = 8p$. In this case, Theorem 3 of [Ivorra 2004] tells us that any elliptic curve with a rational 2-torsion point and the conductor $N = 8p$ satisfies $p \equiv a^2 \pmod{16}$ for $p > 31$. However, by Theorem 2.15, $p \equiv 3 \pmod 4$; therefore there are no elliptic curves with conductor $8p$ and odd congruence number for $p > 31$. Using Cremona's table, we know that the elliptic curve $24A$ is the only elliptic curve with the conductor $8p$ and odd congruence number. Furthermore this curve has rank 0.

We combine all of the results above:

**Theorem 3.8.** *Let $E/\mathbb{Q}$ be an elliptic curve with odd congruence number. Then one of the following is true*:

(1) *$E$ has conductor $p$ and no 2-torsion point, $E$ has supersingular reduction at 2, and $E(\mathbb{R})$ is connected.*

(2) *$E$ has conductor $p$, a rational 2-torsion point, and $p = u^2 + 64$ with $u \equiv 3 \pmod 8$ ($E$ is a Neumann–Setzer curve in this case).*

(3) *$E$ has conductor $2p$ and $p = 2^k - m^2$ for some odd integer $7 \leq k$ and integer $m$, and $E$ is isogenous to*

$$y^2 + xy = x^3 + \frac{m-1}{4}x^2 + 2^{k-6}x.$$

(4) *$E$ has conductor $4p$ and $p = m^2 + 4$ for some integer $m \equiv 1 \pmod 4$, and $E$ is isogenous to*

$$y^2 = x^3 + mx^2 - x.$$

(5) *$E$ has conductor $pq$ with $p$ and $q$ odd primes satisfying $p \equiv q \equiv 3 \pmod 8$ and $p^r - q^s = 16$ for odd integers $r$ and $s$, and $E$ is isogenous to*

$$y^2 + xy = x^3 + \frac{p^r + 15}{4}x^2 + p^r x.$$

(6) *$E$ is one of the exceptional curves $11A$, $15A$, $17A$, $19A$, $21A$, $24A$, $27A$, $32A$, $36A$, $37B$, $49A$, $243B$.*

*In all these cases, $E$ has rank 0, except possibly in case (1). In this case, we know that $E$ has an even analytic rank.*

All the curves in case (6) in the theorem have a nontrivial torsion point. Therefore we have proved that if $E$ has odd congruence number and has a nontrivial torsion point, it has rank 0. Also note that for all the cases above, except for (1), we construct a family of elliptic curves with all the desired torsion structures and conductors. We expect that all these elliptic curves have odd congruence number. This can be proved if, for example, we show that $J_0(N)[\mathfrak{m}] \to E[2]$ is injective and $J[\mathfrak{m}] = C_N[\mathfrak{m}]$. When $E$ is a Neumann–Setzer curve, the results of [Mazur 1977; Mestre and Oesterlé 1989] prove this result. We expect that similar results are true for the other cases; however we, do not yet know of a proof.

## Acknowledgements

This paper would not have been possible without the help of my advisor, Ken Ribet. Specifically, many of the results in Section 2D were suggested to me by him. I thank Frank Calegari, Matt Emerton, William Stein, and Jared Weinstein, with whom I have had many discussions. Manfred Kolster and Romyar Sharifi gave me useful feedback on the first draft of this article. I thank Jovanca Buac for her careful reading of this paper and all her suggestions. Finally, I would like to thank the referees for helpful comments.

## References

[Agashe et al. 2008] A. Agashe, K. A. Ribet, and W. A. Stein, "The modular degree, congruence primes, and multiplicity one", preprint, 2008, available at http://modular.math.washington.edu/papers/ars-congruence/.

[Breuil et al. 2001] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, "On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises", *J. Amer. Math. Soc.* **14**:4 (2001), 843–939. MR 2002d:11058 Zbl 0982.11033

[Calegari and Emerton 2009] F. Calegari and M. Emerton, "Elliptic curves of odd modular degree", *Israel J. Math.* **169** (2009), 417–444. MR 2010k:11092 Zbl 05508747

[Chua and Ling 1997] S.-K. Chua and S. Ling, "On the rational cuspidal subgroup and the rational torsion points of $J_0(pq)$", *Proc. Amer. Math. Soc.* **125**:8 (1997), 2255–2263. MR 98f:11065 Zbl 0891.11036

[Cremona] J. Cremona, "Elliptic curve data", database, available at http://www.warwick.ac.uk/staff/J.E.Cremona/ftp/data/INDEX.html.

[Dummigan 2006] N. Dummigan, "On a conjecture of Watkins", *J. Théor. Nombres Bordeaux* **18**:2 (2006), 345–355. MR 2007j:11072 Zbl 1161.11351

[Emerton 2003] M. Emerton, "Optimal quotients of modular Jacobians", *Math. Ann.* **327**:3 (2003), 429–458. MR 2005g:11100 Zbl 1061.11018

[Frey 1997] G. Frey, "On ternary equations of Fermat type and relations with elliptic curves", pp. 527–548 in *Modular forms and Fermat's last theorem* (Boston, 1995), edited by G. Cornell et al., Springer, New York, 1997. MR 1638494 Zbl 0976.11027

[Gross and Zagier 1986] B. H. Gross and D. B. Zagier, "Heegner points and derivatives of $L$-series", *Invent. Math.* **84**:2 (1986), 225–320. MR 87j:11057 Zbl 0608.14019

[Ivorra 2004] W. Ivorra, "Courbes elliptiques sur $\mathbb{Q}$, ayant un point d'ordre 2 rationnel sur $\mathbb{Q}$, de conducteur $2^N p$", *Dissertationes Math. (Rozprawy Mat.)* **429** (2004), 1–55. MR 2006h:11056 Zbl 1076.11037

[Kolyvagin 1988] V. A. Kolyvagin, "Finiteness of $E(\mathbf{Q})$ and SH$(E, \mathbf{Q})$ for a subclass of Weil curves", *Izv. Akad. Nauk SSSR Ser. Mat.* **52**:3 (1988), 522–540, 670–671. In Russian; translated in *Math. USSR, Izv.* **32**:3 (1989), 523–541. MR 89m:11056

[Ligozat 1975] G. Ligozat, *Courbes modulaires de genre* 1, Mem. Soc. Math. de France (old ser.) **43**, Société Mathématique de France, Paris, 1975. MR 54 #5121 Zbl 0322.14011

[Ling 1997] S. Ling, "On the $\mathbb{Q}$-rational cuspidal subgroup and the component group of $J_0(p^r)$", *Israel J. Math.* **99** (1997), 29–54. MR 98e:11076 Zbl 0934.14022

[Mazur 1977] B. Mazur, "Modular curves and the Eisenstein ideal", *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 33–186. MR 80c:14015 Zbl 0394.14008

[Mestre and Oesterlé 1989] J.-F. Mestre and J. Oesterlé, "Courbes de Weil semi-stables de discriminant une puissance $m$-ième", *J. Reine Angew. Math.* **400** (1989), 173–184. MR 90g:11078 Zbl 0693.14004

[Murty 1999] M. R. Murty, "Bounds for congruence primes", pp. 177–192 in *Automorphic forms, automorphic representations, and arithmetic* (Fort Worth, TX, 1996), Proc. Sympos. Pure Math. **66**, Amer. Math. Soc., Providence, RI, 1999. MR 2000g:11038 Zbl 0933.11024

[Ogg 1974] A. P. Ogg, "Hyperelliptic modular curves", *Bull. Soc. Math. France* **102** (1974), 449–462. MR 51 #514 Zbl 0314.10018

[Ribet 1980] K. A. Ribet, "Twists of modular forms and endomorphisms of abelian varieties", *Math. Ann.* **253**:1 (1980), 43–62. MR 82e:10043 Zbl 0421.14008

[Ribet 1981] K. A. Ribet, "Endomorphism algebras of abelian varieties attached to newforms of weight 2", pp. 263–276 in *Seminaire de Théorie de Nombres* (Paris, 1979–80), edited by M.-J. Bertin, Progr. Math. **12**, Birkhäuser, Boston, 1981. MR 82m:10044 Zbl 0467.14006

[Shimura 1971] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publ. Math. Soc. Japan **11**, Princeton Univ. Press, Princeton, NJ, 1971. MR 47 #3318 Zbl 0872.11023

[Silverman 1992] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1992. MR 95m:11054

[Stein and Watkins 2002] W. A. Stein and M. Watkins, "A database of elliptic curves: first report", pp. 267–275 in *Algorithmic number theory* (Sydney, 2002), Lecture Notes in Comput. Sci. **2369**, Springer, Berlin, 2002. MR 2005h:11113 Zbl 1058.11036

[Stein and Watkins 2004] W. Stein and M. Watkins, "Modular parametrizations of Neumann–Setzer elliptic curves", *Int. Math. Res. Not.* **2004**:27 (2004), 1395–1405. MR 2005c:11070 Zbl 1088.11043

[Taylor and Wiles 1995] R. Taylor and A. Wiles, "Ring-theoretic properties of certain Hecke algebras", *Ann. of Math.* (2) **141**:3 (1995), 553–572. MR 96d:11072 Zbl 0823.11030

[Watkins 2002] M. Watkins, "Computing the modular degree of an elliptic curve", *Experiment. Math.* **11**:4 (2002), 487–502. MR 2004c:11091 Zbl 1162.11349

[Wiles 1995] A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141**:3 (1995), 443–551. MR 96d:11071 Zbl 0823.11029

Communicated by Barry Mazur
Received 2009-11-23 Revised 2010-09-17 Accepted 2010-12-05

syazdani@math.mcmaster.ca *Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4L8, Canada*

ALGEBRA AND NUMBER THEORY 5:1(2011)

msp

# Group algebra extensions of depth one

Robert Boltje and Burkhard Külshammer

A ring extension $A \subseteq B$ is said to have depth one if $B$ is isomorphic to a direct summand of $A^n$ as an $(A, A)$-bimodule, for some positive integer $n$. We prove group-theoretic characterizations of this property in the case $kH \subseteq kG$, where $H$ is a subgroup of a finite group $G$ and $k$ is a field. We determine when the source algebra of a block of $kG$ with defect group $P$ is a depth-one extension of $kP$.

## Introduction

A *ring extension* is a unitary ring homomorphism $f: A \to B$ between two rings $A$ and $B$. In this situation, the ring $B$ can be viewed as an $(A, A)$-bimodule using the map $f$. A ring extension $f: A \to B$ is said to be of *depth one* (or *centrally projective* [Kadison 1999]) if $B$ is isomorphic, as an $(A, A)$-bimodule, to a direct summand of $A^n$ for some positive integer $n$. We write $B \mid A^n$ for this condition. Whenever $A$ is a unitary subring of $B$ and $f: A \to B$ is the inclusion map, we denote the corresponding ring extension by $A \subseteq B$.

To the best of our knowledge, centrally projective ring extensions were first considered by Hirata [1969]. The identification of centrally projective ring extensions with ring extensions of depth one appeared in [Kadison and Szlachányi 2003]. Ring extensions of higher depth were studied in [Kadison 2008], for example.

In this paper we try to answer the question of when a ring extension $kH \subseteq kG$ of group rings has depth one. Here and throughout this paper we denote by $k$ a commutative ring, by $G$ a group and by $H$ a subgroup of finite index in $G$. In [Boltje and Külshammer 2010] we considered the question of when the ring extension $kH \subseteq kG$ has depth two (that is, when $kG \otimes_{kH} kG \mid (kG)^n$ as $(kG, kH)$-bimodules, or equivalently as $(kH, kG)$-bimodules, for some positive integer $n$). It turns out that this is equivalent to $H$ being normal in $G$, independently of $k$. If the ring extension $kH \subseteq kG$ has depth one it also has depth two, since one can apply the functor $kG \otimes_{kH} -$ to the relation $kG \mid (kH)^n$. In particular, $H$ has to be normal in $G$. The converse is not true in general, as our main results, Theorems 1.7 and 1.9,

*MSC2000:* primary 20C05; secondary 19A22, 16D90, 16D20, 20C20.

*Keywords:* depth-one ring extension, centrally projective ring extension, depth-two ring extension, symmetric Frobenius extension, $p$-hypoelementary group, nilpotent block, source algebra, trivial source module.

show. In these theorems we have to assume that $k$ is a field (or a complete discrete valuation ring of characteristic 0 and positive residual characteristic $p$). In both cases the depth-one condition is equivalent to a purely group theoretic condition on the inclusion $H \leqslant G$, namely that $G = HC_G(X)$, for every cyclic subgroup $X$ of $H$ in the characteristic 0 case, and for every $p$-hypoelementary subgroup $X$ of $H$ in the modular case (see Remark 1.8 for a definition of $p$-hypoelementary groups). Therefore, the group theoretic depth-one condition does depend on the base ring $k$. We do not have a group theoretic reformulation in the case $k = \mathbb{Z}$.

At the end we study the depth-one condition for the ring extension $FP \to A$, where $P$ is the defect group of a block of a group algebra over a field $F$ of positive characteristic $p$, and $A$ is a source algebra of the block.

## 1. Depth one for group algebra extensions

**1.1.** Assume that $g$ belongs to $N_G(H)$, the normalizer of $H$ in $G$. Then

$$c_{g^{-1}} : kH \to kH, \quad a \mapsto g^{-1}ag$$

is a $k$-algebra automorphism of $kH$. Restriction along this automorphism defines a functor

$$C_g : {}_{kH}\mathsf{Mod} \to {}_{kH}\mathsf{Mod}, \quad M \mapsto {}^gM$$

on the category of left $kH$-modules. More explicitly, for $M \in {}_{kH}\mathsf{Mod}$, the left $kH$-module ${}^gM$ is defined to be equal to $M$ as an abelian group, and it is endowed with the module structure $a * m := (g^{-1}ag) \cdot m$, where "$\cdot$" denotes the original $kH$-module structure of $M$. The functor $C_g$ maps a homomorphism $f : M_1 \to M_2$ in ${}_{kH}\mathsf{Mod}$ to ${}^gf := f : {}^gM_1 \to {}^gM_2$.

If $C_g$ is naturally equivalent to the identity functor on ${}_{kH}\mathsf{Mod}$, we say that $g$ *acts trivially on* ${}_{kH}\mathsf{Mod}$. For this paper, we say that $G$ *acts trivially on* ${}_{kH}\mathsf{Mod}$ if $H$ is normal in $G$ and $g$ acts trivially on ${}_{kH}\mathsf{Mod}$ for every $g \in G$. Note that $H$ acts trivially on ${}_{kH}\mathsf{Mod}$. This is also an immediate consequence of the next proposition.

The subset $kgH = gkH = kHg$ of $kG$ is a $(kH, kH)$-subbimodule of $kG$. It is isomorphic to ${}^{(g,1)}kH$ if we view $(kH, kH)$-bimodules $M$ as left $k[H \times H]$-modules via

$$(h_1, h_2) \cdot m := h_1 m h_2^{-1} \quad \text{for } h_1, h_2 \in H \text{ and } m \in M.$$

Assume that $R$ is an arbitrary ring and that $\alpha$ is an automorphism of $R$. Let $R_\alpha$ denote the $(R, R)$-bimodule that equals $R$ as abelian group but has the twisted action $axb := \alpha(a)xb$, for $a, x, b \in R$. It is well-known and straightforward to prove that $\alpha$ is an inner automorphism if and only if $R_\alpha$ is isomorphic to $R$ as an $(R, R)$-bimodule. It is also equivalent to $\alpha$ acting trivially on ${}_R\mathsf{Mod}$. The following proposition is a special case, and we leave its proof to the reader.

**Proposition 1.2.** *For $g \in N_G(H)$, the following are equivalent*:

(i) *$kgH \cong kH$ as $(kH, kH)$-bimodules.*

(ii) *There exists a unit $u$ of $kH$ such that $gag^{-1} = uau^{-1}$ for all $a \in kH$.*

(iii) *$g$ acts trivially on ${}_{kH}\mathrm{Mod}$.*

For every subset $X$ of $H$ we denote by $C_G(X)$ the centralizer of $X$ in $G$.

**Corollary 1.3.** *If $G = HC_G(H)$, then the ring extension $kH \subseteq kG$ has depth one. Conversely, if the ring extension $kH \subseteq kG$ has depth one, then $H$ is normal in $G$.*

*Proof.* Suppose first that $G = HC_G(H)$. Then every $g \in G$ satisfies condition (ii). Using condition (i) together with the decomposition $kG = \bigoplus_{gH \in G/H} kgH$ into $(kH, kH)$-subbimodules, the first assertion follows. The second assertion was already observed in the introduction. □

**1.4.** If $\Lambda$ is a $k$-order (that is, a $k$-algebra that is finitely generated and projective as a $k$-module) we say that the *Krull–Schmidt theorem holds* for $\Lambda$-lattices if the following two properties hold for every $\Lambda$-module $M$ that is finitely generated and projective as a $k$-module:

- $M$ has a decomposition $M = U_1 \oplus \cdots \oplus U_r$ into indecomposable $\Lambda$-submodules, and
- if $M = U_1 \oplus \cdots \oplus U_r = V_1 \oplus \cdots \oplus V_s$ are two decompositions into indecomposable $\Lambda$-submodules, then $r = s$ and there exists a permutation $\sigma$ of $\{1, \ldots, r\}$ such that $U_i \cong V_{\sigma(i)}$ for all $i \in \{1, \ldots, r\}$.

If $k$ is a field or a complete discrete valuation ring, then the Krull–Schmidt theorem holds for every $k$-order $\Lambda$ [Curtis and Reiner 1981, Theorem 6.12].

**Proposition 1.5.** *Assume that $G$ is finite and that the Krull–Schmidt theorem holds for $k[H \times H]$-lattices. The following are equivalent*:

(i) *$kH \subseteq kG$ is a ring extension of depth one.*

(ii) *$H$ is normal in $G$ and $kgH \cong kH$ as $(kH, kH)$-bimodules for every $g \in G$.*

(iii) *$H$ is normal in $G$ and $kG \cong (kH)^{[G:H]}$ as $(kH, kH)$-bimodules.*

(iv) *$kG \cong (kH)^{[G:H]}$ as $(kH, kH)$-bimodules.*

*Proof.* (i) ⇒ (ii): Since the ring extension $kH \subseteq kG$ has depth one, Corollary 1.3 implies that $H$ is normal in $G$. Therefore, for every $g \in G$, one has $kgH \mid kG \mid (kH)^n$ as $(kH, kH)$-bimodules for some positive integer $n$. Since the Krull–Schmidt theorem holds for $k[H \times H]$-lattices, every indecomposable direct summand of the $k[H \times H]$-module $kgH$ is isomorphic to an indecomposable direct summand of the $k[H \times H]$-module $kH$. But the indecomposable direct summands of $kH$ are the blocks of $kH$, and they are pairwise nonisomorphic. Since $kgH$ is isomorphic to

${}^{(g,1)}kH$ and since $C_{(g,1)} : {}_{k[H\times H]}\mathrm{Mod} \to {}_{k[H\times H]}\mathrm{Mod}$ is a category equivalence, the indecomposable direct summands of $kgH$ are also pairwise nonisomorphic. Thus, we can decompose $kgH$ and $kH$ multiplicity-free into a direct sum of indecomposable $k[H \times H]$-submodules. The number of these summands coincides, since $kgH \cong {}^{(g,1)}kH$. Since every summand of $kgH$ occurs as a summand of $kH$, we can conclude that $kgH \cong kH$ as $(kH, kH)$-bimodules.

(ii) $\Rightarrow$ (iii): This follows from the decomposition $kG = \bigoplus_{gH\in G/H} kgH$ into $(kH, kH)$-subbimodules.

(iii) $\Rightarrow$ (iv): This is trivial.

(iv) $\Rightarrow$ (i): This is immediate from the definition of depth one. □

**Remark 1.6.** In the proof of Proposition 1.5, the Krull–Schmidt property is only used for permutation $k[H \times H]$-modules. By [Boltje and Glesser 2007, Theorem 1.6], the four conditions in Proposition 1.5 are still equivalent when $k$ is a local domain containing a root of unity of order $e$, where $e$ is defined as the exponent $\exp(H)$ of $H$ in the case that $k$ has characteristic 0, and as the $p'$-part of $\exp(H)$ in the case that $k$ has positive characteristic $p$.

Next we will study the case where $k$ is a field of characteristic 0 and $H$ is a finite group. In this case we will denote by $\overline{k}$ an algebraic closure of $k$ and by $\mathrm{Irr}(H)$ the set of irreducible characters of $\overline{k}H$. Recall that $N_G(H)$ acts from the left on $\mathrm{Irr}(H)$ via $\chi \mapsto {}^g\chi$ for $g \in N_G(H)$, where ${}^g\chi(h) = \chi(g^{-1}hg)$, for $h \in H$. Recall also that $N_G(H)$ acts on the set of conjugacy classes of $H$ via $\mathcal{K} \mapsto {}^g\mathcal{K}$, where ${}^g\mathcal{K} = \{ghg^{-1} \mid h \in \mathcal{K}\}$ for a conjugacy class $\mathcal{K}$ of $H$.

**Theorem 1.7.** *Assume that $k$ is a field of characteristic 0 and that $G$ is finite. The following are equivalent*:

(i) *$kH \subseteq kG$ is a ring extension of depth one.*

(ii) *$H$ is normal in $G$ and $G$ acts trivially on* $\mathrm{Irr}(H)$.

(iii) *$H$ is normal in $G$ and $G$ acts trivially on the set of conjugacy classes of $H$.*

(iv) *For every cyclic subgroup $X$ of $H$ one has $G = HC_G(X)$.*

*Proof.* (i) $\Longleftrightarrow$ (ii): Using Corollary 1.3, we may assume that $H$ is normal in $G$. Using Propositions 1.2 and 1.5, one sees that it suffices to show that, for every $g \in G$, the $(kH, kH)$-bimodules $kgH$ and $kH$ are isomorphic if and only if $g$ acts trivially on $\mathrm{Irr}(H)$. By the Deuring–Noether theorem [Nagao and Tsushima 1989, Theorem II.3.1], $kgH \cong kH$ as $(kH, kH)$-bimodules if and only if $\overline{k}gH \cong \overline{k}H$ as $(\overline{k}H, \overline{k}H)$-bimodules. By Proposition 1.2, the latter is equivalent to $g$ acting trivially on ${}_{\overline{k}H}\mathrm{Mod}$, which implies that $g$ acts trivially on $\mathrm{Irr}(H)$. Conversely, if ${}^g\chi = \chi$ for every $\chi \in \mathrm{Irr}(H)$ then $\overline{k}gH \cong \overline{k}H$ as $\overline{k}[H \times H]$-modules, since the

character of $\overline{k}H$ is equal to $\sum_{\chi\in\mathrm{Irr}(H)}\chi\times\overline{\chi}$ and the character of $\overline{k}gH$ is equal to

$$^{(g,1)}\Big(\sum_{\chi\in\mathrm{Irr}(H)}\chi\times\overline{\chi}\Big)=\sum_{\chi\in\mathrm{Irr}(H)}{}^{g}\chi\times\overline{\chi}=\sum_{\chi\in\mathrm{Irr}(H)}\chi\times\overline{\chi}.$$

(ii) $\Longleftrightarrow$ (iii): This follows immediately from Brauer's permutation lemma [Nagao and Tsushima 1989, Lemma III.2.19].

(iii) $\Rightarrow$ (iv): Let $X=\langle x\rangle$ be a cyclic subgroup of $H$ and let $g$ be an element of $G$. Since $gxg^{-1}$ lies in the same conjugacy class as $x$, there exists $h\in H$ such that $gxg^{-1}=hxh^{-1}$. This implies $h^{-1}g\in C_G(x)$ and $g\in hC_G(x)\subseteq HC_G(X)$.

(iv) $\Rightarrow$ (ii): Condition (iv) implies immediately that $H$ is normal in $G$. Now let $g\in G$, $\chi\in\mathrm{Irr}(H)$ and $x\in H$. Then there exists $h\in H$ and $c\in C_G(x)$ such that $g=hc$. Hence $\chi(gxg^{-1})=\chi(hcxc^{-1}h^{-1})=\chi(hxh^{-1})=\chi(x)$. Thus ${}^{g^{-1}}\chi=\chi$ and $\chi={}^{g}\chi$. □

**Remark 1.8.** Next we study the depth-one condition for the ring extension $kH\subseteq kG$ in the case where $k$ is a field of positive characteristic $p$, or where $k$ is a complete discrete valuation ring of characteristic 0 and positive residual characteristic $p$. We will need the theory of species developed by Benson and Parker [1998, Section 5.5]. For this remark assume that $G$ is finite and that $k$ contains a root of unity whose order is equal to $\exp(G)$ if $k$ has characteristic 0, and to the $p'$-part of $\exp(G)$ if $k$ has characteristic $p$. Let $S$ and $T$ be finite left $G$-sets. We denote the corresponding permutation $kG$-modules by $kS$ and $kT$. The goal of this remark is to derive a criterion for $kS$ being isomorphic to $kT$. Denote by $\mathcal{H}_p(G)$ the set of $p$-hypoelementary subgroups $E$ of $G$, that is, subgroups $E$ that have a normal (Sylow) $p$-subgroup $P$ such that $E/P$ is cyclic. We claim that

$$kS\cong kT \text{ as } kG\text{-modules} \iff |S^E|=|T^E| \text{ for all } E\in\mathcal{H}_p(G), \tag{1.8.a}$$

where $|S^E|$ denotes the cardinality of the set $S^E$ of $E$-fixed points on $S$. In order to see this equivalence, it suffices to show that

$$s_{E,g}(kS)=|S^E| \tag{1.8.b}$$

for all $E\in\mathcal{H}_p(G)$ and all $p'$-elements $g\in E$ such that $gP$ generates $E/P$, where $P$ denotes the Sylow $p$-subgroup of $E$. For a definition of $s_{E,g}$ see [Benson 1998, Section 5.5]. Since $s_{E,g}(kS)=s_{E,g}(\mathrm{Res}^G_E(kS))$, we may assume that $G=E$. Moreover, since $s_{E,g}$ is additive, we may assume that $S$ is a transitive $E$-set, that is, $S=E/D$ for some subgroup $D$ of $E$. Then $kS\cong\mathrm{Ind}_D^E(k)$. If $P$ is not contained in $D$, then no indecomposable direct summand of $kS$ has vertex $P$, and both sides of Equation (1.8.b) are equal to 0. If $P\leqslant D<E$, the Brauer species of $\mathrm{Ind}_D^E(k)$ at $g$ equals 0, since $g\notin D$, and again both sides in Equation (1.8.b) are equal to 0. Finally, if $D=E$, it is immediate that both sides of the equation are equal to 1.

In the next theorem we will apply the criterion in (1.8.a) to the $H \times H$-sets $gH$ and $H$ for $g \in N_G(H)$.

For a subgroup $X$ of $H$, we set $\Delta X := \{(x, x) \mid x \in X\} \leqslant H \times H$.

**Theorem 1.9.** *Assume that $G$ is finite and that $k$ is a field of characteristic $p > 0$ or a complete discrete valuation ring of characteristic 0 with residual characteristic $p > 0$. The following are equivalent*:

(i) *The ring extension $kH \subseteq kG$ has depth one.*

(ii) *$H$ is normal in $G$, and $|(gH)^E| = |H^E|$ for all $g \in G$ and $E \in \mathcal{H}_p(H \times H)$.*

(iii) *$H$ is normal in $G$, and $|(gH)^{\Delta X}| = |H^{\Delta X}|$ for all $g \in G$ and $X \in \mathcal{H}_p(H)$.*

(iv) *For all $X \in \mathcal{H}_p(H)$ one has $G = C_G(X)H$.*

*Proof.* (i) $\Longleftrightarrow$ (ii): By Corollary 1.3, we may assume that $H$ is normal in $G$. Now the equivalence of (i) and (ii) follows immediately from Proposition 1.5 and Remark 1.8 applied to the $k[H \times H]$-modules $kgH$ and $kH$, for $g \in G$. In fact, by [Benson 1998, Corollary 3.11.4(i)] and the Deuring–Noether theorem [Nagao and Tsushima 1989, Theorem II.3.1], we have $kgH \cong kH$ as $k[H \times H]$-modules if and only if $k'gH \cong k'H$ as $k'[H \times H]$-modules, where $k'$ is obtained from $k$ by adjoining a root of unity whose order is equal to $\exp(H)$ if $k$ has characteristic 0, and to the $p'$-part of $\exp(H)$ if $k$ has characteristic $p$.

(ii) $\Rightarrow$ (iii): This is trivial.

(iii) $\Rightarrow$ (iv): Let $g \in G$ and let $X \in \mathcal{H}_p(H)$. Since $1 \in H^{\Delta X}$, the set $(gH)^{\Delta X}$ is nonempty. Let $h \in H$ such that $gh \in (gH)^{\Delta X}$. Then $gh \in C_G(X)$ and $g \in C_G(X)H$.

(iv) $\Rightarrow$ (ii): From (iv) we have immediately that $H$ is normal in $G$. Now let $g \in G$ and let $E \in \mathcal{H}_p(H)$. The $H \times H$-sets $H$ and $gH$ are transitive. The stabilizer of $1 \in H$ is $\Delta H$ and the stabilizer of $g \in gH$ is ${}^{(g,1)}\Delta H$. One has $|H^E| = 0 = |(gH)^E|$ unless $E$ is $H \times H$-conjugate to a subgroup of $\Delta H$ or ${}^{(g,1)}\Delta H$. Since the number of fixed points does not change if we replace $E$ by an $H \times H$-conjugate of $E$, we may assume that $E \leqslant \Delta H$ or $E \leqslant {}^{(g,1)}\Delta H$. We first assume that $E \leqslant \Delta H$. Then $E = \Delta X$ for some $X \in \mathcal{H}_p(H)$. Since $g \in C_G(X)H$, we can write $g = ch$ with $c \in C_G(X)$ and $h \in H$. Then $gH = cH$, and for $h' \in H$ we have

$$h' \in H^{\Delta X} \Longleftrightarrow h' \in C_G(X) \Longleftrightarrow ch' \in C_G(X) \Longleftrightarrow ch' \in (cH)^{\Delta X}.$$

It follows that $|H^{\Delta X}| = |(cH)^{\Delta X}| = |(gH)^{\Delta X}|$. Finally, if $E \leqslant {}^{(g,1)}\Delta H$, then ${}^{(g^{-1},1)}E = \Delta X$ for some $X \in \mathcal{H}_p(H)$. Again we can write $g = ch$ with $c \in C_G(X)$ and $h \in H$. Then $g = h'c$ with $h' = ghg^{-1} \in H$ and $E = {}^{(g,1)}(\Delta X) = {}^{(h'c,1)}(\Delta X) = {}^{(h',1)}(\Delta X)$ is $H \times H$-conjugate to $\Delta X$. By the first case, this implies

$$|(gH)^E| = |(gH)^{\Delta X}| = |H^{\Delta X}| = |H^E|. \qquad \square$$

**Remark 1.10.** In the case $k = \mathbb{Z}$, we do not know if there is a similar equivalence (i) $\Longleftrightarrow$ (iv) as in Theorem 1.9 with $\mathcal{H}_p(H)$ replaced by some other set $\mathcal{S}(H)$ of subgroups of $H$. Even if there existed such a set $\mathcal{S}(H)$, we don't have a good guess what it should be.

If $\mathbb{Z}H \subseteq \mathbb{Z}G$ has depth one, then $kH \subseteq kG$ has depth one for every commutative ring $k$ (by scalar extension). In particular, this implies that $G = HC_G(X)$ for every $p$-hypoelementary subgroup $X$ of $H$ for all primes $p$. We do not know if the converse holds. On the other hand, if $G = HC_G(H)$, then $\mathbb{Z}H \subseteq \mathbb{Z}G$ has depth one by Corollary 1.3. However, the converse is not true. In fact, by [Hertweck 2001, Theorem A], there exist a finite group $H$ (metabelian of order $2^{25} \cdot 97^2$), a noninner automorphism $g$ of $H$, and a unit $u$ of $\mathbb{Z}H$ with $g(a) = uau^{-1}$. We set $G := H \rtimes \langle g \rangle$. By Proposition 1.2, we obtain $\mathbb{Z}g^i H \cong \mathbb{Z}H$ as $(\mathbb{Z}H, \mathbb{Z}H)$-bimodules for every integer $i$. This implies that $\mathbb{Z}G = \bigoplus_{xH \in G/H} \mathbb{Z}xH \cong (\mathbb{Z}H)^{[G:H]}$ and that $\mathbb{Z}H \subseteq \mathbb{Z}G$ has depth one. But $g \notin C_G(H)H$, since $g$ is not an inner automorphism of $H$. This shows that if, for each finite group $H$, there exists a set of subgroups $\mathcal{S}(H)$ of $H$ that replaces $\mathcal{H}_p(H)$ in Theorem 1.9(iv) in the case $k = \mathbb{Z}$, then $H \notin \mathcal{S}(H)$ for Hertweck's group $H$.

## 2. Depth one for source algebras of blocks

**2.1.** Let $G$ be a finite group, let $p$ be a prime, and let $(K, R, F)$ be a $p$-modular system. Thus, $R$ is a complete discrete valuation ring of characteristic zero, $K$ is the field of fractions of $R$, and $F$, the residue field of $R$, has characteristic $p$. We assume that $R$ contains a root of unity of order $\exp(G)$ and that $F$ is algebraically closed. Then $K$ and $F$ are splitting fields for $KG$ and $FG$, respectively. For an $R$-order $A$, we denote by $\overline{A}$ the finite-dimensional $F$-algebra $F \otimes_R A$. In the following, let $k \in \{R, F\}$.

In this section, we will consider the depth-one condition for blocks and source algebras. For general background, we refer to the books [Thévenaz 1995] and [Külshammer 1991]. For the convenience of the reader, we recall some of the basic concepts.

An *interior $G$-algebra* over $k$ consists of a $k$-order $A$ and a group homomorphism $i\colon G \to A^\times$, where $A^\times$ denotes the group of units of $A$. In this case, we will consider the $k$-linear extension $kG \to A$ of $i$ as a ring extension. Two interior $G$-algebras $A_1$ and $A_2$ are called *isomorphic* if there exists an isomorphism $f\colon A_1 \to A_2$ commuting with the structural maps $i_1\colon G \to A_1^\times$ and $i_2\colon G \to A_2^\times$.

If $A$ is an interior $G$-algebra, then a *point* of a subgroup $H$ of $G$ on $A$ is an $(A^H)^\times$-conjugacy class $\beta$ of primitive idempotents in the subalgebra

$$A^H := \{a \in A \mid ha = ah \text{ for all } h \in H\}$$

of $A$. In this case the pair $(H, \beta) =: H_\beta$ is called a *pointed group* on $A$.

The point $\beta$ of $H$ on $A$ is called *local* if $\beta \not\subseteq \mathrm{Tr}_L^H(A^L)$ for every proper subgroup $L$ of $H$; here $\mathrm{Tr}_L^H : A^L \to A^H$, $a \mapsto \sum_{hL \in H/L} hah^{-1}$ is the *relative trace map*. If $\beta$ is a local point of $H$ on $A$ then $H_\beta$ is called a *local pointed group* on $A$. One can show that in this case $H$ has to be a $p$-group.

Let $H_\beta$ and $L_\gamma$ be pointed groups on $A$. We write $L_\gamma \leq H_\beta$ if $L \leqslant H$ and $jAj \subseteq iAi$ for suitable idempotents $i \in \beta$, $j \in \gamma$. This defines a partial order on the set of pointed groups on $A$. The group $G$ acts by conjugation on the set of all pointed groups $H_\beta$ on $A$, and this action is compatible with the partial order relation. We denote by $N_G(H_\beta)$ the stabilizer of $H_\beta$ in $G$. Thus, $N_G(H_\beta)$ is a subgroup of $N_G(H)$.

A *block* of $kG$ is an indecomposable direct summand $B$ of $kG$, considered as a $(kG, kG)$-bimodule. In this case $B$ is a $k$-order in its own right. We consider $B$ as an interior $G$-algebra via the group homomorphism $G \to B^\times$, $g \mapsto g1_B = 1_B g$. Then $\alpha := \{1_B\}$ is a point of $G$ on $B$ and we consider $G_\alpha$ as a pointed group on $B$.

The maximal local pointed groups $P_\gamma \leqslant G_\alpha$ are called *defect pointed groups* of $G_\alpha$ (and of $B$). They are unique up to conjugation in $G$. If $P_\gamma$ is a defect pointed group on $B$, then $P$ is also called a *defect group* of $B$. For $i \in \gamma$, the $k$-order $B_\gamma = iBi = ikGi$ is called a *source algebra* of $B$. One can show that $BiB = B$, so that $B$ and $iBi$ are Morita equivalent $k$-orders via multiplication with $i$. The source algebra $iBi$ will always be considered as an interior $P$-algebra via the map $P \to (iBi)^\times$, $x \mapsto ix = xi$.

The block $B$ is called *nilpotent* if $N_G(Q_\delta)/C_G(Q)$ is a $p$-group for every local pointed group $Q_\delta \leq G_\alpha$ on $B$. (Note that indeed $C_G(Q) \subseteq N_G(Q_\delta)$ here.) Puig [1988] determined the structure of the source algebra of a nilpotent block. It is a consequence of his results that every nilpotent block has a unique simple module in characteristic $p$, up to isomorphism. We will make use of Puig's results in the following theorem.

**Theorem 2.2.** *Let $B$ be a block of $RG$ with defect pointed group $P_\gamma$, and let $B_\gamma$ be a corresponding source algebra. Then the following assertions are equivalent*:

(i) *The ring extension $FP \to \overline{B_\gamma}$ defined by the canonical map $P \to \overline{B_\gamma}^\times$ has depth one.*

(ii) *$B_\gamma$ and $RP$ are isomorphic as interior $P$-algebras.*

(iii) *$B$ is a nilpotent block, and the unique simple $\overline{B}$-module $M$ has a trivial source.*

*Proof.* (i) $\Rightarrow$ (ii): Suppose that the ring extension $FP \to \overline{B_\gamma}$ has depth one. Then $\overline{B_\gamma} \mid (FP)^n$ as an $(FP, FP)$-bimodule, for some positive integer $n$. Thus every indecomposable direct summand of the $(FP, FP)$-bimodule $\overline{B_\gamma}$ is isomorphic to $FP$. Hence [Thévenaz 1995, Theorem 44.3] implies that $N_G(P_\gamma) = PC_G(P)$ and $\overline{B_\gamma} \cong FP$, as an $(FP, FP)$-bimodule; in particular, we have $\mathrm{rk}_R(B_\gamma) = \dim_F \overline{B_\gamma} = |P|$. The same theorem now implies that $B_\gamma \cong RP$ as interior $P$-algebras.

(ii) $\Rightarrow$ (iii): Suppose that $B_\gamma$ and $RP$ are isomorphic interior $P$-algebras. Then a result by Puig [1988, Theorem 1.6] implies that the block $B$ is nilpotent [Thévenaz 1995, Remark 50.10]. We write $\overline{B_\gamma} = iFGi$, where $i$ is a primitive idempotent in $(FG)^P$. Since every block has at least one simple module whose vertices are defect groups of the block, $P$ is a vertex of the unique simple $\overline{B}$-module $M$. By [Thévenaz 1995, Proposition 38.3], $M$ has an $FP$-source $V$ such that $V \mid iM$, as an $FP$-module. Since $\overline{B}$ and $\overline{B_\gamma}$ are Morita equivalent via multiplication with $i$, the $\overline{B_\gamma}$-module $iM$ is simple. Since $\overline{B_\gamma} \cong FP$, $iM$ is trivial as an $FP$-module, and so is $V$.

(iii) $\Rightarrow$ (i): Suppose that $B$ is nilpotent and that the unique simple $\overline{B}$-module $M$ has a trivial source. Then $M$ has vertex $P$, as above, and a result by Puig [Thévenaz 1995, Theorem 50.6] implies that $\overline{B_\gamma} \cong S \otimes_F FP$ as interior $P$-algebras, where $S$ is an interior $P$-algebra that is simple as an $F$-algebra. (The tensor product of two interior $P$-algebras is again an interior $P$-algebra via the diagonal map.) As above, we write $\overline{B_\gamma} = iFGi$, where $i$ is a primitive idempotent in $(FG)^P$. Since $\overline{B}$ and $\overline{B_\gamma}$ are Morita equivalent via multiplication with $i$, the module $iM$ is the unique simple $\overline{B_\gamma}$-module, up to isomorphism. Thus, $S$ and $\mathrm{End}_F(iM)$ are isomorphic interior $P$-algebras; in particular, $S^P \cong \mathrm{End}_{FP}(iM)$ as $F$-algebras. But $S^P$ is a local ring (since $\overline{B_\gamma}^P$ is), so $iM$ is indecomposable as an $FP$-module. On the other hand, [Thévenaz 1995, Proposition 38.3] implies that $iM$ has a direct summand, as an $FP$-module, which is a source of $M$. Thus $\dim_F iM = 1$. Hence $\dim_F S = 1$, so $S \cong F$ and $\overline{B_\gamma} \cong FP$. In particular, the ring extension $FP \to \overline{B_\gamma}$ has depth one. □

**2.3.** It would be interesting to have a similar description of the depth-two condition for source algebras of blocks. The goal of this subsection is to show that $RP \to B_\gamma$ (and also $FP \to \overline{B_\gamma}$) is a *symmetric Frobenius extension*, so that the left and right depth-two conditions are equivalent [Kadison and Szlachányi 2003, Proposition 6.4].

Recall from [Kadison 1999, Theorem I.1.2] that a ring extension $f\colon \Gamma \to \Delta$ is called a *Frobenius extension* if there exist a $(\Gamma, \Gamma)$-bimodule homomorphism $E\colon \Delta \to \Gamma$ and elements $x_j, y_j \in \Delta$, $j = 1, \ldots, n$, such that

$$\sum_{j=1}^{n} x_j E(y_j a) = a = \sum_{j=1}^{n} E(a x_j) y_j \tag{2.3.a}$$

for all $a \in \Delta$. If in addition

$$E(ca) = E(ac) \tag{2.3.b}$$

holds for all $a \in \Delta$ and $c \in C_\Delta(\Gamma)$, then one calls the extension $f\colon \Gamma \to \Delta$ a *symmetric Frobenius extension*.

If $\Gamma \subseteq \Delta$ is a symmetric Frobenius extension and $e$ is an idempotent in $C_\Delta(\Gamma)$, then $e\Gamma e \subseteq e\Delta e$ is a symmetric Frobenius extension. In fact, if $E\colon \Delta \to \Gamma$ satisfies

(2.3.a) and (2.3.b), then it is easy to verify that $\widetilde{E}\colon e\Delta e \to e\Gamma e$, $a \mapsto eE(a)e$ satisfies

$$\sum_{j=1}^{n} ex_j e\widetilde{E}(ey_j ea) = a = \sum_{j=1}^{n} \widetilde{E}(aex_j e)ey_j e$$

for all $a \in e\Delta e$. Moreover, Equation (2.3.b) implies $\widetilde{E}(ca) = \widetilde{E}(ac)$ for all $a \in e\Delta e$ and $c \in C_{e\Delta e}(e\Gamma e) = eC_\Delta(\Gamma)e$.

If $H$ is a subgroup of $G$, then $kH \subseteq kG$ is a symmetric Frobenius extension. In fact, one can choose for $E\colon kG \to kH$ the canonical projection, and for $x_j$ and $y_j$, coset representatives of $G/H$ and their inverses. Thus, if $e$ is an idempotent in $(kG)^H$, then also $ekHe \to ekGe$ is a symmetric Frobenius extension. This holds even over arbitrary commutative rings $k$.

Now our claim follows by specializing to $H = P$ and $e = 1_{B_\gamma}$ (or $e = 1_{\overline{B_\gamma}}$), and noting that $kP \to ekPe$, $a \mapsto eae = ea = ae$ is an isomorphism of $k$-algebras.

By the preceding discussion, we do not need to distinguish between the left and the right depth-two condition in the following proposition.

**Proposition 2.4.** *Let $B$ be the principal block of $RG$, and let $P_\gamma$ be a maximal local pointed group on $B$ (so that $P$ is a Sylow $p$-subgroup of $G$). Set $E := N_G(P_\gamma)/PC_G(P)$. Let $B_\gamma$ be a source algebra of $B$. Then the following assertions are equivalent*:

(i) *The ring extension $FP \to \overline{B_\gamma}$ defined by the structural map $P \to \overline{B_\gamma}^\times$ has depth two.*

(ii) *$B_\gamma$ is isomorphic to a twisted group algebra $R_\sharp[P \rtimes E]$ of the semidirect product $P \rtimes E$, as an interior $P$-algebra.*

*Proof.* (i) $\Rightarrow$ (ii): Suppose that the ring extension $FP \to \overline{B_\gamma}$ has depth two, and write $A := \overline{B_\gamma} = iFGi$, where $i$ is a primitive idempotent in $(FG)^P$. Then there exists a positive integer $n$ such that

$$\mathrm{Res}^A_{FP}\mathrm{Ind}^A_{FP}\mathrm{Res}^A_{FP}(iM) \mid \mathrm{Res}^A_{FP}(iM)^n$$

for every $B$-module $M$. Taking for $M$ the trivial $FG$-module $F$, we obtain

$$A \otimes_{FP} iF \mid (iF)^n$$

in ${}_{FP}\mathsf{Mod}$. Thus, $P$ acts trivially on $A \otimes_{FP} iF$. On the other hand, $A$ is a direct sum of $(FP, FP)$-bimodules of the form $F[PgP]$, for suitable $g \in G$. It is easy to see that $F[PgP] \otimes_{FP} iF \cong \mathrm{Ind}^P_{P\cap gPg^{-1}}(F)$ in ${}_{FP}\mathsf{Mod}$. And if $P$ acts trivially on $\mathrm{Ind}^P_{P\cap gPg^{-1}}(F)$, then $g \in N_G(P)$. Thus $A$ is in fact a direct sum of $(FP, FP)$-bimodules of the form $F[PgP]$, for suitable $g \in N_G(P)$. Hence [Thévenaz 1995, Theorem 44.3], a result by Puig, implies that $\mathrm{rk}_R B_\gamma = \dim_F \overline{B_\gamma} = |P| \cdot |E|$. Thus [Thévenaz 1995, Theorem 45.11], another result by Puig, implies (ii).

(ii) $\Rightarrow$ (i): Suppose that (ii) holds. Since $R_\sharp[P \rtimes E]$ is a strongly $E$-graded ring with 1-component $R_\sharp P \cong RP$, [Boltje and Külshammer 2010, Proposition 1.5] shows that the ring extension $RP \to R_\sharp[P \rtimes E]$ has depth two. Tensoring with $F$, we obtain (i). □

**Remark 2.5.** The implication (ii) $\Rightarrow$ (i) is valid for arbitrary blocks $B$ of $RG$. Also, if (ii) holds, one can show that every simple $\overline{B}$-module $M$ has trivial source by noting that $P$ acts trivially on $iM$.

## References


[Benson 1998] D. J. Benson, *Representations and cohomology, I: Basic representation theory of finite groups and associative algebras*, 2nd ed., Cambridge Studies in Advanced Mathematics **30**, Cambridge University Press, 1998. MR 99f:20001a Zbl 0908.20001

[Boltje and Glesser 2007] R. Boltje and A. Glesser, "On $p$-monomial modules over local domains", *J. Group Theory* **10**:2 (2007), 173–183. MR 2007m:20013 Zbl 1179.20006

[Boltje and Külshammer 2010] R. Boltje and B. Külshammer, "On the depth 2 condition for group algebra and Hopf algebra extensions", *J. Algebra* **323**:6 (2010), 1783–1796. MR 2011a:16043 Zbl 1200.16035

[Curtis and Reiner 1981] C. W. Curtis and I. Reiner, *Methods of representation theory, I*, Wiley, New York, 1981. MR 82i:20001 Zbl 0469.20001

[Hertweck 2001] M. Hertweck, "A counterexample to the isomorphism problem for integral group rings", *Ann. of Math.* (2) **154**:1 (2001), 115–138. MR 2002e:20010 Zbl 0990.20002

[Hirata 1969] K. Hirata, "Separable extensions and centralizers of rings", *Nagoya Math. J.* **35** (1969), 31–45. MR 39 #5636 Zbl 0179.33503

[Kadison 1999] L. Kadison, *New examples of Frobenius extensions*, University Lecture Series **14**, American Mathematical Society, Providence, RI, 1999. MR 2001j:16024 Zbl 0929.16036

[Kadison 2008] L. Kadison, "Finite depth and Jacobson–Bourbaki correspondence", *J. Pure Appl. Algebra* **212**:7 (2008), 1822–1839. MR 2009f:16070 Zbl 1145.16021

[Kadison and Szlachányi 2003] L. Kadison and K. Szlachányi, "Bialgebroid actions on depth two extensions and duality", *Adv. Math.* **179**:1 (2003), 75–121. MR 2004i:16055 Zbl 1049.16022

[Külshammer 1991] B. Külshammer, *Lectures on block theory*, London Mathematical Society Lecture Note Series **161**, Cambridge University Press, 1991. MR 92h:20020 Zbl 0726.20006

[Nagao and Tsushima 1989] H. Nagao and Y. Tsushima, *Representations of finite groups*, Academic Press, Boston, 1989. MR 90h:20008 Zbl 0673.20002

[Puig 1988] L. Puig, "Nilpotent blocks and their source algebras", *Invent. Math.* **93**:1 (1988), 77–116. MR 89e:20023 Zbl 0646.20010

[Thévenaz 1995] J. Thévenaz, *G-algebras and modular representation theory*, Oxford University Press, New York, 1995. MR 96j:20017 Zbl 0837.20015



Communicated by Dave Benson
Received 2010-01-15 Revised 2010-04-22 Accepted 2010-06-06



boltje@ucsc.edu *Department of Mathematics, University of California, Santa Cruz, CA 95064, United States*

kuelshammer@uni-jena.de *Mathematical Institute, Friedrich Schiller University, 07737 Jena, Germany*

ALGEBRA AND NUMBER THEORY 5:1(2011)

msp

# Set-theoretic defining equations of the variety of principal minors of symmetric matrices

Luke Oeding

The variety of principal minors of $n \times n$ symmetric matrices, denoted $Z_n$, is invariant under the action of a group $G \subset \mathrm{GL}(2^n)$ isomorphic to $\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n$. We describe an irreducible $G$-module of degree-four polynomials constructed from Cayley's $2 \times 2 \times 2$ hyperdeterminant and show that it cuts out $Z_n$ set-theoretically. This solves the set-theoretic version of a conjecture of Holtz and Sturmfels. Standard techniques from representation theory and geometry are explored and developed for the proof of the conjecture and may be of use for studying similar $G$-varieties.

## 1. Introduction

The problem of finding the relations among principal minors of a matrix of indeterminates goes back at least to [Nanson 1897], where relations among the principal minors of an arbitrary $4 \times 4$ matrix are given. An expression for the determinant of a matrix in terms of a subset of its principal minors is given in [Stouffer 1928], but the number of independent principal minors was essentially known to that author even earlier [1924], as pointed out in [Griffin and Tsatsomeros 2006]. (In fact, in his 1928 paper Stouffer says that the result was already known to MacMahon in 1893 and later to Muir.)

Subsequently, interest in the subject seems to have diminished, but much more recently, there has been renewed interest in the relations among principal minors and their application to matrix theory, probability, statistical physics and spectral graph theory.

Based upon work supported by the National Science Foundation under the International Research Fellowship Program (IRFP) Award #0853000 and by the US Department of Education, GAANN Award #P200A060298.

*MSC2000:* primary 14M12; secondary 15A69, 15A29, 15A72, 20G05, 13A50, 14L30.

*Keywords:* principal minors, symmetric matrices, hyperdeterminant, G-variety, G-module, representation theory, hyperdeterminantal module, relations among minors, variety of principal minors, determinant.

In response to questions about principal minors of *symmetric* matrices, Holtz and Sturmfels [2007] introduced the algebraic variety of principal minors of symmetric $n \times n$ matrices (denoted $Z_n$ herein — see Section 3B for the precise definition) and asked for generators of its ideal. In the first nontrivial case, those authors showed that $Z_3$ is an irreducible hypersurface in $\mathbb{P}^7$ cut out by a special degree-four polynomial, namely Cayley's hyperdeterminant of format $2 \times 2 \times 2$. In the next case they showed (with the aid of a computer calculation) that the ideal of $Z_4$ is minimally generated by 20 degree-four polynomials, but only 8 of these polynomials are copies of the hyperdeterminant constructed by natural substitutions. The other 12 polynomials were separated into classes based on their multidegrees. This was done in a first draft of [Holtz and Sturmfels 2007], and at that point, the geometric meaning of the remaining polynomials and their connection to the hyperdeterminant was still somewhat mysterious. Because of the symmetry of the hyperdeterminant, Landsberg suggested to Holtz and Sturmfels the following:

**Theorem 1.1** [Holtz and Sturmfels 2007, Theorem 12]. *The variety $Z_n$ is invariant under the action of*

$$\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n.$$

It should be noted that Borodin and Rains [2005] had found a similar result for two other cases: when the matrix is not necessarily symmetric and for a Pfaffian analog. In [Oeding 2009], we showed that $Z_n$ is a linear projection of the well-known Lagrangian Grassmannian, which can also be interpreted as the variety of all minors of a symmetric matrix. We used this projection to prove Theorem 1.1 geometrically.

In [Holtz and Sturmfels 2007], the span of the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-orbit of the $2 \times 2 \times 2$ hyperdeterminant is named *the hyperdeterminantal module* (denoted *HD* herein — see Section 3). It was then understood — and included in the final version of the paper — that the 20 degree-four polynomials are a basis of the hyperdeterminantal module when $n = 4$. This interpretation led to the following:

**Conjecture 1.2** [Holtz and Sturmfels 2007, Conjecture 14]. The prime ideal of the variety of principal minors of symmetric matrices is generated in degree 4 by the hyperdeterminantal module for all $n \geq 3$.

While the first two cases of the conjecture ($n = 3, 4$) were proved using a computer, the dimension of the hyperdeterminantal module and the number of variables both grow exponentially with $n$, making computational methods ineffective already in the case $n = 5$, for which the hyperdeterminantal module has a basis of 250 degree-four polynomials in 32 variables. Our point of departure is the use of the symmetry of $Z_n$ via tools from representation theory and the geometry of $G$-varieties.

Our main purpose is to solve the set-theoretic version of the Holtz–Sturmfels conjecture (see Example 3.3 for the representation-theoretic description of the hyperdeterminantal module in terms of Schur modules used in this statement):

**Theorem 1.3** (Main Theorem). *The variety of principal minors of symmetric $n \times n$ matrices, $Z_n$, is cut out set-theoretically by the hyperdeterminantal module, which is the irreducible* $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$*-module of degree-four polynomials*

$$HD = S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\dots S_{(4)}.$$

The set-theoretic result is sufficient for many applications related to principal minors of symmetric matrices. In particular, set-theoretic defining equations of $Z_n$ are necessary and sufficient conditions for a given vector of length $2^n$ to be expressed as the principal minors of a symmetric matrix. We state this practical membership test as follows:

**Corollary 1.4.** *Suppose $w = [w_{[i_1,\dots,i_n]}] \in \mathbb{C}^{2^n}$ with $i_j \in \{0, 1\}$. Then $w$ represents the principal minors of a symmetric $n \times n$ matrix if and only if $w$ and all images under changes of coordinates by* $\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n$ *are zeros of Cayley's $2 \times 2 \times 2$ hyperdeterminant*

$$\begin{aligned}
&(w_{I_{[0,0,0]}})^2(w_{I_{[1,1,1]}})^2 + (w_{I_{[1,0,0]}})^2(w_{I_{[0,1,1]}})^2 + (w_{I_{[0,1,0]}})^2(w_{I_{[1,0,1]}})^2\\
&+(w_{I_{[0,0,1]}})^2(w_{I_{[1,1,0]}})^2 - 2w_{I_{[0,0,0]}}w_{I_{[1,0,0]}}w_{I_{[0,1,1]}}w_{I_{[1,1,1]}} - 2w_{I_{[0,0,0]}}w_{I_{[0,1,0]}}w_{I_{[1,0,1]}}w_{I_{[1,1,1]}}\\
&\qquad - 2w_{I_{[0,0,0]}}w_{I_{[0,0,1]}}w_{I_{[1,1,0]}}w_{I_{[1,1,1]}} - 2w_{I_{[1,0,0]}}w_{I_{[0,1,0]}}w_{I_{[0,1,1]}}w_{I_{[1,0,1]}}\\
&\qquad - 2w_{I_{[1,0,0]}}w_{I_{[0,0,1]}}w_{I_{[0,1,1]}}w_{I_{[1,1,0]}} - 2w_{I_{[0,1,0]}}w_{I_{[0,0,1]}}w_{I_{[1,0,1]}}w_{I_{[1,1,0]}}\\
&\qquad + 4w_{I_{[0,0,0]}}w_{I_{[0,1,1]}}w_{I_{[1,0,1]}}w_{I_{[1,1,0]}} + 4w_{I_{[0,0,1]}}w_{I_{[0,1,0]}}w_{I_{[1,0,0]}}w_{I_{[1,1,1]}},
\end{aligned}$$

*where $I_{[i_1,i_2,i_3]} = [i_1, i_2, i_3, 0, \dots, 0]$ for $i_j \in \{0, 1\}$.*

A second, unifying purpose of this work is to study $Z_n$ as a prototypical (non-homogeneous) $G$-variety. We aim to show the usefulness of standard constructions in representation theory and geometry, and to further develop general tools for studying geometric and algebraic properties of such varieties. We anticipate these techniques will be applicable to other $G$-varieties in spaces of tensors, such as those that arise naturally in computational complexity [Landsberg 2008; Bürgisser et al. 1997], signal processing [Comon et al. 2008; Comon and Rajih 2006; de Lathauwer and de Baynast 2008], and algebraic statistics [Pachter and Sturmfels 2005; Allman and Rhodes 2008] (see also [Landsberg $\geq$ 2011] for a unified presentation of the use of geometry and representation theory in these areas), and especially to the case of principal minors of arbitrary matrices studied by Lin and Sturmfels [2009] and Borodin and Rains [2005]. In fact, we use techniques similar to those found here as well as Theorem 1.3 in the sequel [Oeding 2011], which investigates a connection between principal minors of symmetric matrices and the tangential variety

to the Segre product of projective spaces, and solves the set-theoretic version of a conjecture of Landsberg and Weyman [2007].

**1A.** ***Outline.*** This paper is organized as follows. Section 2 discusses applications of Theorem 1.3 to statistics, physics and graph theory. In Section 3 we recall basic notions concerning tensors, representations and $G$-varieties. We point out many standard facts from representation theory that we will use to study the ideal of $Z_n$ and the hyperdeterminantal module. In particular, we recall a method used by Landsberg and Manivel to study $G$-modules of polynomials via Schur modules. We also show how to use weights and lowering operators to describe and identify Schur modules. We use these concepts to prove Lemma 7.2, a stepping-stone to Proposition 7.1; the latter, in turn, is crucial to our proof of Theorem 1.3.

Sections 3B, 4, 5, and 6 deal with geometric aspects of the variety of principal minors of symmetric matrices and the zero set of the hyperdeterminantal module. In Section 3B we set up notation and give a precise definition of the variety. We also recall two useful facts; a symmetric matrix is determined up to the signs of its off-diagonal terms by its $1 \times 1$ and $2 \times 2$ principal minors, and the dimension of $Z_n$ is $\binom{n+1}{2}$. In Section 4 we describe the nested structure of $Z_n$. In particular, in Proposition 4.2 we show that $Z_n$ contains all possible Segre products of $Z_p$ and $Z_q$ where $p+q=n$. We use this interpretation in Proposition 7.1.

In Section 5 we study properties of the hyperdeterminantal module. In particular, we point out that it has dimension $\binom{n}{3}5^{n-3}$. In Proposition 5.2 we show that it actually is an irreducible $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module of polynomials that occurs with multiplicity 1 in the space of degree-four homogeneous polynomials. This is a consequence of a more general fact about modules with structure similar to that of the hyperdeterminantal module, which we record in Lemma 5.4. In Proposition 5.5 we record the fact (originally proved in [Holtz and Sturmfels 2007]) that the hyperdeterminantal module is in the ideal of $Z_n$. Then in Proposition 5.6 we generalize the idea to other varieties that have similar structure.

In Section 6 we extract a general property of the hyperdeterminantal module that we call augmentation. We explore properties of augmented modules by means of polarization of tensors, a technique from classical invariant theory used, for example, in the study of secant varieties. Of particular interest is the augmentation lemma (Lemma 6.4), in which we give a geometric description of the zero set of a general augmented module. We apply that lemma to obtain a geometric characterization of the zero set of the hyperdeterminantal module in Lemma 6.6. We use Lemma 6.6 in the proof of Theorem 1.3. Proposition 6.8 is another application of the Augmentation Lemma to polynomials that define Segre products of projective spaces. We use a slightly more complicated version of Proposition 6.8 in the proof of Lemma 7.2.

In Sections 7 and 8 we pull together all of the ideas from the previous parts to prove Theorem 1.3. In particular, we show that any point in the zero set of the hyperdeterminantal module has a symmetric matrix that maps to it under the principal minor map.

In Section 7 we work to understand the case when all principal minors of a symmetric matrix agree with a given vector except possibly the determinant. Of particular importance is Proposition 7.1, which essentially says that for $n \geq 4$, if $z$ is a vector in the zero set of the hyperdeterminantal module, then a specific subset of the coordinates of $z$ determine the rest of its coordinates.

In order to prove Proposition 7.1, we use practically all of the tools from representation theory that we have introduced and developed earlier in the paper. With the aid of Proposition 7.1, we prove Theorem 1.3 in Section 8.

## 2. Applications of Theorem 1.3

We conclude this introduction by describing how Theorem 1.3 answers questions in other areas via three examples: in statistics and the study of negatively correlated random variables, in physics and the study of determinantal point processes, and in spectral graph theory and the study of graph invariants [Holtz and Schneider 2002; Holtz and Sturmfels 2007; Griffin and Tsatsomeros 2006; Borcea et al. 2009; Holtz 1999; Wagner 2008; Mikkonen 2007].

**2A.** ***Application to covariance of random variables.*** Consider a nonsingular real symmetric $n \times n$ matrix $A$. The principal minors of $A$ can be interpreted as values of a function $\omega : \mathcal{P}(\{1, \ldots, n\}) \to [0, \infty)$, where $\mathcal{P}$ is the power set. This function $\omega$, under various restrictions, is of interest to statisticians. In this setting, the off-diagonal entries of the matrix $A^{-1}$ are associated to covariances of random variables. Wagner [2008] asked:

**Question 2.1.** *When is it possible to prescribe the principal minors of the matrix $A$ as well as the off-diagonal entries of $A^{-1}$?*

In [Holtz and Sturmfels 2007, Theorem 6] this question is answered using the hyperdeterminantal equations in degree 4, another set of degree-10 equations, and the strict Hadamard–Fischer inequalities.

Our main result provides an answer to the first part of the question: it is possible to prescribe the principal minors of a symmetric matrix if and only if the candidate principal minors satisfy all the relations given by the hyperdeterminantal module.

For the second part of the question we can give a partial answer. It is not hard to see that the off-diagonal entries of $A^{-1}$ are determined up to sign by the $0 \times 0$, $1 \times 1$ and $2 \times 2$ principal minors, and the rest of the principal minors further restrict the freedom in the choices of signs.

Another useful fact is that if $A$ is invertible, then

$$A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)},$$

where $\operatorname{adj}(A)_{i,j} = ((-1)^{i+j} \det(A_i^j))$ is the adjugate matrix. So up to scale, the vector of principal minors of $A^{-1}$ is the vector of principal minors of $A$ in reverse order. Therefore the determinant and the $(n-1)\times(n-1)$ and $(n-2)\times(n-2)$ principal minors of $A$ determine the off-diagonal entries of $A^{-1}$ up to $\binom{n}{2}$ choices in combinations of signs, and the rest of the principal minors further restrict the choices of combinations of signs.

**2B.** ***Application to determinantal point processes.*** Determinantal point processes were introduced by Macchi in 1975, and subsequently have received significant attention in many areas. A nonzero point $p_S \in \mathbb{C}^{2^n}$ is called determinantal if there is an integer $m$ and an $(n+m)\times(n+m)$ matrix $K$ such that for $S \subset \{1, 2, \ldots, n\}$,

$$p_S = \det_{S\cup\{n+1,\ldots,n+m\}}(K).$$

Borodin and Rains [2005, Theorem 4.6] were able to completely classify all such points for the case $n = 4$ by giving a nice geometric characterization. Lin and Sturmfels [2009] studied the geometric and algebraic properties of the algebraic variety of determinantal points and independently arrived at the same result as Borodin and Rains. Moreover, Lin and Sturmfels gave a complete proof of the claim of [Borodin and Rains 2005] that the ideal of the variety is generated in degree 12 by 718 polynomials.

Consider the case where we impose the restrictions that the matrix $K$ to be symmetric and the integer $m$ equals 0; we call these restricted determinantal points *symmetric determinantal points*.

**Restatement.** *The variety of all symmetric determinantal points is cut out set-theoretically by the hyperdeterminantal module.*

This restatement is useful because it provides a complete list of necessary and sufficient conditions for determining which symmetric determinantal points can possibly exist.

**2C.** ***Application to spectral graph theory.*** A standard construction in graph theory is the following. To a weighted directed graph $\Gamma$ one can assign an adjacency matrix $\Delta(\Gamma)$.

The eigenvalues of $\Delta(\Gamma)$ are invariants of the graph. The first example involves the standard graph Laplacian. Kirchoff's well-known matrix-tree theorem states that any $(n-1)\times(n-1)$ principal minor of $\Delta(\Gamma)$ counts the number of spanning trees of $\Gamma$.

There are many generalizations of the Matrix-Tree Theorem, such as the Matrix-Forest Theorem, which states that $\Delta(\Gamma)^S_S$, the principal minor of the graph Laplacian formed by omitting rows and columns indexed by the set $S \subset \{1, \ldots, n\}$, computes the number of spanning forests of $\Gamma$ rooted at vertices indexed by $S$.

The principal minors of the graph Laplacian are graph invariants. The relations among principal minors are then also relations among graph invariants. Relations among graph invariants are central in the study of the theory of unlabeled graphs. In fact, Mikkonen [2007, p. 1] holds that "the most important problem in graph theory of unlabeled graphs is the problem of determining graphic values of arbitrary sets of graph invariants."

Theorem 1.3 gives relations among the graph invariants that come from principal minors, and in particular, since a graph can be reconstructed from a symmetric matrix, Theorem 1.3 implies the following:

**Restatement.** *There exists an undirected weighted graph $\Gamma$ with invariants $[v] \in \mathbb{P}^{2^n-1}$ specified by the principal minors of a symmetric matrix $\Delta(\Gamma)$ if and only if $[v]$ is a zero of all the polynomials in the hyperdeterminantal module.*

## 3. Background on $G$-varieties in spaces of tensors and their ideals as $G$-modules

An $n \times n$ matrix has $2^n$ principal minors (determinants of submatrices centered on the main diagonal, including the trivial $0 \times 0$ minor), so vectors of principal minors may be considered in the space $\mathbb{C}^{2^n}$. However, the natural ambient space for vectors of principal minors from the point of view of symmetry (Theorem 1.1) is the $n$-fold tensor product $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. With this in mind, in this section we study tensor products of several vector spaces, natural group actions on tensors, representation theory for tensor products, and classical subvarieties in spaces of tensors.

For the sake of the reader not familiar with representation theory, we have chosen to include many definitions and basic concepts that we might have skipped otherwise. For more background, one may consult [Fulton and Harris 1991; Landsberg $\geq$ 2011; Goodman and Wallach 1998; Weyl 1997; Harris 1992; Cox et al. 2007].

If $V$ is a vector space and $G \subset \mathrm{GL}(V)$, a variety $X \subset \mathbb{P}V$ is said to be a *$G$-variety* or *$G$-invariant* if it is preserved by the action of $G$, specifically $g.x \in X$ for every $x \in X$ and $g \in G$. In this article our vector spaces are always assumed to be finite-dimensional. Our study fits into the more general context of arbitrary $G$-varieties for a linearly reductive group $G$, and we sometimes allude to this setting, but for the sake of efficiency and clarity we often present the necessary representation-theoretic concepts only in the case of tensors. The expert reader might try to envision the basic techniques we use in their more general context.

**3A.** ***Examples of classical G-varieties in spaces of tensors.*** Let $V_1, \dots, V_n$ be complex vector spaces and let $V_1 \otimes \cdots \otimes V_n$ denote their tensor product. We give two classic examples of $G$-varieties in the space of tensors $\mathbb{P}(V_1 \otimes \cdots \otimes V_n)$ that happen to show up in our study of the variety of principal minors of symmetric matrices. These definitions can be found in many texts on algebraic geometry, such as [Harris 1992].

The space of all rank-1 tensors (also called decomposable tensors) is the *Segre variety*, defined by the embedding

$$\begin{aligned} \mathrm{Seg} : \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n &\longrightarrow \mathbb{P}(V_1 \otimes \cdots \otimes V_n), \\ ([v_1], \dots, [v_n]) &\longmapsto [v_1 \otimes \cdots \otimes v_n]. \end{aligned}$$

$\mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n)$ is a $G$-variety for $G = \mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_n)$, and it is *homogeneous* (the $G$-orbit of a single point) since $\mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n) = G.[v_1 \otimes \cdots \otimes v_n]$. If $X_1 \subset \mathbb{P}V_1, \dots, X_n \subset \mathbb{P}V_n$ are varieties, let $\mathrm{Seg}(X_1 \times \cdots \times X_n)$ denote their Segre product.

The $r$th *secant variety* to a variety $X \subset \mathbb{P}V$, denoted $\sigma_r(X)$, is the Zariski closure of all embedded secant $\mathbb{P}^{r-1}$'s to X, that is,

$$\sigma_r(X) = \overline{\bigcup_{x_1,\dots,x_r \in X} \mathbb{P}(\mathrm{span}\{x_1, \dots, x_r\})} \subset \mathbb{P}V.$$

Secant varieties inherit the symmetry of the underlying variety. In particular,

$$\sigma_r(\mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n))$$

is a $G$-variety for $G = \mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_n)$. However, homogeneity is not preserved in general.

**3B.** ***The variety of principal minors of symmetric matrices.*** Let $I = [i_1, \dots i_n]$ be a binary multi-index, with $i_k \in \{0, 1\}$ for $k = 1, \dots, n$, and let $|I| = \sum_{k=1}^{n} i_k$. A natural basis of $(\mathbb{C}^2)^{\otimes n}$ is the set of tensors $X^I := x_1^{i_1} \otimes x_2^{i_2} \otimes \cdots \otimes x_n^{i_n}$ for all length-$n$ binary indices $I$. We use this basis to introduce coordinates; if $P = [C_I X^I] \in \mathbb{P}(\mathbb{C}^2)^{\otimes n}$, the coefficients $C_I$ are the homogeneous coordinates of the point $P$. (We use the summation convention that the implied summation is over the index $I$, which appears as a superscript and a subscript.)

Let $S^2\mathbb{C}^n$ denote the space of symmetric $n \times n$ matrices. If $A \in S^2\mathbb{C}^n$, then let $\Delta_I(A)$ denote the principal minor of $A$ formed by taking the determinant of the principal submatrix of $A$ indexed by $I$ in the sense that the submatrix of $A$ is formed by including the $k$th row and column of $A$ whenever $i_k = 1$ and striking the $k$th row and column whenever $i_k = 0$.

The projective variety of principal minors of $n \times n$ symmetric matrices, $Z_n$, is defined by

$$\varphi : \mathbb{P}(S^2\mathbb{C}^n \oplus \mathbb{C}) \dashrightarrow \mathbb{P}(\mathbb{C}^2)^{\otimes n}, \quad [A, t] \mapsto [t^{n-|I|}\Delta_I(A)\, X^I].$$

The map $\varphi$ is defined on the open set where $t \neq 0$. Moreover, $\varphi$ is homogeneous of degree $n$, so it is a well-defined rational map on projective space. The $1 \times 1$ principal minors of a matrix $A$ are the diagonal entries of $A = (a_{i,j})$, and if $A$ is a symmetric matrix, the $1 \times 1$ and $2 \times 2$ principal minors determine the off-diagonal entries of $A$ up to sign in light of the equation

$$a_{i,i}a_{j,j} - a_{i,j}^2 = \Delta_{[0,\ldots,0,1,0,\ldots,0,1,\ldots,0]}(A),$$

where the 1's in $[0, \ldots, 0, 1, 0, \ldots, 0, 1, \ldots, 0]$ occur in positions $i$ and $j$. So $\varphi$ is generically finite-to-one and $Z_n$ is a $\binom{n+1}{2}$-dimensional variety. The affine map (on the set $\{t = 1\}$) defines a closed subset of $\mathbb{C}^{2^n}$ [Holtz and Sturmfels 2007].

**3C.** ***Ideals of G-varieties in spaces of tensors.*** Let $V$ be a finite-dimensional vector space over $\mathbb{C}$. Let $V^*$ denote the dual vector space of linear maps $V \to \mathbb{C}$. Let $S^dV^*$ denote the space of homogeneous degree-$d$ polynomials on $V$, and let $\mathrm{Sym}(V^*) = \bigoplus_d S^dV^*$ denote the polynomial ring.

If $X \subset \mathbb{P}V$ is a projective algebraic variety, let $\mathcal{I}(X) \subset \mathrm{Sym}(V^*)$ denote the ideal of polynomials vanishing on $X$, and let $\widehat{X} \subset V$ denote the cone over $X$. If $M$ is a set of polynomials, let $\mathcal{V}(M)$ denote its zero set. Often algebraic varieties are given via an explicit parameterization by a rational map, but the vanishing ideal may be unknown. A basic question in algebraic geometry is to find generators for the ideal of a given variety. Though there are many known theoretical techniques, this remains a difficult practical problem.

**Fact.** *$X$ is a $G$-variety if and only if $\mathcal{I}(X)$ is a $G$-module.*

This observation, which comes directly from the definitions, is key because it allows us to use the representation theory of $G$-modules to study $\mathcal{I}(X)$.

By definition, all projective varieties are preserved by the action of $\mathbb{C} \setminus \{0\}$ by rescaling. It is well-known that this action induces a grading by degree on the ideal $\mathcal{I}(X) = \bigoplus_d \mathcal{I}_d(X)$, where $\mathcal{I}_d X := S^d(V^*) \cap \mathcal{I}(X)$. In parallel, when a larger, linearly reductive group $G$ acts on $X$, we get a finer decomposition of each module $\mathcal{I}_d(X)$ into a direct sum of irreducible $G$-modules. The irreducible modules in $\mathcal{I}_d(X)$ are a subset of those in $S^dV^*$. This simple observation leads to a useful ideal membership test, which is developed and discussed in [Landsberg and Manivel 2004; Landsberg $\geq$ 2011].

The group $\mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_n)$ acts on $V_1 \otimes \cdots \otimes V_n$ by change of coordinates in each factor. When $V_i$ are all isomorphic, there is also a natural action of the symmetric group $\mathfrak{S}_n$ on $V_1 \otimes \cdots \otimes V_n$ by permuting the factors. With this convention,

one can define a left action of the semidirect product $\mathrm{GL}(V) \ltimes \mathfrak{S}_n$ on $V^{\otimes n}$.

If $V$ is a vector space and $G \subset \mathrm{GL}(V)$, we say that $V$ is a *$G$-module* or *a representation of $G$* if it is preserved by the action of $G \subset \mathrm{GL}(V)$. A $G$-module is said to be *irreducible* if it has no nontrivial $G$-invariant subspaces.

The general linear group $\mathrm{GL}(V)$ has well-understood representation theory. In particular, [Fulton and Harris 1991, Proposition 15.47] says that every $\mathrm{GL}(V)$-module is isomorphic to a Schur module of the form $S_\pi V$, where $\pi$ is a partition of an integer $d$. We refer the reader to the book just cited or to [Landsberg $\geq$ 2011] for general background on Schur modules.

Two common representations (in this language) are the space of symmetric tensors $S^d V = S_{(d)} V$ and the space of skew-symmetric tensors $\bigwedge^d V = S_{(1^d)} V$, where $1^d$ denotes the partition $(1, \ldots, 1)$ with 1 repeated $d$ times.

We will be interested in representations of $\mathrm{SL}(V)$. In light of the isomorphism $\mathrm{SL}(V) \cong \mathrm{GL}(V)/Z(\mathrm{GL}(V))$, where the center $Z(\mathrm{GL}(V)) = \mathbb{C} \setminus \{0\}$ is isomorphic to scalar multiples of the identity, the representation theory of $\mathrm{GL}(V)$ is essentially the same at that of $\mathrm{SL}(V)$. Specifically, if $V$ is $m$-dimensional, two representations $S_\pi V$ and $S_\lambda V$ of $\mathrm{GL}(V)$ are isomorphic as $\mathrm{SL}(V)$ modules if $\pi = \lambda + k^m$, for some $k \in \mathbb{Z}$, where $k^m$ is the partition $(k, \ldots, k)$ with $k$ repeated $m$ times. However, since we care about how the modules we are studying are embedded in the space of polynomials, we will not reduce partitions via this equivalence.

We are interested in the case when $X \subset \mathbb{P}(V_1 \otimes \cdots \otimes V_n)$ is a variety in a space of tensors, and $X$ is invariant under the action of $G = \mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_n)$. To study $\mathcal{I}_d(X)$ as a $G$-module, we need to understand how to decompose the space of homogeneous degree-$d$ polynomials $S^d(V_1^* \otimes \cdots \otimes V_n^*)$ into a direct sum of irreducible $G$-modules. This is a standard computation in representation theory, made explicit, for example, in this result:

**Proposition 3.1** [Landsberg and Manivel 2004, Proposition 4.1]. *Let $V_1, \ldots, V_n$ be vector spaces and let $G = \mathrm{GL}(V_1) \times \cdots \times \mathrm{GL}(V_n)$. Then $S^d(V_1^* \otimes \cdots \otimes V_n^*)$ can be decomposed into a direct sum of irreducible $G$-modules as*

$$S^d(V_1^* \otimes \cdots \otimes V_n^*) = \bigoplus_{|\pi_1|=\cdots=|\pi_n|=d} ([\pi_1] \otimes \cdots \otimes [\pi_n])^{\mathfrak{S}_d} \otimes S_{\pi_1} V_1^* \otimes \cdots \otimes S_{\pi_n} V_n^*,$$

*where the $[\pi_i]$ are representations of the symmetric group $\mathfrak{S}_d$ indexed by partitions $\pi_i$ of $d$, and $([\pi_1] \otimes \cdots \otimes [\pi_n])^{\mathfrak{S}_d}$ denotes the space of $\mathfrak{S}_d$-invariants (that is, instances of the trivial representation) in the tensor product.*

When the vector spaces $V_i^*$ are all isomorphic to the same vector space $V^*$, this decomposition specializes to give the equality of $\mathrm{GL}(V) \times \cdots \times \mathrm{GL}(V)$-modules

$$S^d(V^* \otimes \cdots \otimes V^*) = \bigoplus_{|\pi_1|=\cdots=|\pi_n|=d} (S_{\pi_1} V^* \otimes \cdots \otimes S_{\pi_n} V^*)^{\oplus N_{\pi_1,\ldots,\pi_k}}, \qquad (1)$$

(found in [Landsberg and Manivel 2004]), where the multiplicity $N_{\pi_1,\dots,\pi_k}$ can be computed via characters. The modules $(S_{\pi_1}V^*\otimes\cdots\otimes S_{\pi_n}V^*)^{\oplus N_{\pi_1,\dots,\pi_k}}$ are called *isotypic components*.

The irreducible $\mathrm{SL}(V)^{\times n}\ltimes\mathfrak{S}_n$-modules are constructed by taking an irreducible $\mathrm{SL}(2)^{\times n}$ module $S_{\pi_1}V\otimes\cdots\otimes S_{\pi_n}V$ and summing over all permutations in $\mathfrak{S}_n$ that yield nonredundant modules. When the vector space is understood, we denote this compactly as

$$S_{\pi_1}S_{\pi_2}\dots S_{\pi_n} := \sum_{\sigma\in\mathfrak{S}_n} S_{\pi_{\sigma(1)}}V^*\otimes\cdots\otimes S_{\pi_{\sigma(n)}}V^*.$$

The decomposition formula (1) is essential for understanding the structure of the ideals of $G$-varieties. There is an implementation of (1) in the computer program LiE, and we wrote an implementation in Maple.

The combinatorial description of Schur modules in terms of collections of partitions can be used to construct polynomials in spaces of tensors. We refer the reader to [Landsberg and Manivel 2004; Landsberg $\geq$ 2011] for a complete explanation. A copy of our implementation of these algorithms may be obtained by contacting the author.

**3D.** ***Weights, raising operators, and highest weight vectors.*** The ideas of *weights*, *weight vectors*, *highest weight vectors*, and *raising/lowering operators* are well-known practical tools for studying representations and polynomials in spaces of tensors. Here we recall definitions and concepts that can be found in standard textbooks on representation theory in order to define the terms we use in this paper and to explain our use of these representation-theoretic tools. Many of the concepts in this section are practical reinterpretations of concepts in the previous section.

Choose a basis $\{x_i^0, x_i^1\}$ for each $V_i$ and assign the integer *weight* $-1$ to $x_i^0$ and the weight $+1$ to $x_i^1$. Weights of tensors in the algebra $(V_1\otimes\cdots\otimes V_n)^{\otimes}$ are length-$n$ integer vectors defined first on monomials and extended by linearity. Specifically,

$$(x_1^0)^{\otimes p_1}\otimes(x_1^1)^{\otimes q_1}\otimes(x_2^0)^{\otimes p_2}\otimes(x_2^1)^{\otimes q_2}\otimes\cdots\otimes(x_n^0)^{\otimes p_n}\otimes(x_n^1)^{\otimes q_n}$$

has weight

$$(q_1-p_1, q_2-p_2, \dots, q_n-p_n).$$

A tensor is called a *weight vector* if all of its monomials have the same weight, and this is the only time it makes sense to assign a weight to a tensor. This is the standard assignment of weights for the connected component containing the identity in $\mathrm{SL}(2)^{\times n}\ltimes\mathfrak{S}_n$, and is also known as grading by multidegree.

The Lie algebra of $\mathrm{SL}(2)$ is $\mathfrak{sl}(2)$, the algebra of traceless $2\times 2$ matrices acting as derivations. The raising (respectively lowering) operators can be thought of as upper (respectively lower) triangular matrices when, for example, the lowering

operator in $\mathfrak{sl}_2$ acts on $V = \{x_0, x_1\}$ by sending $x_0$ to a scalar multiple of $x_1$ and sending $x_1$ to 0.

The Lie algebra of $\mathrm{SL}(2)^{\times n}$ is $\mathfrak{sl}_2^{\oplus n}$, where each $\mathfrak{sl}_2$ acts on a single factor of the tensor product $V_1 \otimes \cdots \otimes V_n$. This action is extended to $S^d(V_1 \otimes \cdots \otimes V_n)$ by noting that the differential operators obey the Leibnitz rule. The raising (lowering) operators fix the degree of a polynomial.

A weight vector in a $G$-module is called a *highest weight vector* (respectively *lowest weight vector*) if it is in the kernel of all of the raising (respectively lowering) operators. Consider the irreducible module $S_{\pi_1}V_1 \otimes \cdots \otimes S_{\pi_n}V_n$, with each $\pi_i$ a partition of $d$. Since $V_i \simeq \mathbb{C}^2$ for every $1 \leq i \leq n$, each $\pi_i$ is of the form $(\pi_i^1, \pi_i^2)$ with $\pi_i^1 + \pi_i^2 = d$. A highest weight vector in $S_{\pi_1}V_1 \otimes \cdots \otimes S_{\pi_n}V_n$ has weight $(\pi_1^2 - \pi_1^1, \pi_2^2 - \pi_2^1, \ldots, \pi_n^2 - \pi_n^1)$. If $w$ is the weight of a nonzero vector in $S_{\pi_1}V_1 \otimes \cdots \otimes S_{\pi_n}V_n$, then $-w$ is also the weight of a nonzero vector, and if $w$ is the weight of a highest weight vector in a module then $-w$ is the weight of a lowest weight vector.

**Fact.** *Assume $G$ is a linearly reductive connected algebraic group. Each finite-dimensional irreducible $G$-module is the span of the $G$-orbit of a highest* (*or lowest*) *weight vector.*

**Remark 3.2.** If $T$ is a nonzero homogeneous polynomial on $V_1 \otimes \cdots \otimes V_n$, and $T$ is a highest (or lowest) weight vector, then the degree $d$ and weight $(w_1, w_2, \ldots, w_n)$ of $T$ are sufficient information to determine (up to isomorphism) a module of the form $S_{\pi_1}V_1 \otimes \cdots \otimes S_{\pi_n}V_n$ in which it occurs. We say that we know in which isotypic component the module lives. Specifically, we have $d = \pi_i^1 + \pi_i^2$ and $w_i = \pi_i^2 - \pi_i^1$, so $\pi_i = \frac{1}{2}(d - w_i, d + w_i)$.

In general, the degree and weight of a highest weight polynomial are not enough to find how the module $S_{\pi_1}V_1^* \otimes \cdots \otimes S_{\pi_n}V_n^*$ is embedded in $S^d(V_1^* \otimes \cdots \otimes V_n^*)$ (that is, how it is embedded in the isotypic component). On the other hand, if the found module occurs with multiplicity one in $S^d(V_1^* \otimes \cdots \otimes V_n^*)$, then the degree and weight of a highest weight vector are sufficient information to identify the module.

**Example 3.3.** The hyperdeterminant of format $2 \times 2 \times 2$ is invariant under the action of $\mathrm{SL}(2) \times \mathrm{SL}(2) \times \mathrm{SL}(2)$, and therefore it must have weight $(0, 0, 0)$. This, together with the knowledge that it is a degree-4 polynomial annihilated by each raising operator, immediately tells us that it must be in the module $S_{(2,2)}\mathbb{C}^2 \otimes S_{(2,2)}\mathbb{C}^2 \otimes S_{(2,2)}\mathbb{C}^2$, which occurs with multiplicity one in $S^4(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2)$. Moreover, one can write the $2 \times 2 \times 2$ hyperdeterminant on the variables $X^{[i_1,i_2,i_3,0,\ldots,0]}$. The weight of this polynomial is $(0, 0, 0, -4, \ldots, -4)$ and it is a highest weight vector, and therefore the span of its $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-orbit is the hyperdeterminantal module

$$HD := S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)} \ldots S_{(4)}.$$

**3E.** ***An algorithm to produce a G-module from a polynomial.*** Suppose we can write down a polynomial $h$ in some (unknown) $G$-module $M$. (Again we are assuming that our modules are finite-dimensional and the group $G$ is a linearly reductive connected algebraic group, and we are specifically thinking of the example $G = \mathrm{SL}(V)^{\times n}$.) Since $M$ is a $G$-module, it is also a $\mathfrak{g}$-module, where $\mathfrak{g}$ is the Lie algebra associated to the Lie group $G$. The following algorithm is a standard idea in representation theory and can be used to find more polynomials in $M$, and in fact we will find submodules of $M$. In particular, this procedure is essential in the proof of Lemma 7.2 below.

By successively applying lowering operators, we determine the lowest weight space in which a summand of $h$ can live. The lowest weight vector that we construct will generate a submodule of $M$.

Input: $h \in M$.

Step 0. Choose an ordered basis of lowering operators $\mathfrak{g}_- = \{\alpha_1, \ldots, \alpha_n\}$.

Step 1. Find the largest integer $k_1 \geq 0$ such that $\alpha_1^{k_1}.h \neq 0$, and let $h^{(1)} = \alpha_1^{k_1}.h$.

Step 2. Find the largest integer $k_2 \geq 0$ such that $\alpha_2^{k_2}.h^{(1)} \neq 0$, and let $h^{(2)} = \alpha_2^{k_2}.h^{(1)}$.

Step $n$. Find the largest integer $k_n \geq 0$ such that $\alpha_n^{k_n}.h^{(n-1)} \neq 0$, and let $h^{(n)} = \alpha_n^{k_n}.h^{(n-1)}$.

Output: The vector $h^{(n)}$ is a lowest weight vector in $M$ and $\mathrm{span}\{G.h^{(n)}\}$ is a submodule of $M$ containing $h^{(n)}$.

In the case $\mathfrak{g} = \mathfrak{sl}(2)^{\oplus n}$, the natural ordered basis of $(\mathfrak{sl}(2)^{\oplus n})_-$ is $\{\alpha_1, \ldots, \alpha_n\}$, where $\alpha_i$ is the lowering operator acting on the $V_i^*$ factor in $S^d(V_1^* \otimes \cdots \otimes V_n^*)$.

**Remark 3.4.** In the case that $M$ is irreducible, by the same procedure of applying lowering operators to (this time) a highest weight vector $h$, we can construct a *weight basis* $M$, namely a basis of $M$ consisting of weight vectors in $M$ of every possible weight.

## 4. The nested structure of $Z_n$ via Segre products

**Proposition 4.1.** *The variety* $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}V_n)$ *is a subvariety of* $Z_n$. *In particular, any point of* $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}V_n)$ *is, after a possible change of coordinates, the principal minors of an* $(n-1) \times (n-1)$ *block of an* $n \times n$ *matrix.*

*Proof.* We prove the second claim first. Let $[\eta \otimes v]$ be a point in $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}V_n)$. Then change coordinates in $V_n$ to send $[\eta \otimes v]$ to $[\eta \otimes x_n^0]$. Now $[\eta \otimes x_n^0]$ is in $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}\{x_n^0\})$, which is the image under $\phi$ of matrices of the form

$$\left[\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}, t\right],$$

where $P$ is a symmetric $(n-1) \times (n-1)$ submatrix of an $n \times n$ symmetric matrix.

The first claim then follows from the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-invariance of $Z_n$ and the fact that the $\mathrm{SL}(V_n)$-orbit of $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}\{x_n^0\})$ is $\mathrm{Seg}(Z_{(n-1)} \times \mathbb{P}V_n)$. □

In fact, Proposition 4.1 generalizes as follows.

**Proposition 4.2.** *Let $p+q=n$ and $Z_p \subset \mathbb{P}(V_1 \otimes \cdots \otimes V_p)$ and*

$$Z_q \subset \mathbb{P}(V_{p+1} \otimes \cdots \otimes V_n).$$

*Then $\mathrm{Seg}(Z_p \times Z_q)$ is a subvariety of $Z_n$.*

*Let $U_0 = \{[z] \in \mathbb{P}(V_1 \otimes \cdots \otimes V_n) \mid z = z_I X^I \in V_1 \otimes \cdots \otimes V_n, z_{[0,\ldots,0]} \neq 0\}$. Then $\varphi([A,t]) \in \mathrm{Seg}(Z_p \times Z_q) \cap U_0$ if and only if $A$ is of the form*

$$\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix},$$

*where $P \in S^2\mathbb{C}^p$ and $Q \in S^2\mathbb{C}^q$.*

*Proof.* Let $\varphi^i$ denote the principal minor map on $i \times i$ matrices and let $J$ and $K$ be multi-indices of length $p$ and $q$. Let $[x \otimes y] \in \mathrm{Seg}(Z_p \times Z_q)$ be such that $[x] = \varphi^p([P,r]) = [r^{p-|J|}\Delta_J(P)X^J]$ and $[y] = \varphi^q([Q,s]) = [s^{q-|K|}\Delta_K(Q)X^K]$, with $P \in S^2\mathbb{C}^p$ and $Q \in S^2\mathbb{C}^q$.

Notice that if $r=0$, then $[x] = [0,\ldots,0,\det(P)] \in \mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_p)$, and similarly if $s=0$, then $[y] = [0,\ldots,0,\det(Q)] \in \mathrm{Seg}(\mathbb{P}V_{p+1} \times \cdots \times \mathbb{P}V_{p+q})$. So the cases where $r=0$ or $s=0$ are covered by iterations of Proposition 4.1.

Now assume $r \neq 0, s \neq 0$ so we can set $r=s=1$. Consider a blocked matrix of the form

$$A = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}, \tag{2}$$

where $P \in S^2\mathbb{C}^p$ and $Q \in S^2\mathbb{C}^q$. We claim that $\varphi^{p+q}([A,1]) = [x \otimes y]$. The determinant of a block diagonal matrix is the product of the determinants of the blocks, and principal submatrices of block diagonal matrices are still block diagonal, so

$$\varphi^n([A,1]) = [\Delta_J(P)\Delta_K(Q)\ X^{J,K}],$$

where $X^{J,K} = X^J \otimes X^K$. But we can reorder the terms in the product to find

$$[\Delta_J(P)\Delta_K(Q)X^{J,K}] = [(\Delta_J(P)X^J) \otimes (\Delta_K(Q)X^K)] = [x \otimes y].$$

For the second part of the proposition, notice that for $[x \otimes y] \in \mathrm{Seg}(Z_p \times Z_q) \cap U_0$, we have exhibited a matrix $A$ as in (2) such that $\varphi^n([A,1]) = [x \otimes y]$. But symmetric matrices are determined up to sign by their $1 \times 1$ and $2 \times 2$ principal minors. Any other matrix must have the same blocked form as the one in (2). □

**Remark 4.3.** Proposition 4.2 gives a useful tool in finding candidate modules for $I(Z_n)$: we are forced to consider

$$I(Z_n) \subset \bigcap_{\substack{p+q=n \\ p,q\geq 1}} I(\mathrm{Seg}(Z_p \times Z_q)).$$

## 5. Properties of the hyperdeterminantal module

Because of Theorem 1.1, the defining ideal of $Z_n$, $\mathcal{I}(Z_n) \subset \mathrm{Sym}(V_1^* \otimes \cdots \otimes V_n^*)$, is a $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module. We consider the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module $HD = S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\ldots S_{(4)}$ (called the hyperdeterminantal module in [Holtz and Sturmfels 2007]). In this section we compute the dimension of the hyperdeterminantal module and show that it occurs with multiplicity one in $S^4(V_1^* \otimes \cdots \otimes V_n^*)$. Also, in the course of our observations, we arrive at a practical ideal membership test for a class of varieties that includes the variety of principal minors.

**Observation 5.1.** The module $S_{(2,2)}\mathbb{C}^2$ is 1-dimensional and the module $S_{(4)}\mathbb{C}^2$ is 5-dimensional and therefore the dimension of the hyperdeterminantal module is

$$\dim(S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\ldots S_{(4)}) = \binom{n}{3}5^{n-3}.$$

**Proposition 5.2.** *The module $HD = S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\ldots S_{(4)}$ occurs with multiplicity* 1 *in $S^4(V_1^* \otimes \cdots \otimes V_n^*)$. Moreover, $HD$ is an irreducible* $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$*-module.*

**Remark 5.3.** The fact that $HD$ occurs with multiplicity 1 saves us a lot of work because we do not have to worry about which isomorphic copy of the module occurs in the ideal.

*Proof of Proposition 5.2.* For the "moreover" part, recognize that by definition, $HD = S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\ldots S_{(4)}$ is a direct sum over permutations yielding distinct $\mathrm{SL}(2)^{\times n}$-modules. It is a standard fact that each summand is an irreducible $\mathrm{SL}(2)^{\times n}$-module, and this makes $HD$ an irreducible $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module.

We need to examine the $\mathrm{SL}(2)^{\times n}$-module decomposition of $S^4(V_1^* \otimes \cdots \otimes V_n^*)$. It suffices to prove for any fixed permutation $\sigma$ that $S_{(2,2)}V^*_{\sigma(1)} \otimes S_{(2,2)}V^*_{\sigma(2)} \otimes S_{(2,2)}V^*_{\sigma(3)} \otimes S_{(4)}V^*_{\sigma(4)} \otimes \cdots \otimes S_{(4)}V^*_{\sigma(n)}$ is an $\mathrm{SL}(2)^{\times n}$-module that occurs with multiplicity 1 in the decomposition of $S^4(V_1^* \otimes \cdots \otimes V_n^*)$.

We follow notation and calculations similar to [Landsberg and Manivel 2004]. For a representation $[\pi]$ of the symmetric group $\mathfrak{S}_d$, let $\chi_\pi$ denote its character. The number of occurrences of $S_{\pi_1}V_1^* \otimes \cdots \otimes S_{\pi_n}V_n^*$ in the decomposition of

$$S^d(V_1^* \otimes \cdots \otimes V_n^*)$$

is computed by $\dim(([\pi_1]\otimes\cdots\otimes[\pi_n])^{\mathfrak{S}_d})$, the dimension of the space of $\mathfrak{S}_d$ invariants. This may be computed by the formula

$$\dim(([\pi_1]\otimes\cdots\otimes[\pi_n])^{\mathfrak{S}_d}) = \frac{1}{d!}\sum_{\sigma\in\mathfrak{S}_d}\chi_{\pi_1}(\sigma)\ldots\chi_{\pi_n}(\sigma).$$

In our case, we need to compute

$$\dim\left(([(2,2)]\otimes[(2,2)]\otimes[(2,2)]\otimes[(4)]\otimes\cdots\otimes[(4)])^{\mathfrak{S}_4}\right) = \frac{1}{4!}\sum_{\sigma\in\mathfrak{S}_4}\chi_{(2,2)}(\sigma)\chi_{(2,2)}(\sigma)\chi_{(2,2)}(\sigma)\chi_{(4)}(\sigma)\ldots\chi_{(4)}(\sigma).$$

But $\chi_{(4)}(\sigma)=1$ for every $\sigma\in\mathfrak{S}_4$. So our computation reduces to

$$\dim\left(([(2,2)]\otimes[(2,2)]\otimes[(2,2)]\otimes[(4)]\otimes\cdots\otimes[(4)])^{\mathfrak{S}_n}\right) = \frac{1}{4!}\sum_{\sigma\in\mathfrak{S}_4}\chi_{(2,2)}(\sigma)\chi_{(2,2)}(\sigma)\chi_{(2,2)}(\sigma) = 1,$$

where the last equality is found by direct computation. The module $S_{(2,2)}V_1^*\otimes S_{(2,2)}V_2^*\otimes S_{(2,2)}V_3^*$ occurs with multiplicity 1 in $S^4(V_1^*\otimes V_2^*\otimes V_3^*)$. (The full decomposition of $S^4(V_1^*\otimes V_2^*\otimes V_3^*)$ was computed in [Landsberg and Manivel 2004, Proposition 4.3].) Therefore the module $S_{(2,2)}V_{\sigma(1)}^*\otimes S_{(2,2)}V_{\sigma(2)}^*\otimes S_{(2,2)}V_{\sigma(3)}^*\otimes S_{(4)}V_{\sigma(4)}^*\otimes\cdots\otimes S_{(4)}V_{\sigma(n)}^*$ occurs with multiplicity 1 in $S^4(V_1^*\otimes\cdots\otimes V_n^*)$.

We have seen that each summand of $HD$ is an irreducible $\mathrm{SL}(2)^{\times n}$-module that occurs with multiplicity 1 in $S^4(V_1^*\otimes\cdots\otimes V_n^*)$. Therefore $HD$ is an irreducible $(\mathrm{SL}(2)^{\times n}\ltimes\mathfrak{S}_n)$-module, and it occurs with multiplicity 1 in $S^4(V_1^*\otimes\cdots\otimes V_n^*)$. □

This argument generalizes:

**Lemma 5.4.** *For every collection $\pi_1,\ldots,\pi_n$ of partitions of $d$,*

$$\dim(([\pi_1]\otimes\cdots\otimes[\pi_n])^{\mathfrak{S}_d}) = \dim(([\pi_1]\otimes\cdots\otimes[\pi_n]\otimes[(d)])^{\mathfrak{S}_d}).$$

*In particular, if $M$ is any irreducible $\mathrm{SL}(V_1)\times\cdots\times\mathrm{SL}(V_n)$-module that occurs with multiplicity $m$ in $S^d(V_1^*\otimes\cdots\otimes V_n^*)$, then $M\otimes S^dV_{n+1}^*$ is an irreducible $\left(\mathrm{SL}(V_1)\times\cdots\times\mathrm{SL}(V_n)\times\mathrm{SL}(V_{n+1})\right)$-module that occurs with multiplicity $m$ in $S^d(V_1^*\otimes\cdots\otimes V_n^*\otimes V_{n+1}^*)$.*

*Proof.* Use

$$\dim(([\pi_1]\otimes\cdots\otimes[\pi_n])^{\mathfrak{S}_d}) = \frac{1}{d!}\sum_{\sigma\in\mathfrak{S}_d}\chi_{\pi_1}(\sigma)\ldots\chi_{\pi_n}(\sigma)$$

and note that $\chi_{(d)}(\sigma)=1$ for every $\sigma\in\mathfrak{S}_d$. □

**Proposition 5.5.** *The hyperdeterminantal module is contained in the ideal of the variety of principal minors of symmetric matrices, that is,*

$$HD = S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)}\dots S_{(4)} \subseteq \mathcal{I}(Z_n),$$

*and in particular,* $Z_n \subseteq \mathcal{V}(HD)$.

*Proof.* This statement is proved in [Holtz and Sturmfels 2007]. What follows is a slightly different proof that uses representation theory. Both $HD$ and $\mathcal{I}(Z_n)$ are $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-modules and $HD$ is an irreducible $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module, so we only need to show that the highest weight vector of $HD$ vanishes on all points of $Z_n$. The highest weight vector of $HD$ is the hyperdeterminant of format $2 \times 2 \times 2$ on the variables $X^{[i_1,i_2,i_3,0,\dots,0]}$. The set

$$Z_n \cap \operatorname{span}\{X^{[i_1,i_2,i_3,0,\dots,0]} \mid i_1, i_2, i_3 \in \{0, 1\}\}$$

is the set of principal minors of the upper-left $3 \times 3$ block of $n \times n$ matrices. The highest weight vector of $HD$ vanishes on these principal minors because of the case $n = 3$, so there is nothing more to show. □

**Proposition 5.6.** *Let $V$ and $W$ be complex vector spaces with* $\dim(V) \geq 2$. *Suppose $Y \subset \mathbb{P}W$ and $X \subset \mathbb{P}(V \otimes W)$ are varieties such that* $\mathrm{Seg}(Y \times \mathbb{P}W) \subset X$. *Suppose $M \subset S^dV^*$ is a space of polynomials. Then $M \otimes S^dW^* \subset \mathcal{I}_d(X)$ only if $M \subset \mathcal{I}_d(Y)$.*

*Proof.* By Lemma 5.4, it makes sense to think of $M \otimes S^dW^* \subset S^d(V^* \otimes W^*)$ as a space of polynomials.

There exists a basis of $S^dW^*$ of vectors of the form $\alpha^d$. So $M \otimes S^dW^*$ has a basis of vectors of the form $f \otimes \alpha^d$ with $f \in M$ and $\alpha \in W^*$. It suffices to prove the proposition on this basis.

Suppose $f \otimes \alpha^d$ is a basis vector in $M \otimes S^d(W^*) \subset \mathcal{I}_d(X)$. Then $\mathrm{Seg}(Y \times \mathbb{P}W) \subset X$ implies that $f \otimes \alpha^d \in \mathcal{I}_d(\mathrm{Seg}(Y \times \mathbb{P}W)) \subset S^d(V^* \otimes W^*)$. This means that $f \otimes \alpha^d(y \otimes w) = 0$ for all $y \in Y$ and for all $w \in W$. It is a fact that $\alpha^d(w) = \alpha(w)^d$ (this can be deduced from Lemma 6.3 below, for instance), so we can evaluate

$$f \otimes \alpha^d(y \otimes w) = f(y)\alpha^d(w) = f(y)\alpha(w)^d.$$

Since $\dim(V) \geq 2$, $\mathcal{V}(\alpha)$ is a hyperplane. It is no problem to choose a point that misses a hyperplane, so we can choose a particular $w \in W$ so that $\alpha(w) \neq 0$.

So we have $f(y)\alpha(w)^d = 0$ for all $y \in Y$ and $\alpha(w) \neq 0$, so $f(y) = 0$ for all $y \in Y$ and hence $f \in \mathcal{I}_d(Y)$. We can repeat the argument for any $f \in M$ we choose, so we are done. □

Proposition 5.6 fails to be an if and only if statement. Explicitly, we *cannot* say that every module in the space $\mathcal{I}_d(X)$ occurs as $M \otimes S^dV^*$ for a subset $M \subset I_d(Y)$.

In Section 6 we study the zero sets of modules of the form $I_d(Y) \otimes S^d V^*$, and this sheds light on the failure of the converse of Proposition 5.6.

**Remark 5.7.** Proposition 4.1 says that $\mathrm{Seg}(Z_n \times \mathbb{P}V_{n+1}) \subset Z_{n+1}$. We can use this proposition to study the variety of principal minors in two ways. First, if $M$ is a module in $\mathcal{J}_d(Z_n)$, then $M \otimes S^d V_{n+1}$ is a module in $\mathcal{J}_d(Z_{n+1})$. The second use is the contrapositive version. It gives an easy test for ideal membership for modules that have at least one $S_{(d)}V_i^*$ factor. Suppose we know $\mathcal{J}_d(Z_n)$ for some $n$. If we want to test whether $M = S_{\pi_1}V_1^* \otimes \cdots \otimes S_{\pi_{n+1}}V_{n+1}^*$ is in $\mathcal{J}_d(Z_{n+1})$ and we know that $M$ has at least one $\pi_i = (d)$, then we can remove $S_{\pi_i}V_i^*$ and check whether the module we have left is in $\mathcal{J}_d(Z_n)$.

## 6. A geometric characterization of the zero set of the hyperdeterminantal module via augmentation

The hyperdeterminantal module has a useful inductive description that we would like to be able to exploit. In particular, for $n \geq 3$, the module is always of the form

$$S_{(2,2)}S_{(2,2)}S_{(2,2)}S_{(4)} \dots S_{(4)},$$

where the number of $S_{(4)}$ factors is $n-3$. Then for $n \geq 3$, to construct the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module $HD$ in the case $n = k+1$ from $HD$ in the case $n = k$, we simply append another $S^4$.

More generally, if $M$ is an $\mathrm{SL}(V)$-module, we will call a $(\mathrm{SL}(V) \times \mathrm{SL}(W))$-module of the form $M \otimes S^d W^*$ an *augmentation* or *augmented module*. So for $n \geq 4$, $HD$ can be considered as the sum of augmented modules.

In this section, we study augmented modules and their zero sets in order to arrive at a geometric description of the zero set of an augmented module (Lemma 6.4). By applying this geometric description to the hyperdeterminantal module, we get a geometric description of its zero set (Lemma 6.6). This description is essential in our proof of Theorem 1.3.

**6A.** ***Polarization and its application to augmented modules.*** Augmentation is similar to prolongation, a concept found in the study of the ideals of secant varieties. A difference between the two is that augmentation does not change the degree of the polynomials, whereas prolongation increases the degree.

It is not a surprise that we can get inspiration from the techniques used to study secant varieties when studying augmented modules. In particular, polarization is a tool from classical invariant theory [Weyl 1997, pp. 5–6], and is useful in the study of ideals of secant varieties (see [Landsberg and Manivel 2003; Sidman and Sullivant 2009] for recent examples). In what follows, we use polarization to better understand the polynomials in an augmented module.

Polarization allows for the passage from a homogeneous polynomial to a symmetric multilinear form. Let $x_1, \dots, x_n$ be a basis of $V$, and let $v_i = v_{i,1}x_1 + \cdots + v_{i,n}x_n$ for $1 \le i \le d$. Given a homogeneous degree-$d$ polynomial $f$ in the polynomial ring $\mathbb{C}[x_1, \dots, x_n]$, the *polarization* of $f$ is a symmetric multilinear form $\vec{f} \in S^d V^*$, where we define $\vec{f}(v_1, \dots, v_d)$ to be the coefficient of $t_1 t_2 \dots t_d$ in the expansion of

$$f(t_1 v_1 + \cdots + t_d v_d)$$

considered as a polynomial in $t_1, \dots, t_d$. For example if $f(x_1, x_2) = (x_1)^2 x_2$, one calculates that $\vec{f}(v_1, v_2, v_3) = 2(v_{1,1}v_{2,1}v_{3,2} + v_{1,1}v_{3,1}v_{2,2} + v_{2,1}v_{3,1}v_{1,2})$.

The following useful characterization is straightforward from the definition, and while it is a consequence of classical facts [Weyl 1997], we found it stated in [Sidman and Sullivant 2009, Lemma 2.5(1)].

**Lemma 6.1.** *If $F$ is a homogeneous degree-$d$ polynomial in $x_1, \dots, x_n$, let $\vec{F}$ be its polarization. Let $v = t_1 x_1 + \cdots + t_k x_k$. Then*

$$F(v) = \vec{F}(v, \dots, v) = \sum_{\beta} \frac{1}{\beta!} \boldsymbol{t}^{\beta} \vec{F}(\boldsymbol{x}^{\beta}), \tag{3}$$

*where $\beta = (\beta_1, \dots, \beta_k)$ is a (nonnegative) partition of $d$, $\beta! = \beta_1! \dots \beta_k!$, $\boldsymbol{t}^{\beta} = t_1^{\beta_1} \dots t_k^{\beta_k}$, and $\vec{F}(\boldsymbol{x}^{\beta}) = \vec{F}(x_1^{\beta_1}, \dots, x_k^{\beta_k})$, and $x_i^{\beta_i}$ is to be interpreted as $x_i$ repeated $i$ times.*

Here is an example of the utility of this lemma that we will need later.

**Lemma 6.2.** *A linear space $L = \operatorname{span}\{x_1, \dots, x_k\}$ is a subset of $\mathcal{V}(f)$ if and only if $\vec{f}(\boldsymbol{x}^{\beta}) = 0$ for every partition $\beta$ of $d$.*

*Proof.* A linear space $L = \operatorname{span}\{x_1, \dots, x_k\}$ is in the zero set of $f$ if and only if $f(t_1 x_1 + \cdots + t_k x_k) = 0$ for all choices of $t_i \in \mathbb{C}$. Formula (3) says that

$$f(t_1 x_1 + \cdots + t_k x_k) = \sum_{\beta} \frac{1}{\beta!} \boldsymbol{t}^{\beta} \vec{f}(\boldsymbol{x}^{\beta}), \tag{4}$$

and thus implies that if $\vec{f}(\boldsymbol{x}^{\beta}) = 0$ for all $\beta$, then $f(t_1 x_1 + \cdots + t_k x_k) = 0$ for all $t_i \in \mathbb{C}$.

For the other direction, suppose $f(t_1 x_1 + \cdots + t_k x_k) = 0$ for all $t_i \in \mathbb{C}$. Consider a fixed partition $\beta'$ and take the derivative $\partial / \partial \boldsymbol{t}^{\beta'}$ of (4) to get

$$0 = \vec{f}(\boldsymbol{x}^{\beta'}) + \sum_{\beta > \beta'} \frac{1}{(\beta - \beta')!} \boldsymbol{t}^{\beta - \beta'} \vec{f}(\boldsymbol{x}^{\beta}).$$

Then take limits as $t_i \to 0$ to find that $0 = \vec{f}(\boldsymbol{x}^{\beta'})$. We do this for each $\beta'$ to conclude. □

In general, the polarization of the tensor product of two polynomials is not likely to be the product of the polarized polynomials; however, there is something we can say in the following special case:

**Lemma 6.3.** *Let $F \in S^d(W^*)$ and let $\vec{F}$ denote its polarization. Then for $\gamma \in V^*$ we have*

$$\overrightarrow{F \otimes (\gamma)^d} = \vec{F} \otimes \overrightarrow{(\gamma)^d} = \vec{F} \otimes (\gamma)^d.$$

*Proof.* A standard fact about the polarization is that $\vec{F}$ is a symmetric multilinear form. It is obvious that

$$\overrightarrow{(\gamma)^d} = (\gamma)^d,$$

because $(\gamma)^d$ is already symmetric and multilinear.

So it remains to prove the first equality in the lemma, which we do by induction on the number of terms in $F$. Suppose $F$ is a monomial, $F = \boldsymbol{w}^\alpha = w_1^{\alpha_1} \circ \cdots \circ w_n^{\alpha_n}$. Then use the isomorphism $W^{\otimes d} \otimes V^{\otimes d} \simeq (W \otimes V)^{\otimes d}$, and write $\boldsymbol{w}^\alpha \otimes \gamma^d = (w_1^{\alpha_1} \otimes \gamma^{\alpha_1}) \circ \cdots \circ (w_n^{\alpha_n} \otimes \gamma^{\alpha_n}) = (w_1 \otimes \gamma)^{\alpha_1} \circ \cdots \circ (w_n \otimes \gamma)^{\alpha_n} = (\boldsymbol{w} \otimes \gamma)^\alpha$.

If $F$ is not a monomial, suppose $F = F_1 + F_2$ with $F_i$ nonzero polynomials for $i = 1, 2$, each having strictly fewer monomials than $F$. It is clear that

$$\overrightarrow{F_1 + F_2} = \vec{F}_1 + \vec{F}_2.$$

Also, the operation $\otimes \gamma^d$ is distributive. So $\overrightarrow{F \otimes (\gamma)^d} = \overrightarrow{F_1 \otimes \gamma^d} + \overrightarrow{F_2 \otimes \gamma^d}$. By the induction hypothesis, we know that

$$\overrightarrow{F_i \otimes \gamma^d} = \vec{F}_i \otimes \gamma^d$$

for $i = 1, 2$. We conclude that $\overrightarrow{F_1 \otimes \gamma^d} + \overrightarrow{F_2 \otimes \gamma^d} = (\vec{F}_1 + \vec{F}_2) \otimes \gamma^d = \vec{F} \otimes \gamma^d$. □

The following lemma was inspired by methods found in [Landsberg and Manivel 2003]. It is a geometric description of the zero set of an augmented module.

**Lemma 6.4** (augmentation lemma). *Let $W$ and $V$ be complex vector spaces with $\dim(V) \geq 2$. Let $X \subset \mathbb{P}W$ be a variety and let $\mathcal{J}_d(X) = \mathcal{J}(X) \cap S^d W^*$ be the vector space of degree-$d$ polynomials in the ideal $\mathcal{J}(X)$. Then*

$$\mathcal{V}(\mathcal{J}_d(X) \otimes S^d V^*) = \mathrm{Seg}(\mathcal{V}(\mathcal{J}_d(X)) \times \mathbb{P}V) \cup \bigcup_{L \subset \mathcal{V}(\mathcal{J}_d(X))} \mathbb{P}(L \otimes V), \tag{5}$$

*where $L \subset \mathcal{V}(\mathcal{J}_d(X))$ are linear subspaces.*

Since the linear spaces $L$ can be one-dimensional, we do have

$$\mathrm{Seg}(\mathcal{V}(\mathcal{J}_d(X)) \times \mathbb{P}V) \subset \bigcup_{L \subset \mathcal{V}(\mathcal{J}_d(X))} \mathbb{P}(L \otimes V),$$

and we use Lemma 6.4 with the two terms on the right side of (5) combined, but we keep the two parts separate for emphasis here.

**Remark 6.5.** Note that if $\mathcal{I}(X)$ is generated in degree no larger than $d$, then one can replace $\mathcal{V}(\mathcal{I}_d(X))$ with $X$ in the statement of Lemma 6.4. We use the result of Lemma 6.4 with the induction hypothesis that $\mathcal{V}(HD) = Z_n$ and obtain a description of the zero set $\mathcal{V}(HD \otimes S^4 V_{n+1})$ in terms of the geometry of $Z_n$.

*Proof of Lemma 6.4.* First we prove "$\supseteq$". Suppose $\dim(V) = n \geq 2$. Recall that we can choose a basis of $S^d V^*$ consisting of $d$th powers of linear forms, $\{(\gamma_1)^d, \ldots, (\gamma_r)^d\}$, where $r = \binom{n+d-1}{d}$ and the $\gamma_i$ are in general linear position. It suffices to work on a basis of the vector space $\mathcal{I}_d(X) \otimes S^d V^*$. We choose a basis consisting of polynomials of the form $f \otimes \gamma^d$, with $f \in \mathcal{I}_d(X)$ and $\gamma \in V^*$.

Suppose $[x \otimes a] \in \mathrm{Seg}(\mathcal{V}(\mathcal{I}_d(X)) \times \mathbb{P}V)$ and evaluate $(f \otimes \gamma^d)(x \otimes a) = f(x)\gamma^d(a)$. But $x \in \mathcal{V}(\mathcal{I}_d(X))$, so $f(x) = 0$ for every $f \in \mathcal{I}_d(X)$, and in particular, $[x \otimes a] \in \mathcal{V}(\mathcal{I}_d(X) \otimes S^d V^*)$. So we have established that $\mathcal{V}(\mathcal{I}_d(X) \otimes S^d V^*) \supset \mathrm{Seg}(\mathcal{V}(\mathcal{I}_d(X)) \times \mathbb{P}V)$.

Now suppose $[v] \in \mathbb{P}(L \otimes V)$ for some linear subspace $L = \mathrm{span}\{x_1, \ldots, x_l\} \subset \mathcal{V}(\mathcal{I}_d(X))$. By expanding an expression of $[v]$ in bases and collecting the coefficients of the $x_i$, we can write $[v] = [x_1 \otimes a_1 + \cdots + x_l \otimes a_l]$ for $a_i \in V$ not all zero. Consider $f \otimes \gamma^d \in \mathcal{I}_d(X) \otimes S^d V$. By Lemma 6.3,

$$\overrightarrow{f \otimes \gamma^d} = \vec{f} \otimes \gamma^d,$$

and using the polarization formula (3), we write

$$(f \otimes \gamma^d)(v) = (\vec{f} \otimes \gamma^d)(v, \ldots, v) = \sum_\beta \frac{1}{\beta!} \vec{f}(\boldsymbol{x}^\beta)\gamma^d(\boldsymbol{a}^\beta).$$

The choice of $L \subset \mathcal{V}(\mathcal{I}_d(X))$ means that $L \subset \mathcal{V}(f)$, so by Lemma 6.2, $\vec{f}(\boldsymbol{x}^\beta) = 0$ for all $\beta$. Every term of $(f \otimes \gamma^d)(v)$ vanishes, so $(f \otimes \gamma^d)(v) = 0$, and hence $[v] \in \mathcal{V}(\mathcal{I}_d(X) \otimes S^d V^*)$. So we have established that $\mathcal{V}(\mathcal{I}_d(X) \otimes S^d V^*) \supset \mathbb{P}(L \otimes V)$ for all linear subspaces $L \subset \mathcal{V}(\mathcal{I}_d(X))$.

Now we prove "$\subseteq$". Consider any $[v] \in \mathbb{P}(W \otimes V)$. Choose a basis $\{a_1, \ldots, a_k\}$ of $V$ (by assumption $k \geq 2$). Then expand the expression of $v$ in bases and collect the coefficients of each $a_i$ to find $[v] = [x_1 \otimes a_1 + \cdots + x_k \otimes a_k]$ with $x_1, \ldots, x_k \in W$ and not all $x_i$ zero.

We need to show that $[v] \in \mathbb{P}(L \otimes V)$ for a linear space $L \subset \mathcal{V}(\mathcal{I}_d(X))$. The natural linear space to consider is $L = \mathrm{span}\{x_1, \ldots, x_k\}$. Since we already have an expression $[v] = [x_1 \otimes a_1 + \cdots + x_k \otimes a_k]$, if we can show that $L = \mathrm{span}\{x_1, \ldots, x_k\} \subset \mathcal{V}(\mathcal{I}_d(X))$, we will be done.

For any $f \otimes \gamma^d \in \mathcal{I}_d(X) \otimes S^d V^*$ we can write

$$0 = (f \otimes \gamma^d)(v) = \sum_\beta \frac{1}{\beta!} \vec{f}(\boldsymbol{x}^\beta)\gamma^d(\boldsymbol{a}^\beta). \tag{6}$$

Let $\{\check{a}_1, \ldots, \check{a}_k\}$ be a basis of $V^*$ dual to $\{a_1, \ldots, a_k\}$. Then let $\gamma$ vary continuously

in $V^*$ by writing it as

$$\gamma = t_1\check{a}_1 + \cdots + t_k\check{a}_k,$$

where the parameters $t_i \in \mathbb{C}$ vary. The polynomial $\gamma^d$ is simple enough that we can expand it as follows:

$$\gamma^d(\boldsymbol{a}^\beta) = \gamma^d(a_1^{\beta_1}, \ldots, a_k^{\beta_k}) = \gamma(a_1)^{\beta_1} \ldots \gamma(a_k)^{\beta_k}.$$

But our choices have made it so that $\gamma(a_i) = t_i$, and therefore $\gamma^d(\boldsymbol{a}^\beta) = \boldsymbol{t}^\beta$. So (6) becomes

$$0 = (f \otimes \gamma^d)(v) = \sum_\beta \frac{1}{\beta!}\vec{f}(\boldsymbol{x}^\beta)\boldsymbol{t}^\beta = f(t_1x_1 + \cdots + t_kx_k),$$

where we have used Lemma 6.3. So $f(t_1x_1 + \cdots + t_kx_k) = 0$ for all $t_i \in \mathbb{C}$, and this is an equivalent condition that $L = \operatorname{span}\{x_1, \ldots, x_k\}$ is a subspace of $\mathcal{V}(f)$. Since this was done for arbitrary $f \in \mathcal{I}_d(X)$, we conclude that $L \subset \mathcal{V}(\mathcal{I}_d(X))$. □

Now we can apply this geometric characterization of augmentation to the hyperdeterminantal module. To do this we need to set up more notation.

Assume $n \geq 4$. Let $HD_i$ be the image of the hyperdeterminantal module at stage $n-1$ under the reindexing isomorphism

$$S^4(V_1^* \otimes \cdots \otimes V_{n-1}^*) \longrightarrow S^4(V_1^* \otimes \cdots \otimes V_{i-1}^* \otimes V_{i+1}^* \otimes \cdots \otimes V_n^*),$$

where we still have $n-1$ vector spaces $V_i \simeq \mathbb{C}^2$, but we have shifted the index on the last $n-i$ terms. Then the hyperdeterminantal module at stage $n$ can be expressed as a sum of augmented modules as follows:

$$HD = \sum_{i=1}^n (HD_i \otimes S^4V_i^*).$$

Finally note that if $\dim(V) = k$, then $\sigma_s(\mathbb{P}W \times \mathbb{P}V) = \mathbb{P}(W \otimes V)$ for all $s \geq k$. In the case $V_i \simeq \mathbb{C}^2$, we have $\mathbb{P}(L \otimes V_i) = \sigma_2(\mathbb{P}L \times \mathbb{P}V_i)$. Certainly

$$\operatorname{Seg}(\mathcal{V}(M_i) \times \mathbb{P}V_i) \subset \bigcup_{L \subset V(M_i)} \mathbb{P}(L \otimes V_i),$$

for any modules of polynomials $M_i$. If $L \subset \mathcal{V}(\mathcal{I}_d(X))$, then $\sigma_s(\mathbb{P}L \times \mathbb{P}V) \subseteq \sigma_s(\mathcal{V}(\mathcal{I}_d(X)) \times \mathbb{P}V)$. If $A, B, C$ are vector spaces of polynomials and $C = A + B$, then $\mathcal{V}(C) = \mathcal{V}(A) \cap \mathcal{V}(B)$. Collecting these ideas, we apply the Augmentation Lemma 6.4 to the hyperdeterminantal module to yield the following:

**Lemma 6.6** (characterization lemma). *Consider*

$$\sum_{i=1}^n HD_i \otimes S^dV_i^* \subset S^d(V_1^* \otimes \cdots \otimes V_n^*).$$

*Then*

$$\mathcal{V}\Big(\sum_{i=1}^{n} HD_i \otimes S^d V_i^*\Big) = \bigcap_{i=1}^{n}\Big(\bigcup_{L\subset V(HD_i)} \mathbb{P}(L\otimes V_i)\Big) \subseteq \bigcap_{i=1}^{n}(\sigma_2(\mathcal{V}(HD_i)\times \mathbb{P}V_i)).$$

**Remark 6.7.** A consequence of the characterization lemma is the following test for nonmembership in the zero set of $HD$. Suppose $[z] = [\zeta^1 \otimes x_i^1 + \zeta^2 \otimes x_i^2] \in \mathbb{P}^{2^n-1}$. If either $[\zeta^1]$ or $[\zeta^2]$ is not a vector of principal minors of an $(n-1)\times(n-1)$ symmetric matrix, then $[z]$ is not a zero of the hyperdeterminantal module $HD$ and hence not a vector of principal minors of a symmetric matrix, since $\mathcal{V}(HD) \supset Z_n$. This observation can be iterated, and each iteration cuts the size of the vector in question in half until one need only check honest hyperdeterminants of format $2\times 2\times 2$. This test, while relatively cheap and accessible, is necessary but not sufficient, as is pointed out in [Holtz and Sturmfels 2007].

It is well-known that the ideal of the Segre product of an arbitrary number of projective spaces is generated in degree 2 by the $2\times 2$ minors of flattenings. In essence, this is saying that all of the polynomials in the ideal come from the Segre products of just two projective spaces. The following is a weaker, strictly set-theoretic result in the same spirit. It is another application of the Augmentation Lemma 6.4, and its proof is mimicked in the proof of Lemma 7.11 below.

**Proposition 6.8.** *For $1 \le i \le n$, let $V_i$ be complex vector spaces each with dimension $\ge 2$ and assume $n \ge 2$. If for each $i$, $B_i \subset S^d(V_1^* \otimes \dots V_{i-1}^* \otimes V_{i+1}^* \otimes \cdots \otimes V_n^*)$ is a set of polynomials with the property*

$$\mathcal{V}(M^i) = \mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_{i-1} \times \mathbb{P}V_{i+1} \times \cdots \times \mathbb{P}V_n),$$

*then*

$$\mathcal{V}\Big(\bigoplus_i (M_i \otimes S^d V_i^*)\Big) = \mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n).$$

*Proof.* Work by induction and use Lemma 6.4. It is clear that

$$\mathcal{V}\Big(\bigoplus_i (M_i \otimes S^d V_i^*)\Big) \supset \mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n).$$

All the linear spaces on $\mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n)$ are (up to permutation) of the form $V_1 \otimes \widehat{a_2} \otimes \cdots \otimes \widehat{a_n}$, where $a_i \in V_i$ are nonzero and $\widehat{a_i}$ denotes the line through $a_i$. Then compute the intersection, $\bigcup_{L^i} \bigcap_{i=1}^n \mathbb{P}(L^i \otimes V_i)$, and notice that in the intersection of just 3 factors, all of the resulting linear spaces must live in $\mathrm{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n)$. □

## 7. Understanding the case when two zeros of the hyperdeterminantal module disagree in precisely one coordinate

In the proof of Theorem 1.3 below, we work to construct a matrix whose principal minors are a given point in the zero set of the hyperdeterminantal module. The main difficulty is the following. Suppose we have a point $[z] \in \mathcal{V}(HD)$ and a candidate matrix $A$ that satisfies $\Delta_I(A) = z_I$ for all $I \neq [1, \dots, 1]$. In other words, all of the principal minors of $A$ except possibly for the determinant agree with the entries of $z$. What can we say about $z$?

To answer this question, we must study the points in $\widehat{\mathcal{V}(HD)}$ that have all of their coordinates except one equal. Geometrically, we need to understand the points for which a line in the coordinate direction $X^{[1,\dots,1]}$ above the point $z$ intersects $\widehat{\mathcal{V}(HD)}$ in at least two points. We answer this question in Lemma 7.2 below. Using that lemma, we find the following.

**Proposition 7.1.** *Let $n \geq 4$. Let $z = z_I X^I$ and $w = w_I X^I$ be points in $\widehat{\mathcal{V}(HD)}$. If $z_I = w_I$ for all $I \neq [1, \dots, 1]$ and $z_{[0,\dots,0]} \neq 0$, then $z = w$.*

For the rest of this section we use the following notation. If $K = \{k_1, \dots, k_s\} \subset \{1, \dots, n\}$ and $1 \leq k_j \leq n$ for all $j$, then let $V_K \simeq V_{k_1} \otimes \cdots \otimes V_{k_s}$ for $s \leq n$. We assume $V_k \simeq \mathbb{C}^2$ for all $k$, so that $V_K \simeq (\mathbb{C}^2)^{\otimes s}$. Let $\mathcal{P}^2(\{n_1, \dots, n_n\})$ denote the collection of all partitions of $\{n_1, \dots, n_n\}$ into mutually disjoint subsets of cardinality 2 or less, that is, $\mathcal{P}^2(\{1, \dots, n\})$ consists of the sets $\{K_1, \dots, K_m\}$ such that $K_p \subset \{1, \dots, n\}$ and $|K_p| \leq 2$ for every $1 \leq p \leq m$, $K_p \cap K_q = \varnothing$ whenever $p \neq q$, and $\cup_{p=1}^m K_p = \{1, \dots, n\}$.

**Lemma 7.2.** *Let $n \geq 4$. Let $z = z_I X^I$ and $w = w_I X^I$ be points in $\widehat{\mathcal{V}(HD)}$. If $z_I = w_I$ for all $I \neq [1, \dots, 1]$ but $z_{[1,\dots,1]} \neq w_{[1,\dots,1]}$, then*

$$[z], [w] \in \bigcup_{\{K_1,\dots,K_m\} \in \mathcal{P}^2(\{1,\dots,n\})} \operatorname{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m}) \subset Z_n.$$

Note that the notationally dense Segre product is just a product of $\mathbb{P}^3$'s and $\mathbb{P}^1$'s.

*Proof of Proposition 7.1.* Assume Lemma 7.2. Let $z = z_I X^I$ and $w = w_I X^I$ be points in $\widehat{\mathcal{V}(HD)} \cap \{z \mid z_{[0,\dots,0]} \neq 0\}$. Suppose that $z_I = w_I$ for all $I \neq [1, \dots, 1]$, and suppose for contradiction that $z_{[1,\dots,1]} \neq w_{[1,\dots,1]}$. Lemma 7.2 implies that $[z]$, $[w]$ are in a Segre product of $\mathbb{P}^1$'s and $\mathbb{P}^3$'s.

$Z_1 \simeq \mathbb{P}^1$ and $Z_2 \simeq \mathbb{P}^3$, and Proposition 4.2 implies that a point $[A, t]$ with $t \neq 0$ mapping to $\operatorname{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m})$ with $\{K_1, \dots, K_m\} \in \mathcal{P}^2(\{1, \dots, n\})$ is permutation equivalent to a block diagonal matrix consisting of $1 \times 1$ and $2 \times 2$ blocks. Such a block diagonal matrix is a special case of a symmetric tridiagonal matrix, and therefore none of its principal minors depends on the sign of the off-diagonal terms. So fixing the $0 \times 0$, $1 \times 1$ and $2 \times 2$ principal minors fixes the rest

of the principal minors in such a matrix. If we take $z_{[0,\dots,0]} = w_{[0,\dots,0]} = 1$ and assume the $1 \times 1$ and $2 \times 2$ principal minors agree, then the rest of the principal minors must agree, including the determinants; thus the contradiction.

The assumption $z_{[0,\dots,0]} \neq 0$ is necessary. If $z_{[0,\dots,0]} = 0$, then consider the image of any two matrices $A, B$ with different nonzero determinants under the principal minor map with $t = 0$. Then $\varphi([A, 0]) = [0, \dots, 0, \det(A)] \neq \varphi([B, 0]) = [0, \dots, 0, \det(B)]$. □

**Remark 7.3.** A key point is that we are not making the claim in Proposition 7.1 for $n = 3$. In this case any two zeros of the hyperdeterminant are principal minors of $3 \times 3$ matrices that differ up to sign of the off-diagonal terms. Altering the sign of the off-diagonal terms of a $3 \times 3$ symmetric matrix can change the determinant without changing the other principal minors and without forcing the matrix to be blocked as a $2 \times 2$ block and a $1 \times 1$ block.

**Remark 7.4.** To see that the analog of Proposition 7.1 holds for $Z_n$ with $n \geq 4$ and $t \neq 0$ requires much less work than the case of $\mathcal{V}(HD)$. We used Maple to construct a generic symmetric $4 \times 4$ matrix and computed its principal minors. Then we changed the signs of the off-diagonal terms in every possible combination and compared the number of principal minors that agreed with the principal minors of the original matrix. The result was that the two vectors of principal minors could agree in precisely 11, 13 or 16 entries, but not 15. (Though tedious, the $4 \times 4$ case can also be proved without a computer by analyzing the parity of the various products of the off-diagonal terms in the matrix.) We repeated the experiment in the $5 \times 5$ case and found that the two vectors could agree in precisely 16, 19, 20, 21, 23, 25 or 32 positions, but never 31 positions.

The general case follows from the $4 \times 4$ case by the following. Suppose $n \geq 4$ and $2^n - 1$ of the principal minors of an $n \times n$ symmetric matrix agree with the principal minors of another $n \times n$ symmetric matrix. Then we may assume that the $0 \times 0$, $1 \times 1$ and $2 \times 2$ principal minors of both matrices agree and hence the matrices must agree up to the signs of the off-diagonal terms. Then use the group to move the one position where the principal minors don't agree to be a $4 \times 4$ determinant and use the $4 \times 4$ result for the contradiction.

To prove Lemma 7.2, we will show that if $w_I = z_I$ for all $I \neq [1, \dots, 1]$ and $z_{[1,\dots,1]} \neq w_{[1,\dots,1]}$, then $z$ is a zero of an auxiliary set of polynomials denoted $B$. We then show that the zero set $\mathcal{V}(B)$ is contained in the union of Segre varieties. Finally, Proposition 4.2 provides the inclusion into $Z_n$.

**7A.** ***Reduction to one variable.*** Let $n \geq 4$. Suppose $z = z_I X^I$ and $w = w_I X^I$ are points in $\widehat{\mathcal{V}(HD)}$ such that $z_I = w_I$ for all $I \neq [1, \dots, 1]$. Both points are zeros of every polynomial in $HD$, but the only coordinate in which they can differ

is $[1, \dots, 1]$. Now consider the coordinates $z_I$ $(= w_I)$ as fixed constants for all $I \neq [1, \dots, 1]$, and for $f \in HD$ define $f_z$ by the substitution

$$f(X^{[0,\dots,0]}, \dots, X^{[1,\dots,1]}) \mapsto f(z_{[0,\dots,0]}, \dots, z_{[0,1,\dots,1]}, X^{[1,\dots,1]}) =: f_z(X^{[1,\dots,1]}).$$

Let $HD_{[1,\dots,1]}(z) = \{f_z \mid f \in HD\}$ denote the resulting set of univariate polynomials. Then $z_{[1,\dots,1]}$ and $w_{[1,\dots,1]}$ are two (possibly different) roots of each univariate polynomial $f_z \in HD_{[1,\dots,1]}(z)$.

**Lemma 7.5.** *If $f \in HD$, then the corresponding polynomial $f_z \in HD_{[1,\dots,1]}(z)$ is either degree* 0, 1, *or* 2 *in $X^{[1,\dots,1]}$.*

*Proof.* It suffices to prove the statement for $f \in S_{(2,2)}V_1^* \otimes S_{(2,2)}V_2^* \otimes S_{(2,2)}V_3^* \otimes S_{(4)}V_4^* \otimes \cdots \otimes S_{(4)}V_n^*$. Suppose for contradiction that $f$ has a monomial of the form $(X^{[1,\dots,1]})^3 X^{[i_1,\dots,i_n]}$. Its possible weights are $(2+2i_1, \dots, 2+2i_n)$ with $i_j \in \{0, 1\}$. However, the weight of every polynomial in $S_{(2,2)}V_1^* \otimes S_{(2,2)}V_2^* \otimes S_{(2,2)}V_3^* \otimes S_{(4)}V_4^* \otimes \cdots \otimes S_{(4)}V_n^*$ is of the form $(0, 0, 0, w_4, \dots, w_n)$, where $w_i$ are even integers with $|w_i| \leq 4$ for $4 \leq i \leq n$, a contradiction since obviously $0 \neq 2 + 2i_1$ for any $i_1 \in \{0, 1\}$. Therefore the degree of $f$ is less than 3 in $X^{[1,\dots,1]}$. □

Now we know that $w_{[1,\dots,1]}$ and $z_{[1,\dots,1]}$ are both common zeros of univariate polynomials, all with degree 2 or less. The fact that $w_{[1,\dots,1]}$ and $z_{[1,\dots,1]}$ are both common zeros of more than one univariate polynomial comes from the fact that we have required $n \geq 4$; otherwise there is only one polynomial and what we are about to do would be trivial.

A quadratic (not identically zero) in one variable has at most two solutions, and a linear polynomial (not identically zero) has at most one solution. The only way then for us to have $w \neq z$ and $[w], [z] \in \mathcal{V}(HD)$ is if *all* of the linear polynomials were identically zero and if *all* of the quadratics were scalar multiples of each other.

Therefore, we need to study the points $[z] \in \mathcal{V}(HD)$ for which $HD_{[1,\dots,1]}(z)$ has dimension 1 or less. Define polynomials $a_f$, $b_f$, and $c_f$ (which necessarily do not depend on $X^{[1,\dots,1]}$) for each $f_z \in HD_{[1,\dots,1]}(z)$ by

$$f_z = a_f(z)(X^{[1,\dots,1]})^2 + b_f(z)(X^{[1,\dots,1]}) + c_f(z).$$

The requirement that $HD_{[1,\dots,1]}(z)$ have dimension 1 or less implies the weaker (but still sufficient) condition that $z$ be a root of the polynomials

$$B' := \operatorname{span}\{a_f b_g - a_g b_f \mid f, g \in HD\}.$$

The polynomials in $B'$ have the property that if $h(z) \neq 0$ for a nonzero $h \in B'$, that is, $[z] \notin \mathcal{V}(B')$, then there is a nontrivial pair of polynomials in $HD_{[1,\dots,1]}(z)$ that are not scalar multiples of each other, and thus the zero set of $HD_{[1,\dots,1]}(z)$ is a single point. In this case we must have $w_{[1,\dots,1]} = z_{[1,\dots,1]}$. If, however, $h(z) = 0$ for

all $h \in B'$, that is, $z \in \mathcal{V}(B')$, then it is possible that the polynomials in $HD_{[1,\dots,1]}(z)$ have 2 common roots.

$B'$ is not $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-invariant. Let $B := \mathrm{span}\{(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n).B'\}$ denote the corresponding $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module. If $g.[z] \notin \mathcal{V}(B')$, then by our remarks above, $g.[z] \in Z_n$, and in particular, $[z] \in Z_n$ (because $Z_n$ is a $G$-variety). The following lemma allows us to compare $G$-orbits of points and the zero sets of arbitrary sets of polynomials (not necessarily $G$-modules).

**Lemma 7.6.** *Let $z \in \mathbb{P}V$, let $G \subset \mathrm{GL}(V)$ be a group, and let $M \subset \mathrm{Sym}(V^*)$ be a collection of polynomials ($M$ is not necessarily a $G$-module). Then*

$$G.z \subset \mathcal{V}(M) \text{ if and only if } z \in \mathcal{V}(\mathrm{span}\{G.M\}).$$

*Proof.* $G.z \subset \mathcal{V}(M)$ if and only if $f(g.z) = 0$ for all $g \in G$ and for all $f \in M$. But from the definition of the $G$-action on the dual space, $f(g.z) = (g^{-1}.f)(z)$, so $f(g.z) = 0$ for all $g \in G$ and for every $f \in M$. This happens if and only if $(g.f)(z) = 0$ for all $g \in G$ and for all $f \in M$, but this is the condition that $z \in \mathcal{V}(\mathrm{span}\{G.M\})$. □

We apply Lemma 7.6 to our setting; if $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n).[z] \subset \mathcal{V}(B')$, then $[z] \in \mathcal{V}(B)$ (recall $B := \mathrm{span}\{(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n).B'\}$). So, we need to look at the variety $\mathcal{V}(B)$. We conclude that our construction satisfies the property that if $[z] \in \mathcal{V}(HD)$ but $[z] \notin \mathcal{V}(B)$, then $[z] \in Z_n$.

We need to understand the types of points that can be in $\mathcal{V}(B)$, and the following proposition gives sufficient information about $\mathcal{V}(B)$.

**Proposition 7.7.** *Let $n \geq 4$ and let $B$ be the module of polynomials constructed above. Let $\mathcal{P}^2(\{1,\dots,n\})$ be the collection of all partitions of $\{1,\dots,n\}$ into mutually disjoint subsets of cardinality 2 or less. Then*

$$\mathcal{V}(B) \subset \bigcup_{\{K_1,\dots,K_m\} \in \mathcal{P}^2(\{1,\dots,n\})} \mathrm{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m}) \subset Z_n.$$

*Proof.* Proposition 7.7 is proved in several parts. In Lemma 7.8 we find the module $S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\dots S_{(5)}$ as a submodule of $B$ using the algorithm in Section 3E. In Lemma 7.12 we identify the zero set of this new module. In particular, we show

$$\mathcal{V}(S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\dots S_{(5)}) = \bigcup_{\{K_1,\dots,K_m\} \in \mathcal{P}^2(\{1,\dots,n\})} \mathrm{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m}).$$

We prove this statement by induction on $n$, where we prove the base case $n = 3$ in Lemma 7.9 and the induction step in Lemma 7.11. Finally, each $\mathbb{P}V_{K_i}$ is either a copy of $\mathbb{P}^1 \cong Z_1$ or $\mathbb{P}^3 \cong Z_2$, so we can apply Proposition 4.2 to verify the inclusion

$$\mathrm{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m}) \subset Z_n.$$ □

**Lemma 7.8.** *Suppose $n \geq 4$ and let $B$ be constructed as above. Then*

$$S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\dots S_{(5)} \subset B.$$

*Proof.* Here we have a subset of polynomials in $B$ in an explicit form, and we would like to identify $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-modules in $B$ from this information. To do this we use the ideas presented in Section 3D and particularly the algorithm presented in Section 3E. It suffices to work first with $\mathrm{SL}(2)^{\times n}$-modules and later consider the permutations.

Suppose $f_{k_1,k_2,k_3} \in S_{(2,2)}V_{k_1}^* \otimes S_{(2,2)}V_{k_2}^* \otimes S_{(2,2)}V_{k_3}^* \otimes S_{(4)}V_{k_4}^* \otimes \cdots \otimes S_{(4)}V_{k_n}^*$ is a lowest weight vector. Define $a_{k_1,k_2,k_3}$, $b_{k_1,k_2,k_3}$, $c_{k_1,k_2,k_3}$ by the equation $f_{k_1,k_2,k_3} = a_{k_1,k_2,k_3}(X^{[1,\dots,1]})^2 + b_{k_1,k_2,k_3}(X^{[1,\dots,1]}) + c_{k_1,k_2,k_3}$.

For this proof, we introduce new notation. If $k_1, k_2, k_3$ are fixed, let $X^{I_{p,q,r}}$ denote the coordinate vector with $k_1 = p$, $k_2 = q$, $k_3 = r$ and $k_s = 0$ for $s \geq 4$.

Since $f_{k_1,k_2,k_3}$ is a hyperdeterminant of format $2 \times 2 \times 2$, we find

$$\begin{aligned} a_{k_1,k_2,k_3} &= (X^{I_{0,0,0}})^2. \\ b_{k_1,k_2,k_3} &= -2X^{I_{0,0,0}}(X^{I_{1,0,0}}X^{I_{0,1,1}} + X^{I_{0,1,0}}X^{I_{1,0,1}} + X^{I_{0,1,0}}X^{I_{1,1,0}}) \\ &\quad + 4X^{I_{1,0,0}}X^{I_{0,1,0}}X^{I_{0,0,1}}. \end{aligned}$$

The weight of $a_{k_1,k_2,k_3}$ is (up to permutation) $(-2,-2,-2,2,\dots,2)$, where the $-2$'s actually occur at $\{k_1,k_2,k_3\}$. The weight of $b_{k_1,k_2,k_3}$ is (up to permutation) $(-1,-1,-1,3,\dots,3)$, where the $-1$'s actually occur at $\{k_1,k_2,k_3\}$. Now consider

$$h_{k_1,k_2,k_3,j_1,j_2,j_3} = a_{k_1,k_2,k_3}b_{j_1,j_2,j_3} - a_{j_1,j_2,j_3}b_{k_1,k_2,k_3} \in B.$$

We notice that $h_{k_1,k_2,k_3,j_1,j_2,j_3}$ cannot have $k_1, k_2, k_3$ and $j_1, j_2, j_3$ all equal (this is the zero polynomial). So either two, one or zero pairs of $i$'s and $j$'s match in the indices $k_1, k_2, k_3$ and $j_1, j_2, j_3$. Therefore $h_{k_1,k_2,k_3,j_1,j_2,j_3}$ can have 3 different weights (up to permutation), depending on how $k_1, k_2, k_3$ and $j_1, j_2, j_3$ match up. The three possible weights of $h_{k_1,k_2,k_3,j_1,j_2,j_3}$ are (up to permutation): $(-3,-3,1,1,5,\dots,5)$, $(-3,1,1,1,1,5,\dots,5)$, or $(1,1,1,1,1,1,5,\dots,5)$.

In each case, apply the algorithm in Section 3E and lower $h_{k_1,k_2,k_3,j_1,j_2,j_3}$ to a nonzero vector with the lowest possible weight. We did this calculation in Maple. The output in each case is a vector of weight (up to permutation) $(3,3,3,5,\dots,5)$. Use Remark 3.2 to identify the module with lowest weight $(3,3,3,5,\dots,5)$ as

$$S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\dots S_{(5)},$$

and this must be a submodule of $B$. □

**Lemma 7.9.** *As sets in $\mathbb{P}(V_1 \otimes V_2 \otimes V_3)$,*

$$\begin{aligned} &\mathcal{V}(S_{(4,1)}V_1^* \otimes S_{(4,1)}V_2^* \otimes S_{(4,1)}V_3^*) \\ &\quad = \mathrm{Seg}(\mathbb{P}(V_1 \otimes V_2) \times \mathbb{P}V_3) \cup \mathrm{Seg}(\mathbb{P}(V_1 \otimes V_3) \times \mathbb{P}V_2) \cup \mathrm{Seg}(\mathbb{P}(V_1) \times (V_1 \otimes V_2). \end{aligned}$$

*Proof.* The space $V_1^* \otimes V_2^* \otimes V_3^*$ has seven orbits under the action of $\mathrm{SL}(2)^{\times 3}$ [Gelfand et al. 1994, Example 4.5]. This gives rise to a list of normal forms, which we record below together with the respective $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-orbit closures to which they belong.

- The trivial orbit, $\varnothing$.
- $\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3)$: Normal form $[x] = [a \otimes b \otimes c]$.
- $\tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))_{\mathrm{sing}} = \mathfrak{S}_3.\,\mathrm{Seg}(\mathbb{P}(V_1 \otimes V_2) \times \mathbb{P}V_3)$: Normal form (up to permutation) $[x] = [a \otimes b \otimes c + a' \otimes b' \otimes c]$. This union of 3 irreducible varieties is the singular set of the next orbit.
- $\tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))$: Normal form
$$[x] = [a \otimes b \otimes c + a' \otimes b \otimes c + a \otimes b' \otimes c + a \otimes b \otimes c'].$$
- $\sigma(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))$: Normal form $[x] = [a \otimes b \otimes c + a' \otimes b' \otimes c']$.

The orbit closures are nested:

$$\begin{aligned}\varnothing \subset \mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3) \subset \tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))_{\mathrm{sing}} \\ \subset \tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3)) \subset \sigma(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3)) = \mathbb{P}^7.\end{aligned}$$

The lowest weight vector for $S_{(4,1)}S_{(4,1)}S_{(4,1)}$ is

$$\begin{aligned}f_{(4,1),(4,1),(4,1)} = (X^{[1,1,1]})^2\big(X^{[0,0,0]}(X^{[1,1,1]})^2 + 2X^{[1,0,1]}X^{[0,1,1]}X^{[1,1,0]} \\ - X^{[1,1,1]}(X^{[0,1,1]}X^{[1,0,0]} + X^{[1,0,1]}X^{[0,1,0]} + X^{[1,1,0]}X^{[0,0,1]})\big).\end{aligned}$$

We took a generic point $x \in \tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))_{\mathrm{sing}}$ and evaluated

$$f_{(4,1),(4,1),(4,1)}(x) = 0.$$

So $\tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))_{\mathrm{sing}} \subset \mathcal{V}(S_{(4,1)}S_{(4,1)}S_{(4,1)})$. We could also see this without calculation by noticing that any point of the form $[a \otimes b \otimes c + a' \otimes b' \otimes c]$ lives in $\mathbb{P}(V_1 \otimes V_2 \otimes \widehat{c})$, where $\widehat{c}$ is the line through $c$. But every point in this space is a zero of $S_{(4,1)}S_{(4,1)}S_{(4,1)})$ because $S_{(4,1)}(\widehat{c})^* = 0$.

Next we show that the other two varieties are not in $\mathcal{V}(S_{(4,1)}S_{(4,1)}S_{(4,1)})$. The varieties are nested, so take the point $[x] = \big[X^{[1,1,1]} + X^{[0,1,1]} + X^{[1,0,1]} + X^{[1,1,0]}\big] \in \tau(\mathrm{Seg}(\mathbb{P}V_1 \times \mathbb{P}V_2 \times \mathbb{P}V_3))$. But $f_{(4,1),(4,1),(4,1)}(x) = 2 \neq 0$, so the other two varieties are not in $\mathcal{V}(S_{(4,1)}S_{(4,1)}S_{(4,1)})$. Since we have considered all possible normal forms, we are done. □

**Observation 7.10.** All the linear spaces on $\mathrm{Seg}(\mathbb{P}V_{K_1} \times \cdots \times \mathbb{P}V_{K_m})$ are (up to permutation) contained in one of the form $V_{K_1} \otimes \widehat{v_{K_2}} \otimes \cdots \otimes \widehat{v_{K_m}}$, where $\widehat{v_K}$ denotes the line through $v_{k_1} \otimes \cdots \otimes v_{k_s}$ in $V_K$.

Let $\mathcal{P}_{p,q}(\{n_1,\dots,n_{p+q}\})$ denote the set of partitions of $\{n_1,\dots,n_{p+q}\}$ into two disjoint sets of cardinality $p$ and $q$.

Consider the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module $\tilde{B} = S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\dots S_{(5)}$ that has $n-3$ copies of $S_{(5)}$. We write in $\tilde{B}$ in full detail as

$$\tilde{B} = \bigoplus_{\substack{\{\{k_1,k_2,k_3\},\{k_4,\dots,k_n\}\}\\ \in\mathcal{P}_{3,n-3}(\{1,\dots,n\})}} S_{(4,1)}V_{k_1}^* \otimes S_{(4,1)}V_{k_2}^* \otimes S_{(4,1)}V_{k_3}^* \otimes S_{(5)}V_{k_4}^* \otimes \cdots \otimes S_{(5)}V_{k_n}^*.$$

Let $\tilde{B}_k$ denote the $\mathrm{SL}(2)^{n-1} \ltimes \mathfrak{S}_{n-1}$ module

$$\tilde{B}_k = \bigoplus_{\substack{\{\{k_1,k_2,k_3\},\{k_4,\dots,k_{n-1}\}\}\\ \in\mathcal{P}_{3,n-4}(\{1,\dots,n\}\setminus\{k\})}} S_{(4,1)}V_{k_1}^* \otimes S_{(4,1)}V_{k_2}^* \otimes S_{(4,1)}V_{k_3}^* \otimes S_{(5)}V_{k_4}^* \otimes \cdots \otimes S_{(5)}V_{k_{n-1}}^*.$$

$\tilde{B}\sum_{i=1}^n \tilde{B}_i \otimes S_{(5)}V_i^*$, that is, the $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-module $\tilde{B}$ is constructed as the nonredundant sum over permutations of augmented $\mathrm{SL}(2)^{\times n-1}$-modules.

We want to understand the zero set of this module $\tilde{B}$, and we do this in the next two lemmas by mimicking what we did for Proposition 6.8. We also point out that while notationally more complicated, the resulting Lemma 7.12 is essentially the same idea as Proposition 6.8.

**Lemma 7.11.** *Suppose $n \geq 4$ and let $\tilde{B}$ and $\tilde{B}_k$ be as above. If*

$$\mathcal{V}(\tilde{B}_k) = \bigcup_{\{K_1,\dots,K_m\}\in\mathcal{P}^2(\{1,\dots,n\}\setminus\{k\})} \mathrm{Seg}(\mathbb{P}V_{K_1} \times \mathbb{P}V_{K_2} \times \cdots \times \mathbb{P}V_{K_m}),$$

*then*

$$\mathcal{V}(\tilde{B}_k \otimes S_{(5)}V_k^*) = \bigcup_{\{K_1,\dots,K_m\}\in\mathcal{P}^2(\{1,\dots,n\}\setminus\{k\})} \mathrm{Seg}(\mathbb{P}V_{K_1\cup\{k\}} \times \mathbb{P}V_{K_2} \times \cdots \times \mathbb{P}V_{K_m}). \quad (7)$$

*Proof.* Apply Lemma 6.4 to the left side of (7). It remains to check that

$$\bigcup_{L\subset\mathcal{V}(\tilde{B}_k)} \mathbb{P}(L \otimes V_k) = \bigcup_{\{K_1,\dots,K_m\}\in\mathcal{P}^2(\{1,\dots,n\}\setminus\{k\})} \mathrm{Seg}(\mathbb{P}V_{K_1\cup\{k\}} \times \mathbb{P}V_{K_2} \times \cdots \times \mathbb{P}V_{K_m}),$$

where $L \subset \mathcal{V}(\tilde{B}_k)$ are linear spaces. Because of symmetry and our hypothesis, there is only one type of linear space to consider, $V_{I_1} \otimes \widehat{v_{I_2}} \otimes \cdots \otimes \widehat{v_{I_m}} \otimes V_k = V_{I_1\cup\{k\}} \otimes \widehat{v_{I_2}} \otimes \cdots \otimes \widehat{v_{I_m}}$. It is clear that each of these linear spaces is on one of the Segre varieties on the right side of (7), and moreover every point on the right side of (7) is on one of these linear spaces. □

**Lemma 7.12.** *Let $n \geq 4$ and let $\mathcal{P}^2(\{n_1, \ldots, n_n\})$ denote the collection of all partitions of $\{n_1, \ldots, n_n\}$ into mutually disjoint subsets of cardinality 2 or less. Then*

$$\mathcal{V}(S_{(4,1)}S_{(4,1)}S_{(4,1)}S_{(5)}\ldots S_{(5)}) = \bigcup_{\{K_1,\ldots,K_m\}\in\mathcal{P}^2(\{1,\ldots,n\})} \operatorname{Seg}(\mathbb{P}V_{K_1}\times\cdots\times\mathbb{P}V_{K_m}).$$

*Proof by induction.* The base case is Lemma 7.9. For the induction step, use Lemma 7.11. We need to show that

$$\bigcap_{k=1}^{n}\Big(\bigcup_{\{K_1,\ldots,K_m\}\in\mathcal{P}^2(\{1,\ldots,n\}\setminus\{k\})} \operatorname{Seg}(\mathbb{P}V_{K_1\cup\{k\}}\times\mathbb{P}V_{K_2}\times\cdots\times\mathbb{P}V_{K_m})\Big)$$
$$= \bigcup_{\{K_1,\ldots,K_m\}\in\mathcal{P}^2(\{1,\ldots,n\})} \operatorname{Seg}(\mathbb{P}V_{K_1}\times\cdots\times\mathbb{P}V_{K_m}).$$

It suffices to check that

$$\operatorname{Seg}(\mathbb{P}V_{K_1\cup\{k\}}\times\mathbb{P}V_{K_2}\times\mathbb{P}V_{K_3}\times\cdots\times\mathbb{P}V_{K_m})$$
$$\cap\operatorname{Seg}(\mathbb{P}V_{K_1}\times\mathbb{P}V_{K_2\cup\{k\}}\times\mathbb{P}V_{K_3}\times\cdots\times\mathbb{P}V_{K_m})$$
$$=\operatorname{Seg}(\mathbb{P}V_{K_1}\times\mathbb{P}V_{K_2}\times\mathbb{P}V_k\times\mathbb{P}V_{K_3}\times\cdots\times\mathbb{P}V_{K_m}).$$

This is equivalent to checking that for any vector spaces $V_1$, $V_2$, $V_3$,

$$\operatorname{Seg}(\mathbb{P}(V_1\otimes V_2)\times\mathbb{P}V_3)\cap\operatorname{Seg}(\mathbb{P}V_1\times\mathbb{P}(V_2\otimes V_3)) = \operatorname{Seg}(\mathbb{P}V_1\times\mathbb{P}V_2\times\mathbb{P}V_3)).$$

In this case, let $[T]\in\operatorname{Seg}(\mathbb{P}(V_1\otimes V_2)\times\mathbb{P}V_3)\cap\operatorname{Seg}(\mathbb{P}V_1\times\mathbb{P}(V_2\otimes V_3))$. Then, viewed as a map $T:(V_1\otimes V_2)^*\to V_3$, the image of $T$ must be one-dimensional, and thus $[T]\in\mathbb{P}(V_1\otimes V_2\otimes V_3')$, where $V_3'\subset V_3$ is a one-dimensional subspace. By the same argument using the other Segre variety in the intersection, $[T]\in\mathbb{P}(V_1\otimes V_2'\otimes V_3)$, where $V_2'\subset V_2$ is a one-dimensional subspace. So $T\in\mathbb{P}(V_1\otimes V_2'\otimes V_3')$, but this is a linear space on $\operatorname{Seg}(\mathbb{P}V_1\times\mathbb{P}V_2\times\mathbb{P}V_3))$, so we are done. □

We have established all of the ingredients for the proof of Lemma 7.2.

## 8. Proof of Theorem 1.3

The outline of the proof is the following. Proposition 5.5 says that $Z_n\subseteq\mathcal{V}(HD)$. To show the opposite inclusion, we work by induction. In the cases of $n=3,4$, the (stronger) ideal-theoretic version of Theorem 1.3 was proved with the aid of a computer in [Holtz and Sturmfels 2007]. Since the theorem is already proved for the cases $n=3,4$ we will assume $n\geq 5$. The induction hypothesis is that $\mathcal{V}(HD_i)\simeq Z_{(n-1)}$. We need to show that given a point $[z]\in\mathcal{V}(HD)$, we have $[z]\in Z_n$, that is, that there exists a matrix $A$ such that $\varphi([A,t])=[z]$. The key tools we use in this proof are Proposition 7.1 and Lemma 6.6.

We work on a preferred open set $U_0 = \{[z] = [z_I X^I] \in \mathbb{P}(V_1 \otimes \cdots \otimes V_n) \mid z_{[0,\ldots,0]} \neq 0\}$. Working on this open set causes no loss of generality because of the following:

**Lemma 8.1.** *Let* $U_0 = \{[z] = [z_I X^I] \in \mathbb{P}(V_1 \otimes \cdots \otimes V_n) \mid z_{[0,\ldots,0]} \neq 0\}$. *Then* $\mathcal{V}(HD) \cap U_0 \subset Z_n$ *implies that* $\mathcal{V}(HD) \subset Z_n$.

*Proof.* The result follows from the fact that $Z_n$ and $\mathcal{V}(HD)$ are $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n)$-invariant, and $(\mathrm{SL}(2)^{\times n} \ltimes \mathfrak{S}_n).U_0 = \mathbb{P}(V_1 \otimes \cdots \otimes V_n)$. □

It suffices to work on the section $\{z = z_I Z^I \in V_1 \otimes \cdots \otimes V_n \mid z_{[0,\ldots,0]} = 1\}$ of the cone over projective space, because afterwards we can rescale everything to get the result on the whole open set $U_0$.

Suppose we take a point in the zero set (as described by Lemma 6.6)

$$[z] \in \mathcal{V}(HD) = \bigcap_{i=1}^{n} \bigcup_{L^i \subset \mathcal{V}(HD_i)} \mathbb{P}(L^i \otimes V_i).$$

Since $[z]$ is fixed, we can also fix a single $L^i$ for each $i$ so that $[z] \in \bigcap_{i=1}^n \mathbb{P}(L^i \otimes V_i)$. Work in our preferred section of the cone over projective space and write $n$ different expressions for the point $z$ (one for each $i$):

$$z = z_I X^I = \eta^i \otimes x_i^0 + \nu^i \otimes x_i^1,$$

where $[\eta^i], [\nu^i] \in L^i \subset \mathcal{V}(HD_i)$. (These expressions are possible because each $V_i$ is 2-dimensional.) Choosing $z_{[0,\ldots,0]} = 1$ also implies that $\eta^i_{[0,\ldots,0]} = 1$. The induction hypothesis says that $Z_{(n-1)} \simeq \mathcal{V}(HD_i)$ for $1 \leq i \leq n$. So each $\eta^i$ satisfies $\varphi([A^{(i)}, 1]) = \eta^i$ for a symmetric matrix $A^{(i)} \in S^2\mathbb{C}^{n-1}$. For each $0 \leq j \leq n$ denote by $\mathscr{A}^j$ the subset of matrices

$$\mathscr{A}^j = \{A \in S^2\mathbb{C}^n \mid \Delta_I(A) = z_I \text{ for all } I = [i_1, \ldots, i_n] \text{ with } i_j = 0\}.$$

Each matrix $A \in \mathscr{A}^j$ has the property that the principal submatrix of $A$ formed by deleting the $j$th row and column maps to $\eta^j$ under the principal minor map. Thus each $A \in \mathscr{A}^j$ is a candidate matrix that might satisfy $\varphi([A, 1]) = [z]$, but we don't know if such a matrix will have a submatrix that maps to the other $\eta^i$'s. We claim that there is at least one matrix that satisfies all of these conditions.

**Lemma 8.2.** $\bigcap_{i=1}^{n} \mathscr{A}^i$ *is nonempty.*

*Proof.* By the induction hypothesis, each $\mathscr{A}^i$ is nonempty. Assume $\bigcap_{i=2}^n \mathscr{A}^i$ is nonempty. We show that if $A \in \bigcap_{i=2}^n \mathscr{A}^i$, then $A \in \mathscr{A}^1$. The same argument also proves that if $A \in \bigcap_{i=3}^n \mathscr{A}^i$, then $A \in \mathscr{A}^1$, and so on, so it suffices to check the last, most restrictive case. Also because of the $\mathfrak{S}_n$ action, we don't have to repeat the proof for every permutation.

If $A \in \bigcap_{i=2}^{n} \mathscr{A}^i$, then $\Delta_I A = z_I$ for all $I \neq [0, i_2, \ldots, i_n]$ with $|I| \leq n-2$. The only possible exception we could have is that $\Delta_{[0,1,\ldots,1]}$ might not be equal to $z_{[0,1,\ldots,1]}$. Let $A'$ denote the principal submatrix of $A$ formed by deleting the first row and column of $A$. Now since $n \geq 5$, $|I| \geq 3$, $A'$ is at least as large as $4 \times 4$, and we have determined that all of the principal minors of $A'$ except possibly the determinant agree with a fixed point $\eta^1 \in \mathscr{V}(HD_1)$ (in other words $\Delta(A')_I = \eta^i_I$ for all $I \neq [1, \ldots, 1]$), so we can apply Proposition 7.1 to conclude that the determinant of $A'$ also agrees with $\eta^1$ (that is, $\Delta_{[1,\ldots,1]}(A') = \eta^1_{[1,\ldots,1]}$). Therefore any such $A$ must have $\Delta_{[0,1,\ldots,1]}(A) = z_{[0,1,\ldots,1]}$, and we have shown $A \in \mathscr{A}^1$. □

Lemma 8.2 proves the existence of a symmetric matrix $A$ such that $\Delta_I(A) = z_I$ for all $I \neq [1, \ldots, 1]$. Then since both $z_I X^I$ and $\Delta_I(A) X^I$ are points in $\mathscr{V}(HD)$, Proposition 7.1 implies that $\Delta_{[1,\ldots,1]}(A) = z_{[1,\ldots,1]}$, and this finishes the proof of the main theorem.

**Remark 8.3** (building a matrix). Note that when $n \geq 4$, the proof we gave can be used also to construct a symmetric matrix whose principal minors are prescribed by a point $z \in \widehat{\mathscr{V}(HD)} \cap \{z \mid z_{[0,\ldots,0]} \neq 0\}$. The entries of $z$ corresponding to $1 \times 1$ and $2 \times 2$ principal minors determine a large finite set $\mathscr{A}$ of candidate matrices that could map to $z$. Restrict the set $\mathscr{A}$ to only those matrices whose $3 \times 3$ principal minors agree with the corresponding entries of $z$, that is, keep only the matrices $A$ such that $\Delta(A)_I = z_I$ for all $|I| \leq 3$. We claim that the remaining matrices all map to $z$ under the principal minor map. If $A$ is such that all of the $3 \times 3$ principal minors agree with $z$, then Proposition 7.1 implies that each $4 \times 4$ principal minor of $A$ must agree with $z$ also. Iterate this argument to imply that all of the principal minors of $A$ must agree with $z$.

## Acknowledgments

The author thanks J. M. Landsberg for suggesting this problem as a thesis topic and for his endless support and advice along the way. We thank the two anonymous reviewers who read the first draft of this paper, as well as the third reviewer who read the second draft, for their numerous useful suggestions for revision. We also thank Shaowei Lin, Linh Nguyen, Giorgio Ottaviani, Bernd Sturmfels, and Zach Teitler for useful conversations. Shaowei Lin pointed out the reference [Nanson 1897]. Bernd Sturmfels suggested the addition of Corollary 1.4.

## References

[Allman and Rhodes 2008] E. S. Allman and J. A. Rhodes, "Phylogenetic ideals and varieties for the general Markov model", *Adv. in Appl. Math.* **40**:2 (2008), 127–148. MR 2008m:60145 Zbl 1131.92046

[Borcea et al. 2009] J. Borcea, P. Brändén, and T. M. Liggett, "Negative dependence and the geometry of polynomials", *J. Amer. Math. Soc.* **22**:2 (2009), 521–567. MR 2010b:62215

[Borodin and Rains 2005] A. Borodin and E. M. Rains, "Eynard–Mehta theorem, Schur process, and their Pfaffian analogs", *J. Stat. Phys.* **121**:3-4 (2005), 291–317. MR 2006k:82039 Zbl 1127.82017

[Bürgisser et al. 1997] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic complexity theory*, Grundlehren der Mathematischen Wissenschaften **315**, Springer, Berlin, 1997. MR 99c:68002 Zbl 1087.68568

[Comon and Rajih 2006] P. Comon and M. Rajih, "Blind identification of under-determined mixtures based on the characteristic function", *Signal processing* **86**:9 (2006), 2271–2281. Zbl 1172.65304

[Comon et al. 2008] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank", *SIAM J. Matrix Anal. Appl.* **30**:3 (2008), 1254–1279. MR 2009i:15039 Zbl 1181.15014

[Cox et al. 2007] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms*, 3rd ed., Springer, New York, 2007. MR 2007h:13036 Zbl 1118.13001

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics **129**, Springer, New York, 1991. MR 93a:20069 Zbl 0744.22001

[Gelfand et al. 1994] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Birkhäuser, Boston, 1994. MR 95e:14045 Zbl 0827.14036

[Goodman and Wallach 1998] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*, Encyclopedia of Mathematics and its Applications **68**, Cambridge University Press, 1998. MR 99b:20073 Zbl 0901.22001

[Griffin and Tsatsomeros 2006] K. Griffin and M. J. Tsatsomeros, "Principal minors, II: The principal minor assignment problem", *Linear Algebra Appl.* **419**:1 (2006), 125–171. MR 2008h:15015 Zbl 1110.65035

[Harris 1992] J. Harris, *Algebraic geometry*, Graduate Texts in Mathematics **133**, Springer, New York, 1992. MR 93j:14001 Zbl 0779.14001

[Holtz 1999] O. Holtz, "Not all GKK $\tau$-matrices are stable", *Linear Algebra Appl.* **291**:1-3 (1999), 235–244. MR 2000a:15011 Zbl 0968.15014

[Holtz and Schneider 2002] O. Holtz and H. Schneider, "Open problems on GKK $\tau$-matrices", *Linear Algebra Appl.* **345** (2002), 263–267. MR 1883278 Zbl 1160.15305

[Holtz and Sturmfels 2007] O. Holtz and B. Sturmfels, "Hyperdeterminantal relations among symmetric principal minors", *J. Algebra* **316**:2 (2007), 634–648. MR 2009c:15032 Zbl 1130.15005

[Landsberg 2008] J. M. Landsberg, "Geometry and the complexity of matrix multiplication", *Bull. Amer. Math. Soc.* (*N.S.*) **45**:2 (2008), 247–284. MR 2009b:68055 Zbl 1145.68054

[Landsberg $\geq$ 2011] J. M. Landsberg, "The geometry of tensors: applications to complexity, statistics and other areas", in preparation.

[Landsberg and Manivel 2003] J. M. Landsberg and L. Manivel, "On the projective geometry of rational homogeneous varieties", *Comment. Math. Helv.* **78**:1 (2003), 65–100. MR 2004a:14050 Zbl 1048.14032

[Landsberg and Manivel 2004] J. M. Landsberg and L. Manivel, "On the ideals of secant varieties of Segre varieties", *Found. Comput. Math.* **4**:4 (2004), 397–422. MR 2005m:14101 Zbl 1068.14068

[Landsberg and Weyman 2007] J. M. Landsberg and J. Weyman, "On tangential varieties of rational homogeneous varieties", *J. Lond. Math. Soc.* (2) **76**:2 (2007), 513–530. MR 2008m:14101 Zbl 1127.14045

[de Lathauwer and de Baynast 2008] L. de Lathauwer and A. de Baynast, "Blind deconvolution of DS-CDMA signals by means of decomposition in rank-$(1, L, L)$ terms", *IEEE Trans. Signal Process.* **56**:4 (2008), 1562–1571. MR 2516571

[Lin and Sturmfels 2009] S. Lin and B. Sturmfels, "Polynomial relations among principal minors of a $4 \times 4$-matrix", *J. Algebra* **322**:11 (2009), 4121–4131. MR 2010m:15014 Zbl 1198.14049

[Mikkonen 2007] T. Mikkonen, "The ring of graph invariants — graphic values", preprint, 2007. arXiv 0712.0146

[Nanson 1897] E. J. Nanson, "On the relations between the coaxial minors of a determinant", *Philosophical Magazine* **44**:5 (1897), 362–367. JFM 28.0140.01

[Oeding 2009] L. A. Oeding, *G-varieties and the principal minors of symmetric matrices*, ProQuest LLC, Ann Arbor, MI, 2009. MR 2713672

[Oeding 2011] L. Oeding, "Set-theoretic defining equations of the tangential variety of the Segre variety", *J. Pure Appl. Algebra* **215**:6 (2011), 1516–1527. MR 2769247

[Pachter and Sturmfels 2005] L. Pachter and B. Sturmfels, *Algebraic statistics for computational biology*, Cambridge University Press, New York, 2005. MR 2006i:92002 Zbl 1108.62118

[Sidman and Sullivant 2009] J. Sidman and S. Sullivant, "Prolongations and computational algebra", *Canad. J. Math.* **61**:4 (2009), 930–949. MR 2010h:14084 Zbl 1174.13042

[Stouffer 1924] E. B. Stouffer, "On the independence of principal minors of determinants", *Trans. Amer. Math. Soc.* **26**:3 (1924), 356–368. MR 1501282 JFM 50.0065.01

[Stouffer 1928] E. B. Stouffer, "Expressions for the general determinant in terms of its principal minors", *Amer. Math. Monthly* **35**:1 (1928), 18–21. MR 1521341 JFM 54.0108.06

[Wagner 2008] D. G. Wagner, "Negatively correlated random variables and Mason's conjecture for independent sets in matroids", *Ann. Comb.* **12**:2 (2008), 211-239. MR 2009f:05053 Zbl 1145.05003

[Weyl 1997] H. Weyl, *The classical groups*, Princeton University Press, 1997. MR 98k:01049 Zbl 1024.20501

Communicated by David Eisenbud
Received 2010-01-25 Revised 2010-11-01 Accepted 2010-12-05

oeding@math.berkeley.edu *Dipartimento di Matematica "U. Dini", Università degli Studi di Firenze, Viale Morgagni 67/A, 50134 Firenze, Italy*

*Current address:* *Department of Mathematics, University of California, Berkeley, 970 Evans Hall #3840, Berkeley, CA 94720-3840, United States*

ALGEBRA AND NUMBER THEORY 5:1(2011)

msp

# Frobenius difference equations and algebraic independence of zeta values in positive equal characteristic

Chieh-Yu Chang, Matthew A. Papanikolas and Jing Yu

By analogy with the Riemann zeta function at positive integers, for each finite field $\mathbb{F}_{p^r}$ with fixed characteristic $p$, we consider Carlitz zeta values $\zeta_r(n)$ at positive integers $n$. Our theorem asserts that among the zeta values in the set $\bigcup_{r=1}^{\infty}\{\zeta_r(1), \zeta_r(2), \zeta_r(3), \dots\}$, all the algebraic relations are those relations within each individual family $\{\zeta_r(1), \zeta_r(2), \zeta_r(3), \dots\}$. These are the algebraic relations coming from the Euler–Carlitz and Frobenius relations. To prove this, a motivic method for extracting algebraic independence results from systems of Frobenius difference equations is developed.

## 1. Introduction

**1.1.** ***Motivic transcendence theory.*** Classically, Grothendieck's period conjecture for abelian varieties predicts that the dimension of the Mumford–Tate group of an abelian variety over $\overline{\mathbb{Q}}$ should be equal to the transcendence degree of the field generated by its period matrix over $\overline{\mathbb{Q}}$. Conjecturally, the Mumford–Tate group is the motivic Galois group from Tannakian duality, and therefore Grothendieck's conjecture provides an interpretation of the algebraic relations among periods in question by way of motivic Galois groups.

We are concerned with the algebraic independence of special zeta values over function fields with varying finite constant fields in positive equal characteristic. In the positive characteristic world, there is the concept of $t$-motives introduced by Anderson [1986], dual to the concept of $t$-modules. The third author showed that the structure of $t$-modules is key for proving many interesting linear independence results about special values in this setting [Yu 1997]. The breakthrough in passing from linear independence to algebraic independence by way of $t$-motives began

The first author was supported by NSC and NCTS. The second author was supported by NSF Grants DMS-0600826 and DMS-0903838. The third author was supported by NSC Grant No. 99-2119-M-002-024.

*MSC2000:* primary 11J93; secondary 11M38, 11G09.

*Keywords:* Algebraic independence, Frobenius difference equations, $t$-motives, zeta values.

with Anderson, Brownawell and the second author, and in particular with the linear independence criterion of [Anderson et al. 2004] (the *ABP criterion*).

By introducing a Tannakian formalism for rigid analytically trivial pre-$t$-motives and relating it to the Galois theory of Frobenius difference equations, the second author [Papanikolas 2008] has shown that the Galois group of a rigid analytically trivial pre-$t$-motive is isomorphic to its difference Galois group. Furthermore, the second author has successfully used the ABP criterion to show that the transcendence degree of the field generated by the period matrix of an ABP motive (that is, the pre-$t$-motive comes from a uniformizable abelian $t$-module) is equal to the dimension of its Galois group. More generally, we say that a rigid analytically trivial pre-$t$-motive has the *GP (Grothendieck period) property* if the transcendence degree of the field generated by its period matrix is equal to the dimension of its Galois group (for more details of terminology, see Section 2).

Using a refined version of the ABP criterion proved by the first author [Chang 2009], we observe that there are many pre-$t$-motives that are not ABP motives but that have the GP property. This motivates us to introduce a method of uniformizing the Frobenius twisting operators with respect to different constant fields for those pre-$t$-motives that have the GP property. The pre-$t$-motive we obtain in this way is defined over a larger constant field, but still has the GP property (see Corollary 2.2.4). This technique is very useful when dealing with the problem of determining all the algebraic relations among various special values of arithmetic interest in a fixed positive characteristic. It is used in this paper to study special zeta values. For another application to special arithmetic gamma values, see [Chang et al. 2010].

**1.2.** ***Carlitz zeta values.*** Let $p$ be a prime, and let $\mathbb{F}_{p^r}[\theta]$ be the polynomial ring in $\theta$ over the finite field $\mathbb{F}_{p^r}$ of $p^r$ elements. Our aim is to determine all the algebraic relations among the zeta values

$$\zeta_r(n) := \sum_{\substack{a \in \mathbb{F}_{p^r}[\theta] \\ a \text{ monic}}} \frac{1}{a^n} \in \mathbb{F}_{p^r}((1/\theta)) \subseteq \overline{\mathbb{F}_p}((1/\theta)),$$

where $r$ and $n$ vary over all positive integers. Each $\zeta_r(n)$ lies in $\mathbb{F}_p((1/\theta))$, since it is fixed by the automorphism $\left(\sum a_i(1/\theta)^i \mapsto \sum a_i^p(1/\theta)^i\right) : \overline{\mathbb{F}_p}((1/\theta)) \to \overline{\mathbb{F}_p}((1/\theta))$.

The study of these zeta values was begun by Carlitz [1935]. For a fixed positive integer $r$, he discovered that there is a constant $\tilde{\pi}_r$, algebraic over $\mathbb{F}_{p^r}((1/\theta))$, such that $\zeta_r(n)/\tilde{\pi}_r^n$ lies in $\mathbb{F}_p(\theta)$ if $n$ is divisible by $p^r - 1$. The quantity $\tilde{\pi}_r$ arises as a fundamental period of the Carlitz $\mathbb{F}_{p^r}[t]$-module $C_r$, and Wade [1941] showed that $\tilde{\pi}_r$ is transcendental over $\mathbb{F}_p(\theta)$.

We say that a positive integer $n$ is $(p, r)$-*even* if it is a multiple of $p^r - 1$. Thus the situation of Carlitz zeta values at $(p, r)$-even positive integers is completely

analogous to that of the Riemann zeta function at even positive integers. For these $(p, r)$-even $n$, we call the $\mathbb{F}_p(\theta)$-linear relations between $\zeta_r(n)$ and $\tilde{\pi}_r^n$ the *Euler–Carlitz* relations. Because the characteristic is positive, there are also *Frobenius* $p$-th power relations among these zeta values: for positive integers $m, n$,

$$\zeta_r(p^m n) = \zeta_r(n)^{p^m}.$$

Anderson and Thakur [1990] and Yu [1991; 1997] made several breakthroughs in understanding Carlitz zeta values. Using the $t$-module method, the transcendence of $\zeta_r(n)$ for all positive integers $n$, and in particular for *odd* $n$ (that is, $n$ not divisible by $p^r - 1$), was proved, and it was also proved that the Euler–Carlitz relations are the only $\overline{\mathbb{F}_p(\theta)}$-linear relations among $\{\zeta_r(n), \tilde{\pi}_r^m;\ m, n \in \mathbb{N}\}$. In [Chang and Yu 2007], the first and third authors used ABP motives instead of $t$-modules to show that for fixed $r$, the Euler–Carlitz relations and the Frobenius $p$-th power relations account for all the algebraic relations over $\overline{\mathbb{F}_p(\theta)}$ among the Carlitz zeta values

$$\tilde{\pi}_r, \zeta_r(1), \zeta_r(2), \zeta_r(3), \dots.$$

To complete the story of Carlitz zeta values, the next natural question is what happens if $r$ varies. Denis [1998] proved the algebraic independence of all fundamental periods $\{\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3, \dots\}$ as the constant field varies. Thus, in view of [Chang and Yu 2007], one expects that for the bigger set of zeta values,

$$\bigcup_{r=1}^{\infty} \{\zeta_r(1), \zeta_r(2), \zeta_r(3), \dots\},$$

the Euler–Carlitz relations and the Frobenius $p$-th power relations still account for all the algebraic relations. This is indeed the case, as we find from the following theorem (stated subsequently as Corollary 4.5.2).

**Theorem 1.2.1.** *Given any positive integers $s$ and $d$, the transcendence degree of the field*

$$\overline{\mathbb{F}_p(\theta)}\Big(\bigcup_{r=1}^{d} \{\tilde{\pi}_r, \zeta_r(1), \dots, \zeta_r(s)\}\Big)$$

*over $\overline{\mathbb{F}_p(\theta)}$ is*

$$\sum_{r=1}^{d}\left(s - \left\lfloor \frac{s}{p} \right\rfloor - \left\lfloor \frac{s}{p^r - 1} \right\rfloor + \left\lfloor \frac{s}{p(p^r - 1)} \right\rfloor + 1\right).$$

**1.3.** ***Outline.*** Our strategy is to construct a pre-$t$-motive that has the GP property and whose period matrix accounts for the Carlitz zeta values in question. In [Chang and Yu 2007], an ABP motive has already been constructed for Carlitz zeta values with respect to a fixed constant field. The problem here is one concerning varying the constant fields in a fixed characteristic, and one has to uniformize Frobenius powers in order to apply the method developed in [Papanikolas 2008].

This paper is organized as follows. In Section 2, we review Papanikolas' theory and investigate the pre-$t$-motives that have the GP property. Here we introduce the mechanism of uniformizing Frobenius twisting operators while taking direct sums. Section 3 includes discussions about rigid analytically trivial pre-$t$-motives of type **SV**, that is, whose Galois groups are extensions of split tori by vector groups. The heart of this section is Theorem 3.2.2, where we determine the dimensions of Galois groups of direct sums of pre-$t$-motives of type **SV**. The pre-$t$-motive for Theorem 1.2.1 is constructed in Section 4, and we prove that it satisfies the conditions of Theorem 3.2.2. Finally, we calculate its dimension explicitly in Theorem 4.5.1, which then has Theorem 1.2.1 as direct consequence.

## 2. $t$-motivic Galois groups

### 2.1. *Notation.*

**2.1.1.** *Table of symbols.*
$\mathbb{F}_p :=$ the finite field of $p$ elements, $p$ a prime number.
$k := \mathbb{F}_p(\theta) :=$ the rational function field in the variable $\theta$ over $\mathbb{F}_p$.
$k_\infty := \mathbb{F}_p((1/\theta))$, completion of $k$ with respect to the infinite place.
$\overline{k_\infty} :=$ a fixed algebraic closure of $k_\infty$.
$\overline{k} :=$ the algebraic closure of $k$ in $\overline{k_\infty}$.
$\mathbb{C}_\infty :=$ completion of $\overline{k_\infty}$ with respect to the canonical extension of the infinite place.
$|\cdot|_\infty :=$ a fixed absolute value for the completed field $\mathbb{C}_\infty$ with $|\theta|_\infty = p$.
$\mathbb{C}_\infty[\![t]\!] :=$ the ring of power series in the variable $t$ over $\mathbb{C}_\infty$.
$\mathbb{C}_\infty((t)) :=$ the field of Laurent series in the variable $t$ over $\mathbb{C}_\infty$.
$\mathbb{T} := \{f \in \mathbb{C}_\infty[\![t]\!] \mid f \text{ converges on } |t|_\infty \leq 1\}$. This is known as the Tate algebra.
$\mathbb{L} :=$ the fraction field of $\mathbb{T}$.
$\sigma := \sigma_p := \left(\sum_i a_i t^i \mapsto \sum_i a_i^{1/p} t^i\right) : \mathbb{C}_\infty((t)) \to \mathbb{C}_\infty((t))$.
$\mathbb{G}_a :=$ the additive group.
$\mathrm{GL}_r / F :=$ for a field $F$, the $F$-group scheme of invertible $r \times r$ square matrices.
$\mathbb{G}_m := \mathrm{GL}_1$, the multiplicative group.

**2.1.2.** *Block diagonal matrices.* Let $A_i \in \mathrm{Mat}_{m_i}(\mathbb{L})$ for $i = 1, \ldots, n$, and $m := m_1 + \cdots + m_n$. We define $\bigoplus_{i=1}^n A_i \in \mathrm{Mat}_m(\mathbb{L})$ to be the canonical block diagonal matrix, that is, the matrix with $A_1, \ldots, A_n$ down the diagonal and zeros elsewhere.

**2.1.3.** *$n$-fold twisting.* For $n \in \mathbb{Z}$ and a formal Laurent series $f = \sum_i a_i t^i \in \mathbb{C}_\infty((t))$, we define the $n$-fold twisting $f^{(n)} := \sigma^{-n}(f) := \sum_i a_i^{p^n} t^i$. The $n$-fold twisting operation is an automorphism of the Laurent series field $\mathbb{C}_\infty((t))$ that stabilizes several subrings, for example, $\overline{k}[\![t]\!]$, $\overline{k}[t]$, and $\mathbb{T}$. More generally, for any matrix $B$ with entries in $\mathbb{C}_\infty((t))$, we define $B^{(n)}$ by the rule ${B^{(n)}}_{ij} = {B_{ij}}^{(n)}$.

**2.1.4.** *Entire power series.* A power series $f = \sum_{i=0}^{\infty} a_i t^i \in \mathbb{C}_\infty[\![t]\!]$ that satisfies

$$\lim_{i\to\infty} \sqrt[i]{|a_i|_\infty} = 0$$

and

$$[k_\infty(a_0, a_1, a_2, \dots) : k_\infty] < \infty$$

is called an entire power series. As a function of $t$, such a power series $f$ converges on all $\mathbb{C}_\infty$ and, when restricted to $\overline{k_\infty}$, $f$ takes values in $\overline{k_\infty}$. The ring of the entire power series is denoted by $\mathbb{E}$.

**2.2.** ***Pre-t-motives and the GP property.*** For $r$ a positive integer, let $\bar{k}(t)[\sigma^r, \sigma^{-r}]$ be the noncommutative ring of Laurent polynomials in $\sigma^r$ with coefficients in $\bar{k}(t)$, subject to the relation

$$\sigma^r f := f^{(-r)}\sigma^r \text{ for all } f \in \bar{k}(t).$$

A pre-$t$-motive $M$ of level $r$ is a left $\bar{k}(t)[\sigma^r, \sigma^{-r}]$-module that is finite-dimensional over $\bar{k}(t)$. Letting $\mathbf{m} \in \mathrm{Mat}_{n\times 1}(M)$ comprise a $\bar{k}(t)$-basis of $M$, multiplication by $\sigma^r$ on $M$ is represented by

$$\sigma^r(\mathbf{m}) = \Phi\mathbf{m}$$

for some matrix $\Phi \in \mathrm{GL}_n(\bar{k}(t))$. Furthermore, $M$ is called rigid analytically trivial if there exists $\Psi \in \mathrm{GL}_n(\mathbb{L})$ such that

$$\sigma^r(\Psi) := \Psi^{(-r)} = \Phi\Psi.$$

Such a matrix $\Psi$ is called a rigid analytic trivialization of the matrix $\Phi$. We also say that $\Psi$ is a rigid analytic trivialization of $M$ (with respect to $\mathbf{m}$). Note that if $\Psi' \in \mathrm{GL}_n(\mathbb{L})$ is also a rigid analytic trivialization of $\Phi$, then by [Papanikolas 2008, §4.1.6] we have

$$\Psi'^{-1}\Psi \in \mathrm{GL}_n(\mathbb{F}_{p^r}(t)). \tag{1}$$

Moreover, if we put $\mathbf{m}' := B\mathbf{m}$ for any fixed $B \in \mathrm{GL}_n(\bar{k}(t))$, then $\Phi' := B^{(-1)}\Phi B^{-1}$ represents multiplication by $\sigma^r$ on $M$ with respect to the $\bar{k}(t)$-basis $\mathbf{m}'$ of $M$, and $\Psi' := B\Psi$ is a rigid analytic trivialization of $\Phi'$.

**Definition 2.2.1.** Suppose we are given a rigid analytically trivial pre-$t$-motive $M$ of level $r$ that is of dimension $n$ over $\bar{k}(t)$. If there exists a $\bar{k}(t)$-basis $\mathbf{m} \in \mathrm{Mat}_{n\times 1}(M)$ such that there exists $\Psi \in \mathrm{GL}_n(\mathbb{L}) \cap \mathrm{Mat}_n(\mathbb{E})$ that is a rigid analytic trivialization of $M$ with respect to $\mathbf{m}$ and satisfies

$$\mathrm{tr.deg}_{\bar{k}(t)} \bar{k}(t)(\Psi) = \mathrm{tr.deg}_{\bar{k}} \bar{k}(\Psi(\theta)),$$

then we say that $M$ has the GP property, where $\bar{k}(t)(\Psi)$ (resp. $\bar{k}(\Psi(\theta))$) is the field generated by all entries of $\Psi$ (resp. $\Psi(\theta)$) over $\bar{k}(t)$ (resp. $\bar{k}$). The GP property is independent of the choices of $\Psi$ for a fixed $\mathbf{m}$ because of (1).

Given a rigid analytically trivial pre-$t$-motive $M$ of level $r$ with $(\mathbf{m}, \Phi, \Psi)$ as above, for any $s \in \mathbb{N}$ we define its $s$-th derived pre-$t$-motive $M^{(s)}$, which is a pre-$t$-motive of level $rs$: the underlying space of $M^{(s)}$ is the same as $M$, but it is now regarded as a left $\bar{k}(t)[\sigma^{rs}, \sigma^{-rs}]$-module. Letting

$$\Phi' := \Phi^{(-(s-1)r)} \dots \Phi^{(-r)} \Phi,$$

we have $\sigma^{rs}\mathbf{m} = \Phi'\mathbf{m}$ and $\sigma^{rs}\Psi := \Psi^{(-rs)} = \Phi'\Psi$, and hence $\Psi$ is also a rigid analytic trivialization of $M^{(s)}$.

**Proposition 2.2.2.** *Let $M$ be a rigid analytically trivial pre-$t$-motive of level $r$ that has the GP property. For any positive integer $s$, the $s$-th derived pre-$t$-motive $M^{(s)}$ of $M$ is also rigid analytically trivial and has the GP property.*

**Theorem 2.2.3** [Chang 2009, Theorem 1.2; Papanikolas 2008, Theorem 5.2.2]. *Suppose $\Phi \in \mathrm{Mat}_n(\bar{k}[t])$ defines a rigid analytically trivial pre-$t$-motive $M$ of level $r$ with a rigid analytic trivialization $\Psi \in \mathrm{Mat}_n(\mathbb{T}) \cap \mathrm{GL}_n(\mathbb{L})$. If $\det \Phi(0) \neq 0$ and $\det \Phi(\theta^{1/p^{ri}}) \neq 0$ for all $i = 1, 2, 3, \dots$, then $M$ has the GP property.*

By [Anderson et al. 2004, Proposition 3.1.3], the condition $\det \Phi(0) \neq 0$ implies $\Psi \in \mathrm{Mat}_n(\mathbb{E})$. Combining Theorem 2.2.3 and Proposition 2.2.2, we have:

**Corollary 2.2.4.** *Given an integer $d \geq 2$, we let $\ell := \mathrm{lcm}(1, \dots, d)$. For each $1 \leq r \leq d$, let $\ell_r := \ell/r$ and let $\Phi_r \in \mathrm{Mat}_{n_r}(\bar{k}[t]) \cap \mathrm{GL}_{n_r}(\bar{k}(t))$ define a pre-$t$-motive $M_r$ of level $r$ with a rigid analytic trivialization $\Psi_r \in \mathrm{Mat}_{n_r}(\mathbb{T}) \cap \mathrm{GL}_{n_r}(\mathbb{L})$. Suppose that each $\Phi_r$ satisfies the hypotheses of Theorem 2.2.3 for $r = 1, \dots, d$. Then the direct sum*

$$M := \bigoplus_{r=1}^{d} {M_r}^{(\ell_r)}$$

*is a rigid analytically trivial pre-$t$-motive of level $\ell$ that has the GP property.*

*Proof.* For each $1 \leq r \leq d$, we define

$$\Phi_r' := \Phi_r^{(-(\ell_r-1)r)} \dots \Phi_r^{(-r)} \Phi_r.$$

Moreover, if we define

$$\Phi' := \oplus_{r=1}^{d} \Phi_r', \quad \Psi' := \oplus_{r=1}^{d} \Psi_r,$$

then we have

$$\Psi'^{(-\ell)} = \Phi'\Psi'.$$

Note that the matrix representing multiplication by $\sigma^{\ell}$ on $M$ with respect to the evident $\bar{k}(t)$-basis is given by $\Phi'$.

Our task is to show that $\Phi'$ satisfies the hypotheses of Theorem 2.2.3 (with respect to the operator $\sigma^{\ell}$), whence the result. It is obvious that $\det \Phi'(0) \neq 0$

since $\det \Phi_r(0) \neq 0$ for each $1 \leq r \leq d$. Suppose that $\det \Phi'(\theta^{1/p^{\ell j}}) = 0$ for some $j \in \mathbb{N}$. This implies that there exists $1 \leq r \leq d$ and $0 \leq m \leq \ell_r - 1$ such that

$$\det \Phi_r^{(-rm)}(\theta^{(-\ell j)}) = 0.$$

However, this is equivalent to

$$\det \Phi_r(\theta^{(-(\ell j - rm))}) = 0. \tag{2}$$

Since $0 \leq m \leq \ell_r - 1$ and $r|\ell$, we have that $(\ell j - rm) > 0$ and $r|(\ell j - rm)$. Thus, (2) contradicts the hypothesis that $\det \Phi_r(\theta^{(-rh)}) \neq 0$ for all $h = 1, 2, 3, \ldots$. □

**2.3.** ***Difference Galois groups and transcendence.*** In this section, we review the related theory developed in [Papanikolas 2008]. Let $r$ be a fixed positive integer. The category of pre-$t$-motives of level $r$ forms a rigid abelian $\mathbb{F}_{p^r}(t)$-linear tensor category. Also, the category $\mathcal{R}$ of rigid analytically trivial pre-$t$-motives of level $r$ forms a neutral Tannakian category over $\mathbb{F}_{p^r}(t)$. Given an object $M$ in $\mathcal{R}$, we let $\mathcal{R}_M$ be the strictly full Tannakian subcategory of $\mathcal{R}$ generated by $M$. That is, $\mathcal{R}_M$ consists of all objects of $\mathcal{R}$ isomorphic to subquotients of finite direct sums of

$$M^{\otimes u} \otimes (M^\vee)^{\otimes v} \text{ for various } u, v,$$

where $M^\vee$ is the dual of $M$. By Tannakian duality, $\mathcal{R}_M$ is representable by an affine algebraic group scheme $\Gamma_M$ over $\mathbb{F}_{p^r}(t)$. The group $\Gamma_M$ is called the Galois group of $M$ and it is described explicitly as follows.

Suppose that $\Phi \in \mathrm{GL}_n(\bar{k}(t))$ provides multiplication by $\sigma^r$ on $M$ with respect to a fixed basis $\mathbf{m} \in \mathrm{Mat}_{n\times 1}(M)$ over $\bar{k}(t)$. Let $\Psi \in \mathrm{GL}_n(\mathbb{L})$ be a rigid analytic trivialization for $\Phi$. Let $X := (X_{ij})$ be an $n \times n$ matrix whose entries are independent variables $X_{ij}$, and define a $\bar{k}(t)$-algebra homomorphism $\nu : \bar{k}(t)[X, 1/\det X] \to \mathbb{L}$ such that $\nu(X_{ij}) = \Psi_{ij}$ for all $1 \leq i, j \leq n$. We let

$$\begin{aligned} \Sigma_\Psi &:= \operatorname{im} \nu = \bar{k}(t)[\Psi, 1/\det \Psi] \subseteq \mathbb{L}, \\ Z_\Psi &:= \operatorname{Spec} \Sigma_\Psi. \end{aligned}$$

Then $Z_\Psi$ is a closed $\bar{k}(t)$-subscheme of $\mathrm{GL}_n/\bar{k}(t)$. Let $\Psi_1, \Psi_2 \in \mathrm{GL}_n(\mathbb{L} \otimes_{\bar{k}(t)} \mathbb{L})$ be the matrices satisfying $(\Psi_1)_{ij} = \Psi_{ij} \otimes 1$ and $(\Psi_2)_{ij} = 1 \otimes \Psi_{ij}$ for all $1 \leq i, j \leq n$. Let $\widetilde{\Psi} := \Psi_1^{-1}\Psi_2$. We have an $\mathbb{F}_{p^r}(t)$-algebra homomorphism $\mu : \mathbb{F}_{p^r}(t)[X, 1/\det X] \to \mathbb{L} \otimes_{\bar{k}(t)} \mathbb{L}$ such that $\mu(X_{ij}) = \widetilde{\Psi}_{ij}$ for all $1 \leq i, j \leq n$. Furthermore, we define

$$\begin{aligned} \Delta &:= \operatorname{im} \mu, \\ \Gamma_\Psi &:= \operatorname{Spec} \Delta. \end{aligned} \tag{3}$$

**Theorem 2.3.1** [Papanikolas 2008, Theorems 4.2.11, 4.3.1, 4.5.10]. *The scheme $\Gamma_\Psi$ is a closed $\mathbb{F}_{p^r}(t)$-subgroup scheme of $\mathrm{GL}_n/\mathbb{F}_{p^r}(t)$, which is isomorphic to the Galois group $\Gamma_M$ over $\mathbb{F}_{p^r}(t)$. Moreover, $\Gamma_\Psi$ has the following properties*:

(a) *$\Gamma_\Psi$ is smooth over $\overline{\mathbb{F}_p(t)}$ and is geometrically connected.*

(b) $\dim \Gamma_\Psi = \mathrm{tr.deg}_{\bar{k}(t)} \bar{k}(t)(\Psi)$.

(c) *$Z_\Psi$ is a $\Gamma_\Psi$-torsor over $\bar{k}(t)$.*

*In particular, if $M$ has the GP property, then*

(d) $\dim \Gamma_\Psi = \mathrm{tr.deg}_{\bar{k}} \bar{k}(\Psi(\theta))$.

We call $\Gamma_\Psi$ the Galois group associated to the difference equation $\Psi^{(-r)} = \Phi\Psi$. This $\Gamma_\Psi$ is independent of the analytic trivialization $\Psi$, up to isomorphism over $\mathbb{F}_{p^r}(t)$. Throughout this paper we always identify $\Gamma_M$ with $\Gamma_\Psi$, and regard it as a linear algebraic group over $\mathbb{F}_{p^r}(t)$ because of Theorem 2.3.1(a).

**Remark 2.3.2.** Let $n_1, n_2$ be positive integers and $\mathbf{0} := 0_{n_1\times n_2}$ be the zero matrix of size $n_1 \times n_2$. Suppose that the matrix

$$\Phi := \begin{bmatrix} \Phi_1 & \mathbf{0} \\ \Phi_3 & \Phi_2 \end{bmatrix} \in \mathrm{GL}_{n_1+n_2}(\bar{k}(t))$$

defines a rigid analytically trivial pre-$t$-motive $M$ of level $r$. Then one can always find its rigid analytic trivialization of the form

$$\Psi := \begin{bmatrix} \Psi_1 & \mathbf{0} \\ \Psi_3 & \Psi_2 \end{bmatrix} \in \mathrm{GL}_{n_1+n_2}(\mathbb{L}).$$

By (3), we have that

$$\Gamma_\Psi \subseteq \left\{ \begin{bmatrix} * & \mathbf{0} \\ * & * \end{bmatrix} \right\} \subseteq \mathrm{GL}_{n_1+n_2}/\mathbb{F}_{p^r}(t).$$

Let $N$ be the sub-pre-$t$-motive (of level $r$) of $M$ defined by $\Phi_1 \in \mathrm{GL}_{n_1}(\bar{k}(t))$ with rigid analytic trivialization $\Psi_1$; then by the Tannakian theory we have a natural surjective morphism

$$\begin{aligned} \pi : \Gamma_\Psi(\overline{\mathbb{F}_p(t)}) &\twoheadrightarrow \Gamma_{\Psi_1}(\overline{\mathbb{F}_p(t)}), \\ \gamma &\mapsto \pi(\gamma). \end{aligned} \tag{4}$$

In fact, $\pi(\gamma)$ comes from the restriction of the action of $\gamma$ to the fiber functor of $\mathcal{R}_N$ (which is a full subcategory of $\mathcal{R}_M$). Precisely, $\pi(\gamma)$ is the matrix cut out from the upper left square of $\gamma$ with size $n_1$ (for detailed arguments, see [Papanikolas 2008, §6.2.2]).

## 3. A dimension criterion

**3.1.** ***Pre-t-motives of type*** **SV.** Let $r$ be a fixed positive integer and let $\{n_1, \dots, n_h\}$ be $h$ nonnegative integers. We say that a pre-$t$-motive $M$ of level $r$ is of type **SV** (its Galois group being an extension of a split torus by a vector group) if the multiplication by $\sigma^r$ on $M$ is represented by the matrix

$$\Phi := \bigoplus_{i=1}^{h} A_i, \quad A_i := \begin{bmatrix} a_i & 0 & \cdots & 0 \\ a_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{in_i} & 0 & \cdots & 1 \end{bmatrix} \in \mathrm{GL}_{1+n_i}(\bar{k}(t)),$$

which has rigid analytic trivialization

$$\Psi := \bigoplus_{i=1}^{h} F_i, \quad F_i := \begin{bmatrix} f_i & 0 & \cdots & 0 \\ f_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{in_i} & 0 & \cdots & 1 \end{bmatrix} \in \mathrm{GL}_{1+n_i}(\mathbb{L}).$$

Let $T$ be the Galois group associated to the difference equation

$$\begin{bmatrix} f_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f_h \end{bmatrix}^{(-r)} = \begin{bmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_h \end{bmatrix} \begin{bmatrix} f_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f_h \end{bmatrix},$$

and note that by (3), $T$ is a subtorus of the $h$-dimensional split torus in $\mathrm{GL}_h/\mathbb{F}_{p^r}(t)$. By the same reason as (4), we have the natural projection of Galois groups

$$\Gamma_\Psi \twoheadrightarrow T, \tag{5}$$

given in terms of coordinates by

$$\bigoplus_{i=1}^{h} \begin{bmatrix} x_i & 0 & \cdots & 0 \\ x_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{in_i} & 0 & \cdots & 1 \end{bmatrix} \mapsto [x_1] \oplus \cdots \oplus [x_h].$$

One has also the exact sequence of linear algebraic groups

$$1 \to V \to \Gamma_\Psi \twoheadrightarrow T \to 1, \tag{6}$$

where $V$ is a vector group contained in the $\left(\sum_{i=1}^{h} n_i\right)$-dimensional "coordinate" vector group $G$, which is defined as the set of matrices

$$\bigoplus_{i=1}^{h} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ x_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{in_i} & 0 & \cdots & 1 \end{bmatrix}$$

with usual multiplication. This subgroup $V$ is the unipotent radical of the (solvable) Galois group $\Gamma_\Psi$, and $\dim \Gamma_\Psi = \dim V + \dim T$.

**Definition 3.1.1.** Let $M$ be a rigid analytically trivial pre-$t$-motive of level $r$ that is of **SV** type given as above. We say that its Galois group $\Gamma_\Psi$ is *full* if $\dim V = \sum_{i=1}^{h} n_i$, that is, $V = G$.

**3.2.** ***Criterion for direct sum motives to have full Galois group.*** We continue with the notation of Section 3.1. For each $i$, $1 \le i \le h$, the $1 \times 1$ matrix $[a_i]$ defines a sub-pre-$t$-motive (of level $r$) of $M$, which is one-dimensional over $\bar{k}(t)$. Its rigid analytic trivialization is given by $f_i$ satisfying

$$[f_i]^{(-r)} = [a_i][f_i].$$

We call one such sub-pre-$t$-motive of level $r$ a *diagonal* of the pre-$t$-motive $M$. By Theorem 2.3.1, the Galois group of a diagonal of $M$ is $\mathbb{G}_m$ if and only if the corresponding trivialization $f_i$ is transcendental over $\bar{k}(t)$. In this situation, the canonical projection $T \to \mathbb{G}_m$ on the $i$-th coordinate of $T$ is surjective.

There is a canonical action of $T$ on $G$ given in terms of coordinates by

$$([x_1] \oplus \cdots \oplus [x_h]) * \left( \oplus_{i=1}^{h} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ x_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{in_i} & 0 & \cdots & 1 \end{bmatrix} \right) := \bigoplus_{i=1}^{h} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ x_i x_{i1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_i x_{in_i} & 0 & \cdots & 1 \end{bmatrix}.$$

This action induces an action of $T$ on $V$ compatible with the one coming from (6). Hence, we have that given any element $\gamma \in V$ whose $x_{ij}$-coordinate is nonzero, the orbit inside $V$ given by the action of $T$ on $\gamma$ must be infinite if $f_i$ is transcendental over $\bar{k}(t)$.

**Definition 3.2.1.** Let $M_i$ be a rigid analytically trivial pre-$t$-motive of level $r$ that is of type **SV** for $i = 1, \ldots, d$. We say that this set of pre-$t$-motives $\{M_i\}_{i=1}^{d}$ is diagonally independent if for any $1 \le i, j \le d$, $i \ne j$, the Galois group of $N_i \oplus N_j$ is a two-dimensional torus over $\mathbb{F}_{p^r}(t)$, where $N_i$ is any diagonal of $M_i$ and $N_j$ is any diagonal of $M_j$.

**Theorem 3.2.2.** *Given any positive integer $r$, let $M_1, \dots, M_d$ be rigid analytically trivial pre-$t$-motives of level $r$ that are of type* **SV** *and have full Galois groups. Suppose that the set $\{M_i\}_{i=1}^d$ is diagonally independent. Then the Galois group of $M := \bigoplus_{i=1}^d M_i$ is also full.*

*Proof.* We first explain that without loss of generality, we may assume $d = 2$. We prove the result by induction on $d$. Consider $M = \left(\bigoplus_{i=1}^{d-1} M_i\right) \oplus M_d$. Then the induction hypothesis implies that the Galois group of $\left(\bigoplus_{i=1}^{d-1} M_i\right)$ is full. Since the following argument of the case $d = 2$ can be applied to the computation of the Galois group of the direct sum of $\left(\bigoplus_{i=1}^{d-1} M_i\right)$ and $M_d$, we may assume $d = 2$.

Now we set $d = 2$ and let the pre-$t$-motive $M_i$ be defined by a matrix $\Phi_i$, for $i = 1, 2$, with $\Psi = \Psi_1 \oplus \Psi_2$ a rigid analytic trivialization of $M$, and let $\Gamma_\Psi$, $\Gamma_{\Psi_1}$, $\Gamma_{\Psi_2}$ be the Galois groups. The unipotent radicals of these groups are denoted by $V$, $V_1$, $V_2$ respectively. Let $G$, $G_1$, $G_2$ be the coordinate vector groups containing $V$, $V_1$, $V_2$, respectively. Suppose the matrix $\Phi_1$ has $h$ diagonal blocks, and let the coordinates of $G_1$ be denoted by $x_{ij}$, $i = 1, \dots, h$, $j = 1, \dots, n_i$. Similarly, let $y_{ij}$, $i = 1, \dots, \ell$, $j = 1, \dots, m_i$ denote the coordinates of $G_2$. Any subspace $W \subseteq G$ obtained by setting some of these coordinates to 0 is called a linear coordinate subspace. The hypothesis that the Galois group of $M_i$ is full means exactly that $G_i = V_i$, for $i = 1, 2$. We are going to prove that $G = V$.

Suppose $V$ has codimension $s$ in the coordinate vector group $G$. We can find a linear coordinate subspace $W \subseteq G$ of dimension $s$ such that $W \cap V$ is of dimension 0. Since $W \cap V$ is invariant under $T$, it must be equal to the neutral element of $G$, because the hypothesis that $M_1$ and $M_2$ are diagonally independent implies in particular that the Galois group of any diagonal of $M$ is $\mathbb{G}_m$.

Let $W' \subseteq G$ be the linear coordinate space given by those coordinates disjoint from those of $W$. Then the natural projection from $G$ to $W'$ induces on $V$ an isomorphism of vector groups. Composing the inverse of this isomorphism with the surjective morphism $\pi_1$ in the diagram

$$\begin{array}{ccccccccc} 1 & \longrightarrow & V & \longrightarrow & \Gamma_\Psi & \longrightarrow & T & \longrightarrow & 1 \\ & & \pi_1 \downarrow & & \pi_1 \downarrow & & \downarrow & & \\ 1 & \longrightarrow & G_1 & \longrightarrow & \Gamma_{\Psi_1} & \longrightarrow & T_1 & \longrightarrow & 1, \end{array}$$

we obtain a morphism $\pi_1$ from $W'$ onto $G_1$ that is furthermore a $T$-morphism.

We contend that under the hypothesis that $M_1$ and $M_2$ are diagonally independent, $\pi_1$ maps $G_2 \cap W'$ to zero. This contention results from the following basic lemma by taking any diagonal $N_1$ (resp. $N_2$) of $M_1$ (resp. $M_2$) and considering the restriction of the above morphism $\pi_1$ to a single block. Now since $G = G_1 \times G_2$, it follows that $\pi_1(G_1 \cap W') = G_1$. Thus $G_1 \subseteq W'$. Similarly we also have $G_2 \subseteq W'$, and hence $G = W'$ and $G = V$. □

**Lemma 3.2.3.** *Let $G_1$ (resp. $G_2$) be the vector group with coordinates*

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ x_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \cdots & 1 \end{bmatrix}, \quad \left( \textit{resp.} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ y_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ y_m & 0 & \cdots & 1 \end{bmatrix} \right).$$

*Let $\mathbb{G}_m^2$ act on $G_1$ (resp. $G_2$) by*

$$\mathbb{G}_m^2 \ni (x, y) : x_i \mapsto x x_i, \ 1 \le i \le n, \quad (\textit{resp. } y_j \mapsto y y_j, \ 1 \le j \le m).$$

*If $\pi_1 : G_2 \to G_1$ is a $\mathbb{G}_m^2$-morphism, then $\pi_1 \equiv 0$.*

## 4. Application to zeta values

Throughout this section, we use $ri$ for two subscripts (for example, $D_{ri}$ depends on the two parameters $r$ and $i$), and use $*, rn$ for three subscripts (for example, $L_{\alpha,rn}$ depends on the three parameters $\alpha$, $r$ and $n$).

**4.1.** ***Carlitz theory.*** Throughout Section 4 we fix a positive integer $r$. Recall the Carlitz $\mathbb{F}_{p^r}[t]$-module, denoted by $C_r$, which is given by the $\mathbb{F}_{p^r}$-linear ring homomorphism

$$C_r = (t \mapsto (x \mapsto \theta x + x^{p^r})) : \mathbb{F}_{p^r}[t] \to \mathrm{End}_{\mathbb{F}_{p^r}}(\mathbb{G}_a).$$

Note that when we regard $C_r$ as a Drinfeld $\mathbb{F}_p[t]$-module, it is of rank $r$ [Goss 1996; Thakur 2004]. One has the Carlitz exponential associated to $C_r$:

$$\exp_{C_r}(z) := \sum_{i=0}^{\infty} \frac{z^{p^{ri}}}{D_{ri}}.$$

Here we set

$$D_{r0} := 1, \qquad D_{ri} := \prod_{j=0}^{i-1} (\theta^{p^{ri}} - \theta^{p^{rj}}), \quad i \ge 1.$$

Now $\exp_{C_r}(z)$ is an entire power series in $z$ satisfying the functional equation

$$\exp_{C_r}(\theta z) = \theta \exp_{C_r}(z) + \exp_{C_r}(z)^{p^r}.$$

Moreover one has the product expansion

$$\exp_{C_r}(z) = z \prod_{0 \ne a \in \mathbb{F}_{p^r}[\theta]} \Big(1 - \frac{z}{a\tilde{\pi}_r}\Big),$$

where

$$\tilde{\pi}_r = \theta(-\theta)^{1/(p^r-1)} \prod_{i=1}^{\infty} \big(1 - \theta^{1-p^{ri}}\big)^{-1}$$

is a fundamental period of $C_r$. We fix once and for all a choice of $(-\theta)^{1/(p^r-1)}$ so that $\tilde{\pi}_r$ is a well-defined element in $\overline{\mathbb{F}_p((1/\theta))}$. We also choose these roots in a compatible way so that when $r$ divides $r'$, the number $(-\theta)^{1/(p^r-1)}$ is a power of $(-\theta)^{1/(p^{r'}-1)}$.

The formal inverse of Carlitz exponential is the Carlitz logarithm. It is the power series

$$\log_{C_r}(z) = \sum_{i=0}^{\infty} \frac{z^{p^{ri}}}{L_{ri}},$$

where

$$L_{r0} := 1, \qquad L_{ri} := \prod_{j=1}^{i} (\theta - \theta^{p^{rj}}).$$

As a function in $z$, $\log_{C_r}(z)$ converges for all $z \in \mathbb{C}_\infty$ with $|z|_\infty < |\theta|_\infty^{p^r/(p^r-1)}$. It satisfies the functional equation

$$\theta \log_{C_r}(z) = \log_{C_r}(\theta z) + \log_{C_r}(z^{p^r})$$

whenever the values in question are defined.

For a positive integer $n$, the $n$-th Carlitz polylogarithm associated to $C_r$ is the series

$$\mathrm{Plog}_{rn}(z) := \sum_{i=0}^{\infty} \frac{z^{p^{ri}}}{L_{ri}^n}, \tag{7}$$

which converges $\infty$-adically for all $z \in \mathbb{C}_\infty$ with $|z|_\infty < |\theta|_\infty^{np^r/(p^r-1)}$. Its value at a particular $z = \alpha \neq 0$ is called the $n$-th polylogarithm of $\alpha$ associated to $C_r$. In transcendence theory we are interested in those polylogarithms of $\alpha \in \bar{k}^\times$, as analogous to classical logarithms of algebraic numbers.

**4.2.** ***Algebraic independence of special functions.*** For any positive integer $r$, let

$$\Omega_r(t) := (-\theta)^{-p^r/(p^r-1)} \prod_{i=1}^{\infty} \left(1 - \frac{t}{\theta^{p^{ri}}}\right) \in \overline{k_\infty}[\![t]\!] \subseteq \mathbb{C}_\infty((t)).$$

$\Omega_r \in \mathbb{E}$, since $|\theta^{p^{ri}}|_\infty \to \infty$. Furthermore, $\Omega_r$ satisfies the functional equation

$$\Omega_r^{(-r)}(t) = (t - \theta)\Omega_r(t), \tag{8}$$

and its specialization at $t = \theta$ gives $\Omega_r(\theta) = -1/\tilde{\pi}_r$.

By (8), the function $\Omega_r$ provides a rigid analytic trivialization of the Carlitz motive $\mathscr{C}_r$ of level $r$ that has the GP property (see Theorem 2.2.3). This is the pre-$t$-motive with underlying space $\bar{k}(t)$ itself and $\sigma^r$ acts by $\sigma^r f = (t-\theta) f^{(-r)}$ for $f \in \mathscr{C}_r$.

For any $d \in \mathbb{N}$, we let $\ell := \mathrm{lcm}(1, \ldots, d)$ and $\ell_r := \ell/r$ for $r = 1, \ldots, d$. We let $\mathscr{C}_r^{(\ell_r)}$ be the $\ell_r$-th derived pre-$t$-motive of $\mathscr{C}_r$ that is a rigid analytically

trivial pre-$t$-motive of level $\ell$ (see Section 2.2). Then we define the direct sum $M = M_d := \bigoplus_{r=1}^{d} \mathscr{C}_r^{(\ell_r)}$. By Corollary 2.2.4, $M$ also has the GP property. We note that the canonical rigid analytical trivialization of $M$ is the diagonal matrix $\Psi \in \mathrm{Mat}_d(\mathbb{E}) \cap \mathrm{GL}_d(\mathbb{L})$ with diagonal entries $\Omega_1, \dots, \Omega_d$.

**Lemma 4.2.1.** *Given any positive integer $d \geq 2$, we let $\ell := \mathrm{lcm}(1, \dots, d)$. Let $M = M_d$ be the rigid analytically trivial pre-$t$-motive of level $\ell$ with rigid analytic trivialization $\Psi$ defined as above. Then we have $\dim \Gamma_\Psi = d$. In particular, the functions $\Omega_1, \dots, \Omega_d$ are algebraically independent over $\bar{k}(t)$ and the values $\tilde{\pi}_1, \dots, \tilde{\pi}_d$ are algebraically independent over $\bar{k}$.*

*Proof.* Suppose $\dim \Gamma_\Psi < d$. Since $\Psi$ is a diagonal matrix with diagonal entries $\Omega_1, \dots, \Omega_d$, by (3) we have that $\Gamma_\Psi \subseteq \mathrm{T}$, where T is the split torus of dimension $d$ in $\mathrm{GL}_d / \mathbb{F}_{p^\ell}(t)$. We let $X_1, \dots, X_d$ be the coordinates of T and $\chi_j$ the character of T that projects the $j$-th diagonal position to $\mathbb{G}_m$. Note that $\{\chi_j\}_{j=1}^d$ generates the character group of T. Hence $\Gamma_\Psi$ is the kernel of some characters of T, that is, canonical generators of the defining ideal for $\Gamma_\Psi$ can be of the form $X_1^{m_1} \cdots X_d^{m_d} - 1$ for some integers $m_1, \dots, m_d$, not all zero. By (3) we have that

$$(\Omega_1^{-m_1} \dots \Omega_d^{-m_d}) \otimes (\Omega_1^{m_1} \dots \Omega_d^{m_d}) = 1 \in \mathbb{L} \otimes_{\bar{k}(t)} \mathbb{L},$$

and hence

$$\beta := \Omega_1^{m_1} \dots \Omega_d^{m_d} \in \bar{k}(t)^\times. \tag{9}$$

We recall that $\Omega_r$ has zeros on $\{\theta^{p^{rj}}\}_{j=1}^\infty$. Since $\beta \in \bar{k}(t)^\times$, it has only finitely many zeros and poles, and hence $\mathrm{ord}_{t=\theta^{p^h}}(\beta) = 0$ for $h \gg 0$. Choose a prime number $p'$ sufficiently large that

- $p' > d$, and
- the order of vanishing of $\beta$ at $t = \theta^{p^{p'}}$ is zero.

These conditions imply that $m_1 = 0$ from (9). Iterating this argument, we conclude that $m_1 = \cdots = m_d = 0$, a contradiction. □

**4.3.** ***Algebraic independence of polylogarithms.*** Given $n \in \mathbb{N}$ and $\alpha \in \bar{k}^\times$ with $|\alpha|_\infty < |\theta|_\infty^{np^r/(p^r-1)}$, we consider the power series

$$L_{\alpha,rn}(t) := \alpha + \sum_{i=1}^{\infty} \frac{\alpha^{p^{ri}}}{(t-\theta^{p^r})^n \cdots (t-\theta^{p^{ri}})^n},$$

which as a function on $\mathbb{C}_\infty$ converges on $|t|_\infty < |\theta|_\infty^{p^r}$. We note that $L_{\alpha,rn}(\theta)$ is exactly the $n$-th polylogarithm of $\alpha$ associated to $C_r$, that is,

$$L_{\alpha,rn}(\theta) = \mathrm{Plog}_{rn}(\alpha).$$

Given a collection of such numbers $\alpha$, say $\alpha_1, \dots, \alpha_m$, we define

$$\Phi_{rn}(\alpha_1, \dots, \alpha_m) := \begin{bmatrix} (t-\theta)^n & 0 & \cdots & 0 \\ \alpha_1^{(-r)}(t-\theta)^n & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_m^{(-r)}(t-\theta)^n & 0 & \cdots & 1 \end{bmatrix} \in \mathrm{GL}_{m+1}(\bar{k}(t)) \cap \mathrm{Mat}_{m+1}(\bar{k}[t])$$

and

$$\Psi_{rn}(\alpha_1, \dots, \alpha_m) := \begin{bmatrix} \Omega_r^n & 0 & \cdots & 0 \\ \Omega_r^n L_{\alpha_1, rn} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_r^n L_{\alpha_m, rn} & 0 & \cdots & 1 \end{bmatrix} \in \mathrm{GL}_{m+1}(\mathbb{L}) \cap \mathrm{Mat}_{m+1}(\mathbb{E}).$$

Then one has [Chang and Yu 2007, §3.1.2]

$$\Psi_{rn}(\alpha_1, \dots, \alpha_m)^{(-r)} = \Phi_{rn}(\alpha_1, \dots, \alpha_m)\Psi_{rn}(\alpha_1, \dots, \alpha_m). \tag{10}$$

Hence, $\Phi_{rn}(\alpha_1, \dots, \alpha_m)$ defines a rigid analytically trivial pre-$t$-motive of level $r$ that has the GP property.

In [Chang and Yu 2007], we followed Papanikolas' methods to generalize the algebraic independence of Carlitz logarithms to algebraic independence of polylogarithms. Precisely, by [Chang and Yu 2007, Theorem 3.1] and Theorem 2.2.3 we have:

**Theorem 4.3.1.** *Given any positive integers $r$ and $n$, let $\alpha_1, \dots, \alpha_m \in \bar{k}^\times$ satisfy $|\alpha_i|_\infty < |\theta|_\infty^{np^r/(p^r-1)}$ for $i = 1, \dots, m$. Then*

$$\begin{aligned} \dim_{\mathbb{F}_{p^r}(\theta)} N_{rn} &= \mathrm{tr.deg}_{\bar{k}}\, \bar{k}(\tilde{\pi}_r^n, L_{\alpha_1, rn}(\theta), \dots, L_{\alpha_m, rn}(\theta)) \\ &= \mathrm{tr.deg}_{\bar{k}(t)}\, \bar{k}(t)(\Omega_r^n, L_{\alpha_1, rn}, \dots, L_{\alpha_m, rn}), \end{aligned}$$

*where*

$$N_{rn} := \mathbb{F}_{p^r}(\theta)\text{-}\mathrm{Span}\{\tilde{\pi}_r^n, L_{\alpha_1, rn}(\theta), \dots, L_{\alpha_m, rn}(\theta)\}.$$

### 4.4. *Formulas for zeta values.*

**4.4.1.** *Euler–Carlitz relations.* Fix a positive integer $r$ for this subsection. Carlitz [1935] introduced the power sum

$$\zeta_r(n) := \sum_{\substack{a \in \mathbb{F}_{p^r}[\theta] \\ a \text{ monic}}} \frac{1}{a^n} \in \mathbb{F}_p((1/\theta)) \quad (n \text{ a positive integer}),$$

which are the Carlitz zeta values associated to $\mathbb{F}_{p^r}[\theta]$.

Writing down a $p^r$-adic expansion $\sum_i n_{ri}p^{ri}$ of $n$, we let

$$\Gamma_{r(n+1)} := \prod_{i=0}^{\infty} D_{ri}^{n_{ri}}.$$

We call $\Gamma_{r(n+1)}$ the *Carlitz factorials* associated to $\mathbb{F}_{p^r}[\theta]$. The *Bernoulli–Carlitz numbers* $B_{rn}$ in $\mathbb{F}_p(\theta)$ are given by the following expansions from the Carlitz exponential series

$$\frac{z}{\exp_{C_r}(z)} = \sum_{n=0}^{\infty} \frac{B_{rn}}{\Gamma_{r(n+1)}} z^n.$$

We state the *Euler–Carlitz relations*:

**Theorem 4.4.2** [Carlitz 1935]. *For all positive integers $n$ divisible by $p^r - 1$, one has*

$$\zeta_r(n) = \frac{B_{rn}}{\Gamma_{r(n+1)}} \tilde{\pi}_r^n. \tag{11}$$

We call a positive integer $n$ $(p, r)$-*even* if it is divisible by $p^r - 1$; otherwise we call it $(p, r)$-*odd*. Thus, when $p = 2$ and $r = 1$, all positive integers are even.

**4.4.3.** *The Anderson–Thakur formula.* Anderson and Thakur [1990] introduced the $n$-th tensor power of the Carlitz $\mathbb{F}_{p^r}[t]$-module $C_r$, and they related $\zeta_r(n)$ to the last coordinate of the logarithm associated to the $n$-th tensor power of $C_r$ for each positive integer $n$. More precisely, they interpreted $\zeta_r(n)$ as $\mathbb{F}_{p^r}(\theta)$-linear combinations of $n$-th Carlitz polylogarithms of algebraic numbers:

**Theorem 4.4.4** [Anderson and Thakur 1990]. *Given any positive integers $r$ and $n$, one can find a sequence $h_{0,rn}, \dots, h_{l_{rn},rn} \in \mathbb{F}_{p^r}(\theta)$, $l_{rn} < np^r/(p^r - 1)$, such that*

$$\zeta_r(n) = \sum_{i=0}^{l_{rn}} h_{i,rn} \operatorname{Plog}_{rn}(\theta^i), \tag{12}$$

*where $\operatorname{Plog}_{rn}(z)$ is defined as in (7). In the special case of $n \le p^r - 1$,*

$$\zeta_r(n) = \operatorname{Plog}_{rn}(1).$$

**Definition 4.4.5.** Given any positive integer $r$, for each $n \in \mathbb{N}$, $(p^r - 1) \nmid n$, with $l_{rn}$ as given by (12), we fix a finite subset

$$\{\alpha_{0,rn}, \dots, \alpha_{m_{rn},rn}\} \subseteq \{1, \theta, \dots, \theta^{l_{rn}}\}$$

such that

$$\{\tilde{\pi}_r^n, \mathcal{L}_{0,rn}(\theta), \dots, \mathcal{L}_{m_{rn},rn}(\theta)\} \quad \text{and} \quad \{\tilde{\pi}_r^n, \zeta_r(n), \mathcal{L}_{1,rn}(\theta), \dots, \mathcal{L}_{m_{rn},rn}(\theta)\}$$

are $\mathbb{F}_{p^r}(\theta)$-bases for $N_{rn}$, where $\mathcal{L}_{j,rn}(t) := L_{\alpha_j,rn}(t)$ for $j = 0, \dots, m_{rn}$. This can be done because of (12) (see [Chang and Yu 2007, § 4.1]), and note that $m_{rn} + 2$

is the dimension of $N_{rn}$ over $\mathbb{F}_{p^r}(\theta)$. In the case of $p=2$ and $r=1$, the $\mathbb{F}_p(\theta)$-dimension of $N_{11}$ is 1 and we set $m_{11}:=-1$.

**Definition 4.4.6.** Given any positive integers $s$ and $d$ with $d\geq 2$, for each $1\leq r\leq d$, we define

$$\begin{cases} U_r(s)=\{1\} & \text{if } p=2 \text{ and } r=1;\\ U_r(s)=\{1\leq n\leq s;\ p\nmid n,\ (p^r-1)\nmid n\} & \text{otherwise.}\end{cases}$$

For each $n\in U_r(s)$, we define that if $p=2$ and $r=1$,

$$\Phi_{rn}:=(t-\theta)\in \mathrm{GL}_1(\bar{k}(t)),$$
$$\Psi_{rn}:=\Omega_1\in \mathrm{GL}_1(\mathbb{L}),$$

otherwise

$$\Phi_{rn}:=\Phi_{rn}(\alpha_{0,rn},\dots,\alpha_{m_{rn},rn})\in \mathrm{GL}_{(m_{rn}+2)}(\bar{k}(t)),$$
$$\Psi_{rn}:=\Psi_{rn}(\alpha_{0,rn},\dots,\alpha_{m_{rn},rn})\in \mathrm{GL}_{(m_{rn}+2)}(\mathbb{L}).$$

By Theorem 4.3.1, we have that

$$\dim \Gamma_{\Psi_{rn}}=\mathrm{tr.deg}_{\bar{k}(t)}\,\bar{k}(t)(\Psi_{rn})=m_{rn}+2.$$

Put $\Phi_r:=\bigoplus_{n\in U_r(s)}\Phi_{rn}$; then $\Phi_r$ defines a rigid analytically trivial pre-$t$-motive $M_r$ of level $r$ with rigid analytic trivialization $\Psi_r:=\bigoplus_{n\in U_r(s)}\Psi_{rn}$ (see (10)). Also, $M_r$ is of type **SV**, and by Theorem 2.2.3, $M_r$ has the GP property. The main theorem of [Chang and Yu 2007] is the following:

**Theorem 4.4.7** [Chang and Yu 2007, Theorem 4.5]. *For any positive integers $s$ and $r$, the Galois group $\Gamma_{M_r}$ over $\mathbb{F}_{p^r}(t)$ is full, that is,*

$$\dim \Gamma_{M_r}=1+\sum_{n\in U_r(s)}(m_{rn}+1).$$

**4.5.** ***Proof of Theorem 1.2.1.*** Given any integer $d\geq 2$, we put $\ell:=\mathrm{lcm}(1,\dots,d)$ and $\ell_r:=\ell/r$ for $r=1,\dots,d$. For each $1\leq r\leq d$, let $\mathbf{M}_r:=M_r^{(\ell_r)}$ be the $\ell_r$-th derived pre-$t$-motive of $M_r$ defined as above. Note that $\mathbf{M}_r$ is a rigid analytically trivial pre-$t$-motive of level $\ell$ that is still of type **SV**. By Proposition 2.2.2, each $\mathbf{M}_r$ has the GP property, and by Theorem 4.4.7, its Galois group $\Gamma_{\mathbf{M}_r}$ is full. Further, for each $1\leq r\leq d$ any diagonal of $\mathbf{M}_r$ has canonical rigid analytic trivialization given by $\Omega_r^n$ for some $n\in U_r(s)$, and hence its Galois group is $\mathbb{G}_m$ because $\Omega_r$ is transcendental over $\bar{k}(t)$.

Since by Lemma 4.2.1 the functions $\Omega_1,\dots,\Omega_d$ are algebraically independent over $\bar{k}(t)$, particularly the Galois group of $N_i\oplus N_j$ is a two-dimensional torus over $\mathbb{F}_{p^\ell}(t)$ for any diagonal $N_i$ (resp. $N_j$) of $\mathbf{M}_i$ (resp. $\mathbf{M}_j$) with $i\neq j$, $1\leq i,j\leq d$. Put $\mathbf{M}:=\bigoplus_{r=1}^d\mathbf{M}_r$, and note that $\mathbf{M}$ has the GP property by Corollary 2.2.4. Applying Theorem 3.2.2 to this situation, we obtain the explicit dimension of $\Gamma_{\mathbf{M}}$:

**Theorem 4.5.1.** *Given any positive integers $s$ and $d$ with $d \geq 2$, let $\mathbf{M}$ be defined as above. Then the Galois group $\Gamma_{\mathbf{M}}$ is full, that is,*

$$\dim \Gamma_{\mathbf{M}} = d + \sum_{r=1}^{d} (m_{rn} + 1).$$

As a consequence, we completely determine all the algebraic relations among the families of Carlitz zeta values:

**Corollary 4.5.2.** *Given any positive integers $d$ and $s$, the transcendence degree of the field*

$$\bar{k}\Big( \bigcup_{r=1}^{d} \{\tilde{\pi}_r, \zeta_r(1), \dots, \zeta_r(s)\} \Big)$$

*over $\bar{k}$ is*

$$\sum_{r=1}^{d} \left( s - \left\lfloor \frac{s}{p} \right\rfloor - \left\lfloor \frac{s}{p^r - 1} \right\rfloor + \left\lfloor \frac{s}{p(p^r - 1)} \right\rfloor + 1 \right).$$

*Proof.* We may assume $d \geq 2$, since the case $d = 1$ is already given in [Chang and Yu 2007, Corollary 4.6]. For $1 \leq r \leq d$, let

$$\begin{cases} V_1(s) := \varnothing & \text{if } p = 2; \\ V_r(s) := U_r(s) & \text{otherwise.} \end{cases}$$

Since $\mathbf{M}$ has the GP property, by Theorem 4.5.1 we see that the elements of the set

$$\{\Omega_1(\theta), \dots, \Omega_d(\theta)\} \cup \Big( \bigcup_{r=1}^{d} \bigcup_{n \in V_r(s)} \{\mathscr{L}_{0,rn}(\theta), \dots, \mathscr{L}_{m_{rn},rn}(\theta)\} \Big) \tag{13}$$

are algebraically independent over $\bar{k}$. In particular, by Definition 4.4.5, we have that

$$\{\tilde{\pi}_1, \dots, \tilde{\pi}_d\} \cup \Big( \bigcup_{r=1}^{d} \bigcup_{n \in V_r(s)} \{\zeta_r(n)\} \Big)$$

is an algebraically independent set over $\bar{k}$. Counting the cardinality of $V_r(s)$ for each $1 \leq r \leq d$, we complete the proof. □

## Acknowledgements

We thank D. Thakur for many helpful discussions during the preparation of this manuscript. The first author thanks the NSC and National Center for Theoretical Sciences for financial support, and Texas A&M University for its hospitality.

## References

[Anderson 1986] G. W. Anderson, "$t$-motives", *Duke Math. J.* **53**:2 (1986), 457–502. MR 87j:11042 Zbl 0679.14001

[Anderson and Thakur 1990] G. W. Anderson and D. S. Thakur, "Tensor powers of the Carlitz module and zeta values", *Ann. of Math.* (2) **132**:1 (1990), 159–191. MR 91h:11046 Zbl 0713.11082

[Anderson et al. 2004] G. W. Anderson, W. D. Brownawell, and M. A. Papanikolas, "Determination of the algebraic relations among special Γ-values in positive characteristic", *Ann. of Math.* (2) **160**:1 (2004), 237–313. MR 2005m:11140 Zbl 1064.11055

[Carlitz 1935] L. Carlitz, "On certain functions connected with polynomials in a Galois field", *Duke Math. J.* **1**:2 (1935), 137–168. MR 1545872 Zbl 0012.04904

[Chang 2009] C.-Y. Chang, "A note on a refined version of Anderson–Brownawell–Papanikolas criterion", *J. Number Theory* **129**:3 (2009), 729–738. MR 2010a:11144 Zbl 1187.11022

[Chang and Yu 2007] C.-Y. Chang and J. Yu, "Determination of algebraic relations among special zeta values in positive characteristic", *Adv. Math.* **216**:1 (2007), 321–345. MR 2008j:11114 Zbl 1123.11025

[Chang et al. 2010] C.-Y. Chang, M. A. Papanikolas, D. S. Thakur, and J. Yu, "Algebraic independence of arithmetic gamma values and Carlitz zeta values", *Adv. Math.* **223**:4 (2010), 1137–1154. MR 2011d:11180 Zbl 05671646

[Denis 1998] L. Denis, "Indépendance algébrique de différents $\pi$", *C. R. Acad. Sci. Paris Sér. I Math.* **327**:8 (1998), 711–714. MR 99i:11059 Zbl 0922.11055

[Goss 1996] D. Goss, *Basic structures of function field arithmetic*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **35**, Springer, Berlin, 1996. MR 97i:11062 Zbl 0874.11004

[Papanikolas 2008] M. A. Papanikolas, "Tannakian duality for Anderson–Drinfeld motives and algebraic independence of Carlitz logarithms", *Invent. Math.* **171**:1 (2008), 123–174. MR 2009b:11127 Zbl 05236753

[Thakur 2004] D. S. Thakur, *Function field arithmetic*, World Scientific, River Edge, NJ, 2004. MR 2005h:11115 Zbl 1061.11001

[Wade 1941] L. I. Wade, "Certain quantities transcendental over GF($p^n, x$)", *Duke Math. J.* **8** (1941), 701–720. MR 3,263f Zbl 0063.08101

[Yu 1991] J. Yu, "Transcendence and special zeta values in characteristic $p$", *Ann. of Math.* (2) **134**:1 (1991), 1–23. MR 92g:11075 Zbl 0734.11040

[Yu 1997] J. Yu, "Analytic homomorphisms into Drinfeld modules", *Ann. of Math.* (2) **145**:2 (1997), 215–233. MR 98c:11054 Zbl 0881.11055

Communicated by Brian Conrad
Received 2010-01-27 Revised 2010-10-18 Accepted 2010-11-21

cychang@math.cts.nthu.edu.tw *National Center for Theoretical Sciences, Mathematics Division, National Tsing Hua University, Hsinchu City 30042, Taiwan*
http://math.cts.nthu.edu.tw/~cychang/

map@math.tamu.edu *Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, United States*
http://www.math.tamu.edu/~map/

yu@math.ntu.edu.tw *Department of Mathematics, National Taiwan University, Taipei City 106, Taiwan*
http://www.math.ntu.edu.tw/~yu/

## 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@mathscipub.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 5 No. 1 2011