Algebra & Number Theory

Volume 10

2016

No. 9

msp

Algebra & Number Theory
msp.org/ant

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

The subscription price for 2016 is US $290/year for the electronic version, and $485/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

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

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

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1845

msp

Torsion pour les variétés abéliennes de type I et II

Marc Hindry et Nicolas Ratazzi

Soit A une variété abélienne définie sur un corps de nombres K. Le nombre de points de torsion définis sur une extension finie L est borné polynomialement en terme du degré $[L:K]$ de L sur K. Sous les trois hypothèses suivantes nous calculons l'exposant optimal dans cette borne, en terme de la dimension des sous-variétés abéliennes de A et de leurs anneaux d'endomorphismes. Les trois hypothèses faites sur A sont les suivantes : (1) A est géométriquement isogène à un produit de variétés abéliennes simples de type I ou II dans la classification d'Albert ; (2) A est de "type Lefschetz" c'est-à-dire que le groupe de Mumford–Tate est le groupe des similitudes symplectiques commutant aux endomorphismes ; (3) A vérifie la conjecture de Mumford–Tate. Le résultat est notamment inconditionnel (i.e., ces trois hypothèses sont vérifiées) pour un produit de variétés abéliennes simples de type I ou II et de dimension relative impaire. Par ailleurs nous prouvons, en étendant des résultats de Serre, Pink et Hall, la conjecture de Mumford–Tate pour quelques nouveaux cas de variétés abéliennes de type Lefschetz.

Let A be an abelian variety defined over a number field K. The number of torsion points that are rational over a finite extension L is bounded polynomially in terms of the degree $[L:K]$ of L over K. Under the following three conditions we compute the optimal exponent for this bound, in terms of the dimension of abelian subvarieties and their endomorphism rings. The three hypotheses are the following: (1) A is geometrically isogenous to a product of simple abelian varieties of type I or II, according to the Albert classification; (2) A is of "Lefschetz type," that is, the Mumford–Tate group is the group of symplectic similitudes which commute with the endomorphism ring. (3) A satisfies the Mumford–Tate conjecture. This result is unconditional (i.e., the three hypotheses are satisfied) for a product of simple abelian varieties of type I or II with odd relative dimension. Further, building on work of Serre, Pink and Hall, we also prove the Mumford–Tate conjecture for a few new cases of abelian varieties of Lefschetz type.

MSC2010 : primary 11G10; secondary 11F80, 14K15, 14KXX.
Mots-clefs : abelian varieties, Galois representations, Mumford–Tate group, torsion points, symplectic group.

1. Introduction

Soit A/K une variété abélienne, définie sur un corps de nombres K, de dimension $g \geq 1$. Le classique théorème de Mordell–Weil assure que le groupe $A(F)$ des points F-rationnels de A est de type fini pour toute extension finie F/K. Un problème naturel qui se pose est de déterminer le sous-groupe de torsion $A(F)_{\text{tors}}$. Un premier problème consiste en fait à borner explicitement le cardinal de $A(F)_{\text{tors}}$ lorsque A ou F varient. Comme dans les articles [Ratazzi 2007 ; Hindry et Ratazzi 2012 ; 2010] auxquels ce papier fait suite, nous nous intéressons ici au cas où l'on fixe A et où l'on fait varier F parmi les extensions finies de K ; l'objectif étant d'obtenir une borne avec une dépendance explicite et optimale en le degré $[F : K]$. Introduisons maintenant l'invariant que nous allons étudier.

Définition 1.1. On pose

$$\gamma(A) = \inf\bigl\{x > 0 \mid \forall F/K \text{ finie, } |A(F)_{\text{tors}}| \ll [F : K]^x\bigr\}.$$

La notation $\ll$ signifie qu'il existe une constante C, ne dépendant que de A/K, telle que l'on a $|A(F)_{\text{tors}}| \leq C[F : K]^x$.

On peut traduire la définition en le fait que $\gamma(A)$ est l'exposant le plus petit possible tel que pour tout $\varepsilon > 0$, il existe une constante $C(\varepsilon) = C(A/K, \varepsilon)$ telle que pour toute extension finie F/K on a

$$|A(F)_{\text{tors}}| \leq C(\varepsilon)[F : K]^{\gamma(A)+\varepsilon}.$$

Un résultat général dû à Masser [1986] donne une borne simple :

$$\gamma(A) \leq \dim A.$$

Cette borne est optimale lorsque A est une puissance d'une courbe elliptique avec multiplication complexe ; il est fort probable que la borne de Masser n'est jamais optimale dans les autres cas. L'invariant $\gamma(A)$ est calculé dans [Hindry et Ratazzi 2010] pour un produit de courbes elliptiques et, de manière différente, dans [Ratazzi 2007] pour une variété abélienne de type CM et dans [Hindry et Ratazzi 2012] pour une variété abélienne "générique". Le problème analogue pour les modules de Drinfeld est traité dans [Breuer 2010].

Nous notons $\text{End}(A)$ l'anneau d'endomorphismes de $A/\overline{K}$ et posons $\text{End}^0(A) := \text{End}(A) \otimes \mathbb{Q}$. On étudie dans ce texte les deux cas suivants :

- La variété abélienne A est de type I, i.e., $E := \text{End}(A) \otimes \mathbb{Q}$ est un corps totalement réel de degré e ; en particulier $g = he$ où h est entier.
- La variété abélienne A est de type II, i.e., $D := \text{End}(A) \otimes \mathbb{Q}$ est une algèbre de quaternions totalement indéterminée de centre un corps totalement réel E de degré e ; en particulier $g = 2he$ où h est entier.

Les deux autres cas (type III, i.e., algèbre de quaternions totalement définie, et type IV, i.e., algèbre à division sur un corps CM) sont de nature différente et seront traités dans une autre publication.

Définition 1.2. L'entier h précédent s'appelle la *dimension relative* de A.

On sait que, pour une telle variété abélienne de type I ou II, le groupe de Hodge est toujours contenu dans $\mathrm{Res}_{E/\mathbb{Q}}\,\mathrm{Sp}_{2h}$ et est génériquement égal à ce dernier groupe. Lorsque h est *impair* ou égal à 2, on sait que le groupe de Hodge est effectivement $\mathrm{Res}_{E/\mathbb{Q}}\,\mathrm{Sp}_{2h}$ et, de plus, que la conjecture de Mumford–Tate est vraie (cf. le théorème 3.7 plus loin). Notons que le cas de type I avec $e = 1$ correspond aux variétés abéliennes de type GSp (dans la terminologie de [Hindry et Ratazzi 2012]).

Pour tout premier ℓ, on note $T_\ell(A) = \varprojlim A[\ell^n]$ le module de Tate ℓ-adique de A ; considérant l'action du groupe de Galois $G_K := \mathrm{Gal}(\overline{K}/K)$ sur les points de ℓ^∞-torsion, on associe naturellement à A/K, la représentation ℓ-adique

$$\rho_{\ell^\infty,A} : G_K \to \mathrm{GL}(T_\ell(A)) \simeq \mathrm{GL}_{2g}(\mathbb{Z}_\ell).$$

On pose $V_\ell(A) = T_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$. Rappelons que $V_\ell(A)$ est naturellement le dual de $H^1_{\text{ét}}(A \times \overline{K}, \mathbb{Q}_\ell)$ et que, si l'on note $V = V(A) = H_1(A, \mathbb{Q})$, alors, comme $\mathbb{Q}$-espace vectoriel, $V \cong \mathbb{Q}^{2g}$, et $V \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ s'identifie naturellement à $V_\ell(A)$. Si $\psi : V \times V \to \mathbb{Q}$ est la forme symplectique associée à une polarisation, sa tensorisée par $\mathbb{Q}_\ell$ s'identifie avec la forme symplectique provenant de l'accouplement de Weil associé à la même polarisation $T_\ell(A) \times T_\ell(A) \to \mathbb{Z}_\ell(1) \cong \mathbb{Z}_\ell$.

Définition 1.3. Le *groupe de Lefschetz* de A, noté $L(A)$, est le commutateur [1], en tant que $\mathbb{Q}$-groupe algébrique, de $\mathrm{End}^0(A)$ dans $\mathrm{Sp}(V, \psi) = \mathrm{Sp}_{2g}$.

On note également G_ℓ ou $G_{\ell,A}$ si besoin, l'enveloppe algébrique de l'image de ρ_{ℓ^∞}, et $H_{\ell,A}$ la composante neutre de $G_\ell \cap \mathrm{SL}(V_\ell)$, c'est-à-dire

$$G_{\ell,A} := (\overline{\rho_{\ell,A}(G_K)}^{\mathrm{Zar}})^0 \quad \text{et} \quad H_{\ell,A} := (G_\ell \cap \mathrm{SL}(V_\ell))^0.$$

On note $\mathrm{MT}(A)$ le groupe de Mumford–Tate et $\mathrm{Hdg}(A)$ le groupe de Hodge d'une variété abélienne A (cf. par exemple [Pink 1998; Moonen 1999]). La relation entre les deux groupes est $\mathrm{MT}(A) = \mathbb{G}_m \cdot \mathrm{Hdg}(A)$ et $\mathrm{Hdg}(A)$ est la composante neutre de $\mathrm{MT}(A) \cap \mathrm{SL}(V)$.

Le groupe des homothéties est contenu dans $G_{\ell,A}$; une autre manière de définir le $\mathbb{Q}_\ell$-groupe algébrique $H_{\ell,A}$, pour se ramener dans $\mathrm{Hdg}(A)$, est de le voir comme l'enveloppe algébrique de $\rho(\mathrm{Gal}(\overline{K}/K(\mu_{\ell^\infty}))$, c'est-à-dire la composante neutre de

1. Ce groupe algébrique est connexe, sauf pour les variétés de type III, cf. Murty [1984] et Milne [1999].

la clôture de Zariski de l'image de Galois dans $\mathrm{GL}_{2g,\mathbb{Q}_\ell}$. On sait alors que l'on a toujours

$$H_{\ell,A} \subset \mathrm{Hdg}(A)_{\mathbb{Q}_\ell} \subset L(A)_{\mathbb{Q}_\ell}.$$

Remarquons que ces deux inclusions n'ont pas le même statut : la conjecture de Mumford–Tate prédit que la première inclusion est toujours une égalité, alors que la seconde inclusion peut ne pas être une égalité.

Définition 1.4. Nous dirons que A est de *type Lefschetz* si $\mathrm{Hdg}(A) = L(A)$ et *pleinement de type Lefschetz* si pour tout premier ℓ, on a $H_{\ell,A} = L(A)_{\mathbb{Q}_\ell}$.

Remarque 1.5. Dans la définition précédente, il suffit en fait de supposer que $H_{\ell,A} = L(A)_{\mathbb{Q}_\ell}$ pour un premier ℓ (cf. [Larsen et Pink 1995, Theorem 4.3]).

Ainsi une variété abélienne de type Lefschetz est pleinement de type Lefschetz si elle vérifie la conjecture de Mumford–Tate.

Théorème 1.6. *Soit A une variété abélienne géométriquement simple de type I ou II définie sur un corps de nombres, dont le centre de l'algèbre d'endomorphismes est un corps de nombres totalement réel E de degré $e = [E : \mathbb{Q}]$. Posons $d := \sqrt{[\mathrm{End}^0(A) : E]}$ (i.e., $d = 1$ ou 2). On suppose que A est pleinement de type Lefschetz ; on a alors*

$$\gamma(A) = \frac{2dhe}{1 + 2eh^2 + he} = \frac{2 \dim A}{\dim \mathrm{MT}(A)}.$$

Remarque 1.7. Il est intéressant d'observer que l'exposant $\gamma(A)$ est beaucoup plus petit que la borne $\gamma(A) \leq g$ donnée par Masser ; par exemple, pour toute variété abélienne vérifiant les hypothèses du théorème, on a pour le type I (resp. le type II) $\gamma(A) < 2/3$ (resp. $\gamma(A) < 4/3$).

Pour énoncer la deuxième partie du corollaire suivant, nous introduisons la notation

$$\Sigma = \left\{ g \geq 1 \;\middle|\; \exists k \geq 3,\ \text{impair}, \exists a \geq 1,\ 2g = (2a)^k \text{ ou } 2g = \binom{2k}{k} \right\}. \tag{1}$$

Pour énoncer la troisième partie du corollaire suivant, rappelons que si une variété abélienne A a réduction semi-stable en une place v, la composante neutre de la fibre spéciale est une extension d'une variété abélienne par un tore ; la dimension de ce tore s'appelle la *dimension torique*. Remarquons également (cf. la preuve du théorème 10.7) que si la variété abélienne A est de type I (resp. de type II) et possède mauvaise réduction semi-stable en une place, alors la dimension torique de la fibre spéciale est un multiple de e (resp. de $2e$). En rassemblant les cas où l'on sait démontrer qu'une variété abélienne de type I ou II est pleinement de type Lefschetz (voir théorème 3.7 et section 10) on obtient le corollaire suivant.

Corollaire 1.8. *Soit E un corps de nombres totalement réel de degré $e = [E : \mathbb{Q}]$. Soit A une variété abélienne définie sur un corps de nombres, telle que A est géométriquement simple de type I ou II et le centre de son algèbre d'endomorphismes est E. On suppose de plus que l'une quelconque des trois hypothèses suivantes est satisfaite :*

(1) *La dimension relative h est un nombre impair ou égal à 2.*

(2) *On a $e = 1$ (ou encore $E = \mathbb{Q}$) et la dimension relative h n'appartient pas au sous-ensemble exceptionnel Σ.*

(3) *La variété abélienne A est de type I (resp. II) et possède une place de mauvaise réduction semi-stable avec dimension torique e (resp. avec dimension torique $2e$).*

On a alors

$$\gamma(A) = \frac{2dhe}{1 + 2eh^2 + he} = \frac{2 \dim A}{\dim \mathrm{MT}(A)}.$$

Remarque 1.9. Lorsque $e = 1$ et A de type I, le point (2) correspond au cas "générique" traité dans [Hindry et Ratazzi 2012].

Remarque 1.10. Comme cela est démontré par Noot [1995], pour chaque type de groupe de Hodge, il existe des variétés abéliennes définies sur un corps de nombres, ayant ce groupe de Hodge donné et vérifiant la conjecture de Mumford–Tate. En particulier, pour tout corps de nombres totalement réel E de degré e sur $\mathbb{Q}$ et tout entier h, il existe des variétés abéliennes A définies sur un corps de nombres, de dimension eh (resp. $2eh$) telles que $\mathrm{End}^0(A) = E$ (resp. $\mathrm{End}^0(A)$ est une algèbre de quaternions indéfinie sur E) et qui sont pleinement de Lefschetz. Ces variétés abéliennes vérifient donc les hypothèses du théorème.

Remarque 1.11. La preuve fournit une inégalité légèrement plus précise que $|A(L)_{\mathrm{tor}}| \ll_\varepsilon [L : K]^{\gamma(A)+\varepsilon}$ de la forme

$$|A(L)_{\mathrm{tor}}| \ll [L : K]^{\gamma(A)+c_1/\log\log[L:K]}.$$

On peut récrire ce résultat en terme de minoration du degré de l'extension engendrée par un sous-groupe de torsion ; nous donnons la teneur du résultat ci-dessous dans le cas de l'extension engendrée par un seul point de torsion.

Théorème 1.12. *Soit A/K une variété abélienne géométriquement simple de type I ou II, de dimension relative h et pleinement de type Lefschetz. Il existe une constante $c_1 := c_1(A, K) > 0$ telle que pour tout point de torsion P d'ordre n dans $A(\overline{K})$, on a la minoration*

$$[K(P) : K] \geq c_1^{\omega(n)} n^{2h}.$$

Remarque 1.13. On peut affaiblir légèrement l'énoncé en écrivant l'inégalité sous la forme

$$[K(P):K] \gg n^{2h-c/\log\log n} \quad \text{ou encore} \quad [K(P):K] \gg_{\varepsilon} n^{2h-\varepsilon} \text{ pour tout } \varepsilon > 0.$$

Par ailleurs notons que, toujours pour un point P d'ordre n, on a trivialement $[K(P):K] \leq n^{2g}$. On voit donc que dans le cas totalement générique (appelé "type GSp" dans [Hindry et Ratazzi 2012]) tous les points d'ordre n engendrent des corps de degrés comparables $n^{2g} \geq [K(P):K] \gg n^{2g-\varepsilon}$, mais qu'il n'en est plus de même quand $e \geq 2$.

Nous étendons les résultats au cas d'une variété abélienne géométriquement isogène à un produit de variétés abéliennes de type I ou II.

Théorème 1.14. *Soit A/K une variété abélienne isogène sur $\overline{K}$ à un produit $A_1^{n_1} \times \cdots \times A_d^{n_d}$ avec A_i non $\overline{K}$-isogène à A_j pour $i \neq j$. On suppose que chaque facteur A_i est de type I ou II et est pleinement de type Lefschetz. Pour tout sous-ensemble non vide I de $[1,d]$ on note $A_I := \prod_{i\in I} A_i$. On note e_i la dimension du centre de $\mathrm{End}^0(A_i)$ et h_i la dimension relative de A_i. Enfin on pose $d_i = 1$ (resp. $d_i = 2$) si A_i est de type I (resp. de type II). On a alors*

$$\gamma(A) = \max_I \frac{2\sum_{i\in I} n_i d_i h_i e_i}{1+\sum_{i\in I} 2e_i h_i^2 + h_i e_i} = \max_I \frac{2\sum_{i\in I} n_i \dim A_i}{\dim \mathrm{MT}(A_I)}.$$

Corollaire 1.15. *Soit A une variété abélienne géométriquement isogène à un produit $A_1^{n_1} \times \cdots \times A_d^{n_d}$ avec A_i non $\overline{K}$-isogène à A_j pour $i \neq j$. On suppose que chaque facteur A_i est de type I ou II et vérifie l'une quelconque des trois hypothèses suivantes :*

1. *Le nombre h_i est impair ou égal à 2.*
2. *On a $e_i = 1$ (i.e., $E_i = \mathbb{Q}$) et $h_i \notin \Sigma$.*
3. *La variété abélienne A_i est de type I (resp. II) et possède une place de mauvaise réduction semi-stable avec dimension torique e_i (resp. avec dimension torique $2e_i$).*

On a alors, avec les notations du théorème 1.14,

$$\gamma(A) = \max_I \frac{2\sum_{i\in I} n_i d_i h_i e_i}{1+\sum_{i\in I} 2e_i h_i^2 + h_i e_i} = \max_I \frac{2\sum_{i\in I} n_i \dim A_i}{\dim \mathrm{MT}(A_I)}.$$

Remarque 1.16. Dans le contexte du corollaire 1.15, l'analogue du théorème 1.12 dit simplement que

$$[K(P):K] \geq c_1^{\omega(n)} n^{2h_0},$$

où h_0 est le minimum des h_i.

Réductions. Le problème que nous étudions est clairement invariant par deux modifications : remplacer K par une extension *finie* K' et remplacer A par une variété abélienne A' *isogène* à A. Quitte à effectuer une extension finie de K et à remplacer A par une variété isogène, nous pouvons donc supposer, et nous supposerons dans le reste de l'article, les propriétés suivantes vérifiées par A/K :

(1) L'anneau des endomorphismes définis sur K est égal à l'anneau des endomorphismes définis sur $\overline{K}$; on le notera donc $\mathrm{End}(A)$.

(2) L'anneau des endomorphismes $\mathrm{End}(A)$ est un ordre maximal dans $\mathrm{End}(A)\otimes\mathbb{Q}$.

(3) La variété abélienne s'écrit $A = A_1^{n_1} \times \cdots \times A_r^{n_r}$, avec $1 \le r$, $1 \le n_i$ et des variétés abéliennes A_i absolument simples non isogènes deux à deux.

(4) L'adhérence de Zariski de $\rho_{\ell^\infty,A}(G_K)$ est connexe.

(5) Les représentations $\rho_{\ell^\infty,A}$ sont indépendantes sur K.

En effet la possibilité de l'obtention des trois premières propriétés par extension de K et isogénie découle des propriétés générales des variétés abéliennes, tandis que les points (4) et (5) s'obtiennent par une extension adéquate du corps K, en invoquant deux résultats subtils de Serre [1986a; 2013]. Le point (5) est rappelé, ainsi que la définition de “représentations indépendantes” à la proposition 3.2.

Remarquons qu'on pourrait également songer à imposer que A soit principalement polarisée, mais cela forcerait à renoncer à la propriété (2), il nous a semblé plus commode de préserver cette dernière.

Plan. Le plan de ce texte est le suivant : Dans la section suivante on rassemble un certain nombre de lemmes de théorie des groupes et de combinatoire. Dans les sections 3 et 4 on décrit les accouplements λ-adiques déduits de celui de Weil, ainsi que l'étude des propriétés des représentations galoisiennes qui sont utilisées dans la suite. La section 5 étudie la partie cyclotomique de ces représentations. Les sections 6 et 7 contiennent les preuves des deux théorèmes cités en introduction (théorèmes 1.6 et 1.14), d'abord dans le cas d'une variété abélienne simple et pour un groupe annulé par ℓ puis dans le cas général. Les preuves sont en fait écrites pour ℓ assez grand et on indique dans la huitième section comment on peut modifier les preuves pour traiter les “petits” nombres premiers de manière similaire. La courte neuvième section contient la démonstration de la minoration du degré de l'extension engendrée par un point de torsion (théorème 1.12). La dixième et dernière section est un appendice indépendant du reste de l'article (sauf pour les notations), dans lequel nous montrons comment déduire un énoncé du type “conjecture forte de Mumford–Tate” de l'énoncé usuel de la conjecture de Mumford–Tate et expliquons comment prouver les quelques cas supplémentaires de la conjecture de Mumford–Tate ne figurant pas dans la littérature et énoncés dans les corollaires 1.8 et 1.15.

Notations. Dans tout le reste de cet article, nous utiliserons les notations suivantes :

$$[L:K], \quad (G:H), \quad [G,G]$$

pour désigner respectivement, le degré de l'extension de corps L sur K, l'indice du sous-groupe H de G dans G, le groupe des commutateurs de G.

Par ailleurs nous utiliserons les deux notations supplémentaires suivantes : l'*égalité à indice fini près* $\asymp$ et la *presque égalité* $\doteq$. Commençons par définir l'égalité à indice fini près : Si L_1, L_2 sont des corps de nombres contenus dans un corps L (en pratique nos corps seront tous contenus dans $K(A_{\text{tor}})$ ou, si l'on préfère dans $\overline{\mathbb{Q}}$ ou même $\mathbb{C}$) qui dépendent de A/K et d'un autre ensemble de paramètres Λ, nous écrirons $L_1 \asymp L_2$ pour dire qu'il existe une constante C(A/K) ne dépendant que de A/K telle que les inégalités

$$[L_1 : L_1 \cap L_2] \leq C(A/K) \quad \text{et} \quad [L_2 : L_1 \cap L_2] \leq C(A/K)$$

sont vraies pour toutes valeurs des paramètres dans l'ensemble Λ. De même si G_1 et G_2 sont des sous-groupes d'un même groupe et dépendant d'un ensemble de paramètres Λ, nous écrirons $G_1 \asymp G_2$ pour dire qu'il existe une constante C(A/K) ne dépendant que de A/K telle que les inégalités $(G_1 : G_1 \cap G_2) \leq C(A/K)$ et $(G_2 : G_1 \cap G_2) \leq C(A/K)$ sont vraies pour toutes valeurs des paramètres dans l'ensemble Λ. Nous utiliserons la même notation pour des nombres. Ainsi, si N_1, N_2 sont deux nombres (par exemple des cardinaux de groupes ou des degrés d'extensions) $N_1 \asymp N_2$ signifie qu'il existe deux constantes C_1 et C_2, indépendantes des paramètres, telles que $C_1 N_1 \leq N_2 \leq C_2 N_1$ (voir, par exemple, le lemme 6.4 pour une utilisation de cette dernière notation).

Concernant la presque égalité, dire que $E \doteq F$ signifie par définition que $E \asymp F$ et que de plus il y a en fait égalité $E = F$ pour toutes les valeurs des paramètres dans Λ sauf éventuellement un nombre fini (dépendant éventuellement de A/K).

Un exemple typique d'utilisation de la notation précédente consiste par exemple (cf. proposition 5.2) à écrire, pour A/K une variété abélienne simple de type I ou II de dimension relative h et pleinement de type Lefschetz, que l'on a

$$[\rho_\ell(G_K), \rho_\ell(G_K)] \doteq \mathrm{Sp}_{2h}(\mathcal{O}_E/\ell\mathcal{O}_E),$$

ceci signifiant qu'il existe une constante $\ell_0(A/K)$ dépendant de A/K telle qu'il a égalité pour tout $\ell \geq \ell_0(A/K)$ et que de plus pour tout ℓ les deux groupes sont commensurables, i.e., l'indice de l'intersection des deux groupes dans l'un et l'autre est fini.

Enfin, si L_1, L_2 sont des corps de nombres qui dépendent de A/K et d'un autre ensemble de paramètres Λ, nous écrirons $[L_1 : \mathbb{Q}] \ll [L_2 : \mathbb{Q}]$ pour dire qu'il existe une constante C(A/K) ne dépendant que de A/K telle que l'inégalité $[L_1 : \mathbb{Q}] \leq C(A/K)[L_2 : \mathbb{Q}]$ est vraie uniformément sur Λ.

2. Lemmes de groupes

Soit $E/\mathbb{Q}$ un corps de nombres, d'anneau d'entiers $\mathcal{O}_E$. Si ℓ est un premier et λ une place de $\mathcal{O}_E$ au dessus de ℓ, on note $\mathcal{O}_\lambda$ le complété de $\mathcal{O}_E$ selon λ. De même on note $\mathbb{F}_\lambda$ le corps résiduel correspondant.

Nous rappelons maintenant des objets et notations provenant de [Hindry et Ratazzi 2012] que nous utiliserons ensuite. Soit V un $\mathbb{Q}$-espace vectoriel muni d'une forme symplectique (dans la suite du papier nous utiliserons ceci avec $V = H_1(A(\mathbb{C}), \mathbb{Q})$). Dans les lemmes suivants nous introduisons $(e_1, \ldots, e_{2g})$ une base symplectique de V. (i.e., pour tout $1 \leq i, j \leq g$, $e_i \cdot e_{g+i} = +1$ et $e_i \cdot e_j = 0$ si $|i-j| \neq 0$).

Lemme 2.1. *Soit $0 \leq s \leq r \leq g$ avec $r \geq 1$. Définissons $P_{r,s}$ le sous-groupe algébrique de Sp_{2g} fixant les vecteurs $e_1, \ldots, e_r$ et les vecteurs $e_{g+1}, \ldots, e_{g+s}$, c'est-à-dire*

$$P_{r,s} := \bigl\{M \in \mathrm{Sp}_{2g} \mid Me_i = e_i,\ i \in [1, r] \cup [g+1, g+s]\bigr\}.$$

Alors, $P_{r,s}$ est lisse sur $\mathcal{O}_E$ et sa codimension dans Sp_{2g} est :

$$\operatorname{codim} P_{r,s} = 2sg + 2rg - rs - \frac{r(r-1)}{2} - \frac{s(s-1)}{2}.$$

Démonstration. Il s'agit du lemme 2.24 de [Hindry et Ratazzi 2012] (énoncé sur $\mathbb{Z}$ mais valable par la même preuve sur $\mathcal{O}_E$). □

Lemme 2.2. *Soit λ une place de $\mathcal{O}_E$ au dessus d'un premier ℓ de $\mathbb{Z}$. En introduisant les groupes*

$$D_0 := \left\{ \begin{pmatrix} I & 0 \\ 0 & \alpha I \end{pmatrix} \in \mathrm{GL}_{2g}(\mathcal{O}_\lambda) \;\middle|\; \alpha \in \mathbb{Z}_\ell^\times \right\},$$

$$G_\lambda := \bigl\{M \in \mathrm{GSp}_{2g}(\mathcal{O}_\lambda) \mid \mathrm{mult}(M) \in \mathbb{Z}_\ell^\times\bigr\},$$

on a

$$D_0 \cdot \mathrm{Sp}_{2g}(\mathcal{O}_\lambda) = G_\lambda.$$

Le même énoncé vaut en remplaçant $\mathcal{O}_\lambda$ par $\mathbb{F}_\lambda$ et $\mathbb{Z}_\ell$ par $\mathbb{F}_\ell$.

Démonstration. Il s'agit du lemme 2.12 de [Hindry et Ratazzi 2012] dans sa version λ-adique : soit $M \in G_\lambda$ de multiplicateur $\mathrm{mult}(M)$. La matrice

$$\begin{pmatrix} I_g & 0 \\ 0 & \mathrm{mult}(M) I_g \end{pmatrix} M$$

est dans $\mathrm{Sp}_{2g}(\mathcal{O}_\lambda)$. □

Lemme 2.3. *Soit G_E un sous-groupe algébrique sur E de GL_E. Soit $t \in \mathbb{N}^*$ et soit $\mathcal{G}_{1,E}, \ldots, \mathcal{G}_{t,E}$ une suite de sous-groupes algébriques de G_E. On note G (respectivement $\mathcal{G}_i$) l'adhérence de Zariski de G_E (respectivement de $\mathcal{G}_{i,E}$) dans $\mathrm{GL}_{\mathcal{O}_E}$ sur $\mathcal{O}_E$. Il existe des constantes $C_1 > 0$, $C_2 > 0$ telles que la propriété* *suivante est vraie : soient ℓ un premier de $\mathbb{Z}$ et λ une place de $\mathcal{O}_E$ au-dessus de ℓ. Soient $G_1 \subset G_2 \subset \cdots \subset G_t$ une suite de sous-groupes algébriques sur $\mathcal{O}_\lambda$ de $G_{\mathcal{O}_\lambda}$. On suppose que pour tout i, le groupe G_i est conjugué sur $\mathbb{F}_\lambda$ à $\mathcal{G}_i$. On note $g_i := \dim \mathcal{G}_i = \dim G_i$ et $d_i := \operatorname{codim}_G \mathcal{G}_i = \operatorname{codim}_{G_{\mathcal{O}_\lambda}} G_i$ et on pose, pour toute suite croissante d'entiers $0 = m_0 < m_1 < m_2 < \cdots < m_t$:*

$$H(m_1, \ldots, m_t) = \{M \in G(\mathcal{O}_\lambda) \mid M \in G_i \mod \lambda^{m_i}\}.$$

Pour tous les ℓ tels que G et les $\mathcal{G}_i$ sont lisses sur $\mathbb{F}_\lambda$, on a alors

$$C_1 \times \big(G(\mathcal{O}_\lambda) : H(m_1, \ldots, m_t)\big) \geq \operatorname{Card}(\mathbb{F}_\lambda)^{\sum_{i=1}^t d_i(m_i - m_{i-1})} \geq C_2 \times \big(G(\mathcal{O}_\lambda) : H(m_1, \ldots, m_t)\big).$$

Démonstration. Il s'agit du lemme 2.4 de [Hindry et Ratazzi 2012] dont la preuve reste valable, en remplaçant $\mathbb{Z}$ par $\mathcal{O}_E$, ainsi que ℓ par λ et $\mathbb{F}_\ell$ par $\mathbb{F}_\lambda$. □

Nous donnons dans ce qui suit l'analogue d'un lemme prouvé pour $\mathrm{SL}_2(\mathbb{Z}_\ell)$ par Serre. L'énoncé en vue est le suivant, où les notations suivantes sont utilisées.

Définition 2.4. Nous dirons qu'une sous-algèbre de Lie de gl_m *possède la propriété* $(\mathcal{CN})$ (des carrés nuls) si elle est engendrée, comme espace vectoriel, par des matrices de carré nul.

Exemple 2.5. Les algèbres $\mathfrak{sl}_m$ et $\mathfrak{sp}_{2m}$ ont la propriété $(\mathcal{CN})$ mais pas $\mathfrak{so}_m$. Les matrices E_{ij} ayant un seul coefficient non nul (sur la i-ème ligne et j-ème colonne) sont de carré nul. En dimension 2, on a de même que $\left(\begin{smallmatrix} a & 1 \\ -a^2 & -a \end{smallmatrix}\right)$ est de carré nul, donc

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{pmatrix} a & 1 \\ -a^2 & -a \end{pmatrix} + \begin{pmatrix} 0 & b-1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c+a^2 & 0 \end{pmatrix}$$

est bien somme de matrices de carré nul. Quand m est quelconque, on en tire aisément que les matrices diagonales de trace nulle sont sommes de matrices de carré nul. L'algèbre $\mathfrak{sp}_{2m}$ est l'algèbre des matrices $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ telles que $D + {}^tA = 0$, B et C sont symétriques. On écrit aisément une telle matrice comme somme de matrices :

$$\begin{pmatrix} 0 & S \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ S & 0 \end{pmatrix}, \quad \begin{pmatrix} \delta & \delta \\ -\delta & -\delta \end{pmatrix}, \quad \begin{pmatrix} U & 0 \\ 0 & -{}^tU \end{pmatrix},$$

où S est symétrique, U a trace nulle et δ est la matrice dont le seul terme non nul est dans le coin supérieur gauche et vaut 1 ; les trois premières sont de carré nul et la quatrième est somme de matrices de carrés nuls d'après la propriété pour $\mathfrak{sl}_m$. Enfin notons qu'une matrice anti-symétrique de carré nul est elle-même nulle, donc $\mathfrak{so}_n$ ne possède pas la propriété des carrés nuls.

Notations. On note $\mathcal{O}$ l'anneau d'entiers d'un corps p-adique, ϖ une uniformisante, $\mathbb{F} = \mathcal{O}/\varpi\mathcal{O}$ le corps résiduel et e l'indice de ramification, i.e., $p = \varpi^e u$ avec $u \in \mathcal{O}^\times$.

Lemme 2.6. *Soit G un sous-groupe algébrique lisse de $\mathrm{GL}_m/\mathcal{O}$ et H un sous-groupe fermé de $G(\mathcal{O})$. Considérons, pour tout entier $n \geq 1$, les applications $\pi_n : H \to \mathrm{GL}_m(\mathcal{O}/\varpi^n\mathcal{O})$. Alors on a :*

(1) *Si $\pi_{e+1}(H) = G(\mathcal{O}/\varpi^{e+1}\mathcal{O})$ et $p \geq e+2$, alors $H = G(\mathcal{O})$; si $p \leq e+1$ et π_m surjective avec $m \geq (ep+1)/(p-1)$, alors $H = G(\mathcal{O})$.*

(2) *Si $p \geq 2e+3$, $\pi_e(H) = G(\mathcal{O}/\varpi)$ et $Lie(G_{\mathbb{F}})$ a la propriété $(\mathcal{CN})$ alors $H = G(\mathcal{O})$.*

En particulier, lorsque $G = \mathrm{SL}_m$ ou Sp_m, si $p \geq 5$ et $K/\mathbb{Q}_p$ non ramifié, alors $\pi_1(H) = G(\mathbb{F})$ entraîne $H = G(\mathcal{O})$.

Démonstration. Notons $\mathcal{L} := \mathrm{Lie}(G_{\mathbb{F}})$. Commençons par observer que, d'après l'hypothèse de lissité, si une matrice s'écrit $M = I + \varpi^n B$ alors $\pi_{n+1}(M) \in G(\mathcal{O}/\varpi^{n+1}\mathcal{O})$ si et seulement si $\pi_1(B) \in \mathcal{L}$. On prouve maintenant par récurrence sur n que la projection $H \to G(\mathcal{O}/\varpi^n\mathcal{O})$ est surjective et donc que H est dense dans $G(\mathcal{O})$ et donc égal à ce dernier. Supposons la propriété vraie au cran n et montrons-la au cran $n+1$. Soit donc $A \in G(\mathcal{O})$, on sait donc qu'il existe $A_1 \in H$ telle que $A \equiv A_1[\varpi^n]$ et, quitte à remplacer A par AA_1^{-1} on peut supposer que $A \equiv I[\varpi^n]$ soit encore $A = I + \varpi^n B$ avec donc $\pi_1(B) \in \mathcal{L}$. Par hypothèse, il existe $Z \in H$ telle que $Z \equiv I + \varpi^{n-e}B[\varpi^{n-e+1}]$ ou encore $Z = I + \varpi^{n-e}B + \varpi^{n-e+1}C$. Posons $Y := Z^p$ alors

$$Y = I + p\varpi^{n-e}(B+\varpi C) + \sum_{h=2}^{p-1}\binom{p}{h}\varpi^{h(n-e)}(B+\varpi C)^h + \varpi^{p(n-e)}(B+\varpi C)^p.$$

On a $e + h(n-e) \geq n+1$ si $n \geq e+1$ et $p(n-e) \geq n+1$ si $n \geq (ep+1)/(p-1)$; ainsi si $p \geq e+2$ on voit que $Y \equiv I + \varpi^n uB \mod \varpi^{n+1}$. On peut bien sûr refaire ce calcul en remplaçant B par $u^{-1}B$ et conclure. La surjectivité de π_{e+1} suffit donc pour entraîner $H = G(\mathcal{O})$ si $p \geq e+2$ (resp. la surjectivité de π_m avec $m \geq (ep+1)/(p-1)$ pour $p \leq e$).

Si maintenant $n = e$, on reprend le calcul en supposant d'abord que $\pi_1(B^2) = 0 \in \mathcal{L}$. Observons que, dans ce cas, $B^{2j} \equiv 0 \mod \varpi^j$ donc

$$\begin{aligned}&(B+\varpi C)^p\\ &= B^p + \varpi(B^{p-1}C + \cdots) + \cdots + \varpi^r(\cdots + B^{j_0}CB^{j_1}C\cdots CB^{j_s}) + \cdots + \varpi^p C^p\end{aligned}$$

avec $B^p \equiv 0 \mod \varpi^{(p-1)/2}$ et $\varpi^r B^{j_0}CB^{j_1}C\ldots CB^{j_s} \equiv 0 \mod \varpi^m$ avec

$$m \geq r + \left[\frac{j_0}{2}\right] + \cdots + \left[\frac{j_s}{2}\right] \geq r + \sum_{i=0}^{s}\left(\frac{j_i}{2} - \frac{1}{2}\right) = r + \frac{p-r}{2} - \frac{s+1}{2} \geq \frac{p-1}{2}.$$

Si maintenant $Z = I + B + \varpi C$ et $Y = Z^p$ alors

$$\begin{aligned} Y &= I + p(B+\varpi C) + \sum_{h=2}^{p-1} \binom{p}{h} \varpi^h (B+\varpi C)^{p-h} + (B+\varpi C)^p \\ &\equiv I + pB + (B+\varpi C)^p \mod \varpi^{e+1}. \end{aligned}$$

La condition $(p-1)/2 \geq e+1$ équivaut à $p \geq 2e+3$. On trouve ainsi un élément $Y \in H$ tel que $Y = I + \varpi^e B \mod \varpi^{e+1}$. Pour le cas général, on aura $B = B_1 + \cdots + B_s$ avec $\pi_1(B_i^2) = 0 \in \mathcal{L}$ et $I + \varpi^e B \equiv (I + \varpi^e B_1) \cdots (I + \varpi^e B_s) \mod \varpi^{e+1}$, ce qui permet de conclure. □

Lemme 2.7 [Hindry et Ratazzi 2012, lemme 2.8]. *Soit $d \geq 1$ un entier, et pour tout $i \in \{1, \dots, d\}$, soient $t_i \geq 1$ des entiers. Pour $i \leq d$ et $j \leq t_i$, on se donne également des entiers a_{ij} et b_{ij}, strictement positifs. On a l'égalité*

$$\sup_{\substack{m_{i1} \geq \cdots \geq m_{it_i} \\ 1 \leq i \leq d}} \left\{ \frac{\sum_{i=1}^d \sum_{j=1}^{t_i} a_{ij} m_{ij}}{\sum_{i=1}^d \sum_{j=1}^{t_i} b_{ij} m_{ij}} \right\} = \max_{\substack{1 \leq h_i \leq t_i \\ 1 \leq i \leq d}} \left\{ \frac{\sum_{i=1}^d \sum_{j=1}^{h_i} a_{ij}}{\sum_{i=1}^d \sum_{j=1}^{h_i} b_{ij}} \right\}, \tag{2}$$

le sup dans le membre de gauche étant pris sur les entiers m_{ij} ordonnés pour $1 \leq i \leq d$ par $m_{i1} \geq \cdots \geq m_{it_i}$ et tels que $m_{i1} \neq 0$.

3. Représentations et accouplements λ-adiques

L'étude de la représentation galoisienne adélique se ramène essentiellement à l'étude des représentations ℓ-adiques, grâce au résultat suivant dû à Serre.

Définition 3.1. Une famille de représentations $(\rho_i)_{i \in I} : G_K \to \prod_{i \in I} \mathrm{GL}(V_i)$ indexée par un ensemble I, est dite *indépendante* si

$$(\rho_i)_{i \in I}(G_K) = \prod_{i \in I} (\rho_i(G_K)).$$

Les $(\rho_i)_{i \in I}$ sont *presque indépendantes* s'il existe une extension finie K'/K telles que les restrictions à $G_{K'}$ sont indépendantes.

Proposition 3.2 [Serre 1986c]. *Les représentations ℓ-adiques associées à une variété abélienne A sur un corps de nombres K sont presque indépendantes.*

Démonstration. C'est le théorème 1 de [Serre 1986c], cf. également [Serre 2013, théorème 1 et paragraphe 3.1]. □

Cet énoncé se traduit concrètement en disant que, quitte à remplacer K par une extension finie, pour tous premiers ℓ_1, ℓ_2 distincts, les extensions $K(A[\ell_1^\infty])/K$ et $K(A[\ell_2^\infty])/K$ sont linéairement disjointes.

Nous utiliserons en parallèle les décompositions ℓ-adiques et λ-adiques correspondant aux types I et II ; ces décompositions s'écrivent pour tout ℓ au niveau des

$\mathbb{Q}_\ell$-représentations $V_\ell(A)$ et pour ℓ assez grand (hors d'un ensemble fini de ℓ) pour les $\mathbb{Z}_\ell$-représentations $T_\ell(A)$.

Modules de Tate, représentations ℓ-adiques et λ-adiques. Soit ℓ un premier quelconque. On considère, dans toute la suite de cette section 3, une variété abélienne A/K géométriquement simple de type I ou II, telle que $\mathrm{End}_K(A) = \mathrm{End}_{\overline{K}}(A)$ et telle que $\mathrm{End}_K(A)$ est un ordre maximal de $D := \mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. Nous noterons E le centre de D.

On a la décomposition suivante : $\mathcal{O}_\ell = \prod_{\lambda|\ell} \mathcal{O}_\lambda$, où $\mathcal{O}_\lambda$ est l'anneau des entiers du complété E_λ de E pour la place λ. En notant $f(\lambda) := [E_\lambda : \mathbb{Q}_\ell]$ et $e(\lambda)$ le degré de ramification de $\lambda|\ell$, on a $\sum_{\lambda|\ell} e(\lambda) f(\lambda) = e = [E : \mathbb{Q}]$.

Le module de Tate ℓ-adique, $T_\ell(A)$, est muni d'une action de $\mathcal{O}_\ell$ et se décompose en

$$T_\ell(A) = \prod_{\lambda|\ell} \mathcal{T}_\lambda, \quad \text{où } \mathcal{T}_\lambda := T_\ell(A) \otimes_{\mathcal{O}_\ell} \mathcal{O}_\lambda.$$

En inversant ℓ on obtient les espaces $V_\ell(A)$ et V_λ à partir de $T_\ell(A)$ et de $\mathcal{T}_\lambda$:

$$V_\ell(A) := T_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell \quad \text{et} \quad V_\lambda := \mathcal{T}_\lambda \otimes_{\mathcal{O}_\lambda} E_\lambda.$$

La représentation ℓ-adique étant $\mathcal{O}_E$-linéaire, elle se décompose diagonalement selon les λ divisant ℓ en représentations λ-adiques (cf. [Ribet 1976, paragraphe II]) :

$$\rho_{\ell^\infty} = \left(\prod_{\lambda|\ell} \rho_{\lambda^\infty}\right) : G_K \to \prod_{\lambda|\ell} \mathrm{Aut}(\mathcal{T}_\lambda).$$

La représentation ℓ-adique modulo ℓ étant $\mathcal{O}_E/\ell\mathcal{O}_E$-linéaire, elle se décompose par réduction modulo ℓ diagonalement selon les λ en représentations λ-adiques :

$$\rho_\ell = \left(\prod_{\lambda|\ell} \rho_\lambda\right) : G_K \to \mathrm{Aut}_{\mathcal{O}_\ell/\ell\mathcal{O}_\ell}(A[\ell]) = \prod_{\lambda|\ell} \mathrm{Aut}_{\mathcal{O}_\lambda/\ell\mathcal{O}_\lambda}(\mathcal{T}_\lambda/\ell\mathcal{T}_\lambda).$$

Nous noterons dans toute la suite G_λ le groupe de Galois correspondant à ρ_λ (il est a priori à valeurs dans $\mathrm{GL}_{2h}(\mathcal{O}_\lambda/\ell\mathcal{O}_\lambda)$). La même chose vaut au niveau ℓ-adique et on sait (cf., par exemple, [Chi 1992, p. 319]) que ces représentations λ-adiques sont munies naturellement d'un accouplement de Weil λ-adique provenant de l'accouplement ℓ-adique. Nous rappelons ceci dans le paragraphe suivant.

Accouplements ℓ-adique et λ-adique. Nous supposons ici de plus que A est polarisée par une polarisation ϕ et que ℓ est un premier ne divisant pas $\deg(\phi)$.

Rappelons la construction de l'accouplement λ-adique (cf. [Banaszak et al. 2006, paragraphes 3 et 4], par exemple). On commence pour cela par l'accouplement de Weil ℓ-adique usuel :

$$\phi_{\ell^\infty} : T_\ell(A) \times T_\ell(A) \to \mathbb{Z}_\ell(1) = \varprojlim \mu_{\ell^m}.$$

L'accouplement usuel de Weil est non dégénéré (modulo ℓ^n pour tout $n \geq 1$) car ℓ ne divise pas $\deg(\phi)$. De plus, si † désigne l'involution de Rosati sur $\mathrm{End}^0(A)$ associée à la polarisation définissant l'accouplement on aura $\phi_{\ell^\infty}(ax, y) = \phi_{\ell^\infty}(x, a^\dagger y)$ pour $x, y \in T_\ell(A)$ et $a \in \mathcal{O}_\ell$.

Lemme 3.3. *Notons $\mathcal{O}_\ell^\star$ le dual de $\mathcal{O}_\ell$ pour la dualité donnée par la trace $\mathrm{Tr}_{E_\ell/\mathbb{Q}_\ell}$. Il existe un unique accouplement $\mathcal{O}_\ell$-linéaire, $\phi_{\ell^\infty}^\star : T_\ell(A) \times T_\ell(A) \to \mathcal{O}_\ell^\star(1)$, tel que*

$$\mathrm{Tr}_{E_\ell/\mathbb{Q}_\ell}(\phi_{\ell^\infty}^\star) = \phi_{\ell^\infty}.$$

Démonstration. Il s'agit essentiellement du lemme 3.1 de [Banaszak et al. 2006] (cf. aussi [Deligne et al. 1982, Sublemma 4.7, p. 55]). Le preuve est la suivante : il s'agit de vérifier que le morphisme

$$\mathrm{Hom}_{\mathcal{O}_\ell}\big(T_\ell(A) \otimes_{\mathcal{O}_\ell} T_\ell(A), \mathcal{O}_\ell^\star\big) \to \mathrm{Hom}_{\mathbb{Z}_\ell}\big(T_\ell(A) \otimes_{\mathcal{O}_\ell} T_\ell(A), \mathbb{Z}_\ell\big),$$

donné par la trace, est un isomorphisme. Or ces deux objets sont des $\mathbb{Z}_\ell$-modules libres de même rang et la preuve du lemme 3.1 de [Banaszak et al. 2006] donne la surjectivité. □

Hypothèse. On suppose dans la fin de ce paragraphe que ℓ est de plus non ramifié dans $E/\mathbb{Q}$, i.e., que $\mathcal{O}_\ell^\star = \mathcal{O}_\ell$.

On a dans ce cas l'accouplement $\mathcal{O}_\ell$-linéaire $\phi_{\ell^\infty}^\star : T_\ell(A) \times T_\ell(A) \to \mathcal{O}_\ell(1)$. Par projection on construit alors l'accouplement λ-adique, $\mathcal{O}_\lambda$-linéaire, de la façon suivante :

$$\phi_{\lambda^\infty} : \mathcal{T}_\lambda \times \mathcal{T}_\lambda \to \mathcal{O}_\lambda(1)$$

tel que $\phi_{\ell^\infty}^\star(x, y) \otimes 1 = \phi_{\lambda^\infty}(x \otimes 1, y \otimes 1)$.

Tout comme l'accouplement ℓ-adique, l'accouplement λ-adique est Galois équivariant :

$$\forall \sigma \in G_K, \forall x, y \in \mathcal{T}_\lambda, \quad \phi_{\lambda^\infty}(\sigma x, \sigma y) = \phi_{\lambda^\infty}(x, y)^\sigma,$$

l'action de Galois à gauche se faisant via la représentation λ-adique, et à droite, via le caractère cyclotomique ℓ-adique usuel.

Nous noterons enfin $\phi_{\ell^\infty}^0$ et $\phi_{\lambda^\infty}^0$ les accouplements ℓ-adiques et λ-adiques, définis de manière similaires sur $V_\ell(A)$ et V_λ, à valeurs respectivement dans E_ℓ et E_λ. Ils sont définis sans restriction pour tout ℓ.

Galois pour les variétés de type I et II. Soit ℓ un premier quelconque.

Dans le cas de type I, les $(V_\lambda, \phi_{\lambda^\infty}^0)$ fournissent des représentations irréductibles symplectiques. Dans le cas de type II, on a une décomposition plus fine (cf. [Chi 1990; Milne 1999; Banaszak et al. 2006]) :

$$V_\lambda = W_\lambda(A) \oplus W_\lambda(A),$$

où $(W_\lambda(A), \phi^0_{\lambda\infty|W_\lambda(A)})$ est maintenant irréductible symplectique. Cependant, comme nous le détaillons plus loin, cette décomposition, dans le cas de type II, nécessite, pour le nombre fini de premiers ℓ "ramifiés" pour l'algèbre de quaternions, d'étendre les scalaires à une extension quadratique.

Nous avons $V_\lambda := W_\lambda(A) \oplus W_\lambda(A)$ si A est de type II et nous posons $W_\lambda(A) := V_\lambda$ si A est de type I. Autrement dit en notant $d = 1$ si A est de type I et $d = 2$ si A est de type II, nous aurons toujours

$$V_\lambda = W_\lambda(A)^d,$$

et dans tous les cas, la représentation $W_\lambda(A)$ est irréductible, symplectique et, lorsque A est de type I (resp. de type II), le module $V_\ell(A)$ est isomorphe à $\prod_\lambda W_\lambda(A)$ (resp. à la somme de deux copies de ce produit). On pose ensuite dans tout les cas

$$T_\lambda(A) := \mathcal{T}_\lambda \cap W_\lambda(A).$$

Définition 3.4. Nous noterons dans la suite $S_{\text{ex}}(A)$ l'ensemble fini des ℓ divisant le degré de la polarisation fixée ϕ de A, des ℓ ramifiés dans $\mathcal{O}_E$ et, dans le cas de type II, des ℓ tels que l'algèbre de quaternions D est non décomposée en au moins un $\lambda|\ell$.

Au niveau des $\mathbb{Z}_\ell$-modules, la décomposition perdure, au moins pour ℓ hors de $S_{\text{ex}}(A)$. Nous rassemblons ces énoncés dans la proposition suivante.

Proposition 3.5. *Soit $W_\lambda(A)$ le E_λ-module galoisien symplectique défini ci-dessus, qu'on identifie à un sous-module de $V_\ell(A)$.*

(1) *La représentation $W_\lambda(A)$ est irréductible et symplectique.*

(2) *Si A est de type I, on a une décomposition $V_\ell(A) = \prod_{\lambda|\ell} W_\lambda(A)$, et si A est de type II, $V_\ell(A) = \prod_{\lambda|\ell}(W_\lambda(A) \oplus W_\lambda(A))$. Toutefois, dans le cas de type II et pour les premiers ramifiés de l'algèbre de quaternions, cette décomposition ne s'obtient qu'après tensorisation par une extension quadratique de E.*

(3) *Pour $\ell \notin S_{\text{ex}}(A)$, on a une décomposition analogue pour les $\mathcal{O} \otimes \mathbb{Z}_\ell$-modules : $T_\ell(A) = \prod_{\lambda|\ell} T_\lambda(A)$ si A est de type I, et $T_\ell(A) = \prod_\lambda (T_\lambda(A) \oplus T_\lambda(A))$ si A est de type II.*

Démonstration. Pour A de type I, voir [Ribet 1976]. Pour A de type II, cet énoncé est prouvé pour ℓ assez grand sur $\mathbb{Q}_\ell$ dans [Chi 1990] et sur $\mathbb{Z}_\ell$ dans [Banaszak et al. 2006], le point essentiel étant l'identification de $\text{End}^0(A) \otimes \mathbb{Q}_\ell$ à un produit d'algèbres de matrices $\prod_\lambda M_2(E_\lambda)$, ce qui est possible justement quand ℓ n'est pas ramifié. L'algèbre de quaternions est toujours déployée sur une extension quadratique et l'on peut donc se ramener au cas précédent après tensorisation par une telle extension ; plus précisément, en choisissant F extension quadratique de E telle que $D \otimes_E F \cong M_2(F)$ et en notant $F_\lambda := E_\lambda \otimes_E F$ et $V_\ell \otimes_E F = \prod_\lambda V_{\lambda,F}$, chaque $V_{\lambda,F}$ est un module sur $D\otimes_E F_\lambda \cong M_2(F_\lambda)$ et on peut alors imiter le procédé décrit dans [loc. cit.]. Remarquons que la preuve sur $\mathbb{Q}_\ell$ pour tout ℓ peut aussi être extraite de [Milne 1999]. □

Définition 3.6. Définissons la *dimension relative* de A simple avec $D := \mathrm{End}^0(A)$ et E le centre de D, par la formule :

$$\mathrm{dimrel}(A) := \frac{\dim A}{[E:\mathbb{Q}]\sqrt{[D:E]}}.$$

Ainsi, si $e = [E:\mathbb{Q}]$, la dimension relative h d'une variété abélienne de dimension g de type I (resp. de type II) est $h = g/e$ (resp. $h = g/(2e)$).

On a alors dans le cas I et II une représentation irréductible symplectique de dimension $2h$, et une inclusion

$$H_{\ell,A} \subset \prod_{\lambda|\ell} \mathrm{Sp}(W_\lambda(A), \phi^0_{\lambda\infty}) \cong \prod_{\lambda|\ell} \mathrm{Sp}_{2h,E_\lambda},$$

qu'il convient de mettre en parallèle avec l'inclusion

$$\mathrm{Hdg}(A)_{\mathbb{Q}_\ell} \subset (\mathrm{Res}_{E/\mathbb{Q}}\, \mathrm{Sp}_{2h,E})_{\mathbb{Q}_\ell}.$$

Nous allons nous placer dans le cas générique où l'on a égalité dans les deux inclusions précédentes (la première impliquant d'ailleurs la seconde puisque $H_{\ell,A} \subset \mathrm{Hdg}(A)_{\mathbb{Q}_\ell}$). Le théorème suivant précise des conditions où l'on sait que l'égalité voulue est toujours vraie.

Théorème 3.7 [Banaszak et al. 2006; Pink 1998; Hall 2011]. *Soit A/K une variété abélienne de type I ou II, de dimension relative h ; notons E le centre de* $\mathrm{End}^0(A)$. *Supposons de plus l'une des trois conditions suivantes réalisée :*

(1) *L'entier h est impair ou égal à* 2.

(2) *On a $E = \mathbb{Q}$ et h n'appartient pas à l'ensemble exceptionnel Σ défini par l'équation (1).*

(3) *La variété abélienne A est de type I (resp. II) et possède une place de mauvaise réduction semi-stable avec dimension torique e (resp. avec dimension torique $2e$).*

On a alors $H_{\ell,A} = \prod_\lambda \mathrm{Sp}_{2h,E_\lambda}$.

Démonstration. Le résultat est démontré dans [Banaszak et al. 2006] sous l'hypothèse de l'alinéa (1) (cf. [Lombardo 2016, Remark 2.25] pour le cas $h = 2$) ; le résultat est démontré explicitement dans [Pink 1998] sous l'hypothèse de l'alinéa (2), pour une variété abélienne de type I, mais on peut extraire de [Pink 1998] une preuve pour le type II ; nous indiquons comment en appendice de cet article (cf. section 10). Le résultat est démontré dans [Hall 2011] sous l'hypothèse de l'alinéa (3), pour une variété abélienne de type I vérifiant $e = 1$; nous indiquons en appendice de cet article comment étendre les arguments de [Hall 2011] aux cas énoncés. □

Variantes λ-adiques modulo λ^n. On suppose ici que ℓ est un premier quelconque. Soit $n \geq 1$ un entier. On a par réduction modulo ℓ^n,

$$A[\ell^n] = T_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Z}/\ell^n\mathbb{Z} = T_\ell(A)/\ell^n T_\ell(A).$$

De même par réduction modulo λ^n, on voit que

$$A[\lambda^n] := \mathcal{T}_\lambda \otimes_{\mathcal{O}_\lambda} \mathcal{O}_\lambda/\lambda^n = \mathcal{T}_\lambda/\lambda^n \mathcal{T}_\lambda.$$

On pose ensuite

$$T_\lambda[\lambda^n] := T_\lambda(A) \otimes_{\mathcal{O}_\lambda} \mathcal{O}_\lambda/\lambda^n = T_\lambda(A)/\lambda^n T_\lambda(A).$$

En utilisant que $\ell^n \mathcal{O}_\lambda = \lambda^{e(\lambda)n}$ on voit que par réduction modulo ℓ^n on obtient :

$$A[\ell^n] = T_\ell(A)/\ell^n T_\ell(A) = \prod_{\lambda|\ell} \mathcal{T}_\lambda/\ell^n \mathcal{T}_\lambda,$$

et

$$T_\lambda(A)/\ell^n T_\lambda(A) = T_\lambda(A) \otimes_{\mathcal{O}_\lambda} \mathcal{O}_\lambda/\lambda^{e(\lambda)n} = T_\lambda[\lambda^{e(\lambda)n}].$$

Soit π_λ une uniformisante de λ dans O_λ ($\pi_\lambda = \ell$ dans le cas non ramifié). Pour tout entier $n \geq 0$, les applications

$$i_n : T_\lambda[\lambda^n] \to T_\lambda[\lambda^{n+1}], \quad x \bmod \lambda^n T_\lambda(A) \mapsto \pi_\lambda x \bmod \lambda^{n+1} T_\lambda(A)$$

sont des morphismes, bien définis, injectifs. En prenant le système inductif qu'ils forment, on note $T_\lambda[\lambda^\infty]$ la limite.

Supposons dans la fin de ce paragraphe que ℓ est non ramifié. Dans ce cas le $\mathcal{O}_\ell/\ell^n\mathcal{O}_\ell$-module $A[\ell^n]$ se décompose en le produit des $\mathcal{O}_\lambda/\lambda^n$-modules $T_\lambda[\lambda^n]$ dans le cas de type I, et, pour $\ell \notin S_{\text{ex}}(A)$, en le produit des $T_\lambda[\lambda^n] \oplus T_\lambda[\lambda^n]$ dans le cas de type II. De plus, par projection modulo ℓ^n (ou ce qui revient au même ici, modulo λ^n), on obtient :

$$\phi_{\lambda^n} : T_\lambda[\lambda^n] \times T_\lambda[\lambda^n] \to \mathcal{O}_\lambda/\ell^n\mathcal{O}_\lambda(1), \quad \text{qui vérifie} \quad \phi_{\lambda^n}(\ell x, \ell y) = \phi_{\lambda^{n+1}}(x, y)^\ell.$$

Tout comme l'accouplement λ-adique, on voit par réduction modulo ℓ^n que les accouplements ϕ_{λ^n} sont Galois équivariants, l'action de Galois à gauche se faisant via la représentation λ-adique, et à droite, via le caractère cyclotomique ℓ-adique usuel.

4. Modules isotropes

On considère dans ce paragraphe une variété abélienne A/K géométriquement simple de type I ou II, telle que $\mathrm{End}_K(A) = \mathrm{End}_{\overline{K}}(A)$ et telle que $\mathrm{End}_K(A)$ est un ordre maximal de $\mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. Soit par ailleurs ℓ un premier et λ une place au dessus de ℓ. Notons π_λ une uniformisante de λ. On suppose ici que ℓ est tel que la condition suivante (qui exclut un nombre fini de premiers) est réalisée : on a un accouplement bilinéaire alterné, non dégénéré sur $T_\lambda(A)$ (qui est un $\mathcal{O}_\lambda$-module libre de rang $2h$) et sur le $\mathbb{F}_\lambda$-espace vectoriel $T_\lambda[\lambda] = T_\lambda(A)/\lambda T_\lambda(A)$.

Définition 4.1. Soit $H \subset T_\lambda[\lambda^\infty]$ un sous-groupe fini. Nous dirons que H *est totalement isotrope* si pour tous points P, Q de $H \subset T_\lambda[\lambda^n]$, on a

$$\phi_{\lambda^n}(P, Q) = 1,$$

où ϕ_{λ^n} désigne l'accouplement sur $T_\lambda[\lambda^n]$.

Notons que si H est totalement isotrope au sens précédent, alors son sous-groupe des points de λ-torsion est totalement isotrope dans le $\mathbb{F}_\lambda$-espace vectoriel $T_\lambda[\lambda]$. On retrouve avec cette définition les deux lemmes suivants dont les preuves se reprennent mot pour mot du paragraphe 3.1 de [Hindry et Ratazzi 2012] en remplaçant $\mathbb{Z}$ par $\mathcal{O}_\lambda$ et ℓ par λ.

Lemme 4.2. *Soit $(e_1, \dots, e_h)$ une base d'un sous-$\mathcal{O}_\lambda$-module isotrope maximal H_∞ de $T_\lambda(A)$. Il existe un supplémentaire H'_∞ isotrope maximal et une base $(e_{h+1}, \dots, e_{2h})$ de celui-ci de sorte que dans la décomposition $T_\lambda(A) = H_\infty \oplus H'_\infty$ selon la base $(e_1, \dots, e_{2h})$, la forme symplectique s'écrit comme la forme canonique*

$$J = \begin{pmatrix} 0 & I_h \\ -I_h & 0 \end{pmatrix}.$$

Lemme 4.3. *Soit $n \geq 1$ et $H \subset T_\lambda[\lambda^n]$ un sous-groupe fini, totalement isotrope. Notons $\mathrm{pr}_n : T_\lambda(A) \to T_\lambda[\lambda^n]$ la projection canonique modulo λ^n. Il existe un sous-groupe totalement isotrope H_{ti} de $T_\lambda[\lambda^n]$, contenant H et de même exposant et il existe un sous-$\mathcal{O}_\lambda$-module, H_∞ de $T_\lambda(A)$, totalement isotrope, tel que $\mathrm{pr}_n(H_\infty) = H_{\mathrm{ti}}$.*

Remarque 4.4. Notons que dans [Hindry et Ratazzi 2012] la version correspondante du lemme précédent ne mentionne pas que l'on peut choisir H_{ti} de même exposant que H. Toutefois la construction même de ce H_{ti} fournie dans la preuve du lemme 3.7 de [Hindry et Ratazzi 2012] donne immédiatement cette information supplémentaire.

5. Propriété μ, version λ-adique

On considère dans ce paragraphe une variété abélienne A/K géométriquement simple de type I ou II, telle que $\mathrm{End}_K(A) = \mathrm{End}_{\overline{K}}(A)$ et telle que $\mathrm{End}_K(A)$ est un ordre maximal de $\mathrm{End}_K(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. On suppose par ailleurs ici que $\ell \notin S_{\mathrm{ex}}(A)$.

Propriété μ. Étant donné un sous-groupe H fini de $T_\lambda[\lambda^\infty]$, nous introduisons à présent l'invariant suivant :

$$m_1(H) := \max\left\{k \in \mathbb{N} \mid \exists n \geq 0,\ \exists P, Q \in H \text{ d'ordre } \ell^n,\ \phi_{\lambda^n}(P, Q) \text{ est d'ordre } \ell^k\right\}.$$

Dire que H est totalement isotrope équivaut à dire que $m_1(H) = 0$. De plus on peut noter que, sur la définition, il est évident que $m_1(H)$ est supérieur à la valeur m suivante :

$$m(H) := \max\left\{k \in \mathbb{N} \mid \exists P, Q \in H \text{ d'ordre } \ell^k,\ \phi_{\lambda^k}(P, Q) \text{ est d'ordre } \ell^k\right\}.$$

Lorsque H est de la forme $T_\lambda[\lambda^n]$, nous allons montrer que

$$m_1(T_\lambda[\lambda^n]) = m(T_\lambda[\lambda^n]) = n.$$

Définition 5.1. Nous appelons *propriété μ* pour une variété abélienne le fait d'avoir, pour tout sous-groupe fini $H \subset T_\lambda[\lambda^\infty]$, l'égalité à indice fini près, uniformément en (ℓ, H) :

$$K(\mu_{\ell^{m_1(H)}}) \asymp K(H) \cap K(\mu_{\ell^\infty}).$$

Propriété μ pour $T_\lambda[\lambda^n]$. Soit $n \geq 1$ un entier. La propriété μ pour $T_\lambda[\lambda^n]$ découle essentiellement formellement de la propriété μ pour $A[\ell^n]$ et du fait que le multiplicateur $\mathrm{mult}_\lambda(\rho_\lambda(\sigma))$ est $\chi_\ell(\sigma)$. Plus précisément, on sait que concernant l'image de la représentation λ-adique résiduelle ρ_λ (à valeur dans $\mathbb{F}_\lambda$), on a

Proposition 5.2. *Soit A de type I ou II et pleinement de type Lefschetz. On a les* (*presque*) *égalités suivantes, le produit portant sur les places λ au dessus de ℓ dans l'anneau des entiers $\mathcal{O}_E$ de E :*

(1) $[\rho_\ell(G_K), \rho_\ell(G_K)] \mathrel{\mathring{=}} \prod_{\lambda|\ell} \mathrm{Sp}_{2h}(\mathbb{F}_\lambda) \mathrel{\mathring{=}} \mathrm{Sp}_{2h}(\mathcal{O}_E/\ell\mathcal{O}_E)$.

(2) $\rho_\lambda(G_K) \mathrel{\mathring{=}} \{x \in \mathrm{GSp}_{2h}(\mathbb{F}_\lambda) \mid \mathrm{mult}(x) \in \mathbb{F}_\ell^\times\}$.

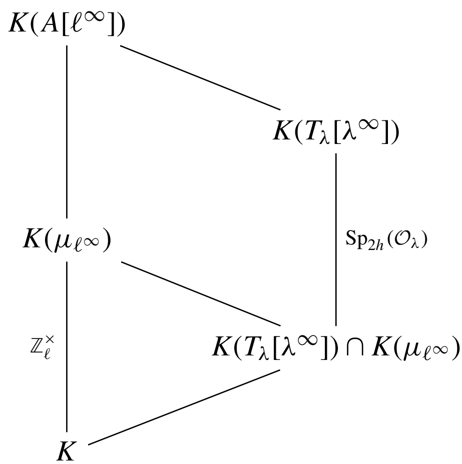
(3) $[\rho_{\ell^\infty}(G_K), \rho_{\ell^\infty}(G_K)] \mathrel{\mathring{=}} \mathrm{Hdg}(A)(\mathbb{Z}_\ell) = \prod_{\lambda|\ell} \mathrm{Sp}_{2h}(\mathcal{O}_\lambda)$.

(4) $\rho_{\lambda^\infty}(G_K) \mathrel{\mathring{=}} \{x \in \mathrm{GSp}_{2h}(\mathcal{O}_\lambda) \mid \mathrm{mult}(x) \in \mathbb{Z}_\ell^\times\}$.

(5) $\rho_{\ell^\infty}(G_K) \mathrel{\mathring{=}} \mathrm{MT}(A)(\mathbb{Z}_\ell) = \left\{(x_\lambda) \in \prod_{\lambda|\ell} \mathrm{GSp}_{2h}(\mathcal{O}_\lambda) \mid \forall\lambda, \exists y \in \mathbb{Z}_\ell^\times, \mathrm{mult}(x_\lambda) = y\right\}$.

Démonstration. L'hypothèse que A est de type Lefschetz signifie que $\mathrm{Hdg}(A) = \mathrm{Res}_{E/\mathbb{Q}} \mathrm{Sp}_{E,2h}$ et $\mathrm{MT}(A) = \mathbb{G}_m \mathrm{Res}_{E/\mathbb{Q}} \mathrm{Sp}_{E,2h}$. L'hypothèse que A est pleinement de Lefschetz signifie que l'image de Galois est d'indice fini dans $\mathrm{MT}(A)(\mathbb{Z}_\ell)$. Comme nous l'expliquons en appendice (théorème 10.1), ceci entraîne que cet indice est borné *indépendamment* de ℓ. En particulier l'indice de $\rho_\lambda(G_K)\cap \mathrm{Sp}_{2h}(\mathbb{F}_\lambda)$ est borné indépendamment de ℓ, disons par c. Observons maintenant que $\mathrm{Sp}_{2h}(\mathbb{F}_{\ell^m})$ ne possède pas de sous-groupe d'indice "petit" (ceci se voit en appliquant les lemmes 2.5 et 2.13 de [Hindry et Ratazzi 2012]), c'est-à-dire que, pour $\ell \geq \ell_0 = \ell_0(c)$, un sous-groupe d'indice inférieur à c est égal au groupe $\mathrm{Sp}_{2h}(\mathbb{F}_\lambda)$ tout entier. D'après le lemme 2.6 nous pouvons conclure que, pour $\ell \geq \ell_0$, nous avons $[\rho_{\lambda^\infty}(G_K), \rho_{\lambda^\infty}(G_K)] = \mathrm{Sp}_{2h}(\mathcal{O}_\lambda)$. Ensuite en utilisant le fait que $\mathrm{mult}_\lambda(\rho_{\lambda^\infty}(\sigma)) = \chi_{\ell^\infty}(\sigma)$ et que le caractère cyclotomique est surjectif sur $\mathbb{Z}_\ell^\times$, toujours pour ℓ assez grand, on conclut que $\rho_{\ell^\infty}(G_K) \mathrel{\mathring{=}} \mathrm{MT}(A)(\mathbb{Z}_\ell)$, comme annoncé. Les autres égalités s'en déduisent aisément. □

On peut déduire de ces considérations l'observation suivante concernant la partie cyclotomique des extensions engendrées, valable pour A pleinement de type Lefschetz, de type I ou II. On peut décrire la situation via la tour d'extensions suivante :



Nous résumons cela dans le corollaire suivant.

Corollaire 5.3. *Soit A de type I ou II, et pleinement de type Lefschetz. On a les (presque) égalités suivantes* :

$$K(T_\lambda[\lambda^\infty]) \cap K(\mu_{\ell^\infty}) \mathrel{\mathring{=}} K(\mu_{\ell^\infty}),$$

On a le même résultat en niveau fini par réduction modulo ℓ^n.

Propriété μ pour $H \subset T_\lambda[\lambda^n]$.

Proposition 5.4. *Soit H un sous-groupe fini de $T_\lambda[\lambda^\infty]$. On a, uniformément en (ℓ, H), l'inégalité*

$$[K(\mu_{\ell^{m_1(H)}}) : \mathbb{Q}] \ll [K(H) \cap K(\mu_{\ell^\infty}) : \mathbb{Q}].$$

Démonstration. Soit $x, y \in H$ deux points de même ordre ℓ^n tels que $\phi_{\lambda^n}(x, y)$ est un élément d'ordre $\ell^{m_1(H)}$. Montrons que l'extension $K(x, y)$ contient "presque" $K(\mu_{\ell^{m_1(H)}})$. Ces deux extensions sont des sous-K-extensions de $K(T_\lambda[\lambda^n])$ et par la description du groupe de Galois de $K(T_\lambda[\lambda^n])/K$, on voit que le groupe de Galois $G_{x,y}$ de $K(T_\lambda[\lambda^n])$ sur $K(x, y)$ est donnée par la presque égalité suivante (valable pour tout ℓ assez grand),

$$G_{x,y} \mathrel{\mathring{=}} \big\{\rho_{\lambda^n}(\sigma) \in \mathrm{GSp}_{2h}(\mathcal{O}_\lambda/\ell^n\mathcal{O}_\lambda) \mid \sigma \in G_K,\ \sigma \cdot x = x,\ \sigma \cdot y = y, \\ \text{et } \chi_{\ell^n}(\sigma) \in (\mathbb{Z}/\ell^n\mathbb{Z})^\times\big\}.$$

Soit donc $\sigma \in G_K$ tel que $\rho_{\lambda^n}(\sigma) \in G_{x,y}$. On a

$$\phi_{\lambda^n}(x, y) = \phi_{\lambda^n}(\rho_{\lambda^n}(\sigma)(x), \rho_{\lambda^n}(\sigma)(y)) = \chi_{\ell^n}(\sigma)\phi_{\lambda^n}(x, y).$$

On en déduit que $\chi_{\ell^n}(\sigma) - 1$ est un multiple de l'ordre de $\phi_{\lambda^n}(x, y)$ dans $\mathcal{O}_\lambda/\ell^n\mathcal{O}_\lambda$, autrement dit que $\chi_{\ell^n}(\sigma) = 1 \bmod \ell^{m_1(H)}$. Or le groupe de Galois de $K(T_\lambda[\lambda^n])$ sur $K(\mu_{\ell^{m_1(H)}})$ est précisément constitué des $\rho_{\lambda^n}(\sigma)$ tels que $\chi_{\ell^n}(\sigma) = 1 \bmod \ell^{m_1(H)}$. On en déduit le résultat. □

Nous pouvons maintenant prouver la propriété μ proprement dite :

Proposition 5.5. *En notant*

$$\delta(H) := \big(\mathbb{Z}_\ell^\times : \mathrm{mult}(G_0(H))\big), \quad \textit{où } G_0(H) = \mathrm{Gal}(K(A[\lambda^\infty])/K(H)),$$

on a, pour tout sous-groupe H fini de $T_\lambda[\lambda^\infty]$, l'égalité à indice fini près, uniformément en (ℓ, H),

$$[K(H) \cap K(\mu_{\ell^\infty}) : K] \asymp \delta(H).$$

De plus, pour tout H sous-groupe fini de $T_\lambda[\lambda^\infty]$, on a l'inclusion suivante, qui est une égalité à un indice fini près uniformément en (ℓ, H) :

$$K(H) \cap K(\mu_{\ell^\infty}) \subset K(\mu_{\ell^{m_1(H)}}) \quad \textit{et} \quad K(H) \cap K(\mu_{\ell^\infty}) \asymp K(\mu_{\ell^{m_1(H)}}).$$

Démonstration. On a la presque égalité $\mathrm{Gal}(K(T_\lambda[\lambda^\infty])/K) \mathrel{\mathring{=}} G_\lambda$ (introduit au lemme 2.2). Le groupe de Galois $\mathrm{Gal}(K(T_\lambda[\lambda^\infty])/K(\mu_{\ell^\infty}))$ s'identifie (c'est une presque égalité) alors avec $\mathrm{SG}_\lambda := G_\lambda \cap \mathrm{Ker}(\mathrm{mult})$. Alors $K(H) \cap K(\mu_{\ell^\infty})$ est la sous-extension fixée par le groupe U engendré par SG_λ et $G_0(H)$. On voit immédiatement que le noyau de $G_\lambda \xrightarrow{\mathrm{mult}} \mathbb{Z}_\ell^\times \to \mathbb{Z}_\ell^\times / \mathrm{mult}(G_0(H))$ est le groupe U, d'où le premier énoncé.

Passons maintenant à la seconde partie de la proposition. Commençons par considérer H_∞ un sous-groupe isotrope maximal de $T_\lambda(A)$. Par le lemme 4.2, on peut supposer que dans une décomposition $T_\lambda(A) = H_\infty \oplus H'_\infty$ la forme symplectique s'écrit comme la forme canonique J. On voit alors aisément que,

$$\begin{aligned}\mathrm{Gal}\big(K(T_\lambda[\lambda^\infty])/K(H_\infty)\big) &\asymp \left\{M=\begin{pmatrix} I & * \\ 0 & * \end{pmatrix} \in \mathrm{GSp}_{2h}(\mathcal{O}_\lambda) \;\middle|\; \mathrm{mult}(M)\in \mathbb{Z}_\ell^\times\right\}\\ &= \left\{M=\begin{pmatrix} I & S \\ 0 & \alpha I \end{pmatrix} \;\middle|\; \alpha\in \mathbb{Z}_\ell^\times \text{ et } S \text{ symétrique}\right\}.\end{aligned}$$

D'après le lemme 2.2, le groupe engendré par ce dernier groupe et par le groupe $\mathrm{Sp}_{2h}(\mathcal{O}_\lambda) \mathrel{\mathring{=}} \mathrm{Gal}(K(T_\lambda[\lambda^\infty])/K(\mu_{\ell^\infty}))$ est $\{x \in \mathrm{GSp}_{2h}(\mathcal{O}_\lambda) \mid \mathrm{mult}(x) \in \mathbb{Z}_\ell^\times\}$ tout entier. Ainsi $K(H_\infty)\cap K(\mu_{\ell^\infty}) \asymp K$. Si H est un sous-groupe fini de $T_\lambda[\lambda^\infty]$ totalement isotrope, dans ce cas le lemme 4.3 et ce qui précède nous permettent de conclure : on a $K(H)\cap K(\mu_{\ell^\infty}) \asymp K$.

Soit maintenant H un sous-groupe fini non isotrope de $T_\lambda[\lambda^\infty]$. Le groupe $[\ell^{m_1(H)}](H)$ est totalement isotrope. En effet si P et Q sont deux points d'ordre ℓ^n dans H, alors, par définition de $m_1(H)$,

$$\phi_{\lambda^{n-m_1(H)}}(\ell^{m_1(H)}P, \ell^{m_1(H)}Q) = \phi_{\lambda^n}(P, Q)^{\ell^{m_1(H)}} = 1.$$

En appliquant le lemme 4.3, on trouve donc un sous-groupe H' contenant $[\ell^{m_1}](H)$ de même exposant et il existe un sous-$\mathcal{O}_\lambda$-module H_∞ totalement isotrope de $T_\lambda(A)$ tel que, si, pour tout entier $n\geq 1$, $\mathrm{pr}_n : T_\lambda(A)\to T_\lambda(A)/\ell^n T_\lambda(A) = T_\lambda[\lambda^n]$ désigne la projection canonique, on a

$$\mathrm{pr}_{r_H}(H_\infty) = H'.$$

Par le lemme 4.2, on peut supposer que dans une décomposition $T_\lambda(A) = H_\infty \oplus H'_\infty$ la forme symplectique s'écrit comme la forme canonique J. Pour tout $n\geq 1$, notons

$$H_n := \mathrm{pr}_n(H_\infty) = H_\infty/H_\infty \cap \ell^n T_\lambda(A).$$

On a pour tout $n\geq 1$, $[\ell]H_{n+1} = H_n$. On peut donc poser

$$H^\infty = \bigcup_{n\geq 1} H_n \subset T_\lambda[\lambda^\infty].$$

De plus, on voit que, dans $K(T_\lambda[\lambda^\infty])$, le groupe de Galois correspondant à H_∞ est le même que celui correspondant à H^∞. On a

$$H\subset [\ell^{m_1(H)}]^{-1}(H') = [\ell^{m_1(H)}]^{-1}(H_{r_H}) \subset [\ell^{m_1(H)}]^{-1}(H^\infty).$$

En considérant la multiplication par $[\ell^{m_1(H)}]$ sur H^∞, on en déduit (car H^∞ est ℓ-divisible) que

$$H\subset H^\infty + \ker[\ell^{m_1(H)}] =: \widetilde{H^\infty},$$

où $[\ell^n]$ est le morphisme de multiplication dans $T_\lambda[\lambda^\infty]$. Ainsi comme dans le cas totalement isotrope, on se ramène à une situation où un lemme de groupe permet de conclure : le groupe de Galois $\mathrm{Gal}\big(K(T_\lambda[\lambda^\infty])/K(\widetilde{H^\infty})\big)$ n'est autre que (il s'agit d'une égalité $\asymp$ à indice fini près)

$$\big\{M\in\mathrm{GSp}_{2h}(\mathcal{O}_\lambda)\,|\,\mathrm{mult}(M)\in\mathbb{Z}_\ell^\times \text{ et } \forall i\le g,\ Me_{g+i}=e_{g+i} \bmod \ell^{m_1(H)},\ Me_i=e_i\,\big\}.$$

La même preuve que celle du corollaire 2.11 de [Hindry et Ratazzi 2012] donne alors le résultat : le groupe engendré par $\mathrm{Gal}\big(K(T_\lambda[\lambda^\infty])/K(\widetilde{H^\infty})\big)$ et $\mathrm{Sp}_{2h}(\mathcal{O}_\lambda)$ est (avec une égalité $\asymp$ à indice fini près)

$$\big\{M\in\mathrm{GSp}_{2h}(\mathcal{O}_\lambda)\mid \mathrm{mult}(M)\in\mathbb{Z}_\ell^\times \text{ et } \mathrm{mult}(M)\equiv 1 \bmod \ell^{m_1(H)}\big\}.$$

Notamment, on a,

$$K(H)\cap K(\mu_{\ell^\infty})\subset K(\widetilde{H^\infty})\cap K(\mu_{\ell^\infty})\subset K(\mu_{\ell^{m_1(H)}}),$$

la seconde inclusion étant aussi une égalité à indice fini près, i.e.,

$$K(\widetilde{H^\infty})\cap K(\mu_{\ell^\infty})\asymp K(\mu_{\ell^{m_1(H)}}).$$

La proposition 5.4 permet de conclure. □

6. Preuve du théorème principal pour $H\subset A[\ell]$

Soit A/K une variété abélienne sur un corps de nombres, telle que $\mathrm{End}_K(A)=\mathrm{End}_{\overline{K}}(A)$. On commence par se ramener au cas ℓ-adique (cf. [Hindry et Ratazzi 2010, proposition 4.1]) grâce à la presque indépendance rappelée à la proposition 3.2 :

Proposition 6.1. *Soit $\alpha>0$. Pour démontrer que $\gamma(A)\le\alpha$, il suffit de montrer que : il existe une constante strictement positive $C(A/K)$ ne dépendant que de A/K telle que pour tout nombre premier ℓ, pour tout sous-groupe fini H_ℓ de $A[\ell^\infty]$, on a*

$$\mathrm{Card}(H_\ell)\le C(A/K)[K(H_\ell):K]^\alpha. \tag{3}$$

Remarque 6.2. Rappelons que l'on a supposé que la variété abélienne A/K est telle que $\mathrm{End}_K(A)=\mathrm{End}_{\overline{K}}(A)$. Concernant notre question de borne sur la torsion, ceci nous permet de supposer que le groupe fini $H\subset A[\ell^n]$ est en fait un $\mathrm{End}_{\overline{K}}(A)$-module. En effet : notons H_E le $\mathrm{End}_{\overline{K}}(A)$-module engendré par H et supposons que l'on ait pour H_E une inégalité de la forme suivante, uniformément en (ℓ, H),

$$|H_E|\ll[K(H_E):K]^\alpha.$$

On a donc $|H|\ll[K(H_E):K]^\alpha$ car H est inclus dans H_E. Mais $\mathrm{End}_K(A)=\mathrm{End}_{\overline{K}}(A)$, donc si $x\in H$ et $f\in\mathrm{End}_{\overline{K}}(A)$ alors $f(x)$ est encore un point de A qui est $K(H)$ rationnel, donc $K(H)=K(H_E)$. En particulier ceci implique que $|H|\ll[K(H):K]^\alpha$ comme annoncé.

Nous nous plaçons dans toute la suite de ce paragraphe dans la situation particulière d'une variété abélienne A définie sur K, géométriquement simple de type I ou II, qui est pleinement de type Lefschetz. Nous supposerons de plus que $\ell \notin S_{\mathrm{ex}}(A)$ de sorte à pouvoir appliquer les techniques développées dans les paragraphes précédents. Enfin nous prenons le cas particulier d'une situation horizontale d'un sous-groupe H de $A[\ell]$ (en particulier il s'agit d'un $\mathbb{F}_\ell$-espace vectoriel). Par la remarque précédente, nous pouvons même supposer que H est un $\mathcal{O}_E/\ell\mathcal{O}_E$-module. Nous avons la décomposition suivante :

$$H = \begin{cases} \prod_{\lambda|\ell} H[\lambda] \subset \prod_{\lambda|\ell} T_\lambda[\lambda] & \text{(Type I)},\\ \prod_{\lambda|\ell} H[\lambda] \oplus H[\lambda] \subset \prod_{\lambda|\ell} T_\lambda[\lambda] \oplus T_\lambda[\lambda] & \text{(Type II)}.\end{cases}$$

On sait par la proposition 5.2 que pour tout ℓ on a,

$$\rho_\lambda(G_K) \mathrel{\mathring{=}} \left\{M \in \mathrm{GSp}_{2h}(\mathbb{F}_\lambda) \mid \mathrm{mult}(M) \in \mathbb{F}_\ell^\times\right\}.$$

Dans notre situation les $H[\lambda] \subset T_\lambda[\lambda]$ sont des $\mathbb{F}_\lambda$-espaces vectoriels. Rappelons que l'on a, uniformément en (ℓ, H), l'égalité à indice fini près $\delta(H[\lambda]) \asymp [K(H[\lambda]) \cap K(\mu_\ell) : K]$. On obtient ainsi :

Lemme 6.3. *Si $H[\lambda]$ est inclus dans un sous-espace totalement isotrope du $\mathbb{F}_\lambda$-espace vectoriel $T_\lambda[\lambda]$ alors, uniformément en (ℓ, H), on a $\delta(H[\lambda]) \asymp 1$. Sinon $\delta(H[\lambda]) \asymp \ell$.*

Lemme 6.4. *Uniformément en (ℓ, H), on a*

$$\delta(H[\lambda]) \asymp (\mathbb{F}_\ell^\times : \mathrm{mult}(G_0(H[\lambda]))), \quad \textit{où } G_0(H[\lambda]) = \mathrm{Gal}(K(T_\lambda[\lambda])/K(H[\lambda])).$$

On a de plus

$$[K(H[\lambda]) : K] \mathrel{\mathring{=}} (\rho_\lambda(G_K) : G_0(H[\lambda])) \mathrel{\mathring{=}} \delta(H[\lambda])(\mathrm{Sp}_{2h}(\mathbb{F}_\lambda) : G(H[\lambda])).$$

Démonstration. Pour le premier point, on a $\mathrm{Gal}(K(T_\lambda[\lambda])/K) \mathrel{\mathring{=}} \rho_\lambda(G_K)$. Le groupe de Galois $\mathrm{Gal}(K(T_\lambda[\lambda])/K(\mu_\ell))$ est alors presque égal à

$$\mathrm{SG}_\lambda := \rho_\lambda(G_K) \cap \mathrm{Ker}(\mathrm{mult}).$$

Alors $K(H[\lambda]) \cap K(\mu_\ell)$ est la sous-extension fixée par le groupe U engendré par SG_λ et $G_0(H[\lambda])$. On voit immédiatement que le noyau de $\rho_\lambda(G_K) \xrightarrow{\mathrm{mult}} \mathbb{F}_\ell^\times \to \mathbb{F}_\ell^\times/\,\mathrm{mult}(G_0(H[\lambda]))$ est le groupe U. Pour le second point : la première égalité est donnée par la théorie de Galois car on a que $\mathrm{Gal}(K(T_\lambda[\lambda])/K) \mathrel{\mathring{=}} \rho_\lambda(G_K)$. La seconde égalité est une chasse au diagramme facile. □

Notons d_λ la dimension de $H[\lambda]$ sur $\mathbb{F}_\lambda$ et $(e_1, \ldots e_{d_\lambda})$ une base que l'on complète en une base $(e_1, \ldots, e_{2h})$ de $T_\lambda[\lambda]$. On définit

$$G(H[\lambda]) = \left\{M \in \mathrm{Sp}_{2h}(\mathbb{F}_\lambda) \mid Me_i = e_i,\ 1 \le i \le d_\lambda\right\}.$$

Notons $(\hat{e}_1, \dots, \hat{e}_{2h})$ une base de $T_\lambda(A)$ relevant la base sur $\mathbb{F}_\lambda$. Introduisons maintenant le groupe algébrique sur $\mathcal{O}_\lambda$ suivant :

$$G_1 := \{M \in \mathrm{Sp}_{2h} \mid M\hat{e}_i = \hat{e}_i,\ 1 \le i \le d_\lambda\}.$$

On voit que

$$G(H[\lambda]) = \{M \in \mathrm{Sp}_{2h}(\mathbb{F}_\lambda) \mid M \in G_1 \bmod \lambda\}.$$

Par changement de base symplectique sur $\mathbb{F}_\lambda$, G_1 est conjugué sur $\mathbb{F}_\lambda$ à l'un des groupes $P_{r,s}$ introduits précédemment. En posant $G = \mathrm{Sp}_{2h}$, et en rappelant que $\mathrm{Card}(\mathbb{F}_\lambda) = \ell^{f(\lambda)}$, on voit que, d'après lemme 2.3 on a

$$[K(H[\lambda]) : K] \asymp \ell^{m(H[\lambda])} \ell^{f(\lambda)\,\mathrm{codim}\, P_{r_\lambda, s_\lambda}},$$

où (r_λ, s_λ) (avec éventuellement $s_\lambda = 0$) est le couple correspondant à $H[\lambda]$.

Utilisant le lemme 6.3, le théorème 6.6 de [Hindry et Ratazzi 2010] s'adapte immédiatement (cf. la proposition 7.3 ci-après) pour donner :

Proposition 6.5. *Avec les notations précédentes, uniformément en* (ℓ, H), *on a*

$$\ell^{m(H)} \asymp [K(H) \cap K(\mu_\ell) : K] \asymp \max_{\lambda|\ell} \ell^{m(H[\lambda])}$$

et

$$[K(H) : K(\mu_{\ell^{m(H)}})] \asymp \prod_{\lambda|\ell} [K(H[\lambda]) : K(\mu_{\ell^{m(H[\lambda])}})].$$

Cas totalement décomposé. Nous supposons ici que $\ell \notin S_{\mathrm{ex}}(A)$ est totalement décomposé dans $\mathcal{O}_E$. Notre situation est alors la suivante :

$$H = \begin{cases} \prod_{\lambda|\ell} H[\lambda] \subset \prod_{\lambda|\ell} T_\lambda[\lambda] & \text{et} \quad \rho_\ell = \prod_{\lambda|\ell} \rho_\lambda \qquad \text{(Type I)}, \\ \prod_{\lambda|\ell} H[\lambda] \oplus H[\lambda] \subset \prod_{\lambda|\ell} T_\lambda[\lambda] \oplus T_\lambda[\lambda] & \text{et} \quad \rho_\ell = \prod_{\lambda|\ell} \rho_\lambda \oplus \rho_\lambda \qquad \text{(Type II)}. \end{cases}$$

De plus

$$\mathrm{Gal}(K(A[\ell])/K(\mu_\ell)) = \prod_{\lambda|\ell} \mathrm{Gal}(K(T_\lambda[\lambda])/K(\mu_\ell)).$$

Du point du vue combinatoire, les formules sont identiques à celles d'un produit de variétés abéliennes de type GSp_{2h} et, les résultats du paragraphe précédent nous indiquent que la combinatoire n'est finalement autre que celle d'un produit de e variétés abéliennes de type GSp_{2h}, deux à deux non isogènes. Nous pouvons donc directement en déduire la valeur de l'exposant $\gamma(A)$.

Définition 6.6. Nous noterons dans la suite : $d = 1$ si A et de type I et $d = 2$ si A est de type II.

Les calculs de [Hindry et Ratazzi 2012, paragraphes 4.1 et 6.2] donnent dans ce cas :

$$\gamma(A) = \sup_{I \subset \{1,\dots,e\}} \frac{2\sum_{\lambda \in I} dh}{1 + (2h^2 + h)|I|}.$$

Ce sup se calcule aisément (le max est atteint pour $I = \{1, \dots, e\}$) et on trouve donc

$$\gamma(A) = \frac{2dhe}{1+(2h^2+h)e} = \frac{2\dim A}{1+\dim \mathrm{Res}_{E/\mathbb{Q}} \mathrm{Sp}_{2h}} = \frac{2\dim A}{\dim \mathrm{MT}(A)}.$$

Cas général. Nous ne supposons plus désormais que ℓ est totalement décomposé, la combinatoire qui résulte est donc différente et il faut dans ce cadre général la refaire explicitement (ceci contient d'ailleurs le cas du sous-paragraphe précédent). On a

$$H[\lambda] = (\mathcal{O}_\lambda/\lambda)^{r_\lambda+s_\lambda} \quad \text{avec } s_\lambda = 0 \Longleftrightarrow H[\lambda] \text{ est inclus dans un Lagrangien.}$$

De plus on a, quitte à réordonner,

$$0 \le s_\lambda \le r_\lambda \le h, \quad \text{où } 2h = \dim_{\mathbb{F}_\lambda} T_\lambda[\lambda],$$
$$\sum_{\lambda|\ell} f(\lambda) = [E:\mathbb{Q}] = e \quad \text{et} \quad dhe = g = \dim A,$$

où l'on note comme précédemment $d = 1$ si A est de type I et $d = 2$ si A est de type II.

On obtient finalement, sous les conditions précédentes, la valeur suivante pour le cardinal de H :

$$\mathrm{Card}(H) = \ell^{d\sum_{\lambda|\ell} f(\lambda)(r_\lambda+s_\lambda)}.$$

Le degré de l'extension $[K(H):K]$ dépend selon que les $H[\lambda]$ sont ou non inclus dans des Lagrangiens. Si l'un des $H[\lambda]$ n'est pas inclus dans un Lagrangien alors nous obtenons

$$[K(H):K] \asymp \ell^{1+\sum_{\lambda|\ell} f(\lambda)\,\mathrm{codim}\, P_{r_\lambda,s_\lambda}}.$$

Si par contre tout les $H[\lambda]$ sont inclus dans un Lagrangien alors nous obtenons

$$[K(H):K] \asymp \ell^{\sum_{\lambda|\ell} f(\lambda)\,\mathrm{codim}\, P_{r_\lambda,0}}.$$

Il reste à interpréter le quotient de l'exposant de ℓ du Card(H) par celui de $[K(H):K]$ pour conclure : c'est l'objet du paragraphe combinatoire suivant.

Combinatoire. Comme précédemment on note $d = 1$ si A est de type I et $d = 2$ si A est de type II. Nous sommes ramenés à calculer la quantité :

$$\frac{1}{d}\gamma := \max_{r_\lambda,s_\lambda} \frac{\sum_{\lambda|\ell} f(\lambda)(r_\lambda+s_\lambda)}{\delta+\sum_{\lambda|\ell} f(\lambda)\,\mathrm{codim}\, P_{r_\lambda,s_\lambda}}$$

où le maximum est pris pour $0 \le s_\lambda \le r_\lambda \le h$ et δ vaut 0 (resp. 1) si tous les s_λ sont nuls (resp. si l'un des s_λ est non nul).

Proposition 6.7. *Soit $\gamma = \gamma(A)$ défini ci-dessus ; alors*

$$\gamma = \frac{2dhe}{1+2eh^2+he} = \frac{2\dim A}{\dim \mathrm{MT}(A)}.$$

Nous donnons ci-dessous, dans le cas particulier de la proposition ci-dessus, une preuve via les interpolateurs de Lagrange. Un argument combinatoire différent, sera donné plus loin dans le cas général de la preuve du lemme 7.5, l'argument suivant n'est donc pas indispensable mais a l'avantage d'être assez direct.

Remarques. (1) On peut réécrire, pour $P_{r,s} \subset \mathrm{Sp}_{2g}$:

$$\mathrm{codim}\, P_{r,s} = 2g(r+s) - rs - \left(\frac{r^2+s^2}{2}\right) + \frac{r+s}{2} = \left(2g + \frac{1}{2} - \frac{r+s}{2}\right)(r+s).$$

On observe en particulier que la dimension ou codimension de $P_{r,s}$ ne dépend que de $r+s$.

(2) Nous allons devoir étudier le sens de variation d'une fraction rationnelle de la forme

$$f(x) = \frac{a-x}{A - x(x-1)/2}$$

dont la dérivée s'écrit

$$f'(x) = -2\frac{(x-a)^2 + 2A + a - a^2}{(2A - x(x-1))^2}$$

et est donc décroissante dès que $2A + a \geq a^2$.

Démonstration de la proposition 6.7. La preuve consiste à appliquer le calcul différentiel à la fonction de variables $(\underline{r}, \underline{s}) := (r_\lambda, s_\lambda)_{\lambda|\ell}$ dont on veut évaluer le maximum :

$$\psi(\underline{r}, \underline{s}) := \frac{N(\underline{r}, \underline{s})}{D(\underline{r}, \underline{s})} := \frac{\sum_{\lambda|\ell} f(\lambda)(r_\lambda + s_\lambda)}{\delta + \sum_{\lambda|\ell} f(\lambda)\, \mathrm{codim}\, P_{r_\lambda, s_\lambda}}$$

(nous écrivons la fonction sous la forme "Numérateur/Dénominateur= N/D"). Commençons par traiter le cas où tous les s_λ sont nuls. Les différentielles des deux fonctions N et D s'écrivent

$$\partial N = (f(\lambda))_{\lambda|\ell} \quad \text{et} \quad \partial D = \left(f(\lambda)\left(2h - r_\lambda + \tfrac{1}{2}\right)\right)_{\lambda|\ell}.$$

Le théorème de Lagrange indique que, en un maximum de N/D, ces deux différentielles sont proportionnelles, donc $2h - r_\lambda + \frac{1}{2}$ est constant, ou encore, $r_\lambda = 2h - \kappa$ (avec $h \leq \kappa \leq 2h$). On obtient alors

$$\frac{N}{D} = \frac{2he - \kappa e}{2\sum_{\lambda|\ell} f(\lambda)h^2 + he - e\kappa(\kappa-1)/2} = \frac{2h-\kappa}{2h^2+h-\kappa(\kappa-1)/2}.$$

La fonction à droite est décroissante avec κ et donc majorée par la valeur en $\kappa = h$, c'est-à-dire $2/(3h+3)$ (noter que $\kappa \geq h$).

On traite ensuite le cas général (avec l'un des s_λ non nul), on pose donc

$$N := \sum_{\lambda|\ell} f(\lambda)(r_\lambda + s_\lambda), \quad D := 1 + \sum_{\lambda|\ell} f(\lambda)(r_\lambda + s_\lambda)\Big(2h + \frac{1}{2} - \frac{r_\lambda + s_\lambda}{2}\Big).$$

Le théorème de Lagrange indique maintenant que, en un maximum de N/D, on aura $s_\lambda + r_\lambda = 2h - \kappa$, avec maintenant $0 \leq \kappa \leq 2h$. En reportant on obtient :

$$\frac{N}{D} \leq \frac{e(2h-\kappa)}{1+2\sum_{\lambda|\ell} f(\lambda)h^2+eh-e\kappa(\kappa-1)/2} = \frac{2h-\kappa}{1/e+2h^2+h-\kappa(\kappa-1)/2}.$$

Cette dernière fonction est décroissante en κ donc majorée par la valeur en $\kappa = 0$, ce qui donne au final :

$$\psi \leq \max\Big\{\frac{2}{3(h+1)}, \frac{2he}{1+2eh^2+eh}\Big\} = \frac{2he}{1+2eh^2+eh}.$$

Observons que $\kappa = 0$ correspond à $r_\lambda + s_\lambda = 2h$ donc à $r_\lambda = s_\lambda = h$. En considérant donc le cas $r_\lambda = s_\lambda = h$, on obtient

$$\psi = \frac{2he}{1+2h^2e+he}. \qquad \square$$

7. Cas d'un groupe H quelconque

Dans ce paragraphe nous allons donner une preuve du résultat principal (théorème 1.14). Rappelons que l'on a supposé que la variété abélienne A/K est un produit $\prod_{i=1}^d A_i^{n_i}$ de variétés abéliennes simples, chacune de type I ou II et chacune pleinement de type Lefschetz. Nous avons déjà indiqué que l'on peut supposer de plus que pour tout i, les A_i sont telles que $\mathrm{End}_{\overline{K}}(A_i) = \mathrm{End}_K(A_i)$ et telles que $\mathrm{End}_K(A_i)$ est un ordre maximal dans $\mathrm{End}_K(A_i) \otimes_{\mathbb{Z}} \mathbb{Q}$.

Définition 7.1. Avec la notation de la définition 3.4, nous noterons dans la suite $S_{\mathrm{ex}} = \bigcup_{i=1}^d S_{\mathrm{ex}}(A_i)$.

Dans la suite de ce paragraphe nous supposerons que $\ell \notin S_{\mathrm{ex}}$.

Soit H un sous-groupe fini de $A[\ell^\infty]$. Par le paragraphe 4.2 de [Hindry et Ratazzi 2010], on peut supposer que H s'écrit sous la forme $H = \prod_{i=1}^d H_i^{n_i}$. De plus, par la remarque 6.2, nous pouvons supposer que chaque H_i est un $\mathcal{O}_{\ell,i}$-module, inclus dans un $A_i[\ell^n]$ pour n convenable (où l'on note $\mathcal{O}_{\ell,i}$ le tensorisé par $\mathbb{Z}_\ell$ de $\mathrm{End}(A_i)$). Notons

$$\mathfrak{I}_\ell := \big\{(\lambda, i) \mid i \in \{1, \dots, d\}, \text{ et } \lambda \text{ une place du centre de } \mathrm{End}(A_i) \otimes \mathbb{Q} \text{ au-dessus de } \ell\big\}.$$

Pour $(\lambda, i) \in \mathfrak{I}_\ell$, posons $\mathcal{O}_{\lambda,i}$ la composante λ-adique de $\mathcal{O}_{\ell,i}$ et posons $X_{\lambda,i}$ le morceau de H_i correspondant à λ. Dans le cas de type II, $X_{\lambda,i}$ se décompose à son tour en deux copies isomorphes : $X_{\lambda,i} = H_{\lambda,i} \oplus H_{\lambda,i}$. Dans le cas de type I, on pose $H_{\lambda,i} := X_{\lambda,i}$. Avec des notations évidentes, les $H_{\lambda,i}$ sont des $\mathcal{O}_{\lambda,i}/\ell^n\mathcal{O}_{\lambda,i}$-sous-modules de $T_{\lambda,i}[\lambda^n]$. On a finalement la décomposition suivante de H :

$$H = \prod_{(\lambda,i)\in\mathfrak{I}_\ell} X_{\lambda,i}^{n_i}.$$

En tant que groupe on sait que pour $(\lambda, i) \in \mathfrak{I}_\ell$,

$$\mathcal{O}_{\lambda,i}/\ell^n\mathcal{O}_{\lambda,i} = (\mathbb{Z}/\ell^n\mathbb{Z})^{f(\lambda)} \quad \text{et} \quad T_{\lambda,i}[\lambda^n] = (\mathbb{Z}/\ell^n\mathbb{Z})^{2h_i f(\lambda)}.$$

On sait également que, uniformément en ℓ et λ, on a,

$$\rho_{A_i,\lambda^\infty}(G_K) \mathrel{\mathring{=}} \left\{M \in \mathrm{GSp}_{2h_i}(\mathcal{O}_{\lambda,i}) \mid \mathrm{mult}(M) \in \mathbb{Z}_\ell^\times\right\}.$$

Dans tous les cas on obtient ainsi une égalité à indice fini près, en réduisant modulo ℓ^n.

Soit donc $H_{\lambda,i}$ un sous-groupe fini de $T_{\lambda,i}[\lambda^\infty]$. On pose

$$G_0(H_{\lambda,i}) := \left\{M \in \mathrm{GSp}_{2h_i}(\mathcal{O}_{\lambda,i}) \mid \mathrm{mult}(M) \in \mathbb{Z}_\ell^\times, \text{ et } \forall x \in H_{\lambda,i},\ Mx = x\right\},$$

et $G(H_{\lambda,i}) := G_0(H_{\lambda,i}) \cap \mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i})$. Comme $\mathcal{O}_{\lambda,i}/\ell^n\mathcal{O}_{\lambda,i}$-module et comme groupe abstrait, $H_{\lambda,i}$ est de la forme

$$H_{\lambda,i} \simeq \prod_{j=1}^{2h_i} \mathcal{O}_{\lambda,i}/\ell^{m_j}\mathcal{O}_{\lambda,i} \simeq \prod_{j=1}^{2h_i} (\mathbb{Z}/\ell^{m_j}\mathbb{Z})^{f(\lambda)},$$

où nous sous-entendons, pour ne pas alourdir plus que de raison les notations, que les nombres m_j dépendent également de (λ, i).

Notons $e^1_{(\lambda,i)}, \dots, e^{2h_i}_{(\lambda,i)}$ un système de générateurs (en tant que $\mathcal{O}_{\lambda,i}/\ell^n\mathcal{O}_{\lambda,i}$-module) ; les $e^j_{(\lambda,i)}$ étant d'ordre respectifs ℓ^{m_j}. Notons de plus $\{\hat{e}^1_{(\lambda,i)}, \dots, \hat{e}^{2h_i}_{(\lambda,i)}\}$ une base du $\mathcal{O}_{\lambda,i}$-module libre $T_{\lambda,i} := T_\lambda(A_i)$ relevant la famille $\{e^j_{(\lambda,i)}\}$, i.e., telle que $e^j_{(\lambda,i)} = \hat{e}^j_{(\lambda,i)} \bmod \ell^{m_j}$ pour tout j. On a

$$G(H_{\lambda,i}) = \left\{M \in \mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i}) \mid M\hat{e}^j_{(\lambda,i)} = \hat{e}^j_{(\lambda,i)} \bmod \ell^{m_j},\ 1 \le j \le 2h_i\right\}.$$

Lemme 7.2. *Notons* $\delta(H_{\lambda,i}) := \left(\mathbb{Z}_\ell^\times : \mathrm{mult}(G_0(H_{\lambda,i}))\right)$. *Uniformément en* (ℓ, H), *on a alors*

$$[K(H_{\lambda,i}) : K] \asymp (\rho_{A_i,\lambda^\infty}(G_K) : G_0(H_{\lambda,i})) \asymp \delta(H_{\lambda,i})(\mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i}) : G(H_{\lambda,i})).$$

Démonstration. Comme le lemme 6.4. □

Quitte à renuméroter on peut supposer que les exposants m_j (correspondants aux $e^j_{(\lambda,i)}$) sont ordonnés dans l'ordre décroissant : $m_1 \geq \cdots \geq m_{2h_i}$. On pose alors

$$m^1 := \max\{m_i \mid m_i \neq 0\} \quad \text{et par récurrence} \quad m^{r+1} = \max\{m_i \mid m_i < m^r\}.$$

On obtient ainsi une suite strictement décroissante $m^1 > \cdots > m^{t_{\lambda,i}} \geq 1$ (avec $t_{\lambda,i} \leq 2h_i$). Le groupe $H_{\lambda,i}$ est isomorphe à $\prod_{j=1}^{t_{\lambda,i}} (\mathbb{Z}/\ell^{m^j}\mathbb{Z})^{f(\lambda)a_j}$, les a_j dépendants de (λ, i). On définit ensuite pour tout $1 \leq r \leq t_{\lambda,i}$, les sous-ensembles emboités

$$I_r = \{j \in \{1, \ldots, 2h_i\} \mid m_j \geq m^r\} \quad \text{de cardinal} \quad |I_r| = \sum_{j=1}^{r} a_j.$$

Introduisons maintenant la suite croissante de groupes algébriques sur $\mathcal{O}_{\lambda,i}$ suivants :

$$\forall 1 \leq r \leq t_{\lambda,i}, \quad G_{r,(\lambda,i)} := \left\{ M \in \mathrm{Sp}_{2h_i} \mid M\hat{e}^j_{(\lambda,i)} = \hat{e}^j_{(\lambda,i)} \;\; \forall j \in I_{t_{\lambda,i}+1-r} \right\}.$$

On voit que

$$G(H_{\lambda,i}) = \left\{ M \in \mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i}) \mid \forall 1 \leq r \leq t_{\lambda,i} \text{ on a } M \in G_{r,(\lambda,i)} \bmod \ell^{m^{t_{\lambda,i}+1-r}} \right\}.$$

Par changement de base symplectique sur $\mathbb{F}_\lambda$, le couple (λ, i) étant fixé, chacun des $G_{j,(\lambda,i)}$ est conjugué sur $\mathbb{F}_\lambda$ à l'un des groupes $P_{r,s}$ introduits au paragraphe 2. En posant $G = \mathrm{Sp}_{2h_i}$, on voit que, avec les notations du lemme 2.3, on a

$$G(H_{\lambda,i}) = H(m^1, \ldots, m^{t_{\lambda,i}}).$$

On va donc pouvoir appliquer le lemme 2.3.

Cas d'un morceau $H_{\lambda,i}$. Le couple (λ, i) étant fixé nous renoterons dans ce paragraphe $t := t_{\lambda,i}$ afin de soulager un peu les notations. On peut appliquer le lemme 2.3, uniformément en (ℓ, H), on a :

$$(\mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i}) : G(H_{\lambda,i})) \gg \ell^{\sum_{j=1}^{t} f(\lambda)\,\mathrm{codim}(G_{j,(\lambda,i)})(m^{t+1-j} - m^{t+1-(j-1)})},$$

où l'on a posé $m^{t+1} = 0$ et où $\mathrm{codim}(G_{j,(\lambda,i)})$ est la codimension de $G_{j,(\lambda,i)}$ dans Sp_{2h_j}. Les groupes algébriques $G_{j,(\lambda,i)}$ étant conjugués sur $\mathbb{F}_\lambda$ aux $P_{r,s}$ (avec éventuellement $s = 0$), $\mathrm{codim}(G_{j,(\lambda,i)})$ est également la codimension du groupe P_{r_j,s_j} correspondant. Par ailleurs, la suite des $(G_{j,(\lambda,i)})_j$ étant croissante, (λ, i) étant fixé, la suite des $(P_{r_j,s_j})_j$ l'est également. Ceci se traduit par

$$r_j \geq r_{j+1} \quad \text{et} \quad s_j \geq s_{j+1} \quad \text{pour tout } j.$$

Il nous reste à calculer la valeur de $\delta(H_{\lambda,i})$ (ou plutôt une minoration de $\delta(H_{\lambda,i})$). Soit donc $h \in \{0, \ldots, t\}$ maximal tel que $s_h \geq 1$ (on pose $h = 0$ si $s_i = 0$ pour tout i). On a donc

$$s_1 \geq \cdots \geq s_h = 1 > 0 = s_{h+1} = \cdots = s_t \quad \text{et} \quad P_{r_1,s_1} \subset \cdots P_{r_h,s_h} \subset P_{r_{h+1},0} \subset \cdots \subset P_{r_t}.$$

Posons

$$\delta_1 = \cdots = \delta_h = 1 \quad \text{et} \quad \delta_{h+1} = \cdots = \delta_t = 0.$$

Posons $m^{t+1} = 0$. On voit (il s'agit d'une somme télescopique) que

$$m^{t+1-h} = m^{t+1-h} - m^{t+1} = \sum_{j=1}^{t} (m^{t+1-j} - m^{t+1-(j-1)})\delta_j.$$

Or P_{r_h,s_h} (avec $s_h \geq 1$) correspond au groupe $G_{h,(\lambda,i)}$ lui même associé à l'ensemble I_{t+1-h}. Il correspond donc à un morceau de $H_{\lambda,i}$ sur lequel on voit qu'il existe P, Q d'ordre $\ell^{m^{t+1-h}}$ tel que l'accouplement de Weil de $\ell^{m^{t-h}}P$ et $\ell^{m^{t-h}}Q$ est une racine primitive ℓ-ème de 1. Ceci se traduit en disant que

$$\delta(H_{\lambda,i}) \geq \phi(\ell^{m^{t+1-h}}) \asymp \ell^{m^{t+1-h}},$$

ceci restant valable pour $h = 0$. Nous obtenons ainsi la minoration

$$[K(H_{\lambda,i}) : K] \gg \ell^{\sum_{j=1}^{t_{\lambda,i}} (m^{t_{\lambda,i}+1-j} - m^{t_{\lambda,i}+1-(j-1)})(\delta_j + f(\lambda)\,\mathrm{codim}\, P_{r_j,s_j})}.$$

De plus, pour tout entier $k \in \{1, \ldots t\}$,

$$r_{t+1-k} + s_{t+1-k} = |I_k| = \sum_{j=1}^{k} a_j.$$

Invariant $\gamma(A)$ pour $H \subset A[\ell^\infty]$. Nous sommes ici dans la situation présentée au début de cette section avec $H = \prod_{(\lambda,i)\in\mathfrak{I}_\ell} X_{\lambda,i}$. Avec les notations introduites dans le cas d'un $H_{\lambda,i}$ (i.e., au paragraphe précédent, page 1874), on peut, pour tout $(\lambda, i) \in \mathfrak{I}_\ell$, écrire

$$H_{\lambda,i} = \prod_{j=1}^{2h_i} (\mathbb{Z}/\ell^{m_j}\mathbb{Z})^{f(\lambda)} = \prod_{j=1}^{t_{\lambda,i}} (\mathbb{Z}/\ell^{m^j}\mathbb{Z})^{a_j f(\lambda)},$$

où $(m^j)_{j\geq 1}$ est une suite strictement décroissante (le couple (λ, i) étant fixé).

Nous allons utiliser un résultat galoisien que nous avons démontré dans [Hindry et Ratazzi 2010]. Dans le théorème 6.6 de [Hindry et Ratazzi 2010] nous donnons une preuve pour $A = \prod_i A_i$ et avec $M_i = T_\ell(A_i)$ (cf. notations ci-dessous). En fait la même preuve reprise mot pour mot donne :

Proposition 7.3. *Soient $T_\ell(A) = \bigoplus_{j\in J} M_j^{\alpha_j}$ une décomposition de $\mathbb{Z}_\ell$-modules galoisiens vérifiant les deux propriétés suivantes (où $M_j[\ell^\infty]$ désigne $\bigcup_n M_j/\ell^n M_j$) :*

(1) *Pour tout $j \in J$ et tout groupe fini $H_j \subset M_j[\ell^\infty]$, il existe $w_j = w_j(H_j)$ tels qu'on a, uniformément en (ℓ, H_j),*

$$K(H_j) \cap K(\mu_{\ell^\infty}) \asymp K(\mu_{\ell^{w_j}}).$$

(2) *Uniformément en ℓ, on a l'identité*

$$\mathrm{Gal}(K(A[\ell^\infty])/K(\mu_{\ell^\infty})) \asymp \prod_{j\in J} \mathrm{Gal}(K(M_j[\ell^\infty]))/K(\mu_{\ell^\infty})).$$

Alors si $w := \max w_j$, pour tout groupe fini $H = \prod_j H_j \subset A[\ell^\infty]$, uniformément en (ℓ, H) on a, $K(H)\cap K(\mu_{\ell^\infty}) \asymp K(\mu_{\ell^w})$ et

$$[K(H) : K(\mu_{\ell^w})] \asymp \prod_{j\in J}[K(H_j) : K(\mu_{\ell^{w_j}})].$$

Nous allons appliquer ceci avec l'ensemble $J = \mathfrak{I}_\ell$, et pour $j = (\lambda, i) \in \mathfrak{I}_\ell$, avec $M_j = T_{\lambda,i}(A_i)$, ainsi que $\alpha_j = n_i$ si A_i est de type I et $\alpha_j = 2n_i$ si A_i est de type II. Enfin nous l'utiliserons avec $H_j = H_{\lambda,i}$.

Par le lemme 7.2 on a

$$[K(H_{\lambda,i}) : K] \asymp \delta(H_{\lambda,i})(\mathrm{Hdg}(A_i)(\mathbb{Z}_\ell) : G(H_{\lambda,i})).$$

Or on sait que dans notre situation on a uniformément en (ℓ, H) :

$$\mathrm{Gal}(K(A_i[\ell^\infty])/K(\mu_{\ell^\infty})) \asymp \prod_{(\lambda,i)\in\mathfrak{I}_\ell} \mathrm{Gal}(K(T_{\lambda,i}[\lambda^\infty])/K(\mu_{\ell^\infty})).$$

On peut appliquer la proposition 7.3 et on obtient, uniformément en (ℓ, H),

$$[K(H) : K] \asymp \delta(H) \prod_{(\lambda,i)\in\mathfrak{I}_\ell} (\mathrm{Sp}_{2h_i}(\mathcal{O}_{\lambda,i}) : G(H_{\lambda,i})).$$

Notons $\mathrm{cd}^k_{(\lambda,i)}$ la codimension du groupe algébrique $G_{k,(\lambda,i)}$. Dans la situation d'un $H_{\lambda,i}$ fixé nous avions introduit au paragraphe précédent des notations

$$m_j \text{ et } a_j,\ 1 \le j \le 2h_i, \quad \text{ainsi que} \quad m^r,\ 1 \le r \le t_{\lambda,i}.$$

Afin de rendre claire les diverses dépendances nous utiliserons ci-dessous les notations un peu plus lourdes suivantes en lieu et place des précédentes :

$$m_j(\lambda, i) \text{ et } a_j(\lambda, i),\ 1 \le j \le 2h_i, \quad \text{ainsi que} \quad m^r_{\lambda,i},\ 1 \le r \le t_{\lambda,i}.$$

Les calculs effectués dans le cas d'un $H_{\lambda,i}$ nous donnent, uniformément en (ℓ, H),

$$\big(\mathrm{Hdg}(A)(\mathbb{Z}_\ell) : G(H)\big) \asymp \exp\Bigg(\sum_{(\lambda,i)\in\mathfrak{I}_\ell}\sum_{k=1}^{t_{\lambda,i}} f(\lambda)\, \mathrm{cd}^k_{(\lambda,i)}(m_{\lambda,i}^{t_{\lambda,i}+1-k} - m_{\lambda,i}^{t_{\lambda,i}+1-(k-1)})\log\ell\Bigg).$$

De plus, il existe un (λ_1, i_1) tel que $\delta(H) = \delta(H_{\lambda_1,i_1})$. Quitte à renuméroter on peut supposer que $i_1 = 1$. On note alors $(\delta(\lambda_1)_j)_j$ la suite de 0 et de 1 relative à $\delta(H_{\lambda_1,1})$

définie au paragraphe précédent. On a, uniformément en (ℓ, H),

$$\delta(H) \gg \exp\left(\sum_{j=1}^{t_{\lambda_1,1}} (m_{\lambda_1,1}^{t_{\lambda_1,1}+1-j} - m_{\lambda_1,1}^{t_{\lambda_1,1}+1-(j-1)})\delta(\lambda_1)_j \log \ell\right).$$

On pose par ailleurs $\delta(\lambda)_j = 0$ pour tout j si $(\lambda, i) \neq (\lambda_1, 1)$. Avec ces notations, on trouve en suivant les calculs du cas d'un $H_{\lambda,i}$, la minoration suivante (au sens $\gg$, uniformément en (ℓ, H)) pour $[K(H) : K]$:

$$\exp\Bigg(\sum_{(\lambda,i)\in\mathfrak{I}_\ell} \sum_{j=1}^{t_{\lambda,i}} m_{\lambda,i}^j \Big[\big(\delta(\lambda)_{t_{\lambda,i}+1-j} - \delta(\lambda)_{t_{\lambda,i}+1-(j-1)}\big) + f(\lambda)\big(\mathrm{cd}_{(\lambda,i)}^{t_{\lambda,i}+1-j} - \mathrm{cd}_{(\lambda,i)}^{t_{\lambda,i}+1-(j-1)}\big)\Big] \log \ell \Bigg),$$

et

$$|H| = \exp\left(\sum_{(\lambda,i)\in\mathfrak{I}_\ell} n_i d_i \sum_{j=1}^{t_{\lambda,i}} m_{\lambda,i}^j f(\lambda) a_j(\lambda, i) \log \ell \right),$$

où d_i vaut 1 (respectivement 2) si A_i est de type I (respectivement de type II) et où l'on rappelle que $A = \prod_{i=1}^d A_i^{n_i}$.

Notons

$$b_{\lambda,i}^j := (\delta(\lambda)_{t_{\lambda,i}+1-j} - \delta(\lambda)_{t_{\lambda,i}+1-(j-1)}) + f(\lambda)(\mathrm{cd}_{\lambda,i}^{t_{\lambda,i}+1-j} - \mathrm{cd}_{\lambda,i}^{t_{\lambda,i}+1-(j-1)}),$$

et posons de plus

$$a_{\lambda,i}^j := n_i d_i a_j(\lambda, i).$$

Avec ces notations, on aura donc, uniformément en (ℓ, H), l'inégalité $|H| \ll [K(H) : K]^\gamma$ si

$$\gamma \geq \max \frac{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} \sum_{j=1}^{t_{\lambda,i}} m_{\lambda,i}^j f(\lambda) a_{\lambda,i}^j}{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} \sum_{j=1}^{t_{\lambda,i}} m_{\lambda,i}^j b_{\lambda,i}^j},$$

le max étant pris sur les $m_{\lambda,i}^1 \geq \cdots \geq m_{\lambda,i}^{t_{\lambda,i}}$ pour $(\lambda, i) \in \mathfrak{I}_\ell$.

Ainsi, en invoquant le lemme combinatoire (lemme 2.7) et en suivant les notations et calculs du cas d'un $H_{\lambda,i}$, on voit que l'inégalité $|H| \ll [K(H) : K]^\gamma$ est vraie uniformément en (ℓ, H), si

$$\gamma \geq \max \frac{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} n_i d_i f(\lambda)(r(\lambda, i)_{t_{\lambda,i}+1-h_{\lambda,i}} + s(\lambda, i)_{t_{\lambda,i}+1-h_{\lambda,i}})}{\delta(\lambda_1)_{t_{\lambda_1,1}+1-h_{\lambda_1,1}} + \sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda) \operatorname{codim} P_{r(\lambda,i)_{t_{\lambda,i}+1-h_{\lambda,i}}, s(\lambda,i)_{t_{\lambda,i}+1-h_{\lambda,i}}}}.$$

Ce dernier max se réécrit sous la forme

$$\max_{\substack{1\le r_{\lambda,i}\\ 0\le s_{\lambda,i}\le r_{\lambda,i}\le h_i}} \frac{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} n_i d_i f(\lambda)(r_{\lambda,i}+s_{\lambda,i})}{\delta+\sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda)(r_{\lambda,i}+s_{\lambda,i})\big(2h_i-\frac{1}{2}(r_{\lambda,i}+s_{\lambda,i}-1)\big)},$$

et où en reprenant la définition de $\delta(\lambda_1)_{t_{\lambda_1,1}+1-h_{\lambda_1,1}}$ on voit que $\delta=0$ si tout les $s_{\lambda,i}$ sont nuls et $\delta=1$ si l'un des $s_{\lambda,i}$ est non nul.

Il y a en fait deux évaluations à faire selon que $\delta=1$ ou que $\delta=0$.

1. Si $\delta=1$ alors le max à évaluer se réécrit naturellement sous la forme suivante :

$$\max_{1\le x_{\lambda,i}\le 2h_i} \frac{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} m_i d_i f(\lambda)x_{\lambda,i}}{1+\sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda)x_{\lambda,i}\big(2h_i-\frac{1}{2}(x_{\lambda,i}-1)\big)}.$$

Posons

$$\rho_1(\underline{x}) := \frac{\sum_{(\lambda,i)\in\mathfrak{I}_\ell} m_i d_i f(\lambda)x_{\lambda,i}}{1+\sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda)x_{\lambda,i}\big(2h_i-\frac{1}{2}(x_{\lambda,i}-1)\big)}.$$

On rappelle que l'on veut comparer $\rho_1(\underline{x})$ avec la quantité

$$\alpha(A) := \max_{I\subset\{1,\dots,r\}} \frac{2\sum_{i\in I} m_i g_i}{1+\sum_{i\in I} 2e_i h_i^2+e_i h_i}.$$

Lemme 7.4. *Pour tout $i\in\{1,\dots,r\}$, on a $\alpha(A)\ge (m_i d_i)/(h_i+1)$.*

Démonstration. C'est un calcul immédiat. □

Lemme 7.5. *On a $\max_{1\le x_{\lambda,i}\le 2h_i}\rho_1(\underline{x})\le\alpha(A)$.*

Démonstration. L'inégalité $\rho_1(\underline{x})\le\alpha(A)$ se réécrit

$$\sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda)\Big[x_{\lambda,i}^2-\Big(4h_i+1-\frac{2m_i d_i}{\alpha(A)}\Big)x_{\lambda,i}\Big]\le 2.$$

Par le lemme précédent, on voit dans la somme dans le membre de gauche de l'inégalité, que les indices tels que $x_{\lambda,i}\le 2h_i-1$ contribuent via un terme négatif à la somme. Autrement dit, la valeur $\rho_1(\underline{x})-\alpha(A)$ est maximale quand pour tout les indices on a $x_{\lambda,i}=2h_i$. Mais dans ce cas, en utilisant que

$$\sum_{\lambda \text{ place de } \mathrm{End}(A_i)\otimes\mathbb{Z}_\ell} f(\lambda)=e_i,$$

on a

$$\rho_1(2h_i,\ldots,2h_i)-\alpha(A)\le 0 \iff \sum_{(\lambda,i)\in\mathfrak{I}_\ell} f(\lambda)\Big[4h_i^2-\Big(4h_i+1-\frac{2m_id_i}{\alpha(A)}\Big)2h_i\Big]\le 2$$
$$\iff \sum_{i=1}^{r}\Big(-4h_i^2-2h_i+\frac{4m_ih_id_i}{\alpha(A)}\Big)\sum_{\lambda} f(\lambda)\le 2$$
$$\iff \sum_{i=1}^{r}\Big(-4h_i^2e_i-2h_ie_i+\frac{4m_ih_ie_id_i}{\alpha(A)}\Big)\le 2$$
$$\iff \frac{2\dim A}{\dim \mathrm{MT}(A)}\le \alpha(A).$$

La dernière assertion de la série d'équivalences est vraie, ce qui conclut. □

2. Si $\delta=0$ alors dans ce cas un calcul du même type (plus facile) permet également de conclure.

8. Petites valeurs de ℓ exceptionnelles

Dans ce paragraphe nous indiquons quelles modifications apporter pour les valeurs exceptionnelles de ℓ (en nombre fini). On se place dans le cadre d'une variété abélienne géométriquement simple A/K de type I ou II et pleinement de type Lefschetz, telle que $\mathrm{End}_K(A)=\mathrm{End}_{\overline{K}}(A)$ et telle que $\mathrm{End}_K(A)$ est un ordre maximal de $D:=\mathrm{End}_K(A)\otimes_{\mathbb{Z}}\mathbb{Q}$. Nous notons E le centre de D et nous notons enfin $\phi: A\to A^\vee$ une polarisation fixée avec A (les diverses constantes intervenant dépendant de A, dépendent aussi du degré de cette polarisation).

Notons que, dans le cas où nous nous sommes placés (pleinement de type Lefschetz), on sait que la conjecture de Mumford–Tate est vraie, donc que l'on a l'inclusion suivante qui est une égalité à indice fini près (dépendant éventuellement de ℓ mais peu importe ici car on travaille uniquement avec un nombre fini de valeurs de ℓ) :

$$\rho_{\lambda^\infty}(G_K)\subset\{x\in \mathrm{GSp}_{2h}(\mathcal{O}_\lambda)\mid \mathrm{mult}(x)\in\mathbb{Z}_\ell^\times\}.$$

Nous indiquons dans ce qui suit les petites modifications à faire pour pouvoir traiter les ℓ qui sont dans l'ensemble fini exceptionnel $S_{\mathrm{ex}}(A)$ introduit dans la définition 3.4.

Si ℓ est ramifié dans $\mathcal{O}_E$. On suppose dans ce paragraphe que ℓ ne divise pas $\deg(\phi)$ et, dans le cas de type II, que ℓ est tel que l'algèbre de quaternions D est décomposée en $\lambda|\ell$. On suppose par contre que ℓ est ramifié dans $\mathcal{O}_E$:

$$\ell\mathcal{O}_E=\prod_{\lambda|\ell}\lambda^{e(\lambda)}.$$

Nous notons toujours $f(\lambda)$ le degré du corps résiduel en la place λ. Le lemme 3.3 produit l'accouplement $\phi^{\star}_{\ell^\infty}$ sur $T_\ell(A) \times T_\ell(A)$ à valeurs dans $\mathcal{O}^{\star}_\ell$.

Hypothèse. Supposons pour l'instant pour simplifier que $\phi^{\star}_{\ell^\infty}$ est en fait à valeurs dans $\mathcal{O}_\ell$. Nous verrons plus loin comment faire dans le cas général.

Rappelons que dans cette situation on a la décomposition

$$T_\ell(A) = \begin{cases} \prod_{\lambda|\ell} T_\lambda(A) & \text{(Type I)}, \\ \prod_{\lambda|\ell} T_\lambda(A) \oplus T_\lambda(A) & \text{(Type II)}. \end{cases}$$

Par réduction modulo λ^n, on obtient alors pour tout entier $+\infty \geq n \geq 1$, comme dans le cas non ramifié,

$$\phi_{\lambda^n} : T_\lambda(A)/\lambda^n T_\lambda(A) \times T_\lambda(A)/\lambda^n T_\lambda(A) \to O_\lambda/\lambda^n(1).$$

Notons par ailleurs que $\ell O_\lambda = \lambda^{e(\lambda)}$, donc par réduction modulo ℓ^n on a

$$A[\ell^n] = T_\ell(A)/\ell^n T_\ell(A) = \begin{cases} \prod_{\lambda|\ell} T_\lambda[\lambda^{e(\lambda)n}] & \text{(Type I)}, \\ \prod_{\lambda|\ell} T_\lambda[\lambda^{e(\lambda)n}] \oplus T_\lambda[\lambda^{e(\lambda)n}] & \text{(Type II)} \end{cases}$$

et

$$\phi_{\lambda^{e(\lambda)n}} : T_\lambda[\lambda^{e(\lambda)n}] \times T_\lambda[\lambda^{e(\lambda)n}] \to \mathcal{O}_\lambda/\ell^n \mathcal{O}_\lambda(1).$$

De plus on vérifie que

$$\phi_{\lambda^{e(\lambda)n}}(\ell x, \ell y) = \phi_{\lambda^{e(\lambda)(n+1)}}(x, y)^\ell$$

et on voit que l'action de Galois sur les accouplements $\phi_{\lambda^{e(\lambda)n}}$ se fait via le caractère cyclotomique χ_{ℓ^n}. Par ailleurs, le $\mathbb{Z}_\ell$-module $\mathcal{O}_\lambda$ étant libre de rang $e(\lambda)f(\lambda)$, on a

$$\begin{aligned} O_\lambda/\ell^n \mathcal{O}_\lambda &\simeq (\mathbb{Z}/\ell^n\mathbb{Z})^{e(\lambda)f(\lambda)}, \\ T_\lambda[\lambda^{e(\lambda)n}] \simeq (O_\lambda/\ell^n\mathcal{O}_\lambda)^{2h} &\simeq (\mathbb{Z}/\ell^n\mathbb{Z})^{2he(\lambda)f(\lambda)}. \end{aligned} \tag{4}$$

Finalement en travaillant en lieu de place de ϕ_{λ^n} avec $\phi_{\lambda^{e(\lambda)n}}$, en décomposant H selon les $H_\lambda \subset T_\lambda[\lambda^{e(\lambda)n}] \subset A[\ell^n]$, on peut reprendre tout ce qui a été fait dans le cas non ramifié. Notamment la propriété μ pour les H_λ est toujours vérifiée avec la modification évidente suivante dans la définition de $m_1(H_\lambda)$ (et de même pour $m(H_\lambda)$) : on pose

$$m_1(H_\lambda) := \max\left\{m \mid \exists P, Q \in H_\lambda \text{ d'ordre } \ell^k \text{ tels que } \phi_{\lambda^{e(\lambda)k}}(P, Q) \text{ est d'ordre } \ell^m\right\}.$$

Pour prouver la propriété μ dans ce cadre on peut toujours utiliser le paragraphe sur les modules isotropes sur $(\mathcal{O}_\lambda, \mathbb{F}_\lambda)$, puisque les accouplements λ-adiques ϕ_{λ^r} sont construits modulo λ^r pour tout entier $r \geq 1$ et non pas uniquement modulo ℓ^r (cette réduction modulo ℓ^r étant celle utilisée pour définir les H_λ).

En reprenant mot pour mot les calculs combinatoires déjà effectués, on voit avec (4) que les calculs restent inchangés sous réserve de remplacer partout $f(\lambda)$

par $f(\lambda)e(\lambda)$. La contrainte $\sum f(\lambda) = [E : \mathbb{Q}]$ étant remplacée par la contrainte $\sum e(\lambda) f(\lambda) = [E : \mathbb{Q}]$, on voit que la valeur γ reste la même dans ce cadre, ce que l'on voulait prouver.

Suppression de l'hypothèse. On ne suppose désormais plus que $\phi^{\star}_{\ell^\infty}$ est à valeurs dans $\mathcal{O}_\ell$.

On sait que $\mathcal{O}^{\star}_\lambda$ est un idéal fractionnaire de E_λ, donc de la forme $\pi_\lambda^{-m_\lambda}\mathcal{O}_\lambda$ pour un certain entier m_λ (avec π_λ une uniformisante). On choisit

$$m_0 := \text{pgcd}\{m_\lambda \mid \lambda|\ell,\ \ell \text{ ramifié dans } \mathcal{O}_E\}.$$

On a ainsi $\ell^{m_0}\mathcal{O}^{\star}_\ell \subset \mathcal{O}_\ell$ pour tout les ℓ que l'on considère. On fait alors les modifications suivantes :

(1) On remplace au départ le module $T_\ell(A)$ par $T'_\ell := \ell^{m_0} T_\ell(A)$. Dans ce cas, le lemme 3.3 produit l'acccouplement $\phi^{\star}_{\ell^\infty}$ sur $T'_\ell \times T'_\ell$ à valeurs dans $\mathcal{O}_\ell$ (c'est pour arriver dans $\mathcal{O}_\ell$ que l'on a remplacé $T_\ell(A)$ par T'_ℓ).

(2) On travaille avec $H' = \ell^{m_0} H$ en lieu et place de H. La raison de cette modification est que le groupe H n'est a priori pas contenu dans la réduction modulo ℓ^n de T'_ℓ. Par contre H' l'est.

À la déperdition près d'indice en ℓ^{2hm_0} près, on peut reprendre la preuve déjà effectuée et on obtient, uniformément en H :

$$|H'| \ll [K(H') : K]^{\gamma(A)} \leq [K(H) : K]^{\gamma(A)}.$$

De plus, on a visiblement $|H| \leq |H'| \cdot |A[\ell^{m_0}]|$. Les premiers problématiques ℓ étant en nombre bornés, il en est de même pour le cardinal des divers $A[\ell^{m_0}]$ et on voit donc que l'on obtient ainsi, uniformément en H,

$$|H| \ll [K(H) : K]^{\gamma(A)}.$$

Si ℓ divise le degré de la polarisation. On suppose désormais que ℓ est un premier quelconque divisant $\deg(\phi)$ et, dans le cas de type II, que ℓ est tel que l'algèbre de quaternions D est décomposée en $\lambda|\ell$. Notons m_0 l'entier maximal tel que ℓ^{m_0} divise $\deg(\phi)$. Dans ce cas, toute les constructions faites jusqu'à présent continuent encore à s'appliquer à condition de faire au départ les modifications suivantes :

(1) On travaille avec $T'_\ell(A) := \ell^{m_0} T_\ell(A)$ en lieu et place de $T_\ell(A)$.

(2) On travaille avec l'accouplement $\phi^{\phi}_\ell : T'_\ell(A) \times T'_\ell(A) \to T'_\ell(A) \times T'_\ell(A^\vee) \to \mathbb{Z}_\ell(1)$ défini par $(x, y) \mapsto \phi_\ell(x, \phi(y))$, en lieu et place de l'accouplement ϕ_ℓ. Ce choix ainsi que le point précédent sont faits de sorte à avoir un accouplement bilinéaire alterné sur $T'_\ell(A)$, non dégénéré modulo ℓ^n pour tout $n \geq 1$.

(3) On travaille avec $H' = \ell^{m_0} H$ en lieu et place de H. La raison de cette modification est que le groupe H n'est a priori pas contenu dans la réduction modulo ℓ^n de $T'_\ell(A)$. Par contre H' l'est.

Avec ces modifications on peut reprendre la preuve déjà effectuée et on obtient, uniformément en H :

$$|H'| \ll [K(H') : K]^{\gamma(A)} \leq [K(H) : K]^{\gamma(A)}.$$

De plus, on a visiblement $|H| \leq |H'| + |A[\ell^{m_0}]|$. Les premiers ℓ étant en nombre fini, le cardinal des divers $A[\ell^{m_0}]$ est donc borné et l'on obtient ainsi uniformément en H,

$$|H| \ll [K(H) : K]^{\gamma(A)}.$$

Petites valeurs exceptionnelles pour le type II. Dans le cas d'une variété abélienne de type II, il y a encore un nombre fini de valeurs ℓ exceptionnelles à traiter : les premiers ℓ tels que l'algèbre D est non décomposée en $\lambda|\ell$. Ce cas des ℓ ramifiés dans l'algèbre de quaternions doit être traité avec une légère modification : la décomposition $V_\ell(A) = \prod_\lambda (W_\lambda(A) \oplus W_\lambda(A))$ n'existant que après tensorisation par une extension quadratique (voir proposition 3.5). Avec les notations du début du section 7 : au lieu d'avoir sur $\mathcal{O}_{\lambda,i}$ la décomposition $X_{\lambda,i} = H_{\lambda,i} \oplus H_{\lambda,i}$, on a, en travaillant sur une extension quadratique de $\mathcal{O}_{\lambda,i}$, la décomposition $X_{\lambda,i} = H_{\lambda,i} \oplus \overline{H}_{\lambda,i}$, où $\overline{H}_{\lambda,i}$ est conjugué à $H_{\lambda,i}$. Ainsi, au lieu de comparer le cardinal de $H_{\lambda,i}$ avec le degré de l'extension $[K(H_{\lambda,i}) : K]$, on reprend la même preuve en travaillant directement avec $X_{\lambda,i}$, comparant le cardinal de $X_{\lambda,i}$ et le degré de $[K(X_{\lambda,i}) : K]$.

9. Ordre d'un point et degré de l'extension qu'il engendre

Nous donnons dans ce paragraphe la preuve du théorème 1.12. Nous pouvons pour cela supposer (et nous le faisons) que tous les $\overline{K}$-endomorphismes de A sont définis sur K.

Soit P un point de torsion et H_P le End(A)-module engendré par P. On a clairement $K(P) = K(H_P)$. En remarquant que codim $P_{1,0} = 2h$ (dans Sp_{2h}), les arguments des paragraphes 6 et 7 précédents montrent qu'un point P_λ d'ordre ℓ^n dans $T_\lambda[\lambda^\infty]$ engendre (uniformément en (ℓ, P)) une extension de degré

$$[K(P_\lambda) : K] \gg \ell^{2hn}.$$

Si ensuite $P = \sum_{\lambda|\ell} P_\lambda$ avec P_λ point de $T_\lambda[\lambda^\infty]$ et d'ordre ℓ^{n_λ}, de sorte que P est d'ordre ℓ^n avec $n = \max n_\lambda$, alors, uniformément en (ℓ, P) on a

$$[K(P) : K] \gg \ell^{2h \sum_\lambda n_\lambda} \geq \ell^{2hn}.$$

Enfin si P est d'ordre m quelconque avec $m=\prod_{i=1}^r \ell_i^{n_i}$, on peut écrire $P=\sum_{i=1}^r P_i$ avec P_i d'ordre $\ell_i^{n_i}$. L'indépendance des représentations ℓ-adiques (proposition 3.2) permet d'écrire, uniformément en (ℓ, P),

$$[K(P):K]=[K(P_1,\ldots,P_r):K]\gg\prod_{i=1}^r[K(P_i):K]\geq\prod_{i=1}^r c_1\ell^{2hn_i}=c_1^{\omega(m)}m^{2h}.$$

10. Appendice : compléments autour de la conjecture de Mumford–Tate

Indice de l'image de Galois dans le groupe de Mumford–Tate. La conjecture de Mumford–Tate dit que l'inclusion $G^0_{\ell^\infty}\subset \mathrm{MT}\otimes\mathbb{Q}_\ell$ est une égalité, ou encore que, quitte a avoir effectué une extension finie du corps de base K, l'image de la représentation ℓ-adique galoisienne $\rho(G_K)$ est contenue et ouverte dans $\mathrm{MT}(\mathbb{Q}_\ell)$, ou encore comme $\mathrm{GL}(T_\ell(A))\cong\mathrm{GL}_{2g}(\mathbb{Z}_\ell)$ est compact, la conjecture équivaut à dire que, quitte a avoir effectué une extension finie du corps de base K, $\rho(G_K)$ est d'indice fini dans $\mathrm{MT}(\mathbb{Z}_\ell)$. Une forme légèrement plus forte, suggérée par Serre, affirme que cet indice est borné indépendamment de ℓ. Clarifions tout d'abord ce point en montrant que la conjecture de Mumford–Tate entraîne la forme "forte".

Théorème 10.1. *Si A/K vérifie la conjecture de Mumford–Tate alors l'indice de $\rho(G_K)$ dans $\mathrm{MT}(A)(\mathbb{Z}_\ell)$ est borné indépendamment de ℓ.*

La preuve consiste à réunir un résultat de Serre [1986a] (resp. de Wintenberger [2002]) concernant la partie torique centrale (resp. la partie semi-simple) des groupes ℓ-adiques et des groupes de Mumford–Tate. Plus précisément, notons $S=S_A$ le groupe dérivé du groupe de Mumford–Tate de A ou, ce qui revient au même, du groupe de Hodge et notons C la composante neutre du centre du groupe $\mathrm{MT}(A)$; ce sont des $\mathbb{Q}$-groupes algébriques. Notons similairement $S_\ell=S_{\ell,A}$ le groupe dérivé de $G_{\ell,A}$ ou, ce qui revient au même, du groupe $H_{\ell,A}$ et notons C_ℓ la composante neutre du centre du groupe $G_{\ell,A}$; ce sont des $\mathbb{Q}_\ell$-groupes algébriques.

On sait par les travaux de Borovoĭ [1974], Deligne [1982, Exposé I, 2.9, 2.11], et Pjateckiĭ-Šapiro [1971] que

$$C_\ell\subset C_{\mathbb{Q}_\ell}\quad\text{et}\quad S_\ell\subset S_{\mathbb{Q}_\ell}. \tag{5}$$

En fait on sait même que la première inclusion est une égalité, essentiellement d'après la théorie abélienne de Serre [1998], une preuve est détaillée dans [Vasiu 2008] et reprise dans [Ullmo et Yafaev 2013]. On ne sait pas, en général si la deuxième inclusion est une égalité, en fait l'égalité est équivalente à la conjecture de Mumford–Tate. D'après Faltings, les deux groupes réductifs ont le même commutant, leur égalité est aussi équivalente à l'égalité des *rangs* des deux groupes semi-simples.

Posons $V_\ell(A) = T_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$. Les groupes C et S sont des sous-groupes du groupe de Mumford–Tate $\mathrm{MT} = \mathrm{MT}_A$. En voyant $\mathrm{MT}_{\mathbb{Q}_\ell}$ comme un sous-groupe algébrique de $\mathrm{GL}_{V_\ell(A)} \cong \mathrm{GL}_{2g,\mathbb{Q}_\ell}$, on peut étendre ces groupes sur $\mathbb{Z}_\ell$ en prenant leur adhérence de Zariski dans $\mathrm{GL}_{T_\ell(A)}$. Avec un léger abus de notation nous noterons $C(\mathbb{Z}_\ell)$ (resp. $S(\mathbb{Z}_\ell)$) le groupe des $\mathbb{Z}_\ell$-points de C (resp. S) vu comme groupe algébrique sur $\mathbb{Z}_\ell$. Le même procédé nous permet d'étendre C_ℓ (resp. S_ℓ) en un groupe sur $\mathbb{Z}_\ell$.

Concernant la partie torique centrale, nous savons donc que $\rho_{\ell^\infty}(G_K) \cap C(\mathbb{Z}_\ell)$ est d'indice fini dans $C(\mathbb{Z}_\ell)$. Le résultat suivant de Serre précise ce point et est un des deux points clef pour la preuve du théorème 10.1.

Proposition 10.2 [Serre 1986a, p. 60]. *L'indice $\big(C(\mathbb{Z}_\ell) : C_\ell(\mathbb{Z}_\ell) \cap \rho_{\ell^\infty}(G_K)\big) =: c_\ell$ est fini borné indépendamment de ℓ.*

Notons que la preuve de Serre [1986a] est rédigée dans le cas $\mathrm{End}(A) = \mathbb{Z}$ et esquissée dans le cas général. Pour la commodité du lecteur nous donnons ci-dessous une description du centre et de la relation avec l'image de Galois.

Concernant la partie semi-simple S_ℓ de G_ℓ, suivant Wintenberger [2002], notons $S_{\ell,\mathrm{sc}} \to S_\ell$ le revêtement universel de S_ℓ (sur $\mathbb{Z}_\ell$) et posons

$$S_\ell(R)_u \text{ l'image de } S_{\ell,\mathrm{sc}}(R) \text{ dans } S_\ell(R), \text{ pour } R \in \{\mathbb{Z}_\ell, \mathbb{Q}_\ell, \mathbb{F}_\ell\}.$$

Si $\ell \geq 5$ alors le groupe $S_\ell(\mathbb{F}_\ell)_u$ est le sous-groupe de $S_\ell(\mathbb{F}_\ell)$ engendré par les éléments unipotents. C'est également le groupe des commutateurs de $S_\ell(\mathbb{F}_\ell)$. Le point clef que nous utilisons peut s'énoncer ainsi.

Proposition 10.3 [Wintenberger 2002]. *L'indice $S_\ell(\mathbb{Z}_\ell)_u$ dans $S_\ell(\mathbb{Z}_\ell)$ est borné indépendamment de ℓ. Pour tout premier ℓ assez grand, le groupe $S_\ell(\mathbb{Z}_\ell)_u$ est contenu dans G_ℓ. En particulier, on a les inclusions $S_\ell(\mathbb{Z}_\ell)_u \subset G_\ell \cap S_\ell(\mathbb{Z}_\ell) \subset S_\ell(\mathbb{Z}_\ell)$ avec indices finis, bornés indépendamment de ℓ.*

Cet énoncé découle de l'énoncé plus précis suivant.

Proposition 10.4 [Wintenberger 2002]. *On a l'égalité $S_\ell(\mathbb{Z}_\ell)_u = S_\ell(\mathbb{Z}_\ell) \cap S_\ell(\mathbb{Q}_\ell)_u$. De plus, si le centre $Z(S_{\ell,\mathrm{sc}})$ est de cardinal premier à ℓ alors $S_\ell(\mathbb{Z}_\ell)_u$ est l'image réciproque de $S_\ell(\mathbb{F}_\ell)_u$ par $\pi_\ell : S_\ell(\mathbb{Z}_\ell) \to S_\ell(\mathbb{F}_\ell)$, le morphisme de réduction modulo ℓ. Pour tout premier ℓ assez grand, le groupe $S_\ell(\mathbb{Z}_\ell)_u$ est contenu dans G_ℓ. Si ℓ est assez grand, alors l'indice $(S_\ell(\mathbb{Z}_\ell) : S_\ell(\mathbb{Z}_\ell)_u)$ est majoré par $c(2\dim A) :=$* $\mathrm{ppcm}\{n \mid n \leq 2\dim A\}$.

La démonstration du théorème 10.1 est maintenant immédiate à partir des propositions 10.2 et 10.3. Comme S (resp. C) est le groupe dérivé (resp. la composante neutre du centre) de $\mathrm{MT}(A)$ on a $\mathrm{MT}(A) = C \cdot S$ et on en tire aisément que $(\mathrm{MT}(A)(\mathbb{Z}_\ell) : C(\mathbb{Z}_\ell) \cdot S(\mathbb{Z}_\ell))$ est borné indépendamment de ℓ. La proposition 10.2 fournit un sous-groupe C_1 de $C(\mathbb{Z}_\ell) \cap \rho_{\ell^\infty}(G_K)$ d'indice fini dans

$C(\mathbb{Z}_\ell)$, tandis que la proposition 10.3, jointe à l'hypothèse $S_{\mathbb{Q}_\ell} = S_\ell$ fournit un sous-groupe S_1 de $S(\mathbb{Z}_\ell) \cap \rho_{\ell^\infty}(G_K)$ d'indice fini dans $S(\mathbb{Z}_\ell)$. On conclut bien alors que $(MT(A)(\mathbb{Z}_\ell) : \rho(G_K)) \leq (MT(A)(\mathbb{Z}_\ell) : C_1 \cdot S_1)$ est borné indépendamment de ℓ.

Donnons maintenant la description promise du centre du groupe de Mumford–Tate.

Notons $L = \prod_i L_i$ le centre de $\mathrm{End}^0(A)$; chaque L_i est un corps de nombres et la décomposition correspond à la décomposition de A à isogénie près en composantes isotypiques, i.e., $A \cong \prod_i A_i$ avec $A_i = B_i^{m_i}$ et B_i absolument simple. On pose aussi $T_L = \prod_i \mathrm{Res}_{L_i/\mathbb{Q}}(\mathbb{G}_{m,L_i})$ et on note $\det_{L_i} : \mathrm{Res}_{L_i/\mathbb{Q}}(\mathrm{GL}_{V_i,L_i}) \to \mathrm{Res}_{L_i/\mathbb{Q}}(\mathbb{G}_{m,L_i})$ et $\det_L = \prod_i \det_{L_i}$. La restriction de $\det_L$ au tore T_L donne une isogénie $\delta : T_L \to T_L$ qui peut être explicitée comme l'application $x = (x_i)_{i \in I} \mapsto (x_i^{d_i})_{i \in I}$, où $d_i := 2 \dim A_i / [L_i : \mathbb{Q}]$. Introduisons une extension auxiliaire $\tilde{L}$ finie, galoisienne sur $\mathbb{Q}$ et contenant les L_i ; on définit ensuite la “norme” (cf. [Ichikawa 1991, p. 135]) :

$$\Psi_i : \mathrm{Res}_{\tilde{L}/\mathbb{Q}} \mathbb{G}_{m,\tilde{L}} \to \mathrm{Res}_{L_i/\mathbb{Q}} \mathbb{G}_{m,L_i},$$
$$\Psi = \prod_i \Psi_i : \mathrm{Res}_{\tilde{L}/\mathbb{Q}} \mathbb{G}_{m,\tilde{L}} \to \prod_i \mathrm{Res}_{L_i/\mathbb{Q}} \mathbb{G}_{m,L_i}.$$

On a alors une description de la composante neutre du centre du groupe de Mumford–Tate comme le sous-tore de T_L vérifiant $\delta(C) = \mathrm{Im}\,\Psi$ (aux notations près, c'est la proposition 1.2.1 de [Ichikawa 1991]).

Proposition 10.5 (cf. [Ichikawa 1991; Serre 1986a]). *Le tore $C = C_A$ est le sous-tore de T_L tel que $\delta(C) = \Psi(T_L)$.*

Le lien avec les représentations ℓ-adiques peut être décrit ainsi (cf. [Serre 1986b]).

Chaque morceau $V_\ell(A_i)$ est libre de rang d_i sur $L_i \otimes \mathbb{Q}_\ell$; si l'on pose $L_\ell := L \otimes \mathbb{Q}_\ell$, alors $V_\ell(A)$ est un L_ℓ-module et on peut définir $\det_{L_\ell}(V_\ell(A))$ qui est libre de rang 1, ce qui fournit une représentation ℓ-adique abélienne à valeurs dans T_L :

$$\phi_\ell : G_K \to L_\ell^\times = (L \otimes \mathbb{Q}_\ell)^\times = T_L(\mathbb{Q}_\ell).$$

Par la théorie abélienne de Serre il existe un module $\mathfrak{m}$ et un homomorphisme de groupes algébriques associé $S_\mathfrak{m} \to T_L$ induisant la représentation de la manière suivante ; rappelons (cf. [Serre 1998]) que $S_\mathfrak{m}$ est un $\mathbb{Q}$-groupe algébrique extension du groupe fini $C_\mathfrak{m}$ des classes d'idèles modulo $\mathfrak{m}$ par un tore $T_\mathfrak{m}$. Si on note $\varepsilon_\ell : G_K \to S_\mathfrak{m}(\mathbb{Q}_\ell)$ la représentation définie dans [Serre 1998], alors $\phi_\ell = \phi \circ \varepsilon_\ell$.

Le lien avec le centre du groupe de Hodge est que $\Psi(T_L)$ est la composante neutre de $\phi(S_\mathfrak{m})$. Ce fait appelé “exercice embêtant” dans [Serre 1986b] est équivalent à l'égalité de la composante connexe des centres de G_ℓ et $\mathrm{MT}(A)_{\mathbb{Q}_\ell}$ citée ci-dessus.

Quelques cas de la conjecture de Mumford–Tate. Pour énoncer le premier résultat en vue, nous rappelons la notation suivante, cf. (1) :

$$\Sigma = \left\{ g \geq 1 \,\middle|\, \exists k \geq 3,\ \text{impair}\ , \exists a \geq 1,\ 2g = (2a)^k \text{ ou } 2g = \binom{2k}{k} \right\}. \tag{6}$$

Théorème 10.6. *Soit h un entier tel que $h \notin \Sigma$ et soit A/K une variété abélienne de type II telle que le centre de* $\mathrm{End}(A) \otimes \mathbb{Q}$ *est réduit à* $\mathbb{Q}$. *On suppose que A est de dimension $g = 2h$. La conjecture de Mumford–Tate est vraie pour A et le groupe de Mumford–Tate associé à A est* $\mathrm{GSp}_{2h,\mathbb{Q}}$.

Démonstration. Il s'agit tout simplement d'appliquer la proposition 4.7 de Pink [1998]. Précisément : soit ℓ un nombre premier suffisamment grand ; décomposons le module de Tate V_ℓ en somme de 2 copies isomorphes W_ℓ sur $\mathbb{Q}_\ell$. Les W_ℓ donnent des représentations de dimension $2h$, de type Mumford–Tate fort (puisque c'est le cas par [Pink 1998, theorem 5.10] pour les V_ℓ), fidèles, symplectiques, absolument irréductibles du groupe dérivé de G_ℓ. On peut donc appliquer la proposition 4.7 de Pink qui nous dit que si h est en dehors de l'ensemble exceptionnel Σ alors G_ℓ est isomorphe à $\mathrm{GSp}_{2h,\mathbb{Q}}$. □

Pour énoncer le résultat suivant, rappelons que si une variété abélienne A a réduction semi-stable en une place v, la composante neutre de la fibre spéciale est une extension d'une variété abélienne B_0 par un tore que nous noterons T_0 ; la dimension de ce tore s'appelle la *dimension torique*. De plus, si ce tore est non trivial (i.e., s'il y a mauvaise réduction) l'anneau des endomorphismes de A agit fidèlement sur ce tore. En conséquence, si A est de type I et $e = [\mathrm{End}^0(A) : \mathbb{Q}]$ (resp. de type II et $4e = [\mathrm{End}^0(A) : \mathbb{Q}]$) alors la dimension de T_0 est un multiple de e (resp. de $2e$).

Théorème 10.7. *Soit A une variété abélienne simple de type I ou II, définie sur un corps de nombres K. On note e le degré sur $\mathbb{Q}$ du centre de* $\mathrm{End}^0(A)$ *et $d := 1$, si A est de type I, et $d := 2$ si A est de type II. Supposons de plus que :*

(T) *Il existe une place de K où A possède réduction semi-stable de dimension torique de.*

Alors A est pleinement de type Lefschetz.

Commençons par rappeler quelques résultats de Grothendieck décrits dans [SGA 7_{I} 1972]. On étudie la réduction modulo un idéal premier donc on peut passer à un anneau local que l'on peut d'ailleurs compléter et qu'on notera $\mathcal{O}$; dans le langage de [SGA 7_{I} 1972], le spectre de $\mathcal{O}$ est un *trait*. Tout schéma quasi-fini $X/\mathcal{O}$ se décompose en

$$X = X^{\mathrm{f}} \sqcup X',$$

où X^{f} est un schéma fini sur $\mathcal{O}$ et X' est un schéma égal à sa fibre générique. On choisit un nombre premier ℓ assez grand pour qu'il ne divise pas le cardinal du groupe des composantes de la fibre spéciale du modèle de Néron sur $\mathcal{O}$ de A. En décomposant ainsi le schéma quasi-fini $A[\ell^n]$ et en prenant la limite on obtient la *partie fixe* $T_\ell(A)^{\mathrm{f}} \subset T_\ell(A)$. En considérant la partie torique de la fibre spéciale A_0^0 du modèle de Néron

$$0 \to T_0 \to A_0^0 \to B_0 \to 0,$$

on obtient la *partie torique* $T_\ell(A)^{\mathrm{t}} \subset T_\ell(A)^{\mathrm{f}}$, qui est de rang $\mu := \dim T_0$ (i.e., égal à la dimension torique). Le résultat clef de Grothendieck décrit ces sous-modules en termes de l'action du groupe d'inertie, noté I et s'énonce ainsi (le premier point est la proposition 2.2.5 de [SGA 7_{I} 1972], le deuxième le théorème 2.4 de [SGA 7_{I} 1972]).

Théorème 10.8 [SGA 7_{I} 1972]. *Avec les notations précédentes, on a :*

(1) $T_\ell(A)^{\mathrm{f}} = T_\ell(A)^I$.

(2) *Soit $\check{A}$ la duale de A, désignons par $\perp$ l'orthogonal au sens de l'accouplement canonique de Weil $T_\ell(A) \times T_\ell(\check{A}) \to T_\ell(\mathbb{G}_m)$. Alors*

$$T_\ell(A)^{\mathrm{t}} = T_\ell(A)^{\mathrm{f}} \cap (T_\ell(\check{A})^{\mathrm{f}})^\perp = T_\ell(A)^I \cap (T_\ell(\check{A})^I)^\perp.$$

De plus, si on note $U = V_\ell(A) = T_\ell(A) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$, $U_1 = V_\ell(A)^{\mathrm{f}}$ et $U_2 = V_\ell(A)^{\mathrm{t}}$, et si l'on identifie $V_\ell(A)$ et $V_\ell(\check{A})$ via une polarisation fixée, on a donc $U_2 = U_1 \cap U_1^\perp$. L'inertie agit de plus trivialement sur U_2 et sur U/U_2 (mais non trivialement sur U si l'on suppose qu'il y a mauvaise réduction).

Démonstration du théorème 10.7. Notons comme précédemment E le centre de $\mathrm{End}^0(A)$. Il suffit de montrer que, pour un ℓ, le groupe H_ℓ est le produit des $\mathrm{Sp}_{2h,E_\lambda}$, quand λ parcourt les idéaux premiers de E au dessus de ℓ. On choisit un ℓ assez grand et totalement décomposé dans $E/\mathbb{Q}$; d'après le lemme de relèvement (lemme 2.6), il suffit de voir que l'image de Galois modulo ℓ contient le produit des $\mathrm{Sp}_{2h,\mathbb{F}_\ell}$. Chris Hall [2011] montre dans le cas où $\mathrm{End}(A) = \mathbb{Z}$ (i.e., $e = d = 1$) que l'hypothèse (T) entraîne ceci. Nous expliquons comment adapter ses arguments au cas plus général.

Considérons les sous-modules décrits ci-dessus $U_2 \subset U_1 \subset V_\ell(A)$ et rappelons que $\mu = \mathrm{rang}\, U_2$ est la dimension torique. L'inertie opère non trivialement sur $V_\ell(A)$ puisqu'il y a mauvaise réduction mais trivialement sur U_2 et $V_\ell(A)/U_2$. Si g_I désigne un générateur topologique du quotient maximal pro-ℓ-fini de I, sa matrice $\rho_\ell(g_I)$ est donc conjuguée à une matrice $\begin{pmatrix} I_\mu & B \\ 0 & I_{2g-\mu} \end{pmatrix}$.

On peut "découper" les U_i, qui sont des $\mathbb{Q}_\ell$-espaces vectoriels, en introduisant e_λ l'idempotent de $E_\ell = E \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ qui projette E_ℓ sur E_λ et en notant $U_{i,\lambda}$ l'image $e_\lambda U_i$. On obtient alors la décomposition cherchée.

Lemme 10.9. *Les sous-modules galoisiens U_1 "partie fixe" et U_2 "partie torique" sont stables par* $\mathrm{End}(A)$ *et se décomposent ainsi :*

(1) (*Type I*) *Les U_i se décomposent en $U_i = \prod_\lambda U_{i,\lambda}$.*

(2) (*Type II*) *Les* U_i *se décomposent en* $U_i = \prod_\lambda (U_{i,\lambda} \oplus U_{i,\lambda})$.

Démonstration (du lemme). L'action des endomorphismes commute avec celle du groupe de Galois et en particulier du groupe d'inertie, ce qui entraîne la première affirmation. De plus, pour tout endomorphisme α, on a

$$< \alpha v, v' >=< v, \alpha^\dagger v' >,$$

où $< \cdot, \cdot >$ désigne l'accouplement de Weil et † l'involution de Rosati associée à la polarisation choisie ; ainsi les décompositions de la représentation galoisienne décrites dans la proposition 3.5 induisent celles sur U_1 et U_2. □

L'hypothèse (T) se traduit en disant que $de = \operatorname{rang} U_2 = d \sum_\lambda \operatorname{rang} U_{i,\lambda}$, ce qui impose $\operatorname{rang} U_{i,\lambda} = 1$. Si l'on écrit maintenant la matrice de $\rho(g_I)$ par blocs, on voit que chaque bloc est une transvection. Les arguments de [Hall 2011] s'appliquent alors, permettant de montrer que chaque bloc de la représentation modulo ℓ contient dans son image $\mathrm{Sp}_{2h}(\mathbb{F}_\ell)$, ce qui achève la preuve du théorème 10.7. □

Théorème 10.10. *Soit* $s \geq 1$ *un entier et soient* $A_1, \dots, A_s$ *des variétés abéliennes, deux à deux non isogènes, chaque* A_i *étant pleinement de type Lefschetz, de type I ou II, de dimension relative un entier* h_i. *On a dans ce cas*

$$\mathrm{Hdg}\Big(\prod_{i=1}^{s} A_i\Big) = \prod_{i=1}^{s} \mathrm{Hdg}(A_i) = \prod_{i=1}^{s} \mathrm{Res}_{E_i/\mathbb{Q}} \mathrm{Sp}_{2h_i,E_i},$$

$$\textit{pour tout } \ell, \quad H_\ell\Big(\prod_{i=1}^{s} A_i\Big) = \prod_{i=1}^{s} H_\ell(A_i) = \prod_{i=1}^{s} \prod_{\lambda_i \mid \ell} \mathrm{Sp}_{2h_i,E_{\lambda_i}},$$

où λ_i *parcourt les places de* E_i *au dessus de* ℓ.

Démonstration. Nous expliquons la preuve dans le cadre ℓ-adique qui s'appuie sur l'article de Lombardo [2016] (le cas complexe est plus simple et s'appuie de manière parallèle sur l'article antérieur d'Ichikawa [1991]). Il s'agit en fait d'appliquer le théorème 4.1 de [Lombardo 2016]. Pour cela il nous suffit de vérifier que les hypothèses de ce théorème sont satisfaites. Nos hypothèses entraînent que pour tout entier i, on a $\mathrm{Hdg}(A_i) \times \mathbb{C}_\ell = \mathrm{Sp}_{2h_i,\mathbb{C}_\ell}^{[E_i:\mathbb{Q}]}$, où E_i est le centre de $\mathrm{End}(A_i) \otimes \mathbb{Q}$. Or les automorphismes de (l'algèbre de Lie de) $\mathrm{Sp}_{2h_i,\mathbb{C}_\ell}$ sont intérieurs et les automorphismes intérieurs préservent les plus haut poids (cf. remarque 3.8 de [Lombardo 2016]), donc le point (3) des hypothèses du théorème 4.1 de [Lombardo 2016] est vérifié. Les points (1) et (2) sont immédiats dans notre situation, d'où la conclusion. □

Corollaire 10.11. *Soit* $s \geq 1$ *un entier et soient* $A_1, \dots, A_s$ *des variétés abéliennes, deux à deux non isogènes, chaque* A_i *étant de type I ou II, de dimension relative un entier* h_i. *Pour chaque* i, *on suppose que l'une des hypothèses suivantes est vérifiée* :

(1) *Le centre de* $\mathrm{End}(A)\otimes\mathbb{Q}$ *est réduit à* $\mathbb{Q}$ *et l'entier* h_i *n'est pas dans l'ensemble exceptionnel* Σ.

(2) *L'entier* h_i *est égal à deux ou est impair.*

(3) *La variété abélienne* A_i *est de type I* (*resp. II*) *et possède une place de mauvaise réduction semi-stable avec dimension torique* e_i (*resp. avec dimension torique* $2e_i$).

Sous ces hypothèses on a alors

$$\mathrm{Hdg}\Big(\prod_{i=1}^{s} A_i\Big)=\prod_{i=1}^{s}\mathrm{Hdg}(A_i)=\prod_{i=1}^{s}\mathrm{Res}_{E_i/\mathbb{Q}}\,\mathrm{Sp}_{2h_i,E_i},$$

$$\textit{pour tout } \ell,\qquad H_\ell\Big(\prod_{i=1}^{s} A_i\Big)=\prod_{i=1}^{s} H_\ell(A_i)=\prod_{i=1}^{s}\prod_{\lambda_i\mid\ell}\mathrm{Sp}_{2h_i,E_{\lambda_i}},$$

où λ_i *parcourt les places de* E_i *au dessus de* ℓ.

Démonstration. Soit i un entier dans l'ensemble $\{1,\ldots,s\}$. Si l'hypothèse (1) est vérifiée alors par le théorème 5.14 de [Pink 1998] (dans le cas de type I) ou par le théorème 10.6 ci-dessus (dans le cas de type II) on en déduit que A est pleinement de type Lefschetz, de groupe de Hodge, Sp_{2h_i}. Si l'hypothèse (2) est vérifiée, là encore la même conclusion vaut en appliquant cette fois le théorème A de [Banaszak et al. 2006] (cf. [Lombardo 2016, remarque 2.25] pour le cas de dimension relative 2). Il suffit donc d'appliquer le théorème 10.10 ci-dessus pour conclure. □

Remarque 10.12. On peut bien sûr énoncer le théorème et le corollaire précédents en remplaçant $\prod_{i=1}^{s} A_i$ par $\prod_{i=1}^{s} A_i^{m_i}$ puisque $\mathrm{Hdg}\big(\prod_{i=1}^{s} A_i^{m_i}\big)=\mathrm{Hdg}\big(\prod_{i=1}^{s} A_i\big)$ et $H_\ell\big(\prod_{i=1}^{s} A_i^{m_i}\big)=H_\ell\big(\prod_{i=1}^{s} A_i\big)$.

Remerciements

Les auteurs remercient le rapporteur, dont la lecture attentive et les corrections et suggestions judicieuses ont permis d'améliorer la présentation de ce texte. Une partie de ces recherches a bénéficié du soutien de l'ANR "HAMOT".

Bibliographie

[Banaszak et al. 2006] G. Banaszak, W. Gajda et P. Krasoń, "On the image of l-adic Galois representations for abelian varieties of type I and II", *Doc. Math.* Extra Vol. (2006), 35–75. MR 2290584

[Borovoĭ 1974] M. V. Borovoĭ, "О действии группы Галуа на рациональные классы когомологий (p, p) абелевых многообразий", *Mat. Sb.* (*N.S.*) **94 (136)**:4 (1974), 649–652. Traduction anglaise : "On the Galois action on rational cohomology classes of type $(p,\ p)$ of abelian varieties,", *Math. USSR Sb.* **23** :4, 613–616. MR 0352101

[Breuer 2010] F. Breuer, "Torsion bounds for elliptic curves and Drinfeld modules", *J. Number Theory* **130**:5 (2010), 1241–1250. MR 2607312 Zbl 1227.11072

[Chi 1990] W. C. Chi, “On the Tate modules of absolutely simple abelian varieties of type II”, *Bull. Inst. Math. Acad. Sinica* **18**:2 (1990), 85–95. MR 1071226 Zbl 0729.14033

[Chi 1992] W. C. Chi, “l-adic and λ-adic representations associated to abelian varieties defined over number fields”, *Amer. J. Math.* **114**:2 (1992), 315–353. MR 1156568 Zbl 0795.14024

[Deligne et al. 1982] P. Deligne, J. S. Milne, A. Ogus et K.-y. Shih, *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Math. **900**, Springer, Berlin, 1982. MR 654325 Zbl 0465.00010

[Hall 2011] C. Hall, “An open-image theorem for a general class of abelian varieties”, *Bull. Lond. Math. Soc.* **43**:4 (2011), 703–711. MR 2820155 Zbl 1225.11083

[Hindry et Ratazzi 2010] M. Hindry et N. Ratazzi, “Torsion dans un produit de courbes elliptiques”, *J. Ramanujan Math. Soc.* **25**:1 (2010), 81–111. MR 2643390 Zbl 1206.11075

[Hindry et Ratazzi 2012] M. Hindry et N. Ratazzi, “Points de torsion sur les variétés abéliennes de type GSp”, *J. Inst. Math. Jussieu* **11**:1 (2012), 27–65. MR 2862374 Zbl 1311.11050

[Ichikawa 1991] T. Ichikawa, “Algebraic groups associated with abelian varieties”, *Math. Ann.* **289**:1 (1991), 133–142. MR 1087242 Zbl 0697.14031

[Larsen et Pink 1995] M. Larsen et R. Pink, “Abelian varieties, l-adic representations, and l-independence”, *Math. Ann.* **302**:3 (1995), 561–579. MR 1339927 Zbl 0867.14019

[Lombardo 2016] D. Lombardo, “On the ℓ-adic Galois representations attached to nonsimple Abelian varieties”, *Ann. Inst. Fourier* (*Grenoble*) **66**:3 (2016), 1217–1245. MR 3494170

[Masser 1986] D. Masser, lettre à Daniel Bertrand, 10 novembre 1986.

[Milne 1999] J. S. Milne, “Lefschetz classes on abelian varieties”, *Duke Math. J.* **96**:3 (1999), 639–675. MR 1671217 Zbl 0976.14009

[Moonen 1999] B. Moonen, “Notes on Mumford–Tate groups”, lecture notes, 1999, Voir http://www.math.ru.nl/~bmoonen/Lecturenotes/CEBnotesMT.pdf.

[Murty 1984] V. K. Murty, “Exceptional Hodge classes on certain abelian varieties”, *Math. Ann.* **268**:2 (1984), 197–206. MR 744607 Zbl 0521.14004

[Noot 1995] R. Noot, “Abelian varieties—Galois representation and properties of ordinary reduction”, *Compositio Math.* **97**:1-2 (1995), 161–171. MR 1355123 Zbl 0868.14021

[Piatetski-Shapiro 1971] I. I. Piatetski-Shapiro, “Взаимоотношения между гипотезами Тайта и Ходжа для абелевых многообразий”, *Mat. Sb.* (*N.S.*) **85 (127)** (1971), 610–620. Traduction anglaise : “Interrelations between the Tate and Hodge conjectures for abelian varieties”, *Math. USSR Sb.* **14** :4, 615–624. MR 0294352

[Pink 1998] R. Pink, “l-adic algebraic monodromy groups, cocharacters, and the Mumford–Tate conjecture”, *J. Reine Angew. Math.* **495** (1998), 187–237. MR 1603865

[Ratazzi 2007] N. Ratazzi, “Borne sur la torsion dans les variétés abéliennes de type CM”, *Ann. Sci. École Norm. Sup.* (4) **40**:6 (2007), 951–983. MR 2419854 Zbl 1140.14041

[Ribet 1976] K. A. Ribet, “Galois action on division points of Abelian varieties with real multiplications”, *Amer. J. Math.* **98**:3 (1976), 751–804. MR 0457455 Zbl 0348.14022

[Serre 1986a] J.-P. Serre, lettre à Ken Ribet, 7 mars 1986. §138 dans ses *Œuvres, IV, 1985–1998*, Springer, Berlin, 2000.

[Serre 1986b] J.-P. Serre, lettre à Marie-France Vignéras, 10 février 1986. §137 dans ses *Œuvres, IV, 1985–1998*, Springer, Berlin, 2000.

[Serre 1986c] J.-P. Serre, "Résumé des cours au Collège de France de 1985–1986", 1986. §136 dans ses *Œuvres, IV : 1985–1998*, Springer, Berlin, 2000.

[Serre 1998] J.-P. Serre, *Abelian l-adic representations and elliptic curves*, Research Notes in Mathematics **7**, A K Peters, Wellesley, MA, 1998. Revised reprint of the 1968 original. MR 1484415 Zbl 0902.14016

[Serre 2013] J.-P. Serre, "Un critère d'indépendance pour une famille de représentations ℓ-adiques", *Comment. Math. Helv.* **88**:3 (2013), 541–554. MR 3093502 Zbl 1317.14040

[SGA 7_{I} 1972] A. Grothendieck, *Groupes de monodromie en géométrie algébrique, I: Exposés I–II, VI–IX* (Séminaire de Géométrie Algébrique du Bois Marie 1967–1969), Lecture Notes in Math. **288**, Springer, Berlin, 1972. MR 0354656 Zbl 0237.00013

[Ullmo et Yafaev 2013] E. Ullmo et A. Yafaev, "Mumford–Tate and generalised Shafarevich conjectures", *Ann. Math. Qué.* **37**:2 (2013), 255–284. MR 3117743 Zbl 06350059

[Vasiu 2008] A. Vasiu, "Some cases of the Mumford–Tate conjecture and Shimura varieties", *Indiana Univ. Math. J.* **57**:1 (2008), 1–75. MR 2400251 Zbl 1173.11039

[Wintenberger 2002] J.-P. Wintenberger, "Démonstration d'une conjecture de Lang dans des cas particuliers", *J. Reine Angew. Math.* **553** (2002), 1–16. MR 1944805 Zbl 1048.14008

Communicated by Joseph H. Silverman
Received 2015-05-25 Revised 2016-07-14 Accepted 2016-08-12

marc.hindry@imj-prg.fr *Institut de Mathématiques Jussieu – Paris Rive Gauche, Université Paris Diderot, Bâtiment Sophie Germain, 8 place Aurélie Nemours, 75013 Paris, France*

nicolas.ratazzi@math.u-psud.fr *Equipe d'Arithmétique et Géométrie Algébrique, Université Paris-Sud, Bâtiment 425, 91405 Orsay, France*

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1893

Galois-generic points on Shimura varieties

Anna Cadoret and Arno Kret

We discuss existence and abundance of Galois-generic points for adelic representations attached to Shimura varieties. First, we show that, for Shimura varieties of abelian type, ℓ-Galois-generic points are Galois-generic; in particular, adelic representations attached to such Shimura varieties admit ("lots of") closed Galois-generic points. Next, we investigate further the distribution of Galois-generic points and show the André–Pink conjecture for them: if S is a connected Shimura variety associated to a $\mathbb{Q}$-simple reductive group, then every infinite subset of the generalized Hecke orbit of a Galois-generic point is Zariski-dense in S. Our proof follows the approach of Pink for Siegel Shimura varieties. Our main contribution consists in showing that there are only finitely many Hecke operators of bounded degree on (adelic and connected) Shimura varieties. Compared with other approaches of this result, our proof, which relies on Bruhat–Tits theory, is effective and works for arbitrary Shimura varieties.

1. Introduction

Given a smooth, separated and geometrically connected scheme S over a field k and a point $s \in S$, let $\sigma_s : \pi_1(s) \to \pi_1(S)$ denote the morphism induced by functoriality of the étale fundamental group.[1] Given an algebraic group G over $\mathbb{Q}$ and an adelic representation $\rho : \pi_1(S) \to G(\mathbb{A}_f)$, let $\rho_\ell : \pi_1(S) \to G(\mathbb{A}_f) \to G(\mathbb{Q}_\ell)$ denote its ℓ-adic component. We say that $s \in S$ is *Galois-generic* with respect to $\rho : \pi_1(S) \to G(\mathbb{A}_f)$ if the image of $\rho \circ \sigma_s$ is open in the image of ρ, and *ℓ-Galois-generic* if the image of $\rho_\ell \circ \sigma_s$ is open in the image of ρ_ℓ.

To a Shimura datum (G, X) and a neat compact open subgroup $K_0 \subset G(\mathbb{A}_f)$, we can attach a representation $\rho_{K_0} : \pi_1(S[K_0]) \to K_0 \subset G(\mathbb{A}_f)$ of the étale fundamental group, where $S[K_0] \subset \mathrm{Sh}_{K_0}(G, X)$ is a geometrically connected component (defined over a finite extension $E[K_0]$ of the reflex field $E = E(G, X)$ of (G, X)). This representation group-theoretically encodes the tower of étale covers $\mathrm{Sh}_K(G, X) \to \mathrm{Sh}_{K_0}(G, X)$ indexed by open subgroups $K \subset K_0$. For a

MSC2010: primary 11G18; secondary 20G35, 14F20.
Keywords: Shimura varieties, Hecke orbits, Adelic representations of étale fundamental group, Galois generic points.

[1] Recall that $\pi_1(s)$ identifies with the absolute Galois group of the residue field $k(s)$ at s.

point $s[K_0] \in S[K_0]$ and a field extension F of $E[K_0]$, we say that the induced point $s[K_0]_F \in S[K_0]_F$ is Galois-generic if it is Galois-generic with respect to $\rho_{K_0}|_{\pi_1(S[K_0]_F)} : \pi_1(S[K_0]_F) \to G(\mathbb{A}_f)$, and ℓ-Galois-generic if it is ℓ-Galois-generic with respect to $\rho[K_0]_\ell|_{\pi_1(S[K_0]_F)} : \pi_1(S[K_0]_F) \to G(\mathbb{Q}_\ell)$. We say that a point s in $\mathrm{Sh}(G, X)_F$ is (ℓ-)Galois-generic if for some (equivalently, any) neat compact open subgroup $K_0 \subset G(\mathbb{A}_f)$, the image $s[K_0]$ of s in $\mathrm{Sh}_{K_0}(G, X)_F$ is (ℓ-)Galois-generic.

In the context of Shimura varieties, the terminology "Galois-generic" was introduced by Pink [2005, Definition 6.3]. The definition of Pink does not resort to the formalism of étale fundamental groups and is seemingly stronger than ours. Namely, if E^{ab} denotes the maximal abelian extension of E, a point $s[K_0] \in S[K_0]$ is Galois-generic in the sense of Pink if and only if the induced point $s[K_0]_{E^{\mathrm{ab}}} \in S[K_0]_{E^{\mathrm{ab}}}$ is Galois-generic in our sense. However, using the facts that $\rho(\pi_1(S[K_0]_{\overline{E}})) = \Gamma_0^-$, where $\Gamma_0^- \subset G(\mathbb{A}_f)$ denotes the adelic closure of $\Gamma_0 := G(\mathbb{Q}) \cap K$ in $G(\mathbb{A}_f)$, and that every open subgroup of Γ_0^- has finite abelianization (Theorem 5.4), we show in Proposition 6.1.1 that the two definitions coïncide.

1.1. ***Existence.*** Given a scheme S smooth, separated and geometrically connected over a field k, and an adelic representation $\rho : \pi_1(S) \to G(\mathbb{A}_f)$, the first question which arises is whether there exists Galois-generic points (other than the generic point) with respect to ρ. While ℓ-adic specialization techniques give rise to "lots of" closed ℓ-Galois-generic points (see 3.3.1), they fail to ensure the existence of a single closed point which is ℓ-Galois-generic for every prime ℓ (hence a fortiori which is Galois-generic).

However, for adelic representations attached to motives, the ℓ-adic Tate conjectures say that a point that is ℓ-Galois-generic for one prime ℓ is ℓ-Galois-generic for every prime ℓ, and the modulo-ℓ variant of the Tate conjectures even predict that a point which is ℓ-Galois-generic for one prime ℓ is Galois-generic (see 3.3.2).

By works of Faltings (e.g., [Faltings and Wüstholz 1984]), partial forms of the modulo-ℓ variant of the Tate conjectures are available for abelian varieties; this is enough to ensure that for adelic representations attached to the Tate module of abelian schemes, a point which is ℓ-Galois-generic for one prime ℓ is Galois-generic [Cadoret 2015, Theorem 1.2] (see Theorem 3.3.2.2). The first main result of this paper is the extension of this statement to adelic representations attached to Shimura varieties of abelian type (see 6.3.2 for the definition of "abelian type").

Theorem A. *Assume (G, X) is a Shimura datum of abelian type. Then a point $s \in \mathrm{Sh}(G, X)$ is ℓ-Galois-generic if and only if it is Galois-generic.*

The bridge between [Cadoret 2015, Theorem 1.2] and adelic representations attached to Shimura varieties is provided by the moduli description of Siegel Shimura varieties. The remaining parts of the argument rely on the general machinery of Shimura varieties and group-theoretical arguments. Our approach fails to handle the case of Shimura data (G, X) which are not of abelian type, though it seems reasonable to expect that Theorem A should also hold for such representations.

Theorem A and the description of Galois-generic points in terms of adelic representations are also used in [Cadoret and Moonen 2015] to prove that, for motives parametrized by Shimura varieties of abelian type (e.g., abelian varieties, $K3$ surfaces) the integral and adelic variants of the Mumford–Tate conjecture follow from the standard (ℓ-adic) Mumford–Tate conjecture.

1.2. ***Equidistribution and the André–Pink conjecture.*** Theorem A implies that adelic representations attached to Shimura varieties of abelian type admit "lots of" Galois-generic points since they admit "lots of" ℓ-Galois-generic points. For instance, combining Theorem A and the abundance result for ℓ-Galois-generic points of [Cadoret and Tamagawa 2013] (see Fact 3.3.1.2), on can show the following. Say that an irreducible curve $C \hookrightarrow S[K_0]$ is Galois-generic if its generic point is. Then the set of irreducible Galois-generic curves defined over a number field is Zariski-dense in $S[K_0]$, and for each such curve $C \hookrightarrow S[K_0]$, defined over a finite extension E_C of $E[K_0]$ and integer $d \geq 1$, all but finitely many closed points $t \in C$ with $[k(t) : E_C] \leq d$ are Galois-generic. In particular, closed Galois-generic points are Zariski-dense, which is not surprising once their existence is proved: being Galois-generic is preserved by Hecke operators and Hecke orbits are Zariski dense. But the restrictions on the degree show more. Indeed, it follows from the definition of Galois-generic points and the fact (see Theorem 7.2.2) that there are only finitely many Hecke operators of bounded degree on a connected Shimura variety, that for every Galois-generic point $t \in S[K_0]$ and integer $d \geq 1$ there are only finitely many t' in the Hecke orbit of t with $[k(t') : E] \leq d$. Thus there are infinitely many Hecke orbits of closed Galois-generic points on Shimura varieties of abelian type, and even infinitely many Hecke orbits of closed Galois-generic points intersecting a Galois-generic curve defined over a number field.

Using equidistribution techniques, we can strengthen these results as follows. Let (G, X) be a Shimura datum (which we no longer assume to be of abelian type). Let $K_0 \subset G(\mathbb{A}_f)$ be a neat compact open subgroup and let $X^+ \subset X$ be a connected component. Write $\Gamma_0 := K_0 \cap G(\mathbb{Q})_+$, where $G(\mathbb{Q})_+ \subset G(\mathbb{Q})$ denotes the stabilizer of X^+. Eventually, let $S[K_0] =: \mathrm{Sh}_{\Gamma_0}(G, X^+) \subset \mathrm{Sh}_{K_0}(G, X)$ denote the geometrically connected component containing the image of $X^+ \times \{1\}$ (that is, $\mathrm{Sh}_{\Gamma_0}(G, X^+)_{\mathbb{C}} \simeq \Gamma_0 \setminus X^+$). Write $\mathrm{Aut}(G, X^+)$ for the group of automorphisms of G defined over $\mathbb{Q}$ and stabilizing X^+. For every $S[K_0] \in \mathrm{Sh}_{\Gamma_0}(G, X^+)$, write

$$\widehat{T}_{\Gamma_0}(s[K_0]) := \bigcup_{\phi \in \mathrm{Aut}(G, X^+)} T_{\Gamma_0,\phi}(s[K_0])$$

for the (full) generalized Hecke orbit of $s[K_0]$, where $T_{\Gamma_0,\phi}$ denotes the generalized Hecke operator induced by ϕ on $\mathrm{Sh}_{\Gamma_0}(G, X^+)$ (see Section 7.1).

Theorem B (André–Pink conjecture for Galois-generic points [André 1989, Chapter X, Problem 3; Pink 2005, Conjecture 1.6; Orr 2013, Conjecture 1.3]). *Assume G is almost $\mathbb{Q}$-simple. Then for every Galois-generic point $s[K_0] \in \mathrm{Sh}_{\Gamma_0}(G, X^+)$, every infinite subset of $\widehat{T}_{\Gamma_0}(s[K_0])$ is Zariski-dense in $\mathrm{Sh}_{\Gamma_0}(G, X^+)$.*

For Shimura varieties of abelian type a consequence of Theorem A, Theorem B and [Cadoret and Tamagawa 2013] is that if $C \hookrightarrow S$ is an irreducible Galois-generic curve defined over a number field, then C is cut by infinitely many Hecke orbits of closed Galois-generic points, and each of these Hecke orbits cuts C in only finitely many points.

Theorem B extends a previous result of Pink [2005, Theorem 7.6] for the Siegel Shimura varieties; our proof follows that of Pink but with some technical adjustments required to deal with non-simply connected groups G. More precisely, the main ingredient in Pink's argument is an equidistribution result of Clozel, Oh and Ullmo [Clozel et al. 2001, Theorem 1.6, Remark (3)] for GSp_{2g}. To deal with arbitrary Shimura varieties, one needs a generalization of [Clozel et al. 2001, Theorem 1.6] for adjoint groups G and arithmetic (not only congruence) subgroups $\Gamma \subset G(\mathbb{Q})$. Such a generalization was proved by Eskin and Oh [2006, Theorem 1.2] following an idea of Burger and Sarnak [1991] (see Section 7.2). To apply Eskin and Oh's result to our situation, we have to ensure that, for an arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, there are only finitely many Hecke operators T_a, $a \in G(\mathbb{Q})$ with $|\Gamma \backslash \Gamma a \Gamma|$ bounded (see Theorem 7.2.2). We provide a proof of this result in Section 8, proceeding in two steps: First, we prove Theorem 8.2.1, the adelic variant of Theorem 7.2.2; here the "natural" tools are avatars of the Bruhat–Tits decomposition, which give explicit estimates for the local degrees (see Section 8.2 for details). Then, we deduce Theorem 7.2.2 from this adelic variant by reducing to the simply connected case, where we can apply strong approximation.

After a first version of this paper was released, we were informed by Hee Oh that Theorem 7.2.2 could also be proved by equidistribution arguments like those used in [Eskin and Oh 2006], but this proof does not seem to be effective, nor does it work for adelic Hecke operators (see Section 9.1 for details).

If we restrict to connected Shimura varieties of abelian type, Theorem B can easily be recovered from Orr's thesis [2013, Theorem 1.5(ii)] arguing as follows. Let $s[K_0] \in \mathrm{Sh}_{\Gamma_0}(G, X^+)$ be a Galois-generic point, let A be an infinite subset of $\widehat{T}_{\Gamma_0}(s[K_0])$, and let Z be some irreducible component of the Zariski-closure of A containing $s[K_0]$. Since $s[K_0]$ is Galois-generic, it is Hodge-generic (see Proposition 6.2.1). Hence, the smallest special subvariety of $\mathrm{Sh}_{\Gamma_0}(G, X^+)$ containing $s[K_0]$ is equal to the smallest special subvariety of $\mathrm{Sh}_{\Gamma_0}(G, X^+)$ containing Z (because in both cases, this subvariety has to be $\mathrm{Sh}_{\Gamma_0}(G, X^+)$ itself). By construction, $A \cap Z$ is Zariski-dense in Z; hence, from [Orr 2013, Theorem 1.5(ii)], Z is weakly special. But as Z contains the Hodge-generic point $s[K_0]$, this forces

$Z = \mathrm{Sh}_{\Gamma_0}(G, X^+)$. Orr's approach to Theorem B relies on different techniques than ours (Masser–Wüstholz isogeny bound for abelian varieties and o-minimality); it is more general since it works for Hodge-generic points but, due to the use of the Masser–Wüstholz isogeny bound, it only applies to Shimura varieties of abelian type. For connected Shimura varieties of abelian type, one can also give a proof of Theorem 7.2.2 based on the Masser–Wüstholz isogeny bound and the existence of closed Galois-generic points (see Section 9.2 for details).

In Section 3, we review the notion of Galois-generic points attached to adelic representations of the étale fundamental group and some of their basic properties. We also recall the main existence and abundance results for ℓ-Galois-generic points and discuss in more details the relation between ℓ-Galois generic points and Galois-generic points for motivic representations. In Section 4, we construct the adelic representations attached to Shimura varieties and review some of their basic properties. Section 5 is technical and gathers the group-theoretical results about the adelic closure of arithmetic subgroups of semisimple groups that we need to prove Theorem A. In Section 6, we focus on Galois-generic points attached to Shimura varieties, show that our definition coïncides with the one of Pink, that Galois-generic-points are Hodge generic, and complete the proof of Theorem A. Section 7 is devoted to the proof of Theorem B. The proof of Theorem 7.2.2 is postponed to Section 8 as it can be read independently of the rest of the paper and involves techniques of a different nature. In the final Section 9, we discuss alternative approaches to Theorem 7.2.2: a noneffective approach based on equidistribution (see Section 9.1) and an effective approach (limited to connected Shimura varieties of abelian type) based on the Masser–Wüstholz isogeny theorem (see Section 9.2).

2. Notation and conventions

The fields in this paper, when of characteristic 0, will always be assumed to be embedded into the field $\mathbb{C}$ of complex numbers. For such fields, compositum, Galois, abelian, algebraic closures, etc., will always mean with respect to the given embedding into $\mathbb{C}$.

Given schemes S and T over a field k, unless there is a risk of confusion, we write $S_T := S \times_k T$ (that is, we omit the notation for the base field k). When $T = \mathrm{spec}(F)$ for a field extension $k \subset F$, we write $S_F := S_{\mathrm{spec}(F)}$. However, when $S =: s = \mathrm{spec}(E)$ for a field extension $k \subset E$, and $k \subset F$ is another field extension with E, F embedded into $\mathbb{C}$, we write $s_F := \mathrm{spec}(EF)$; that is, we implicitly pick the connected component of $s \times_{\mathrm{spec}(k)} \mathrm{spec}(F)$ corresponding to the given embeddings of E, F in $\mathbb{C}$.

Given a scheme S of finite type over a field k and a point $s \in S$, we write $k(s)$ for the residue field (a finitely generated extension of k) of S at s. We identify points $s \in S$ and the corresponding morphisms of k-schemes $s : \mathrm{spec}(k(s)) \to S$. For a point

$s \in S$, a geometric point above s is a morphism $\bar{s} : \mathrm{spec}(\Omega) \to S$ factorizing through $s : \mathrm{spec}(k(s)) \to S$ and such that Ω is an algebraically closed field. In general, we do not specify the algebraically closed field Ω in the notation for geometric points (see below) and for a point $s \in S$, unless otherwise specified $\bar{s}$ will always denote a geometric point over s. For every $s \in S$ and geometric point $\bar{s}$ over s, let $F_{\bar{s}}$ denote the associated fiber functor from étale covers of S to sets. Recall that, by definition, the étale fundamental group of S based at $\bar{s}$ is the automorphism group $\pi_1(S, \bar{s})$ of $F_{\bar{s}}$ and that, if S is connected, for every $s, s' \in S$ there always exists isomorphisms of fiber functors $\alpha : F_{\bar{s}} \tilde{\to} F_{\bar{s}'}$, and the set of such isomorphisms is a $\pi_1(S, \bar{s})$-torsor. In particular, for every étale cover $X \to S$, α yields a bijection $\alpha : F_{\bar{s}}(X) \tilde{\to} F_{\bar{s}'}(X)$ which is equivariant with respect to the isomorphism of étale fundamental groups

$$\pi_1(S, \bar{s}) \tilde{\to} \pi_1(S, \bar{s}'), \quad \sigma \mapsto \alpha\sigma\alpha^{-1}.$$

Thus, unless it helps to understand the situation, we will omit the base-point $\bar{s}$ in our notation for étale fundamental groups. Given a field k, we often shorten $\pi_1(\mathrm{spec}(k))$ in $\pi_1(k)$, which identifies with the absolute Galois group of k. For a point $s \in S$, we also write $\pi_1(s) = \pi_1(k(s))$.

For an algebraic group G, we let $G^{\mathrm{der}} \subset G$ denote its derived subgroup, $Z(G) \subset G$ its center, $p^{\mathrm{ab}} : G \to G^{\mathrm{ab}} := G/G^{\mathrm{der}}$ its abelianization and $p^{\mathrm{ad}} : G \to G^{\mathrm{ad}} := G/Z(G)$ its adjoint quotient. If G is semisimple, we write $p^{\mathrm{sc}} : G^{\mathrm{sc}} \to G$ for its simply connected cover and set $\mu_G := \ker(p^{\mathrm{sc}})$. Let $\mathbb{A}_f$ denote the ring of finite adeles of $\mathbb{Q}$. Given a subgroup $\Gamma \subset G(\mathbb{A}_f) \subset \prod_\ell G(\mathbb{Q}_\ell)$, we write $\Gamma_\ell \subset G(\mathbb{Q}_\ell)$ for the projection of Γ into $G(\mathbb{Q}_\ell)$.

3. Galois-generic and strictly Galois-generic points

Let k be a field of characteristic 0 and let S be a smooth, separated and geometrically connected scheme over k with generic point η.

3.1. ***Galois-generic points.*** Let Γ be a topological group and $\rho : \pi_1(S) \to \Gamma$ a continuous group morphism. Write

$$\Pi_\rho := \mathrm{im}(\rho), \quad \overline{\Pi}_\rho := \rho(\pi_1(S_{\bar{k}})).$$

Every point $s \in S$ induces by functoriality of the étale fundamental group a morphism of profinite groups $\sigma_s : \pi_1(s) \to \pi_1(S)$ which is a section of the canonical projection $\pi_1(S_{k(s)}) \twoheadrightarrow \pi_1(k(s))$. Write

$$\Pi_{\rho,s} := \mathrm{im}(\rho \circ \sigma_s) \subset \Pi_\rho, \quad \overline{\Pi}_{\rho,s} := \overline{\Pi}_\rho \cap \Pi_{\rho,s}.$$

Definition 3.1.1. We say that $s \in S$ is *Galois-generic* with respect to ρ if $\Pi_{\rho,s}$ is open in Π_ρ, and *strictly Galois-generic* with respect to ρ if $\Pi_{\rho,s} = \Pi_\rho$.

We use this terminology when $\Gamma = G(\mathbb{A}_f)$ for some algebraic group G over $\mathbb{Q}$. For every prime ℓ, write

$$\rho_\ell : \pi_1(S) \xrightarrow{\rho} G(\mathbb{A}_f) \subset \prod_\ell G(\mathbb{Q}_\ell) \to G(\mathbb{Q}_\ell)$$

for the ℓ-adic component of $\rho : \pi_1(S) \to G(\mathbb{A}_f)$. If $\rho : \pi_1(S) \to G(\mathbb{A}_f)$ is clear from the context, we omit the subscript $(-)_\rho$ in the notation Π_ρ, $\overline{\Pi}_\rho$, $\Pi_{\rho,s}$, etc., and simply say that $s \in S$ is *Galois-generic* (resp. *ℓ-Galois-generic*) if $s \in S$ is Galois-generic with respect to ρ (resp. ρ_ℓ). Similarly, we say that $s \in S$ is *strictly Galois-generic* (resp. *strictly ℓ-Galois-generic*) if $s \in S$ is strictly Galois-generic with respect to ρ (resp. ρ_ℓ).

3.2. *Elementary properties of Galois-generic and strictly Galois-generic points.*

3.2.1. As S is normal, $\eta \in S$ is strictly Galois-generic.

3.2.2. Let $k \subset \tilde{k}$ be a finitely generated field extension. Then $s \in S$ is Galois-generic with respect to ρ if and only if $s_{\tilde{k}} \in S_{\tilde{k}}$ is Galois-generic with respect to $\rho|_{\pi_1(S_{\tilde{k}})}$. This follows from the fact that the images of the canonical morphisms $\pi_1(s_{\tilde{k}}) \to \pi_1(s)$ and $\pi_1(S_{\tilde{k}}) \to \pi_1(S)$ are open.

3.2.3. As S is geometrically connected over k, we have a short exact sequence

$$1 \to \pi_1(S_{\bar{k}}) \to \pi_1(S) \to \pi_1(k) \to 1$$

and as the image of $\pi_1(s) \xrightarrow{\sigma_s} \pi_1(S) \to \pi_1(k)$ is open in $\pi_1(k)$, we see that $s \in S$ is Galois-generic with respect to ρ if and only if $\overline{\Pi}_s$ is open in $\overline{\Pi}$. Another way to formulate this observation is the following. Let $k \subset \tilde{k} \subset \bar{k}$ denote the (in general infinite) Galois subextension corresponding to the image of $\rho^{-1}(\overline{\Pi}) \subset \pi_1(S)$ by $\pi_1(S) \twoheadrightarrow \pi_1(k)$. Then $s \in S$ is Galois-generic with respect to ρ if and only if $s_{\tilde{k}} \in S_{\tilde{k}}$ is Galois-generic with respect to $\rho|_{\pi_1(S_{\tilde{k}})}$. Under additional assumptions, one can enlarge $\tilde{k}$. For instance:

Lemma 3.2.4. *Assume that every open subgroup of $\overline{\Pi}$ has finite abelianization. With the above notation, let $\tilde{k} \subset \tilde{k}^{\mathrm{ab}} \subset \bar{k}$ denote the maximal abelian extension of $\tilde{k}$ in $\bar{k}$. Then $s \in S$ is Galois-generic with respect to ρ if and only if $s_{\tilde{k}^{\mathrm{ab}}} \in S_{\tilde{k}^{\mathrm{ab}}}$ is Galois-generic with respect to $\rho|_{\pi_1(S_{\tilde{k}^{\mathrm{ab}}})}$.*

Proof. The nontrivial implication is the "only if" one. So, assume that $s \in S$ is Galois-generic with respect to ρ. Then $\Pi_{s_{\tilde{k}}} \subset \overline{\Pi}$ is open. In particular, it has finite abelianization. Thus its quotient $\Pi_{s_{\tilde{k}}} \twoheadrightarrow \Pi_{s_{\tilde{k}}}/\Pi_{s_{\tilde{k}^{\mathrm{ab}}}}$ (which, being a quotient of $\mathrm{Gal}(\tilde{k}^{\mathrm{ab}}|\tilde{k})$, is abelian) is finite. □

3.2.5. If $S' \to S$ is a dominant morphism of finite type with S' connected (for instance, a connected étale cover) and $s' \in S'$ a point above $s \in S$, then $s \in S$ is Galois-generic with respect to ρ if and only if $s' \in S'$ is Galois-generic with respect to $\rho|_{\pi_1(S')}$.

3.2.6. Given a Galois-generic point $s \in S$, one can always find a connected étale cover $S' \to S$ and $s' \in S'$ above s such that $s' \in S'$ is strictly Galois-generic with respect to $\rho|_{\pi_1(S')}$. Indeed, let $s \in S$ be a Galois-generic point and let $U \subset \Pi$ be any open subgroup contained in Π_s. Let $S_U \to S$ denote the connected étale cover corresponding to the open subgroup $\rho^{-1}(U) \subset \pi_1(S)$ and let $k(s) \hookrightarrow k(s)_U$ denote the finite field extension corresponding to the open subgroup $(\rho \circ \sigma_s)^{-1}(U) \subset \pi_1(s)$. Then $s \in S$ lifts to a $k(s)_U$-rational point $s_U \in S_U$ which is strictly Galois-generic with respect to $\rho|_{\pi_1(S_U)}$.

By definition, if $s \in S$ is strictly Galois-generic, then for every open subgroup $U \subset \Pi$ and corresponding connected étale cover $S_U \to S$, $\pi_1(s)$ acts transitively on the geometric fiber of $S_U \to S$ over s.

3.3. ***The ℓ-GG ⇔ GG problem.*** We now assume that $\Gamma = G(\mathbb{A}_f)$. Let $S^{\mathrm{gg}} \subset S$ and $S^{\mathrm{gg}}_\ell \subset S$ denote the sets of Galois- and ℓ-Galois-generic points, respectively. Write

$$S^{\mathrm{gg}}_\infty := \bigcap_\ell S^{\mathrm{gg}}_\ell, \qquad S^{\mathrm{gg}\,\infty} := \bigcup_\ell S^{\mathrm{gg}}_\ell.$$

Clearly,

$$S^{\mathrm{gg}} \subset S^{\mathrm{gg}}_\infty \subset S^{\mathrm{gg}\,\infty}.$$

3.3.1. *ℓ-Galois-generic points.* One can show by ℓ-adic specialization techniques that S^{gg}_ℓ is nonempty and even "large in an arithmetical sense" provided k satisfies some reasonable assumptions. More precisely, we have the following facts.

Fact 3.3.1.1 [Serre 1989, §10.6]. *Assume k is Hilbertian. Then there exist an integer $d \geq 1$ and infinitely many closed strictly ℓ-Galois-generic points $s \in S$ with $[k(s) : k] \leq d$.*

Fact 3.3.1.2 [Cadoret and Tamagawa 2013, Theorem 1.1]. *Assume k is finitely generated, S is a curve and every open subgroup of $\overline{\Pi}_\ell$ has finite abelianization. Then for every integer $d \geq 1$, all but finitely many $s \in S$ with $[k(s) : k] \leq d$ are ℓ-Galois-generic.*

Note that ℓ-adic motivic representations (see Section 3.3.2) satisfy the assumption of Fact 3.3.1.2 [Cadoret and Tamagawa 2012, §5.2]. The ℓ-adic components of adelic representations attached to Shimura varieties also satisfy this assumption (see Section 4.2 and Theorem 5.4).

These results rely heavily on the fact that $\overline{\Pi}_\ell$ and Π_ℓ are compact ℓ-adic Lie groups: the key point in the proof of Fact 3.3.1.1 is that the Frattini subgroup $\Phi(\Pi)$ of a compact ℓ-adic Lie group Π is open in Π. This property is also crucial in the proof of Fact 3.3.1.2, which also involves finer structural results about compact ℓ-adic Lie groups.

3.3.2. *Motivic representations.* The above results are ℓ-adic in nature and fail to ensure that S_∞^{gg} contains points other than the generic point.

Definition 3.3.2.1. We say that an adelic representation $\rho:\pi_1(S)\to G(\mathbb{A}_f)$ satisfies the (ℓ-GG $\Leftrightarrow$ GG)-property if $S^{\mathrm{gg}} = S_\infty^{\mathrm{gg}} = S^{\mathrm{gg}\,\infty}$.

When an adelic representation satisfies the (ℓ-GG $\Leftrightarrow$ GG)-property, the abundance results of Section 3.3.1 automatically hold for Galois-generic points.

Conjecturally, motivic representations — those of the form

$$\rho_X^w : \pi_1(S) \to \prod_\ell \mathrm{GL}(\mathrm{H}^w(X_{\bar\eta}, \mathbb{Q}_\ell))$$

for some smooth projective scheme $f : X \to S$ — should satisfy the (ℓ-GG $\Leftrightarrow$ GG)-property. More precisely, the equality $S_\infty^{\mathrm{gg}} = S^{\mathrm{gg}\,\infty}$ follows from the Tate conjectures [Serre 1994, §9; André 2004, §7.3] while the equality $S^{\mathrm{gg}} = S_\infty^{\mathrm{gg}}$ follows from the modulo-ℓ variant of the Tate conjectures proposed by Serre [1994, §10].

For abelian schemes, partial forms of the modulo-ℓ variants of the Tate conjectures were proved by Faltings; see, e.g., [Faltings and Wüstholz 1984]. These are enough to show that adelic motivic representations attached to abelian schemes satisfy the (ℓ-GG $\Leftrightarrow$ GG)-property. More precisely, given an abelian scheme $f : X \to S$, recall that $R^w f_*\mathbb{Z}_\ell = \Lambda^w R^1 f_*\mathbb{Z}_\ell \simeq \Lambda^w T_\ell(X)^\vee$, where

$$T_\ell(X) := \varprojlim_n X[\ell^n]$$

denotes the ℓ-adic Tate module (here $X[N]$ denotes the kernel of multiplication by N on X; as k has characteristic 0, this is an étale cover of S). Thus, the (ℓ-GG $\Leftrightarrow$ GG)-property for $\rho_X^w : \pi_1(S) \to \prod_\ell \mathrm{GL}(\mathrm{H}^w(X_{\bar\eta}, \mathbb{Q}_\ell))$ boils down to the following statement.

Theorem 3.3.2.2 [Cadoret 2015, Theorem 1.2]. *The representation*

$$(\rho_X^1)^\vee : \pi_1(S) \to \mathrm{GL}(T(X)_{\bar\eta})$$

satisfies the (ℓ-GG $\Leftrightarrow$ GG)-property.

4. Adelic representations attached to Shimura varieties

Let (G, X) be a Shimura datum. We assume throughout that G is the generic Mumford–Tate group on X. In this case, conditions 2.1.1.1–2.1.1.5 of [Deligne 1979] are satisfied and $Z(\mathbb{Q})$ is discrete in $Z(\mathbb{A}_f)$ [Deligne 1979, 2.1.11]; for details see also [Ullmo and Yafaev 2013, Lemma 5.13].

Let $K_0 \subset G(\mathbb{A}_f)$ be a neat compact open subgroup. If $K \subset K_0$ is an open subgroup, the induced morphism $p_{K,K_0} : \mathrm{Sh}_K(G, X) \to \mathrm{Sh}_{K_0}(G, X)$ on Shimura varieties is finite étale. If, moreover, K is normal in K_0, this morphism is Galois with group K_0/K. Fix a point $s \in \mathrm{Sh}(G, X)$ and for every open subgroup $K \subset K_0$, let $s[K]$ denote the image of s in $\mathrm{Sh}_K(G, X)$, and let $S[K, s] \subset \mathrm{Sh}_K(G, X)$ denote the geometrically connected component of $s[K]$ and $E[K] := E[K](G, X)$ its field of definition (a finite abelian extension of the reflex field $E := E(G, X)$). Let $\tilde{S}[K, s] \subset \mathrm{Sh}_K(G, X)_{E[K_0]}$ be the connected component of $s[K]$ in $\mathrm{Sh}_K(G, X)_{E[K_0]}$ (explicitly, $\tilde{S}[K, s]$ is the union of the $\mathrm{Gal}(E[K]|E[K_0])$-conjugate of $S[K, s]$). The tower of (connected) pointed Galois covers

$$p_{K,K_0,s} : (\tilde{S}[K, s], s[K]) \to (S[K_0, s], s[K_0])$$

corresponds to a projective system of continuous group morphisms

$$\pi_1(S[K_0, s], \bar{s}[K_0]) \twoheadrightarrow \mathrm{Aut}(p_{K,K_0,s}) \subset K_0/K.$$

As the intersection of all open normal subgroups $K \subset K_0$ is trivial, passing to the limit we obtain a continuous group morphism

$$\rho[K_0, s] : \pi_1(S[K_0, s], \bar{s}[K_0]) \twoheadrightarrow \varprojlim_K K_0/K = K_0.$$

By construction, this morphism satisfies the following properties.

4.1. ***Functoriality.*** Let $f : (G_2, X_2) \to (G_1, X_1)$ be a morphism of Shimura data and $K_i \subset G_i(\mathbb{A}_f)$, $i = 1, 2$ neat compact open subgroups such that $f(K_2) \subset K_1$; these induces morphisms of Shimura varieties $\mathrm{Sh}(G_2, X_2) \to \mathrm{Sh}(G_1, X_1)_{E_2}$ and $\mathrm{Sh}_{K_2}(G_2, X_2) \to \mathrm{Sh}_{K_1}(G_1, X_1)_{E_2}$ over the reflex field $E_2 := E(G_2, X_2)$. Fix $s_2 \in \mathrm{Sh}(G_2, X_2)$ and set $s_1 := f(s_2) \in \mathrm{Sh}(G_1, X_1)$. Then the following diagram commutes:

$$\begin{array}{ccc} K_2 & \xrightarrow{\quad f \quad} & K_1 \\ {\scriptstyle \rho[K_2,s_2]}\uparrow & & \uparrow{\scriptstyle \rho[K_1,s_1]} \\ \pi_1(S[K_2, s_2], \bar{s}_2[K_2]) & \longrightarrow & \pi_1(S[K_1, s_1]_{E[K_2]}, \bar{s}_1[K_1]) \end{array}$$

where, as above, $E[K_2] := E[K_2](G_2, X_2)$ denotes the field of definition of the geometrically connected component $S[K_2, s_2]$ of $s_2[K_2]$ in $\mathrm{Sh}_{K_2}(G_2, X_2)$.

4.2. ***Change of connected component.*** Assume the $\mathbb{C}$-valued point

$$s_{\mathbb{C}} \in \mathrm{Sh}(G, X)(\mathbb{C}) = G(\mathbb{Q}) \setminus X \times G(\mathbb{A}_f)$$

corresponding to s is of the form $s_{\mathbb{C}} = G(\mathbb{Q})(x, 1)$, let $a \in G(\mathbb{A}_f)$ and write $sa \in \mathrm{Sh}(G, X)$ for the point corresponding to $s_{\mathbb{C}}a = G(\mathbb{Q})(x, a)$. For a neat compact open subgroup $K \subset G(\mathbb{A}_f)$, write $K^a := K \cap aKa^{-1}$. As the Hecke-operator $-a : \mathrm{Sh}(G, X) \xrightarrow{\sim} \mathrm{Sh}(G, X)$ is defined over the reflex field, the following diagram commutes:

$$\begin{array}{ccc}
\pi_1(S[K_0,s],\bar{s}[K_0]) & \xrightarrow{\rho[K_0,s]} & K_0 \\
\uparrow & & \uparrow \\
\pi_1(S[K_0^a,s],\bar{s}[K_0^a]) & \xrightarrow{\rho[K_0^a,s]} & K_0^a \\
\downarrow \simeq & & a^{-1}-a \downarrow \simeq \\
\pi_1(S[K_0^{a^{-1}},sa],\overline{sa}[K_0^{a^{-1}}]) & \xrightarrow{\rho[K_0^{a^{-1}},sa]} & K_0^{a^{-1}} \\
\downarrow & & \downarrow \\
\pi_1(S[K_0,s],\overline{sa}[K_0]) & \xrightarrow{\rho[K_0,sa]} & K_0
\end{array}$$

where the downwards vertical arrow on the left is induced by the isomorphism $-a : S[K_0^a, s] \tilde{\to} S[K_0^{a^{-1}}, sa]$ (mapping $s[K_0^a]$ to $sa[K_0^{a^{-1}}]$), and the upper and lower left arrows are open embeddings.

Let $X^+ \subset X$ denote the connected component of $x \in X$ and $G(\mathbb{Q})_+ \subset G(\mathbb{Q})$ the stabilizer of X^+ in $G(\mathbb{Q})$. Given an open subgroup $K \subset K_0$, write $\Gamma := G(\mathbb{Q})_+ \cap K$ and $\Gamma_0 := G(\mathbb{Q})_+ \cap K_0$. Then $\tilde{S}[K, s]_{E_K}$ splits into a disjoint union of geometrically connected components isomorphic to the connected component $\mathrm{Sh}_\Gamma(G, X^+)$ of $\mathrm{Sh}_K(G, X)_{E_K}$ containing the image of $X^+ \times \{1\}$. As $\mathrm{Sh}_\Gamma(G, X^+)_{\mathbb{C}} \simeq \Gamma \setminus X^+$, it follows that $\rho_{K_0,s} : \pi_1(S_{K_0,s}, \bar{s}_{K_0}) \to K_0$ restricts to a surjective continuous group morphism

$$\rho[K_0, s] : \pi_1(S[K_0, s]_{\overline{E}}, \bar{s}[K_0]) \twoheadrightarrow \Gamma_0^- .$$

Here we implicitly identify the closure Γ_0^- of Γ_0 in $G(\mathbb{A}_f)$ with $\varprojlim \Gamma_0/\Gamma$, where the projective limit is taken over all normal congruence subgroups of Γ_0.

Actually, as $E[K]$ is contained in the maximal abelian extension E^{ab} of the reflex field E, the above shows that $\rho[K_0, s] : \pi_1(S[K_0, s], \bar{s}[K_0]) \to K_0$ already restricts to a surjective continuous group morphism

$$\rho[K_0, s] : \pi_1(S[K_0, s]_{E^{\mathrm{ab}}}, \bar{s}[K_0]) \twoheadrightarrow \Gamma_0^- ,$$

which is completely determined by the tower of connected étale covers over E^{ab}:

$$p_{\Gamma,\Gamma_0} : \mathrm{Sh}_\Gamma(G, X^+)_{E^{\mathrm{ab}}} \to \mathrm{Sh}_{\Gamma_0}(G, X^+)_{E^{\mathrm{ab}}},$$

where Γ describes all normal congruence subgroups of Γ_0.

4.3. ***Change of base point.*** Let $s, s' \in \mathrm{Sh}(G, X)$ be two points lying in the same geometrically connected component; write $S[K_0, s] = S[K_0, s'] =: S[K_0]$. Then every étale path $\alpha : \bar{s}[K_0] \to \bar{s}'[K_0]$ mapping $\bar{s}$ to $\bar{s}'$ induces an isomorphism of profinite groups $\alpha : \pi_1(S[K_0], \bar{s}[K_0]) \tilde{\to} \pi_1(S[K_0], \bar{s}'[K_0])$, which makes the following diagram commute:

$$\begin{array}{ccc} \pi_1(S[K_0], \bar{s}[K_0]) & \xrightarrow{\rho[K_0,s]} & K_0 \\ \simeq \Big\downarrow \alpha & \nearrow_{\rho[K_0,s']} & \\ \pi_1(S[K_0], \bar{s}'[K_0]) & & \end{array}$$

4.4. ***Galois-generic points.*** For $s \in \mathrm{Sh}(G, X)$ the following assertions are equivalent (see 3.2.2 and 3.2.5):

(1) There exists a neat compact open subgroup $K \subset G(\mathbb{A}_f)$ such that $s[K] \in S[K, s]$ is (ℓ-)Galois-generic with respect to $\rho[K, s] : \pi_1(S[K, s], \bar{s}[K]) \to K \subset G(\mathbb{A}_f)$.

(2) For every neat compact open subgroup $K \subset G(\mathbb{A}_f)$, $s[K] \in S[K, s]$ is (ℓ-)Galois-generic with respect to $\rho[K, s] : \pi_1(S[K, s], \bar{s}[K]) \to K \subset G(\mathbb{A}_f)$.

In this case we say that $s \in \mathrm{Sh}(G, X)$ is (ℓ-)*Galois-generic*. In view of Section 4.2, we also see that for every $a \in G(\mathbb{A}_f)$, $s \in \mathrm{Sh}(G, X)$ is Galois-generic if and only if $sa \in \mathrm{Sh}(G, X)$ is Galois-generic. So, in the following, we always assume that $s_{\mathbb{C}} = G(\mathbb{Q})(x, 1)$ and write $X^+ \subset X$ for the connected component of x. In particular, with the notation of Section 4.2, $S[K_0, s] = \mathrm{Sh}_{\Gamma_0}(G, X^+)$ and the restriction $\rho[K_0, s] : \pi_1(S[K_0, s]_{E^{\mathrm{ab}}}, \bar{s}[K_0]) \twoheadrightarrow \Gamma_0^-$ is completely determined by the tower of connected étale covers over E^{ab},

$$p_{\Gamma,\Gamma_0} : \mathrm{Sh}_{\Gamma}(G, X^+)_{E^{\mathrm{ab}}} \to \mathrm{Sh}_{\Gamma_0}(G, X^+)_{E^{\mathrm{ab}}}.$$

Also, in view of Section 4.3, we shall omit the reference to s in the notation (e.g., write $S[K_0]$, $\rho[K_0]$, $\pi_1(S[K_0])$ instead of $S[K_0, s]$, $\rho[K_0, s]$, $\pi_1(S[K_0, s], \bar{s}[K_0])$, etc.) unless it plays a part in the discussion.

4.5. ***Adelic representation attached to Siegel Shimura varieties.*** Let (GSp_{2g}, X) denote the Siegel Shimura datum [Deligne 1971, Exemple 1.6]. Using the moduli description of the attached Shimura variety [Deligne 1971, Exemple 4.16], one easily shows that for a neat compact open subgroup $K_0 \subset \mathrm{GSp}_{2g}(\mathbb{A}_f)$ and geometrically connected component $S[K_0] \subset \mathrm{Sh}_{K_0}(\mathrm{GSp}_{2g}, X)$, the corresponding adelic representation $\rho[K_0] : \pi_1(S[K_0]) \to K \subset G(\mathbb{A}_f)$ identifies with the representation $\rho : \pi_1(S[K_0]) \to \mathrm{GL}(T(A)_{\bar{\eta}})$ on the adelic Tate module of the universal abelian scheme $A \to S[K_0]$. (See also [Ullmo and Yafaev 2013].)

5. Group-theoretical preliminaries

In this section, we gather technical group-theoretical results about the adelic closure of arithmetic subgroups of semisimple algebraic groups. These will be used in the proof of Theorem A to deduce the case of Shimura data of abelian type from the case of Shimura data of Hodge type.

Let G be a group. Recall that two subgroups $K, K' \subset G$ are said to be commensurable if $K \cap K'$ is of finite index in both K and K'. Commensurability is an equivalence relation, which we denote by $\equiv$, on the set of subgroups of G.

For an algebraic group G over $\mathbb{Q}$, a faithful $\mathbb{Q}$-linear representation $G \hookrightarrow \mathrm{GL}(V)$ and a $\mathbb{Z}$-lattice $L \subset V$, write G_L for the subgroups of elements $g \in G(\mathbb{Q})$ stabilizing L. If $G \hookrightarrow \mathrm{GL}(V)$, $G \hookrightarrow \mathrm{GL}(V')$ are two faithful $\mathbb{Q}$-linear representations and $L \subset V$, $L' \subset V'$ are $\mathbb{Z}$-lattices then $G_L \equiv G_{L'} \subset G$. Thus the class of commensurability of G_L does not depend on the choices of $G \hookrightarrow \mathrm{GL}(V)$ and $L \subset V$; the groups in this class are the arithmetic subgroups of G. Arithmetic subgroups have the following properties.

Fact 5.1. (1) *For an algebraic group G over $\mathbb{Q}$, a faithful $\mathbb{Q}$-linear representation $G \hookrightarrow \mathrm{GL}(V)$ and an arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, there exists a Γ-invariant $\mathbb{Z}$-lattice $L \subset V$.*

(2) *For a surjective morphism $f : G_2 \to G_1$ of algebraic groups over $\mathbb{Q}$ and an arithmetic subgroup $\Gamma \subset G_2(\mathbb{Q})$, the subgroup $f(\Gamma) \subset G_1(\mathbb{Q})$ is again arithmetic.*

Let G be a semisimple algebraic group over $\mathbb{Q}$ and $\Gamma \subset G(\mathbb{Q})$ an arithmetic subgroup. Then:

(3) *Γ is finitely presented as an abstract group.*

(4) *If, furthermore, G is of noncompact type then Γ is Zariski-dense in G.*

An algebraic group G over $\mathbb{Q}$ is said to be of compact type if $G(\mathbb{R})$ is compact [Platonov and Rapinchuk 1994, Definition p. 205]. A semisimple algebraic group over $\mathbb{Q}$ is said to be of noncompact type if none of its simple factors is of compact type.

Proof. For assertions (1), (2) and (3), see [Platonov and Rapinchuk 1994, Proposition 4.2, Theorems 4.1 and 4.2], respectively. Assertion (4) is the Borel density theorem [Borel 1966]; see also [Platonov and Rapinchuk 1994, Theorem 4.10]. □

For an algebraic group G over $\mathbb{Q}$ and a subgroup $\Gamma \subset G(\mathbb{Q})$, let $\Gamma^- \subset G(\mathbb{A}_f)$ denote the adelic closure of Γ in $G(\mathbb{A}_f)$.

Lemma 5.2. *If $\Gamma \subset G(\mathbb{Q})$ is an arithmetic subgroup, then Γ^- is profinite and the collection of subgroups Γ'^-, for $\Gamma' \subset \Gamma$ a normal subgroup of finite index, is a fundamental system of open neighborhoods of 1 in Γ^-.*

For an algebraic subgroup $G \subset \mathrm{GL}_{n,\mathbb{Q}}$ and a ring A of characteristic 0, write $G(A) := G(\mathbb{Q}) \cap \mathrm{GL}_n(A)$.

Proof. Let $G \hookrightarrow \mathrm{GL}(V)$ be a faithful $\mathbb{Q}$-rational representation of G and $L \subset V$ a Γ-invariant $\mathbb{Z}$-lattice (see Fact 5.1(1)). Fixing a $\mathbb{Z}$-basis of L we get an isomorphism $\mathrm{GL}(V) \simeq \mathrm{GL}_{n,\mathbb{Q}}$ such that $\Gamma \subset G(\mathbb{Z})$. Then Γ^- is a closed subgroup of the profinite group $\prod_\ell G(\mathbb{Z}_\ell)$, hence is profinite. In particular, a subgroup of Γ^- is open if and only if it is closed of finite index in Γ^-. As for any subgroup $\Gamma' \subset \Gamma$ of finite index, $\Gamma'^- \subset \Gamma^-$ is a closed subgroup of finite index $\leq [\Gamma : \Gamma']$, one already sees that the Γ'^-, with $\Gamma' \subset \Gamma$ of finite index, are open neighborhoods of 1 in Γ^-. Moreover, for every open subgroup $U \subset \Gamma^-$, the intersection $\Gamma \cap U \subset \Gamma$ has finite index $\leq [\Gamma^- : U]$ and, by construction, $(\Gamma \cap U)^- \subset U^- = U$. Eventually, if $\Gamma' \subset \Gamma$ is a subgroup of finite index, then

$$\bigcap_{\gamma \in \Gamma} \gamma \Gamma' \gamma^{-1} \subset \Gamma'$$

is normal and again of finite index $\leq [\Gamma : \Gamma']!$ in Γ. □

For a closed subgroup $U \subset \mathrm{GL}_n(\mathbb{Z}_\ell)$, let $U^+ \subset U$ denote the (normal) subgroup generated by the ℓ-Sylow subgroups of U.

Lemma 5.3. *Let $G \subset \mathrm{GL}_{n,\mathbb{Q}}$ be an algebraic subgroup and $U \subset G(\mathbb{A}_f)$ a closed subgroup such that $U_\ell \subset G(\mathbb{Z}_\ell)$ for $\ell \gg 0$. Assume that $U_\ell = U_\ell^+$ for $\ell \gg 0$. Then there exists an open subgroup $U' \subset U$ such that $U' = \prod_\ell U'_\ell$.*

Proof. This follows from a combination of results about almost-ℓ independency in the sense of Serre [2013]. More precisely, given an infinite set of primes L, a family G_ℓ, $\ell \in L$, of ℓ-adic Lie groups and a profinite group $\Delta \subset \prod_\ell G_\ell$, one says that Δ is ℓ-independent (as a subgroup of $\prod_\ell G_\ell$) if $\Delta = \prod_\ell \Delta_\ell$, and that Δ is almost ℓ-independent if there exists an open subgroup $\Delta' \subset \Delta$ which is ℓ-independent as a subgroup of $\prod_\ell G_\ell$. With these definitions, the following hold:

(1) [Serre 2013, Lemma 1] If for $\ell \neq \ell'$ no simple quotient of Δ_ℓ is isomorphic to a simple quotient of $\Delta_{\ell'}$, then $\Delta \subset \prod_\ell G_\ell$ is ℓ-independent.

(2) [Serre 2013, Lemma 3] If there exists a finite subset $F \subset L$ such that the image of Δ in $\prod_{L \setminus F} G_\ell$ is almost ℓ-independent, then $\Delta \subset \prod_\ell G_\ell$ is almost ℓ-independent.

For every prime ℓ, let Σ_ℓ denote the set of all (isomorphism classes of) finite groups which are either a simple group of Lie type in characteristic ℓ (see [Serre 2013, §6.1]) or $\mathbb{Z}/\ell$.

(3) [Serre 2013, Theorem 4] Every finite simple subquotient of $\mathrm{GL}_n(\mathbb{Z}_\ell)$ of order divisible by ℓ is in Σ_ℓ for $\ell \gg 0$ (depending on n).

(4) [Serre 2013, Theorem 5] For $\ell, \ell' \geq 5$ with $\ell \neq \ell'$, one has $\Sigma_\ell \cap \Sigma_{\ell'} = \varnothing$.

From (2), it is enough to show that the image U_L of U in $\prod_{\ell \in L} G(\mathbb{Q}_\ell)$ is almost ℓ-independent for a set L containing all but finitely many primes. In particular, one may assume that $U_\ell \subset G(\mathbb{Z}_\ell)$, that $U_\ell = U_\ell^+$ for every $\ell \in L$ and that the conclusions of (3) and (4) hold for $\ell \in L$ and $\ell \neq \ell' \in L$, respectively. As every simple quotient of $U_\ell^\#(=U_\ell)$ is in Σ_ℓ for $\ell \in L$, (4) and (1) show that U_L is ℓ-independent, as requested. □

Theorem 5.4. *Let G be a semisimple algebraic group over $\mathbb{Q}$. Let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic subgroup. Then every open subgroup U of $\Gamma^- \subset G(\mathbb{A}_f)$ has finite abelianization.*

Proof. Fix an embedding $G \hookrightarrow \mathrm{GL}_{n,\mathbb{Q}}$. We let again G denote the Zariski closure of G in $\mathrm{GL}_{n,\mathbb{Z}}$; this is a semisimple group over some nonempty open subscheme of $\mathrm{spec}(\mathbb{Z})$.

• *Reduction to the case where U is of the form Γ^- for some normal arithmetic subgroup* $\Gamma \subset G(\mathbb{Z})$. From Lemma 5.2, there exists a subgroup $\Gamma' \subset \Gamma$ normal, of finite index in Γ and such that $\Gamma'^- \subset U$. From the finiteness of U/Γ'^- and the exact sequence

$$(\Gamma'^-)^{\mathrm{ab}} \to U^{\mathrm{ab}} \to (U/\Gamma'^-)^{\mathrm{ab}} \to 0,$$

it is enough to perform the proof for Γ'^-; that is, we may assume U is of the form Γ^- for some arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$. Next, as Γ is finitely generated as an abstract group (Fact 5.1(3)), it has only finitely many subgroups of bounded index $\leq [\Gamma : G(\mathbb{Z}) \cap \Gamma]$. In particular, the group

$$\Delta := \bigcap_{g \in G(\mathbb{Z})\Gamma} g G(\mathbb{Z}) \cap \Gamma g^{-1} \subset G(\mathbb{Q})$$

is again an arithmetic subgroup, contained and normal in both Γ and $G(\mathbb{Z})$. So the conclusion follows from the finiteness of Γ/Δ and the exact sequence

$$\Delta^{-\,\mathrm{ab}} \to \Gamma^{-\,\mathrm{ab}} \to (\Gamma^-/\Delta^-)^{\mathrm{ab}} \to 0.$$

• *Reduction to the case where G is of noncompact type.* Let $G^{\mathrm{nc}} \subset G$ denote the largest (normal) algebraic subgroup of G of noncompact type, and $p : G \twoheadrightarrow G/G^{\mathrm{nc}}$ the canonical projection. Write

$$\Gamma^{\mathrm{nc}} := \Gamma \cap G^{\mathrm{nc}}(\mathbb{Q}) \subset G^{\mathrm{nc}}(\mathbb{Q}).$$

As $p(\Gamma) \subset G/G^{\mathrm{nc}}(\mathbb{Q})$ is again an arithmetic subgroup (Fact 5.1(3)) and G/G^{nc} is of compact type, $p(\Gamma)$ is finite; in particular, $\Gamma^-/\Gamma^{\mathrm{nc}\,-}$ is finite. Also, by construction, Γ^{nc} is contained and normal in $G^{\mathrm{nc}}(\mathbb{Z})$ and $[G^{\mathrm{nc}}(\mathbb{Z}) : \Gamma^{\mathrm{nc}}] \leq [G(\mathbb{Z}) : \Gamma]$, which shows that $\Gamma^{\mathrm{nc}} \subset G^{\mathrm{nc}}(\mathbb{Q})$ is again an arithmetic subgroup. Thus the conclusion follows from the finiteness of $\Gamma^-/\Gamma^{\mathrm{nc}\,-}$ and the exact sequence

$$\Gamma^{\mathrm{nc}\,-\,\mathrm{ab}} \to \Gamma^{-\,\mathrm{ab}} \to (\Gamma^-/\Gamma^{\mathrm{nc}\,-})^{\mathrm{ab}} \to 0.$$

• *Reduction to the case where $\Gamma_\ell^- = G(\mathbb{Z}_\ell)^+$ for $\ell \gg 0$.* For a prime ℓ, let $p_\ell : \mathrm{GL}_n(\mathbb{Z}) \twoheadrightarrow \mathrm{GL}_n(\mathbb{F}_\ell)$ denote the reduction modulo-ℓ morphism. Then, as

Γ is finitely generated as an abstract group (Fact 5.1(3)) and Zariski-dense in G (Fact 5.1(4)), we have

$$G(\mathbb{F}_\ell)^+ \subset p_\ell(\Gamma) \subset G(\mathbb{F}_\ell)$$

for $\ell \gg 0$ depending only on n [Nori 1987, Theorem 5.1]. Here $G(\mathbb{F}_\ell)^+ \subset G(\mathbb{F}_\ell)$ denotes the (normal) subgroup generated by the order-ℓ elements in $G(\mathbb{F}_\ell)$ (or, equivalently, the ℓ-Sylow subgroups as soon as $\ell > n$). As G is semisimple, $G(\mathbb{F}_\ell)/G(\mathbb{F}_\ell)^+$ is abelian of order $\leq 2^{n-1}$. In particular, there exists an integer $N \geq 1$ such that for every prime ℓ the subgroup

$$\Delta[\ell] := p_\ell^{-1}(G(\mathbb{F}_\ell)^+) \cap \Gamma \subset \Gamma$$

is normal and of index $\leq N$ in Γ. As Γ is finitely generated, it has only finitely many subgroups of index $\leq N$. So

$$\Delta := \bigcap_\ell \Delta[\ell] \subset \Gamma$$

is again a subgroup normal and of finite index in Γ. For $\ell > [\Gamma : \Delta]$ and $\ell \gg 0$ such that $G(\mathbb{F}_\ell)^+ \subset \overline{\Gamma}_\ell$, we have

$$[G(\mathbb{F}_\ell)^+ : p_\ell(\Delta)] = [p_\ell(\Delta[\ell]) : p_\ell(\Delta)] \leq [\Delta[\ell] : \Delta] \leq [\Gamma : \Delta] < \ell.$$

As $G(\mathbb{F}_\ell)^+$ is generated by its order-ℓ elements, this forces $p_\ell(\Delta) = G(\mathbb{F}_\ell)^+$. Then the finiteness of Γ^-/Δ^- and the exact sequence

$$\Delta^{-\,\mathrm{ab}} \to \Gamma^{-\,\mathrm{ab}} \to (\Gamma^-/\Delta^-)^{\mathrm{ab}} \to 0$$

show that it is enough to prove that $\Delta^{-\,\mathrm{ab}}$ is finite. That is, without loss of generality, we may replace Γ with Δ, and hence assume that $p_\ell(\Gamma) = G(\mathbb{F}_\ell)^+$ for $\ell \gg 0$. As $p_\ell^-(\Gamma_\ell^-) = p_\ell(\Gamma_\ell) = G(\mathbb{F}_\ell)^+$ and $\ker(p_\ell^-) \subset G(\mathbb{Z}_\ell)$ is a pro-ℓ group, this implies $\Gamma_\ell^- = (\Gamma_\ell^-)^+$, and hence $\Gamma_\ell^- \subset G(\mathbb{Z}_\ell)^+$. Assume also that $\ell \gg 0$, so that $G_{\mathbb{Z}_\ell}$ is semisimple over $\mathbb{Z}_\ell$. Then the reduction-modulo-ℓ morphism $p_\ell^- : G(\mathbb{Z}_\ell) \twoheadrightarrow G(\mathbb{F}_\ell)$ is surjective and, up to increasing ℓ, we may assume that the induced surjective morphism

$$p_\ell^-|_{G(\mathbb{Z}_\ell)^+} : G(\mathbb{Z}_\ell)^+ \twoheadrightarrow G(\mathbb{F}_\ell)^+$$

is Frattini [Cadoret 2015, Fact 2.4, Lemma 2.5]; that is, $G(\mathbb{Z}_\ell)^+$ is the unique closed subgroup $X \subset G(\mathbb{Z}_\ell)^+$ mapping subjectively onto $G(\mathbb{F}_\ell)^+$. This shows that $\Gamma_\ell^- = G(\mathbb{Z}_\ell)^+$.

• *End of the proof.* As $\Gamma_\ell^- = (\Gamma_\ell^-)^+$ for $\ell \gg 0$, there exists (Lemma 5.3) an open subgroup $U \subset \Gamma^-$ such that $U = \prod_\ell U_\ell$. As Γ^- is profinite, U is of finite index in Γ^-, and thus $[\Gamma_\ell^- : U_\ell] \leq [\Gamma^- : U]$ for every ℓ. On the other hand, as $\Gamma_\ell^- = (\Gamma_\ell^-)^+$, all subgroups of Γ_ℓ^- have index $\geq \ell$ in Γ_ℓ^-. This forces $U_\ell = \Gamma_\ell^-$ for $\ell \gg 0$. Also, up to replacing U_ℓ by

$$\bigcap_{\gamma_\ell \in \Gamma_\ell^-} \gamma_\ell U_\ell \gamma_\ell^{-1}$$

for small ℓ, we may assume that U is normal in $\prod_\ell \Gamma_\ell^-$ (hence a fortiori in Γ^-). Then the exact sequence

$$U^{\mathrm{ab}} \to \Gamma^{-\,\mathrm{ab}} \to (\Gamma^-/U)^{\mathrm{ab}} \to 0$$

shows that it is enough to prove that $U^{\mathrm{ab}} = \prod_\ell U_\ell^{\mathrm{ab}}$ is finite; equivalently,

(1) U_ℓ^{ab} is finite for every ℓ;

(2) $U_\ell^{\mathrm{ab}} = 0$ for $\ell \gg 0$ (or, equivalently, $(G(Z_\ell)^+)^{\mathrm{ab}} = 0$ for $\ell \gg 0$).

Proof of (1): If (1) were false, U_ℓ would have an infinite abelian quotient $U_\ell \twoheadrightarrow Z_\ell$. As U_ℓ is an ℓ-adic Lie group, so is Z_ℓ and Z_ℓ has dimension ≥ 1 as an ℓ-adic Lie group. From exactness of the Lie algebra functor on the category of ℓ-adic Lie groups, we obtain a surjective morphism of Lie algebras

$$\mathrm{Lie}(U_\ell) \twoheadrightarrow \mathrm{Lie}(Z_\ell).$$

On the other hand, since G is semisimple, we also have $\mathrm{Lie}(U_\ell) = \mathrm{Lie}(\Gamma_\ell^-) = \mathrm{Lie}(G) \otimes \mathbb{Q}_\ell$, which has no abelian quotient as a Lie algebra.

Proof of (2): Since, as noted above,

$$p_\ell^-|_{G(\mathbb{Z}_\ell)^+} : G(\mathbb{Z}_\ell)^+ \twoheadrightarrow G(\mathbb{F}_\ell)^+$$

is Frattini and $[G(\mathbb{Z}_\ell)^+, G(\mathbb{Z}_\ell)^+]^-$ maps surjectively onto $[G(\mathbb{F}_\ell)^+, G(\mathbb{F}_\ell)^+]$, it is enough to show that $[G(\mathbb{F}_\ell)^+, G(\mathbb{F}_\ell)^+] = G(\mathbb{F}_\ell)^+$; that is, $G(\mathbb{F}_\ell)^{+\,\mathrm{ab}} = 0$. This fact is probably well-known to specialists. However, for lack of a suitable reference, we sketch the argument.

Without loss of generality, we may assume G is semisimple over $\mathbb{Z}$. Then, for $\ell \gg 0$, $G_{\mathbb{F}_\ell}$ coincides with the algebraic envelope (in the sense of Nori [1987]) $G(\mathbb{F}_\ell)^\sim \subset \mathrm{GL}_{n,\mathbb{F}_\ell}$ of $G(\mathbb{F}_\ell)$ in $\mathrm{GL}_{n,\mathbb{F}_\ell}$. More precisely, if $P_1, \ldots, P_r \in \mathbb{Z}[X_{i,j}, Y]$ are polynomial equations defining $G \subset \mathrm{GL}_{n,\mathbb{Z}} \simeq M_n(\mathbb{Z})[1/\det] \subset \mathbb{Z}^{n^2+1}$ for every $g \in G(\mathbb{F}_\ell)$ of order ℓ, the polynomial $P_{i,g}(T) := P_i\big(\exp(T\log(g))\big) \in \mathbb{F}_\ell[T]$ has degree bounded from above by a constant δ_i independent of ℓ. As $P_{i,g}(T)$ has at least ℓ distinct roots, this forces $P_{i,g}(T) = 0$ as soon as $\ell > \delta_i$. In other words, $G_{\mathbb{F}_\ell}$ contains the one-parameter subgroup

$$e_g : \mathbb{A}^1_{\mathbb{F}_\ell} \to \mathrm{GL}_{n,\mathbb{F}_\ell}, \quad t \mapsto \exp(t\log(g)).$$

So, for $\ell > \max\{\delta_1, \dots, \delta_r\}$, $G_{\mathbb{F}_\ell}$ contains $G(\mathbb{F}_\ell)^\sim$. On the other hand [Nori 1987, Theorem B], for $\ell \gg 0$,

$$G(\mathbb{F}_\ell)^\sim(\mathbb{F}_\ell)^+ = G(\mathbb{F}_\ell)^+. \tag{†}$$

As $[G(\mathbb{F}_\ell)^\sim(\mathbb{F}_\ell) : G(\mathbb{F}_\ell)^\sim(\mathbb{F}_\ell)^+] \le 2^{n-1}$ (because $G(\mathbb{F}_\ell)^\sim$ is exponentially generated) and $[G(\mathbb{F}_\ell) : G(\mathbb{F}_\ell)^+] \le 2^{n-1}$ (because $G_{\mathbb{F}_\ell}$ is semisimple) [Nori 1987, Remark 3.6], (†) implies that $G(\mathbb{F}_\ell)^\sim$ and $G_{\mathbb{F}_\ell}$ have the same dimension [Nori 1987, Lemma 3.5]. As $G_{\mathbb{F}_\ell}$ is connected, this eventually yields $G(\mathbb{F}_\ell)^\sim = G_{\mathbb{F}_\ell}$ as claimed.

As $G(\mathbb{F}_\ell)^\sim = G_{\mathbb{F}_\ell}$ is semisimple, $G(\mathbb{F}_\ell)^\sim$ acts semisimply on $\mathbb{F}_\ell^{\oplus n}$ for $\ell \gg 0$ [Jantzen 1997, Proposition 3.2]. Since for $\ell \gg 0$ the $G(\mathbb{F}_\ell)^\sim$-submodules and $G(\mathbb{F}_\ell)^+$-submodules of $\mathbb{F}_\ell^{\oplus n}$ coincide, this in turn implies that $G(\mathbb{F}_\ell)^+$ acts semisimply on $\mathbb{F}_\ell^{\oplus n}$. The conclusion then follows from [Cadoret and Tamagawa 2014, Lemma 3.4].

This concludes the proof of Theorem 5.4. □

Lemma 5.5. *Let $f : G_2 \to G_1$ be an isogeny of connected semisimple algebraic groups over $\mathbb{Q}$. Let $\Gamma_1 \subset G_1(\mathbb{Q})$ and $\Gamma_2 \subset G_2(\mathbb{Q})$ be arithmetic subgroups such that $f(\Gamma_2) \subset \Gamma_1$. Then for every closed subgroup $\Delta \subset \Gamma_2^-$, the image $f(\Delta) \subset \Gamma_1^-$ is open if and only if $\Delta \subset \Gamma_2^-$ is open.*

Proof. This follows from Lemma 5.2 and the fact that a profinite group is compact and that its open subgroups are exactly its closed subgroups of finite index. As $f(\Gamma_2) \subset G_1(\mathbb{Q})$ is arithmetic, $f(\Gamma_2)$ and Γ_1 are commensurable. This implies that $f(\Gamma_2)^-$ and Γ_1^- are commensurable as well, and hence that $f(\Gamma_2)^-$ is open in Γ_1^-. But as Γ_2^- is compact, $f(\Gamma_2^-) = f(\Gamma_2)^-$. This shows that $f : \Gamma_2^- \to \Gamma_1^-$ is a morphism of profinite groups with open image and finite kernel (since $f : G_2 \to G_1$ is an isogeny). In particular, if $\Delta \subset \Gamma_2^-$ has finite index then $f(\Delta) \subset \Gamma_1^-$ also has finite index: $[\Gamma_1^- : f(\Delta)] \le [\Gamma_1^- : f(\Gamma_2^-)][\Gamma_2^- : \Delta]$. Conversely, if $f(\Delta) \subset \Gamma_1^-$ has finite index then $\Delta \subset \Gamma_2^-$ also has finite index:

$$[\Gamma_2^- : \Delta] \le |\ker(f)|[f(\Gamma_2^-) : f(\Delta)] \le |\ker(f)|[\Gamma_1^- : f(\Delta)]. \qquad \square$$

6. Galois-generic points for adelic representations attached to Shimura varieties

Let (G, X) be a Shimura datum.

6.1. ***Comparison with Pink's definition.*** Let $E := E(G, X)$ denote the reflex field. A point $s \in \mathrm{Sh}(G, X)$ is Galois-generic in the sense of Pink [2005, Definition 6.3] if and only if $s_{E^{\mathrm{ab}}} \in \mathrm{Sh}(G, X)_{E^{\mathrm{ab}}}$ is Galois-generic in the sense of Section 4.4. With the notation of Lemma 3.2.4, one has $\tilde{E} \subset E^{\mathrm{ab}}$, but in general the extension $\tilde{E} \subset E^{\mathrm{ab}}$ is not finite (when the reciprocity map describing the action of $\pi_1(E^{\mathrm{ab}})$ on $\pi_0(\mathrm{Sh}(G, X))$ has infinite kernel). Still, the two notions of Galois-genericity coincide. More precisely, for $s \in \mathrm{Sh}(G, X)$ let $s^{\mathrm{ad}} \in \mathrm{Sh}(G^{\mathrm{ad}}, X^{\mathrm{ad}})$ denote its image by the canonical morphism

$$\mathrm{Sh}(G, X) \to \mathrm{Sh}(G^{\mathrm{ad}}, X^{\mathrm{ad}}).$$

Proposition 6.1.1. *For every $s \in \mathrm{Sh}(G, X)$, the following properties are equivalent*:

(1) *$s \in \mathrm{Sh}(G, X)$ is Galois-generic* (*resp. ℓ-Galois-generic*).

(2) *$s_{E^{\mathrm{ab}}} \in \mathrm{Sh}(G, X)_{E^{\mathrm{ab}}}$ is Galois-generic* (*resp. ℓ-Galois-generic*).

(3) *$s^{\mathrm{ad}} \in \mathrm{Sh}(G^{\mathrm{ad}}, X^{\mathrm{ad}})$ is Galois-generic* (*resp. ℓ-Galois-generic*).

(4) *$s^{\mathrm{ad}}_{E^{\mathrm{ab}}} \in \mathrm{Sh}(G^{\mathrm{ad}}, X^{\mathrm{ad}})_{E^{\mathrm{ab}}}$ is Galois-generic* (*resp. ℓ-Galois-generic*).

Proof. The equivalences (1) $\Leftrightarrow$ (2) and (3) $\Leftrightarrow$ (4) follow from Lemma 3.2.4, from the fact that $\overline{\Pi} = \Gamma^- \subset G(\mathbb{A}_f)$ and from Theorem 5.4. For (2) $\Leftrightarrow$ (4), we may assume that the connected component of $s_{\mathbb{C}}$ is of the form $G(\mathbb{Q}) \setminus X^+ \times \{1\}$ (see Section 4.4). Fix neat compact open subgroups $K \subset G(\mathbb{A}_f)$ and $K^{\mathrm{ad}} \subset G^{\mathrm{ad}}(\mathbb{A}_f)$ such that $p^{\mathrm{ad}}(K) \subset K^{\mathrm{ad}}$. Write $\Gamma := K \cap G(\mathbb{Q})$ and $\Gamma^{\mathrm{ad}} := K^{\mathrm{ad}} \cap G^{\mathrm{ad}}(\mathbb{Q})$; we may assume $\Gamma \subset G^{\mathrm{der}}(\mathbb{Q})$. As K is neat, Γ maps injectively into Γ^{ad}. Then the geometrically connected component S of $s[K]$ in $\mathrm{Sh}_K(G, X)$ is an étale cover of the geometrically connected component S^{ad} of $s^{\mathrm{ad}}[K^{\mathrm{ad}}]$ in $\mathrm{Sh}_{K^{\mathrm{ad}}}(G^{\mathrm{ad}}, X^{\mathrm{ad}})$. The functoriality of adelic representations attached to Shimura varieties yields the following commutative diagram:

$$\begin{array}{ccccccc}
\pi_1(s[K]_{E^{\mathrm{ab}}}) & \longrightarrow & \pi_1(S_{E^{\mathrm{ab}}}) & \xrightarrow{\rho[K,s]} & \Gamma^- & \hookrightarrow & G(\mathbb{A}_f) \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\pi_1(s[K^{\mathrm{ad}}]^{\mathrm{ad}}_{E^{\mathrm{ab}}}) & \longrightarrow & \pi_1(S_{E^{\mathrm{ab}}}) & \xrightarrow[\rho[K^{\mathrm{ad}},s^{\mathrm{ad}}]]{} & (\Gamma^{\mathrm{ad}})^- & \hookrightarrow & G^{\mathrm{ad}}(\mathbb{A}_f)
\end{array}$$

The conclusion then follows from the fact that the left vertical arrow has open image, and from Lemma 5.5 applied to the isogeny $G^{\mathrm{der}} \to G^{\mathrm{ad}}$ and to $\Delta = \Pi_{s[K]_{E^{\mathrm{ab}}}} \subset \Gamma^-$. The proof for ℓ-Galois-generic points is similar. $\square$

6.2. *Galois-generic versus Hodge-generic points.* We say that $x \in X$ is *Hodge-generic* if $\mathrm{MT}(x) = G$. Let $X^{\mathrm{hg}} \subset X$ denote the subset of Hodge-generic points. The set X^{hg} is analytically dense in X and $G(\mathbb{Q})X^{\mathrm{hg}} = X^{\mathrm{hg}}$. Let $K \subset G(\mathbb{A}_f)$ be a neat compact open subgroup. We say that $s \in \mathrm{Sh}(G, X)$ is *Hodge-generic* if $s_{\mathbb{C}} \in G(\mathbb{Q}) \setminus X^{\mathrm{hg}} \times G(\mathbb{A}_f) \subset \mathrm{Sh}(G, X)(\mathbb{C})$, and $s[K] \in \mathrm{Sh}_K(G, X)$ is *Hodge-generic* if $s[K]_{\mathbb{C}} \in G(\mathbb{Q}) \setminus X^{\mathrm{hg}} \times G(\mathbb{A}_f)/K \subset \mathrm{Sh}_K(G, X)(\mathbb{C})$.

Proposition 6.2.1. *All ℓ-Galois-generic points are Hodge-generic.*

Proof. (See also [Pink 2005, Proposition 6.7].) Let $s \in \mathrm{Sh}(G, X)$ be ℓ-Galois-generic and $(x, g) \in X \times G(\mathbb{A}_f)$ lifting $s_{\mathbb{C}} \in \mathrm{Sh}(G, X)(\mathbb{C})$. We may assume $g = 1$ (see

Section 4.4). Let X_x denote the $\mathrm{MT}(x)(\mathbb{R})$-conjugacy class of $x : \mathbb{S} \to \mathrm{MT}(x)_{\mathbb{R}}$. The inclusion $\mathrm{MT}(x) \hookrightarrow G$ induces a morphism of Shimura data $(\mathrm{MT}(x), X_x) \to (G, X)$, and hence a morphism of Shimura varieties $\mathrm{Sh}(\mathrm{MT}(x), X_x) \to \mathrm{Sh}(G, X)$ and, for every neat compact open subgroup $K \subset G(\mathbb{A}_f)$, a morphism $\mathrm{Sh}_{K_x}(\mathrm{MT}(x), X_x) \to \mathrm{Sh}_K(G, X)$, where we set $K_x := K \cap \mathrm{MT}(x)(\mathbb{A}_f)$. Assume $K = \prod_\ell K_\ell$ and write $\Gamma := K \cap G(\mathbb{Q})_+$. Let $s_x \in \mathrm{Sh}(\mathrm{MT}(x), X_x)$ be the point corresponding to the image of x and let $E[K_x]$ denote the field of definition of the geometrically connected component $S[K_x, s_x]$ of $s_x[K_x]$ in $\mathrm{Sh}_{K_x}(\mathrm{MT}(x), X_x)$. The following diagram commutes (Section 4.1):

$$\begin{array}{ccccc} \pi_1(s[K_x]) & \longrightarrow & \pi_1(S[K_x, s_x]) & \longrightarrow & K_x \\ \downarrow & & \downarrow & & \downarrow \\ \pi_1(s[K]) & \longrightarrow & \pi_1(S[K, s]_{E[K_x]}) & \longrightarrow & K \end{array}$$

As $s \in \mathrm{Sh}(G, X)$ is ℓ-Galois-generic and the image of

$$\pi_1(s[K_x]) \to \pi_1(s[K]) \to \pi_1(S[K, s]_{E[K_x]})$$

is open in the image of $\pi_1(s[K]) \to \pi_1(S[K, s]_{E[K_x]})$, the commutativity of the above diagram implies that $K[x, \ell] \subset \mathrm{MT}(x)(\mathbb{Q}_\ell)$ contains an open subgroup of Γ_ℓ^-. As such a subgroup is Zariski-dense in $G^{\mathrm{der}}_{\mathbb{Q}_\ell}$, this shows that $\mathrm{MT}(x)_{\mathbb{Q}_\ell}$ contains $G^{\mathrm{der}}_{\mathbb{Q}_\ell}$, and hence that $\mathrm{MT}(x)$ contains G^{der}. In particular, $\mathrm{MT}(x)$ is normal in G. So every $G(\mathbb{R})$-conjugate of $x : \mathbb{S} \to \mathrm{MT}(x)_{\mathbb{R}}$ factors through $\mathrm{MT}(x)_{\mathbb{R}}$. Since G is the generic Mumford–Tate group of (G, X), this forces $\mathrm{MT}(x) = G$. □

6.2.2. Pink [2005, Conjecture 6.8] conjectures that Hodge-generic points are Galois-generic. Thus, Theorem A reduces Pink's conjecture to proving that every Hodge-generic point is ℓ-Galois generic which, in the case of abelian schemes, is precisely the statement of the standard (ℓ-adic) Mumford–Tate conjecture.

6.3. *Proof of Theorem A.*

6.3.1. *Shimura data of Hodge type.* Recall that a Shimura datum (G_2, X_2) (as well as the associated Shimura variety) is said to be of Hodge type if there exists an embedding of Shimura data $f : (G_2, X_2) \hookrightarrow (G_1, X_1)$ with (G_1, X_1) a Siegel Shimura datum. Let $K_i \subset G_i(\mathbb{A}_f)$ $(i = 1, 2)$ be neat compact open subgroups such that $f(K_2) \subset K_1$. Let $s_2 \in \mathrm{Sh}(G_2, X_2)$ and write $s_1 := f(s_2) \in \mathrm{Sh}(G_1, X_1)$. For simplicity, write $S_i := S[K_i, s_i]$ for the geometrically connected component of $s_i[K_i]$ and let E_i denote its field of definition for $i = 1, 2$. Let $A \to S_1$ denote the universal abelian scheme over S_1. Then the adelic representation

$$\pi_1(S_2) \xrightarrow{f} \pi_1(S_{1E_2}) \xrightarrow{\rho[K_1, s_1]} K_1$$

coincides with the adelic representation attached to the abelian scheme $A|_{S_2} := A \times_{S_{1E_2}} S_2 \to S_2$. But, as $f : K_1 \hookrightarrow K_2$ is injective, one sees from Section 4.1 that

$$\pi_1(S_2) \xrightarrow{f} \pi_1(S_{1E_2}) \xrightarrow{\rho[K_1,s_1]} K_1$$

and $\rho[K_2, s_2] : \pi_1(S_2) \to K_2 \hookrightarrow K_1$ have the same Galois-generic and ℓ-Galois-generic points. So, consider the following assertions for $s_2[K_2] \in S_2$:

(1) Galois-generic with respect to $\rho[K_2, s_2] : \pi_1(S_2) \to K_2 \subset G_2(\mathbb{A}_f)$;

(2) ℓ-Galois-generic with respect to $\rho[K_2, s_2] : \pi_1(S_2) \to K_2 \subset G_2(\mathbb{A}_f)$;

(3) Galois-generic with respect to the adelic representation attached to $A|_{S_2} \to S_2$.

(4) ℓ-Galois-generic with respect to the adelic representation attached to $A|_{S_2} \to S_2$.

Then, from the above, (1) $\Leftrightarrow$ (3), (2) $\Leftrightarrow$ (4) and from Theorem 3.3.2.2, (3) $\Leftrightarrow$ (4). This shows Theorem A for Shimura data of Hodge type.

6.3.2. *Shimura data of abelian type.* Recall that a Shimura datum (G_1, X_1) (as well as the associated Shimura variety) is said to be of abelian type if there exists a Shimura datum (G_2, X_2) of Hodge type and an isogeny $f : G_2^{\mathrm{der}} \to G_1^{\mathrm{der}}$ which induces an isomorphism of adjoint Shimura data $f : (G_2^{\mathrm{ad}}, X_2^{\mathrm{ad}}) \tilde{\to} (G_1^{\mathrm{ad}}, X_1^{\mathrm{ad}})$. We refer to [Deligne 1971, §1.2, 1.3 and 2.7] and [Milne 2013, §10] for a detailed account of Shimura data of abelian type. These include essentially all Shimura data (G, X) except those for which G has simple factors of type E_6, E_7 and certain type D.

We can now conclude the proof of Theorem A. Start from an ℓ-Galois-generic point $s_1 \in \mathrm{Sh}(G_1, X_1)$. We may assume (Section 4.4) that the image s_1^{ad} of s_1 in $\mathrm{Sh}(G_1^{\mathrm{ad}}, X_1^{\mathrm{ad}}) \simeq \mathrm{Sh}(G_2^{\mathrm{ad}}, X_2^{\mathrm{ad}})$ lies in the image of $\mathrm{Sh}(G_2, X_2) \to \mathrm{Sh}(G_2^{\mathrm{ad}}, X_2^{\mathrm{ad}})$. Fix $s_2 \in \mathrm{Sh}(G_2, X_2)$ above s_1^{ad}. Because $s_1 \in \mathrm{Sh}(G_1, X_1)$ is ℓ-Galois-generic, $s_1^{\mathrm{ad}} \in \mathrm{Sh}(G_1^{\mathrm{ad}}, X_1^{\mathrm{ad}})$ is ℓ-Galois-generic (by (1) $\Rightarrow$ (4) in Proposition 6.1.1) and $s_2 \in \mathrm{Sh}(G_2, X_2)$ is ℓ-Galois-generic (by (4) $\Rightarrow$ (1) in Proposition 6.1.1). As (G_2, X_2) is of Hodge type, $s_2 \in \mathrm{Sh}(G_2, X_2)$ is Galois-generic by Section 6.3.1. Thus $s_1^{\mathrm{ad}} \in \mathrm{Sh}(G_1^{\mathrm{ad}}, X_1^{\mathrm{ad}})$ is Galois-generic (by (1) $\Rightarrow$ (4) in Proposition 6.1.1) and $s_1 \in \mathrm{Sh}(G_1, X_1)$ is Galois-generic (by (4) $\Rightarrow$ (1) in Proposition 6.1.1). □

7. Proof of Theorem B

7.1. ***Generalized Hecke operators.*** Let (G, X) be a Shimura datum. Let $X^+ \subset X$ be a connected component and let $K \subset G(\mathbb{A}_f)$ be a neat compact open subgroup. Write $\Gamma := G(\mathbb{Q})_+ \cap K$. For every $a \in G(\mathbb{A}_f)$, let T_a denote the Hecke operator $\cdot a^{-1} : \mathrm{Sh}(G, X)_{\mathbb{C}} \tilde{\to} \mathrm{Sh}(G, X)_{\mathbb{C}}$ and $T_{a,K}$ the corresponding algebraic correspondence

$$\mathrm{Sh}_K(G, X)_{\mathbb{C}} \leftarrow \mathrm{Sh}(G, X)_{\mathbb{C}} \overset{\cdot a^{-1}}{\tilde{\to}} \mathrm{Sh}(G, X)_{\mathbb{C}} \to \mathrm{Sh}_K(G, X)_{\mathbb{C}}.$$

It is part of the definition of a canonical model that $T_{a,K}$ is defined over the reflex field $E := E(G, X)$. For $a \in G(\mathbb{Q})_+ K$, $T_{a,K}$ restricts to an algebraic correspondence $T_{a,\Gamma}$ (of degree 1 if $a \in K$) on $\mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C}$, defined over $E_{K^{a^{-1}}}$ (see Section 4.2 for the notation):

$$\mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C} \leftarrow \mathrm{Sh}(G, X^+)_\mathbb{C} \overset{a\cdot}{\tilde{\to}} \mathrm{Sh}(G, X^+)_\mathbb{C} \to \mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C}.$$

Here, we write

$$\mathrm{Sh}(G, X^+)_\mathbb{C} := \varprojlim_\Gamma \mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C},$$

where the projective limit is taken over all congruence subgroups $\Gamma \subset G(\mathbb{Q})$. Recall that $\mathrm{Sh}(G, X^+)_\mathbb{C}$ identifies with the connected component of $\mathrm{Sh}(G, X)$ containing the image of $X^+ \times \{1\}$ [Deligne 1971, §1.8]. For $s = \Gamma x \in \mathrm{Sh}_\Gamma(G, X^+)(\mathbb{C})$ we set

$$T_{a,\Gamma}(s) = \{\Gamma a \gamma x \mid \gamma \in \Gamma\}.$$

More generally, for a subset $A \subset G(\mathbb{Q})_+$, we set

$$T_{A,\Gamma}(s) := \bigcup_{a\in A} T_{a,\Gamma}(s) \subset \mathrm{Sh}_\Gamma(G, X^+)(\mathbb{C})$$

for its A-Hecke orbit. For $A = G(\mathbb{Q})_+$ we simply write $T_{A,\Gamma}(s) =: T_\Gamma(s)$ for the full Hecke orbit of s.

Let $\mathrm{Aut}(G, X^+)$ denote the group automorphisms of G defined over $\mathbb{Q}$ and stabilizing X^+. For every $\phi \in \mathrm{Aut}(G, X^+)$, the corresponding generalized Hecke operator T_ϕ is the algebraic correspondence

$$\mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C} \leftarrow \mathrm{Sh}(G, X^+)_\mathbb{C} \overset{\phi\cdot}{\tilde{\to}} \mathrm{Sh}(G, X^+)_\mathbb{C} \to \mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C}.$$

For $s = \Gamma x \in \mathrm{Sh}_\Gamma(G, X^+)(\mathbb{C})$ we set

$$T_{\phi,\Gamma}(s) = \{\Gamma \phi(\gamma)\phi(x) \mid \gamma \in \Gamma\}.$$

More generally, for a subset $\Phi \subset \mathrm{Aut}(G, X^+)$, we set

$$T_{\Phi,\Gamma}(s) := \bigcup_{\phi\in\Phi} T_{\phi,\Gamma}(s) \subset \mathrm{Sh}_\Gamma(G, X^+)(\mathbb{C})$$

for its Φ-Hecke orbit. The usual full Hecke orbit $T_\Gamma(s)$ coincides with the Φ-Hecke orbit $T_{\Phi,\Gamma}(s)$ for Φ the image of $G(\mathbb{Q})_+ \to \mathrm{Aut}(G, X^+)$ given by inner automorphisms. For $\Phi = \mathrm{Aut}(G, X^+)$ we simply write $T_{\Phi,\Gamma}(s) =: \widehat{T}_\Gamma(s)$ for the full generalized Hecke orbit of s.

If Γ is obvious from the context, we will omit it from the notation. The above definitions of Hecke orbits extend as they are to arithmetic subgroups $\Gamma \subset G(\mathbb{Q})_+$.

For the comparison between usual Hecke orbits and generalized Hecke orbits, see [Orr 2013, §4.1.1]; let us only point out the following observation, which will be used in the proof of Theorem B.

Assume G is adjoint. For an arithmetic subgroup $\Gamma \subset G(\mathbb{Q})^+$, write again $\mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C}$ for the complex algebraic variety underlying $\Gamma \setminus X^+$ [Baily and Borel 1966].

Lemma 7.1.1. *For every congruence (resp. arithmetic) subgroup $\Gamma \subset G(\mathbb{Q})^+$ there exists a congruence (resp. an arithmetic) subgroup $\Gamma' \subset \Gamma$ such that, for every subset $\Phi \subset \mathrm{Aut}(G, X^+)$ and $s \in \mathrm{Sh}_\Gamma(G, X^+)$, the inverse image of $T_{\Phi,\Gamma}(s)$ by $p_{\Gamma',\Gamma} : \mathrm{Sh}_{\Gamma'}(G, X^+) \to \mathrm{Sh}_\Gamma(G, X^+)$ is contained in a finite union of usual Hecke orbits on $\mathrm{Sh}_{\Gamma'}(G, X^+)$.*

Proof. As G is adjoint, the quotient $\mathrm{Aut}(G, X^+)/G(\mathbb{Q})^+$ is finite. Choose a system of representatives $\phi_1, \dots, \phi_r$ for $\mathrm{Aut}(G, X^+)/G(\mathbb{Q})^+$ and set

$$\Gamma' := \bigcap_{1 \le i \le r} \phi_i(\Gamma) \subset \Gamma, \qquad \Gamma'' := \bigcap_{1 \le i \le r} \phi_i(\Gamma') \subset \Gamma'.$$

Note that, by construction, if Γ is a congruence (resp. an arithmetic) subgroup then Γ' and Γ'' are again congruence (resp. arithmetic) subgroups. Fix systems of representatives

- γ_j, $j = 1, \dots, s$, of Γ/Γ';
- α_k, $k = 1, \dots, t$, of Γ/Γ'';
- and $\alpha_{i,l}$, $l = 1, \dots, t_i$, of $\phi_i(\Gamma')/\Gamma''$ for $i = 1, \dots, r$.

For an arbitrary element $\phi = a - a^{-1} \circ \phi_i \in \mathrm{Aut}(G, X^+)$, we can then compute explicitly

$$p^{-1}_{\Gamma'',\Gamma}(T_{\phi,\Gamma}(s)) = \bigcup_{1 \le j \le s} \bigcup_{1 \le k \le t} \bigcup_{1 \le l \le t_i} T_{\alpha_k a \alpha_{l,i},\Gamma''}(\phi_i(\gamma_j x)).$$

This shows that for every subset $\Phi \subset \mathrm{Aut}(G, X^+)$ and point $s \in \mathrm{Sh}_\Gamma(G, X^+)$, we have that $p^{-1}_{\Gamma'',\Gamma}(T_{\Phi,\Gamma}(s))$ is contained in a finite union of usual Hecke orbits on $\mathrm{Sh}_{\Gamma''}(G, X^+)$. □

7.2. *Equidistribution.* Let G be a connected $\mathbb{Q}$-simple algebraic group of noncompact type and let $G(\mathbb{R})^+ \subset G(\mathbb{R})$ denote the connected component of 1 in $G(\mathbb{R})$. Set $G(\mathbb{Q})^+ := G(\mathbb{Q}) \cap G(\mathbb{R})^+$. Fix an arithmetic subgroup $\Gamma \subset G(\mathbb{Q})^+$. Then Γ is a lattice in $G(\mathbb{R})^+$ [Borel and Harish-Chandra 1962]; let μ denote the normalized Haar measure on $\Gamma \setminus G(\mathbb{R})^+$. For a function $f : \Gamma \setminus G(\mathbb{R})^+ \to \mathbb{C}$ and an element $a \in G(\mathbb{Q})^+$, define its Hecke transform $T_a(f) : \Gamma \setminus G(\mathbb{R})^+ \to \mathbb{C}$ by

$$T_a(f)(y) = \frac{1}{\deg_\Gamma(a)} \sum_{y' \in T_a(y)} f(y'),$$

where $\deg_\Gamma(a) = |\Gamma a \Gamma/\Gamma| = [\Gamma : \Gamma \cap a\Gamma a^{-1}]$.

Fact 7.2.1 ([Eskin and Oh 2006, Theorem 1.2]; see also [Burger and Sarnak 1991, "Theorem 5.2"]). *For every sequence $\underline{a} = (a_n)$ of elements in $G(\mathbb{Q})^+$ with*

$$\lim_{n\to+\infty} \deg_\Gamma(a_n) = +\infty$$

and for every continuous bounded function $f : \Gamma \setminus G(\mathbb{R})^+ \to \mathbb{C}$ and $y \in \Gamma \setminus G(\mathbb{R})^+$, we have

$$\lim_{n\to+\infty} T_{a_n}(f)(y) = \int_{\Gamma\setminus G(\mathbb{R})^+} f\, d\mu.$$

In particular, for every $y \in \Gamma \setminus G(\mathbb{R})^+$ the set

$$T_{\underline{a}}(y) := \bigcup_{n\geq 1} T_{a_n}(y)$$

is dense in $\Gamma\setminus G(\mathbb{R})^+$. To exploit Fact 7.2.1, we need the following general finiteness result about Hecke operators of bounded degree.

Theorem 7.2.2. *Let G be a connected semisimple group over $\mathbb{Q}$ of noncompact type*[2] *and let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic subgroup. Then for every integer $d \geq 1$ there are only finitely many double-classes $\Gamma a \Gamma \in \Gamma \setminus G(\mathbb{Q})/\Gamma$ with $\deg_\Gamma(a) \leq d$.*

Theorem 7.2.2 will be proved in Section 8.

Let (G, X) be a Shimura datum. Fix a connected component $X^+ \subset X$ and a neat arithmetic subgroup $\Gamma \subset G^{\mathrm{der}}(\mathbb{Q})_+$.

Corollary 7.2.3 (compare with [Pink 2005, Theorem 7.5]). *Assume G is almost $\mathbb{Q}$-simple. Then for every $s_\Gamma \in \mathrm{Sh}_\Gamma(G, X^+)$ and for every sequence $\underline{\phi} = (\phi_n)$ in $\mathrm{Aut}(G, X^+)$, the set $T_{\underline{\phi},\Gamma}(s_\Gamma)$ is either finite or Zariski-dense in $\mathrm{Sh}_\Gamma(G, X^+)$.*

Proof. Let X^{ad} denote the $G^{\mathrm{ad}}(\mathbb{R})$-conjugacy class of $p^{\mathrm{ad}} \circ x : \mathbb{S} \to G^{\mathrm{ad}}_{\mathbb{R}}$ for one (equivalently every) $x \in X$. Then $p^{\mathrm{ad}} : X \hookrightarrow X^{\mathrm{ad}}$ identifies X with a union of connected components of X^{ad}. As $\Gamma \subset G^{\mathrm{der}}(\mathbb{Q})_+$, $p^{\mathrm{ad}} : G \to G^{\mathrm{ad}}$ maps Γ bijectively onto its image $\Gamma^{\mathrm{ad}} := p^{\mathrm{ad}}(\Gamma)$. Thus the morphism of Shimura data $p^{\mathrm{ad}} : (G, X) \to (G^{\mathrm{ad}}, X^{\mathrm{ad}})$ induces an isomorphism of schemes over $\mathbb{C}$

$$p^{\mathrm{ad}} : \mathrm{Sh}_\Gamma(G, X^+) \xrightarrow{\sim} \mathrm{Sh}_{\Gamma^{\mathrm{ad}}}(G^{\mathrm{ad}}, X^+),$$

and this isomorphism maps $T_{\underline{\phi},\Gamma}(s_\Gamma)$ bijectively onto $T_{p^{\mathrm{ad}}(\underline{\phi}),\Gamma^{\mathrm{ad}}}(p^{\mathrm{ad}}(s_\Gamma))$, where we write again

$$p^{\mathrm{ad}} : \mathrm{Aut}(G, X^+) \to \mathrm{Aut}(G^{\mathrm{ad}}, X^+)$$

for the morphism of groups induced by $p^{\mathrm{ad}} : G \to G^{\mathrm{ad}}$. Thus, we may assume G is adjoint. Next, for every arithmetic subgroup $\Gamma' \subset \Gamma$, the quotient map

[2] If we remove the assumption that G is of noncompact type, Theorem 7.2.2 becomes trivially false. Indeed, if G is $\mathbb{Q}$-simple of compact type then Γ is always finite while $G(\mathbb{Q})$ is always infinite.

$p_{\Gamma',\Gamma} : \mathrm{Sh}_{\Gamma'}(G, X^+) \to \mathrm{Sh}_\Gamma(G, X^+)$ is a finite cover so it is enough to prove that $p^{-1}_{\Gamma',\Gamma}(T_{\underline{\phi},\Gamma}(s_\Gamma))$ is finite. In particular, by Lemma 7.1.1 we may assume that $\phi_n = a_n \in G(\mathbb{Q})_+$, $n \geq 0$. Then, assume that the Zariski-closure Z of $T_{\underline{a},\Gamma}(s_\Gamma)$ in $\mathrm{Sh}_\Gamma(G, X^+)$ is a strict closed subscheme. So $Z(\mathbb{C}) \subsetneq \mathrm{Sh}_\Gamma(G, X^+)(\mathbb{C}) = \Gamma \setminus X^+$ is a strict closed analytic subset. As the canonical map $p_\Gamma : \Gamma \setminus G(\mathbb{R})^+ \twoheadrightarrow \Gamma \setminus X^+$ is analytic and surjective, $p_\Gamma^{-1}(Z(\mathbb{C})) \subsetneq \Gamma \setminus G(\mathbb{R})^+$ is again a strict closed analytic subset. As $p_\Gamma^{-1}(T_{\underline{a},\Gamma}(s_\Gamma)) = T_{\underline{a}}(1) \subset p_\Gamma^{-1}(Z(\mathbb{C}))$, we have

$$T_{\underline{a}}(1)^- \subset p_\Gamma^{-1}(Z(\mathbb{C}))^- = p_\Gamma^{-1}(Z(\mathbb{C})) \subsetneq \Gamma \setminus G(\mathbb{R})^+.$$

Then by Fact 7.2.1, $\deg_\Gamma(a_n)$ is bounded which, in turn, implies by Theorem 7.2.2 that the set

$$\{\Gamma a_n \Gamma \mid n \geq 0\} \subset \Gamma \setminus G(\mathbb{Q}) / \Gamma$$

is finite. Hence $T_{\underline{a},\Gamma}(s_\Gamma) = p_\Gamma(T_{\underline{a}}(1))$ is finite as well. □

7.3. *Proof of Theorem B.* (Compare with [Pink 2005, proof of Theorem 7.6]). Fix a Galois-generic point $s \in \mathrm{Sh}(G, X)$ such that $s_\mathbb{C} = G(\mathbb{Q})(x, 1)$, and let $X^+ \subset X$ denote the connected component of x. Setting $\Gamma_0 := K_0 \cap G(\mathbb{Q})_+$, we then have $S[K_0, s] = \mathrm{Sh}_{\Gamma_0}(G, X^+)$. Recall from Proposition 6.1.1 that $s_{E^{\mathrm{ab}}} \in \mathrm{Sh}(G, X)_{E^{\mathrm{ab}}}$ is again Galois-generic. So, to prove Theorem B, we may and will work over E^{ab} (without mentioning it explicitly in the notation, for simplicity). Then $\rho[K_0] : \pi_1(\mathrm{Sh}_{\Gamma_0}(G, X^+)) \twoheadrightarrow \Gamma_0^-$ is completely determined by the tower of connected étale covers

$$p_{\Gamma,\Gamma_0} : \mathrm{Sh}_\Gamma(G, X^+) \to \mathrm{Sh}_{\Gamma_0}(G, X^+).$$

Let $Z \hookrightarrow \mathrm{Sh}_\Gamma(G, X^+)_\mathbb{C}$ be a closed subvariety containing an infinite subset T of $\widehat{T}_\Gamma(s)$.

• *Reduction to the case where $s[K_0] \in \mathrm{Sh}_{\Gamma_0}(G, X^+)$ is strictly Galois-generic and Z is defined over the residue field $k(s[K_0])$ of $s[K_0]$.* As all the points in $\widehat{T}_{\Gamma_0}(s)$ are defined over the algebraic closure of $k(s[K_0])$, up to replacing Z by the irreducible components of the Zariski closure of T in $\mathrm{Sh}_{\Gamma_0}(G, X^+)$, we may assume Z is defined over a finite field extension F of $k(s[K_0])$. As $s[K_0]_F \in \mathrm{Sh}_{\Gamma_0}(G, X^+)_F$ is again Galois-generic, up to replacing $s[K_0]$ by $s[K_0]_F$, we may assume $F = k(s[K_0])$. Then we can find a congruence subgroup $\Gamma := K \cap G(\mathbb{A}_f) \subset \Gamma_0$ with $\Gamma^- \subset \Pi_{s[K_0]}$. Hence $s[K_0]$ lifts to a strictly Galois-generic point $s[K] \in \mathrm{Sh}_\Gamma(G, X^+)_F$; see 3.2.6. Then $p^{-1}_{\Gamma,\Gamma_0} Z \hookrightarrow \mathrm{Sh}_\Gamma(G, X^+)_F$ is again a closed subvariety which contains an infinite subset of $\widehat{T}_\Gamma(s[K])$ (just observe that, for every $\phi \in \mathrm{Aut}(G, X^+)$,

$$p^{-1}_{\Gamma,\Gamma_0}(T_{\phi,\Gamma_0}(s[K_0])) = \bigcup_{\gamma_1,\gamma_2 \in \Gamma_0/\Gamma} T_{\phi_{\gamma_1,\gamma_2},\Gamma}(s[K]_F),$$

where $\phi_{\gamma_1,\gamma_2} \in \mathrm{Aut}(G, X^+)$ is defined by $\phi_{\gamma_1,\gamma_2}(g) = \gamma_1\phi(\gamma_2)\phi(g)(\gamma_1\phi(\gamma_2))^{-1}$, $g \in G$). As it is enough to show that $p^{-1}_{\Gamma,\Gamma_0} Z = \mathrm{Sh}_\Gamma(G, X^+)_F$, up to replacing E^{ab} with F, $\mathrm{Sh}_{\Gamma_0}(G, X^+)$ with $\mathrm{Sh}_\Gamma(G, X^+)_F$, $s[K_0]$ with $s[K]$ and Z with $p^{-1}_{\Gamma,\Gamma_0} Z$, without loss of generality we may assume that $s[K_0] \in \mathrm{Sh}_{\Gamma_0}(G, X^+)$ is strictly Galois-generic and that Z is defined over $F = k(s[K_0])$.

• *Conclusion.* As $s[K_0]$ is strictly Galois-generic, for every generalized Hecke operator $\phi \in \mathrm{Aut}(G, X^+)$, the group $\pi_1(k(s[K_0]))$ acts transitively on $T_{\phi,\Gamma_0}(s[K_0]) \subset \mathrm{Sh}_{\Gamma_0}(G, X^+)$. In particular, for each $\phi \in \mathrm{Aut}(G, X^+)$ such that $Z \cap T_{\phi,\Gamma_0}(s[K_0]) \neq \varnothing$ we have $T_{\phi,\Gamma_0}(s[K_0]) \subset Z$ (recall that we assume Z is defined over $k(s[K_0])$). Now, consider a sequence $\underline{\phi} = (\phi_n)$ of elements in $\mathrm{Aut}(G, X^+)$ such that Z contains a point $s_n \in T_{\phi_n,\Gamma_0}(s[K_0])$, $n \geq 1$ with $s_n \neq s_m$ for $n \neq m$. Then

$$T_{\underline{\phi},\Gamma_0}(s[K_0]) = \bigcup_{n\geq 1} T_{\phi_n,\Gamma_0}(s[K_0]) \subset Z$$

is infinite by construction. So, the conclusion follows from Corollary 7.2.3. □

8. Degree of Hecke operators

8.1. *Formal lemmas.* Let G be a group. For every subgroup $K, K' \subset G$ such that $K \equiv K'$, set

$$[K : K'] := \frac{[K : K \cap K']}{[K' : K \cap K']} \in \mathbb{Q}.$$

For a subgroup $K \subset G$, let $K^{\equiv}$ be the set of all $g \in G$ such that K and gKg^{-1} are commensurable. Then $K^{\equiv_G} \subset G$ is a subgroup containing K and for every subgroup $K' \subset G$ such that $K' \equiv K$, we have $K^{\equiv} = K'^{\equiv}$. For $g \in K^{\equiv}$, define the degree of g with respect to K as

$$\deg_K(g) = [K : K \cap gKg^{-1}] = |K \setminus KgK|.$$

For subgroups $K \subset U \subset G$, let $\mathrm{Cor}_U(K) := \bigcap_{u\in U} uKu^{-1} \subset K$ denote the largest subgroup of K which is normalized by U. Equivalently, $\mathrm{Cor}_U(K)$ is the kernel of U acting on U/K by left translation. In particular, if $[U : K]$ is finite then $[U : \mathrm{Cor}_U(K)]$ is also finite and $[U : \mathrm{Cor}_U(K)] \leq [U : K]!$.

Lemma 8.1.1. *For subgroups $K, K' \subset G$ such that $K \equiv K'$ and $a \in K^{\equiv}$, we have $\deg_K(a) \leq C_{K,K'} \deg_{K'}(a)$, where $C_{K,K'} = [K : \mathrm{Cor}_K(K \cap K')][K' : \mathrm{Cor}_K(K \cap K')]$.*

Proof. Observe first that

(i) if $K' \subset K$ then $\deg_{K'}(a) \leq [K : K'] \deg_K(a)$, and

(ii) if furthermore $K' \subset K$ is normal, then $\deg_K(a) \leq [K : K'] \deg_{K'}(a)$.

The proof of (i) is straightforward. As for the proof of (ii), if $K' \subset K$ is normal then

$$\deg_{K'}(ka) = |K'/K' \cap kaK'(ka)^{-1}| = |K'/k^{-1}K'k \cap aK'a^{-1}| = \deg_{K'}(a) \quad (*)$$

and similarly, $\deg_{K'}(ak) = \deg_{K'}(a)$. Let $R \subset K$ be a set of representatives for the left cosets of K' in K. By normality, R is also a set of representatives for the right cosets of K' in K. Hence,

$$\begin{aligned}[K : K'] \deg_K(a) &= |K' \setminus K||K \setminus KaK| = |K' \setminus KaK| \\ &\leq \sum_{x,y \in R} |K' \setminus K'xayK'| = \sum_{x,y\in R} \deg(xay) = [K : K']^2 \deg_{K'}(a),\end{aligned}$$

where the last equality follows from (∗). The assertion in Lemma 8.1.1 now follows from the combination of (i) and (ii):

$$\begin{aligned}\deg_K(a) &\leq [K : \mathrm{Cor}_K(K \cap K')] \deg_{\mathrm{Cor}_K(K\cap K')}(a) && \text{(by (i))} \\ &\leq [K : \mathrm{Cor}_K(K \cap K')][K' : \mathrm{Cor}_K(K \cap K')] \deg_{K'}(a) && \text{(by (ii)).} \quad \square\end{aligned}$$

Definition 8.1.2. For a subgroup $K \subset G$, we say that property $\star(G, K)$ holds if for every integer $d \geq 1$ there are only finitely many K-double classes $KaK \in K \setminus G/K$ with $\deg_K(a) \leq d$.

Lemma 8.1.3. *Let $\varphi\colon G' \to G$ be a morphism of groups and $K' \subset G'$, $K \subset G$ two subgroups. Assume that* $\ker(\varphi)$, $[G : \varphi(G')]$ *are finite and $K \equiv \varphi(K') \subset G$. Then $\star(G, K)$ holds if and only if $\star(G', K')$ holds.*

Proof. First, consider the case where $G' = G$ and φ is the identity. As the situation is symmetric in K, K', it is enough to show the implication $\star(G, K) \Rightarrow \star(G', K')$. Set $K'' := K \cap K'$ and fix an integer $d \geq 1$. Let $a \in K^{\equiv}(= K'^{\equiv})$ with $\deg_{K'}(a) \leq d$. From Lemma 8.1.1(ii), $\deg_K(a) \leq C_{K',K}d$. From $\star(G, K)$, there are only finitely many possibilities for the K-double class $KaK \in K \setminus G/K$. But, then there are also only finitely many possibilities for the K'-double class $K'aK' \in K' \setminus G/K'$ since the induced maps $K'' \setminus G/K'' \to K \setminus G/K$ and $K'' \setminus G/K'' \to K' \setminus G/K'$ are both surjective with finite fibers.

In particular, $\star(G, K)$ holds if and only if $\star(G, \varphi(K'))$ holds. So, in the following, we may and will assume that $K = \varphi(K')$.

Then the assumption that $\ker(\varphi)$ is finite ensures that the induced map

$$K' \setminus G'/K' \to K \setminus G/K$$

has finite fibers. The implication $\star(G, K) \Rightarrow \star(G', K')$ then follows from the inequality $\deg_K(\varphi(a')) \leq \deg_{K'}(a')$, $a' \in G'$.

For the implication $\star(G', K') \Rightarrow \star(G, K)$, observe that

$$\deg_K(ab) \leq \deg_K(a)\deg_K(b), \quad a, b \in G$$

(just note that $KabK \subset KaKbK$). Let Δ denote a (finite) set of representatives of left cosets of $\varphi(G')$ in G. Then for every $a \in G$ there exists (a unique) $\delta_a \in \Delta$ such that $a\delta_a^{-1} = \varphi(a')$ for some $a' \in G'$. In particular,

$$\deg_{K'}(a') \leq \deg_K(\varphi(a')) \leq \min\{\deg_K(\delta^{-1}) \mid \delta \in \Delta\} \deg_K(a). \qquad \square$$

As a result, to prove Theorem 7.2.2, we may assume that $\Gamma \subset G(\mathbb{Q})$ is a congruence subgroup. So, let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup such that $\Gamma = K \cap G(\mathbb{Q})$. By shrinking K, we may assume K is of the form

$$K = K_{\mathcal{P}} K^{\mathcal{P}}, \tag{8.1.4}$$

where $\mathcal{P}$ is a finite set of primes containing the primes where G ramifies, where $K_{\mathcal{P}} = \prod_{p \in \mathcal{P}} K_p$ with $K_p \subset G(\mathbb{Q}_p)$ compact open and where $K^{\mathcal{P}} = \prod_{p \notin \mathcal{P}} K_p$ with $K_p \subset G(\mathbb{Q}_p)$ hyperspecial.

8.2. *Degree of adelic Hecke operators.* In this section, we reduce the adelic variant of Theorem 7.2.2, Theorem 8.2.1, to statements (Lemmas 8.2.2, 8.2.5) about the degree of local (p-adic) Hecke operators.

Theorem 8.2.1. *Let G be a connected semisimple group over $\mathbb{Q}$. Then $\star(G(\mathbb{A}_f), K)$ holds for every compact open subgroup $K \subset G(\mathbb{A}_f)$.*

Lemma 8.2.2. *Let G be a connected semisimple group over $\mathbb{Q}_p$. Then $\star(G(\mathbb{Q}_p), K)$ holds for every compact open subgroup $K \subset G(\mathbb{Q}_p)$.*

Remark 8.2.3. More precisely, let $n_d(G(\mathbb{Q}_p), K)$ denote the number of double classes KaK for $a \in G(\mathbb{Q}_p)$ with $\deg_K(a) \leq d$, and let $B \subset \mathbb{G}(\mathbb{Q}_p)$ be an Iwahori subgroup. Then

$$n_d(G(\mathbb{Q}_p), K) \leq \left| \left\{ w \in W_G \;\middle|\; \ell(w) \leq \frac{\ln(C_{B,K} d)}{\ln(p)} \right\} \right|,$$

where W_G denotes the affine Weil group of G, $\ell : W_G \to \mathbb{Z}_{\geq 0}$ the length function on it and $C_{B,K}$ is the constant of Lemma 8.1.1.

Lemma 8.2.4 (definition of $\iota(-)$). *Let $G' \to G$ be an isogeny of algebraic groups over $\mathbb{Q}_p$. Assume that the degree N of its kernel μ is at most p. Then there exists a constant $\iota(N)$ depending only on N (but not on p) such that*

$$|\operatorname{coker}(G'(\mathbb{Q}_p) \to G(\mathbb{Q}_p))| \leq \iota(N).$$

Proof. By the long exact sequence in Galois cohomology, it is enough to show that there exists a constant $\iota(N)$ depending only on N such that $|\mathrm{H}^1(\mathbb{Q}_p, \mu(\overline{\mathbb{Q}}_p))| \leq \iota(N)$. For this, let $E/\mathbb{Q}_p$ denote the finite Galois extension corresponding to the kernel of the action of $\pi_1(\mathbb{Q}_p)$ on $\mu(\overline{\mathbb{Q}}_p)$. We have $[E:\mathbb{Q}_p] \leq N!$. As p is prime to $N!$, the number of extensions of degree $\leq N!N$ of $\mathbb{Q}_p$ is

$$c(N) := \sum_{n=1}^{N!N} \sum_{d|n} \frac{n}{d}.$$

Let E_N/E denote the finite Galois extension corresponding to the open subgroup $\bigcap_{E'} \pi_1(E') \subset \pi_1(E)$, where E'/E describes all Galois extensions of degree $\leq N$ in $\overline{\mathbb{Q}}_p/E$. We have $[E_N : \mathbb{Q}_p] \leq N!c(N)N$. By construction, the restriction map $\mathrm{H}^1(\mathbb{Q}_p, \mu(\overline{\mathbb{Q}}_p)) \to \mathrm{H}^1(E_N, \mu(\overline{\mathbb{Q}}_p))$ is trivial. So, by the inflation-restriction exact sequence, we get an isomorphism $\mathrm{H}^1(E_N|\mathbb{Q}_p, \mu(\overline{\mathbb{Q}}_p)) \tilde{\to} \mathrm{H}^1(\mathbb{Q}_p, \mu(\overline{\mathbb{Q}}_p))$ and we can take $\iota(N) := N^{N!c(N)N}$. □

Lemma 8.2.5. *Let G be a connected semisimple group over $\mathbb{Q}_p$ and let $K \subset G(\mathbb{Q}_p)$ be a maximal special compact subgroup. Assume that $p > \iota(|\mu_G|)$. Then, for every $a \in G(\mathbb{Q}_p)$, either $a \in K$ or $\deg_K(a) \geq p$.*

8.2.6. *Proof of Theorem 8.2.1.* We may assume K is of the form (8.1.4). Then for every $a \in G(\mathbb{A}_f)$ we have $\deg_K(a) = \prod_p \deg_{K_p}(a)$. Assume $\deg_K(a) \leq d$. If $p \notin \mathcal{P}$ and $p > \max\{d, \iota(|\mu_G|)\}$, Lemma 8.2.5 shows that $a \in K_p$. But then the conclusion follows from Lemma 8.2.2 applied to the finitely many p which are in $\mathcal{P}$ or $\leq \max\{d, \iota(|\mu_G|)\}$. □

Remark 8.2.7. Assume that K is of the form (8.1.4) and let $n_d(G(\mathbb{A}_f), K)$ denote the number of double classes KaK for $a \in G(\mathbb{A}_f)$ with $\deg_K(a) \leq d$. Then

$$n_d(G(\mathbb{A}_f), K) \leq \prod_{p \in \mathcal{P} \text{ or } p \leq d} n_d(G(\mathbb{Q}_p), K_p).$$

8.3. ***Degree of local Hecke operators.*** This section is devoted to the proof of Lemmas 8.2.2 and 8.2.5. For an anisotropic group G over $\mathbb{Q}_p$, the group $G(\mathbb{Q}_p)$ is compact [Prasad 1982]. So it is enough to prove Lemmas 8.2.2 and 8.2.5 for isotropic groups G. Then we can use (avatars of) the Bruhat–Tits decomposition attached to the euclidean building of $G(\mathbb{Q}_p)$, which expresses explicitly the degree of local Hecke operators in terms of the extended affine Weyl group of $G(\mathbb{Q}_p)$. For Lemma 8.2.2, we may assume that G is simply connected (Lemma 8.1.3). Under this assumption, the parametrizing group of the Bruhat–Tits decomposition is a Coxeter group (the affine Weyl group) and computations are easy. For Lemma 8.2.5, we can no longer resort to Lemma 8.1.3 and thus have to handle the Bruhat–Tits decomposition for possibly non-simply connected G. There, the parametrizing group of the Bruhat–Tits decomposition (the extended affine Weyl group) is a semidirect product of a Coxeter group by a finite group of nonpreserving type automorphisms, which make computations slightly more technical. As this is possibly less known to nonexperts, we include an expository section, which we try to keep as self-contained as possible.

8.3.1. *Review of Bruhat–Tits theory.* Let G be a connected semisimple isotropic group over $\mathbb{Q}_p$ with $\mathbb{Q}_p$-split maximal torus S. The principle of Bruhat–Tits theory is to attach to $G(\mathbb{Q}_p)$ a discrete euclidean building endowed with a strongly transitive action of $G(\mathbb{Q}_p)$. The existence of this building is essentially equivalent to the datum of a generalized Tits system $(G(\mathbb{Q}_p), B, N)$, and once the existence of $(G(\mathbb{Q}_p), B, N)$ is established, the axioms of Tits systems gives a combinatorial description of the compact subgroups of $G(\mathbb{Q}_p)$ containing B in terms of the extended affine Weyl group $N/N \cap B$.

We assume the reader is familiar with the formalism of Tits systems and buildings [Bourbaki 1968, Chapitre IV; Bruhat and Tits 1972, §1, §2; Brown 1989; Garrett 1997]. We review below the construction of the euclidean Bruhat–Tits building attached to $G(\mathbb{Q}_p)$ and summarize in 8.3.1.6 the consequences of this construction which we will need. We follow closely the survey [Rousseau 2009], which provides a synthetic introduction to the classical [Tits 1979; Bruhat and Tits 1972; 1984a; 1984b].

Given a group H acting on a set X and an element $x \in X$, we write H_x for the stabilizer of x in H.

Let $X^*(S)$ and $X_*(S)$ denote the groups of characters and cocharacters of S, respectively. These are free $\mathbb{Z}$-modules of rank $r = \dim(S)$, dual to each other via the evaluation pairing $X^*(S) \times X_*(S) \to \mathbb{Z}$. Set $V(S) := X_*(S) \otimes_{\mathbb{Z}} \mathbb{R}$ and $N := \mathrm{Nor}_G(S)(\mathbb{Q}_p)$, $Z := \mathrm{Cen}_G(S)(\mathbb{Q}_p)$. The action of N on S by conjugation yields

$$\nu^v : N \twoheadrightarrow N/Z \hookrightarrow \mathrm{Aut}_{\mathrm{GrAlg}\,/\mathbb{Q}_p}(S) \hookrightarrow \mathrm{Aut}_{\mathbb{Z}}(X_*(S)) \hookrightarrow \mathrm{GL}(V(S)).$$

8.3.1.1. *The vectorial part W^v of the extended affine Weyl group.* The torus S acts on the Lie algebra $\mathfrak{g}$ of G, which decomposes accordingly as $\mathfrak{g} = \bigoplus_{\alpha\in X^*(S)} g_\alpha$, where $\mathfrak{g}_\alpha \subset \mathfrak{g}$ denotes the eigenspace corresponding to the character $\alpha \in X^*(S)$. Let Φ denote the set of all $0 \neq \alpha \in X^*(S)$ such that $\mathfrak{g}_\alpha \neq 0$. Then $\Phi \subset V(S)^*$ is a (not necessarily reduced) root system in the usual sense. Let W^v denote the Weyl group of Φ; we endow $V(S)$ with a W^v-invariant scalar product. For every $\alpha \in \Phi$, let $r_\alpha \in W^v$ denote the orthogonal reflexion fixing $\ker(\alpha)$. The morphism $\nu^v : N \to \mathrm{GL}(V(S))$ induces an isomorphism $\nu^v : N/Z \tilde{\to} W^v$. We can describe more precisely elements corresponding to the reflexions r_α, $\alpha \in \Phi$. For every $\alpha \in \Phi$ there exists a unique connected unipotent group $U_\alpha \hookrightarrow G$ normalized by $\mathrm{Cen}_G(S)$ and such that the induced embedding of Lie algebras $\mathfrak{u}_\alpha \hookrightarrow \mathfrak{g}$ identifies $\mathfrak{u}_\alpha$ with $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{2\alpha}$. Also, for every $1 \neq u \in U_\alpha(\mathbb{Q}_p)$, the intersection $U_{-\alpha}(\mathbb{Q}_p)uU_{-\alpha}(\mathbb{Q}_p) \cap N$ consists of a single element $m(u)$, and $m(u)$ has the property that $\nu^v(m(u)) = r_\alpha \in N/Z \simeq W^v$ [Borel and Tits 1965, §5].

8.3.1.2. *The extended affine Weyl group* $\widehat{W}$. As the restriction map

$$X^*(\mathrm{Cen}_G(S)) \to X^*(S)$$

has finite cokernel of order, say, $m \geq 1$, for every $z \in Z$ there exists a unique $\nu(z) \in V(S)$ such that $\chi(\nu(z)) = -\frac{1}{m}v_p(m\chi(z))$, $\chi \in X^*(S)$. This defines a morphism $\nu : Z \to V(S)$ characterized by the fact that $\chi(\nu(s)) = -v_p(\chi(s))$ for $\chi \in X^*(S)$, $s \in S(\mathbb{Q}_p)$. Set $\widehat{T} := \nu(Z) \subset V(S)$ and $Z_0 := \ker(\nu) \subset Z$, which is a compact open subgroup containing $S(\mathbb{Z}_p) \simeq (\mathbb{Z}_p^\times)^r$. The morphism $\nu : Z \to V(S)$ extends [Rousseau 2009, Proposition 11.3] to a morphism $\nu : N \to \mathrm{GA}(V(S))$ with values in the group $\mathrm{GA}(V(S))$ of affine transformations of $V(S)$ and which makes the following diagram commute:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & Z & \longrightarrow & N & \longrightarrow & N/Z & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \nu} & & \downarrow{\scriptstyle \nu} & & \downarrow{\scriptstyle \nu^v} & & \\ 1 & \longrightarrow & V(S) & \longrightarrow & \mathrm{GA}(V(S)) & \longrightarrow & \mathrm{GL}(V(S)) & \longrightarrow & 1 \end{array}$$

where the left and right vertical arrows are the morphisms defined above. The extension $\nu : N \to \mathrm{GA}(V(S))$ is unique up to $\mathrm{GA}(V(S))$-conjugacy. Write $\widehat{W} := \nu(N)$.

8.3.1.3. *The Weyl group W and the standard apartment $A(S)$.* For every $\alpha \in \Phi$ and $1 \neq u \in U_\alpha(\mathbb{Q}_p)$, $\nu(m(u)) \in \widehat{W}$ is an orthogonal reflexion with hyperplane $H(u)$ of direction $\ker(\alpha)$, and

$$W := \langle \nu(m(u)) \mid 1 \neq u \in U_\alpha(\mathbb{Q}_p),\ \alpha \in \Phi \rangle \lhd \widehat{W}$$

is a discrete affine reflexion group. Let $A(S) := (V(S), W)$ denote the corresponding apartment (see [Bruhat and Tits 1972, §1.3; Rousseau 2009, Part I]) and $\mathcal{F}(S)$ the set of its facets. For a special point $x \in A(S)$, we have $W \simeq T \rtimes W_x$ with $T \subset \widehat{T}$ and $W_x \simeq \widehat{W}_x \simeq W^v$. Fix from now on a special point $0 \in A(S)$.

8.3.1.4. *Parahoric and parabolic subgroups.* For every $\alpha \in \Phi$ and $1 \neq u \in U_\alpha(\mathbb{Q}_p)$, there exists a unique $\phi_\alpha(u) \in \mathbb{R}$ such that $H(u) = \alpha^{-1}(-\phi_\alpha(u))$. Set $\phi_\alpha(1) = +\infty$. This defines a map $\phi_\alpha : U_\alpha(\mathbb{Q}_p) \to]-\infty, +\infty]$ with the property that $U_{\alpha,\lambda} := \phi_\alpha^{-1}([\lambda, +\infty]) \subset U_\alpha(\mathbb{Q}_p)$ is a subgroup for every $\lambda \in \mathbb{R}$. For every $x \in A(S)$, let $P_x \subset G(\mathbb{Q}_p)$ denote the subgroup generated by Z_0 and the $U_{\alpha,-\alpha(x)}$, $\alpha \in \Phi$. When $F = C$ is a chamber, $P_0(C)$ is called an *Iwahori subgroup*. For every facet $F \subset A(S)$, define the *parahoric subgroup of type F* to be the subgroup $P_0(F) \subset G(\mathbb{Q}_p)$ generated by Z_0 and the $U_{\alpha,\lambda}$ for $\alpha \in \Phi$, $\lambda \in \mathbb{R}$ such that $F \subset \phi_\alpha^{-1}([-\lambda, +\infty])$. Also let $N_F \subset N$ denote the pointwise stabilizer of F in N. Then the group $P(F) := N_F P_0(F) \subset G(\mathbb{Q}_p)$ is called the *parabolic subgroup of type F*.

8.3.1.5. *The building $\mathcal{I}^a(G, \mathbb{Q}_p)$.* Let $\mathcal{I}^a(G, \mathbb{Q}_p)$ be the quotient of $G(\mathbb{Q}_p) \times A(S)$ by the equivalence relation $(g, x) \sim (g', x')$ if and only if there exists $n \in N$ such that $x' = \nu(n)x$ and $g^{-1}g'n \in P_x N_x$. The action of $G(\mathbb{Q}_p)$ by left multiplication on the first factor of $G(\mathbb{Q}_p) \times A(S)$ induces an action on $\mathcal{I}^a(G, \mathbb{Q}_p)$. One shows that $\mathcal{I}^a(G, \mathbb{Q}_p)$ is a euclidean building — the *Bruhat–Tits building of G over* $\mathbb{Q}_p$ — with set of apartments $gA(S)$, $g \in G(\mathbb{Q}_p)$ and set of facets $g\mathcal{F}(S)$, $g \in G(\mathbb{Q}_p)$. The pointwise stabilizer of a facet gF is $gP(F)g^{-1}$, the stabilizer of $gA(S)$ is gNg^{-1} and the pointwise stabilizer of $gA(S)$ is gZ_0g^{-1}.

For the following, we refer to [Tits 1979, §3.4, 3.5]. Let F be a facet in $\mathcal{I}^a(G, \mathbb{Q}_p)$ and let $G(\mathbb{Q}_p)_F \subset G(\mathbb{Q}_p)$ denote the stabilizer of F in $G(\mathbb{Q}_p)$. Then there exists a unique smooth affine group scheme $\mathcal{G}_F$ over $\mathbb{Z}_p$ with generic fiber G and with the property that $\mathcal{G}_F(\mathcal{O}_k) = G(k)_F$ for every finite unramified extension $k/\mathbb{Q}_p$ (here we implicitly identify $\mathcal{I}^a(G, \mathbb{Q}_p)$ with its image in $\mathcal{I}^a(G, k)$). Write $\mathcal{G}^{\mathrm{red}}_{F,\mathbb{F}_p} := \mathcal{G}^{\circ}_{F,\mathbb{F}_p}/R_u(\mathcal{G}_{F,\mathbb{F}_p})$ for the connected reductive part of the reduction modulo p of $\mathcal{G}_F$; as G is residually quasisplit, $\mathcal{G}^{\mathrm{red}}_{F,\mathbb{F}_p}$ is quasisplit. The link of F in $\mathcal{I}^a(G, \mathbb{Q}_p)$ is the spherical building of $\mathcal{G}^{\mathrm{red}}_{F,\mathbb{F}_p}$. In particular, when F is of codimension 1, $\mathcal{G}^{\mathrm{red}}_{F,\mathbb{F}_p}$ has semisimple $\mathbb{F}_p$-rank 1 and its spherical building is 0-dimensional with vertices corresponding to its minimal parabolic subgroups. More precisely, let R denote the canonical set of generators of W (the reflexions with respect to the walls of a chamber). If F is of type $R \setminus \{r\}$, let d_r denote the dimension of $\mathcal{G}^{\mathrm{red}}_{F,\mathbb{F}_p}/P$, where P is a minimal parabolic subgroup. Then the number of vertices in the link of F is $p^{d_r} + 1$. The classification shows that $d_r = 1, 2, 3$ but we only need the fact that $d_r \geq 1$. This number can also be interpreted as the number of chambers in $\mathcal{I}^a(G, \mathbb{Q}_p)$ containing F; that is, $\deg_B(r) + 1$.[3]

8.3.1.6. *Non-type-preserving automorphisms.* The action of $G(\mathbb{Q}_p)$ on $\mathcal{I}^a(G, \mathbb{Q}_p)$ is strongly transitive[3] but not type-preserving in general. More precisely, let $G_\circ \subset G(\mathbb{Q}_p)$ denote the subgroup acting on $\mathcal{I}^a(G, \mathbb{Q}_p)$ by type-preserving automorphisms; this is the subgroup generated by $N_\circ := \nu^{-1}(W)$ and the $U_\alpha(\mathbb{Q}_p)$, $\alpha \in \Phi$. The action of $G_\circ$ on $\mathcal{I}^a(G, \mathbb{Q}_p)$ remains strongly transitive [Garrett 1997, §17.7]. Set $B := P_0(C)$ for a chamber C in $A(S)$; note that $B \subset G_\circ$. Then $(G_\circ, B, N_\circ)$ is the Tits system induced by the strongly transitive, type-preserving action of $G_\circ$ on $\mathcal{I}^a(G, \mathbb{Q}_p)$ [Garrett 1997, §5.2]. In particular we get the standard Bruhat–Tits decompositions [Garrett 1997, §5.1–5.4]. The formalism of Tits systems (or BN-pairs) extends (formally) to the "generalized" Tits system $(G(\mathbb{Q}_p), B, N)$. This is explained in [Bourbaki 1968, Chapitre IV, §2, Exercice 8; Garrett 1997, §5.5, §14.7] (see also [Borel 1976, §3.1; Tits 1979, §2.5]), which we briefly summarize.

[3] This follows from the fact that the action of $G(\mathbb{Q}_p)$ on $\mathcal{I}^a(G, \mathbb{Q}_p)$ is strongly transitive (that is, transitive on the set of embeddings of a chamber into an apartment): if C and rC are two chambers in a given apartment A with wall F and if C' is another chamber in $\mathcal{I}^a(G, \mathbb{Q}_p)$ containing F then there exists an apartment A' containing C, C' and $g \in G(\mathbb{Q}_p)$ such that $gA = A'$, $gC = C$. So $g \in B$. Also, the chambers $C = gC$ and grC have the same wall $gF = F$ in A', which forces $grC = C'$. This shows that the chambers containing F in $\mathcal{I}^a(G, \mathbb{Q}_p)$ are C and the chambers brC, $b \in B$.

(1) (structure of $\widehat{W}$) Set $Z_\circ := Z \cap G_\circ$. Since $N_\circ = \nu^{-1}(W)$, we have $W \simeq N_\circ / B \cap N_\circ$ and the following commutative diagram of short exact sequences:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & T_\circ := Z_\circ / B \cap N_\circ & \longrightarrow & W = N_\circ / B \cap N_\circ & \longrightarrow & W^v & \longrightarrow & 1 \\ & & \downarrow & \square & \downarrow & & \downarrow \simeq & & \\ 1 & \longrightarrow & \widehat{T} = Z / B \cap N & \longrightarrow & \widehat{W} = N / B \cap N & \longrightarrow & W^v & \longrightarrow & 1 \end{array}$$

Let $\Psi := Z / Z_\circ$ denote the cokernel of the right (and middle) vertical arrows. As the extension $1 \to T_\circ \to \widehat{T} \to \Psi \to 1$ splits, $\widehat{W} = W \rtimes \Psi$ and an explicit complement of W in $\widehat{W}$ is

$$\Psi = \frac{P(C) \cap N}{B \cap N} \simeq \frac{N_C(B \cap N)}{B \cap N}.$$

The order of Ψ is bounded from above by a constant which only depends on the Coxeter diagram $\Delta(W, R)$ of (W, R). Indeed, Ψ injects into $\mathrm{Aut}(\Delta(W, S))$ and stabilizes the connected components of $\Delta(W, R)$. In our case, we can also show that Ψ is abelian. This follows for instance from the explicit description given in [Tits 1979, §2.5]. Namely, if $S^{\mathrm{sc}} \subset G^{\mathrm{sc}}$ is a $\mathbb{Q}_p$-split maximal torus mapping into S, $Z^{\mathrm{sc}} := \mathrm{Cen}_{G^{\mathrm{sc}}}(S^{\mathrm{sc}})(\mathbb{Q}_p)$ and $Z' := \phi(\mathbb{Q}_p)(Z^{\mathrm{sc}})$, then $\Psi \simeq Z / Z' Z_0$. In particular, a rough upper bound for $|\Psi|$ is $|\Psi| \leq \iota(|\mu_G|)$.

(2) (double cosets) The map $\hat{w} \to B \hat{w} B$ induces a bijection

$$\widehat{W} \xrightarrow{\sim} B \setminus G(\mathbb{Q}_p) / B.$$

(3) (subgroups of $G(\mathbb{Q}_p)$ containing B) Recall that R denotes the canonical set of generators of W, and let $\mathcal{R}$ denote the set of pairs (Ψ', R') with $\Psi' \subset \Psi$ a subgroup and $R' \subset R$ a subset normalized by Ψ'. For a subset $R' \subset R$, let $W_{R'} \subset W$ denote the subgroup generated by R'; if $R' \subsetneq R$ the group $W_{R'}$ is finite. Then the map

$$(\Psi', R') \to P_{(\Psi', R')} := B W_{R'} \Psi' B := \bigsqcup_{\hat{w} \in W_{R'} \Psi'} B \hat{w} B$$

induces a bijection from $\mathcal{R}$ to the set of subgroups of $G(\mathbb{Q}_p)$ containing B. Furthermore, a subgroup $P_{(\Psi', R')} \subsetneq G(\mathbb{Q}_p)$ is a maximal compact subgroup if and only if $R' \subsetneq R$, $\Psi' = \mathrm{Nor}_\Psi(R')$ and Ψ' acts transitively on $R \setminus R'$.

8.3.2. *Proof of Lemma 8.2.2.* The exact sequence

$$1 \to \mu_G(\mathbb{Q}_p) \to G^{\mathrm{sc}}(\mathbb{Q}_p) \to G(\mathbb{Q}_p) \to \mathrm{H}^1(\mathbb{Q}_p, \mu_G(\overline{\mathbb{Q}_p}))$$

shows that $p^{\mathrm{sc}}(\mathbb{Q}_p) : G^{\mathrm{sc}}(\mathbb{Q}_p) \to G(\mathbb{Q}_p)$ has finite kernel and finite cokernel. So, from Lemma 8.1.3, we may assume that G is simply connected and that $K = B \subset G(\mathbb{Q}_p)$ is an Iwahori subgroup. Then, from 8.3.1.6(3), we have the

Bruhat–Tits decomposition

$$G(\mathbb{Q}_p) = \bigsqcup_{w \in W} BwB.$$

As (W, R) is a Coxeter system, every $w \in W$ can be written as $w = r_1 \cdots r_{\ell(w)}$ with $r_1, \ldots, r_{\ell(w)} \in R$ and $\ell(w) \in \mathbb{Z}_{\geq 0}$ minimal (called the length of w); the elements $r_1, \ldots, r_\ell(w)$ are then unique up to permutation. Furthermore, we have

$$\deg_B(w) = \deg_B(r_1) \cdots \deg_B(r_{\ell(w)})$$

with $\deg_B(r) = p^{d_r}$ and $d_r = 1, 2, 3$ for every $r \in R$ (see 8.3.1.5). In particular, $\deg_B(w) \geq p^{\ell(w)}$. The conclusion then follows from the fact that there are only finitely many elements of bounded length in a Coxeter group. □

8.3.3. *Proof of Lemma 8.2.5.* From 8.3.1.6(3), $K \subset G(\mathbb{Q}_p)$ is of the form

$$K = \bigsqcup_{\hat{w} \in W_{R'}\Psi'} B\hat{w}B$$

for a subset $R' \subsetneq R$ and a subgroup $\Psi' \subset \Psi$ such that $\Psi' = \mathrm{Nor}_\Psi(R')$ and Ψ' acts transitively on $R \setminus R'$. Write $\widehat{W}_K := W_{R'}\Psi'$. Then, for every $\hat{w} = w\psi \in \widehat{W}$ we can compute, using $BwBw'B = Bww'B$,

$$\begin{aligned} K\hat{w}K &= \Psi' BW_{R'}Bw\psi BW_{R'}B\Psi' = \Psi' BW_{R'}BwB\psi W_{R'}\psi^{-1}B\psi\Psi' \\ &= \Psi' BW_{R'}w\psi W_{R'}\psi^{-1}B\psi\Psi' = B\widehat{W}_K\hat{w}\widehat{W}_K B. \end{aligned}$$

That is,

$$K\hat{w}K = \bigsqcup_{\lambda \in \widehat{W}_K\hat{w}\widehat{W}_K} B\lambda B.$$

As a result,

$$|B \setminus K\hat{w}K| = \sum_{\lambda \in \widehat{W}_K\hat{w}\widehat{W}_K} \deg_B(\lambda).$$

On the other hand,

$$|B \setminus K\hat{w}K| = \deg_K(\hat{w})[K : B] = \deg_K(\hat{w}) \sum_{\mu \in \widehat{W}_K} \deg_B(\mu).$$

From this, we get

$$\deg_K(\hat{w}) = \frac{\sum_{\lambda \in \widehat{W}_K\hat{w}\widehat{W}_K} \deg_B(\lambda)}{\sum_{\mu \in \widehat{W}_K} \deg_B(\mu)}.$$

Now, we can compute explicitly

$$\sum_{\mu \in \widehat{W}_K} \deg_B(\mu) = \sum_{\psi \in \Psi} \sum_{w \in W_{R'}} \deg_B(w) = |\Psi| \Bigl(1 + \sum_{1 \neq w \in W_{R'}} \deg_B(w)\Bigr) \equiv |\Psi|\,[p].$$

Write

$$\widehat{W}_K \setminus \widehat{W}_K \hat{w} \widehat{W}_K = \mathcal{S}_1 \sqcup \mathcal{S}_2,$$

where $\mathcal{S}_1$ denotes the set of left cosets $\widehat{W}_K\lambda$ such that $\widehat{W}_K\lambda \cap \Psi \neq \varnothing$. Then, for $C \in \mathcal{S}_1$ we have

$$\sum_{\lambda \in C} \deg_B(\lambda) = \sum_{\lambda \in \widehat{W}_K} \deg_B(\lambda),$$

and hence

$$\deg_K(\hat{w}) = |\mathcal{S}_1| + \frac{\sum_{C \in \mathcal{C}_2} \sum_{\lambda \in C} \deg_B(\lambda)}{\sum_{\mu \in \widehat{W}_K} \deg_B(\mu)}.$$

But for $C \in \mathcal{S}_2$ and $\lambda \in C$ we have $\lambda = w\psi$ with $1 \neq w \in W$ and $\psi \in \Psi$. In particular, $\deg_B(\lambda) = \deg_B(w)$ is a power of p. As $\deg_K(\hat{w})$ is an integer and $\sum_{\mu \in \widehat{W}_K} \deg_B(\mu) \equiv |\Psi|[p]$ is nonzero (this is where we use the assumption $p > \iota(|\mu_G|)(\geq |\Psi|)$), it is thus enough to prove that if $\hat{w} \notin \widehat{W}_K$ then $\mathcal{S}_2 \neq \varnothing$. This follows from the maximality of K. Indeed, otherwise, we may assume that $\hat{w} = \psi \in \Psi \setminus \Psi'$. Then the following holds:

$$\text{for every } w \in \widehat{W}_K \text{ there exists } \psi(w, \psi) \in \Psi \setminus \Psi',\ w(w, \psi) \in \widehat{W}_K \text{ such that } \psi w = w(w, \psi)\psi(w, \psi). \qquad (*)$$

But then,

$$\widehat{W}_K \psi(w, \psi) \widehat{W}_K = \widehat{W}_K \psi \widehat{W}_K.$$

Thus $\psi(w, \psi)$ satisfies again $(*)$. This shows that the subgroup generated by $\widehat{W}_K$ and ψ in $\widehat{W}$ is of the form $\widehat{W}_K \rtimes \Psi''$ for some subgroup $\Psi' \subsetneq \Psi'' \subset \Psi$, which contradicts the maximality of K. □

8.4. *Proof of 7.2.2.* Recall that we may assume that $\Gamma = G(\mathbb{Q}) \cap K$ for a compact open subgroup K as in (8.1.4). So we will take $K = \prod_p K_p$ with $K_p \subset G(\mathbb{Q}_p)$ compact open (which we could even assume to be maximal) for every p and hyperspecial for $p \notin \mathcal{P}$, where $\mathcal{P}$ denotes the finite set of primes where G ramifies. Also, recall that $\Gamma^- \subset G(\mathbb{A}_f)$ denotes the closure of Γ for the adelic topology.

Lemma 8.4.1. *For every $a \in G(\mathbb{Q})$ the canonical map $\varphi_a : \Gamma \setminus \Gamma a \Gamma \to \Gamma^- \setminus \Gamma^- a \Gamma^-$ is bijective. In particular* $\deg_\Gamma(a) = \deg_{\Gamma^-}(a)$.

Proof. We first prove that φ_a is injective. Let $\Gamma a\gamma, \Gamma a\gamma' \in \Gamma \setminus \Gamma a \Gamma$ be such that $\Gamma^- a\gamma = \Gamma^- a\gamma'$; that is, there exists $\gamma^- \in \Gamma^-$ such that $a\gamma = \gamma^- a\gamma'$. But then $\gamma^- = a\gamma\gamma'^{-1}a^{-1} \in \Gamma^- \cap G(\mathbb{Q}) = \Gamma$. Thus $\Gamma a\gamma = \Gamma a\gamma'$. We now show that φ_a is surjective. As $\varphi_a : \Gamma \setminus \Gamma a \Gamma \hookrightarrow \Gamma^- \setminus \Gamma^- a \Gamma^-$ is injective and both sets are finite, it is enough to prove that $\deg_{\Gamma^-}(a) \leq \deg_\Gamma(a)$. For this, fix a set of representatives R

of $\Gamma \setminus \Gamma a \Gamma$ and observe that

$$\Gamma^{-} a \Gamma^{-} = (\Gamma a \Gamma)^{-} = \left(\bigsqcup_{b \in R} \Gamma b \right)^{-} = \bigcup_{b \in R} (\Gamma b)^{-} = \bigcup_{b \in R} \Gamma^{-} b.$$

Here, the first equality follows from the fact that Γ^{-} is compact (to prove that $(\Gamma a \Gamma)^{-} \subset \Gamma^{-} a \Gamma^{-}$). □

Lemma 8.4.2. *The canonical map $\varphi : \Gamma \setminus G(\mathbb{Q})/\Gamma \to K \setminus G(\mathbb{A}_f)/K$ has finite fibers. More precisely, for every $a \in G(\mathbb{Q})$, $|\varphi^{-1}(\varphi(a))| \le \deg_K(a)$.*

Proof. Let $a \in G(\mathbb{Q})$ and let $R \subset G(\mathbb{Q})$ be a set of representatives for $\varphi^{-1}(KaK)$. Since $\Gamma = G(\mathbb{Q}) \cap K$, the map $\Gamma \setminus G(\mathbb{Q}) \to K \setminus G(\mathbb{A}_f)$ is injective, hence restricts to an injective map

$$\Gamma \setminus \bigsqcup_{b \in R} \Gamma b \Gamma = \bigsqcup_{b \in R} \Gamma \setminus \Gamma b \Gamma \to K \setminus KaK.$$

Because the union is disjoint, we get

$$|R| \le \left| \bigsqcup_{b \in R} \Gamma \setminus \Gamma b \Gamma \right| \le |K \setminus KaK| = \deg_K(a). \qquad \square$$

For simply connected groups G of noncompact type, the proof is now complete: let $n_d(G(\mathbb{Q}), \Gamma)$ denote the number of double classes $\Gamma a \Gamma$ with $a \in G(\mathbb{Q})$ and $\deg_\Gamma(a) \le d$. By strong approximation, $\Gamma^{-} = K$. So $\deg_\Gamma(a) = \deg_K(a)$ (Lemma 8.4.1) and

$$n_d(G(\mathbb{Q}), \Gamma) \le d n_d(G(\mathbb{Q}), K) \le d n_d(G(\mathbb{A}_f), K)$$

(Lemma 8.4.2). So the conclusion follows from Theorem 8.2.1.

In the non-simply connected case, we can no longer apply Lemma 8.4.1 directly.

Lemma 8.4.3. *There exists an integer $r(G, K) \ge 1$ and a compact normal subgroup $H \subset K$ such that $H \equiv \Gamma^{-}$ and for every p, $H_p \subset G(\mathbb{Q}_p)$ is compact open and $[K_p : H_p] \le r(G, K)$.*

Proof. Set $K_p^{\mathrm{sc}} = (p^{\mathrm{sc}})^{-1}(K_p) \subset G^{\mathrm{sc}}(\mathbb{Q}_p)$ and $K^{\mathrm{sc}} := \prod_p K_p^{\mathrm{sc}} \subset G(\mathbb{A}_f)$. Then $K_p^{\mathrm{sc}} \subset G^{\mathrm{sc}}(\mathbb{Q}_p)$ is compact open for every p and hyperspecial for $p \notin \mathcal{P}$. Set $H_p := \mathrm{Cor}_{K_p}(p^{\mathrm{sc}}(K_p^{\mathrm{sc}}))$ for $p \in \mathcal{P}$ and $H_p := p^{\mathrm{sc}}(K_p^{\mathrm{sc}})$ otherwise. We claim that $H = \prod_p H_p \subset K$ works. Indeed, for $p \notin \mathcal{P}$ the short exact sequence

$$1 \longrightarrow \mu_G \longrightarrow G^{\mathrm{sc}} \xrightarrow{p^{\mathrm{sc}}} G \longrightarrow 1$$

extends to a short exact sequence of smooth group schemes over $\mathbb{Z}_p$ with $K_p^{\mathrm{sc}} = G^{\mathrm{sc}}(\mathbb{Z}_p)$, $K_p = G(\mathbb{Z}_p)$. Taking the reduction modulo p and $\overline{\mathbb{F}}_p$-points, we obtain a short exact sequence of finite groups

$$1 \longrightarrow \mu_G(\overline{\mathbb{F}}_p) \longrightarrow G^{\mathrm{sc}}(\overline{\mathbb{F}}_p) \xrightarrow{p^{\mathrm{sc}}} G(\overline{\mathbb{F}}_p) \longrightarrow 1.$$

Since $\mathbb{Z}_p^{\mathrm{ur}}$ (the integral closure of $\mathbb{Z}_p$ in the maximal unramified extension $\mathbb{Q}_p^{\mathrm{ur}}$ of $\mathbb{Q}_p$) is Henselian, by smoothness (see [Platonov and Rapinchuk 1994, Lemma 6.5, p. 295] for details) we get a short exact sequence

$$1 \longrightarrow \mu_G(\mathbb{Z}_p^{\mathrm{ur}}) \longrightarrow G^{\mathrm{sc}}(\mathbb{Z}_p^{\mathrm{ur}}) \xrightarrow{p^{\mathrm{sc}}} G(\mathbb{Z}_p^{\mathrm{ur}}) \longrightarrow 1.$$

Taking the $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p)$-invariants, we obtain

$$H_p = \ker\bigl(K_p \to \mathrm{H}^1(\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p), \mu_G(\mathbb{Z}_p^{\mathrm{ur}}))\bigr).$$

In particular, H_p is normal in K_p (thus H is normal in K as required) and

$$[K_p : H_p] \le \left|\mathrm{H}^1(\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p), \mu_G(\mathbb{Z}_p^{\mathrm{ur}}))\right| \le |\mu_G|.$$

The last inequality comes from the fact that $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p) \simeq \hat{\mathbb{Z}}$ is procyclic, so that by [Serre 1968, Chapitre XIII, §1, Proposition 1],

$$\mathrm{H}^1(\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p), \mu_G(\mathbb{Z}_p^{\mathrm{ur}})) = \mu_G(\mathbb{Z}_p^{\mathrm{ur}})_{\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p)}$$

(the maximal trivial $\mathrm{Gal}(\mathbb{Q}_p^{\mathrm{ur}}/\mathbb{Q}_p)$-quotient of $\mu_G(\mathbb{Z}_p^{\mathrm{ur}})$). So we can take

$$r(G, K) := \max\{[K_p : H_p] \mid p \in \mathcal{P} \text{ or } p < p_K\} \cup \{|\mu_G|\}.$$

It remains to show that $\Gamma^-, H \subset G(\mathbb{A}_f)$ are commensurable. As $\Gamma^{\mathrm{sc}} \subset G^{\mathrm{sc}}(\mathbb{Q})$ is arithmetic and $p^{\mathrm{sc}} : G^{\mathrm{sc}} \to G$ is surjective, the group $p^{\mathrm{sc}}(\Gamma^{\mathrm{sc}}) \subset G(\mathbb{Q})$ is again arithmetic [Platonov and Rapinchuk 1994, Theorem 4.1], and hence $[\Gamma : p^{\mathrm{sc}}(\Gamma^{\mathrm{sc}})]$ is finite. Thus $[\Gamma^- : p^{\mathrm{sc}}((\Gamma^{\mathrm{sc}})^-)](\le [\Gamma : p^{\mathrm{sc}}(\Gamma^{\mathrm{sc}})])$ is finite as well. But, by strong approximation, $(\Gamma^{\mathrm{sc}})^- = K^{\mathrm{sc}}$ and by the continuity of p^{sc} and the compactness of $(\Gamma^{\mathrm{sc}})^-$,

$$p^{\mathrm{sc}}(\Gamma^{\mathrm{sc}})^- = p^{\mathrm{sc}}((\Gamma^{\mathrm{sc}})^-) = p^{\mathrm{sc}}(K^{\mathrm{sc}}) = H. \qquad \square$$

Lemma 8.4.4. *For every integer $d \ge 1$ there are only finitely many double classes $KaK \in K \backslash G(\mathbb{Q})/K$ with $\deg_{\Gamma^-}(a) \le d$.*

Proof. Let $a \in G(\mathbb{Q})$ with $\deg_{\Gamma^-}(a) \le d$ and let H be as in Lemma 8.4.3. Then $dC_{H,\Gamma^-} \ge \deg_H(a) = \prod_p \deg_{H_p}(a)$ and $\deg_{H_p}(a) \ge \deg_{K_p}(a)/r(G, K)$ (Lemma 8.1.1). Set

$$\mu(G, K, d) := \max\{\iota(|\mu_G|), r(G, K)dC_{H,\Gamma^-}\}.$$

Then by Lemma 8.2.5, for $p \notin \mathcal{P}$, $p > \mu(G, K, d)$, we have $a \in K_p$. Set $\nu(G, K, d) := \iota(|\mu_G|)dr(G, K)C_{H,\Gamma^-}$ and

$$N(G, K, d) := |\mathcal{P}| + |\{p \notin \mathcal{P} : p \le \nu(G, K, d)\}|.$$

Then by Lemma 8.1.1,

$$\deg_K(a) = \prod_{p\in\mathcal{P}} \deg_{K_p}(a) \prod_{p\notin\mathcal{P},\, p\le\nu(G,K,d)} \deg_{K_p}(a)$$
$$\le r(G,K)^{N(G,K,d)} \deg_H(a) \le r(G,K)^{N(G,K,d)} dC_{H,\Gamma^-}.$$

The conclusion thus follows from Theorem 8.2.1. □

8.4.5. *End of the proof of Theorem 7.2.2.* Let $a \in G(\mathbb{Q})$ such that $\deg_\Gamma(a) \le d$. Then $\deg_{\Gamma^-}(a) = \deg_\Gamma(a) \le d$ (Lemma 8.4.1). So there are only finitely many possibilities for the set of double classes $KaK \in K\setminus G(\mathbb{Q})/K$ (Lemma 8.4.4), and hence only finitely many possibilities for the set of double classes $\Gamma a\Gamma \in \Gamma\setminus G(\mathbb{Q})/\Gamma$ (Lemma 8.4.2). □

Remark 8.4.6. Fix from now on $K_0 \subset G(\mathbb{A}_f)$ compact open of the form (8.1.4). Let $K \subset G(\mathbb{A}_f)$ be an arbitrary compact open subgroup and set $\Gamma := K \cap G(\mathbb{Q})$, $\Gamma_0 := K_0 \cap G(\mathbb{Q})$. Then, the proof yields an explicit estimate

$$n_d(G(\mathbb{Q}),\Gamma) \le A[\Gamma_0 : \Gamma\cap\Gamma_0]^2 (dC_{\Gamma_0,\Gamma}) B^{\alpha(dC_{\Gamma_0,\Gamma})\ln(dC_{\Gamma_0,\Gamma})+\beta(dC_{\Gamma_0,\Gamma})+\gamma},$$

where A, B, α, β, γ are absolute constants depending only on the group-theoretical data Γ_0, $\mathcal{P}$, W_G, etc., but not on d nor on Γ.

9. Alternative approaches to Theorem 7.2.2

9.1. ***An ergodic proof of Theorem 7.2.2.*** The argument below was explained to us by Hee Oh. We use the notation of [Eskin and Oh 2006, Theorem 1.2]. If we have infinitely many distinct $\Gamma a_i \Gamma$ with degree bounded by d, then the associated $\Delta(G)$-invariant measures $\tilde{\nu}_{a_i}$ have a weak limit $\tilde{\nu}$. There are two possibilities for $\tilde{\nu}$ as discussed in the proof: either $\tilde{\nu}$ is supported in the closed $\Delta(G)$-invariant measure or is a $G\times G$ invariant measure. In the first case, the proof shows that the sequence $[(e,a_i)]\Delta(G)$ has a constant subsequence; or equivalently, that the sequence $\Gamma a_i\Gamma$ has a constant subsequence, contradicting the assumption that they are distinct. The second case where $\tilde{\nu}$ is $G\times G$-invariant cannot happen, since this is equivalent to $\Gamma\setminus\Gamma a_i\Gamma$ being equidistributed in $\Gamma\setminus G(\mathbb{R})$; but $\Gamma\setminus\Gamma a_i\Gamma$ has at most d points, so cannot possibly be equidistributed.

We do not know whether Theorem 8.2.1 can be recovered from Theorem 7.2.2, and so be proved by ergodic techniques. In any case, our proof relies on different arguments (Bruhat–Tits, strong approximation) and is effective.

It was also mentioned to us by an anonymous referee that, when G is $\mathbb{Q}$-split, [Gan and Oh 2003, Proposition 6.1] gives an effective bound for the degree of Hecke operators; the proof of [Gan and Oh 2003, Proposition 6.1] involves elements of Bruhat–Tits theory in the split case.

9.2. ***Masser–Wüstholz isogeny theorem.*** One key ingredient of the proof of [Orr 2013, Theorem 1.5(ii)] is (a generalization to finitely generated fields of characteristic 0 of) the Masser–Wüstholz isogeny theorem [Masser and Wüstholz 1993; Orr 2013, Theorem 5.2]. Using it, the existence of closed Galois-generic points on Shimura varieties of abelian type (Theorem A) and technical arguments from Orr's thesis, we can give an alternative (and, again, effective) proof of the fact that on a connected Shimura variety of adjoint abelian type, there are only finitely many Hecke operators of bounded degree.

Lemma 9.2.1. *Let (G, X^+) be a connected Shimura datum of abelian type. Then for every arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, $s \in \mathrm{Sh}_\Gamma(G, X^+)$ with residue field $k = k(s)$ and integer $d \geq 1$, there are only finitely many $t \in \widehat{T}_\Gamma(s)$ with $[k(t) : k] \leq d$.*

Proof. It is enough to prove the assertion when G is adjoint and for an ordinary Hecke orbit. By [Orr 2013, Theorem 4.6], this case, in turn, reduces to the case of an ordinary Hecke orbit in the Siegel moduli space $\mathcal{A}_g := \mathrm{Sh}_{\mathrm{GSp}_{2g}(\mathbb{Z})}(\mathrm{GSp}_{2g}, \mathcal{H}_g^+)$. This allows one to use the modular interpretation of $\mathcal{A}_g$ as a coarse moduli space for g-dimensional principally polarized abelian varieties and the fact that Hecke orbits on $\mathcal{A}_g$ correspond to isogeny classes of such objects (here we say that (A, λ_A) is isogenous to (B, λ_B) if there is an isogeny $f : A \to B$ and an integer $N \geq 1$ such that $f^\vee \circ \lambda_B \circ f = N\lambda_A$). Let $a \in \mathcal{A}_g$ with residue field $k = k(a)$ and $b \in T_\Gamma(a)$ with $[k(b) : k] \leq d$. Over $\bar{k}$, a and b correspond to isogenous g-dimensional principally polarized abelian varieties $(A_{\bar{k}}, \lambda_{A\bar{k}})$ and $(B_{\bar{k}}, \lambda_{B\bar{k}})$. Let δ denote the order of $\mathrm{GSp}_{2g}(\mathbb{F}_3)$. Then[4] $(A_{\bar{k}}, \lambda_{A\bar{k}})$ admits a model (A, λ_A) over a finite field extension L of k with $[L : k] \leq \delta$ and $(B_{\bar{k}}, \lambda_{B\bar{k}})$ admits a model (B, λ_B) over a finite field extension $L(b)$ of $k(b)$ with $[L(b) : k(b)] \leq \delta$. From [Orr 2013, Theorem 5.2] (see also [Masser and Wüstholz 1993]), there exist constants $c(A, L)$ and $\kappa(g)$ (independent of B) and an isogeny $f : A_{\bar{k}} \to B_{\bar{k}}$ of degree

$$\deg(f) \leq c(A, L)[L.L(b) : L]^{\kappa(g)} \leq (c(A, L)\delta^{2\kappa(g)})d^{\kappa(g)}.$$

As $f : A_{\bar{k}} \to B_{\bar{k}}$ is uniquely determined by its kernel, there are, up to $\bar{k}$-isomorphism, only finitely many possibilities for $B_{\bar{k}}$, and thus for $(B_{\bar{k}}, \lambda_{B\bar{k}})$ [Milne 1986, Theorem 18.1]. □

Now, let (G, X^+) be a connected Shimura datum of adjoint abelian type, and let $\Gamma \subset G(\mathbb{Q})$ be a neat congruence subgroup. Fix a closed Galois-generic point $s \in \mathrm{Sh}_\Gamma(G, X^+)$ with residue field $k(s) = k$. Up to replacing Γ by a smaller congruence subgroup, we may assume that s is strictly Galois-generic. Then,

---

[4] Explicitly, let $\mathrm{GSp}_{2g}(3)$ denote the kernel of the reduction modulo-3 morphism $\mathrm{GSp}_{2g}(\mathbb{Z}) \to \mathrm{GSp}_{2g}(\mathbb{F}_3)$ and write $\mathcal{A}_{g,3} := \mathrm{Sh}_{\mathrm{GSp}_{2g}(\mathbb{Z})(3)}(\mathrm{GSp}_{2g}, \mathcal{H}_g^+)$. Then L (resp. $L(b)$) can be taken to be the residue field of any point in the fiber over a (resp. b) of $\mathcal{A}_{g,3} \to \mathcal{A}_g$.

- As s is strictly Galois-generic, for every $a \in G(\mathbb{Q})_+$ and $t \in T_{\Gamma,a}(s)$ we have $\deg_\Gamma(a) = [k(t):k]$.
- As Γ is neat, for every $a, b \in G(\mathbb{Q})_+$, the following properties are equivalent:
  (1) $T_{\Gamma,a}(s) \cap T_{\Gamma,b}(s) \neq \varnothing$;
  (2) $T_{\Gamma,a}(s) = T_{\Gamma,b}(s)$;
  (3) $\Gamma a \Gamma = \Gamma b \Gamma$.

Combining these observations with Lemma 9.2.1, we see that the number of Hecke operators of bounded degree $\leq d$ on $\mathrm{Sh}_\Gamma(G, X^+)$ is finite and bounded from above by the number of $t \in T_\Gamma(s)$ with $[k(t):k] \leq d$.

Acknowledgements

This work began while the authors stayed at the Institute for Advanced Study; they are very grateful to the institute for its hospitality. The authors also thank Akio Tamagawa for pointing out a gap in an earlier proof of Theorem A and Ziyang Gao for telling us that Theorem B follows from Orr's result in the case of Shimura data of abelian type. Cadoret thanks the Institute for Advanced Study (Princeton) and the Research Institute for Mathematical Sciences (Kyoto). During her time at the IAS she was supported by the NSF grant No DMS-1155. She was also partly supported by the project ANR-10-JCJC 0107 from the ANR. Kret thanks the Institute for Advanced Study (Princeton), the Mathematical Sciences Research Institute (Berkeley), and the Max Planck Institute for Mathematics (Bonn), where parts of this work were carried out. During his time at the IAS (resp. MSRI) he was supported by the NSF under Grant No. DMS 1128155 (resp. No. 0932078000). Finally, he thanks CMLS at the École Polytechnique for a pleasant stay.

References

[André 1989] Y. André, *G-functions and geometry*, Aspects of Mathematics **E13**, Vieweg, Braunschweig, 1989. MR Zbl

[André 2004] Y. André, *Une introduction aux motifs (motifs purs, motifs mixtes, périodes)*, Panoramas et Synthèses **17**, Société Mathématique de France, Paris, 2004. MR Zbl

[Baily and Borel 1966] W. L. Baily, Jr. and A. Borel, "Compactification of arithmetic quotients of bounded symmetric domains", *Ann. of Math.* (2) **84** (1966), 442–528. MR Zbl

[Borel 1966] A. Borel, "Density and maximality of arithmetic subgroups", *J. Reine Angew. Math.* **224** (1966), 78–89. MR Zbl

[Borel 1976] A. Borel, "Admissible representations of a semi-simple group over a local field with vectors fixed under an Iwahori subgroup", *Invent. Math.* **35** (1976), 233–259. MR Zbl

[Borel and Harish-Chandra 1962] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups", *Ann. of Math.* (2) **75** (1962), 485–535. MR Zbl

[Borel and Tits 1965] A. Borel and J. Tits, "Groupes réductifs", *Inst. Hautes Études Sci. Publ. Math.* 27 (1965), 55–150. MR Zbl

[Bourbaki 1968] N. Bourbaki, *Éléments de mathématique, XXXIV: Groupes et algèbres de Lie, chapitres 4 à 6*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. MR Zbl

[Brown 1989] K. S. Brown, *Buildings*, Springer, New York, 1989. MR Zbl

[Bruhat and Tits 1972] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local", *Inst. Hautes Études Sci. Publ. Math.* 41 (1972), 5–251. MR Zbl

[Bruhat and Tits 1984a] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local, II: Schémas en groupes. Existence d'une donnée radicielle valuée", *Inst. Hautes Études Sci. Publ. Math.* 60 (1984), 197–376. MR Zbl

[Bruhat and Tits 1984b] F. Bruhat and J. Tits, "Schémas en groupes et immeubles des groupes classiques sur un corps local", *Bull. Soc. Math. France* **112**:2 (1984), 259–301. MR Zbl

[Burger and Sarnak 1991] M. Burger and P. Sarnak, "Ramanujan duals, II", *Invent. Math.* **106**:1 (1991), 1–11. MR Zbl

[Cadoret 2015] A. Cadoret, "An open adelic image theorem for abelian schemes", *Int. Math. Res. Not.* **2015**:20 (2015), 10208–10242. MR Zbl

[Cadoret and Moonen 2015] A. Cadoret and B. Moonen, "Integral and adelic aspects of the Mumford–Tate conjecture", preprint, 2015. arXiv

[Cadoret and Tamagawa 2012] A. Cadoret and A. Tamagawa, "A uniform open image theorem for ℓ-adic representations, I", *Duke Math. J.* **161**:13 (2012), 2605–2634. MR Zbl

[Cadoret and Tamagawa 2013] A. Cadoret and A. Tamagawa, "A uniform open image theorem for ℓ-adic representations, II", *Duke Math. J.* **162**:12 (2013), 2301–2344. MR Zbl

[Cadoret and Tamagawa 2014] A. Cadoret and A. Tamagawa, "On the geometric image of $\mathbb{F}_\ell$-linear representations of étale fundamental groups", preprint, 2014, http://www.math.polytechnique.fr/perso/cadoret/JordanNorilIndepFinal.pdf.

[Clozel et al. 2001] L. Clozel, H. Oh, and E. Ullmo, "Hecke operators and equidistribution of Hecke points", *Invent. Math.* **144**:2 (2001), 327–351. MR Zbl

[Deligne 1971] P. Deligne, "Travaux de Shimura", exposé no. 389, 123–165 in *Séminaire Bourbaki, 1970/1971*, Lecture Notes in Mathematics **244**, Springer, Berlin, 1971. MR Zbl

[Deligne 1979] P. Deligne, "Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques", pp. 247–289 in *Automorphic forms, representations and L-functions, Part 2* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proceedings of Symposia in Pure Mathematics **33**, American Mathematical Society, Providence, RI, 1979. MR Zbl

[Eskin and Oh 2006] A. Eskin and H. Oh, "Ergodic theoretic proof of equidistribution of Hecke points", *Ergodic Theory Dynam. Systems* **26**:1 (2006), 163–167. MR Zbl

[Faltings and Wüstholz 1984] G. Faltings and G. Wüstholz (editors), *Rational points* (Max-Planck-Institut für Mathematik, Bonn, 1983/1984), Aspects of Mathematics **E6**, Vieweg, Braunschweig, 1984. MR Zbl

[Gan and Oh 2003] W. T. Gan and H. Oh, "Equidistribution of integer points on a family of homogeneous varieties: a problem of Linnik", *Compositio Math.* **136**:3 (2003), 323–352. MR Zbl

[Garrett 1997] P. Garrett, *Buildings and classical groups*, Chapman & Hall, London, 1997. MR Zbl

[Jantzen 1997] J. C. Jantzen, "Low-dimensional representations of reductive groups are semisimple", pp. 255–266 in *Algebraic groups and Lie groups*, Austral. Math. Soc. Lect. Ser. **9**, Cambridge University Press, 1997. MR Zbl

[Masser and Wüstholz 1993] D. Masser and G. Wüstholz, "Isogeny estimates for abelian varieties, and finiteness theorems", *Ann. of Math.* (2) **137**:3 (1993), 459–472. MR Zbl

[Milne 1986] J. S. Milne, "Abelian varieties", pp. 103–150 in *Arithmetic geometry* (Storrs, CT, 1984), Springer, New York, 1986. MR Zbl

[Milne 2013] J. S. Milne, "Shimura varieties and moduli", pp. 467–548 in *Handbook of moduli, Vol. II*, edited by G. Farkas and I. Morrison, Adv. Lect. Math. **25**, International Press, Somerville, MA, 2013. MR Zbl

[Nori 1987] M. V. Nori, "On subgroups of $\mathrm{GL}_n(\mathbf{F}_p)$", *Invent. Math.* **88**:2 (1987), 257–275. MR Zbl

[Orr 2013] M. Orr, *La conjecture d'André–Pink: Orbites de Hecke et sous-variétés faiblement spéciales*, Ph.D. thesis, Université Paris-Sud, 2013, https://tel.archives-ouvertes.fr/tel-00879010.

[Pink 2005] R. Pink, "A combination of the conjectures of Mordell–Lang and André–Oort", pp. 251–282 in *Geometric methods in algebra and number theory*, edited by F. Bogomolov and Y. Tschinkel, Progr. Math. **235**, Birkhäuser, Boston, 2005. MR Zbl

[Platonov and Rapinchuk 1994] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics **139**, Academic Press, Boston, 1994. MR Zbl

[Prasad 1982] G. Prasad, "Elementary proof of a theorem of Bruhat–Tits–Rousseau and of a theorem of Tits", *Bull. Soc. Math. France* **110**:2 (1982), 197–202. MR Zbl

[Rousseau 2009] G. Rousseau, "Euclidean buildings", pp. 77–116 in *Géométries à courbure négative ou nulle, groupes discrets et rigidités*, edited by L. Bessières et al., Sémin. Congr. **18**, Société Mathématique de France, Paris, 2009. MR Zbl

[Serre 1968] J.-P. Serre, *Corps locaux*, Hermann, Paris, 1968. MR

[Serre 1989] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, Aspects of Mathematics **E15**, Vieweg, Braunschweig, 1989. MR Zbl

[Serre 1994] J.-P. Serre, "Propriétés conjecturales des groupes de Galois motiviques et des représentations l-adiques", pp. 377–400 in *Motives* (Seattle, 1991), edited by U. Jannsen et al., Proceedings of Symposia in Pure Mathematics **55**, American Mathematical Society, Providence, RI, 1994. MR Zbl

[Serre 2013] J.-P. Serre, "Un critère d'indépendance pour une famille de représentations ℓ-adiques", *Comment. Math. Helv.* **88**:3 (2013), 541–554. MR Zbl

[Tits 1979] J. Tits, "Reductive groups over local fields", pp. 29–69 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), vol. 1, edited by A. Borel and W. Casselman, Proceedings of Symposia in Pure Mathematics **33**, American Mathematical Society, Providence, RI, 1979. MR Zbl

[Ullmo and Yafaev 2013] E. Ullmo and A. Yafaev, "Mumford–Tate and generalised Shafarevich conjectures", *Ann. Math. Qué.* **37**:2 (2013), 255–284. MR Zbl

Communicated by Bjorn Poonen
Received 2015-09-13 Revised 2016-06-29 Accepted 2016-08-12

anna.cadoret@polytechnique.edu *Centre de Mathématiques Laurent Schwartz, Ecole Polytechnique, 91128 Palaiseau, France*

arnokret@gmail.com *Faculty of Science, Korteweg–de Vries Instituut, Postbus 94248, 1090 GE Amsterdam, Netherlands*

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1935

Extremality of loci of hyperelliptic curves with marked Weierstrass points

Dawei Chen and Nicola Tarasca

The locus of genus-two curves with n marked Weierstrass points has codimension n inside the moduli space of genus-two curves with n marked points, for $n \leq 6$. It is well known that the class of the closure of the divisor obtained for $n = 1$ spans an extremal ray of the cone of effective divisor classes. We generalize this result for all n: we show that the class of the closure of the locus of genus-two curves with n marked Weierstrass points spans an extremal ray of the cone of effective classes of codimension n, for $n \leq 6$. A related construction produces extremal nef curve classes in moduli spaces of pointed elliptic curves.

Every smooth curve of genus two has a unique map of degree two to the projective line, ramified at six points, called *Weierstrass points*. It follows that the locus $\mathcal{H}yp_{2,n}$ of curves of genus two with n marked Weierstrass points has codimension n inside the moduli space $\mathcal{M}_{2,n}$ of smooth curves of genus two with n marked points, for $1 \leq n \leq 6$. In this paper, we study the classes of the closures of the loci $\mathcal{H}yp_{2,n}$ inside the moduli space of stable curves $\overline{\mathcal{M}}_{2,n}$.

The cone of effective codimension-one classes on $\overline{\mathcal{M}}_{2,1}$ is explicitly described in [Rulla 2001; 2006], and encodes the rational contractions of $\overline{\mathcal{M}}_{2,1}$. It is thus natural to study cones of effective classes of higher codimension. The following is one of the first results in this direction.

Theorem 1. *For $1 \leq n \leq 6$, the class of $\overline{\mathcal{H}yp}_{2,n}$ is rigid and extremal in the cone of effective classes of codimension n in $\overline{\mathcal{M}}_{2,n}$.*

Theorem 1 motivates the computation of the classes of the loci $\overline{\mathcal{H}yp}_{2,n}$. The class of the divisor $\overline{\mathcal{H}yp}_{2,1}$ has been computed in [Eisenbud and Harris 1987], and the class of the codimension-two locus $\overline{\mathcal{H}yp}_{2,2}$ has been computed in [Tarasca 2015]. In Section 5 we study the next nontrivial case.

During the preparation of this article, Chen was partially supported by NSF CAREER grant DMS-1350396.

MSC2010: primary 14H99; secondary 14C99.

Keywords: subvarieties of moduli spaces of curves, effective cones, higher codimensional cycles.

Theorem 2. *In $A^3(\overline{\mathcal{M}}_{2,3})$, we have*

$$[\overline{\mathcal{H}yp}_{2,3}] = \big((3\omega_1 - \lambda - \delta_1)\cdot(3\omega_2 - \lambda - \delta_1) - (\delta_{0:\{1,2\}} + \delta_{0:3})\cdot(3\omega_1 - \lambda - \delta_1) - \gamma_{1:\varnothing} - \gamma_{1:\{3\}}\big) \cdot (3\omega_3 - \lambda - \delta_1 - \delta_{0:\{1,3\}} - \delta_{0:\{2,3\}}) - \gamma_{1:\{1\}}\cdot(2\psi_1 - \delta_{1:\{1\}}) - \gamma_{1:\{2\}}\cdot(2\psi_2 - \delta_{1:\{2\}}) - \gamma_{1:\varnothing}\cdot(\psi_1 - \delta_{0:\{1,3\}}).$$

For elliptic curves, the difference of two ramification points of a degree-2 map to the projective line can be regarded as a 2-torsion point. In a somewhat similar fashion, we consider in general the locus of points on elliptic curves whose pairwise differences are k-torsion points. More precisely, for $k \geq 2$ and $2 \leq n \leq k^2$, consider the following one-dimensional locus in $\mathcal{M}_{1,n}$:

$$\mathcal{T}or^k_{1,n} := \{[C, p_1, \ldots, p_n] \in \mathcal{M}_{1,n} \mid kp_1 \sim \cdots \sim kp_n\}.$$

Note that $\mathcal{T}or^k_{1,n}$ might be reducible: $\mathcal{T}or^d_{1,n}$ is a subcurve of $\mathcal{T}or^k_{1,n}$ for all divisors d of k.

The class of the divisor $\overline{\mathcal{T}or}^2_{1,2}$ is in the interior of the two-dimensional cone of effective divisor classes in $\overline{\mathcal{M}}_{1,2}$ and spans an extremal ray of the cone of nef divisor classes in $\overline{\mathcal{M}}_{1,2}$ [Rulla 2001].

Theorem 3. *For $k \geq 2$ and $2 \leq n \leq k^2$, the class of $\overline{\mathcal{T}or}^k_{1,n}$ spans an extremal ray of the cone of nef curve classes in $\overline{\mathcal{M}}_{1,n}$, and this ray does not dependent on k.*

Structure of the paper. The proof of Theorem 3 is in Section 1 — this section is independent from the rest of the paper. In Section 2 we collect some classical results on classes of hyperelliptic loci which are needed later on. The proof of Theorem 1 in the case $n = 2$ is in Section 4.2 and is based on the explicit description of the codimension-two class $[\overline{\mathcal{H}yp}_{2,2}]$ presented in Section 4.1. In Section 3 we prove a recursive argument that works in a more general context, and we thus complete the proof of Theorem 1. Finally, we prove Theorem 2 in Section 5 using the description of the classes $[\overline{\mathcal{H}yp}_{2,1}]$ and $[\overline{\mathcal{H}yp}_{2,2}]$ from Section 4.1.

Notation. We use throughout the following notation for divisor classes on $\overline{\mathcal{M}}_{g,n}$. The class ψ_i is the cotangent class at the marked point i, and the class ω_i is the pullback of the class ψ_i via the map $\rho_i : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,1}$ obtained by forgetting all marked points but the point i. The class δ_{irr} is the class of the closure of the locus of nodal irreducible curves. We denote by λ the pullback of the first Chern class of the Hodge bundle over $\overline{\mathcal{M}}_g$. For $i \in \{0, \ldots, g\}$ and $J \subseteq \{1, \ldots, n\}$, we denote by $\delta_{i:J}$ the class of the divisor $\Delta_{i:J}$ whose general element has a component of genus i containing the points marked by indices in J and meeting a component of genus $g - i$ containing the remaining marked points. One has $\delta_{i:J} = \delta_{g-i:J^c}$. We denote by $\delta_{i:j}$ the sum of all distinct divisor classes $\delta_{i:J}$ such that $|J| = j$, and by δ_i the sum of all distinct classes $\delta_{i:j}$ for all possible values of j. Let $\pi_k : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,n-1}$ be the map obtained by forgetting the k-th marked point. Note that $\pi_k^*(\delta_i) = \delta_i$ on $\overline{\mathcal{M}}_{g,n}$ for $n \geq 2$.

We also use the following codimension-two tautological classes on $\overline{\mathcal{M}}_{g,n}$. For $J \subseteq \{1, \ldots, n\}$, let $\gamma_{1:J}$ be the class of the locus $\Gamma_{1:J}$ of curves whose general element has an elliptic component containing exactly the points marked by indices in J, and meeting in two points a component of genus $g-2$ containing the remaining marked points.

Throughout we work over an algebraically closed field of characteristic 0. All cycle classes are stack fundamental classes, and all cohomology and Chow groups are taken with rational coefficients. We implicitly assume real coefficients when we consider nef classes and closures of cones of effective classes.

1. Extremal nef curve classes on $\overline{\mathcal{M}}_{1,n}$

In this section we show that the class of the one-dimensional locus $\overline{\mathcal{T}or}^{k}_{1,n}$ spans an extremal ray of the cone $\mathrm{Nef}^{n-1}(\overline{\mathcal{M}}_{1,n})$ of nef curve classes on $\overline{\mathcal{M}}_{1,n}$, for $k \geq 2$ and $2 \leq n \leq k^2$.

By definition, the cone $\mathrm{Nef}^{n-1}(\overline{\mathcal{M}}_{1,n})$ is dual to the cone of pseudoeffective divisors $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$. A subcone S of $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$ is *extremal in* $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$ if whenever $E_1 + E_2 \in S$, then $E_1, E_2 \in S$. We first show that the cone generated by the boundary divisor classes is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$.

Lemma 4. *The cone generated by the classes $\delta_{0:J}$ for $J \subseteq \{1, \ldots, n\}$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$ for $n \geq 2$.*

Proof. We follow the strategy of [Rulla 2001, Corollary 1.4.7], where Rulla shows that the cone generated by the boundary divisor classes on $\overline{\mathcal{M}}_g$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_g)$.

We first show that the classes $\delta_{0:\{i,j\}}$ are extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$, for $n \geq 3$ and $i, j \in \{1, \ldots, n\}$. Consider the one-dimensional family of curves $C_{\{i,j\}}$ obtained by attaching a rational component containing the points with markings i and j at a moving point of an elliptic curve containing the remaining $n-2$ marked points. The curve $C_{\{i,j\}}$ is a moving curve in $\Delta_{0:\{i,j\}}$, and one has $C_{\{i,j\}} \cdot \delta_{0:\{i,j\}} < 0$. It follows that $\delta_{0:\{i,j\}}$ is rigid and extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$. Moreover, $C_{\{i,j\}}$ has empty intersection with $\delta_{0:J}$ for $|J| = 2$ and $J \neq \{i, j\}$. From [Rulla 2001, Lemma 1.4.6], the cone generated by the classes $\delta_{0:J}$ for $|J| = 2$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$.

We then use the following recursion on k, for $3 \leq k \leq n-1$. Suppose that the cone generated by all classes $\delta_{0:J}$ with $|J| < k < n$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$. For $i_1, \ldots, i_k \in \{1, \ldots, n\}$, consider the one-dimensional family $C_{\{i_1,\ldots,i_k\}}$ obtained by attaching a rational component containing the points with markings $i_1, \ldots, i_k$ at a moving point of an elliptic curve containing the remaining $n-k$ marked points. One has $C_{\{i_1,\dots,i_k\}} \cdot \delta_{0:\{i_1,\dots,i_k\}} < 0$ and $C_{\{i_1,\dots,i_k\}} \cdot \delta_{0:J} = 0$ for $|J| \leq k$ and $J \neq \{i_1, \dots, i_k\}$. Again from [Rulla 2001, Lemma 1.4.6], the cone generated by the classes $\delta_{0:J}$ for $|J| \leq k$ is thus extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$.

Finally, let us consider the class $\delta_{0:\{1,\dots,n\}}$ for $n \geq 2$. Consider the one-dimensional family E obtained by attaching a rational curve containing all marked points to a base point of a pencil of plane cubics. One has $E \cdot \delta_{0:\{1,\dots,n\}} < 0$ and $E \cdot \delta_{0:J} = 0$ for $|J| < n$. Hence, the cone generated by $\delta_{0:\{1,\dots,n\}}$ and $\delta_{0:J}$ with $|J| < n$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$. □

We are now ready to prove Theorem 3.

Proof of Theorem 3. Singular elements in $\overline{\mathcal{T}or}^k_{1,n}$ do not have rational tails. Indeed, consider a singular pointed curve $[C, p_1, \dots, p_n]$ inside the closure of $\mathcal{T}or^k_{1,n}$. The condition $kp_i \sim kp_j$ means that there exists an admissible cover $\pi : C \to \mathbb{P}^1$ of degree k totally ramified at p_i and p_j. Suppose C has a rational tail R containing p_i and p_j. By the Riemann–Hurwitz formula, R does not contain any other ramification point of π. Since $\overline{C \setminus R}$ has arithmetic genus 1, one has $\deg(\pi|_{\overline{C\setminus R}}) > 1$. Hence, the tail R has to meet the other components of C in more than one point, a contradiction. It follows that $[\overline{\mathcal{T}or}^k_{1,n}]$ has zero intersection with all divisor classes $\delta_{0:J}$. Note that by the projection formula relative to the natural map $\overline{\mathcal{M}}_{1,n} \to \overline{\mathcal{M}}_{1,1}$, the class $[\overline{\mathcal{T}or}^k_{1,n}]$ has positive intersection with the divisor class λ.

From Lemma 4, the cone generated by the classes $\delta_{0:J}$ is extremal in $\overline{\mathrm{Eff}}^1(\overline{\mathcal{M}}_{1,n})$, hence by duality the class $[\overline{\mathcal{T}or}^k_{1,n}]$ spans an extremal ray of $\mathrm{Nef}^{n-1}(\overline{\mathcal{M}}_{1,n})$, and this ray does not depend on k. □

2. On hyperelliptic loci

In the following, we collect some well-known facts about classes of hyperelliptic loci which we will use later. For $g \geq 2$ and $0 \leq n \leq 2g+2$, let

$$\mathcal{H}yp_{g,n} := \{[C, p_1, \dots, p_n] \in \mathcal{M}_{g,n} \mid C \text{ is hyperelliptic and } h^0(C, \mathcal{O}_C(2p_i)) \geq 2, \text{ for } i = 1, \dots, n\}$$

be the locus of hyperelliptic curves of genus g with n marked Weierstrass points. The locus $\mathcal{H}yp_{g,n}$ has codimension $g-2+n$ in the moduli space $\mathcal{M}_{g,n}$ of smooth curves of genus g with n marked points. The class of the closure $\overline{\mathcal{H}yp}_{g,n}$ is tautological on the moduli of stable curves $\overline{\mathcal{M}}_{g,n}$ [Faber and Pandharipande 2005].

Let $\mathcal{M}^{\mathrm{rt}}_{g,n}$ be the moduli space of curves with rational tails. From [Faber and Pandharipande 2005] or [Graber and Vakil 2005], the tautological group $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$ is one-dimensional. When $n = 0$, $R^{g-2}(\mathcal{M}_g)$ is one-dimensional and is generated by the class of the hyperelliptic locus $\mathcal{H}yp_g$, or equivalently the class κ_{g-2} [Looijenga 1995; Faber 1999]. Let $\mathcal{H}yp^{\mathrm{rt}}_{g,n}$ be the restriction of $\overline{\mathcal{H}yp}_{g,n}$ to $\mathcal{M}^{\mathrm{rt}}_{g,n}$. Since the pushforward of $[\mathcal{H}yp^{\mathrm{rt}}_{g,n}]$ via the natural map $\mathcal{M}^{\mathrm{rt}}_{g,n} \to \mathcal{M}_g$ is a positive multiple of $[\mathcal{H}yp_g]$, it follows that $[\mathcal{H}yp^{\mathrm{rt}}_{g,n}]$ is nonzero and generates $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$.

Equivalently, $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$ is generated by the decorated class $\delta_{g,\psi^{g-1}}$ defined in the following way. Consider the gluing map

$$\xi : \overline{\mathcal{M}}_{g,1} \to \overline{\mathcal{M}}_{g,n}$$

obtained by attaching a chain of $n-1$ rational components at the marked point of an element in $\overline{\mathcal{M}}_{g,1}$. We fix the markings in an increasing order, from the inner rational component to the outer one. From the rational equivalence of points in $\overline{\mathcal{M}}_{0,n+1}$, the classes of the loci obtained by permuting the markings on the image of ξ are all rationally equivalent. The class $\delta_{g,\psi^{g-1}}$ is defined as the pushforward of the class ψ^{g-1} in $R^{g-1}(\overline{\mathcal{M}}_{g,1})$ via the map ξ.

Let $\pi : \overline{\mathcal{M}}_{g,1} \to \overline{\mathcal{M}}_g$ and $\pi_n : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,n-1}$ be the natural maps. Note that $(\pi_n)_*\delta_{g,\psi^{g-1}} = \delta_{g,\psi^{g-1}}$ in $R^{g-2+n-1}(\mathcal{M}_{g,n-1})$ for $n \geq 3$, and $(\pi_2)_*\delta_{g,\psi^{g-1}} = \psi^{g-1}$. Since $\kappa_{g-2} := \pi_*(\psi^{g-1})$ is nonzero in $R^{g-2}(\mathcal{M}_g)$, we conclude $\delta_{g,\psi^{g-1}}$ is nonzero in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$.

Example. In the case $g=2$, for $2 \leq n \leq 6$ we have

$$[\mathcal{H}yp^{\mathrm{rt}}_{2,n}] = \frac{6!}{2\cdot(6-n)!}\,\delta_{2,\psi} \in R^n(\mathcal{M}^{\mathrm{rt}}_{2,n}).$$

Indeed, let us write $[\mathcal{H}yp^{\mathrm{rt}}_{2,n}] = \alpha\delta_{2,\psi}$ in $R^n(\mathcal{M}^{\mathrm{rt}}_{2,n})$. In order to determine the coefficient α, we intersect both sides of the equation with a test space. Let C be a smooth curve of genus 2, and let $C[n]$ be the n-th Fulton–MacPherson compactification of the space of n distinct points of C [Fulton and MacPherson 1994]. The natural map $C[n+1] \to C[n]$ gives an n-dimensional family of genus-two curves with rational tails. Weierstrass points on C are ramification points of the hyperelliptic double covering. Analyzing the Hurwitz space of admissible double coverings, it is easy to see that the intersection $[\mathcal{H}yp^{\mathrm{rt}}_{2,n}]\cdot C[n]$ corresponds to all ordered n-tuples of Weierstrass points in C, and is transversal. We deduce that $[\mathcal{H}yp^{\mathrm{rt}}_{2,n}]\cdot C[n] = 6!/(6-n)!$. On the other hand, one has $\delta_{2,\psi}\cdot C[n] = \psi\cdot\xi^*(C[n]) = \psi\cdot C[1] = 2$, whence the statement.

3. A recursive argument

Let $N^k(\overline{\mathcal{M}}_{g,n})$ be the group of codimension-k cycles on $\overline{\mathcal{M}}_{g,n}$ modulo numerical equivalence. We denote by $\mathrm{Eff}^k(\overline{\mathcal{M}}_{g,n}) \subset N^k(\overline{\mathcal{M}}_{g,n})$ the *cone of effective cycle classes*, and by $\mathrm{REff}^k(\overline{\mathcal{M}}_{g,n}) \subseteq \mathrm{Eff}^k(\overline{\mathcal{M}}_{g,n})$ the *subcone of effective tautological classes* (see [Faber and Pandharipande 2005] for tautological classes on $\overline{\mathcal{M}}_{g,n}$).

A cycle class E inside a cone $K \subset N^k(\overline{\mathcal{M}}_{g,n})$ is called *extremal in K* if whenever two cycle classes E_1 and E_2 in K are such that $E = E_1 + E_2$, then both E_1 and

E_2 lie in the ray spanned by E. An effective cycle class E is called *rigid* if any effective cycle with class mE is supported on the support of E.

Theorem 5. *Given $g \geq 2$, if $[\overline{\mathcal{H}yp}_{g,2}]$ is rigid and extremal in* $\mathrm{REff}^g(\overline{\mathcal{M}}_{g,2})$, *then $[\overline{\mathcal{H}yp}_{g,n}]$ is rigid and extremal in* $\mathrm{REff}^{g-2+n}(\overline{\mathcal{M}}_{g,n})$, *for $3 \leq n \leq 2g+2$.*

Proof. Let $n \geq 3$ and assume that the statement is true for $\overline{\mathcal{H}yp}_{g,n-1}$. Suppose that

$$[\overline{\mathcal{H}yp}_{g,n}] = \sum_i a_i [X_i], \tag{1}$$

with $a_i > 0$, X_i irreducible, tautological, effective of codimension n, and $[X_i]$ not proportional to $[\overline{\mathcal{H}yp}_{g,n}]$, for all i.

Since $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$ is generated by $\delta_{g,\psi^{g-1}}$ (see Section 2), we can express the class of each X_i as

$$[X_i] = c_i \delta_{g,\psi^{g-1}} + B_i,$$

where c_i is a nonnegative coefficient, and B_i is a (not necessarily effective) cycle class in $R^{g-2+n}(\overline{\mathcal{M}}_{g,n})$ with $B_i = 0$ in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$. Let $\pi_j : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,n-1}$ be the map obtained by forgetting the point j. Applying $(\pi_j)_*$ to (1), we have

$$(2g+2-(n-1))[\overline{\mathcal{H}yp}_{g,n-1}] = \sum_i a_i (\pi_j)_*[X_i]. \tag{2}$$

Pick a locus X_i appearing on the right side of (1). Consider two cases. First, suppose $(\pi_1)_*[X_i] = \cdots = (\pi_n)_*[X_i] = 0$. Note that $(\pi_j)_*\delta_{g,\psi^{g-1}} = \delta_{g,\psi^{g-1}}$ in $R^{g-2+n-1}(\mathcal{M}^{\mathrm{rt}}_{g,n-1})$, for all $j = 1, \ldots, n$. Since $B_i = 0$ in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$, using the exact sequence

$$A^{g-2+n}(\overline{\mathcal{M}}_{g,n} \setminus \mathcal{M}^{\mathrm{rt}}_{g,n}) \to A^{g-2+n}(\overline{\mathcal{M}}_{g,n}) \to A^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n}) \to 0,$$

we can assume that B_i is represented by a linear combination of cycle classes supported in $\overline{\mathcal{M}}_{g,n} \setminus \mathcal{M}^{\mathrm{rt}}_{g,n}$. An element in the support of such a cycle does not have an irreducible and smooth component of genus g, hence $(\pi_j)_*B_i = 0$ in $A^{g-2+n-1}(\mathcal{M}^{\mathrm{rt}}_{g,n-1})$, for all $j = 1, \ldots, n$. We deduce that $c_i = 0$, that is, $[X_i] = 0$ in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$.

For the other case, suppose $(\pi_1)_*[X_i]$ is nonzero. Since the class $[\overline{\mathcal{H}yp}_{g,n-1}]$ is rigid and extremal in $\mathrm{REff}^{g+n-3}(\overline{\mathcal{M}}_{g,n-1})$, from (2) we deduce that $(\pi_1)_*[X_i]$ is a positive multiple of the class of $\overline{\mathcal{H}yp}_{g,n-1}$ and, moreover, $X_i \subset (\pi_1)^{-1}\overline{\mathcal{H}yp}_{g,n-1}$. This implies that $(\pi_2)_*[X_i], \ldots, (\pi_n)_*[X_i]$ are also nonzero. It follows that X_i is in the intersection of all the $(\pi_j)^{-1}\overline{\mathcal{H}yp}_{g,n-1}$, for $j = 1, \ldots, n$. In particular, any $n-1$ marked points in a general element of X_i are distinct Weierstrass points. Hence, all n marked points must be distinct Weierstrass points. (Note that $n \geq 3$.) This forces $[X_i]$ to be a positive multiple of $[\overline{\mathcal{H}yp}_{g,n}]$, a contradiction.

Finally, the above steps show that $[X_i] = 0$ in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$, for all i. This yields a contradiction since $[\overline{\mathcal{H}yp}_{g,n}] \neq 0$ in $R^{g-2+n}(\mathcal{M}^{\mathrm{rt}}_{g,n})$ (see Section 2), hence $[\overline{\mathcal{H}yp}_{g,n}]$ is extremal in $\mathrm{REff}^{g-2+n}(\overline{\mathcal{M}}_{g,n})$.

Suppose that $E := m[\overline{\mathcal{H}yp}_{g,n}]$ is effective. Since

$$(\pi_j)_*(E) = (2g-n+3)m[\overline{\mathcal{H}yp}_{g,n-1}],$$

and since $[\overline{\mathcal{H}yp}_{g,n-1}]$ is rigid, $(\pi_j)_*(E)$ is supported on $\overline{\mathcal{H}yp}_{g,n-1}$, for $j = 1,\ldots,n$. This implies that E is supported on the intersection of all the $(\pi_j)^{-1}\overline{\mathcal{H}yp}_{g,n-1}$, for $j = 1,\ldots,n$. We conclude that E is supported on $\overline{\mathcal{H}yp}_{g,n}$ and $[\overline{\mathcal{H}yp}_{g,n}]$ is rigid. □

Remark 6. The classes $[\overline{\mathcal{H}yp}_3]$, $[\overline{\mathcal{H}yp}_{3,1}]$, and $[\overline{\mathcal{H}yp}_4]$ are known to be extremal in $\mathrm{Eff}^1(\overline{\mathcal{M}}_3)$ [Rulla 2001], $\mathrm{Eff}^2(\overline{\mathcal{M}}_{3,1})$, and $\mathrm{Eff}^2(\overline{\mathcal{M}}_4)$ [Chen and Coskun 2015], respectively. It is natural to wonder whether $[\overline{\mathcal{H}yp}_{g,n}]$ is extremal in $\mathrm{REff}^{g-2+n}(\overline{\mathcal{M}}_{g,n})$, for all $g \geq 2$ and $0 \leq n \leq 2g+2$. By Theorem 5, it is enough to study the cases $n \leq 2$.

4. Loci of Weierstrass points on curves of genus 2

In this section, we complete the proof of Theorem 1. It is enough to show that $[\overline{\mathcal{H}yp}_{2,n}]$ is rigid and extremal in $\mathrm{REff}^n(\overline{\mathcal{M}}_{2,n})$, and to use the fact that for small values of n, $\mathrm{REff}^*(\overline{\mathcal{M}}_{2,n}) = \mathrm{Eff}^*(\overline{\mathcal{M}}_{2,n})$. Indeed, according to [Petersen and Tommasi 2014; Petersen 2016, Theorem 3.8], all even cohomology of $\overline{\mathcal{M}}_{2,n}$ is tautological for $n < 20$. Note that the Betti numbers of $\overline{\mathcal{M}}_{2,n}$ for $n \leq 7$ have been computed in [Getzler 1998; Bergström 2009].

4.1. ***The classes for n = 1, 2.*** When $n = 1$, the class of the divisor $\overline{\mathcal{H}yp}_{2,1}$ in $\overline{\mathcal{M}}_{2,1}$ is

$$[\overline{\mathcal{H}yp}_{2,1}] = 3\omega - \tfrac{1}{10}\delta_{\mathrm{irr}} - \tfrac{6}{5}\delta_1 = 3\omega - \lambda - \delta_1 \in \mathrm{Pic}(\overline{\mathcal{M}}_{2,1}) \tag{3}$$

[Eisenbud and Harris 1987, Theorem 2.2], and $[\overline{\mathcal{H}yp}_{2,1}]$ is rigid and extremal in $\mathrm{Eff}^1(\overline{\mathcal{M}}_{2,1})$ [Rulla 2001]. When $n = 2$, the class of the double ramification locus $\overline{\mathcal{H}yp}_{2,2}$ in $\overline{\mathcal{M}}_{2,2}$ is

$$\begin{aligned}[\overline{\mathcal{H}yp}_{2,2}] = 6\psi_1\cdot\psi_2 - \tfrac{3}{2}(\psi_1^2+\psi_2^2)&\\ -(\psi_1+\psi_2)\cdot\left(\tfrac{21}{10}\delta_{1:1} + \tfrac{3}{5}\delta_{1:0} + \tfrac{1}{20}\delta_{\mathrm{irr}}\right) &\in A^2(\overline{\mathcal{M}}_{2,2}) \end{aligned}\tag{4}$$

and $\overline{\mathcal{H}yp}_{2,2}$ is not a complete intersection [Tarasca 2015]. Expressing products of divisor classes in terms of decorated boundary strata classes, we have

$$[\overline{\mathcal{H}yp}_{2,2}] = 5\delta_{2,w} + 9\delta_{11|} + \tfrac{5}{8}\delta_{01|} - \tfrac{1}{8}(\delta_{01|1} + \delta_{01|2} + \delta_{01|12}) + 2\gamma_{1:\varnothing} + \tfrac{1}{24}\delta_{00}. \tag{5}$$

Here, $\delta_{2,w}$ is the class of the locus of curves with a rational tail containing both marked points attached at a Weierstrass point on a component of genus 2; $\delta_{11|}$ is the class of the locus of curves whose general element has two elliptic tails attached at a rational component containing both marked points; $\delta_{01|}$, $\delta_{01|1}$, $\delta_{01|2}$, $\delta_{01|12}$ are the classes of the loci of curves whose general element has an elliptic tail attached at a nodal rational component with the points distributed in the following way: for the class $\delta_{01|}$ both marked points are on the rational component, for $\delta_{01|i}$ the point i is on the elliptic component and the other marked point is on the rational component, and for $\delta_{01|12}$ both marked points are on the elliptic component; $\gamma_{1:\varnothing}$ is the class of the locus of curves with an elliptic component meeting in two points a rational component containing both marked points; and finally, δ_{00} is the class of the locus whose general element is a rational curve with two nondisconnecting nodes.

In Section 4.2, we will use in a crucial way the expression in (5). In Section 5, we will also use the following description. Let $\pi_i : \overline{\mathcal{M}}_{2,2} \to \overline{\mathcal{M}}_{2,1}$ be the map obtained by forgetting the point i, for $i = 1, 2$.

Lemma 7. *The following equality holds in* $A^2(\overline{\mathcal{M}}_{2,2})$:

$$\pi_1^*[\overline{\mathcal{H}yp}_{2,1}] \cdot \pi_2^*[\overline{\mathcal{H}yp}_{2,1}] = [\overline{\mathcal{H}yp}_{2,2}] + \gamma_{1:\varnothing} + \delta_{2,w}.$$

In particular, $\pi_1^{-1}(\overline{\mathcal{H}yp}_{2,1}) \cap \pi_2^{-1}(\overline{\mathcal{H}yp}_{2,1})$ *is the union of the supports of* $[\overline{\mathcal{H}yp}_{2,2}]$, $\gamma_{1:\varnothing}$, *and* $\delta_{2,w}$.

Proof. The desired equality follows from (3) and (4). Since the supports of $[\overline{\mathcal{H}yp}_{2,2}]$, $\gamma_{1:\varnothing}$, and $\delta_{2,w}$ are contained in $\pi_i^{-1}(\overline{\mathcal{H}yp}_{2,1})$, for $i = 1, 2$, the statement follows. □

Note that

$$\delta_{2,w} = \delta_{0:2} \cdot \pi_1^*[\overline{\mathcal{H}yp}_{2,1}] = \delta_{0:2} \cdot (3\omega_1 - \lambda - \delta_1).$$

Hence, we can write

$$\begin{aligned}[\overline{\mathcal{H}yp}_{2,2}] &= \pi_1^*(3\omega_2 - \lambda - \delta_1) \cdot \pi_2^*(3\omega_1 - \lambda - \delta_1) - \delta_{0:2} \cdot (3\omega_1 - \lambda - \delta_1) - \gamma_{1:\varnothing} \\ &= (3\omega_2 - \lambda - \delta_1) \cdot (3\omega_1 - \lambda - \delta_1) - \delta_{0:2} \cdot (3\omega_1 - \lambda - \delta_1) - \gamma_{1:\varnothing}. \qquad (6)\end{aligned}$$

4.2. ***The extremality for n = 2.*** By Theorem 5, in order to show that $[\overline{\mathcal{H}yp}_{2,n}]$ is rigid and extremal in $\mathrm{REff}^n(\overline{\mathcal{M}}_{2,n})$ for $2 \le n \le 6$, it is enough to show that $[\overline{\mathcal{H}yp}_{2,2}]$ is rigid and extremal in $\mathrm{Eff}^2(\overline{\mathcal{M}}_{2,2})$.

Theorem 8. $[\overline{\mathcal{H}yp}_{2,2}]$ *is rigid and extremal in* $\mathrm{Eff}^2(\overline{\mathcal{M}}_{2,2})$.

Proof. Suppose that

$$[\overline{\mathcal{H}yp}_{2,2}] = \sum_i a_i [X_i], \qquad (7)$$

where $a_i > 0$ and X_i is an irreducible codimension-two effective cycle on $\overline{\mathcal{M}}_{2,2}$ with $[X_i]$ not proportional to $[\overline{\mathcal{H}yp}_{2,2}]$, for all i. Let $\pi_j : \overline{\mathcal{M}}_{2,2} \to \overline{\mathcal{M}}_{2,1}$ be the map forgetting the point j, for $j = 1, 2$, and $\pi : \overline{\mathcal{M}}_{2,2} \to \overline{\mathcal{M}}_2$ be the map forgetting both marked points. Applying $(\pi_j)_*$ to (7), we obtain

$$5[\overline{\mathcal{H}yp}_{2,1}] = \sum_i a_i (\pi_1)_*[X_i]. \qquad (8)$$

Pick a locus X_i appearing on the right side of (7). If $(\pi_1)_*[X_i] = (\pi_2)_*[X_i] = 0$, then either X_i is contained in the inverse image via π of a codimension-two effective cycle on $\overline{\mathcal{M}}_2$, or a general point of X_i contains a smooth rational component with two marked points and two singular points. Note that a codimension-two locus in $\overline{\mathcal{M}}_2$ is a curve, and the cone of effective curves in $\overline{\mathcal{M}}_2$ is known to be spanned by the two one-dimensional boundary strata. We deduce that $[X_i]$ is in the cone generated by the boundary strata classes $\delta_{01|}$, $\delta_{01|1}$, $\delta_{01|2}$, $\delta_{01|12}$, δ_{00}, $\delta_{11|}$, and $\gamma_{1:\varnothing}$.

Suppose $(\pi_1)_*[X_i]$ is nonzero. Since $[\overline{\mathcal{H}yp}_{2,1}]$ is rigid and extremal in $\mathrm{Eff}^1(\overline{\mathcal{M}}_{2,1})$, from (8) we deduce that $X_i \subset \pi_1^{-1}(\overline{\mathcal{H}yp}_{2,1})$. Hence, $(\pi_2)_*[X_i]$ is also nonzero, and $X_i \subset \pi_1^{-1}(\overline{\mathcal{H}yp}_{2,1}) \cap \pi_2^{-1}(\overline{\mathcal{H}yp}_{2,1})$. From Lemma 7 we conclude that $[X_i]$ is supported on the locus of curves with a rational tail containing both marked points attached at a Weierstrass point of a genus-two curve, hence $[X_i]$ is a positive multiple of $\delta_{2,w}$.

From (5), the class of $\overline{\mathcal{H}yp}_{2,2}$ lies outside the cone generated by $\delta_{2,w}$, $\delta_{01|}$, $\delta_{01|1}$, $\delta_{01|2}$, $\delta_{01|12}$, δ_{00}, $\delta_{11|}$, and $\gamma_{1:\varnothing}$. Indeed, the coefficient of $\delta_{01|1} + \delta_{01|2} + \delta_{01|12}$ is negative. Hence $[\overline{\mathcal{H}yp}_{2,2}]$ is extremal in $\mathrm{Eff}^2(\overline{\mathcal{M}}_{2,2})$.

The rigidity of $[\overline{\mathcal{H}yp}_{2,2}]$ follows from a similar argument. Suppose that for some positive m, $E := m[\overline{\mathcal{H}yp}_{2,2}]$ is effective. Since $(\pi_j)_*(E) = 5m[\overline{\mathcal{H}yp}_{2,1}]$ and $[\overline{\mathcal{H}yp}_{2,1}]$ is rigid, we have that $(\pi_j)_*(E)$ is supported on $\overline{\mathcal{H}yp}_{2,1}$, for $j = 1, 2$. Hence, the support of E is in $\pi_1^{-1}(\overline{\mathcal{H}yp}_{2,1}) \cap \pi_2^{-1}(\overline{\mathcal{H}yp}_{2,1})$. From Lemma 7, E is supported on the union of the loci $\overline{\mathcal{H}yp}_{2,2}$, $\Gamma_{1:\varnothing}$, and $\Delta_{2,w}$. Since $E = m[\overline{\mathcal{H}yp}_{2,2}]$, E is supported only on $\overline{\mathcal{H}yp}_{2,2}$, and the statement follows. □

5. The class of $\overline{\mathcal{H}yp}_{2,3}$

The aim of this section is to compute the class of $\overline{\mathcal{H}yp}_{2,3}$ in $A^3(\overline{\mathcal{M}}_{2,3})$. We first discuss a recursive relation between the classes of a partial closure of $\mathcal{H}yp_{2,n}$ and $\mathcal{H}yp_{2,n-1}$.

Recall the map $\pi_i : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,n-1}$ obtained by forgetting the i-th marked point, and the map $\rho_i : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{M}}_{g,1}$ obtained by forgetting all but the i-th marked point.

5.1. ***A recursive relation.*** Let $\overline{\mathcal{M}}^o_{g,n}$ be the open locus in $\overline{\mathcal{M}}_{g,n}$ of stable curves with at most one nondisconnecting node. Let $\overline{\mathcal{H}yp}^o_{g,n}$ be the closure of $\mathcal{H}yp_{g,n}$ in $\overline{\mathcal{M}}^o_{g,n}$. For $2 \le n \le 6$, we note the following identity in $A^n(\overline{\mathcal{M}}^o_{2,n})$:

$$\pi_n^*(\overline{\mathcal{H}yp}^o_{2,n-1}) \cdot \rho_n^*(\overline{\mathcal{H}yp}_{2,1}) \equiv \overline{\mathcal{H}yp}^o_{2,n} + \sum_{i=1}^{n-1} \pi_n^*(\overline{\mathcal{H}yp}^o_{2,n-1}) \cdot \delta_{0:\{i,n\}}. \tag{9}$$

Indeed, the intersection on the left-hand side consists of genus-two curves with a choice of n ordered Weierstrass points, the first $n-1$ being distinct. The component $\overline{\mathcal{H}yp}^o_{2,n}$ corresponds to curves with all n points distinct, and the component $\pi_n^*(\overline{\mathcal{H}yp}^o_{2,n-1}) \cdot \delta_{0:\{i,n\}}$ corresponds to curves with the n-th point coinciding with the i-th point, for $i = 1, \dots, n-1$. A Weierstrass point on a smooth hyperelliptic curve of genus g has weight $g(g-1)/2$. This explains the coefficient of $\overline{\mathcal{H}yp}^{o}_{2,n}$. Since the right-hand side is symmetric with respect to the first $n-1$ points, it is clear that all the components $\pi_n^*(\overline{\mathcal{H}yp}^{o}_{2,n-1}) \cdot \delta_{0:\{i,n\}}$ have equal multiplicity, which, forgetting the point n, must equal 1.

Using (9), one can recursively express the class of $\overline{\mathcal{H}yp}^{o}_{2,n}$ in terms of products of divisor classes. In the following, we derive a complete expression for the class of $\overline{\mathcal{H}yp}_{2,3}$ in $A^3(\overline{\mathcal{M}}_{2,3})$.

5.2. ***A set-theoretic description.*** To extend (9) with $n = 3$ over $\overline{\mathcal{M}}_{2,3}$, we need to consider loci of curves with at least two nondisconnecting nodes. Let Ξ_i be the closure of the locus of curves with an elliptic component $[E, p_i, x, y]$ such that $2p_i \sim x + y$, and a rational component containing the other two marked points p_j, p_k, and meeting E at the points x, y. Let Θ be the closure of the locus of curves whose general element has a rational component $[R, p_1, p_2, p_3, x, y]$ such that $2p_1 \sim 2p_2 \sim x + y$, and an elliptic component meeting R at the points x, y.

Proposition 9. *We have*

$$\pi_3^*(\overline{\mathcal{H}yp}_{2,2}) \cap \rho_3^*(\overline{\mathcal{H}yp}_{2,1}) = \overline{\mathcal{H}yp}_{2,3} \cup \pi_3^*(\overline{\mathcal{H}yp}_{2,2})_{|\Delta_{0:\{1,3\}}} \cup \pi_3^*(\overline{\mathcal{H}yp}_{2,2})_{|\Delta_{0:\{2,3\}}} \cup \Xi_1 \cup \Xi_2 \cup \Theta.$$

Proof. The intersection

$$\pi_3^*(\overline{\mathcal{H}yp}_{2,2}) \cap \rho_3^*(\overline{\mathcal{H}yp}_{2,1})$$

consists of stable curves $[C, p_1, p_2, p_3]$ with three marked Weierstrass points and with p_1 and p_2 corresponding to two different Weierstrass points. If the three points correspond to three different Weierstrass points, then $[C, p_1, p_2, p_3]$ is in $\overline{\mathcal{H}yp}_{2,3}$. If p_3 and p_1 correspond to the same Weierstrass point, then $[C, p_1, p_2, p_3]$ is in the restriction of $\pi_3^*(\overline{\mathcal{H}yp}_{2,2})$ to $\Delta_{0:\{1,3\}}$. The case when p_3 and p_2 correspond to the same Weierstrass point is similar. Finally, restricting the intersection to the codimension-two boundary strata and using admissible covers to describe Weierstrass points on singular curves, we deduce that Ξ_1, Ξ_2, and Θ are the only additional components contained in the intersection, hence the statement. □

5.3. ***The multiplicities.*** Since the left-hand side of the expression in Proposition 9 is symmetric with respect to the first two marked points, we conclude that the contributions of $\pi_3^*[\overline{\mathcal{H}yp}_{2,2}] \cdot \delta_{0:\{1,3\}}$ and $\pi_3^*[\overline{\mathcal{H}yp}_{2,2}] \cdot \delta_{0:\{2,3\}}$ on the right-hand side are equal. Similarly for Ξ_1 and Ξ_2. Hence we have, for some coefficients $\alpha, \beta, \gamma, \delta$,

$$\pi_3^*[\overline{\mathcal{H}yp}_{2,2}] \cdot \rho_3^*[\overline{\mathcal{H}yp}_{2,1}] = \alpha[\overline{\mathcal{H}yp}_{2,3}] + \beta(\delta_{0:\{1,3\}} + \delta_{0:\{2,3\}}) \cdot \pi_3^*[\overline{\mathcal{H}yp}_{2,2}] + \gamma([\Xi_1] + [\Xi_2]) + \delta[\Theta]. \quad (10)$$

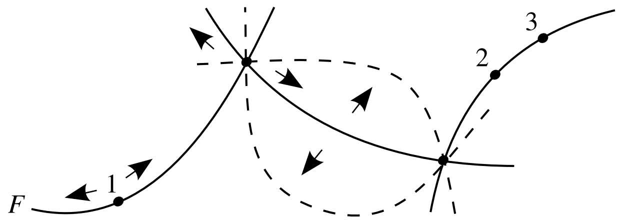


Figure 1. How the general element of the family moves.

Forgetting the first marked point in (10), the left-hand side is

$$5(\pi_1^*[\overline{\mathcal{H}yp}_{2,1}]\cdot\pi_2^*[\overline{\mathcal{H}yp}_{2,1}]) = 5([\overline{\mathcal{H}yp}_{2,2}]+\gamma_{1:\varnothing}+\delta_{2,w})$$

by Lemma 7, hence we have

$$5([\overline{\mathcal{H}yp}_{2,2}]+\gamma_{1:\varnothing}+\delta_{2,w}) = (4\alpha+\beta)[\overline{\mathcal{H}yp}_{2,2}]+(4\gamma+\delta)\gamma_{1:\varnothing}+5\beta\delta_{2,w}.$$

We deduce $\alpha=\beta=1$ and $4\gamma+\delta=5$.

In order to compute γ and δ, we consider the restriction of (10) to the following three-dimensional test family. Attach at two points of an elliptic curve E a rational tail containing the points marked by 2 and 3, and an elliptic tail F containing the point marked by 1. Consider the family obtained by varying E in a pencil of degree 12, by varying the point of intersection with the rational tail on the central elliptic component in which it lies, and by varying the point marked by 1 on F (see Figure 1). The base of this family is $Y\times F$, where Y is the blow-up of $\mathbb{P}^2$ at the nine points of intersection of two general cubics.

Let H be the pullback of the hyperplane class in $\mathbb{P}^2$, let Σ be the sum of the nine exceptional divisors, and let E_0 be one of them. Denote by $\pi : Y\times F\to F$ the natural projection, and let $q=E\cap F$ be the singular point on F. The divisor classes on $\overline{\mathcal{M}}_{2,3}$ restrict as follows:

$$\begin{aligned}
\psi_1 &= \pi^*(q),\\
\delta_{\mathrm{irr}} &= 36H-12\Sigma = 12\lambda,\\
\delta_{0:\{2,3\}} &= -3H+\Sigma-E_0,\\
\delta_{1:\{1\}} &= -3H+\Sigma-E_0-\pi^*(q),\\
\delta_{1:0} &= E_0+\pi^*(q).
\end{aligned}$$

From (3) and (4), it follows that

$$\begin{aligned}
\rho_3^*[\overline{\mathcal{H}yp}_{2,1}]\cdot\pi_3^*[\overline{\mathcal{H}yp}_{2,2}] &= -\big(3\delta_{0:\{2,3\}}+\tfrac{1}{10}\delta_{\mathrm{irr}}+\tfrac{6}{5}(\delta_{1:\{1\}}+\delta_{1:0})\big)\\
&\quad\cdot\big(-6\psi_1\cdot\delta_{0:\{2,3\}}-\tfrac{3}{2}\delta_{0:\{2,3\}}^2-(\psi_1-\delta_{0:\{2,3\}})\cdot\big(\tfrac{21}{10}\delta_{1:\{1\}}+\tfrac{3}{5}\delta_{1:0}+\tfrac{1}{20}\delta_{\mathrm{irr}}\big)\big)=27,
\end{aligned}$$

and similarly,

$$\pi_3^*[\overline{\mathcal{H}yp}_{2,2}] \cdot \delta_{0:\{2,3\}} = -9.$$

Note that this family meets Ξ_1 when E degenerates to one of the 12 rational nodal fibers, the rational tail is attached at a point colliding with the nondisconnecting node, and the point marked by 1 differs from q in F by a nontrivial torsion point of order 2 in $\mathrm{Pic}^0(F)$. The intersection is transverse at each point, hence we have

$$\Xi_1 = 12 \cdot 3.$$

All other classes in (10) are disjoint from this family. We deduce the relation

$$27 = -9\beta + 36\gamma,$$

and hence conclude that $\alpha = \beta = \gamma = \delta = 1$. We have thus proved the following statement.

Proposition 10. *One has*

$$[\overline{\mathcal{H}yp}_{2,3}] = \pi_3^*[\overline{\mathcal{H}yp}_{2,2}] \cdot (\rho_3^*[\overline{\mathcal{H}yp}_{2,1}] - \delta_{0:\{1,3\}} - \delta_{0:\{2,3\}}) - [\Xi_1] - [\Xi_2] - [\Theta].$$

5.4. ***The boundary components.*** It remains to compute the classes of Ξ_1, Ξ_2, and Θ. Recall the classes $\gamma_{1:J}$ defined in the introduction.

Lemma 11. *The following equalities hold in* $A^3(\overline{\mathcal{M}}_{2,3})$:

$$[\Xi_i] = (2\psi_i - \delta_{1:\{i\}}) \cdot \gamma_{1:\{i\}} \qquad \text{for } i = 1, 2,$$
$$[\Theta] = \gamma_{1:\varnothing} \cdot (\psi_1 - \delta_{0:\{1,3\}}) = \gamma_{1:\varnothing} \cdot (\psi_2 - \delta_{0:\{2,3\}}).$$

Proof. Consider the divisor D_i of curves $[E, p_i, x, y]$ in $\mathcal{M}_{1,3}$ such that $2p_i \sim x + y$. From [Belorousski and Pandharipande 2000, §2.6] or [Chen and Coskun 2014, Proposition 3.1], one has $[\overline{D}_i] = 2\psi_i - \delta_{1:\{i\}}$ in $\mathrm{Pic}(\overline{\mathcal{M}}_{1,3})$. The locus Ξ_i is the pushforward of $\overline{D}_i \times \overline{\mathcal{M}}_{0,4} \subset \overline{\mathcal{M}}_{1,3} \times \overline{\mathcal{M}}_{0,4}$ via the natural map

$$\overline{\mathcal{M}}_{1,3} \times \overline{\mathcal{M}}_{0,4} \to \Gamma_{1:\{i\}} \subset \overline{\mathcal{M}}_{2,3}.$$

Similarly, consider the map $\xi : \overline{\mathcal{M}}_{1,2} \times \overline{\mathcal{M}}_{0,5} \to \Gamma_{1:\varnothing} \subset \overline{\mathcal{M}}_{2,3}$ defined as

$$([E, x_1, y_1], [R, p_1, p_2, p_3, x_2, y_2]) \mapsto [E \cup_{x_1 \sim x_2, y_1 \sim y_2} R, p_1, p_2, p_3].$$

The locus Θ is the pushforward via ξ of the locus $\overline{\mathcal{M}}_{1,2} \times \pi_3^*(\text{point}) \subset \overline{\mathcal{M}}_{1,2} \times \overline{\mathcal{M}}_{0,5}$, where $\pi_3 : \overline{\mathcal{M}}_{0,5} \to \overline{\mathcal{M}}_{0,4}$ is the map obtained by forgetting the point p_3. The statement follows. □

From (3), (6), and Lemma 11, the resulting expression in Proposition 10 gives the statement in Theorem 2.

Acknowledgements

We would like to thank Izzet Coskun for helpful discussions on related ideas, Dan Petersen for pointing our attention to the results in [Petersen and Tommasi 2014; Petersen 2016], and the referee for suggesting many improvements on the exposition.

References

[Belorousski and Pandharipande 2000] P. Belorousski and R. Pandharipande, "A descendent relation in genus 2", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **29**:1 (2000), 171–191. MR Zbl

[Bergström 2009] J. Bergström, "Equivariant counts of points of the moduli spaces of pointed hyperelliptic curves", *Doc. Math.* **14** (2009), 259–296. MR Zbl

[Chen and Coskun 2014] D. Chen and I. Coskun, "Extremal effective divisors on $\overline{\mathcal{M}}_{1,n}$", *Math. Ann.* **359**:3-4 (2014), 891–908. MR Zbl

[Chen and Coskun 2015] D. Chen and I. Coskun, "Extremal higher codimension cycles on moduli spaces of curves", *Proc. Lond. Math. Soc.* (3) **111**:1 (2015), 181–204. MR Zbl

[Eisenbud and Harris 1987] D. Eisenbud and J. Harris, "The Kodaira dimension of the moduli space of curves of genus ≥ 23", *Invent. Math.* **90**:2 (1987), 359–387. MR Zbl

[Faber 1999] C. Faber, "A conjectural description of the tautological ring of the moduli space of curves", pp. 109–129 in *Moduli of curves and abelian varieties*, edited by C. Faber and E. Looijenga, Aspects Math. **E33**, Vieweg, Braunschweig, 1999. MR Zbl

[Faber and Pandharipande 2005] C. Faber and R. Pandharipande, "Relative maps and tautological classes", *J. Eur. Math. Soc.* (*JEMS*) **7**:1 (2005), 13–49. MR Zbl

[Fulton and MacPherson 1994] W. Fulton and R. MacPherson, "A compactification of configuration spaces", *Ann. of Math.* (2) **139**:1 (1994), 183–225. MR Zbl

[Getzler 1998] E. Getzler, "Topological recursion relations in genus 2", pp. 73–106 in *Integrable systems and algebraic geometry* (Kobe/Kyoto, 1997), edited by M.-H. Saito et al., World Sci. Publ., River Edge, NJ, 1998. MR Zbl

[Graber and Vakil 2005] T. Graber and R. Vakil, "Relative virtual localization and vanishing of tautological classes on moduli spaces of curves", *Duke Math. J.* **130**:1 (2005), 1–37. MR Zbl

[Looijenga 1995] E. Looijenga, "On the tautological ring of $\mathcal{M}_g$", *Invent. Math.* **121**:2 (1995), 411–419. MR Zbl

[Petersen 2016] D. Petersen, "Tautological rings of spaces of pointed genus two curves of compact type", *Compos. Math.* **152**:7 (2016), 1398–1420. MR Zbl

[Petersen and Tommasi 2014] D. Petersen and O. Tommasi, "The Gorenstein conjecture fails for the tautological ring of $\overline{\mathcal{M}}_{2,n}$", *Invent. Math.* **196**:1 (2014), 139–161. MR Zbl

[Rulla 2001] W. F. Rulla, *The birational geometry of moduli space M(3) and moduli space M(2,1)*, Ph.D. thesis, University of Texas at Austin, 2001, Available at http://search.proquest.com/docview/304719193. MR

[Rulla 2006] W. F. Rulla, "Effective cones of quotients of moduli spaces of stable n-pointed curves of genus zero", *Trans. Amer. Math. Soc.* **358**:7 (2006), 3219–3237. MR Zbl

[Tarasca 2015] N. Tarasca, "Double total ramifications for curves of genus 2", *Int. Math. Res. Not.* **2015**:19 (2015), 9569–9593. MR Zbl

Communicated by David Eisenbud
Received 2015-11-09 Revised 2016-06-17 Accepted 2016-09-10

dawei.chen@bc.edu *Department of Mathematics, Boston College, Maloney Hall, 140 Commonwealth Ave, Chestnut Hill, MA 02467-3806, United States*

tarasca@math.utah.edu *Department of Mathematics, University of Utah, 155 S 1400 E, Salt Lake City, UT 84112, United States*

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1949

$\overline{\mathcal{R}}_{15}$ is of general type

Gregor Bruns

We prove that the moduli space $\overline{\mathcal{R}}_{15}$ of Prym curves of genus 15 is of general type. To this end we exhibit a virtual divisor $\overline{\mathcal{D}}_{15}$ on $\overline{\mathcal{R}}_{15}$ as the degeneracy locus of a globalized multiplication map of sections of line bundles. We then proceed to show that this locus is indeed of codimension one and calculate its class. Using this class, we can conclude that $K_{\overline{\mathcal{R}}_{15}}$ is big. This complements a 2010 result of Farkas and Ludwig: now the spaces $\overline{\mathcal{R}}_g$ are known to be of general type for $g \geq 14$.

1. Introduction

The study of Prym varieties has a long history, going back to work of Riemann, Wirtinger, Schottky and Jung in the late 19th and early 20th century. Of particular interest is the correspondence between moduli of étale double covers of curves of genus g and abelian varieties of dimension $g-1$, given by the Prym map $\mathcal{P}_g : \mathcal{R}_g \to \mathcal{A}_{g-1}$. Here we denote by $\mathcal{R}_g$ the moduli space of pairs $[C, \eta]$ where $[C] \in \mathcal{M}_g$ is a smooth genus g curve and $\eta \in \operatorname{Pic}^0(C)$ is a 2-torsion point (or equivalently an étale double cover of C).

It turns out that every principally polarized abelian variety (ppav) up to dimension 5 is a Prym variety. This generalizes the well-known fact that the general ppav of dimension at most 3 is the Jacobian of a curve. In dimension greater than 5, Prym varieties are no longer dense in the moduli space of ppavs, but their locus is still of geometric interest.

It is natural to ask for a modular compactification of $\mathcal{R}_g$ in order to study degenerations of Prym varieties and the birational geometry of their families. Explicit constructions were put forward in [Beauville 1977; Bernstein 1999] and in [Ballico et al. 2004], where the compactification is given in terms of admissible covers and Prym curves, respectively.

Much is already known about the birational geometry of $\mathcal{R}_g$. It is a rational variety for $g \leq 4$, unirational for $g \leq 7$ and uniruled for $g \leq 8$ (see [Farkas and Verra

The author was supported by the GRK 1800 of the Deutsche Forschungsgemeinschaft.
MSC2010: primary 14H10; secondary 14E08, 14H40, 14K10.
Keywords: Prym variety, Kodaira dimension, genus 15 curve, moduli space.

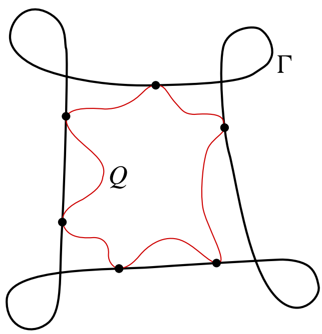
2016] for a discussion). The availability of a modular compactification has sparked interest in the Kodaira dimension of $\overline{\mathcal{R}}_g$ for higher g. Farkas and Ludwig [2010] prove that $\overline{\mathcal{R}}_g$ is of general type for $g \geq 14$ and $g \neq 15$. The Kodaira dimension of $\overline{\mathcal{R}}_{12}$ is shown to be nonnegative.

In this note we close the gap at $g = 15$.

Theorem 1.1. *The moduli space $\overline{\mathcal{R}}_{15}$ is of general type.*

We briefly outline the strategy of the proof. In order to show that the canonical class of $\overline{\mathcal{R}}_{15}$ is big, we construct an effective divisor $\mathcal{D}_{15}$ such that $K_{\overline{\mathcal{R}}_{15}}$ can be written as a positive linear combination of the Hodge class, the class of $\overline{\mathcal{D}}_{15}$ and other effective divisor classes.

To motivate the construction of $\mathcal{D}_{15}$, consider first the case of genus 6. A general curve $[C] \in \mathcal{M}_6$ possesses a finite number of complete $\mathfrak{g}^2_6$. Any such $L \in W^2_6(C)$ induces a birational map to a plane sextic curve Γ with 4 nodes. If there is a plane conic Q totally tangent to Γ, i.e., $Q \cdot \Gamma = 2D$ where D is effective of degree 6, then $\eta = \mathcal{O}_\Gamma(-1) \otimes \mathcal{O}_\Gamma(D)$ is 2-torsion.



The existence of such a totally tangent conic is equivalent to the failure of the map

$$\operatorname{Sym}^2 H^0(C, L \otimes \eta) \to \frac{H^0(C, L^{\otimes 2})}{\operatorname{Sym}^2 H^0(C, L)}$$

to be injective. It turns out that the closure of the locus of pairs $[C, \eta] \in \mathcal{R}_6$ where this injectivity fails is a divisor, i.e., the condition to possess a totally tangent conic to a plane sextic model gives a divisorial condition on $\mathcal{R}_6$. This divisor can also be identified with the closure of the ramification divisor of the Prym map $\mathcal{R}_6 \to \mathcal{A}_5$. For details, see [Farkas et al. 2014].

We generalize this condition and adapt it to genus 15. A general genus 15 curve C carries a finite number of complete $\mathfrak{g}^4_{16}$ linear series. For any such $L \in W^4_{16}(C)$ we can consider the multiplication map

$$\mu_{[C,L]} : \operatorname{Sym}^2 H^0(C, L) \to H^0(C, L^{\otimes 2}).$$

The vector spaces on the left- and right-hand sides are of dimensions 15 and 18, respectively, and the map is injective for the general pair $[C, L]$. We can make use of a torsion bundle η to get the remaining three sections:

$$\mu_{[C,\eta,L]} : \mathrm{Sym}^2 H^0(C, L) \oplus \mathrm{Sym}^2 H^0(C, L \otimes \eta) \to H^0(C, L^{\otimes 2}). \tag{1}$$

We consider the locus of Prym curves carrying a $\mathfrak{g}^4_{16}$ such that this map fails to be an isomorphism. Unlike in genus 6, such curves are not directly characterized by having a totally tangent quadric hypersurface, although on those that have, the map (1) certainly fails to be injective.

It turns out that $\mu_{[C,\eta,L]}$ is bijective for all L on the general pair $[C, \eta] \in \mathcal{R}_{15}$ and the failure locus is in codimension one. We may therefore consider the divisor

$$\mathcal{D}_{15} = \{[C, \eta] \in \mathcal{R}_{15} \mid \exists L \in W^4_{16}(C) \text{ such that } \mu_{[C,\eta,L]} \text{ is not an isomorphism}\}.$$

In order to show that (1) is indeed bijective for all η and L on a general curve C, we first construct in Section 3A a single example, using a curve that carries a theta characteristic with a large number of sections. Afterwards we prove that the moduli space $\mathfrak{G}^{4,(2)}_{16}$ of triples $[C, \eta, L]$ is irreducible, allowing us to specialize the general triple to the constructed example. More generally, we obtain the following result:

Proposition 1.2. *Assume $g \geq 3$ and the Brill–Noether number $\rho(g, r, d) = 0$. If either $r \leq 2$ or $g - d + r - 1 \leq 2$ then $\mathfrak{G}^{r,(2)}_d$ is irreducible.*

Taking the closure $\overline{\mathcal{D}}_{15}$ of $\mathcal{D}_{15}$ in an appropriate partial compactification $\overline{\mathcal{R}}^0_{15}$ of $\mathcal{R}_{15}$, we can calculate the class of $\overline{\mathcal{D}}_{15}$ using a determinantal description coming from globalizing the map (1) to a morphism of vector bundles.

Theorem 1.3. *The locus $\overline{\mathcal{D}}_{15}$ is a divisor in $\overline{\mathcal{R}}^0_{15}$ and we have the expression*

$$[\overline{\mathcal{D}}_{15}] + E \equiv 31020\Big(\frac{3127}{470}\lambda - (\delta'_0 + \delta''_0) - \frac{3487}{1880}\delta^{\mathrm{ram}}_0\Big),$$

where E is an effective class on $\overline{\mathcal{R}}^0_{15}$.

A suitable positive linear combination of $\overline{\mathcal{D}}_{15}$ and another divisor $\overline{\mathcal{D}}_{15:2}$, which was described in [Farkas and Ludwig 2010], then shows that the canonical class of $\overline{\mathcal{R}}_{15}$ is big.

To be able to calculate the class of $\overline{\mathcal{D}}_{15}$, various technical difficulties have to be overcome. In Section 2 we closely follow the setup of [Farkas 2009; Farkas and Ludwig 2010] to construct partial compactifications of suitable open subsets of $\mathcal{M}_g$ and $\mathcal{R}_g$ and to extend the moduli stacks of linear series there. We also make use of a result in [Farkas and Ludwig 2010] showing that all pluricanonical forms defined on the smooth part of $\overline{\mathcal{R}}_g$ extend to any resolution of singularities.

2. The moduli space of Prym curves

We follow the techniques and notations introduced in [Farkas and Ludwig 2010, Section 1]. First we recall the basic definitions.

A *smooth Prym curve* is a pair $[C, \eta]$ where $[C] \in \mathcal{M}_g$ is a smooth curve and $\eta \in \mathrm{Pic}^0(C) \setminus \{\mathcal{O}_C\}$ is such that $\eta^{\otimes 2} \cong \mathcal{O}_C$. To such a pair we can naturally associate an étale double cover $f : C' \to C$, where C' is given as $\mathcal{S}pec(\mathcal{O}_C \oplus \eta)$. Conversely, every étale double cover determines a unique 2-torsion bundle η on C.

We denote by $\mathcal{R}_g$ the moduli space of smooth Prym curves of genus g and by $\pi : \mathcal{R}_g \to \mathcal{M}_g$ the forgetful morphism $[C, \eta] \mapsto [C]$ of degree $2^{2g} - 1$. The corresponding morphism on stacks is étale and denoted by $\pi : \mathsf{R}_g \to \mathsf{M}_g$ as well.

2A. ***Compactifying the space of Prym curves.*** In order to compactify $\mathcal{R}_g$, we make the following definitions. We say that a smooth rational component of a nodal curve is *exceptional* if it meets the other components in exactly two points. A nodal curve is called *quasistable* if every smooth rational component meets the rest of the curve in at least two points, and the exceptional components are pairwise disjoint.

Definition 2.1. A *Prym curve* of genus g is a triple (C, η, β) consisting of a quasistable curve C of genus g, a line bundle $\eta \in \mathrm{Pic}^0(C)$ and a sheaf homomorphism $\beta : \eta^{\otimes 2} \to \mathcal{O}_C$, subject to the following conditions:

(1) For each exceptional component E of C we have $\eta|_E = \mathcal{O}_E(1)$.

(2) For each nonexceptional component the morphism β is not the zero morphism.

A *family of Prym curves* over a scheme S is a triple $(\mathcal{C} \to S, \eta, \beta)$, where $\mathcal{C} \to S$ is a flat family of quasistable curves, η is a line bundle on $\mathcal{C}$ and $\beta : \eta^{\otimes 2} \to \mathcal{O}_{\mathcal{C}}$ is a sheaf homomorphism such that for each fiber $C_s = \mathcal{C}(s)$ the triple $(C_s, \eta|_{C_s}, \beta|_{C_s})$ is a Prym curve.

If there is no danger of confusion, we usually omit the morphism β from the data to describe a Prym curve. We denote by $\overline{\mathsf{R}}_g$ the (nonsingular Deligne–Mumford) stack of Prym curves of genus g and its coarsening by $\overline{\mathcal{R}}_g$. The locus $\mathcal{R}_g$ of smooth Prym curves is contained in $\overline{\mathcal{R}}_g$ as an open subset and the forgetful map π extends to a ramified covering $\overline{\mathcal{R}}_g \to \overline{\mathcal{M}}_g$, which we also denote by π. Note that by blowing down all exceptional components of a quasistable curve we obtain a stable curve. It should also be remarked that there is a close relationship between the Prym curves discussed here and admissible covers in the sense of Beauville [1977]. For a detailed treatment of the previous statements, see [Ballico et al. 2004; Bernstein 1999].

2B. ***Boundary divisors.*** We study the boundary components of $\overline{\mathcal{R}}_g$. They lie over the boundary of $\overline{\mathcal{M}}_g$, so we can study the components lying over Δ_i for $i = 0, \dots, \lfloor g/2 \rfloor$. As is customary, we denote by δ_i the corresponding divisor classes.

The divisors Δ_i, Δ_{g-i}, $\Delta_{g:i}$ for $i \geq 1$. First consider $i \geq 1$ and let $X \in \Delta_i$ be general, i.e., $X = C \cup D$ is the union of two curves of genera i and $g-i$ meeting transversally in a single node. The line bundle $\eta \in \mathrm{Pic}^0(X)$ on the corresponding Prym curve is determined by its restrictions $\eta_C = \eta|_C$ and $\eta_D = \eta|_D$ satisfying $\eta_C^{\otimes 2} = \mathcal{O}_C$ and $\eta_D^{\otimes 2} = \mathcal{O}_D$.

Either one of η_C and η_D (but not both) can be trivial, so $\pi^*(\Delta_i)$ splits into three irreducible components

$$\pi^*(\Delta_i) = \Delta_i + \Delta_{g-i} + \Delta_{i:g-i},$$

where the general element in Δ_i is $[C \cup D, \eta_C \neq \mathcal{O}_C, \mathcal{O}_D]$, the generic point of Δ_{g-i} is of the form $[C \cup D, \mathcal{O}_C, \eta_D \neq \mathcal{O}_D]$ and the generic point of $\Delta_{i:g-i}$ looks like $[C \cup D, \eta_C \neq \mathcal{O}_C, \eta_D \neq \mathcal{O}_D]$.

The divisor Δ_0''. Now let $i = 0$. The generic point of Δ_0 in $\overline{\mathcal{M}}_g$ is a one-nodal irreducible curve C of geometric genus $g-1$. We first consider points of the form $[C, \eta]$ lying over C, i.e., without an exceptional component. Denote by $\nu : \widetilde{C} \to C$ the normalization and by p, q the preimages of the node. Then we have an exact sequence

$$0 \to \mathbb{C}^* \to \mathrm{Pic}^0(C) \xrightarrow{\nu^*} \mathrm{Pic}^0(\widetilde{C}) \to 0,$$

which restricts to

$$0 \to \mathbb{Z}/2\mathbb{Z} \to \mathrm{Pic}^0(C)[2] \xrightarrow{\nu^*} \mathrm{Pic}^0(\widetilde{C})[2] \to 0$$

on the 2-torsion part. The group $\mathbb{Z}/2\mathbb{Z}$ represents the two possible choices of gluing of the fibers at p and q for each line bundle in $\mathrm{Pic}^0(\widetilde{C})[2]$. For the case $\nu^*\eta = \mathcal{O}_{\widetilde{C}}$ there is exactly one possible choice of $\eta \neq \mathcal{O}_C$. These curves $[C, \eta]$ correspond to the classical *Wirtinger covers*

$$\widetilde{C}_1 \amalg \widetilde{C}_2/(p_1 \sim q_2, p_2 \sim q_1) \xrightarrow{2:1} \widetilde{C}/(p \sim q) = C.$$

We denote by Δ_0'' the closure of the locus of Wirtinger covers.

The divisor Δ_0'. On the other hand, there are $2^{2(g-1)} - 1$ nontrivial elements in the group $\mathrm{Pic}^0(\widetilde{C})[2]$. For each of them there are two choices of gluing, so we have a total of $2 \cdot (2^{2g-2} - 1)$ choices for η such that $\nu^*\eta \neq \mathcal{O}_{\widetilde{C}}$. We let Δ_0' be the closure of the locus of pairs $[C, \eta]$ such that $\nu^*\eta \neq \mathcal{O}_{\widetilde{C}}$.

The divisor Δ_0^{ram}. Let us turn to the case of curves of the form $[X = \widetilde{C} \cup_{p,q} E, \eta]$, where E is an exceptional component. The stabilization of such a curve is again a one-nodal curve C. Denote by β the morphism $\eta^{\otimes 2} \to \mathcal{O}_X$. Since $\eta|_E = \mathcal{O}_E(1)$, we must have $\beta_{E \setminus \{p,q\}} = 0$ and $\deg(\eta^{\otimes 2}|_{\widetilde{C}}) = -2$. It follows that $\eta^{\otimes 2}|_{\widetilde{C}} = \mathcal{O}_{\widetilde{C}}(-p-q)$. There are $2^{2(g-1)}$ choices of square roots of $\mathcal{O}_{\widetilde{C}}(-p-q)$ and each of these determines uniquely a Prym curve $[X, \eta]$ of this form. We denote the closure of the locus of such curves by Δ_0^{ram}.

As a result of the local analysis carried out for instance in [Chiodo et al. 2013], we see that π is simply ramified over Δ_0^{ram} and unramified everywhere else. This gives the relation

$$\pi^*(\delta_0) = \delta_0' + \delta_0'' + 2\delta_0^{\mathrm{ram}}.$$

2C. *The canonical class.* In order to show that $\overline{\mathcal{R}}_g$ is of general type, we need to show the canonical class is big for some desingularization $\widehat{\mathcal{R}}_g$ of $\overline{\mathcal{R}}_g$. Using the following extension result we see that all pluricanonical differentials on the smooth part of $\overline{\mathcal{R}}_g$ extend to $\widehat{\mathcal{R}}_g$.

Theorem 2.2 [Farkas and Ludwig 2010, Theorem 6.1]. *Let $g \geq 4$ and $\widehat{\mathcal{R}}_g \to \overline{\mathcal{R}}_g$ be any desingularization. Then every pluricanonical form defined on the smooth locus $\overline{\mathcal{R}}_g^{\mathrm{reg}}$ of $\overline{\mathcal{R}}_g$ extends holomorphically to $\widehat{\mathcal{R}}_g$; that is, for all integers $l \geq 0$ we have isomorphisms*

$$H^0\big(\overline{\mathcal{R}}_g^{\mathrm{reg}}, K_{\overline{\mathcal{R}}_g}^{\otimes l}\big) \cong H^0\big(\widehat{\mathcal{R}}_g, K_{\widehat{\mathcal{R}}_g}^{\otimes l}\big).$$

Furthermore, one has the expression

$$K_{\overline{\mathcal{R}}_g} = 13\lambda - 2(\delta_0' + \delta_0'') - 3\delta_0^{\mathrm{ram}} - 2\sum_{i=1}^{\lfloor g/2 \rfloor} (\delta_i + \delta_{g-i} + \delta_{i:g-i}) - (\delta_1 + \delta_{g-1} + \delta_{1:g-1})$$

for the canonical class $K_{\overline{\mathcal{R}}_g}$ in terms of the divisor classes introduced before (see for example [Farkas and Ludwig 2010, Theorem 1.5]). Here we have abused notation and set $\lambda = \pi^*(\lambda)$, the pullback of the Hodge class from $\overline{\mathcal{M}}_g$. It is therefore enough to exhibit an effective divisor D of the form

$$D = a\lambda - (b_0'\delta_0' + b_0''\delta_0'') - b_0^{\mathrm{ram}}\delta_0^{\mathrm{ram}} - \sum_{i=1}^{\lfloor g/2 \rfloor} (b_i\delta_i + b_{g-i}\delta_{g-i} + b_{i:g-i}\delta_{i:g-i})$$

such that

$$\frac{a}{\gamma} < \frac{13}{2} \quad \text{for all } \gamma \in \{b_0', b_0''\} \cup \{b_i, b_{g-i}, b_{i:g-i} \mid i = 1, \dots, \lfloor g/2 \rfloor\}$$

as well as

$$\frac{a}{\gamma} < \frac{13}{3} \quad \text{for all } \gamma \in \{b_0^{\mathrm{ram}}, b_1, b_{g-1}, b_{1:g-1}\}.$$

Remark 2.3. Actually, the situation turns out to be simpler. Proposition 1.9 of [Farkas and Ludwig 2010] shows that for $g \leq 23$ it is enough to consider the coefficients of λ, δ_0', δ_0'' and δ_0^{ram}. If they satisfy the inequalities above, the other boundary divisor coefficients are automatically suitably bounded. We will make full use of the fact that we do not have to consider singular curves of compact type.

2D. ***The universal Prym curve.*** Since we are only concerned with the boundary divisors Δ_0', Δ_0'' and Δ_0^{ram}, we partially compactify $\mathcal{M}_g$ by adding the open sublocus $\widetilde{\Delta}_0 \subset \Delta_0$ of one-nodal irreducible curves. Set

$$\widetilde{\mathcal{M}}_g = \mathcal{M}_g \cup \widetilde{\Delta}_0$$

and let $\widetilde{\mathcal{R}}_g = \pi^{-1}(\widetilde{\mathcal{M}}_g)$. We also set

$$\mathcal{Z} = \widetilde{\mathsf{R}}_g \times_{\widetilde{\mathsf{M}}_g} \widetilde{\mathsf{M}}_{g,1}.$$

This is not yet the universal Prym curve over $\widetilde{\mathsf{R}}_g$, since the points on exceptional components of curves in Δ_0^{ram} are not present. We have to blow up $\mathcal{Z}$ along the locus V of points

$$(X, \eta_X, p = q) \in \Delta_0^{\mathrm{ram}} \times_{\widetilde{\mathsf{M}}_g} \widetilde{\mathsf{M}}_{g,1}, \quad X = C \cup_{\{p,q\}} E \to C/p \sim q, \quad \eta_E = \mathcal{O}_E(1).$$

Set $\mathcal{X}_g = \mathrm{Bl}_V(\mathcal{Z})$ and let $f : \mathcal{X}_g \to \widetilde{\mathsf{R}}_g$ be the induced universal family of Prym curves. The family $\mathcal{X}_g$ comes equipped with a Poincaré bundle $\mathcal{P}$ such that $\mathcal{P}|_{f^{-1}([X,\eta,\beta])} = \eta$. We need the following result from [Farkas and Ludwig 2010, Proposition 1.6]:

Lemma 2.4. *In* $\mathrm{Pic}(\widetilde{\mathcal{R}}_g)$ *we have* $f_*(c_1^2(\mathcal{P})) = -\delta_0^{\mathrm{ram}}/2$ *and* $f_*(c_1(\mathcal{P})c_1(\omega_\chi)) = 0$.

2E. ***Moduli spaces of linear series over the Prym moduli space.*** To compute the classes of divisors on $\overline{\mathcal{R}}_g$, a viable method is to give them a determinantal description, i.e., exhibit them as degeneracy loci of morphisms of vector bundles. To obtain these vector bundles, we consider the stack $\mathfrak{G}_d^{r,(2)}$ parametrizing triples $[C, \eta, L]$ where $[C, \eta] \in \mathcal{R}_g$ and $L \in G_d^r(C)$. Note that in the case $\rho(g, r, d) = 0$ in which we are interested, the forgetful map $\mathfrak{G}_d^{r,(2)} \to \mathcal{R}_g$ is a generically finite cover of degree

$$N = g! \,\frac{1!\,2! \cdots r!}{(g-d+r)! \cdots (g-d+2r)!}.$$

We want to first restrict this construction to an open subset of $\mathcal{R}_g$ such that various pushforwards of the Poincaré bundles on the universal curve are indeed vector bundles on $\mathfrak{G}_d^{r,(2)}$. Then we shall extend the stack over a suitable partial compactification to be able to also determine the behavior on the boundary.

Let M_g^0 be the open substack of M_g classifying curves C with $W_d^{r+1}(C) = \varnothing$ and $W_{d-1}^r(C) = \varnothing$. A general such curve indeed has a finite amount of $\mathfrak{g}_d^r$ linear series and all of them are very ample. Observe that both

$$\rho(g, r+1, d) = -(g - d + 2(r+1)) \le -2, \quad \rho(g, r, d-1) = -(r+1) \le -2,$$

so the codimension of the complement of $\mathcal{M}_g^0$ in $\mathcal{M}_g$ is at least 2, for instance by results in [Eisenbud and Harris 1989]. Therefore, restricting to $\mathcal{M}_g^0$ does not change divisor class calculations.

To partially compactify M_g^0, add the locus Δ_0^0 of Brill–Noether general irreducible one-nodal curves, i.e., $[C/p \sim q]$ with $[C] \in \mathcal{M}_{g-1}$ satisfying the Brill–Noether theorem. Denote by $\overline{\mathsf{M}}_g^0 = \mathsf{M}_g^0 \cup \Delta_0^0$ the resulting partial compactification. Over $\overline{\mathsf{M}}_g^0$ we consider the stack of pairs $[C, L]$ where $L \in G_d^r(C)$. We denote this stack by $\overline{\mathfrak{G}}_d^r$. Pulling back the universal curve $\overline{\mathsf{M}}_{g,1}^0$ to $\overline{\mathfrak{G}}_d^r$, we get a universal family

$$f_d^r : \overline{\mathfrak{C}}_d^r = \overline{\mathfrak{G}}_d^r \times_{\overline{\mathsf{M}}_g^0} \overline{\mathsf{M}}_{g,1}^0 \to \overline{\mathfrak{G}}_d^r$$

and we choose a Poincaré bundle, i.e., an $\mathcal{L} \in \mathrm{Pic}(\overline{\mathfrak{C}}_d^r)$ such that $\mathcal{L}|_{(f_d^r)^{-1}([C,L])} = L$ for every $[C, L] \in \overline{\mathfrak{G}}_d^r$.

We are now ready to pull these constructions back to Prym curves. We let $\overline{\mathsf{R}}_g^0 = \pi^{-1}(\overline{\mathsf{M}}_g^0)$ and

$$\sigma : \overline{\mathfrak{G}}_d^{r,(2)} = \overline{\mathfrak{G}}_d^r \times_{\overline{\mathsf{M}}_g^0} \overline{\mathsf{R}}_g^0 \to \overline{\mathsf{R}}_g^0$$

be the stack parametrizing triples $[C, \eta, L]$ for $[C, \eta] \in \overline{\mathsf{R}}_g^0$ and $L \in W_d^r(C)$. We also have the universal curve

$$\chi : \overline{\mathfrak{C}}_d^{r,(2)} = \mathcal{X}_g \times_{\overline{\mathsf{R}}_g^0} \overline{\mathfrak{G}}_d^{r,(2)} \to \overline{\mathfrak{G}}_d^{r,(2)}.$$

By pulling back from $\overline{\mathsf{R}}_g^0$ and $\overline{\mathfrak{G}}_d^{r,(2)}$, respectively, this comes equipped with two Poincaré bundles $\mathcal{P}$ and $\mathcal{L}$. We can also use σ to pull back the boundary classes Δ_0', Δ_0'' and Δ_0^{ram} from $\overline{\mathsf{R}}_g^0$ to $\overline{\mathfrak{G}}_d^{r,(2)}$. Slightly abusing notation, the pullbacks will be denoted by the same symbols.

3. A new divisor on $\overline{\mathcal{R}}_{15}$

As before, we denote by $\chi : \overline{\mathfrak{C}}_{16}^{4,(2)} \to \overline{\mathfrak{G}}_{16}^{4,(2)}$ the universal curve and let $\mathcal{L}$ be a Poincaré bundle on $\overline{\mathfrak{C}}_{16}^{4,(2)}$. Furthermore, let ω_χ be the relative dualizing sheaf of χ and $\sigma : \overline{\mathfrak{G}}_{16}^{4,(2)} \to \overline{\mathcal{R}}_{15}^0$ be the generically finite cover of degree $N = 6006$.

By construction of our moduli stacks and Grauert's theorem, the pushforwards of $\mathcal{L}$ and $\mathcal{L}^{\otimes 2}$ by χ are vector bundles on $\overline{\mathfrak{G}}_{16}^{4,(2)}$ of ranks 5 and 18, respectively. The sheaf $\chi_*(\mathcal{L} \otimes \mathcal{P})$ is possibly not a vector bundle, but at least it is torsion-free. By excluding the subvariety (of codimension at least two) where it fails to be locally free we can assume it is in fact a vector bundle of rank 2. Divisor class calculations will not be affected.

We may then consider the following morphism of vector bundles of the same rank:

$$\phi : \mathrm{Sym}^2 \chi_*(\mathcal{L}) \oplus \mathrm{Sym}^2 \chi_*(\mathcal{L} \otimes \mathcal{P}) \to \chi_*(\mathcal{L}^{\otimes 2}).$$

On the fiber over the class of a triple $[C, \eta, L]$ it is given by the multiplication map of sections

$$\mu_{[C,\eta,L]} : \mathrm{Sym}^2 H^0(C, L) \oplus \mathrm{Sym}^2 H^0(C, L \otimes \eta) \to H^0(C, L^{\otimes 2}). \qquad (2)$$

The closure of the locus

$$\mathcal{D}_{15} = \{[C, \eta] \in \mathcal{R}_{15} \mid \exists L \in W^4_{16}(C) \text{ such that } \mu_{[C,\eta,L]} \text{ is not an isomorphism}\}$$

therefore has a determinantal description as the pushforward of the first degeneracy locus of the map ϕ. Its expected codimension is one and we obtain a virtual divisor. Note that while the vector bundles involved in defining ϕ clearly depend on the choice of the Poincaré bundle $\mathcal{L}$, the resulting morphism ϕ does not (cf. the remark before Theorem 2.1 in [Farkas 2009]).

3A. ***Proof of divisoriality of $\mathcal{D}_{15}$.*** We now prove that $\overline{\mathcal{D}}_{15}$ is a genuine divisor, that is, $\mu_{[C,\eta,L]}$ is an isomorphism for every $L \in W^4_{16}(C)$ on the general Prym curve $[C, \eta]$. We will prove in Section 3B that $\mathfrak{G}^{4,(2)}_{16}$ over the whole space $\mathcal{R}_{15}$ is irreducible. Hence it will be enough to exhibit a single smooth curve C and two line bundles $L \in W^4_{16}(C)$ and $\eta \in \mathrm{Pic}^0(C)[2]$ such that the multiplication map (2) is bijective. We can then specialize the general element of $\mathfrak{G}^{4,(2)}_{16}$ to this particular example and conclude by semicontinuity.

We start with a smooth nonhyperelliptic curve $C \in \mathcal{M}_{15}$ possessing a theta characteristic ϑ with an exactly 5-dimensional space of global sections, i.e., $|\vartheta| \in G^4_{14}(C)$ and $\vartheta^{\otimes 2} \cong \omega_C$. In order to construct an L such that $\mu_{[C,\eta,L]}$ is bijective, C should in fact be half-canonically embedded by ϑ such that the image does not lie on any quadric hypersurface in $\mathbb{P}^4$.

Explicit models of such curves can be obtained as hyperplane sections of smooth canonical surfaces $S \subseteq \mathbb{P}^5$ with $p_g = 6$ and $K^2_S = 14$. To construct such a surface, one can employ the method described by Catanese [1997].

Lemma 3.1. *There exists a smooth projective surface S of general type with invariants $(K^2_S, p_g, q) = (14, 6, 0)$, canonically embedded in $\mathbb{P}^5$, not lying on any quadric hypersurface.*

Proof. The surfaces S arise from Pfaffian resolutions

$$0 \to \mathcal{O}_{\mathbb{P}^5}(-7) \to \mathcal{O}_{\mathbb{P}^5}(-4)^{\oplus 7} \xrightarrow{\alpha} \mathcal{O}_{\mathbb{P}^5}(-3)^{\oplus 7} \xrightarrow{p} \mathcal{I}_S \to 0 \tag{3}$$

of the ideal sheaf $\mathcal{I}_S$, where α is a 7×7 antisymmetric matrix with linear entries and p is the map given by the Pfaffians of 6×6 principal submatrices of α.

Using the projective resolution (3) and Serre duality for Ext sheaves, we see that S is canonically embedded. We also see that S is a regular surface (i.e., $q = 0$) and $p_g = 6$, which combines to give $\chi(\mathcal{O}_S) = 7$. Again using (3), the Hilbert polynomial of $\mathcal{O}_S$ is $P_S(t) = 7t^2 - 7t + 7$, which tells us $\deg(S) = 14$, and because S is canonically embedded we have $K^2_S = 14$. □

A general hyperplane section $C = H \cap S$ of S has, by the adjunction formula,

$$\omega_C \cong (\mathcal{O}_S(1) \otimes \omega_S)|_C \cong \omega_S^{\otimes 2}|_C, \quad 2g - 2 = 2K_S \cdot K_S = 28,$$

so $C \hookrightarrow \mathbb{P}^4$ is half-canonically embedded of degree 14 and genus 15. Using the exact sequence

$$0 \to \mathcal{I}_S(2) \to \mathcal{O}_{\mathbb{P}^5}(2) \to \mathcal{O}_S(2) \to 0$$

and $h^0(S, \omega_S^{\otimes 2}) = 21$ by Riemann–Roch, we get $H^0(\mathbb{P}^5, \mathcal{I}_S(2)) = 0$, so S does not lie on a quadric hypersurface of $\mathbb{P}^5$. The same then applies for C in $\mathbb{P}^4$. A moduli count shows that hyperplane sections of such S form a 32-dimensional family.

Remark 3.2. This is not the only way in which such curves arise. Iliev and Markushevich [2000] also obtain a 32-dimensional family (i.e., an irreducible component of the expected dimension of the locus $\mathcal{T}^4_{15}$ of curves of genus 15 having a theta-characteristic with 5 independent global sections) as vanishing loci of sections of rank 2 ACM bundles on quartic 3-folds in $\mathbb{P}^4$.

Lemma 3.3. *For a half-canonically embedded curve C in $\mathbb{P}^4$ not lying on a quadric hypersurface, the multiplication map $\mu_{[C,\eta,L]}$ is bijective.*

Proof. Set $\vartheta = \mathcal{O}_C(1)$. Of course $\mathcal{O}_C(2) = \omega_C$. The fact that C does not lie on a quadric hypersurface is equivalent to the bijectivity of the multiplication map

$$\mu_\vartheta : \operatorname{Sym}^2 H^0(C, \vartheta) \to H^0(C, \omega_C).$$

We now choose any closed point $x \in C$. Using that ϑ is very ample we get

$$h^0(C, \vartheta(-2x)) = h^0(C, \vartheta) - 2.$$

By Serre duality this implies $h^0(C, \vartheta(2x)) = h^0(C, \vartheta)$. Let $L = \vartheta(2x)$, so L is a complete $\mathfrak{g}^4_{16}$ and $2x$ is contained in the base locus of L. In particular, we have $H^0(C, L) \cong H^0(C, \vartheta)$ and $|L| = |\vartheta| + 2x$. Taking symmetric powers, we get

$$\operatorname{Sym}^2 H^0(C, L) \cong \operatorname{Sym}^2 H^0(C, \vartheta) \cong H^0(C, \omega_C).$$

The space $H^0(C, L^{\otimes 2})$ is 18-dimensional, and it decomposes via the inclusion $H^0(C, \vartheta^{\otimes 2}) \hookrightarrow H^0(C, L^{\otimes 2})$ as

$$H^0(C, L^{\otimes 2}) \cong H^0(C, \omega_C) \oplus V \cong \operatorname{Sym}^2 H^0(C, L) \oplus V,$$

where $\dim V = 3$. The sections in $\operatorname{Sym}^2 H^0(C, L)$ vanish to orders at least 4 at x. By Riemann–Roch, the space $H^0(C, L^{\otimes 2})$ does contain sections vanishing to orders 0, 1 and 2 at x. By the previous analysis, they must span V.

Choose a two-torsion bundle $\eta \in \operatorname{Pic}^0(C)[2]$ such that $H^0(C, \vartheta \otimes \eta) = 0$. Since $\operatorname{Pic}^0(C)[2]$ acts transitively on the theta-characteristics, such an η always exists by a result of Mumford [1966]. Then we have

$$h^0(C, L \otimes \eta) = h^0(C, \vartheta(2x) \otimes \eta) \le h^0(C, \vartheta \otimes \eta) + 2 = 2.$$

By Riemann–Roch we must in fact have $h^0(C, L\otimes\eta) = 2$. By construction,

$$H^0(C, (L\otimes\eta)(-2x)) = H^0(C, \vartheta\otimes\eta) = 0,$$

so the two sections of $L\otimes\eta$ vanish to orders 0 and 1 at x. We conclude that the map

$$\operatorname{Sym}^2 H^0(C, L\otimes\eta) \to H^0(C, L^{\otimes 2})$$

is injective and its image is precisely V. □

3B. ***Irreducibility of some spaces of linear series.*** We now want to prove the irreducibility of $\mathfrak{G}^{4,(2)}_{16}$, i.e., the moduli space of triples $[C, \eta, L]$ where $[C, \eta] \in \mathcal{R}_{15}$ and $L \in W^4_{16}(C)$. This will show that for the general triple $[C, \eta, L]$, the map $\mu_{[C,\eta,L]}$ is an isomorphism. Notice that the pair $[C, L]$ constructed in Section 3A is *not* Petri general, so we need more than the existence of a unique component of $\mathfrak{G}^{4,(2)}_{16}$ dominating $\mathcal{M}_{15}$. Nonetheless, this fact is what we are going to establish first in greater generality:

Proposition 3.4. *Let $g \geq 3$ and $\rho(g, r, d) = 0$. Then there is a unique irreducible component of $\mathfrak{G}^{r,(2)}_d$ dominating $\mathcal{M}_g$, i.e., containing the Petri general triple $[C, \eta, L]$.*

Proof. If $r = g-1$, the only $\mathfrak{g}^r_d$ on a curve is the canonical bundle, so $\mathfrak{G}^{r,(2)}_d \cong \mathcal{R}_g$ is irreducible. Otherwise, set $k = g-d+r+1 \geq 3$. We recall that the locus of Petri general pairs $[C, L]$ is a connected smooth open subset U of one irreducible component of $\mathfrak{G}^r_d$ [Eisenbud and Harris 1987]. The restriction of $\mathfrak{G}^{r,(2)}_d$ to the preimage $U^{(2)}$ of U is smooth, so in order to show $U^{(2)}$ is irreducible we only have to show it is connected.

Take a general k-gonal curve $[D, A]$. We then have $h^0(D, A^{\otimes l}) = l+1$ for all $l \leq r+1$ (see [Ballico 1989]). So there is a rational map

$$\Psi : \mathfrak{G}^{1,(2)}_k \dashrightarrow \mathfrak{G}^{r,(2)}_d$$

defined by $[D, \eta, A] \mapsto [D, \eta, A^{\otimes r}]$. We claim $A^{\otimes r}$ is Petri general, i.e., the map

$$\mu_{A^{\otimes r}} : H^0(D, A^{\otimes r}) \otimes H^0(D, \omega_D \otimes A^{\otimes(-r)}) \to H^0(D, \omega_D)$$

is injective. The aforementioned result of Ballico implies

$$h^0(D, \omega_D \otimes A^{\otimes(-j)}) = (k-1)(r+1-j)$$

for all $j \leq r+1$. Note also that $g = (k-1)(r+1)$. By counting dimensions we find that $\mu_{A^{\otimes r}}$ is injective if and only if it is surjective.

We write down the beginning of the long exact sequence coming from the base point free pencil trick:

$$0 \to H^0(\omega_D \otimes A^{\otimes(-j-1)}) \to H^0(A) \otimes H^0(\omega_D \otimes A^{\otimes(-j)}) \to H^0(\omega_D \otimes A^{\otimes(-j+1)}).$$

Comparing dimensions we find that the map on the right is surjective for all $j \leq r$. Now note that $h^0(D, A^{\otimes r}) = r+1$ is equivalent to $H^0(D, A^{\otimes r}) \cong \mathrm{Sym}^r H^0(D, A)$. The chain of surjective maps

$$H^0(A)^{\otimes r} \otimes H^0(\omega_D \otimes A^{\otimes(-r)}) \twoheadrightarrow H^0(A)^{\otimes(r-1)} \otimes H^0(\omega_D \otimes A^{\otimes(-r+1)}) \twoheadrightarrow \cdots$$
$$\cdots \twoheadrightarrow H^0(A) \otimes H^0(\omega_D \otimes A^{-1})$$

then implies that the Petri map

$$\mu_{A^{\otimes r}} : \mathrm{Sym}^r H^0(D, A) \otimes H^0(D, \omega_D \otimes A^{\otimes(-r)}) \to H^0(D, \omega_D)$$

is surjective as well. So $[D, \eta, A^{\otimes r}]$ lies in $U^{(2)}$.

In [Biggers and Fried 1986] it is shown that the Hurwitz space $\mathfrak{G}_k^{1,(2)}$ is irreducible for $k \geq 3$. Hence Ψ maps to the smooth locus of a unique component Z of $\mathfrak{G}_d^{r,(2)}$ and its image is an irreducible subset consisting generically of Petri general curves. Since the image is closed under monodromy of 2-torsion, it follows that $U^{(2)}$ must be connected. □

We employ this result to prove irreducibility of $\mathfrak{G}_d^{r,(2)}$ under special circumstances:

Corollary 3.5. *Assume $g \geq 3$ and $\rho(g, r, d) = 0$. If $r \leq 2$ or $r' = g-d+r-1 \leq 2$, then $\mathfrak{G}_d^{r,(2)}$ is irreducible.*

Proof. Note that the Serre dual of a $\mathfrak{g}_d^r$ is a $\mathfrak{g}_{2g-2-d}^{r'}$, so the space $\mathfrak{G}_d^{r,(2)}$ is irreducible if and only if $\mathfrak{G}_{2g-2-d}^{r',(2)}$ is. As mentioned above, if $r = 0$ or, equivalently, $r' = g-1$, the unique $\mathfrak{g}_d^r$ on a curve is its canonical bundle, so $\mathfrak{G}_d^{r,(2)} \cong \mathcal{R}_g$ is irreducible. The case $r = 1$ is just the aforementioned result by Biggers and Fried [1986] about the irreducibility of Hurwitz spaces.

In the remaining case $r = 2$ a general $\mathfrak{g}_d^2$ maps C birationally to a nodal curve in $\mathbb{P}^2$. Thus we get a dominant rational map

$$V^{d,g} \dashrightarrow \mathfrak{G}_d^2$$

from the Severi variety $V^{d,g}$ of irreducible plane curves of degree d and arithmetic genus g. The Severi varieties are irreducible, as proven in [Harris 1986], so $\mathfrak{G}_d^2$ is irreducible as well.

Étale maps preserve dimension, so all components of $\mathfrak{G}_d^{2,(2)}$ have dimension $3g-3+\rho(g,r,d) = 3g-3$. Each component is generically smooth, which implies that the general element has injective Petri map. But by Proposition 3.4 there is only one such component. □

In particular, $\mathfrak{G}^{4,(2)}_{16}$ is irreducible. We may therefore specialize a general triple $[C, \eta, L] \in \mathfrak{G}^{4,(2)}_{16}$ to the previously constructed explicit example. This proves that the locus $\overline{\mathcal{D}}_{15}$ is a genuine divisor. We proceed to calculate its class.

3C. *Calculation of the divisor class.* Recall that we are considering the morphism

$$\phi : \mathrm{Sym}^2 \chi_*(\mathcal{L}) \oplus \mathrm{Sym}^2 \chi_*(\mathcal{L} \otimes \mathcal{P}) \to \chi_*(\mathcal{L}^{\otimes 2})$$

between vector bundles of the same rank. To calculate the Chern classes of these bundles we will employ Grothendieck–Riemann–Roch. For this we study the contribution coming from $R^1\chi_*(\mathcal{L} \otimes \mathcal{P})$.

Lemma 3.6. *Let $[C, \eta] \in \Delta_0''$ be general and $L \in W^4_{16}(C)$. Then $h^0(C, L \otimes \eta) = 4$.*

Proof. Let $\nu : \widetilde{C} \to C$ be the normalization of C and x be the node. Then $\nu^*\eta = \mathcal{O}_{\widetilde{C}}$ and $\nu^*L \in W^4_{16}(\widetilde{C})$, since $\widetilde{C}$ is Brill–Noether general. From the exact sequence

$$0 \to \mathcal{O}_C \to \nu_*\mathcal{O}_{\widetilde{C}} \xrightarrow{e} \mathbb{C}_x \to 0$$

we get

$$0 \to L \otimes \eta \to \nu_*\nu^*L \xrightarrow{e'} L \otimes \eta|_x \to 0,$$

and by the long exact sequence in cohomology we obtain

$$0 \to H^0(C, L \otimes \eta) \to H^0(\widetilde{C}, \nu^*L) \xrightarrow{H^0(e')} \mathbb{C}.$$

Now $H^0(e)$ is the zero map, hence $H^0(e')$ must be nonzero and we get

$$h^0(C, L \otimes \eta) = h^0(\widetilde{C}, \nu^*L) - 1 = 4. \qquad \square$$

This implies that the dimension of $h^0(C, L \otimes \eta)$ jumps by two on the boundary component Δ_0''. Hence $R^1\chi_*(\mathcal{L} \otimes \mathcal{P})$ is supported at least on Δ_0'', and there it is of rank 2.

Remark 3.7. In fact, Δ_0'' seems to be the only divisor where $R^1\chi_*(\mathcal{L} \otimes \mathcal{P})$ is supported. Since a proof of this would take long, and is not strictly necessary to achieve the goal of the article, we do not assume this fact here and will discuss it in greater generality in future work.

Denote $\mathfrak{d} = c_1(R^1\chi_*(\mathcal{L} \otimes \mathcal{P}))$.

Proposition 3.8. *The pushforward to $\overline{\mathcal{R}}^0_{15}$ of the class of the degeneracy locus $Z_1(\phi)$ is*

$$[\overline{\mathcal{D}}_{15}]^{\mathrm{virt}} \equiv 31020\Big(\frac{3127}{470}\lambda - (\delta_0' + \delta_0'') - \frac{3487}{1880}\delta_0^{\mathrm{ram}}\Big) - 3\sigma_*(\mathfrak{d}),$$

and $[\overline{\mathcal{D}}_{15}]^{\mathrm{virt}} - n[\overline{\mathcal{D}}_{15}]$ is an effective class supported on the boundary for some $n \geq 1$.

Proof. We introduce the following classes in $A^1(\overline{\mathfrak{G}}^{4,(2)}_{16})$:

$$\mathfrak{a} = \chi_*(c_1^2(\mathcal{L})), \quad \mathfrak{b} = \chi_*(c_1(\mathcal{L}) \cdot c_1(\omega_\chi)), \quad \mathfrak{c} = c_1(\chi_*(\mathcal{L})).$$

By Porteous' formula, the class of the first degeneracy locus $Z_1(\phi)$ of ϕ is given by

$$Z_1(\phi) = c_1(\chi_*(\mathcal{L}^{\otimes 2})) - c_1(\operatorname{Sym}^2 \chi_*\mathcal{L}) - c_1(\operatorname{Sym}^2 \chi_*(\mathcal{L} \otimes \mathcal{P})).$$

For a vector bundle $\mathcal{G}$ we have the elementary fact

$$c_1(\operatorname{Sym}^2 \mathcal{G}) = (\operatorname{rk}(\mathcal{G}) + 1)c_1(\mathcal{G}).$$

Furthermore, for every $[C, \eta] \in \overline{\mathcal{R}}^0_g$ and every $L \in W^4_{16}(C)$ we have $H^1(C, L^{\otimes 2}) = 0$, so $R^1\chi_*(\mathcal{L}^{\otimes 2}) = 0$. We can then apply Grothendieck–Riemann–Roch and express everything in terms of the classes $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ and $\mathfrak{d}$. For instance we have

$$\begin{aligned} c_1(\chi_*(\mathcal{L}^{\otimes 2})) &= \Big[\chi_*\big(1 + c_1(\mathcal{L}^{\otimes 2}) + \tfrac{1}{2}c_1^2(\mathcal{L}^{\otimes 2})\big) \\ &\qquad \cdot \big(1 - \tfrac{1}{2}c_1(\omega_\chi) + \tfrac{1}{12}(c_1^2(\omega_\chi) + c_2(\Omega_\chi))\big)\Big]_1 \\ &= \lambda + 2\mathfrak{a} - \mathfrak{b}, \end{aligned}$$

where $[-]_1$ denotes the degree 1 part of an expression. We have used Mumford's formula to calculate $\chi_*(c_1^2(\omega_\chi) + c_2(\Omega_\chi)) = 12\lambda$. Similarly, also using Lemma 2.4, we find

$$c_1(\chi_*(\mathcal{L} \otimes \mathcal{P})) = \lambda + \tfrac{1}{2}\mathfrak{a} - \tfrac{1}{2}\mathfrak{b} - \tfrac{1}{4}\delta_0^{\mathrm{ram}} + \mathfrak{d}.$$

Using the results of [Farkas 2009], in particular Lemmata 2.6 and 2.13 as well as Proposition 2.12, we can calculate the pushforwards of $\mathfrak{a}$, $\mathfrak{b}$ and $\mathfrak{c}$ by σ:

$$\begin{aligned} \sigma_*(\mathfrak{a}) &= -146784\lambda + 20856(\delta_0' + \delta_0'') + 41712\delta_0^{\mathrm{ram}}, \\ \sigma_*(\mathfrak{b}) &= 4224 + 264(\delta_0' + \delta_0'') + 528\delta_0^{\mathrm{ram}}, \\ \sigma_*(\mathfrak{c}) &= -48279 + 6930(\delta_0' + \delta_0'') + 13860\delta_0^{\mathrm{ram}}, \end{aligned}$$

and of course $\sigma_*(\lambda) = N\lambda$, $\sigma_*(\delta_0^{\mathrm{ram}}) = N\delta_0^{\mathrm{ram}}$, where $N = 6006$ is the degree of σ. Putting everything together, we obtain the result. The difference between $[\overline{\mathcal{D}}_{15}]^{\mathrm{virt}}$ and $[\overline{\mathcal{D}}_{15}]$ arises from the boundary components where ϕ is degenerate. □

Theorem 3.9. $\overline{\mathcal{R}}_{15}$ *is of general type.*

Proof. The contribution of $\sigma_*(\mathfrak{d})$ to $[\overline{\mathcal{D}}_{15}]$ only improves the ratio between the coefficients of λ and the boundary components. The same goes for the boundary components where ϕ is degenerate. Hence we may as well work with the class $[\overline{\mathcal{D}}_{15}]^{\mathrm{virt}} + 3\sigma_*(\mathfrak{d})$. Then we take an appropriate linear combination of $\overline{\mathcal{D}}_{15}$ and the divisor $\overline{\mathcal{D}}_{15:2}$ from [Farkas and Ludwig 2010] having class

$$[\overline{\mathcal{D}}_{15:2}] = 5808\lambda - 924(\delta_0' + \delta_0'') - 990\delta_0^{\mathrm{ram}}$$
$$= 924\big(\tfrac{44}{7}\lambda - (\delta_0' + \delta_0'') - \tfrac{15}{14}\delta_0^{\mathrm{ram}}\big).$$

For instance we have

$$\beta\overline{\mathcal{D}}_{15:2} + \gamma\overline{\mathcal{D}}_{15} = \epsilon\lambda - 2(\delta_0' + \delta_0'') - 3\delta_0^{\mathrm{ram}},$$

where

$$\beta = \frac{667}{680394}, \quad \gamma = \frac{4}{113399}, \quad \epsilon = \frac{10288}{793}.$$

Here $\epsilon < 13$, hence the canonical class is big. □

Remark 3.10. The map

$$\mathrm{Sym}^2 H^0(C, L\otimes\eta) \to H^0(C, L^{\otimes 2})/\,\mathrm{Sym}^2 H^0(C, L)$$

is identically zero along the boundary component Δ_0''. Hence the morphism ϕ is degenerate with order 3 along Δ_0''. It follows that we can subtract $3\delta_0''$ from $Z_1(\phi)$ and still obtain an effective class.

Acknowledgements

This project was generously supported by the Graduiertenkolleg 1800 of the Deutsche Forschungsgemeinschaft and the Berlin Mathematical School. I would like to thank my PhD advisor Gavril Farkas for his suggestion to study the circle of problems to which this question belongs, as well as for many helpful discussions. I also thank the anonymous referee for careful reading and helpful comments.

References

[Ballico 1989] E. Ballico, "A remark on linear series on general k-gonal curves", *Boll. Un. Mat. Ital. A* (7) **3**:2 (1989), 195–197. MR 1008591 Zbl 0702.14026

[Ballico et al. 2004] E. Ballico, C. Casagrande, and C. Fontanari, "Moduli of Prym curves", *Doc. Math.* **9** (2004), 265–281. MR 2117416 Zbl 1072.14029

[Beauville 1977] A. Beauville, "Prym varieties and the Schottky problem", *Invent. Math.* **41**:2 (1977), 149–196. MR 0572974 Zbl 0333.14013

[Bernstein 1999] M. Bernstein, *Moduli of curves with level structure*, Ph.D. thesis, Harvard University, 1999, Available at http://search.proquest.com/docview/304503454.

[Biggers and Fried 1986] R. Biggers and M. Fried, "Irreducibility of moduli spaces of cyclic unramified covers of genus g curves", *Trans. Amer. Math. Soc.* **295**:1 (1986), 59–70. MR 831188 Zbl 0601.14022

[Catanese 1997] F. Catanese, "Homological algebra and algebraic surfaces", pp. 3–56 in *Algebraic geometry* (Santa Cruz, 1995), edited by J. Kollár et al., Proc. Sympos. Pure Math. **62**, American Mathematical Society, Providence, RI, 1997. MR 1492517 Zbl 0937.14024

[Chiodo et al. 2013] A. Chiodo, D. Eisenbud, G. Farkas, and F.-O. Schreyer, "Syzygies of torsion bundles and the geometry of the level ℓ modular variety over $\overline{\mathcal{M}}_g$", *Invent. Math.* **194**:1 (2013), 73–118. MR 3103256 Zbl 1284.14006

[Eisenbud and Harris 1987] D. Eisenbud and J. Harris, "Irreducibility and monodromy of some families of linear series", *Ann. Sci. École Norm. Sup.* (4) **20**:1 (1987), 65–87. MR 892142 Zbl 0625.14013

[Eisenbud and Harris 1989] D. Eisenbud and J. Harris, "Irreducibility of some families of linear series with Brill–Noether number -1", *Ann. Sci. École Norm. Sup.* (4) **22**:1 (1989), 33–53. MR 985853 Zbl 0691.14006

[Farkas 2009] G. Farkas, "Koszul divisors on moduli spaces of curves", *Amer. J. Math.* **131**:3 (2009), 819–867. MR 2530855 Zbl 1176.14006

[Farkas and Ludwig 2010] G. Farkas and K. Ludwig, "The Kodaira dimension of the moduli space of Prym varieties", *J. Eur. Math. Soc.* **12**:3 (2010), 755–795. MR 2639318 Zbl 1193.14043

[Farkas and Verra 2016] G. Farkas and A. Verra, "Prym varieties and moduli of polarized Nikulin surfaces", *Adv. Math.* **290** (2016), 314–328. MR 3451926 Zbl 1332.14035

[Farkas et al. 2014] G. Farkas, S. Grushevsky, R. Salvati Manni, and A. Verra, "Singularities of theta divisors and the geometry of A_5", *J. Eur. Math. Soc.* **16**:9 (2014), 1817–1848. MR 3273309 Zbl 06377393

[Harris 1986] J. Harris, "On the Severi problem", *Invent. Math.* **84**:3 (1986), 445–461. MR 837522 Zbl 0596.14017

[Iliev and Markushevich 2000] A. Iliev and D. Markushevich, "Quartic 3-fold: Pfaffians, vector bundles, and half-canonical curves", *Michigan Math. J.* **47**:2 (2000), 385–394. MR 1793633 Zbl 1077.14551

[Mumford 1966] D. Mumford, "On the equations defining abelian varieties. I", *Invent. Math.* **1** (1966), 287–354. MR 0204427 Zbl 0219.14024

Communicated by David Eisenbud
Received 2015-12-18 Revised 2016-04-19 Accepted 2016-08-30

gregor.bruns@hu-berlin.de *Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany*

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1965

A vanishing theorem for weight-one syzygies

Lawrence Ein, Robert Lazarsfeld and David Yang

We give a criterion for the vanishing of the weight-one syzygies associated to a line bundle B in a sufficiently positive embedding of a smooth complex projective variety of arbitrary dimension.

Introduction

Inspired by the methods of Voisin [2002; 2005], Ein and Lazarsfeld [2015] recently proved the gonality conjecture of [Green and Lazarsfeld 1986], asserting that one can read off the gonality of an algebraic curve C from the syzygies of its ideal in any one embedding of sufficiently large degree. They deduced this as a special case of a vanishing theorem for the asymptotic syzygies associated to an arbitrary line bundle B on C, and conjectured that an analogous statement should hold on a smooth projective variety of any dimension. The purpose of this note is to prove the conjecture in question.

Turning to details, let X be a smooth complex projective variety of dimension n, and set

$$L_d = dA + P,$$

where A is ample and P is arbitrary. We always assume that d is sufficiently large so that L_d is very ample, defining an embedding

$$X \subseteq \mathbb{P}H^0(X, L_d) = \mathbb{P}^{r_d}.$$

Given an arbitrary line bundle B on X, we wish to study the weight-one syzygies of B with respect to L_d for $d \gg 0$. More precisely, let $S = \operatorname{Sym} H^0(X, L_d)$ be the homogeneous coordinate ring of $\mathbb{P}H^0(X, L_d)$, and put

$$R = R(X, B; L_d) = \bigoplus_m H^0(X, B + mL_d).$$

The research of Ein was partially supported by NSF grant DMS-1501085. Lazarsfeld's research was partially supported by NSF grant DMS-1439285.
MSC2010: primary 14J99; secondary 13D02.
Keywords: Syzygies.

Thus R is a finitely generated graded S-module, and hence has a minimal graded free resolution $E_\bullet = E_\bullet(B; L_d)$:

$$0 \longrightarrow E_{r_d} \longrightarrow \cdots \longrightarrow E_2 \longrightarrow E_1 \longrightarrow E_0 \longrightarrow R \longrightarrow 0,$$

where $E_p = \bigoplus S(-a_{p,j})$. As customary, denote by $K_{p,q}(X, B; L_d)$ the finite-dimensional vector space of degree $p+q$ minimal generators of E_p, so that

$$E_p(B; L_d) = \bigoplus_q K_{p,q}(X, B; L_d) \otimes_{\mathbb{C}} S(-p-q).$$

We refer to elements of this group as p-th syzygies of B with respect to L_d of weight q. When $B = \mathcal{O}_X$ we write simply $K_{p,q}(X; L_d)$, which — provided that d is large enough so that L_d is normally generated — are the vector spaces describing the syzygies of the homogeneous ideal $I_X \subset S$ of X in $\mathbb{P}H^0(X, L_d)$.

The question we address involves fixing B and asking when it happens that

$$K_{p,q}(X, B; L_d) = 0 \quad \text{for } d \gg 0.$$

When $q = 0$ or $q \geq 2$ the situation is largely understood thanks to results of Green [1984a; 1984b] and Ein and Lazarsfeld [1993; 2012]. (See Remark 1.10 for a summary.) Moreover in this range the statements are uniform in nature, in that they don't depend on the geometry of X or B. However as suggested in [Ein and Lazarsfeld 2012, Problem 7.1], for $K_{p,1}(X, B; L_d)$ one can anticipate more precise asymptotic results that do involve geometry. This is what we establish here.

Recall that a line bundle B on a smooth projective variety X is said to be *p-jet very ample* if for every effective zero-cycle

$$w = a_1x_1 + \cdots + a_sx_s$$

of degree $p+1 = \sum a_i$ on X, the natural map

$$H^0(X, B) \to H^0(X, B \otimes \mathcal{O}_X/\mathfrak{m}_1^{a_1} \cdots \mathfrak{m}_s^{a_s})$$

is surjective, where $\mathfrak{m}_i \subseteq \mathcal{O}_X$ is the ideal sheaf of x_i. So for example if $p = 1$ this is simply asking that B be very ample. When $\dim X = 1$ the condition is the same as requiring that B be p-very ample — i.e., that every subscheme $\xi \subset X$ of length $p+1$ imposes independent conditions in $H^0(X, B)$ — but in higher dimensions it is a stronger condition.

Our main result is

Theorem A. *If B is p-jet very ample, then*

$$K_{p,1}(X, B; L_d) = 0 \quad \textit{for } d \gg 0.$$

The statement was conjectured in [Ein and Lazarsfeld 2015, Conjecture 2.4], where the case $\dim X = 1$ was established.

It is not clear whether one should expect that p-jet amplitude is equivalent to the vanishing of $K_{p,1}(X, B; L_d)$ for $d \gg 0$. However we prove:

Theorem B. *Suppose that there is a* reduced $(p+1)$*-cycle* w *on* X *that fails to impose independent conditions on* $H^0(X, B)$*. Then*

$$K_{p,1}(X, B; L_d) \neq 0 \quad \textit{for } d \gg 0.$$

In general, the proof of Theorem A will show that if $H^1(X, B) = 0$, then the jet amplitude hypothesis on B is equivalent when $d \gg 0$ to the vanishing of a group that contains $K_{p,1}(X, B; L_d)$ as a subspace (Remark 1.8).

When $B = K_X$ is the canonical bundle of X, Theorem A translates under a mild additional hypothesis into a statement involving the syzygies of L_d itself.

Corollary C. *Assume that* $H^i(X, \mathcal{O}_X) = 0$ *for* $0 < i < n$*, or equivalently that* $X \subseteq \mathbb{P}^{r_d}$ *is projectively Cohen–Macaulay for* $d \gg 0$.

(i) *The canonical bundle* K_X *of* X *is very ample if and only if*

$$K_{r_d-n-1,n}(X; L_d) = 0 \quad \textit{for } d \gg 0.$$

(ii) *If* K_X *is* p*-jet very ample, then*

$$K_{r_d-n-p,n}(X; L_d) = 0 \quad \textit{for } d \gg 0.$$

When $n = \dim X = 1$, this (together with Theorem B) implies $K_{r_d-c,1}(X; L_d) \neq 0$ for $d \gg 0$ if and only if X admits a branched covering $X \to \mathbb{P}^1$ of degree $\leq c$, which is the statement of the gonality conjecture established in [Ein and Lazarsfeld 2015].

The proof of Theorem A occupies Section 1. It follows very closely the strategy of [Ein and Lazarsfeld 2015], which in turn was inspired by the ideas of Voisin [2002; 2005]. However instead of working on a Hilbert scheme or symmetric product, we work on a Cartesian product of X, using an idea that goes back in a general way to Green [1984b]. For the benefit of nonexperts, we outline now the approach in some detail in the toy case $p = 0$.[1]

Keeping notation as above, it follows from the definition that $K_{0,1}(X, B; L_d) = 0$ if and only if the multiplication map

$$H^0(X, B) \otimes H^0(X, L_d) \to H^0(X, B \otimes L_d) \tag{$*$}$$

[1]This was in fact the train of thought that led us to the arguments here and in [Ein and Lazarsfeld 2015].

is surjective: in fact, $K_{0,1}$ is its cokernel. A classical way to study such maps is to pass to the product $X \times X$ and then restrict to the diagonal. Specifically, (∗) is identified with the homomorphism

$$H^0(X \times X, \mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L_d) \to H^0(X \times X, \mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L_d \otimes \mathcal{O}_\Delta) \qquad (**)$$

arising from this restriction. Thus the vanishing $K_{0,1}(X, B; L_d) = 0$ is implied by the surjectivity of (∗∗). Green [1984b] observed that there is a similar way to tackle the $K_{p,1}$ for $p \geq 1$: one works on the $(p+2)$-fold product $X^{p+2} = X \times X^{p+1}$ and restricts to a suitable union of pairwise diagonals (Proposition 1.1). This is explained in Section 1, and forms the starting point of our argument. Although not strictly necessary we give a new proof of Green's result here that clarifies its relation to other approaches.

There remains the issue of actually proving the surjectivity of (∗∗) for $d \gg 0$ provided that B is 0-jet very ample, i.e., globally generated. For this one starts with the restriction

$$\mathrm{pr}_1^* B \to \mathrm{pr}_1^* B \otimes \mathcal{O}_\Delta$$

of sheaves on $X \times X$ and pushes down to X via pr_2. There results a map of vector bundles

$$\mathrm{ev}_B : H^0(X, B) \otimes_{\mathbb{C}} \mathcal{O}_X \to B$$

on X which is given by evaluation of sections of B. Note that ev_B is surjective as a map of bundles if and only if B is globally generated. The surjectivity in (∗) or (∗∗) is then equivalent to the surjectivity on global sections of the map

$$H^0(X, B) \otimes L_d \to B \otimes L_d$$

obtained from twisting ev_B by L_d.

Suppose now that B is 0-jet very ample. The setting $M_B = \ker(\mathrm{ev}_B)$, we get an exact sequence

$$0 \to M_B \to H^0(X, B) \otimes_{\mathbb{C}} \mathcal{O}_X \to B \to 0$$

of sheaves on X. Serre vanishing implies that

$$H^1(X, M_B \otimes L_d) = 0$$

for $d \gg 0$, and by what we have just said this means that $K_{0,1}(X, B; L_d) = 0$. The proof of Theorem A in general proceeds along analogous lines. We construct a torsion-free sheaf $\mathcal{E}_B = \mathcal{E}_{p+1,B}$ of rank $p+1$ on X^{p+1} whose fiber at $(x_1, \ldots, x_{p+1})$ is identified with

$$H^0(X, B \otimes \mathcal{O}_X / \mathfrak{m}_1 \cdots \mathfrak{m}_{p+1}),$$

where $\mathfrak{m}_i \subseteq \mathcal{O}_X$ is the ideal sheaf of x_i. This comes with an evaluation map

$$\mathrm{ev}_{p+1,B} : H^0(X, B) \otimes_{\mathbb{C}} \mathcal{O}_{X^{p+1}} \to \mathcal{E}_B$$

which is surjective (as a map of sheaves) if and only if B is p-jet very ample.[2] Green's criterion for the vanishing of $K_{p,1}(X, B; L_d)$ turns out to be equivalent to the surjectivity of the map on global sections resulting from twisting $\mathrm{ev}_{p+1,B}$ by a suitable ample divisor $\mathcal{N}_d$ on X^{p+1} deduced from L_d, and this again follows from Serre vanishing.

Returning to the case $p = 0$, the argument just sketched actually proves more. Namely for arbitrary B one has an exact sequence

$$0 \to \ker(\mathrm{ev}_B) \to H^0(X, B) \otimes_{\mathbb{C}} \mathcal{O}_X \to B \to \mathrm{coker}(\mathrm{ev}_B) \to 0,$$

and so Serre vanishing shows conversely that if B is not 0-jet very ample then

$$K_{p,1}(X, B; L_d) = H^0(X, \mathrm{coker}(\mathrm{ev}_B) \otimes L_d) \neq 0 \tag{$***$}$$

for $d \gg 0$. Unfortunately this does not generalize when $p \geq 1$ because the computations on X^{p+1} lead to groups that contain $K_{p,1}(X, B; L_d)$ as summands, but may contain other terms as well. (Said differently, Green's criterion is sufficient but not necessary for the vanishing of $K_{p,1}$.) To prove a nonvanishing statement such as Theorem B, one needs a geometric interpretation of $K_{p,1}$ itself. Voisin [2002; 2005] achieves this by working on a Hilbert scheme — which has the advantage of being smooth when $\dim X = 2$ — while Yang [2014] passes in effect to the symmetric product.[3] We follow the latter approach for Theorem B: we exhibit a sheaf on $\mathrm{Sym}^{p+1}(X)$ whose twisted global sections compute $K_{p,1}(X, B; L_d)$, and we show that it is nonzero provided that there is a reduced cycle that fails to impose independent conditions on $H^0(X, B)$. Then we can argue much as in the case $p = 0$ just described. This is the content of Section 2.

1. Proof of Theorem A

This section is devoted to the proof of Theorem A from the introduction.

We start by describing the set-up. As above, X is a smooth complex projective variety of dimension n, and we consider the $(p+2)$-fold product

$$Y \stackrel{\text{def}}{=} X \times X^{p+1}$$

of X with itself. For $0 \leq i < j \leq p+1$ denote by

$$\pi_{i,j} : Y \to X \times X$$

[2] Note that the fiber of B at a point $x \in X$ is identified with $B \otimes \mathcal{O}_X/\mathfrak{m}_x$, i.e., $\mathcal{E}_{1,B} = B$.

[3] Roughly speaking, one is picking out $K_{p,1}$ inside Green's construction as the space of invariants under a suitable action of the symmetric group.

the projection of Y onto the product of the i and j factors. We write $\Delta_{i,j} \subseteq Y$ for the pull-back of the diagonal $\Delta \subseteq X \times X$ under $\pi_{i,j}$, so that $\Delta_{i,j}$ consists of those points $y = (x_0, x_1, \ldots, x_{p+1}) \in Y$ with $x_i = x_j$.

The basic idea — which goes back to Green [1984b] and has been used repeatedly since (e.g., [Inamdar 1997; Bertram et al. 1991; Lazarsfeld et al. 2011; Hwang and To 2013; Yang 2014]) — is to relate syzygies on X to a suitable union of pairwise diagonals on Y. Specifically, let

$$Z = Z_{p+1} = \Delta_{0,1} \cup \cdots \cup \Delta_{0,p+1} \subseteq X \times X^{p+1} \tag{1-1}$$

be the union of the indicated pairwise diagonals, considered as a reduced subscheme. We denote by

$$q : Z \to X, \quad \sigma : Z \to X^{p+1} \tag{1-2}$$

the indicated projections.

The importance of this construction for us is given by:

Proposition 1.1. *Let L and B be respectively base-point-free and arbitrary line bundles on X, and assume (for simplicity) that $H^1(X, L) = 0$. If the restriction homomorphism*

$$H^0(Y, B \boxtimes L^{\boxtimes p+1}) \to H^0(Y, (B \boxtimes L^{\boxtimes p+1}) \mid Z) \tag{1-3}$$

is surjective, then

$$K_{p,1}(X, B; L) = 0.$$

The Proposition was essentially established for instance in [Yang 2014], but it is instructive to give a direct argument. We start with a lemma that will also be useful later:

Lemma 1.2. *Writing $I_{Z/Y}$ for the ideal sheaf of Z in Y, one has*

$$I_{Z/Y} = I_{\Delta_{0,1}/Y} \cdot I_{\Delta_{0,2}/Y} \cdots I_{\Delta_{0,p+1}/Y} = \bigotimes_{j=1}^{p+1} \pi_{0,j}^* I_{\Delta/X \times X}.$$

Sketch of Proof. This is implicit in [Li 2009, Theorem 1.3], but does not appear there explicitly so we very briefly indicate an argument. The statement is étale local, so we can assume $X = \mathbb{A}^n$. By looking at a suitable subtraction map, as in [Lazarsfeld et al. 2011, (1-3)], it then suffices to prove the analogous statement for $Y = X^{p+1}$ with Z being the union of the "coordinate planes"

$$L_i = \{(x_1, \ldots, x_i, \ldots, x_{p+1}) \mid x_i = 0 \in \mathbb{A}^n\} = \{\mathbb{A}^n\} \times \ldots \times \{0\} \times \ldots \times \{\mathbb{A}^n\} \subseteq Y$$

$(1 \le i \le p+1)$. For this one can proceed by induction on p, writing out explicitly the equations defining each L_i. □

Remark 1.3. This is the essential place where we use the hypothesis that X is smooth. We do not know whether the statement of the Lemma remains true for singular X.

Proof of Proposition 1.1. To begin with, it is well known (see [Green 1984a]) that $K_{p,1}(X, B; L)$ is the cohomology of the Koszul-type complex

$$\Lambda^{p+1}H^0(L)\otimes H^0(B)\to \Lambda^p H^0(L)\otimes H^0(B+L)\to \Lambda^{p-1}H^0(L)\otimes H^0(B+2L).$$

Moreover this cohomology can in turn be interpreted geometrically in terms of the vector bundle M_L on X defined (as in the introduction) as the kernel of the evaluation map

$$\mathrm{ev}_L : H^0(L)\otimes_{\mathbb{C}} \mathcal{O}_X \to L.$$

Specifically, M_L sits in an exact sequence of vector bundles

$$0\to M_L\to H^0(L)\otimes \mathcal{O}_X\to L\to 0 \tag{$\star$}$$

on X, and then $K_{p,1}(X, B; L) = 0$ if and only if the sequence

$$0\to \Lambda^{p+1}M_L\otimes B\to \Lambda^{p+1}H^0(L)\otimes B\to \Lambda^p M_L\otimes L\otimes B\to 0 \tag{$\star\star$}$$

deduced from ($\star$) is exact on global sections. (See for instance [Green and Lazarsfeld 1986, Lemma 1.10] or [Lazarsfeld 1989].)

On the other hand, consider on $X\times X$ the exact sequence

$$0\to I_\Delta\otimes \mathrm{pr}_2^*L\to \mathrm{pr}_2^*L\to L\otimes\mathcal{O}_\Delta\to 0.$$

As in the introduction, this pushes down via pr_1 to ($\star$). Therefore one finds from Lemma 1.2 and the Künneth formula that

$$q_*(I_{Z/Y}\otimes B\boxtimes L^{\boxtimes p+1}) = \Big(\bigotimes^{p+1} M_L\Big)\otimes B,$$

and moreover the R^1q_* vanishes thanks to our hypothesis that $H^1(L)=0$.[4]

Writing

$$\mathcal{N} = \mathrm{coker}\Big(\bigotimes^{p+1} M_L\to \bigotimes^{p+1} H^0(L)\Big),$$

[4]The Künneth theorem in play here is the following: let $V_1\to S, \dots, V_r\to S$ be mappings of schemes over a field, and suppose that F_i is a quasicoherent sheaf on V_i that is flat over S. Write

$$p_i : V_1\times_S\cdots\times_S V_r\to V_i,\quad p : V_1\times_S\cdots\times_S V_r\to S$$

for the natural maps. Then

$$p_*(p_1^*F_1\otimes\cdots\otimes p_r^*F_r) = p_{1,*}F_1\otimes\cdots\otimes p_{r,*}F_r$$

as sheaves on S, with analogous Künneth-type computations of the $R^j p_*$. See for instance [Kempf 1980, Theorem 14] for a simple proof when S is affine, and [EGA III$_2$ 1963, Theorem 6.7.8] for the general case.

it follows that

$$q_*((B \boxtimes L^{\boxtimes p+1}) \mid Z) = \mathcal{N} \otimes B,$$

and hence the surjectivity of (1-3) is equivalent to asking that

$$0 \to \Big(\bigotimes^{p+1} M_L\Big) \otimes B \to \Big(\bigotimes^{p+1} H^0(L)\Big) \otimes B \to \mathcal{N} \otimes B \to 0 \qquad (\star\star\star)$$

be exact on global sections. But since we are in characteristic zero, the exact sequence $(\star\star)$ is a summand of this, and the lemma follows. □

Remark 1.4. The argument just completed shows that if in addition $H^1(X, B) = 0$ then (1-3) is surjective if and only if

$$H^1\Big(X, \Big(\bigotimes^{p+1} M_L\Big) \otimes B\Big) = 0.$$

It remains to relate these considerations to the jet-amplitude of B. To this end, keeping notation as in (1-2), set

$$\mathcal{E}_B = \mathcal{E}_{p+1,B} = \sigma_*(q^* B).$$

This is a torsion-free sheaf of rank $p+1$ on X^{p+1} (since it is the push forward of a line bundle under a finite mapping of degree $p+1$), and one has:

Lemma 1.5. (i) *Fix a point*

$$\xi = (x_1, \dots, x_{p+1}) \in X^{p+1},$$

the x_i being (possibly nondistinct) points of X, and denote by $\mathcal{E}_B|\xi$ the fiber of $\mathcal{E}_B$ at ξ. Then there is a natural identification

$$\mathcal{E}_B|\xi = H^0(X, B \otimes \mathcal{O}_X/\mathfrak{m}_1 \cdots \mathfrak{m}_{p+1}),$$

where $\mathfrak{m}_i \subseteq \mathcal{O}_X$ is the maximal ideal of x_i.

(ii) *There is a canonical injection*

$$H^0(X, B) \hookrightarrow H^0(X^{p+1}, \mathcal{E}_B),$$

giving rise to a homomorphism

$$\mathrm{ev}_B = \mathrm{ev}_{p+1,B} : H^0(X, B) \otimes_{\mathbb{C}} \mathcal{O}_{X^{p+1}} \to \mathcal{E}_B.$$

of sheaves on X^{p+1}. Under the identification in (i), ev_B *is given fiberwise by the natural map*

$$H^0(X, B) \to H^0\big(X, B \otimes \mathcal{O}_X/(\mathfrak{m}_1 \cdots \mathfrak{m}_{p+1})\big).$$

(iii) *The homomorphism* (1-3) *is identified with the map on global sections arising from the sheaf homomorphism*

$$H^0(X, B) \otimes_{\mathbb{C}} L^{\boxtimes p+1} \to \mathcal{E}_B \otimes L^{\boxtimes p+1}. \tag{1-4}$$

on X^{p+1} *determined by twisting* ev_B *by* $L^{\boxtimes p+1}$.

(iv) *The mapping* $\mathrm{ev}_{p+1,B}$ *is surjective as a homomorphism of sheaves on* X^{p+1} *if and only if* B *is* p*-jet very ample.*

Proof. For (i), consider the diagram

$$\begin{array}{ccc} Z & \xrightarrow{\subseteq} & X \times X^{p+1} \\ & {\scriptstyle\sigma}\searrow \quad \swarrow{\scriptstyle \mathrm{pr}_2} & \\ & X^{p+1} & \end{array} \tag{1-5}$$

and fix $\xi = (x_1, \ldots, x_{p+1}) \in X^{p+1}$. The scheme-theoretic fiber $\sigma^{-1}(\xi)$ lives naturally as a subscheme of X, and Lemma 1.2 implies that it is in fact the scheme defined by the ideal sheaf

$$\mathfrak{m}_1 \cdots \mathfrak{m}_{p+1} \subseteq \mathcal{O}_X.$$

Therefore, thanks to the projection formula, the fiber $\mathcal{E}_B|\xi$ of $\mathcal{E}_B$ at ξ is identified with

$$\mathrm{pr}_{2,*}\big(B \otimes \mathcal{O}_X/(\mathfrak{m}_1 \cdot \ldots \cdot \mathfrak{m}_{p+1})\big) = H^0\big(X, B \otimes \mathcal{O}_X/(\mathfrak{m}_1 \cdots \mathfrak{m}_{p+1})\big),$$

as claimed. For (ii), note that in any event

$$H^0(X^{p+1}, \mathcal{E}_B) = H^0(X^{p+1}, \sigma_* q^* B) = H^0(Z, q^* B).$$

On the other hand, each of the irreducible components of Z maps via projection onto X, and this gives an inclusion

$$q^* : H^0(X, B) \to H^0(Z, q^* B) = H^0\big(X \times X^{p+1}, (\mathrm{pr}_1^* B)|Z\big).$$

It is evident from the construction that fiber by fiber ev_B is as described, and (iv) is then a consequence of the fact that a morphism of sheaves is surjective if and only it is so on each fiber. Finally, statement (iii) follows from the construction of $\mathcal{E}_B$ and ev_B. □

Remark 1.6. Using the resolution of $\mathcal{O}_Z$ appearing in [Yang 2014, p. 4], one can show that in fact

$$H^0(X^{p+1}, \mathcal{E}_B) = H^0(X, B).$$

However this isn't necessary for the argument.

Remark 1.7. The reader familiar with [Ein and Lazarsfeld 2015] or Voisin's Hilbert schematic approach to syzygies will recognize that $Z \to X^{p+1}$ plays the role of the universal family over the Hilbert scheme, and that Proposition 1.1 is the analogue of [Voisin 2002, Lemma 1, p. 369]. The sheaf $\mathcal{E}_{p+1,B}$ plays the role of the vector bundle $E_{p+1,B}$ on the symmetric product appearing in [Ein and Lazarsfeld 2015].

Just as in [Ein and Lazarsfeld 2015], the main result now follows immediately from Serre vanishing.

Proof of Theorem A. Assuming that B is p-jet very ample, so that $\mathrm{ev}_{p+1,B}$ is surjective, let $\mathcal{M}_{p+1,B}$ denote its kernel:

$$0 \to \mathcal{M}_{p+1,B} \to H^0(B) \otimes \mathcal{O}_{X^{p+1}} \to \mathcal{E}_B \to 0.$$

To show that $K_{p,1}(X, B; L_d) = 0$ it suffices, thanks to Proposition 1.1 and its interpretation in terms of (1-4), to prove that

$$H^1(X^{p+1}, \mathcal{M}_{p+1,B} \otimes L_d^{\boxtimes p+1}) = 0$$

for $d \gg 0$. But this follows immediately from Serre vanishing. □

Remark 1.8. It follows from the argument just completed that if $L = L_d$ then the surjectivity in Proposition 1.1 holds for $d \gg 0$ if and *only if* B is p-jet very ample. In particular, in view of Remark 1.4 this means that if $H^1(X, B) = 0$ then the p-jet very amplitude of B is equivalent to the vanishing

$$H^1\Big(X, \Big(\bigotimes^{p+1} M_{L_d}\Big) \otimes B\Big) = 0 \quad \text{for } d \gg 0.$$

Proof of Corollary C. Under the stated hypothesis on X, the groups in question are Serre dual to $K_{p,1}(X, B; L_d)$ for $d \gg 0$ (see [Green 1984a, §2]). If B fails to be very ample, then a simple argument as in [Eisenbud et al. 2006, Theorem 1.1] shows that $K_{1,1}(X, B; L_d) \neq 0$ for $d \gg 0$, and therefore the corollary follows from the main theorem. □

Remark 1.9. In the case of curves, Rathmann [2016] has given a very interesting argument that leads to an essentially optimal effective version of the asymptotic results of [Ein and Lazarsfeld 2015]: in fact, it suffices that $H^1(L) = H^1(L - B) = 0$. In this spirit, it would be very interesting to find an effective estimate for the positivity of L to guarantee the vanishing of $K_{p,1}(X, B; L)$ when B is p-jet very ample.

Remark 1.10 (other Koszul cohomology groups). To conclude this section, we briefly summarize what is known about the groups $K_{p,q}(X, B; L_d)$ for $q \neq 1$. Specifically, fix B. Then for $d \gg 0$:

(i) $K_{p,q}(X, B; L_d) = 0$ for $q \geq n + 2$.

(ii) One has

$$K_{p,0}(X, B; L_d) \neq 0 \Longleftrightarrow 0 \le p \le r(B), \text{ and}$$
$$K_{p,n+1}(X, B; L_d) \neq 0 \Longleftrightarrow r_d - n - r(K_X - B) \le p \le r_d - n.$$

(iii) If $q \ge 2$, then $K_{p,q}(X, B; L_d) = 0$ when $p \le O(d)$.

(iv) For $1 \le q \le n$, $K_{p,q}(X, B; L_d) \neq 0$ for

$$O(d^{q-1}) \le p \le r_d - O(d^{n-1}).$$

Statement (i) is a consequence of Castelnuovo–Mumford regularity, while (ii) is due to Green and others. (See [Green 1984a, §3; Ein and Lazarsfeld 2012, Corollary 3.3, §5].) Assertion (ii) follows for instance from [Ein and Lazarsfeld 1993], while (iv) is the main result of [Ein and Lazarsfeld 2012]. Furthermore, it is conjectured in [Ein and Lazarsfeld 2012] that if $q \ge 2$, then $K_{p,q}(X, B; L_d) = 0$ for $p \le O(d^{q-1})$.

2. A nonvanishing theorem

This section is devoted to the proof of Theorem B from the introduction. Recall the statement:

Theorem 2.1. *Assume that the nonsingular projective variety X carries an effective $(p+1)$-cycle $w = x_1 + \cdots + x_{p+1}$ consisting of $p+1$* distinct *points $x_1, \ldots, x_{p+1} \in X$ that fail to impose independent conditions on $H^0(X, B)$. Then*

$$K_{p,1}(X, B; L_d) \neq 0 \quad \textit{for } d \gg 0.$$

The argument is somewhat technical, so before launching into it we would like to outline the rough strategy. As in the case $p = 0$ discussed in the introduction, in principle we would like to find a map of sheaves on $\mathrm{Sym}^{p+1}(X)$ depending on B — say $a_B : \mathcal{A}_1 \to \mathcal{A}_2$ — having the property that

$$K_{p,1}(X, B; L_d) = \mathrm{coker}\big(H^0(\mathcal{A}_1 \otimes \mathcal{N}(L_d)) \to H^0(\mathcal{A}_2 \otimes \mathcal{N}(L_d))\big), \qquad (2\text{-}1)$$

where $\mathcal{N}(L_d)$ is a line bundle whose positivity grows suitably with d. Ideally — as in equation $(***)$ from the introduction — we would be able to see that a_B cannot be surjective as a map of sheaves if B is not p-jet very ample, and then one could hope to apply Serre vanishing to conclude that $K_{p,1}(X, B; L_d)$ cannot vanish for $d \gg 0$. Unfortunately we do now know whether such a construction is possible. Instead, what we do in effect is to use the ideas of Yang [2014] to construct a map α_B, and show that the nonvanishing of $K_{p,1}$ is implied by the nonvanishing of a certain quotient sheaf of $\mathcal{A}_2$. We show that a *reduced* $(p+1)$-cycle that fails to impose independent conditions on $H^0(X, B)$ must appear in the support of this quotient, and this leads to the stated nonvanishing.

We start by recalling the results of Yang [2014] interpreting $K_{p,1}$ as an equivariant cohomology group. Consider then a very ample line bundle L on the smooth complex projective variety X. Then the symmetric group S_{p+1} acts in two ways on the bundle $L^{\boxtimes p+1}$ on X^{p+1}, namely via the symmetric and the alternating characters. Denote these S_{p+1}-bundles on X^{p+1} by

$$L^{\boxtimes p+1,\mathrm{sym}} \quad \text{and} \quad L^{\boxtimes p+1,\mathrm{alt}}, \tag{2-2}$$

respectively. Now let S_{p+1} act on $X \times X^{p+1}$ via the trivial action on the first factor, so that the union of pairwise diagonals $Z \subseteq X \times X^{p+1}$ defined in (1-1) becomes an S_{p+1}-subspace. It is established in [Yang 2014, Theorem 3] that if

$$H^i(X, mL) = H^i(X, B + mL) = 0 \quad \text{for } i, m > 0, \tag{2-3}$$

then $K_{p,1}(X, B; L)$ is identified with the cokernel of the restriction mapping

$$H^0_{S_{p+1}}(X \times X^{p+1}, \mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}}) \to H^0_{S_{p+1}}(Z, \mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}} \otimes \mathcal{O}_Z) \tag{2-4}$$

on S_{p+1}-equivariant cohomology groups.[5] One can think of this as a precision and strengthening of Proposition 1.1. Following the line of attack of Section 1, the plan is to study these groups by modding out by the symmetric group and pushing down to the symmetric product.

To this end, denote by $\mathrm{Sym}^{p+1}(X)$ the $(p+1)$-st symmetric product of X, which we view as parametrizing zero-cycles of degree $p+1$, and write

$$\pi : X^{p+1} \to \mathrm{Sym}^{p+1}(X)$$

[5]For the theory of equivariant cohomology groups and pushforwards, see [Grothendieck 1957, Chapter 5]. What we need can be summarized as follows. Let G be a finite group acting on a complex projective variety V, and suppose a coherent sheaf F on X together with an action of G on F are given. Then one can define equivariant cohomology groups $H^j_G(V, F)$. While this isn't how they are initially constructed, one can show that

$$H^j_G(V, F) = H^j(V, F)^G,$$

the group on the right being the G-invariant subspace of $H^j(V, F)$ under the natural action of G on this cohomology group [Grothendieck 1957, p. 202], and for practical purposes one can take this as the definition. Writing

$$\pi : V \to V/G \overset{\mathrm{def}}{=} W,$$

one also has an action of G on $\pi_* F$. The G-equivariant direct image of F can be interpreted as

$$\pi_*^G F = (\pi_*(F))^G,$$

and one can show that

$$H^j_G(V, F) = H^j(W, \pi_*^G F).$$

for the quotient map. The equivariant pushforward of the line bundles in (2-2) determine respectively a line bundle and torsion-free sheaf of rank one

$$\mathcal{S}_{p+1}(L) \stackrel{\text{def}}{=} \pi_*^{S_{p+1}}(L^{\boxtimes p+1,\text{sym}}), \quad \mathcal{N}_{p+1}(L) \stackrel{\text{def}}{=} \pi_*^{S_{p+1}}(L^{\boxtimes p+1,\text{alt}})$$

on $\text{Sym}^{p+1}(X)$. One has

$$H^0\big(\text{Sym}^{p+1}(X), \mathcal{S}_{p+1}(L)\big) = \text{Sym}^{p+1} H^0(X, L),$$
$$H^0\big(\text{Sym}^{p+1}(X), \mathcal{N}_{p+1}(L)\big) = \Lambda^{p+1} H^0(X, L),$$

and for any line bundle A on X:

$$\mathcal{S}(L \otimes A) = \mathcal{S}(L) \otimes \mathcal{S}(A), \quad \mathcal{N}(L \otimes A) = \mathcal{N}(L) \otimes \mathcal{S}(A). \tag{2-5}$$

Moreover $\mathcal{S}(A)$ is ample if A is.

Proof of Theorem 2.1. Note to begin with that the symmetric group S_{p+1} acts on each of the spaces appearing in diagram (1-5). Taking the quotients yields the diagram

$$\begin{array}{ccc} \bar{Z} = X \times \text{Sym}^p(X) & \xhookrightarrow{\ \subseteq\ } & X \times \text{Sym}^{p+1}(X) \\ & {\scriptstyle \bar{\sigma}} \searrow \quad \swarrow {\scriptstyle p_2} & \\ & \text{Sym}^{p+1}(X) & \end{array} \tag{2-6}$$

where $\bar{\sigma}$ is the addition map,

$$p_1 : X \times \text{Sym}^{p+1}(X) \to X, \quad p_2 : X \times \text{Sym}^{p+1}(X) \to \text{Sym}^{p+1}(X)$$

are the projections, and the inclusion on the top line is given by $(x, w) \mapsto (x, x+w)$. One has

$$(1 \times \pi)_*^{S_{p+1}}(\text{pr}_1^* B \otimes \text{pr}_2^* L^{\boxtimes p+1,\text{alt}}) = p_1^* B \otimes p_2^* \mathcal{N}(L).$$

Now define

$$\mathcal{G}(B; L) = (1 \times \pi)_*^{S_{p+1}}(\text{pr}_1^* B \otimes \text{pr}_2^* L^{\boxtimes p+1,\text{alt}} \otimes \mathcal{O}_Z).$$

Pushing forward the restriction to Z gives rise to a natural surjective mapping

$$\varepsilon(B; L) : p_1^* B \otimes p_2^* \mathcal{N}(L) \to \mathcal{G}(B; L) \tag{2-7}$$

of sheaves on $X \times \text{Sym}^{p+1}(X)$. Thanks to [Grothendieck 1957, §5.2], the groups appearing in (2-4) are given by the global sections of the sheaves in (2-7), and hence under the vanishing hypothesis (2-3), $K_{p,1}(X, B; L)$ is computed as the cokernel

$$K_{p,1}(X, B; L) = \text{coker}\big(H^0(\varepsilon(B; L))\big)$$

on global sections determined by $\varepsilon(B; L)$. Hence we are reduced to showing that under the hypothesis of the theorem, $\varepsilon(B; L_d)$ cannot be surjective on global sections when $d \gg 0$.

The next step is to form and study the push-forward of (2-7) to $\mathrm{Sym}^{p+1}(X)$. To begin with, define $\mathcal{F}(B; L) = p_{2,*}\mathcal{G}(B; L)$ and $\delta(B; L) = p_{2,*}\varepsilon(B; L)$. This gives rise to a morphism

$$\delta(B; L) : H^0(X, B) \otimes_{\mathbb{C}} \mathcal{N}(L) \to \mathcal{F}(B; L)$$

of sheaves on $\mathrm{Sym}^{p+1}(X)$ with the property that

$$K_{p,1}(X, B; L) = \mathrm{coker}\big(H^0(\delta(B, L))\big).$$

We wish to study the geometry of this mapping assuming that B does not impose independent conditions on all reduced cycles. We assert that there is a natural homomorphism

$$t : \big(p_1^* B \otimes p_2^* \mathcal{N}(L) \otimes \mathcal{O}_{\bar{Z}}\big) \to \mathcal{G}(B; L) \tag{2-8}$$

which is an isomorphism on the smooth locus of $\bar{Z}$. Grant this for the time being. By the projection formula one has

$$p_{2,*}\big(p_1^* B \otimes p_2^* \mathcal{N}(L) \otimes \mathcal{O}_{\bar{Z}}\big) = \bar{\sigma}_*(p_1^* B) \otimes \mathcal{N}(L),$$

and then taking direct images in (2-7) and (2-8), one arrives at a diagram

$$\begin{array}{ccc} H^0(X, B) \otimes_{\mathbb{C}} \mathcal{N}(L) & \xrightarrow{e \otimes 1} & \bar{\sigma}_*(p_1^* B) \otimes \mathcal{N}(L) \\ & \searrow^{\delta(B; L)} & \downarrow s \\ & & \mathcal{F}(B; L) \end{array} \tag{2-9}$$

where s is an isomorphism over the smooth locus of $\mathrm{Sym}^{p+1}(X)$.

Now fix a reduced zero-cycle $w \in \mathrm{Sym}^{p+1}(X)$ that fails to impose independent conditions on $H^0(X, B)$, and let $\xi \subseteq X$ be the corresponding subscheme of length $p+1$. We can identify the fiber of the morphism

$$e : H^0(X, B) \otimes \mathcal{O}_{\mathrm{Sym}^{p+1}(X)} \to \bar{\sigma}(p_1^* B)$$

appearing in (2-9) at w with the evaluation $H^0(X, B) \to H^0(X, B \otimes \mathcal{O}_\xi)$. Writing $\mathcal{K} = \mathrm{coker}(e)$, so that

$$\mathrm{coker}(e \otimes 1) = \mathcal{K} \otimes \mathcal{N}(L),$$

it follows that $w \in \mathrm{supp}(\mathcal{K} \otimes \mathcal{N}(L))$. But thanks to (2-5) and Serre vanishing, $\bar{\sigma}_*(p_1^* B) \otimes \mathcal{N}(L_d)$ is globally generated and

$$H^0\big(\bar{\sigma}(p_1^* B) \otimes \mathcal{N}(L_d)\big) \to H^0\big(\mathcal{K} \otimes \mathcal{N}(L_d)\big)$$

is surjective when $d \gg 0$. On the other hand, s is an isomorphism in a neighborhood of w since w is reduced, and it then follows that the map

$$H^0\big(\mathcal{K} \otimes \mathcal{N}(L_d)\big) \to H^0\big(\mathrm{coker}(\delta(B;\, L_d))\big)$$

is nonzero when $d \gg 0$. In other words, we have a commutative diagram

$$\begin{array}{ccccc} H^0(B) \otimes H^0(\mathcal{N}(L_d)) & \longrightarrow & H^0\big(\bar{\sigma}(p_1^* B) \otimes \mathcal{N}(L_d)\big) & \twoheadrightarrow & H^0\big(\mathcal{K} \otimes \mathcal{N}(L_d)\big) \\ & \searrow^{\delta(B;L_d)} & \downarrow & \searrow & \downarrow{\scriptstyle \neq 0} \\ & & H^0(\mathcal{F}(B;\, L_d)) & \longrightarrow & H^0\big(\mathrm{coker}(\delta(B;\, L_d))\big), \end{array}$$

with exact top row, in which the right-hand diagonal mapping, and hence also the bottom homomorphism, are nonzero. Therefore $\delta(B;\, L_d)$ cannot be surjective on global sections when $d \gg 0$, as required.

It remains to construct the homomorphism t appearing in (2-8). To this end, let

$$\widehat{Z} = \bar{Z} \times_{X \times \mathrm{Sym}^{p+1}(X)} (X \times X^{p+1}).$$

The projection formula gives an isomorphism

$$(1 \times \pi)_*\big((\mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}}) \otimes \mathcal{O}_{\widehat{Z}}\big) = \big((1 \times \pi)_*(\mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}})\big) \otimes \mathcal{O}_{\bar{Z}},$$

which, upon taking S_{p+1} invariants, yields

$$(1 \times \pi)_*^{S_{p+1}}\big(\mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}} \otimes \mathcal{O}_{\widehat{Z}}\big) = \big(p_1^* B \otimes p_2^* \mathcal{N}(L)\big) \otimes \mathcal{O}_{\bar{Z}}.$$

On the other hand, as a set $\widehat{Z}$ consists of those points

$$(x_0, (x_1, \ldots, x_{p+1})) \in X \times X^{p+1}$$

having the property that x_0 appears in the cycle $x_1 + \cdots + x_{p+1}$. In other words, $\widehat{Z}$ and Z coincide set-theoretically. Since Z is reduced this implies that $Z = \widehat{Z}_{\mathrm{red}}$, and in particular Z is a subscheme of $\widehat{Z}$. Thus there is a natural surjective map

$$(\mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}}) \otimes \mathcal{O}_{\widehat{Z}} \to (\mathrm{pr}_1^* B \otimes \mathrm{pr}_2^* L^{\boxtimes p+1,\mathrm{alt}}) \otimes \mathcal{O}_Z$$

which is an isomorphism over the smooth locus of Z, and taking direct images gives (2-8). □

Remark 2.2. It would be very interesting to give a necessary and sufficient condition for the nonvanishing of $K_{p,1}(X, B;\, L_d)$ for $d \gg 0$. Keeping the notation of the previous proof, the issue is to determine when $\delta(B;\, L)$ has a nonzero cokernel. It is conceivable that the failure of B to be p-jet very ample suffices, but the question seems somewhat difficult to analyze. Already the case $\dim X = 2$ would be interesting.

Remark 2.3. Recall that a line bundle B on a smooth variety X is said to be p-very ample if every finite subscheme of length $p+1$ imposes independent conditions on $H^0(X, B)$. When $\dim X = 1$ this is the same as jet-amplitude, but when $\dim X \geq 2$ it is a strictly weaker condition in general. A quick way to see this is to recall that if A is an ample line bundle on a smooth surface X, then $B_p = K_X + (p+3)A$ is always p-very ample thanks to a theorem of Beltrametti, Francia and Sommese [1989]. On the other hand, the p-jet amplitude of B_p for $p \gg 0$ would imply that the Seshadri constant $\varepsilon(A; x)$ is very close to 1 for every point $x \in X$ (see [Lazarsfeld 1997, Proposition 5.10]). Hence any line bundle A for which there exist points with small Seshadri constant gives rise to examples of the required sort.

Acknowledgements

We are grateful to Claudiu Raicu and Bernd Sturmfels for valuable discussions. We particularly profited from conversations with B. Purnaprajna, who suggested to us that one could work on a Cartesian rather than a symmetric product to establish the vanishing we wanted.

References

[Beltrametti et al. 1989] M. Beltrametti, P. Francia, and A. J. Sommese, "On Reider's method and higher order embeddings", *Duke Math. J.* **58**:2 (1989), 425–439. MR 1016428 Zbl 0702.14010

[Bertram et al. 1991] A. Bertram, L. Ein, and R. Lazarsfeld, "Vanishing theorems, a theorem of Severi, and the equations defining projective varieties", *J. Amer. Math. Soc.* **4**:3 (1991), 587–602. MR 1092845 Zbl 0762.14012

[EGA III$_2$ 1963] A. Grothendieck, "Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, II", *Inst. Hautes Études Sci. Publ. Math.* **17** (1963), 5–91. MR 0163911 Zbl 0122.16102

[Ein and Lazarsfeld 1993] L. Ein and R. Lazarsfeld, "Syzygies and Koszul cohomology of smooth projective varieties of arbitrary dimension", *Invent. Math.* **111**:1 (1993), 51–67. MR 1193597 Zbl 0814.14040

[Ein and Lazarsfeld 2012] L. Ein and R. Lazarsfeld, "Asymptotic syzygies of algebraic varieties", *Invent. Math.* **190**:3 (2012), 603–646. MR 2995182 Zbl 1262.13018

[Ein and Lazarsfeld 2015] L. Ein and R. Lazarsfeld, "The gonality conjecture on syzygies of algebraic curves of large degree", *Publ. Math. Inst. Hautes Études Sci.* **122** (2015), 301–313. MR 3415069 Zbl 1342.14070

[Eisenbud et al. 2006] D. Eisenbud, C. Huneke, and B. Ulrich, "The regularity of Tor and graded Betti numbers", *Amer. J. Math.* **128**:3 (2006), 573–605. MR 2230917 Zbl 1105.13017

[Green 1984a] M. L. Green, "Koszul cohomology and the geometry of projective varieties", *J. Differential Geom.* **19**:1 (1984), 125–171. MR 739785 Zbl 0559.14008

[Green 1984b] M. L. Green, "Koszul cohomology and the geometry of projective varieties, II", *J. Differential Geom.* **20**:1 (1984), 279–289. MR 772134 Zbl 0559.14009

[Green and Lazarsfeld 1986] M. Green and R. Lazarsfeld, "On the projective normality of complete linear series on an algebraic curve", *Invent. Math.* **83**:1 (1986), 73–90. MR 813583 Zbl 0594.14010

[Grothendieck 1957] A. Grothendieck, "Sur quelques points d'algèbre homologique", *Tôhoku Math. J.* (2) **9** (1957), 119–221. MR 0102537 Zbl 0118.26104

[Hwang and To 2013] J.-M. Hwang and W.-K. To, "Syzygies of compact complex hyperbolic manifolds", *J. Algebraic Geom.* **22**:1 (2013), 175–200. MR 2993051 Zbl 1262.14008

[Inamdar 1997] S. P. Inamdar, "On syzygies of projective varieties", *Pacific J. Math.* **177**:1 (1997), 71–76. MR 1444773 Zbl 0898.14015

[Kempf 1980] G. R. Kempf, "Some elementary proofs of basic theorems in the cohomology of quasicoherent sheaves", *Rocky Mountain J. Math.* **10**:3 (1980), 637–645. MR 590225

[Lazarsfeld 1989] R. Lazarsfeld, "A sampling of vector bundle techniques in the study of linear series", pp. 500–559 in *Lectures on Riemann surfaces* (Trieste, 1987), edited by M. Cornalba et al., World Sci., Teaneck, NJ, 1989. MR 1082360 Zbl 0800.14003

[Lazarsfeld 1997] R. Lazarsfeld, "Lectures on linear series", pp. 161–219 in *Complex algebraic geometry* (Park City, UT, 1993), edited by J. Kollár, IAS/Park City Math. Ser. **3**, Amer. Math. Soc., Providence, RI, 1997. MR 1442523

[Lazarsfeld et al. 2011] R. Lazarsfeld, G. Pareschi, and M. Popa, "Local positivity, multiplier ideals, and syzygies of abelian varieties", *Algebra Number Theory* **5**:2 (2011), 185–196. MR 2833789 Zbl 1239.14035

[Li 2009] L. Li, "Wonderful compactification of an arrangement of subvarieties", *Michigan Math. J.* **58**:2 (2009), 535–563. MR 2595553 Zbl 1187.14060

[Rathmann 2016] J. Rathmann, "An effective bound for the gonality conjecture", preprint, 2016. arXiv 1604.06072

[Voisin 2002] C. Voisin, "Green's generic syzygy conjecture for curves of even genus lying on a $K3$ surface", *J. Eur. Math. Soc.* (*JEMS*) **4**:4 (2002), 363–404. MR 1941089 Zbl 1080.14525

[Voisin 2005] C. Voisin, "Green's canonical syzygy conjecture for generic curves of odd genus", *Compos. Math.* **141**:5 (2005), 1163–1190. MR 2157134 Zbl 1083.14038

[Yang 2014] D. H. Yang, "S_n-equivariant sheaves and Koszul cohomology", *Res. Math. Sci.* **1** (2014), Art. 10, 6. MR 3375645

Communicated by David Eisenbud
Received 2015-12-23 Revised 2016-06-29 Accepted 2016-08-21

ein@uic.edu Department of Mathematics, University of Illinois, Chicago, Chicago, IL 60607, United States

robert.lazarsfeld@stonybrook.edu Department of Mathematics, Stony Brook University, Stony Brook, NY 11794, United States

dhy@mit.edu Department of Mathematics, MIT, Building 2, 182 Memorial Dr., Cambridge, MA 02139, United States

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.1983

Effective cones of cycles on blowups of projective space

Izzet Coskun, John Lesieutre and John Christian Ottem

In this paper we study the cones of higher codimension (pseudo)effective cycles on point blowups of projective space. We determine bounds on the number of points for which these cones are generated by the classes of linear cycles and for which these cones are finitely generated. Surprisingly, we discover that for (very) general points the higher codimension cones behave better than the cones of divisors. For example, for the blowup X^n_r of $\mathbb{P}^n$, $n > 4$ at r very general points, the cone of divisors is not finitely generated as soon as $r > n+3$, whereas the cone of curves is generated by the classes of lines if $r \leq 2^n$. In fact, if X^n_r is a Mori dream space then all the effective cones of cycles on X^n_r are finitely generated.

1. Introduction

In recent years, the theory of cones of cycles of higher codimension has been the subject of increasing attention [Chen and Coskun 2015; Debarre et al. 2011; 2013; Fulger and Lehmann 2014a; 2014b]. However, these cones have been computed only for a very small number of examples, mainly because the current theory is hard to apply in practice. The goal of this paper is to provide some much-needed examples.

Let Γ be a set of r distinct points on $\mathbb{P}^n$. Let X^n_Γ denote the blowup of $\mathbb{P}^n$ along Γ. When Γ is a set of r very general points, we denote X^n_Γ by X^n_r. For a smooth variety Y, we write $\overline{\mathrm{Eff}}^k(Y)$ for the pseudoeffective cone of codimension-k cycles on Y, and $\overline{\mathrm{Eff}}_k(Y)$ for the pseudoeffective cone of dimension-k cycles. In this paper, we study the cones $\overline{\mathrm{Eff}}_k\left(X^n_\Gamma\right)$ when the points of Γ are either in linearly general or very general position. We also investigate the cones when Γ contains points in certain special configurations.

During the preparation of this article the first author was partially supported by the NSF CAREER grant DMS-0950951535 and NSF grant DMS-1500031, the second author was partially supported by an NSF RTG grant, and the third author was supported by RCN grant 250104.

MSC2010: primary 14C25, 14C99; secondary 14E07, 14E30, 14M07, 14N99.

Keywords: Cones of effective cycles, higher codimension cycles, blowups of projective space, Mori dream space.

Cones of positive divisors on X^n_Γ provide an important source of examples in the study of positivity. These cones are particularly attractive since they have concrete interpretations in terms of subvarieties of projective space, yet still have very complicated structure. However, even the cones of divisors on blowups of $\mathbb{P}^2$ at 10 or more points are far from well-understood, and several basic questions remain open, including the Nagata conjecture [1959] and the Segre–Harbourne–Gimigliano–Hirschowitz (SHGH) conjecture [Gimigliano 1987; Harbourne 1986; Hirschowitz 1989]. We expect the cones of higher codimension cycles on X^n_Γ to be an equally rich source of examples.

Surprisingly, these cones are simpler than one might expect. Effective cones of low-dimensional cycles are generated by the classes of linear spaces for r well into the range for which X^n_r ceases to be a Mori dream space. For example, $\overline{\mathrm{Eff}}_1(X^n_r)$ is generated by classes of lines for $r \le 2^n$ even though $\overline{\mathrm{Eff}}^1(X^n_r)$ is not finitely generated for $r \ge n+4$ when $n \ge 5$. We now describe our results in greater detail.

Definition 1.1. We say that $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is *linearly generated* if it is the cone spanned by the classes of k-dimensional linear spaces in the exceptional divisors and the strict transforms of k-dimensional linear subspaces of $\mathbb{P}^n$, possibly passing through the points of Γ. We say $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is *finitely generated* if it is a rational polyhedral cone.

Theorem 3.1. *Let Γ be a set of r points in $\mathbb{P}^n$ in linearly general position. If $r \le \max(n+2, n+n/k)$, then $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is linearly generated.*

There exist configurations of $2n+2-k$ points in linearly general position in $\mathbb{P}^n$ for which $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is not linearly generated (see Example 3.4). In particular, Theorem 3.1 is sharp for 1-cycles. We expect that this bound can be improved to $r \le 2n+1-k$, and prove this in the case that Γ is a very general configuration of points (Theorem 4.5). We obtain the following consequence.

Corollary 4.7. *If X^n_r is a Mori dream space, then $\overline{\mathrm{Eff}}_k(X^n_r)$ is finitely generated.*

In general, Mori dream spaces may have effective cones of intermediate dimensional cycles which are not finitely generated; the corollary shows that this does not happen for blowups of $\mathbb{P}^n$. A good example is [Debarre et al. 2011, Example 6.10], attributed to Tschinkel. Let X_b be the blowup of $\mathbb{P}^4$ along a smooth quartic K3 surface $Y_b \subset \mathbb{P}^3 \subset \mathbb{P}^4$. Then X_b is Fano, hence, by [Birkar et al. 2010], a Mori dream space. On the other hand, $\overline{\mathrm{Eff}}_2(X_b)$ has infinitely many extremal rays when $\overline{\mathrm{Eff}}_1(Y_b)$ does. Quartic K3 surfaces may have infinitely many (-2)-curves or even a round cone of curves. This example also shows that the property of having finitely generated higher codimension cones can fail countably many times in a family.

The bounds can be exponentially improved (at least for 1-cycles) if we assume that Γ is a set of very general points.

Proposition 4.1. *The cone $\overline{\mathrm{Eff}}_1(X^n_r)$ is linearly generated if and only if $r \le 2^n$.*

As a consequence of Proposition 4.1, we conclude that $\overline{\mathrm{Eff}}_k(X_r^n)$ is not linearly generated if $r \geq 2^{n-k+1}+k$ (Corollary 4.2). This specializes to the fact that the cone of divisors of X_r^n is not linearly generated as soon as $r > n+2$ (see Theorem 2.7).

Mukai [2004] shows that the cone of divisors of X_r^n is not finitely generated if $r \geq n+4$ and $n \geq 5$ (one needs $r \geq 9$ for $n = 2$ or 4, and $r \geq 8$ for $n = 3$). Mukai explicitly constructs infinitely many extremal divisors on $\overline{\mathrm{Eff}}^1(X_r^n)$ as the orbit of one of the exceptional divisors under the action of Cremona transformations. However, in higher codimensions it is more difficult to prove that the corresponding cones become infinite.

Many questions about cones of higher codimension cycles appear to be intractable, quickly reducing to difficult questions about cones of divisors. For example, the interesting part of the cone of curves of $\mathbb{P}^3$ blown up at 9 points is given by curves lying on the unique quadric Q through the 9 points. The blowup of Q is isomorphic to the blowup of $\mathbb{P}^2$ at 10 points, and the curves which are extremal on X_9^3 are certain K_Q-positive ones contained in Q. Hence understanding $\overline{\mathrm{Eff}}_1(X_9^3)$ requires understanding the $K_{X_{10}^2}$-positive part of $\overline{\mathrm{Eff}}_1(X_{10}^2)$, running immediately into the SHGH conjecture (see Conjecture 5.1). We are able to show this nonfiniteness only for cones of codimension-2 cycles, and then assuming the SHGH conjecture on the cone of curves of $\mathbb{P}^2$ blown up at 10 points.

Corollary 5.7. *Assume the SHGH conjecture holds for blowups of $\mathbb{P}^2$ at 10 points, then $\overline{\mathrm{Eff}}^2(X_r^n)$ is not finitely generated if $r \geq n+6$ and $n \geq 3$.*

Finally, in the last section, we compute $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ when Γ is a set of points in certain special positions. Using these computations, we show that linear and finite generation of $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ are neither open nor closed in families (see Corollaries 6.6 and 6.7). This generalizes analogous jumping behavior exhibited for divisors and Mori dream spaces to all codimensions.

The organization of the paper. In Section 2, we collect basic facts concerning the cohomology of X_Γ^n, cones of divisors, the action of Cremona transformations, and some preliminary lemmas. In Section 3, we prove Theorem 3.1 and study the linear generation of the cones $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ when Γ is a linearly general set of points. In Section 4, we study the linear generation of the cones $\overline{\mathrm{Eff}}_k(X_r^n)$. In Section 5, we prove that $\overline{\mathrm{Eff}}^2(X_r^n)$ is not finitely generated for $r \geq n+6$ assuming the SHGH conjecture. In Section 6, we discuss the cones X_Γ^n when Γ contains points in certain special configurations and study the variation of $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ in families.

2. Preliminaries

In this section, we recall basic facts about the cohomology of X_Γ^n and cones of codimension-1 cycles. We will work over the complex numbers $\mathbb{C}$.

The cohomology of X^n_Γ. Let Γ be a set of r points $p_1, \dots, p_r$ in $\mathbb{P}^n$, and let

$$\pi : X^n_\Gamma = \mathrm{Bl}_\Gamma \mathbb{P}^n \to \mathbb{P}^n$$

denote the blowup of $\mathbb{P}^n$ along Γ. Let H denote the pullback of the hyperplane class and let E_i denote the class of the exceptional divisor over p_i. The exceptional divisor E_i is isomorphic to $\mathbb{P}^{n-1}$ and $\mathcal{O}_{E_i}(E_i) \cong \mathcal{O}_{\mathbb{P}^{n-1}}(-1)$. Consequently, we have the following intersection formulas:

$$H^n = (-1)^{n-1} E_i^n = 1, \qquad H \cdot E_i = 0, \qquad E_i \cdot E_j = 0, \ i \neq j.$$

Notation 2.1. In order to simplify notation, we make the convention that H_k is the class of a k-dimensional linear space in $\mathbb{P}^n$ and $E_{i,k}$ is the class of a k-dimensional linear space contained in the exceptional divisor E_i. We then have the relations

$$H^{n-k} = H_k, \qquad (-1)^{n-k+1} E_i^{n-k} = E_{i,k}, \qquad E_i \cdot E_{i,k} = -E_{i,k-1}.$$

On X^n_Γ homological, numerical, and rational equivalence coincide. For $0 < k < n$, we write $N_k(X^n_\Gamma)$ for the $\mathbb{R}$-vector space of k-dimensional cycles on X^n_Γ, modulo numerical equivalence. Dually, $N^k(X^n_\Gamma)$ denotes the space of codimension-k cycles modulo numerical equivalence. They are both $(r+1)$-dimensional vector spaces.

A class in $N_k(X^n_\Gamma)$ is said to be pseudoeffective if it is the limit of classes of effective cycles. We write $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ for the closed convex cone in $N_k(X^n_\Gamma)$ containing pseudoeffective classes. If V is an (irreducible) k-dimensional subvariety of X^n_Γ, we write $[V]$ for the class of V in $N_k(X^n_\Gamma)$, although when confusion seems unlikely we omit the brackets.

A set of points in $\mathbb{P}^n$ is said to be *linearly general* if no $k+2$ points are contained in a linear subspace $\mathbb{P}^k \subset \mathbb{P}^n$ for $1 \leq k \leq n-1$. A claim holds for a *very general* configuration of points if it holds for all points in the complement of a countable union of proper configurations of points.

Convention 2.2. It is occasionally useful to compare the cones $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ and $\overline{\mathrm{Eff}}_k(X^m_\Delta)$, where X^n_Γ and X^m_Δ are the blowups of $\mathbb{P}^n$ and $\mathbb{P}^m$ along sets of points Γ and Δ, respectively. If $n > k$, we can identify $N_k(X^n_\Gamma)$ with the abstract vector space spanned by H_k and $E_{i,k}$ for $1 \leq i \leq r$, irrespective of n and Γ provided that Γ has cardinality r. We can thus view the cones $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ as cones in the same abstract vector space and compare the effective cones of different blowups after this identification. In the rest of the paper, we will do so without further comment.

We will often use the following easy lemma implicitly.

Lemma 2.3. *Let $Y \subset X^n_\Gamma$ be a k-dimensional subvariety*:

(1) *If $Y \subset E_i$ for some $1 \leq i \leq r$, then $[Y] = b_i E_{i,k}$ for $b_i > 0$.*

(2) *Otherwise*, $[Y] = aH_k - \sum_{i=1}^r b_i E_{i,k}$ *with $a \geq b_i \geq 0$. The coefficient b_i is equal to the multiplicity of $\pi(Y)$ in $\mathbb{P}^n$ at the point p_i.*

Proof. If $Y \subset E_i$, then Y is a subvariety of $E_i \cong \mathbb{P}^{n-1}$. Hence its class is a positive multiple of the class of a k-dimensional linear space. The linear system $H - E_i$ defines the projection from the point p_i and is a basepoint-free linear system. Hence the intersection of k general members of $H - E_i$ with Y is either empty or finitely many points. Therefore, $(H - E_i)^k \cdot [Y] = a - b_i \geq 0$. Similarly, the intersection $Y \cap E_i$ is a (possibly empty) effective cycle of dimension $k - 1$ contained in E_i. Hence by the first part of the lemma, $b_i \geq 0$. That b_i in fact coincides with the multiplicity is [Fulton 1998, Corollary 6.7.1]. □

The cones $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ satisfy a basic semicontinuity property under specialization.

Lemma 2.4. *Suppose that $V \subset \mathbb{P}^n \times T$ is a closed subvariety, flat over T, with fibers of dimension k, and let $p : T \to \mathbb{P}^n$ be a section. Then* $\mathrm{mult}_{p(t)}(V_t)$ *is an upper semicontinuous function on T.*

Proof. It suffices to prove this in the case that T has dimension 1. Let $\pi : Y \to \mathbb{P}^n \times T$ be the blowup along $p(T)$, with exceptional divisor E, and let $\widetilde{V}$ be the strict transform of V on Y. Since $\widetilde{V}$ is irreducible and dominates T, this family is flat. The intersection of a flat family of cycles with a Cartier divisor is constant in t [Fulton 1998, Proposition 10.2.1], and so $(-1)^{k+1} E^k \cdot \widetilde{V}_t$ is independent of t.

The general fiber $\widetilde{V}_t$ is irreducible, but a special fiber $\widetilde{V}_0$ may have additional components in the exceptional divisor E_0. Write $\widetilde{V}_0 = V^0 \cup \bigcup_i W_i$, where the W_i are contained in E_0. Then $(-1)^{k+1} E^k \cdot V^0 = \mathrm{mult}_{p(0)} V_0$. The class $(-1)^{k+1} E_0^k$ is a linear space in E_0, and so $(-1)^{k+1} E_0^k \cdot W_i \geq 0$. This shows that $\mathrm{mult}_{p(0)} V_0 \geq (-1)^{k+1} E^k \cdot \widetilde{V}_t = \mathrm{mult}_{p(t)} V_t$, and so the multiplicity is upper semicontinuous. □

Corollary 2.5. *Let Γ be a configuration of r distinct points on $\mathbb{P}^n$. Then*

$$\overline{\mathrm{Eff}}_k(X^n_r) \subseteq \overline{\mathrm{Eff}}_k(X^n_\Gamma).$$

Proof. Let Γ_t be a very general one-parameter family of configurations of points in $\mathbb{P}^n$ with $\Gamma_0 = \Gamma$. If a k-cycle class W is effective for very general T, then by a Hilbert scheme argument there exists a flat family $V_t \subset \mathrm{Bl}_{\Gamma_t} \mathbb{P}^n$ over T with $[V_t] = W$ for general T. Since the multiplicity of W_t can only increase at $t = 0$ by Lemma 2.4, the class W is also effective on X_Γ. □

Cones. Taking cones will be a useful method to generate interesting cycles. Let Γ' be a very general configuration of $r + 1$ points $p'_0, \dots, p'_r$ in $\mathbb{P}^{n+1}$. The projection of the points $p'_1, \dots, p'_r$ from p'_0 is then a set Γ of r very general points $p_1, \dots, p_r$ in $\mathbb{P}^n$. Suppose that V is a k-cycle on X^n_Γ, with class $aH_k - \sum_{i=1}^r b_i E_{i,k}$. The image of V in $\mathbb{P}^n$ has degree a and multiplicity b_i at the points of Γ. We may form the cone CV over V inside $\mathbb{P}^{n+1}$ with vertex at p'_0. This is a $(k+1)$-dimensional variety of degree a. It has multiplicity a at the cone point and multiplicity b_i along the lines spanned by p_i and p'_i for $1 \leq i \leq r$. In particular, the cycle CV has degree a

and multiplicities $a, b_1, \ldots, b_r$ at the points of Γ'. Its proper transform has class $aH_{k+1} - aE_{0,k+1} - \sum_{i=1}^{r} b_i E_{i,k+1}$.

We define a map $C : N_k(X_r^n) \to N_{k+1}(X_{r+1}^{n+1})$ by $C(H_k) = H_{k+1} - E_{0,k+1}$ and $C(E_{i,k}) = E_{i,k+1}$. With this definition, $C([V])$ is the class of the cone over V with vertex p'_0, and so $C(\overline{\mathrm{Eff}}_k(X_r^n)) \subseteq \overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$ with respect to the identification discussed in Convention 2.2.

The following computation of a dual cone will be useful on a number of occasions.

Lemma 2.6. *Suppose that* $\mathbf{v} = (a, -b_1, \ldots, -b_r) \in \mathbb{Z}^{r+1}$ *is a vector satisfying*

(1) $a, b_i \geq 0$,

(2) $a \geq b_i$ *for every* i,

(3) $na \geq \sum_{i=1}^{r} b_i$.

Then $\mathbf{v}$ *is a positive linear combination of the vectors* e_i, *for* $1 \leq i \leq r$, *and* $h_I = e_0 - \sum_{i\in I} e_i$ *with* $|I| \leq n$. *When* $r \geq n$, *we may assume each term has* $|I| = n$.

Proof. Note first that the vectors $h_I = e_0 - \sum_{i\in I} e_i$ with $|I| < n$ are positive linear combinations of the given vectors. We now proceed by induction on a. The case $a = 1$ is clear since by (2) $a \geq b_i$ for each i, each b_i is either 0 or 1. By (3) there are at most n nonzero b_i, and the vector is of the form claimed.

Suppose that $a > 1$. Let J be the set of indices i such that $b_i > 0$, and let $j = \min(n, |J|)$. Let I be a set of j indices $\{i_1, \ldots, i_j\}$ such that $b_{i_1} \geq \cdots \geq b_{i_j} \geq b_i$ for any $i \notin I$. Then the vector h_I is a nonnegative linear combination of the given vectors. Set $v' = (a', -b'_1, \ldots, -b'_r) = \mathbf{v} - h_I$. If $j \geq n$, then v' still satisfies all of the inequalities in question since $a' = a - 1$ and $\sum b'_i = \sum b_i - n$. If $j < n$, then in view of inequality (2), the inequality (3) can be improved to $ja \geq \sum_{i=1}^{r} b_i$. Then v' satisfies these improved inequalities. This completes the proof by induction on a. □

Lemma 2.6 implies that the cone $\overline{\mathrm{Eff}}_k(X_r^n)$ is linearly generated if and only if the class $(k+1)H_{n-k} - \sum_{i=1}^{r} E_{i,n-k}$ is nef.

The codimension-1 cones and Cremona actions. We are primarily interested in the question of when the cones of cycles on X_Γ^n are linearly or finitely generated. For cones of divisors, the answers to these questions were worked out by Castravet and Tevelev [2006] and Mukai [2004].

Theorem 2.7 [Castravet and Tevelev 2006; Mukai 2004]. *Let* Γ *be a set of* r *very general points in* $\mathbb{P}^n$. *The cone* $\overline{\mathrm{Eff}}^1(X_\Gamma^n)$ *is linearly generated if and only if* $r \leq n+2$, *and finitely generated if and only if*

(1) $n = 2$ *and* $r \leq 8$,

(2) $n = 3$ *and* $r \leq 7$,

(3) $n = 4$ *and* $r \leq 8$,

(4) $n \geq 5$ *and* $r \leq n+3$.

The characterization of cases when the effective cone of divisors is finitely generated is based on the study of the action of Cremona transformations on the pseudoeffective cone. The Coxeter group W corresponding to a T-shaped Dynkin diagram of type $T_{2,n+1,r-n-1}$ acts on $N^1(X^n_\Gamma)$ and preserves the pseudoeffective cone $\overline{\mathrm{Eff}}^1(X^n_\Gamma)$. This is an infinite group if $\frac{1}{2}+\frac{1}{n+1}+\frac{1}{r-n-1}<1$, which happens as soon as $n\geq 5$ and $r\geq n+4$. (When $n=2$ or 4, we need $r\geq 9$; while when $n=3$, we require $r\geq 8$). The orbit of a single exceptional divisor class gives an infinite set of divisors spanning other extremal rays. For details on this group action, we refer to [Dolgachev 2011] (see also [Coble 1982]).

Unfortunately, there does not seem to be a simple way to use the Cremona action to understand cones of cycles of higher codimension. The standard Cremona involution acts on X^n_r by a map with codimension 2 indeterminacy, so it does not define an action preserving the cone $\overline{\mathrm{Eff}}^k(X^n_r)$ for any $k>1$. For example, suppose that L is a line through two blown up points. The class of L defines an extremal ray on $\overline{\mathrm{Eff}}_1(X^n_r)$. The strict transform of L under a Cremona transformation centered at $n+1$ other points is a rational normal curve in $\mathbb{P}^n$ passing through $n+3$ points. If $n\geq 3$, this is no longer an extremal ray on $\overline{\mathrm{Eff}}_1(X^n_r)$, since it is in the interior of the subcone generated by classes of lines through 2 of the $n+3$ points.

One might attempt to construct interesting codimension-2 cycles on X^r_n by taking the intersections of a fixed divisor with an infinite sequence of (-1)-divisors (i.e., divisors in the orbit of E_i under the action of W) of increasing degree. However, the next lemma shows that the intersection of a (-1)-divisor with any other effective divisor on X^n_r is in the span of the classes of codimension-2 linear cycles.

Lemma 2.8. *Suppose that D_1 is a (-1)-divisor and that D_2 is an irreducible effective divisor distinct from D_1. Then $[D_1\cap D_2]$ is in the span of linear codimension-2 cycles.*

Proof. Consider the pairing on $N^1(X^n_r)$ defined by $(H,H)=n-1$, $(H,E_i)=0$, $(E_i,E_i)=-1$, and $(E_i,E_j)=0$ if $i\neq j$. This pairing is invariant under the action of W on $N^1(X^n_r)$ [Mukai 2004; Dolgachev 2011].

We first show that $(D_1,D_2)\geq 0$. Since the pairing $(\ ,\ )$ is invariant under the action of W on $N^1(X^n_r)$, we may apply a suitable element of W and assume that $D_1=E_1$ is an exceptional divisor. If $D_2=E_j$ is an exceptional divisor different from E_1, then $(D_1,D_2)=0$. Otherwise, $[D_2]=aH-\sum_{i=1}^r b_iE_i$, with $b_i\geq 0$, in which case $(D_1,D_2)=b_i\geq 0$.

For the second part, write $D_1=aH-\sum_{i=1}^r b_iE_i$ and $D_2=cH-\sum_{i=1}^r d_iE_i$. That $(D_1,D_2)\geq 0$ yields

$$(n-1)ac\geq\sum_{i=1}^r b_id_i.$$

By Lemma 2.6, this means that the codimension-2 cycle $[D_1 \cap D_2] = acH - \sum_{i=1}^{r} b_i d_i$ is contained in the span of linear cycles. □

Easy lemmas. Here we collect a couple of geometric lemmas that we will use repeatedly.

Lemma 2.9. *Suppose that E is an effective divisor and that P is a nef divisor. If Y is an irreducible, effective variety of dimension k which is not contained in E, then $P^{k-1} \cdot E \cdot Y \geq 0$.*

Proof. The intersection $E \cdot Y$ is a (possibly empty) cycle of dimension $k-1$ by assumption. Since P is nef, it follows that $P^{k-1} \cdot E \cdot Y \geq 0$. □

Lemma 2.10. *Let $Y \subset X^n_\Gamma$ be an irreducible variety of dimension k, not contained in any exceptional divisor E_i, with class $aH_k - \sum_{i=1}^{e} b_i E_{i,k}$. If $b_i + b_j > a$ for two indices $i \neq j$, then Y contains the line through p_i and p_j with multiplicity at least $b_i + b_j - a$.*

Proof. The base locus of the linear system $|H - E_i - E_j|$ is the line $l_{i,j}$ spanned by p_i and p_j. Consequently, the intersection $(H - E_i - E_j)^{k-1} \cdot Y$ is an effective 1-cycle Z. Express

$$Z = \alpha l_{i,j} + u,$$

where u is a 1-cycle not containing $l_{i,j}$. Since

$$-\alpha \leq (H - E_i - E_j) \cdot Z = a - b_i - b_j < 0,$$

we conclude that $\alpha \geq b_i + b_j - a$. Hence Y must have multiplicity at least $b_i + b_j - a$ at every point of $l_{i,j}$. □

3. Points in linearly general position

In this section, we study $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ when the cardinality of Γ is small and the points of Γ are in linearly general position. Our main theorem is the following.

Theorem 3.1. *Let Γ be a set of r points in $\mathbb{P}^n$ in linearly general position. If*

$$r \leq \max\left(n+2, n+\frac{n}{k}\right),$$

then $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is linearly generated.

The proof will be by induction on k and n. We first single out the case $k = 1$.

Lemma 3.2. *Let Γ be a set of $r \leq 2n$ points in $\mathbb{P}^n$ in linearly general position. Then $\overline{\mathrm{Eff}}_1(X^n_\Gamma)$ is linearly generated.*

Proof. Let B be an irreducible curve. By Lemma 2.3, we may assume that B is not contained in any of the exceptional divisors and has class $aH_1 - \sum_{i=1}^r b_i E_{i,1}$ with $a \geq b_i \geq 0$. Any $r \leq 2n$ points in linearly general position are cut out by quadrics [Harris 1995, Lecture 1]. Consequently, there is a quadric whose proper transform has class $[Q] = 2H - \sum_{i=1}^r E_i$ in X_Γ^n and does not contain B. Hence B has nonnegative intersection with Q and satisfies $2a \geq \sum_{i=1}^r b_i$. By Lemma 2.6, the class of B is spanned by the classes of lines. □

Next, we study the case when $r \leq n+1$. In this case, X_Γ^n is toric and the effective cones are generated by torus-invariant cycles (see e.g., [Li 2015, Proposition 3.1]). For the reader's convenience we will give a simple independent proof.

Lemma 3.3. *Let Γ be a set of $r \leq n+1$ linearly general points in $\mathbb{P}^n$. The cone $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ is linearly generated for any k.*

Proof. Let $\Gamma' \subset \Gamma$ be two sets with cardinality r and $n+1$, respectively. Then the proper transform of any effective cycle in $X_{\Gamma'}^n$ is an effective cycle in X_Γ^n. Consequently, if $\overline{\mathrm{Eff}}_k(X_\Gamma^n)$ is linearly generated, then $\overline{\mathrm{Eff}}_k(X_{\Gamma'}^n)$ is also linearly generated. Hence, without loss of generality, we may assume that $r = n+1$. Let Y be an irreducible k-dimensional variety in X_Γ^n with class

$$[Y] = aH_k - \sum_{i=1}^{n+1} b_i E_{i,k}.$$

By Lemma 2.3, we may assume that Y is not contained in an exceptional divisor and that $a \geq b_i \geq 0$. We proceed by induction on k and n. After reordering the points, we may assume $b_1 \geq b_2 \geq \cdots \geq b_{n+1}$. Let L be the proper transform of the $\mathbb{P}^{n-1}$ spanned by the first n points. First, suppose Y is contained in L. Since L is isomorphic to the blowup of $\mathbb{P}^{n-1}$ in n points, by induction on n with base case Theorem 2.7, we conclude that the class of Y is in the span of linear spaces. Otherwise, $Y \cap L$ is an effective cycle of dimension $k-1$ in L. Write $H_{L,k-1}$ and $E_{L,i,k-1}$ for the restriction of H_k to L and the $(k-1)$-dimensional linear space in the exceptional divisor $E_{L,i}$ of the blowup of L at p_i. Then we have

$$[Y \cap L] = aH_{L,k-1} - \sum_{i=1}^{n} b'_i E_{L,i,k-1},$$

with $b'_i \geq b_i$. By induction on n with base case Lemma 3.2, $Y \cap L$ is in the span of linear spaces. In particular, $ka \geq \sum_{i=1}^n b'_i$. Hence $(k+1)a \geq \sum_{i=1}^{n+1} b_i$. By Lemma 2.6, the class of Y is in the span of linear spaces. This concludes the proof. □

We can now complete the proof of Theorem 3.1.

Proof of Theorem 3.1. We preserve the notation from the proof of Lemma 3.3 and argue similarly. It suffices to check the result in the case $r = \max(n+2, n+n/k)$.

Suppose that Y is an irreducible k-dimensional variety on X^n_Γ with class

$$[Y] = aH_k - \sum_{i=1}^{r} b_i E_{i,k}.$$

We may assume that Y is not contained in an exceptional divisor and, by reordering the points, we have that

$$a \geq b_1 \geq \cdots \geq b_r \geq 0.$$

Let L be the $\mathbb{P}^{n-1}$ passing through the points $p_1, \ldots, p_n$. If Y is contained in L, then its class is linearly generated by Lemma 3.3. Otherwise, $Y \cap L$ is an effective cycle of dimension $k-1$ with class

$$[Y \cap L] = aH_{L,k-1} - \sum_{i=1}^{n} b'_i E_{L,i,k-1},$$

with $b'_i \geq b_i$. This class and hence $aH_{L,k-1} - \sum_{i=1}^{n} b_i E_{L,i,k-1}$ is linearly generated by Lemma 3.3. Therefore, it can be written as a combination of linear classes $H_{L,k-1} - \sum_{|I|=k} E_{L,i,k-1}$ and $E_{L,i,k-1}$,

$$\sum_{j=1}^{a} \alpha_j \left(H_{L,k-1} - \sum_{|I|=k} E_{L,i,k-1} \right) + \sum_{j=1}^{n} \beta_j E_{L,j,k-1}.$$

Each of the classes in this sum is effective, with those on the left the classes of $\mathbb{P}^{k-1}$ through k of the points in L. By taking cones over these classes, we obtain a $\mathbb{P}^k$ on X, passing through an additional one of the points p_i with $i > n$. Since there are a planes available, if $\sum_{i=n+1}^{r} b_i \leq a$, the class Y can be expressed as a sum of linear cycles.

Observe that

$$ak \geq \sum_{i=1}^{n} b_i \geq nb_n, \qquad \text{and so} \qquad b_j \leq b_n \leq \frac{ak}{n} \text{ for } j \geq n.$$

This implies that if $(r-n)k/n \leq 1$ or equivalently if $r \leq n + n/k$, the classes of all effective cycles are in the span of the classes of linear spaces.

If $k \leq n/2$, then $n+2 \leq n+n/k$ and the theorem is proved. If $k > n/2$, then $n+1 < n+n/k < n+2$ and we need to settle the case $r = n+2$. There is a rational normal curve through any $n+3$ points in linearly general position in $\mathbb{P}^n$ [Harris 1995, Lecture 1]. Consequently, given an effective divisor D, there exists a rational normal curve C containing the points but not contained in D. Hence $C \cdot D \geq 0$ and all effective divisors satisfy $na \geq \sum_{i=1}^{n+2} b_i$. We recover the linear generation result of Theorem 2.7. By Lemma 3.2, the curve classes are also linearly generated. By induction assume that for all $m < n$ and all $k < m$, the effective cone of k cycles of the blowup of $\mathbb{P}^m$ in $m+2$ linearly general points is linearly generated. We carry out the inductive step for $\mathbb{P}^n$. Let Y, L be as above. By Lemma 3.3, we may assume that Y is not contained in L. If $b_{n+1}+b_{n+2} \le a$, then we already proved that the class of Y is linearly generated. If $b_{n+1}+b_{n+2} > a$, then, by Lemma 2.10, Y contains the line $l_{n+1,n+2}$ spanned by p_{n+1} and p_{n+2} with multiplicity at least $b_{n+1}+b_{n+2}-a$. Let p_0 denote the point of intersection $L \cap l_{n+1,n+2}$. Then the proper transform of $L \cap Y$ is an effective cycle in the blowup of L in $p_0, p_1, \dots, p_n$ with class

$$aH_{L,k-1} - (b_{n+1}+b_{n+2}-a+c)E_{L,0,k-1} - \sum_{i=1}^{n} b_i E_{L,i,k-1},$$

where $c \ge 0$. By induction on n, this class is linearly generated. Hence

$$ka \ge b_{n+1}+b_{n+2}-a+c+\sum_{i=1}^{n} b_i, \qquad \text{therefore} \qquad (k+1)a \ge \sum_{i=1}^{n+2} b_i.$$

By Lemma 2.6, the class of Y is linearly generated. This concludes the proof. □

Example 3.4. Lemma 3.2 is sharp in the sense that there exist sets Γ of $r > 2n$ points in general linear position such that $\overline{\mathrm{Eff}}_1(X^n_\Gamma)$ is not linearly generated. For example, let Γ be $r > 2n$ points on a rational normal curve C in $\mathbb{P}^n$. Points on a rational normal curve are in general linear position [Harris 1995]. Then the class of the proper transform of C is

$$nH_1 - \sum_{i=1}^{r} E_{i,1}.$$

Since $r > 2n$, this class cannot be in the span of the classes of lines. In the next section we will see that we can improve the bounds for linear generation exponentially if instead of assuming that Γ is linearly general we assume Γ is a set of very general points in $\mathbb{P}^n$.

More generally, let Y be the cone over a rational normal curve of degree $n-k+1$ with vertex V a $\mathbb{P}^{k-2}$. Let Γ be the union of a set of $k-1$ general points $p_1, \dots, p_{k-1}$ in V and a set of $r-k+1$ general points $p_k, \dots, p_r$ on Y. Then Γ is in general linear position. The class of the proper transform of Y is

$$(n-k+1)H_k - \sum_{i=1}^{k-1}(n-k+1)E_{i,k} - \sum_{i=k}^{r} E_{i,k},$$

which cannot be in the span of linear spaces if $r > 2n-k+1$. Consequently, we conclude the following.

Proposition 3.5. *There exist sets Γ of $r > 2n-k+1$ points in general linear position in $\mathbb{P}^n$ such that $\overline{\mathrm{Eff}}_k\big(X^n_\Gamma\big)$ is not linearly generated.*

In view of Proposition 3.5, it is natural to ask whether the bound in Theorem 3.1 can be improved to $r \leq 2n-k+1$.

Question 3.6. Assume that Γ is a set of r linearly general points in $\mathbb{P}^n$ such that

$$\max\Big(n+2, n+\frac{n}{k}\Big) < r \leq 2n-k+1.$$

Is $\overline{\mathrm{Eff}}_k\big(X^n_\Gamma\big)$ linearly generated?

The answer is affirmative for curves and divisors. We will shortly check that for 2-cycles in $\mathbb{P}^4$ the answer is also affirmative. In Theorem 4.5 we will see that the answer is also affirmative if the points are very general. In view of this evidence, we expect the answer to Question 3.6 to be affirmative.

Remark 3.7. The dimension of the space $\mathcal{S}_{n-k,k+1}(\mathbb{P}^n)$ of scrolls of dimension $n-k$ and degree $k+1$ in $\mathbb{P}^n$ is

$$2n+2nk-k^2-2$$

[Coskun 2008, Lemma 2.4]. There are scrolls in $\mathcal{S}_{n-k,k+1}(\mathbb{P}^n)$ passing through $2n-k+2$ points (see [Coskun 2006] for the surface case). Hence the family of scrolls passing through $2n-k+1$ points covers $\mathbb{P}^n$. By Lemma 2.6, an affirmative answer to Question 3.6 is equivalent to the statement that every effective k-dimensional cycle intersects the proper transform of a scroll passing through the $2n-k+1$ points nonnegatively.

Question 3.8. Let Γ be $2n-k+1$ linearly general points in $\mathbb{P}^n$. For every effective k-cycle Y in X^n_Γ, does there exist a scroll S of dimension $n-k$ and degree $k+1$ such that the proper transform S in X^n_Γ intersects Y in finitely many points?

By Remark 3.7, an affirmative answer to Question 3.8 implies an affirmative answer to Question 3.6.

Effective 2-cycles on the blowup of $\mathbb{P}^4$ at 7 points. We now verify that the answer to Question 3.6 is affirmative for two-cycles in $\mathbb{P}^4$. The argument is subtle because we need to verify linear generation for *every* configuration of 7 points in linear general position, rather than just very general configurations of points.

Theorem 3.9. *Let Γ be 7 linearly general points on $\mathbb{P}^4$. Then the cone $\overline{\mathrm{Eff}}_2\big(X^4_\Gamma\big)$ is linearly generated.*

Proof. There is a unique rational normal quartic curve R containing 7 linearly general points in $\mathbb{P}^4$ [Harris 1995]. The secant variety, $\mathrm{Sec}(R)$, to R is a cubic hypersurface which has multiplicity two along R. Hence its proper transform $\overline{\mathrm{Sec}}(R)$

in X^4_Γ has class $3H - \sum_{i=1}^7 2E_i$. In fact, this secant variety is a (-1)-divisor on X^4_Γ; it is in the Cremona orbit of one of the exceptional divisors.

Let Y be an irreducible surface in X^4_Γ. Without loss of generality, we may assume that Y is not contained in an exceptional divisor and has class $aH_2 - \sum_{i=1}^7 b_i E_{i,2}$ with $a \geq b_i \geq 0$. First, suppose that Y is not contained in $\overline{\mathrm{Sec}}(R)$. The class of a quadric $[Q] = 2H - \sum_{i=1}^7 E_i$ is nef, and so by Lemma 2.9 we have

$$Y \cdot Q \cdot \overline{\mathrm{Sec}}(R) = 6a - \sum_{i=1}^{7} 2b_i \geq 0.$$

Lemma 2.6 implies that $[Y]$ is in the span of the classes of planes.

We are reduced to showing that if $Y \subset \overline{\mathrm{Sec}}(R)$, then $[Y]$ is in the span of the classes of planes. Let $\mathcal{S}_3$ denote the space of cubic surface scrolls containing the points of Γ. We will show the following.

Theorem 3.10. *The proper transform $\bar{S}$ of a general member $S \in \mathcal{S}_3$ intersects $\overline{\mathrm{Sec}}(R)$ in an irreducible curve B whose projection to* $\mathrm{Sec}(R)$ *is a degree* 9 *curve with multiplicity two at the points of* Γ. *Furthermore, the curve B can be made to pass through a general point of* $\overline{\mathrm{Sec}}(R)$.

Assume Theorem 3.10. Let $p \in \overline{\mathrm{Sec}}(R)$ be a general point not contained in Y. Hence an irreducible curve B passing through p intersects Y in finitely many points. Let $\bar{S}$ be the proper transform of a scroll $S \in \mathcal{S}_3$ containing p and intersecting $\overline{\mathrm{Sec}}(R)$ in an irreducible curve. We conclude that $\bar{S}$ and Y intersect in finitely many points, hence their intersection number is nonnegative. Therefore,

$$[\bar{S}] \cdot [Y] = \left(3H_2 - \sum_{i=1}^{7} E_{i,2}\right) \cdot \left(aH_2 - \sum_{i=1}^{7} b_i E_{i,2}\right) = 3a - \sum_{i=1}^{7} b_i \geq 0.$$

By Lemma 2.6, we conclude that $[Y]$ is in the span of the classes of planes.

There remains to prove Theorem 3.10, which we will do via a series of claims. We first set some notation.

Notation 3.11. Let $l_{i,j}$ denote the line spanned by $p_i, p_j \in \Gamma$ and let $\Pi_{i_1,\dots,i_l}$ denote the linear space spanned by $p_{i_1}, \dots, p_{i_l} \in \Gamma$. Let $\Gamma_{i_1,\dots,i_l}$ denote the set of points $p_{i_1}, \dots, p_{i_l}$. Let l be the line of intersection $\Pi_{1,2,3,4} \cap \Pi_{5,6,7}$ and, for $5 \leq i < j \leq 7$, let $z_{i,j}$ denote the point of intersection $\Pi_{1,2,3,4} \cap l_{i,j}$. Since the points are in linearly general position, the line l does not intersect the lines $l_{i,j}$ for $1 \leq i < j \leq 4$ and intersects the planes $\Pi_{i,j,k}$, for $1 \leq i < j < k \leq 4$, in a unique point different from $z_{i,j}$.

Next, we recall a compactification $\bar{\mathcal{S}}_3$ of $\mathcal{S}_3$. Every irreducible cubic scroll induces a degree 3 rational curve in the Grassmannian $\mathbb{G}(1,4)$ of lines in $\mathbb{P}^4$. We can compactify the space of degree 3 rational curves in $\mathbb{G}(1,4)$ via the Kontsevich moduli space. Hence we can take the closure of $\mathcal{S}_3$ in the Kontsevich moduli space (see [Coskun 2006, §3] for details). More precisely, let $\overline{\mathcal{M}}_{0,7}(\mathbb{G}(1,4),3)$ denote the Kontsevich moduli space of 7-pointed genus-0 maps of degree 3 to $\mathbb{G}(1,4)$. It is equipped with 7 evaluation morphisms $\mathrm{ev}_i : \overline{\mathcal{M}}_{0,7}(\mathbb{G}(1,4),3) \to \mathbb{G}(1,4)$, $1 \le i \le 7$. Define

$$\bar{\mathcal{S}}_3 = \bigcap_{i=1}^{7} \mathrm{ev}_i^{-1}(\Sigma_3(p_i)),$$

where $\Sigma_3(p_i)$ denotes the Schubert variety of lines containing p_i.

Claim 3.12. *The space $\bar{\mathcal{S}}_3$ is irreducible of dimension* 4.

Proof. The locus $\overline{\mathcal{T}} = \cap_{i=1}^{4} \mathrm{ev}_i^{-1}(p_i)$ in the Kontsevich moduli space $\overline{\mathcal{M}}_{0,4}(\mathbb{P}^3,3)$ of 4-pointed genus 0 maps of degree 3 to $\mathbb{P}^3$ provides a compactification of the space of twisted cubic curves in $\Pi_{1,2,3,4}$ containing $\Gamma_{1,2,3,4}$. Since $\overline{\mathcal{M}}_{0,4}(\mathbb{P}^3,3)$ is irreducible of dimension 16, every component of $\overline{\mathcal{T}}$ has dimension at least 4.

If a twisted cubic T is irreducible, then any finite set of points on T is linearly general. Furthermore, given 6 linearly general points in $\mathbb{P}^3$, there is a unique twisted cubic curve containing them. Consider the incidence correspondence $I = \{(T, q_1, q_2) | q_1, q_2 \in T\}$, where T is a twisted cubic curve containing the set of points $\Gamma_{1,2,3,4}$ and q_1, q_2 are points such that $\Gamma_{1,2,3,4} \cup \{q_1, q_2\}$ are in linearly general position. The incidence correspondence I is irreducible of dimension 6 since it is isomorphic to an open subset of $\mathbb{P}^3 \times \mathbb{P}^3$. It dominates the space of twisted cubic curves containing $p_1, \dots, p_4$ via the first projection. Since the fibers of the first projection are two-dimensional, we conclude that the space of irreducible twisted cubics containing $\Gamma_{1,2,3,4}$ is irreducible of dimension 4.

Since there are no connected curves of degree two or one containing 4 points in linearly general position in $\mathbb{P}^3$, any map in $\overline{\mathcal{T}}$ is birational to its image. If there is a reducible curve of degree 3 containing $\Gamma_{1,2,3,4}$, either a degree two curve must contain 3 of the points or a line must contain two of the points. In either case, it is easy to see that there is a 3-dimensional family of reducible cubics containing $\Gamma_{1,2,3,4}$. Hence these cannot form a component of $\overline{\mathcal{T}}$ and $\overline{\mathcal{T}}$ is irreducible.

Furthermore, 2 additional points q_1 and q_2 impose independent conditions on twisted cubics unless they are coplanar with 3 of the points in $\Gamma_{1,2,3,4}$ or 1 of the points is collinear with 2 of the points in $\Gamma_{1,2,3,4}$. If q_1 is collinear with p_1 and p_2, then there is a 1-parameter family of reducible cubics containing the line $l_{1,2}$. Similarly, if q_1 and q_2 are in $\Pi_{1,2,3}$ but no 4 of the points are collinear, then there is a 1-parameter family of reducible cubics containing the conic through $\Gamma_{1,2,3} \cup \{q_1, q_2\}$. If q_1 and q_2 are collinear with p_1 and p_2, there is a 3-parameter family of reducible cubics containing $l_{1,2}$. Recall that $l = \Pi_{1,2,3,4} \cap \Pi_{5,6,7}$. In particular, the subset of $\overline{\mathcal{T}}$ that parametrizes twisted cubics incident or secant to l has dimension 3 or 2 respectively, since any pair of distinct points impose independent conditions on twisted cubics. Similarly, the locus of twisted cubics in $\overline{\mathcal{T}}$ passing through $z_{5,6}$ has dimension 2.

Since $\overline{\mathcal{M}}_{0,7}(\mathbb{G}(1,4),3)$ is irreducible of dimension 25 [Coskun 2006, §2], every irreducible component of $\bar{S}$ has dimension at least 4. Let T be a twisted cubic curve containing $\Gamma_{1,2,3,4}$, not secant to the line l, and not containing the points $z_{5,6}$, $z_{5,7}$ and $z_{6,7}$. Then there is a unique cubic scroll S containing T and passing through p_5, p_6, p_7 [Coskun 2006, Example A1]. Briefly, take a general $\mathbb{P}^3$ containing $\Pi_{5,6,7}$. This $\mathbb{P}^3$ intersects T in 3 points r_1, r_2, r_3. There is a unique twisted cubic curve T' containing r_1, r_2, r_3 and $\Gamma_{5,6,7}$. The curves T and T' are both isomorphic to $\mathbb{P}^1$ and there is a unique isomorphism ϕ taking $r_i \in T$ to $r_i \in T'$. Then the surface $S_{T,T'}$ swept out by lines joining the points that correspond under ϕ is the unique cubic scroll containing T and $\Gamma_{5,6,7}$. If T contains the point $z_{5,6}$ or is secant to the line l, then there is a 1-parameter family of choices for T'. Once we fix T and T', the scroll is uniquely determined by a similar construction. Since the locus of T containing $z_{5,6}$ or secant to l has codimension 2, this locus cannot form a component of $\bar{S}_3$. Finally, reducible cubic surfaces containing Γ must contain a plane through 3 of the points and a quadric surface through the remaining 4 points. There is a 2-dimensional family of such surfaces and they do not give rise to a component in $\bar{S}_3$ (see [Coskun 2006]). We conclude that $\bar{S}_3$ is irreducible of dimension 4. □

Claim 3.13. *There exists a dense open set $U \subset \bar{S}_3$ such that $S \not\subset \mathrm{Sec}(R)$ for $S \in U$. Furthermore, S can be made to pass through a general point of* $\mathrm{Sec}(R)$.

Proof. It suffices to exhibit one $S \in \bar{S}_3$ such that $S \not\subset \mathrm{Sec}(R)$. Given 7 points in general linear position and 2 general additional points, [Coskun 2006, Example A1] shows that there are 2 cubic scrolls containing these nine points. In particular, if we take one of the two additional points outside $\mathrm{Sec}(R)$, we obtain a scroll not contained in $\mathrm{Sec}(R)$. Furthermore, a general twisted cubic in $\Pi_{1,2,3,4}$ containing $\Gamma_{1,2,3,4}$ intersects $\mathrm{Sec}(R)$ in a another point q. Consequently, the construction in the proof of Claim 3.12 exhibits a cubic scroll containing q and not contained in $\mathrm{Sec}(R)$. Since the space $\bar{S}_3$ is irreducible, the general scroll containing a general point of $\mathrm{Sec}(R)$ and Γ will not be contained in $\mathrm{Sec}(R)$. □

Claim 3.14. *There exists a dense open set $U \subset \bar{S}_3$ such that for $S \in U$*:

1. *The intersection $\bar{S} \cap \overline{\mathrm{Sec}}(R) \cap E_i$ is a finite set of points in X^4_Γ for every $1 \le i \le 7$.*
2. *The scroll S does not contain any lines $l_{i,j}$ for $1 \le i < j \le 7$.*
3. *The scroll S does not contain any conics through 3 of the points in Γ.*
4. *The scroll S does not contain a twisted cubic curve through 5 of the points of Γ.*
5. *The scroll S does not contain the rational normal quartic R.*
6. *The scroll S does not contain a quintic curve double at one of the points of Γ and passing through the others.*
7. *The directrix of the scroll does not contain any of the points in Γ.*

Proof. Since each of these conditions are closed conditions and $\bar{\mathcal{S}}_3$ is irreducible, it suffices to exhibit one element $S \in \bar{\mathcal{S}}_3$ satisfying each condition. For (1), there exists a twisted cubic containing $\Gamma_{1,2,3,4}$ with any tangent line at p_1 (for example, the reducible twisted cubic consisting of any line through p_1 and a conic through $\Gamma_{2,3,4}$). Hence the tangent spaces to the scrolls at p_1 sweep out E_1 and there exists S such that $\bar{S} \cap E_1 \not\subset \overline{\operatorname{Sec}}(R)$. By permuting indices, we conclude (1).

For (2) and (3), take the scroll $S_{T,T'}$ constructed in the proof of Claim 3.12. Since $\Pi_{1,2,3,4} \cap S_{T,T'} = T$, this scroll does not contain any of the linear $l_{i,j}$ with $1 \le i < j \le 4$ or any conic passing through any of the three points in $\Gamma_{1,2,3,4}$. By permuting indices, we conclude (2) and (3).

Since a twisted cubic curve spans a $\mathbb{P}^3$ and the points are in linearly general position (4) is clear. For (5), (6) and (7), it is more convenient to exhibit a reducible scroll satisfying these properties. Let S be the union of the plane $\Pi_{5,6,7}$ and a general quadric surface Q containing l and $\Gamma_{1,2,3,4}$. After choosing a point of l, this surface determines a point p of $\bar{\mathcal{S}}_3$ [Coskun 2006]. The directrix line is then the unique line on the quadric Q intersecting l at p. Hence (7) holds. Since R is irreducible and nondegenerate, it cannot be contained in this surface. Suppose there is a quintic curve F in S containing Γ and double at p_1. Since p_5, p_6, p_7 are not collinear, F must intersect $\Pi_{5,6,7}$ in a curve of degree at least 2. Hence F intersects Q in a curve of degree 3 containing $\Gamma_{2,3,4}$ and double at p_1. Any cubic double at p_1 must contain the line of ruling through p_1. Since $\Gamma_{1,2,3,4}$ are linearly general there cannot be a degree 2 curve through these points on Q. After permuting indices, we conclude (6) holds. □

Claim 3.15. *There exists a dense open set $U \subset \bar{\mathcal{S}}_3$ such that for $S \in U$ the intersection $S \cap \operatorname{Sec}(R)$ is an irreducible degree 9 curve double along Γ.*

Proof. By Claims 3.13 and 3.14, we can find a scroll $S \not\subset \operatorname{Sec}(R)$ satisfying the conclusions of Claim 3.14. The intersection $S \cap \operatorname{Sec}(R)$ is a curve B of degree 9 double along Γ. We need to show that B is irreducible. Recall that a smooth cubic scroll is isomorphic to the blowup of $\mathbb{P}^2$ at a point. Its Picard group is generated by the directrix e (the curve of self-intersection -1) and the class of a fiber line f. The intersection numbers are

$$e^2 = -1, \quad e \cdot f = 1, \quad f^2 = 0.$$

The effective cone is spanned by e and f. The canonical class is $-2e - 3f$ and the class of B is $3e + 6f$. The degree of a curve $ke + mf$ is $k + m$. If $k > m$, then any representative contains e with multiplicity $k - m$. By adjunction, the arithmetic genus of a curve in the classes $e + mf$, $2e + mf$, and $3e + mf$ are 0, $m - 2$, and $2m - 5$, respectively.

It is now straightforward, but somewhat tedious, to check that B cannot be

class	the number of double points of Γ	the number of remaining points of Γ, respectively	reason
e	0	0	Claim 3.14 (7)
$e+f$	0	2	Claim 3.14 (3,2)
$e+2f$	0 or 1	4 or 1	Claim 3.14 (4,3,2)
$e+3f$	0, 1 or 2	6, 3 or 0	Claim 3.14 (7,5,4,3,2)
$e+4f$	0, 1 or 2	7, 5 or 2	Claim 3.14 (7,5,4,3,2)

Table 1. Possible curves with class $e+mf$.

reducible. Indeed, suppose that B is reducible. Write $B = B_1 \cup B_2$, where the class of B_1 is $ke+mf$ with $2 \le k \le 3$ and assume that B_1 does not contain any fibers as components. Furthermore, if $k = 3$, we may assume that B_1 is irreducible. Otherwise, we can regroup a component with class $e+m'f$ with B_2. Then the class of B_2 is $(3-k)e+(6-m)f$ and every fiber component of B is included in B_2. By Claim 3.14 (2), a curve with class mf can be double at most in $0 \le d \le m/2$ points of Γ in which case it can contain at most $m-2d$ of the remaining points of Γ. We tabulate the possibilities for curves with class $e+mf$, see Table 1.

First, suppose B_1 has class $3e+mf$. By assumption, it is irreducible and by arithmetic genus considerations can have at most $2m-5$ nodes. On the other hand, B_2 can pass through at most $6-m$ of the points. We have $2m-5+6-m = m+1 < 7$ if $m < 6$. Hence such a curve cannot be double at all the points of Γ.

We may therefore assume that the class of B_1 is $2e+mf$ and the class of B_2 is $e+(6-m)f$. If B_1 is reducible, then it can have at most 2 components with classes $e+m_1f$ and $e+m_2f$. An inspection of Table 1 shows that it is not possible to make B double at all points of Γ. If B_1 is irreducible, then $m \ge 2$ and its arithmetic genus is $m-2$. Hence the maximal number of double points on B_1 is $m-2$. If $m = 2$, then B_1 can contain at most 6 of the points of Γ by Claim 3.14 (5) and it is smooth at those points. Hence B cannot be made double at all points of Γ by the last line of the table. If $m = 3$ and B_1 has a double point, then by Claim 3.14 (6) B_1 contains at most 5 other points of Γ. By the second to last row of the table, B cannot be double at all points of Γ. If $m \ge 4$, an easy inspection of the first three rows of the table show that B can have at most 6 double points. Hence B is irreducible. □

This concludes the proof of Theorem 3.10 and consequently of Theorem 3.9.

4. Nonlinearly generated cones

Recall that X^n_r denotes the blowup of $\mathbb{P}^n$ in r very general points. In this section, we study the cones of effective cycles on X^n_r. Our first result completely characterizes when the cone of curves is linearly generated.

Proposition 4.1. *The cone* $\overline{\mathrm{Eff}}_1(X_r^n)$ *is linearly generated if and only if* $r \le 2^n$.

Proof. We first observe that the linear system of quadrics through 2^n very general points is nef. Choose n general quadrics $Q_1, \dots, Q_n$ in $\mathbb{P}^n$. By Bertini's theorem, the intersection of these quadrics is a set of 2^n points in $\mathbb{P}^n$. Let X_0 be the blowup of $\mathbb{P}^n$ at these points. We claim that $D = 2H - \sum_{i=1}^{2^n} E_i$ is nef on X_0. Note that the proper transforms of Q_i have class D and D has positive degree on curves contained in exceptional divisors E_i. Since the intersection $Q_1 \cap \dots \cap Q_n$ is finite, if B is a curve on X_0 not contained in an exceptional divisor, there is a quadric Q_i whose proper transform does not contain B. Consequently, $D \cdot B \ge 0$ and D is nef. By [Lazarsfeld 2004, Proposition 1.4.14], $2H - \sum_{i=1}^{2^n} E_i$ is nef for very general configurations of 2^n points as well. We conclude that if $r \le 2^n$, an effective curve class in X_r^n satisfies the inequalities in the assumptions of Lemma 2.6, and so every curve class is a linear combination of classes of lines.

The top self-intersection of the class $Q = 2H - \sum_{i=1}^{r} E_i$ on X_r^n is given by $2^n - r$. Hence if $r > 2^n$, the top self-intersection of Q is negative and Q cannot be nef by Kleiman's Theorem [Lazarsfeld 2004, Theorem 1.4.9]. Suppose the class of every effective curve is in the span of the classes of lines. The cone generated by the classes of lines is a closed cone. Hence the effective and the pseudoeffective cones coincide. Since every line has nonnegative intersection with Q, we conclude that Q is nef. This contradiction shows that there must exist effective curves whose classes are not spanned by the classes of lines. □

Corollary 4.2. *If* $r \ge 2^{n-k+1} + k$, *then* $\overline{\mathrm{Eff}}_k(X_r^n)$ *is not linearly generated.*

Proof. Let Γ be a set of r very general points. Project Γ from the first $k-1$ points $p_1, \dots, p_{k-1}$ and let Γ' be the set of points in $\mathbb{P}^{n-k+1}$ consisting of the images of the remaining points. Then Γ' is a set of $r-k+1$ very general points in $\mathbb{P}^{n-k+1}$. If $r-k+1 > 2^{n-k+1}$, the cone $\overline{\mathrm{Eff}}_1(X_{r-k+1}^{n-k+1})$ is not linearly generated by Proposition 4.1. Fix a 1-cycle B with class $aH_1 - \sum_{i=k}^{r} b_i E_{i,1}$ that is not in the span of linear spaces. In particular, $2a < \sum_{i=k}^{r} b_i$. Then the class

$$aH_k - \sum_{i=1}^{k-1} aE_{i,k} - \sum_{i=k}^{r} b_i E_{i,k}$$

is represented in X_r^n by the proper transform of the cone over B with vertex the span of $p_1, \dots, p_{k-1}$. The resulting k-cycle is not in the span of k-dimensional linear spaces since $(k+1)a < (k-1)a + \sum_{i=k}^{r} b_i$. □

Question 4.3. If $r < 2^{n-k+1} + k$, is $\overline{\mathrm{Eff}}_k(X_r^n)$ linearly generated?

Remark 4.4. The answer to Question 4.3 is affirmative for curves and divisors. For cycles of intermediate dimension, we do not know any examples with $r = 2^{n-k+1} + k - 1$ where the cone is linearly generated.

There has been a great deal of interest in the construction of cycles that are nef but not pseudoeffective. Such cycles were constructed on abelian varieties in [Debarre et al. 2011], and on hyperkähler varieties in [Ottem 2015]. If Question 4.3 has an affirmative answer, this would give many examples of nef classes that are not pseudoeffective. For example, if $\overline{\mathrm{Eff}}_3(X^6_r)$ is linearly generated for $16 < r < 19$, then the class $4H_3 - \sum_{i=1}^r E_{i,3}$ would be nef but not pseudoeffective; indeed, the self-intersection of this class is negative.

We can, however, give a linear bound.

Theorem 4.5. *The cone $\overline{\mathrm{Eff}}_k(X^n_r)$ is linearly generated if $r \le 2n - k + 1$.*

Proof. The theorem is true for $k = 1$ by Proposition 4.1 and for divisors by Theorem 3.1. We will prove the general case by induction on n. Assume that the theorem is true for $\overline{\mathrm{Eff}}_k(X^m_r)$ for $r \le 2m - k + 1$ and all $k < m < n$. Let Γ be a set of r points such that Γ consists of $r-2$ very general points $p_1, \dots, p_{r-2}$ in a hyperplane $L = \mathbb{P}^{n-1}$ and two very general points p_{r-1}, p_r not contained in L. Let L' denote the proper transform of L in X^n_Γ. Note that $L' \cong X^{n-1}_{r-2}$. Let Y be an irreducible k-dimensional subvariety of X^n_Γ not contained in an exceptional divisor with class $aH_k + \sum_{i=1}^r b_i E_i$ on X^n_Γ. If Y is contained in L', then $Y \subset X^{n-1}_{r-2}$. Since $r - 2 \le 2(n-1) - k + 1$, by the induction hypothesis the class of Y is linearly generated and $(k+1)a \ge \sum_{i=1}^r b_i$. If Y is not contained in L', then $Z = Y \cap L'$ is an effective cycle of dimension $k-1$. Let p_0 denote the intersection of the line $l_{1,2}$ spanned by p_{r-1} and p_r with L. Let $\beta = \max(0, b_{r-1} + b_r - a)$. Consider the blowup X^{n-1}_{r-1} of L along $p_0, p_1, \dots, p_{r-2}$. Then the proper transform of Z is an effective cycle in X^{n-1}_{r-1} with class

$$aH_{k-1} - (\beta + c)E_{0,k-1} - \sum_{i=1}^{r-2} b_i E_{i,k-1},$$

for some $c \ge 0$. Since $r \le 2n - k + 1$, the inductive hypothesis $r - 1 \le 2(n-1) - (k-1) + 1$ is satisfied. We conclude that this class is linearly generated. Consequently,

$$ka \ge \beta + \sum_{i=1}^{r-2} b_i \qquad \text{and hence} \qquad (k+1)a \ge \sum_{i=1}^{r} b_i.$$

By Lemma 2.6, the class of Y is linearly generated. By Corollary 2.5, $\overline{\mathrm{Eff}}_k(X^n_r) \subset \overline{\mathrm{Eff}}_k(X^n_\Gamma)$ and $\overline{\mathrm{Eff}}_k(X^n_r)$ is linearly generated. This concludes the induction and the proof of the theorem. □

The cone $\overline{\mathrm{Eff}}_2(X^4_8)$. In this subsection we prove that $\overline{\mathrm{Eff}}_2(X^4_8)$ is linearly generated.

Theorem 4.6. *The cone $\overline{\mathrm{Eff}}_2(X^4_8)$ is linearly generated.*

The fact that the points are now very general means that we are in position to apply degeneration arguments. These sorts of arguments work because of the semicontinuity of multiplicities in families; a surface S on the very general X_8^4 violating the inequality $3a \geq \sum b_i$ will specialize to an effective 2-cycle with the same property (Lemma 2.4).

To illustrate the range of applicable techniques we give two different degeneration arguments to prove Theorem 4.6.

Proof 1. Let Γ be a configuration of 8 points in $\mathbb{P}^4$ such that $p_1, \dots, p_7$ are very general points and p_8 is a general point on $\mathrm{Sec}(R)$, where $\mathrm{Sec}(R)$ is the secant variety of the rational normal curve R through the points $p_1, \dots, p_7$. Let Y be an irreducible surface in X_Γ. If Y is not contained in the strict transform $\overline{\mathrm{Sec}}(R)$, then $Y \cap \overline{\mathrm{Sec}}(R)$ is a curve and the class $3aH_1 - \sum_{i=1}^{7} 2b_i E_{i,1} - b_8 E_{8,1}$ is effective. We claim that the linear system $Q = 2H - \sum_{i=1}^{7} E_i - 2E_8$ is nef on $\overline{\mathrm{Sec}}(R)$. Granting this, $Q \cdot \overline{\mathrm{Sec}(R)} \cdot Y = 6a - 2\sum_{i=1}^{8} b_i \geq 0$, and so Y is in the span of planes. We may therefore assume that $Y \subset \overline{\mathrm{Sec}}(R)$.

To prove that Q is nef on $\overline{\mathrm{Sec}}(R)$ it suffices to show that the linear system of quadrics double at p_8 and passing through $p_1, \dots, p_7$ in $\mathbb{P}^4$ has base locus consisting of 8 lines, none of which are contained in $\mathrm{Sec}(R)$. Then Q restricts to a semiample, and in particular, nef class on $\overline{\mathrm{Sec}}(R)$. Since p_8 is general, the only line through p_8 incident to R and contained in $\mathrm{Sec}(R)$ is the unique secant line l to R through p_8. Suppose there was another line $p_8 \in l' \subset \mathrm{Sec}(R)$ incident to R, then the plane spanned by l and l' would intersect $\mathrm{Sec}(R)$ in a completely reducible cubic curve singular along the three points of intersection with R and at p_8. This is clearly impossible.

Since p_8 is general, we may assume that the secant line l does not contain any of the points $p_1, \dots, p_7$. Any effective member of the linear system Q is a quadric cone with vertex at p_8. Let $q_1, \dots, q_7$ denote the projection of $p_1, \dots, p_7$ through p_8. These are very general 7 points in $\mathbb{P}^3$. The base locus of the linear system of quadrics passing through $q_1, \dots, q_7$ in $\mathbb{P}^3$ is 8 points $q, q_1, \dots, q_7$, contained in the smooth locus of the projection R' of R. (The curve R' is a complete intersection of 2 quadrics. There is a three-dimensional linear system of quadrics passing through $q_1, \dots, q_7$. Any quadric in this linear system not containing R' intersects R' at a further point q in the smooth locus of R'.) By taking cones over these quadrics with vertex p_8, we see that the base locus of Q in $\mathbb{P}^4$ are the 8 lines spanned by p_8 and one of $q, q_1, \dots, q_7$. Since these are lines through p_8 incident to R and distinct from the secant line containing p_8, none of them are contained in $\mathrm{Sec}(R)$.

To prove the theorem we need to show that $S \cdot Y = 3a - \sum_{i=1}^{8} b_i \geq 0$ for a cubic scroll S. It suffices to show that there is some scroll S which intersects Y in finitely many points. We may further assume that $Y \subset \overline{\mathrm{Sec}}(R)$.

Let us consider the family of cubic scrolls through these 8 points. Through the first 7 points, $p_1, \dots, p_7$ there is a 4-dimensional space of scrolls. There is an open set of this parameter space parametrizing scrolls S such that the intersection $\mathrm{Sec}(R) \cap S = B$ is an irreducible curve. There exists such a curve through a general eighth point, by Claim 3.13, and so there must in fact be an irreducible curve through a general eighth point. In all, picking p_8 general, we have that the space of cubic scrolls containing $p_1, \dots, p_8$ is of dimension 2 and there is a nonempty open subset of scrolls such that the proper transform of a scroll $\bar{S}$ intersects $\mathrm{Sec}(R)$ in an irreducible curve.

Now, taking a general point q of $\mathrm{Sec}(R)$, there exists two scrolls containing the 9 points $q, p_1, \dots, p_8$ [Coskun 2006]. Since varying q gives a 2-dimensional family of scrolls, we conclude that by choosing S generic, the curve B can be made to pass through a general ninth point of $\mathrm{Sec}(R)$. In particular, we can choose a scroll so that B is not contained in Y. As the curve B is irreducible and not contained in Y, we see that $\bar{S}$ and Y generically intersect in finitely many points and so their intersection number is nonnegative. We conclude that the effective cone is linearly generated. □

Proof 2. We will degenerate to a configuration where the 8 points lie on two rank 3 quadrics.

Let $x_0, \dots, x_4$ be coordinates on $\mathbb{P}^4$ and let $p_1 = (1, 0, 0, 0, 0), \dots, p_5 = (0, 0, 0, 0, 1)$ denote the coordinate points. Consider the two quadrics $q_1 = \{x_0x_1 + x_0x_2 + x_1x_2 = 0\}$ and $q_2 = \{x_2x_3 + x_2x_4 + x_3x_4 = 0\}$. Here q_1 is a cone over a smooth conic in the $x_0x_1x_2$-plane with the vertex being the line $\{x_0 = x_1 = x_2 = 0\}$, and similarly for q_2. Note that q_1 and q_2 both contain the points $p_1, \dots, p_5$. Moreover, q_1 and q_2 respectively contain p_4, p_5 and p_1, p_2 with multiplicity two. We now choose the remaining three points p_6, p_7, p_8 to be general on the intersection $q_1 \cap q_2$.

On the blowup at these points the strict transform of q_1 is an irreducible divisor Q_1 with class $2H - \sum_{i=1}^{8} E_i - E_4 - E_5$. Consider the divisor

$$D_1 = 3H - 2\sum_{i=1}^{8} E_i + E_4 + E_5;$$

it satisfies $[Q_1] \cdot [D_1] = 2(3H^2 - \sum_{i=1}^{8} E_{i,2})$. A computation gives that the linear system $|D_1|$ is 2-dimensional. Moreover, the base locus of D_1 is 1-dimensional and has 18 components: 15 lines and 3 quartic normal curves. One checks that none of these curves lie on Q_1. Indeed, these statements are easy to verify for one particular configuration (and thus it follows a general 8-tuple as above). In particular, $D_1|_{Q_1}$ has only finitely many base-points, and hence is nef on Q_1.

Now, suppose that $Y \subset X$ is an irreducible surface with class $aH^2 - \sum_{i=1}^{8} b_i E_{i,2}$. Then, if Y is not contained in Q_1, the intersection i^*Y is represented by an effective

1-cycle on Q_1 (here $i : Q_1 \to X$ is the inclusion). As $D_1|_{Q_1}$ is nef, we have

$$0 \leq D_1|_{Q_1} \cdot i^*Y = 6a - \sum_{i=1}^{8} 2b_i,$$

as desired. A symmetric argument (with $Q_2 = 2H - \sum_{i=1}^{8} E_i - E_1 - E_2$ and $D_2 = 3H - 2\sum_{i=1}^{8} E_i + E_1 + E_2$), shows that the conclusion also holds if Y is not contained in Q_2.

We therefore reduce to the case where $Y \subseteq Q_1 \cap Q_2$. Note that $Q_1 \cap Q_2$ is an irreducible surface, so $Y = Q_1 \cap Q_2$. Now

$$[Q_1][Q_2] = 4H^2 - 2E_{1,2} - 2E_{2,2} - 2E_{3,2} - 2E_{4,2} - E_{5,2} - E_{6,2} - E_{7,2} - E_{8,2},$$

which is equivalent to a sum of 4 planes. This completes the proof. □

We immediately deduce the following corollary.

Corollary 4.7. *If X_r^n is a Mori dream space, then $\overline{\mathrm{Eff}}_k(X_r^n)$ is finitely generated.*

Remark 4.8. Combining Theorem 4.6 with the degeneration argument of Theorem 4.5, it follows that $\overline{\mathrm{Eff}}_2(X_r^n)$ is linearly generated for $r \leq 2n$ as long as $n \geq 4$.

We will see in the next section that $\overline{\mathrm{Eff}}_2(X_{10}^4)$ is not finitely generated, assuming the SHGH conjecture holds for blowups of $\mathbb{P}^2$ at 10 points. The only remaining case in dimension 4 is

Question 4.9. Is the cone $\overline{\mathrm{Eff}}_2(X_9^4)$ linearly generated?

It is not easy to find explicit curves in X_r^n which are not in the span of lines. The following example gives a construction in the case of 9 very general points in $\mathbb{P}^3$.

Example 4.10. The class $C_{CM} = 57H_1 - \sum_{i=1}^{10} 18E_{i,1}$ on X_{10}^2 is represented by a unique irreducible plane curve of genus 10, by [Ciliberto and Miranda 2011]. On X_9^3 there is a unique divisor Q in the class $2H_1 - \sum_{i=1}^{9} E_{i,1}$, given by the strict transform of the unique quadric through the 9 points. There is a morphism $i : X_{10}^2 \to X_9^3$ identifying the proper transform of Q with the blowup of $\mathbb{P}^2$ at 10 points.

A quick calculation shows that the pushforward of the class of C_{CM} to X_9^3 is

$$i_*(C_{CM}) = 78H_1 - 21E_{1,1} - \sum_{i=2}^{9} 18E_{i,1}.$$

We have $21 + 8(18) = 165$, while $2 \cdot 78 = 156$. Hence this curve is not in the span of the lines. It does not, however, define an extremal ray on $\overline{\mathrm{Eff}}_1(X_9^3)$. In the next section we will use a similar construction to show that, assuming the SHGH conjecture, the cone $\overline{\mathrm{Eff}}_1(X_9^3)$ is not finitely generated.

By repeatedly taking cones over $i_*(C_{CM})$, we obtain explicit nonlinearly generated codimension-two cycles on X_{n+6}^n for every $n \geq 3$.

Complete intersections also provide examples of nonlinearly generated pseudoeffective curve classes, provided that the number of points is large.

Example 4.11. Assume that $d^n \geq r > 2d^{n-1}$ for some integer $d > 2$. Then the divisor class $D = dH - \sum_{i=1}^{r} E_i$ is nef on X_r^n by the argument given in the proof of Proposition 4.1. The $(n-1)$-fold self-intersection of the class is

$$D^{n-1} = d^{n-1} H_1 - \sum_{i=1}^{r} E_{i,1}.$$

Since $r > 2d^{n-1}$, this class is not in the span of lines. On the other hand, the class is pseudoeffective. A small perturbation of D is ample. Hence a sufficiently high multiple is very ample and the $(n-1)$-fold self-intersection is an effective curve. It follows that the class D^{n-1} is pseudoeffective.

5. Nonfinitely generated cones

The cone of curves of the blowup of $\mathbb{P}^2$ at 10 or more very general points is not entirely understood, and we will find it useful to assume the following standard conjecture.

Conjecture 5.1 (Segre–Harbourne–Gimigliano–Hirschowitz (SHGH) conjecture [Gimigliano 1987]). Suppose that $r \geq 10$ and that $m_1 \geq m_2 \geq \cdots \geq m_r$ and $d > m_1 + m_2 + m_3$. Then

$$H^0\left(X_r^2, dH_1 - \sum_{i=1}^{r} m_i E_{i,1}\right) = \binom{d+2}{2} - \sum_{i=1}^{r} \binom{m_i+1}{2}.$$

We next prove that the cone of codimension-2 cycles on X_r^n is not finitely generated for $r \geq n+6$, assuming the SHGH conjecture. The calculation relies on the following observation of de Fernex.

Theorem 5.2 [de Fernex 2010, Proposition 3.4]. *Assume the SHGH conjecture holds for* 10 *points. Let $P \subset N^1(X_{10}^2)$ be the positive cone*

$$P = \{D \in N^1\left(X_{10}^2\right) : D^2 \geq 0,\, D \cdot H \geq 0\},$$

where H is an ample divisor. Then

$$\overline{\mathrm{Eff}}_1\left(X_{10}^2\right) \cap K_{\geq 0} = P \cap K_{\geq 0}.$$

Let $Q \subset X_9^3$ be the strict transform of the unique quadric passing through the 9 points and let i denote the inclusion of Q in X_9^3. Note that Q is isomorphic to X_{10}^2, and so Conjecture 5.1 provides some information about the cone $\overline{\mathrm{Eff}}_1(Q)$. However, the map $N_1(Q) \to N_1\left(X_9^3\right)$ is not injective, since the two rulings of the quadric both map to the class of a line in $\mathbb{P}^3$. The next lemma gives a criterion to show that certain extremal rays on $\overline{\mathrm{Eff}}_1(Q)$ nevertheless push forward to extremal rays on $\overline{\mathrm{Eff}}_1(X_9^3)$.

Write r_1 and r_2 for the classes of the two rulings on the quadric, and let $f_i = E_i|_Q$ be the exceptional curves. Let $\ell_{ij} = r_1 - f_i - f_j \in N_1(Q)$, this class is not effective on Q, but $i_*\ell_{ij}$ is effective in $N_1(X_9^3)$ since it is the class of a line through the points p_i and p_j.

Theorem 5.3. *Suppose that D is a class in $N_1(Q)$ which satisfies*:

(1) *D is nef, and if $\gamma \in \overline{\mathrm{Eff}}_1(Q)$ has $D \cdot \gamma = 0$, then γ is a multiple of D.*

(2) *$D \cdot r_1 = D \cdot r_2$.*

(3) *$D \cdot \ell_{ij} > 0$ for all i and j.*

Then i_*D is nonzero and spans an extremal ray on $\overline{\mathrm{Eff}}_1(X_9^3)$.

Proof. That i_*D is nonzero follows from the fact that D is nef, since if H is ample then so is i^*H and then $i_*D \cdot H = D \cdot i^*H > 0$. We claim next that D lies on a two-dimensional extremal face of the cone

$$\Sigma = \overline{\mathrm{Eff}}_1(Q) + \sum_{i,j} \mathbb{R}_{\geq 0}[\ell_{ij}] + \mathbb{R}[r_1 - r_2] \subset N_1(Q).$$

More precisely, if $D = \alpha + \beta$ with $\alpha, \beta \in \Sigma$, then

$$\alpha = a_1 D + b_1(r_1 - r_2) \quad \text{and} \quad \beta = a_2 D + b_2(r_1 - r_2),$$

where a_1 and a_2 are positive. Note that D is nef, $D \cdot (r_1 - r_2) = 0$, $D \cdot \ell_{ij} > 0$ for all i, j, and $D \cdot f_k > 0$. Hence D is contained in the dual cone of Σ. By conditions (1) and (2), the classes in Σ with $D \cdot C = 0$ are precisely $\mathbb{R}_{\geq 0}D + \mathbb{R}(r_1 - r_2)$.

We claim next that

$$\overline{\mathrm{Eff}}_1(X_9^3) = i_*\,\overline{\mathrm{Eff}}_1(Q) + \sum_{i,j} \mathbb{R}_{\geq 0} i_*[\ell_{ij}] = i_*\Sigma.$$

Suppose that Γ is an irreducible effective cycle on $\overline{\mathrm{Eff}}_1(X_9^3)$. If $Q \cdot \Gamma < 0$, then Γ must be contained in Q, and so $[\Gamma]$ is contained in $i_*\,\overline{\mathrm{Eff}}_1(Q)$. If $Q \cdot \Gamma \geq 0$, then Γ satisfies $2a \geq \sum_i b_i$, which means that $[\Gamma]$ is in the span of classes of lines $i_*[\ell_{ij}]$ and lines $i_*[f_k]$ in the exceptional divisors, by Lemma 2.6. Each f_k is numerically equivalent to a curve in the quadric.

Suppose now that $i_*D = \alpha + \beta$, where α and β are pseudoeffective classes on $\overline{\mathrm{Eff}}_1(X_9^3)$. Using the decomposition above, we can write $\alpha = i_*\alpha_Q + \sum c_{ij} i_*[\ell_{ij}]$ and $\beta = i_*\beta_Q + \sum d_{ij} i_*[\ell_{ij}]$, where α_Q and β_Q are classes in $\overline{\mathrm{Eff}}_1(Q)$.

We claim now that

$$D = \alpha_Q + \beta_Q + \sum (c_{ij} + d_{ij})\ell_{ij} + f(r_1 - r_2),$$

for some constant f. Indeed, the two sides differ by an element of the kernel of $i_*: N_1(Q) \to N_1(X_9^3)$, which is generated by $r_1 - r_2$, giving rise to the constant f.

Since $D^2 = 0$, $D \cdot (r_1 - e_i - e_j) > 0$ for any i and j, and $D \cdot (r_1 - r_2) = 0$, it must be that $c_{ij} = d_{ij} = 0$ for all i and j. We conclude that

$$\alpha_Q = a_1 D + b_1(r_1 - r_2) \quad \text{and} \quad \beta_Q = a_2 D + b_2(r_1 - r_2).$$

Hence $\alpha = i_*\alpha_Q = a_1 i_* D$ and $\beta = i_*\beta_Q = a_2 i_* D$. This shows that $i_* D$ is extremal. □

The requirement that D is nef makes it difficult to exhibit classes D on X_{10}^2 with the necessary properties without assuming the description of $\overline{\mathrm{Eff}}_1(X_{10}^2)$ provided by the SHGH conjecture.

Theorem 5.4. *Assume that the SHGH conjecture holds for blowups of $\mathbb{P}^2$ at 10 very general points. Then there exist infinitely many classes D satisfying the hypotheses of Theorem 5.3.*

Proof. It is convenient to fix an identification $Q \cong X_{10}^2$ and rewrite the hypotheses in the basis for $N^1(Q)$ arising from this identification. The strict transforms of the two rulings through the point p_1 give disjoint (-1)-curves on Q, and these can be contracted. The other 8 exceptional curves f_i can then be contracted to give a map to $\mathbb{P}^2$. Let e_0 and e_1 be the first two (-1)-curves contracted, and let $e_j = f_j$ for $2 \le j \le 8$. With respect to this new basis, we have $r_1 - f_1 = e_0$ and $r_2 - f_1 = e_1$, and $f_1 = h - e_0 - e_1$, where h denotes the class of a line on $\mathbb{P}^2$.

While the first condition in Theorem 5.3 is independent of the basis, the last two can be rewritten as:

(2′) $D \cdot e_0 = D \cdot e_1$.

(3′) $D \cdot e_0 > D \cdot e_j$ for any $j > 1$, and $D \cdot (h - e_1 - e_i - e_j) > 0$ for any $i, j > 1$.

The first part of (3′) arises when $i = 1 < j$, while the second case is when $1 < i < j$.

Fix any $\frac{1}{\sqrt{10}} < \delta < \frac{1}{3}$, and let $\delta' = \sqrt{\frac{1}{8}(1 - 2\delta^2)}$. Observe that $\frac{3}{10} < \frac{1}{6}\sqrt{\frac{7}{2}} < \delta' < \delta$ for δ in this range. Consider the divisor

$$D_\delta = h - \delta(e_0 + e_1) - \delta' \sum_{j=2}^{9} e_j.$$

We check each of the hypotheses in turn. To simplify notation, for the rest of this proof set $X = X_{10}^2$.

(1) First we check that D_δ is nef. The cone theorem implies that

$$\overline{\mathrm{Eff}}_1(X) = \overline{\mathrm{Eff}}_1(X)_{K_X \ge 0} + \sum_i \mathbb{R}_{\ge 0}[C_i],$$

where the C_i are K_X-negative curves. According to Theorem 5.2,

$$\overline{\mathrm{Eff}}_1(X) \cap \overline{\mathrm{Eff}}_1(X)_{K_X \geq 0} = P \cap K_{\geq 0}.$$

Hence it suffices to show that $D_\delta \cdot C \geq 0$ if C is K_X-negative, and that $D_\delta \cdot C \geq 0$ if C has $C^2 \geq 0$ and $C \cdot H > 0$.

First, suppose that C is a pseudoeffective class with $K_X \cdot C < 0$. We have

$$3D_\delta - K_X = (3\delta - 1)(e_0 + e_1) + (3\delta' - 1)\sum_{j=2}^{9} e_j,$$

and so

$$3D_\delta \cdot C = K_X \cdot C + ((3\delta - 1)(e_0 + e_1) + (3\delta' - 1)\sum_{j=2}^{9} e_j) \cdot C.$$

However, since $\delta < \frac{1}{3}$, the number $3\delta - 1$ is negative. It is easy to check that $\delta' < \frac{1}{3}$ as well, and so the divisor on the right is a sum of exceptional divisors with negative coefficients. If C is any curve other than one of the e_i, then both terms on the right are negative. If C is one of the curves e_i, then $D_\delta \cdot C > 0$ because δ and δ' are both positive.

It remains to check that $D_\delta \cdot C > 0$ if C is a class with positive self-intersection. This follows from the Cauchy–Schwartz inequality. Suppose that $d^2 \geq \sum a_i^2$ and $e^2 \geq \sum b_i^2$, then $de \geq \sum a_i b_i$. Moreover, equality is achieved if and only if C is a multiple of D_δ.

(2′) Since D_δ is of the form $h - \delta e_0 - \delta e_1 - \cdots$, we have $D \cdot e_0 = D \cdot e_1$.

(3′) Because $\delta > \delta'$, we have $D \cdot e_0 > D \cdot e_j$ for any $j > 1$. We also have

$$D \cdot (h - e_1 - e_i - e_j) = 1 - \delta - 2\delta' > 0,$$

since $\delta' < \delta < \frac{1}{3}$. □

Remark 5.5. One can even arrange that D_δ is a rational class through judicious choice of δ. For example,

$$D_{\frac{226}{692}} = h - \frac{226}{692}(e_0 + e_1) - \frac{217}{692}\sum_{i=2}^{9} e_i.$$

In general such classes are not expected to have any effective representatives.

Assuming the SHGH conjecture, we can now conclude that $\overline{\mathrm{Eff}}^2(X_r^n)$ is not finitely generated if $r \geq n + 6$. We need the following lemma, which guarantees that cones over extremal classes in $\overline{\mathrm{Eff}}_k(X_r^n)$ are extremal in $\overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$.

Lemma 5.6. *Suppose that* $D = aH_k - \sum_{i=1}^r a_i E_{i,k}$ *spans an extremal ray on* $\overline{\mathrm{Eff}}_k(X_r^n)$. *Then* $CD = aH_{k+1} - aE_{0,k+1} - \sum_{i=1}^r a_i E_{i,k+1}$ *spans an extremal ray on* $\overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$. *In particular, if* $\overline{\mathrm{Eff}}_k(X_r^n)$ *has infinitely many extremal rays, then so does* $\overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$.

Lemma 5.6 immediately implies the following.

Corollary 5.7. *Assume the SHGH conjecture for the blowup of* $\mathbb{P}^2$ *at* 10 *points. Then* $\overline{\mathrm{Eff}}^2(X_r^n)$ *is not finitely generated if* $r \geq n+6$.

Proof of Lemma 5.6. Given $r+1$ very general points $p_0, \dots, p_r$ in $\mathbb{P}^{n+1}$, their projection from p_0 gives r very general points in $\mathbb{P}^n$. Let D_i be effective cycles arbitrarily close to D in $\overline{\mathrm{Eff}}_k(X_r^n)$. Then the classes of the cones CD_i over D_i converge to CD. Hence $CD \in \overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$.

Conversely, we claim that if $CD = aH_{k+1} - aE_0 - \sum_{i=1}^r b_i E_{i,k+1}$ is a pseudoeffective $(k+1)$-cycle on X_{n+1}^{r+1}, then $D = aH_k - \sum_{i=1}^r b_i E_{i,k}$ is a pseudoeffective k-cycle on X_n^r. The class $CD + \epsilon H_{k+1}$ is effective for any $\epsilon > 0$. Let V_ϵ be a (rational) cycle representing the class $CD + \epsilon H_{k+1}$. Let ℓ_{0j} denote the strict transform on X_{n+1}^{r+1} of the line through p_0 and p_j. By Lemma 2.10, V_ϵ contains the line ℓ_{0j} with multiplicity $\beta_j \geq b_j - \epsilon$. Let $L \subset X_{r+1}^{n+1}$ be the proper transform of a general hyperplane in $\mathbb{P}^{n+1}$. The lines l_{0j} intersect L in r very general points p. The proper transform of the intersection $L \cap V_\epsilon$ gives an effective cycle with class $(a+\epsilon)H_k - \sum_{i=1}^j \beta_j E_{j,k}$. Letting ϵ tend to 0, we see that $aH_k - \sum_{i=1}^r b_i E_{i,k}$ is pseudoeffective in X_r^n, as required.

Now, suppose that $D = aH_k - \sum_{i=1}^r b_i E_{i,k}$ spans an extremal ray of $\overline{\mathrm{Eff}}_k(X_r^n)$. We claim that $CD = aH_{k+1} - aE_{0,k+1} - \sum_{i=1}^r b_i E_{i,k+1}$ spans an extremal ray of $\overline{\mathrm{Eff}}_{k+1}(X_{r+1}^{n+1})$. Suppose $CD = \alpha + \beta$, where α and β are both pseudoeffective $(k+1)$-cycles on X_{r+1}^{n+1}. Since any pseudoeffective class has $a \geq b_0$, it must be that

$$\alpha = cH_{k+1} - cE_{0,k+1} - \sum_{i=1}^r c_i E_{i,k+1} \quad \text{and} \quad \beta = dH_{k+1} - dE_{0,k+1} - \sum_{i=1}^r d_i E_{i,k+1}.$$

Then

$$\alpha_0 = cH_k - \sum_{i=1}^r c_i E_{i,k} \quad \text{and} \quad \beta_0 = dH_k - \sum_{i=1}^r e_i E_{i,k}.$$

are pseudoeffective on X_r^n. Hence α_0 and β_0 are proportional to D. It follows that α and β are proportional to CD and CD is extremal. □

There are several interesting remaining questions concerning the finite generation of cones of higher codimension.

Question 5.8. Can one show that $\overline{\mathrm{Eff}}_{n-2}(X_r^n)$ is not finitely generated for $r \geq n+6$ independently of the SHGH conjecture?

Question 5.9. Fix n and k. Does there exist r for which $\overline{\mathrm{Eff}}_k(X^n_r)$ is not finitely generated? How does r depend on n and k?

In particular, we have the following fundamental question:

Question 5.10. For every n, does there exist r for which $\overline{\mathrm{Eff}}_1(X^n_r)$ is not finitely generated?

If $\overline{\mathrm{Eff}}_1(X^n_r)$ is not finitely generated for $r \geq r_0$, then by Lemma 5.6 $\overline{\mathrm{Eff}}_k(X^{n+k-1}_{r+k-1})$ is not finitely generated for $r \geq r_0$. Hence an affirmative answer to Question 5.10 implies an affirmative answer to Question 5.9.

6. Blowups at points in special position

Until now we have considered blowups of $\mathbb{P}^n$ at linearly general or very general points. It is also interesting to consider cones of effective cycles on blowups of $\mathbb{P}^n$ at special configurations of points. The dependence of the cones on the position of the points can be subtle, which makes degeneration arguments difficult. We will see that the property of the effective cone being finite is neither an open nor closed condition, even in families where the vector space of numerical classes of k-dimensional cycles has constant dimension.

Proposition 6.1. *Let Γ be a set of r points whose span is $\mathbb{P}^m \subset \mathbb{P}^n$. Let X^n_Γ and X^m_Γ denote the blowup of $\mathbb{P}^n$ and $\mathbb{P}^m$ along Γ, respectively, then*

(1) *$\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ is linearly generated for $m \leq k \leq n-1$ and*

(2) *$\overline{\mathrm{Eff}}_k(X^n_\Gamma) = \overline{\mathrm{Eff}}_k(X^m_\Gamma)$ for $k < m$.*

Proof. Since X^m_Γ embeds in X^n_Γ as the proper transform of the $\mathbb{P}^m$ spanned by Γ, any effective cycle $Z \subset X^m_\Gamma$ is also an effective cycle in X^n_Γ with the same class. Hence $\overline{\mathrm{Eff}}_k(X^m_\Gamma) \subseteq \overline{\mathrm{Eff}}_k(X^n_\Gamma)$ for $k < m$. Conversely, suppose that $k < m$. Let Z be an effective cycle in $\mathbb{P}^n$ of dimension k with class $[Z]$. We may assume that Z is not contained in an exceptional divisor. Choose a general point p. Let q_i denote the projection of p_i from p and let Z' be the projection of Z from p. Then Z' and Z have the same degree and the multiplicities of Z' at q_i are greater than or equal to the multiplicities of Z at p_i. Repeatedly projecting Z to $\mathbb{P}^m$ from general points, we obtain an effective cycle contained in $\mathbb{P}^m$. Since $[Z]$ differs from the class of this cycle by a positive combination of exceptional linear spaces $E_{i,k}$, we conclude that $[Z]$ is effective on X^m_Γ. Taking closures, we obtain the reverse inclusion $\overline{\mathrm{Eff}}_k(X^n_\Gamma) \subseteq \overline{\mathrm{Eff}}_k(X^m_\Gamma)$.

If $k \geq m$, let L be a k-dimensional linear space containing Γ. Then the proper transform $\bar{L}$ of L has class $H_k - \sum_{i=1}^r E_{i,k}$. Since a k-dimensional variety not contained in an exceptional divisor has class $a H_k - \sum_{i=1}^r b_i E_{i,k}$, with $a \geq b_i \geq 0$, we conclude that any k-dimensional effective cycle is a nonnegative linear combination of $[\bar{L}]$ and $E_{i,k}$, $1 \leq i \leq r$. ☐

By taking $m = 1$, we obtain the following corollary.

Corollary 6.2. *Suppose* Γ *is a set of* r *collinear points in* $\mathbb{P}^n$. *Then* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is linearly generated for every* $1 \le k \le n-1$.

Remark 6.3. It was shown in [Ottem 2011] that the blowup of $\mathbb{P}^2$ in collinear points is a Mori dream space (and indeed its Cox ring can be computed). Consequently, the cone of curves and divisors are finite polyhedral. The previous corollary generalizes this to all cycles.

Let $L \cong \mathbb{P}^{n-1}$ be a hyperplane in $\mathbb{P}^n$. Let $\Gamma' \subset L$ be a set of points $p_1, \dots, p_r$ and let $p_0 \in \mathbb{P}^n$ be a point not contained in L. Let $\Gamma = \Gamma' \cup \{p_0\}$. Let X^n_Γ and $X^{n-1}_{\Gamma'}$ denote the blowup of $\mathbb{P}^n$ and $\mathbb{P}^{n-1}$ along Γ and Γ', respectively. Taking cones with vertex at p_0, we generate a subcone $\mathrm{C}\overline{\mathrm{Eff}}_k(X^{n-1}_{\Gamma'}) \subset \overline{\mathrm{Eff}}_{k+1}(X^n_\Gamma)$.

Proposition 6.4. *The cone* $\overline{\mathrm{Eff}}_{k+1}(X^n_\Gamma)$ *is generated by* $\mathrm{C}\overline{\mathrm{Eff}}_k(X^{n-1}_{\Gamma'})$, $\overline{\mathrm{Eff}}_{k+1}(X^{n-1}_{\Gamma'})$, *and* $E_{0,k+1}$. *Furthermore, the extremal rays of* $\mathrm{C}\overline{\mathrm{Eff}}_k(X^{n-1}_{\Gamma'})$ *are also extremal in* $\overline{\mathrm{Eff}}_{k+1}(X^n_\Gamma)$.

Proof. Let $Z = aH_{k+1} - \sum_{i=0}^r b_i E_{i,k+1}$ be an irreducible $(k+1)$-dimensional variety in X^n_Γ. We may assume that Z is not contained in any exceptional divisors. Otherwise, its class is a positive multiple of $E_{i,k+1}$. The proper transform of L in X^n_Γ is isomorphic to $X^{n-1}_{\Gamma'}$. If Z is contained in $X^{n-1}_{\Gamma'}$, then the class of Z is in $\overline{\mathrm{Eff}}_{k+1}(X^{n-1}_{\Gamma'})$. Otherwise, $Z \cap X^{n-1}_{\Gamma'}$ is an effective k cycle with class $\alpha' = aH_k - \sum_{i=1}^r b'_i E_{i,k}$, where $b'_i \ge b_i$. Consequently, the class $\alpha = aH_k - \sum_{i=1}^r b_i E_{i,k}$ is effective. Then the cone $C(\alpha)$ is an effective class in X^n_Γ and, since $b_0 \le a$, $[Z]$ is in the span of $E_{0,k+1}$ and $C(\alpha)$.

Let Z be a cycle that generates an extremal ray of $\mathrm{C}\overline{\mathrm{Eff}}_k(X^{n-1}_{\Gamma'})$. Suppose $[Z] = \alpha + \beta$ in $\overline{\mathrm{Eff}}_{k+1}(X^n_\Gamma)$. Since $b_0 \le a$ holds on $\overline{\mathrm{Eff}}_{k+1}(X^n_\Gamma)$ and $b_0 = a$ on $\mathrm{C}\overline{\mathrm{Eff}}_k(X^{n-1}_{\Gamma'})$, we must have that both α and β satisfy $b_0 = a$. We can perturb α and β by ϵH_{k+1} to obtain rational effective classes. Since the coefficient of $E_{0,k+1}$ of any class contained in $\overline{\mathrm{Eff}}_{k+1}(X^{n-1}_{\Gamma'})$ is 0, the coefficients of any component of the class contained in $\overline{\mathrm{Eff}}_{k+1}(X^{n-1}_{\Gamma'})$ are bounded by ϵ. Taking cones over the classes of the hyperplane sections of the remaining subvarieties and letting ϵ tend to zero, we see that both α and β are contained in $\mathrm{C}\overline{\mathrm{Eff}}_k\, g(X^{n-1}_{\Gamma'})$. By the extremality of Z, we conclude that they are both proportional to $[Z]$. $\square$

Corollary 6.5. *Let* Γ *be a set of points* $\{q_1, \dots, q_9, p_1, \dots, p_{s-1}\}$ *such that the* q_i *are general points in a plane* $P \subset \mathbb{P}^n$ *and the* p_i *are linearly general points with span disjoint from* P. *Then* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is not finitely generated for* $k \le s$ *and is linearly generated for* $k > s$.

Proof. When $r \ge 9$, the blowup of $\mathbb{P}^2$ at r general points has infinitely many (-1)-curves, which span extremal rays of the effective cone of curves. Applying Proposition 6.4 $(k-1)$-times, the cones over the classes of (-1)-curves with vertex

$p_1, \dots, p_{k-1}$ provide infinitely extremal rays of $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ for $k \le s$. The linear generation of $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ for $k > s$ follows from Proposition 6.1. □

Corollary 6.6. (1) *Linear generation of* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is not closed in smooth families.*

(2) *Finite generation of* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is not closed in smooth families.*

Proof. Let $n \ge k+8$. Take a general smooth curve B in $(\mathbb{P}^n)^{k+8}$ which avoids all the diagonals and contains a point $0 \in B$ where 9 of the points are general points in a plane P and the remaining points are in linearly general position with span not intersecting P. Such curves exist by Bertini's Theorem since the diagonals have codimension $n \ge 2$. Consider the family $\mathcal{X} \to B$, where $\mathcal{X}_b$ is the blowup of $\mathbb{P}^n$ in the $k+8$ points Γ_b parametrized by $b \in B$. If the points in Γ_b are in linearly general position, then by Lemma 3.3 the cone $\overline{\mathrm{Eff}}_k(X_{\Gamma_b})$ is linearly generated. In particular, the cone is finite. However, by Corollary 6.5, $\overline{\mathrm{Eff}}_k(X_{\Gamma_0})$ is not finitely generated. In particular, the cone is not linearly generated. □

Corollary 6.7. (1) *Linear generation of* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is not open in smooth families.*

(2) *Finite generation of* $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ *is not open in smooth families.*

Proof. Let B be a smooth curve parametrizing 9 general points in a plane P becoming collinear at $0 \in B$. Let Γ' be $k-1$ points in general linear position in $\mathbb{P}^n$ whose span is disjoint from P. Let Γ_b be the union of Γ' and the points parametrized by b. Consider the family $\mathcal{X} \to B$ obtained by blowing up $\mathbb{P}^n$ along Γ_b. When the points are collinear, $\overline{\mathrm{Eff}}_k(X_{\Gamma_0})$ is linearly generated. However, for the general point of B, $\overline{\mathrm{Eff}}_k(X_{\Gamma_b})$ is not finitely generated by Corollary 6.5. □

Remark 6.8. Corollary 6.7 is well-known for cones of divisors. For example, Castravet and Tevelev [2006] prove that the blowup of $\mathbb{P}^n$ at points on a rational normal curve is a Mori dream space. In particular, if we specialize a large number of points to lie on a rational normal curve, we see that being a Mori dream space is not an open condition.

One can ask for the finite/linear generation of $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ for Γ any special set of points. Perhaps the following question is the most interesting among them.

Question 6.9. Let Γ be a set of points contained in a rational normal curve in $\mathbb{P}^n$. Is $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ finitely generated? Is $\overline{\mathrm{Eff}}_k(X^n_\Gamma)$ generated by the classes of cones over secant varieties of projections of the rational normal curve?

By results of Castravet and Tevelev, the answer to Question 6.9 is affirmative for curves and divisors. The cone of curves $\overline{\mathrm{Eff}}_1(X^n_\Gamma)$ is generated by the class of the proper transform of the rational normal curve $nH_1 - \sum_{i=1}^r E_{i,1}$ and the classes of lines. The rational normal curve is cut out by quadrics. If a curve B has positive intersection with a quadric containing the points, then by Lemma 2.6 the class of B is spanned by the classes of lines. Otherwise, B is contained in the base locus of all the quadrics containing the rational normal curve. Hence B is a multiple of the rational normal curve. Castravet and Tevelev [2006] show that the classes of divisors are generated by linear spaces and codimension-1 cones over secant varieties of the projection of the rational normal curve. We do not know whether $\overline{\mathrm{Eff}}_2(X^n_\Gamma)$ is generated by the classes of planes and cones over the rational normal curve with vertex a point of Γ.

Acknowledgements

We would like to thank Dawei Chen, Lawrence Ein, Mihai Fulger, Joe Harris, Brian Lehmann, Kristian Ranestad, and Kevin Tucker for helpful discussions on cones of higher codimension cycles. We also thank the referees for numerous helpful suggestions.

References

[Birkar et al. 2010] C. Birkar, P. Cascini, C. D. Hacon, and J. McKernan, "Existence of minimal models for varieties of log general type", *J. Amer. Math. Soc.* **23**:2 (2010), 405–468. MR 2601039 Zbl 1210.14019

[Castravet and Tevelev 2006] A.-M. Castravet and J. Tevelev, "Hilbert's 14th problem and Cox rings", *Compos. Math.* **142**:6 (2006), 1479–1498. MR 2278756 Zbl 1117.14048

[Chen and Coskun 2015] D. Chen and I. Coskun, "Extremal higher codimension cycles on moduli spaces of curves", *Proc. Lond. Math. Soc.* (3) **111**:1 (2015), 181–204. MR 3404780 Zbl 06464911

[Ciliberto and Miranda 2011] C. Ciliberto and R. Miranda, "Homogeneous interpolation on ten points", *J. Algebraic Geom.* **20**:4 (2011), 685–726. MR 2819673 Zbl 1230.14040

[Coble 1982] A. B. Coble, *Algebraic geometry and theta functions*, American Mathematical Society Colloquium Publications **10**, American Mathematical Society, Providence, R.I., 1982. MR 733252

[Coskun 2006] I. Coskun, "Degenerations of surface scrolls and the Gromov–Witten invariants of Grassmannians", *J. Algebraic Geom.* **15**:2 (2006), 223–284. MR 2199064 Zbl 1105.14072

[Coskun 2008] I. Coskun, "Gromov–Witten invariants of jumping curves", *Trans. Amer. Math. Soc.* **360**:2 (2008), 989–1004. MR 2346480 Zbl 1222.14033

[Debarre et al. 2011] O. Debarre, L. Ein, R. Lazarsfeld, and C. Voisin, "Pseudoeffective and nef classes on abelian varieties", *Compos. Math.* **147**:6 (2011), 1793–1818. MR 2862063 Zbl 1234.14008

[Debarre et al. 2013] O. Debarre, Z. Jiang, and C. Voisin, "Pseudo-effective classes and pushforwards", *Pure Appl. Math. Q.* **9**:4 (2013), 643–664. MR 3263971 Zbl 1319.14010

[Dolgachev 2011] I. V. Dolgachev, "Cremona special sets of points in products of projective spaces", pp. 115–134 in *Complex and differential geometry*, edited by W. Ebeling et al., Springer Proc. Math. **8**, Springer, Heidelberg, 2011. MR 2964472 Zbl 1227.14020

[de Fernex 2010] T. de Fernex, "On the Mori cone of blow-ups of the plane", preprint, 2010. arXiv 1001.5243

[Fulger and Lehmann 2014a] M. Fulger and B. Lehmann, "Kernels of numerical pushforwards", preprint, 2014. arXiv 1407.6455

[Fulger and Lehmann 2014b] M. Fulger and B. Lehmann, "Positive cones of dual cycle classes", preprint, 2014. arXiv 1408.5154

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik (3) **2**, Springer, Berlin, 1998. MR 1644323 Zbl 0885.14002

[Gimigliano 1987] A. Gimigliano, *On linear systems of plane curves*, Ph.D. thesis, Queen's University, Ann Arbor, MI, 1987. MR 2635606

[Harbourne 1986] B. Harbourne, "The geometry of rational surfaces and Hilbert functions of points in the plane", pp. 95–111 in *Proceedings of the 1984 Vancouver conference in algebraic geometry*, edited by J. Carrell et al., CMS Conf. Proc. **6**, Amer. Math. Soc., Providence, RI, 1986. MR 846019 Zbl 0611.14002

[Harris 1995] J. Harris, *Algebraic geometry*, Graduate Texts in Mathematics **133**, Springer, New York, 1995. MR 1416564

[Hirschowitz 1989] A. Hirschowitz, "Une conjecture pour la cohomologie des diviseurs sur les surfaces rationnelles génériques", *J. Reine Angew. Math.* **397** (1989), 208–213. MR 993223

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry I*, Ergebnisse der Mathematik (3) **48**, Springer, Berlin, 2004. MR 2095471 Zbl 1093.14501

[Li 2015] Q. Li, "Pseudo-effective and nef cones on spherical varieties", *Math. Z.* **280**:3-4 (2015), 945–979. MR 3369360 Zbl 06468909

[Mukai 2004] S. Mukai, "Geometric realization of T-shaped root systems and counterexamples to Hilbert's fourteenth problem", pp. 123–129 in *Algebraic transformation groups and algebraic varieties*, edited by V. L. Popov, Encyclopaedia Math. Sci. **132**, Springer, Berlin, 2004. MR 2090672 Zbl 1108.13300

[Nagata 1959] M. Nagata, "On the 14-th problem of Hilbert", *Amer. J. Math.* **81** (1959), 766–772. MR 0105409 Zbl 0192.13801

[Ottem 2011] J. C. Ottem, "On the Cox ring of $\mathbb{P}^2$ blown up in points on a line", *Math. Scand.* **109**:1 (2011), 22–30. MR 2869458 Zbl 1221.14007

[Ottem 2015] J. C. Ottem, "Nef cycles on some hyperkähler fourfolds", 2015. arXiv 1505.01477

Communicated by János Kollár
Received 2016-03-15 Revised 2016-07-11 Accepted 2016-08-21

coskun@math.uic.edu — *Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, United States*

jdl@uic.edu — *Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, United States*

johnco@math.uio.no — *Department of Mathematics, University of Oslo, Blindern, 0316 Oslo, Norway*

ALGEBRA AND NUMBER THEORY 10:9 (2016)
dx.doi.org/10.2140/ant.2016.10.2015

Cluster algebras and category $\mathcal{O}$ for representations of Borel subalgebras of quantum affine algebras

David Hernandez and Bernard Leclerc

Let $\mathcal{O}$ be the category of representations of the Borel subalgebra of a quantum affine algebra introduced by Jimbo and the first author. We show that the Grothendieck ring of a certain monoidal subcategory of $\mathcal{O}$ has the structure of a cluster algebra of infinite rank, with an initial seed consisting of prefundamental representations. In particular, the celebrated Baxter relations for the 6-vertex model get interpreted as Fomin–Zelevinsky mutation relations.

1. Introduction

Let $U_q(\mathfrak{g})$ be an untwisted quantum affine algebra (we assume throughout this paper that $q \in \mathbb{C}^*$ is not a root of unity). M. Jimbo and Hernandez [2012] introduced a category $\mathcal{O}$ of representations of a Borel subalgebra $U_q(\mathfrak{b})$ of $U_q(\mathfrak{g})$. Finite-dimensional representations of $U_q(\mathfrak{g})$ are objects in this category as well as the infinite-dimensional prefundamental representations of $U_q(\mathfrak{b})$ constructed in

Hernandez is supported in part by the European Research Council under the European Union's Framework Programme H2020 with ERC Grant Agreement number 647353 QAffine.
MSC2010: primary 17B37; secondary 13F60, 17B10, 82B23.
Keywords: quantum affine algebras, cluster algebras, category $\mathcal{O}$, monoidal categorification, Baxter relations.

[Hernandez and Jimbo 2012]. They are obtained as asymptotic limits of Kirillov–Reshetikhin modules, which form a family of simple finite-dimensional representations of $U_q(\mathfrak{g})$. These prefundamental representations, denoted by $L^+_{i,a}$ and $L^-_{i,a}$, are simple $U_q(\mathfrak{b})$-modules parametrized by a complex number $a \in \mathbb{C}^*$ and $1 \le i \le n$, where n is the rank of the underlying finite-dimensional simple Lie algebra. Such prefundamental representations were first constructed explicitly for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ by Bazhanov, Lukyanov, and Zamolodchikov [Bazhanov et al. 1997], for $\widehat{\mathfrak{sl}}_3$ by Bazhanov, Hibberd, and Khoroshkin [Bazhanov et al. 2002] and for $\widehat{\mathfrak{sl}}_n$ with $i = 1$ by Kojima [2008].

The category $\mathcal{O}$ and the prefundamental representations were used by E. Frenkel and Hernandez [2015] to prove a conjecture of Frenkel–Reshetikhin on the spectra of quantum integrable systems. Let us recall that the partition function Z of a quantum integrable system is crucial to understanding its physical properties. It may be written in terms of the eigenvalues λ_j of the transfer matrix T. Therefore, to compute Z one needs to find the spectrum of T. In his seminal paper, Baxter [1972] tackled this question for the 6-vertex and 8-vertex models. He observed moreover that the eigenvalues λ_j of T have the very remarkable form

$$\lambda_j = A(z)\frac{Q_j(zq^2)}{Q_j(z)} + D(z)\frac{Q_j(zq^{-2})}{Q_j(z)}, \tag{1-1}$$

where q, z are parameters of the model (quantum and spectral), the functions $A(z)$, $D(z)$ are universal (in the sense that they are the same for all eigenvalues), and Q_j is a polynomial. The above relation is now called *Baxter's relation* (or *Baxter's TQ relation*). Frenkel and Reshetikhin [1999] conjectured that the spectra of more general quantum integrable systems constructed from a representation V of a quantum affine algebra $U_q(\mathfrak{g})$ have a similar form. (In this framework, the 6-vertex model is the particular case when $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ and V is irreducible of dimension 2.) One of the main steps in the proof of this conjecture, given in [Frenkel and Hernandez 2015], is to interpret the expected generalized Baxter relations as algebraic identities in the Grothendieck ring of the category $\mathcal{O}$ for $U_q(\mathfrak{b})$ (see [Hernandez 2015] for a short overview). For example, if $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ and V is the 2-dimensional simple representation of $U_q(\mathfrak{g})$ with q-character $\chi_q(V) = Y_{1,a} + Y^{-1}_{1,aq^2}$, one gets the following categorical version of Baxter's relation (1-1):

$$[V \otimes L^+_{1,aq}] = [\omega_1][L^+_{1,aq^{-1}}] + [-\omega_1][L^+_{1,aq^3}]. \tag{1-2}$$

(Here, $[\omega_1]$ and $[-\omega_1]$ denote the classes of certain one-dimensional representations of $U_q(\mathfrak{b})$; see Definition 3.4 below.)

In another direction, the notion of monoidal categorification of cluster algebras was introduced in [Hernandez and Leclerc 2010]. The cluster algebra $\mathcal{A}(Q)$ attached to a quiver Q is a commutative $\mathbb{Z}$-algebra with a distinguished set of generators called cluster variables and obtained inductively via the Fomin–Zelevinsky procedure of mutation [2002]. By definition, the rank of $\mathcal{A}(Q)$ is the number of vertices of Q (finite or infinite). A monoidal category $\mathcal{C}$ is said to be a *monoidal categorification* of $\mathcal{A}(Q)$ if there exists a ring isomorphism

$$\mathcal{A}(Q) \xrightarrow{\sim} K_0(\mathcal{C})$$

which induces a bijection between cluster variables and classes of simple modules which are *prime* (without nontrivial tensor factorization) and *real* (with simple tensor square). Various examples of monoidal categorifications have been established in terms of quantum affine algebras [Hernandez and Leclerc 2010; 2016], perverse sheaves on quiver varieties [Nakajima 2011; Kimura and Qin 2014; Qin 2015], and Khovanov–Lauda–Rouquier algebras [Kang et al. 2014; 2015].

In this paper, we propose new monoidal categorifications of cluster algebras in terms of the category $\mathcal{O}$ of a Borel subalgebra $U_q(\mathfrak{b})$ of an untwisted quantum affine algebra $U_q(\mathfrak{g})$. More precisely, in [Hernandez and Leclerc 2016] we attached to $\mathfrak{g}$ a semi-infinite quiver G^- and we proved that the cluster algebra $\mathcal{A}(G^-)$ is isomorphic to the Grothendieck ring of a monoidal category $\mathcal{C}_{\mathbb{Z}}^-$ of finite-dimensional representations of $U_q(\mathfrak{g})$. Moreover, the classes of the Kirillov–Reshetikhin modules in $\mathcal{C}_{\mathbb{Z}}^-$ are the images under this isomorphism of a subset of the cluster variables. Let Γ be the doubly infinite quiver corresponding to G^-, as defined in [Hernandez and Leclerc 2016, Section 2.1.2]. The main result of this paper (Theorem 4.2) is that the completed cluster algebra $\mathcal{A}(\Gamma)$ attached to this doubly infinite quiver is isomorphic to the Grothendieck ring of a certain monoidal subcategory $\mathcal{O}_{2\mathbb{Z}}^+$ of $\mathcal{O}$. This subcategory $\mathcal{O}_{2\mathbb{Z}}^+$ is generated by finite-dimensional representations and positive prefundamental representations whose spectral parameters satisfy an integrality condition (see Definitions 3.8 and 4.1 and Proposition 5.16). Moreover, the classes of the positive prefundamental representations form the cluster variables of an initial seed of $\mathcal{A}(\Gamma)$. In particular, when $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ the counterparts (1-2) of Baxter's relations (1-1) get interpreted as instances of Fomin–Zelevinsky mutation relations.[1] For general types, the one-step mutation relations are interpreted as other remarkable relations in the Grothendieck ring $K_0(\mathcal{O}^+)$ (Formula (6-14)).

Along the way we get interesting additional results, for instance, (i) the construction of new asymptotic representations beyond the case of prefundamental representations (Theorem 7.6), and (ii) the tensor factorization of arbitrary simple modules of $\mathcal{O}^+$ into prime modules when $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ (Theorem 7.9). We conjecture that the category $\mathcal{O}_{2\mathbb{Z}}^+$ is a monoidal categorification of the cluster algebra $\mathcal{A}(\Gamma)$ (Conjecture 7.2). We prove this conjecture for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ (Theorem 7.11). An essential

[1] The generalized Baxter relations of [Frenkel and Hernandez 2015] can also be regarded as relations in the cluster algebra, but not as mutation relations since they involve more than 2 terms in general.

tool in several of our proofs is a duality D between the Grothendieck rings of certain subcategories $\mathcal{O}^+$ and $\mathcal{O}^-$ of $\mathcal{O}$ (Theorem 5.17), which maps classes of simple objects to classes of simple objects (Theorem 7.7).

The paper is organized as follows. In Section 2 we give some background on quantum affine (or loop) algebras and their Borel subalgebras. In Section 3 we review the main properties of the category $\mathcal{O}$ introduced in [Hernandez and Jimbo 2012] and we introduce the subcategories $\mathcal{O}^+$ and $\mathcal{O}^-$ of interest for this paper (Definition 3.9). In Section 4 we state the main result on the isomorphism between $\mathcal{A}(\Gamma)$ and the Grothendieck ring of $\mathcal{O}^+_{2\mathbb{Z}}$. In Section 5 we establish relevant properties of $\mathcal{O}^+$; in particular, we introduce and study the duality D between $\mathcal{O}^+$ and $\mathcal{O}^-$. The proof of Theorem 4.2 is given in Section 6. In Section 7 we present the conjecture on monoidal categorifications and we give various evidence supporting it, in particular the existence of asymptotic representations (Section 7B). To conclude we present additional conjectural relations in $K_0(\mathcal{O}^+)$, extending the generalized Baxter relations of [Frenkel and Hernandez 2015] (Conjecture 7.15).

The main results of this paper were presented in several conferences (the Oberwolfach workshop "Enveloping algebras and geometric representation theory" in May 2015, "Categorical representation theory and combinatorics" in Seoul in December 2015, and "A bridge between representation theory and physics" in Canterbury in January 2016). An announcement was also published in the Oberwolfach Report [Hernandez 2015].

2. Quantum loop algebra and Borel algebras

2A. ***Quantum loop algebra.*** Let $C = (C_{i,j})_{0\le i,j\le n}$ be an indecomposable Cartan matrix of nontwisted affine type. We denote by $\mathfrak{g}$ the Kac–Moody Lie algebra associated with C. Set $I = \{1, \dots, n\}$, and denote by $\dot{\mathfrak{g}}$ the finite-dimensional simple Lie algebra associated with the Cartan matrix $(C_{i,j})_{i,j\in I}$. Let $\{\alpha_i\}_{i\in I}$, $\{\alpha_i^\vee\}_{i\in I}$, $\{\omega_i\}_{i\in I}$, $\{\omega_i^\vee\}_{i\in I}$, $\dot{\mathfrak{h}}$ be the simple roots, the simple coroots, the fundamental weights, the fundamental coweights and the Cartan subalgebra of $\dot{\mathfrak{g}}$, respectively. We set

$$Q = \bigoplus_{i\in I} \mathbb{Z}\alpha_i, \quad Q^+ = \bigoplus_{i\in I} \mathbb{Z}_{\ge 0}\alpha_i, \quad P = \bigoplus_{i\in I} \mathbb{Z}\omega_i.$$

We will also use $P_{\mathbb{Q}} = P \otimes \mathbb{Q}$ with its partial ordering defined by $\omega \le \omega'$ if and only if $\omega' - \omega \in Q^+$. Let $D = \operatorname{diag}(d_0, \dots, d_n)$ be the unique diagonal matrix such that $B = DC$ is symmetric and the d_i are relatively prime positive integers. We denote by $(\ ,\ ) : Q \times Q \to \mathbb{Z}$ the invariant symmetric bilinear form such that $(\alpha_i, \alpha_i) = 2d_i$. We use the numbering of the Dynkin diagram as in [Kac 1990]. Let $a_0, \dots, a_n$ be the Kac labels as on pages 55–56 of the same work. We have $a_0 = 1$ and we set $\alpha_0 = -(a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n)$.

Throughout this paper, we fix a nonzero complex number q which is not a root of unity. We set $q_i = q^{d_i}$. We fix once and for all $h \in \mathbb{C}$ such that $q = e^h$, and we define $q^r = e^{rh}$ for any $r \in \mathbb{Q}$. Since q is not a root of unity, for $r, s \in \mathbb{Q}$ we have that $q^r = q^s$ if and only if $r = s$.

We will use the standard symbols for q-integers:

$$[m]_z = \frac{z^m - z^{-m}}{z - z^{-1}}, \quad [m]_z! = \prod_{j=1}^{m} [j]_z, \quad \begin{bmatrix} s \\ r \end{bmatrix}_z = \frac{[s]_z!}{[r]_z![s-r]_z!}.$$

The *quantum loop algebra* $U_q(\mathfrak{g})$ is the $\mathbb{C}$-algebra defined by generators $e_i, f_i, k_i^{\pm 1}$ $(0 \le i \le n)$ and the following relations for $0 \le i, j \le n$:

$$k_i k_j = k_j k_i, \qquad [e_i, f_j] = \delta_{i,j} \frac{k_i - k_i^{-1}}{q_i - q_i^{-1}},$$

$$k_0^{a_0} k_1^{a_1} \cdots k_n^{a_n} = 1, \qquad 0 = \sum_{r=0}^{1-C_{i,j}} (-1)^r e_i^{(1-C_{i,j}-r)} e_j e_i^{(r)} \quad (i \ne j),$$

$$k_i e_j k_i^{-1} = q_i^{C_{i,j}} e_j,$$

$$k_i f_j k_i^{-1} = q_i^{-C_{i,j}} f_j, \qquad 0 = \sum_{r=0}^{1-C_{i,j}} (-1)^r f_i^{(1-C_{i,j}-r)} f_j f_i^{(r)} \quad (i \ne j).$$

Here we have set $x_i^{(r)} = x_i^r / [r]_{q_i}!$ $(x_i = e_i, f_i)$. The algebra $U_q(\mathfrak{g})$ has a Hopf algebra structure given by

$$\Delta(e_i) = e_i \otimes 1 + k_i \otimes e_i, \qquad S(e_i) = -k_i^{-1} e_i,$$

$$\Delta(f_i) = f_i \otimes k_i^{-1} + 1 \otimes f_i, \qquad S(f_i) = -f_i k_i,$$

$$\Delta(k_i) = k_i \otimes k_i, \qquad S(k_i) = k_i^{-1},$$

where $i = 0, \dots, n$.

The algebra $U_q(\mathfrak{g})$ can also be presented in terms of the Drinfeld generators [Drinfeld 1987; Beck 1994]

$$x_{i,r}^{\pm} \quad (i \in I,\ r \in \mathbb{Z}), \qquad \phi_{i,\pm m}^{\pm} \quad (i \in I,\ m \ge 0), \qquad k_i^{\pm 1} \quad (i \in I).$$

Example 2.1. For $\dot{\mathfrak{g}} = \mathfrak{sl}_2$, we have $e_1 = x_{1,0}^{+},\ e_0 = k_1^{-1} x_{1,1}^{-},\ f_1 = x_{1,0}^{-},\ f_0 = x_{1,-1}^{+} k_1$.

We shall use the generating series $(i \in I)$

$$\phi_i^{\pm}(z) = \sum_{m \ge 0} \phi_{i,\pm m}^{\pm} z^{\pm m} = k_i^{\pm 1} \exp\Bigl(\pm (q_i - q_i^{-1}) \sum_{m > 0} h_{i,\pm m} z^{\pm m}\Bigr).$$

We also set $\phi_{i,\pm m}^{\pm} = 0$ for $m < 0,\ i \in I$.

The algebra $U_q(\mathfrak{g})$ has a $\mathbb{Z}$-grading defined by $\deg(e_i) = \deg(f_i) = \deg(k_i^{\pm 1}) = 0$ for $i \in I$ and $\deg(e_0) = -\deg(f_0) = 1$. It satisfies $\deg(x_{i,m}^{\pm}) = \deg(\phi_{i,m}^{\pm}) = m$ for

$i \in I,\ m \in \mathbb{Z}$. For $a \in \mathbb{C}^*$, there is a corresponding automorphism $\tau_a : U_q(\mathfrak{g}) \to U_q(\mathfrak{g})$ such that $\tau_a(x) = a^m x$ for every $x \in U_q(\mathfrak{g})$ of degree $m \in \mathbb{Z}$.

By Proposition 1.6 of [Chari 1995], there exists an involutive automorphism $\hat{\omega} : U_q(\mathfrak{g}) \to U_q(\mathfrak{g})$ defined by ($i \in I,\ m, r \in \mathbb{Z},\ r \neq 0$)

$$\hat{\omega}(x^{\pm}_{i,m}) = -x^{\mp}_{i,-m}, \quad \hat{\omega}(\phi^{\pm}_{i,\pm m}) = \phi^{\mp}_{i,\mp m}, \quad \hat{\omega}(h_{i,r}) = -h_{i,-r}.$$

Let $U_q(\mathfrak{g})^{\pm}$ (resp. $U_q(\mathfrak{g})^0$) be the subalgebra of $U_q(\mathfrak{g})$ generated by the $x^{\pm}_{i,r}$, where $i \in I,\ r \in \mathbb{Z}$ (resp. by the $\phi^{\pm}_{i,\pm r}$, where $i \in I,\ r \geq 0$). We have a triangular decomposition [Beck 1994]

$$U_q(\mathfrak{g}) \simeq U_q(\mathfrak{g})^- \otimes U_q(\mathfrak{g})^0 \otimes U_q(\mathfrak{g})^+. \tag{2-3}$$

2B. ***Borel algebra.***

Definition 2.2. The Borel algebra $U_q(\mathfrak{b})$ is the subalgebra of $U_q(\mathfrak{g})$ generated by e_i and $k_i^{\pm 1}$ with $0 \leq i \leq n$.

This is a Hopf subalgebra of $U_q(\mathfrak{g})$. The algebra $U_q(\mathfrak{b})$ contains the Drinfeld generators $x^+_{i,m},\ x^-_{i,r},\ k_i^{\pm 1},\ \phi^+_{i,r}$, where $i \in I,\ m \geq 0$ and $r > 0$.

Let $U_q(\mathfrak{b})^{\pm} = U_q(\mathfrak{g})^{\pm} \cap U_q(\mathfrak{b})$ and $U_q(\mathfrak{b})^0 = U_q(\mathfrak{g})^0 \cap U_q(\mathfrak{b})$. Then we have

$$U_q(\mathfrak{b})^+ = \langle x^+_{i,m} \rangle_{i \in I, m \geq 0}, \quad U_q(\mathfrak{b})^0 = \langle \phi^+_{i,r}, k_i^{\pm 1} \rangle_{i \in I, r > 0}.$$

It follows from [Beck 1994; Damiani 1998] that we have a triangular decomposition

$$U_q(\mathfrak{b}) \simeq U_q(\mathfrak{b})^- \otimes U_q(\mathfrak{b})^0 \otimes U_q(\mathfrak{b})^+. \tag{2-4}$$

3. Representations of Borel algebras

In this section we review results on representations of the Borel algebra $U_q(\mathfrak{b})$, in particular on the category $\mathcal{O}$ defined in [Hernandez and Jimbo 2012] and on finite-dimensional representations of $U_q(\mathfrak{g})$. We also introduce the subcategories $\mathcal{O}^+$ and $\mathcal{O}^-$ of particular interest for this paper.

3A. ***Highest ℓ-weight modules.*** For a $U_q(\mathfrak{b})$-module V and $\omega \in P_{\mathbb{Q}}$, we have the weight space

$$V_\omega = \{v \in V \mid k_i v = q_i^{\omega(\alpha_i^\vee)} v \text{ for all } i \in I\}. \tag{3-5}$$

We say that V is *Cartan-diagonalizable* if $V = \bigoplus_{\omega \in P_{\mathbb{Q}}} V_\omega$.

For V a Cartan-diagonalizable $U_q(\mathfrak{b})$-module, we define the structure of a $U_q(\mathfrak{b})$-module on its graded dual $V^* = \bigoplus_{\beta \in P_{\mathbb{Q}}} V^*_\beta$ by

$$(xu)(v) = u(S^{-1}(x)v) \quad (u \in V^*,\ v \in V,\ x \in U_q(\mathfrak{b})).$$

Definition 3.1. A series $\mathbf{\Psi} = (\Psi_{i,m})_{i \in I, m \geq 0}$ of complex numbers such that $\Psi_{i,0} \in q_i^{\mathbb{Q}}$ for all $i \in I$ is called an ℓ-weight.

We denote by P_ℓ the set of ℓ-weights. Identifying $(\Psi_{i,m})_{m\geq 0}$ with its generating series, we shall write

$$\boldsymbol{\Psi} = (\Psi_i(z))_{i\in I}, \quad \Psi_i(z) = \sum_{m\geq 0} \Psi_{i,m} z^m.$$

We will often use the involution $\boldsymbol{\Psi} \mapsto \overline{\boldsymbol{\Psi}}$ on P_ℓ, where $\overline{\boldsymbol{\Psi}}$ is obtained from $\boldsymbol{\Psi}$ by replacing every pole and zero of $\boldsymbol{\Psi}$ by its inverse.

Since each $\Psi_i(z)$ is an invertible formal power series, P_ℓ has a natural group structure by componentwise multiplication. We have a surjective morphism of groups $\varpi : P_\ell \to P_{\mathbb{Q}}$ given by $\Psi_i(0) = q_i^{\varpi(\boldsymbol{\Psi})(\alpha_i^\vee)}$. For a $U_q(\mathfrak{b})$-module V and $\boldsymbol{\Psi} \in P_\ell$, the linear subspace

$$V_{\boldsymbol{\Psi}} = \{v \in V \mid \text{there exists a } p \geq 0 \text{ such that, for all } i \in I \text{ and } m \geq 0,\ (\phi^+_{i,m} - \Psi_{i,m})^p v = 0\} \tag{3-6}$$

is called the ℓ-weight space of V of ℓ-weight $\boldsymbol{\Psi}$. Note that since $\phi^+_{i,0} = k_i$, we have $V_{\boldsymbol{\Psi}} \subset V_\omega$, where $\omega = \varpi(\boldsymbol{\Psi})$.

Definition 3.2. A $U_q(\mathfrak{b})$-module V is said to be of highest ℓ-weight $\boldsymbol{\Psi} \in P_\ell$ if there is $v \in V$ such that $V = U_q(\mathfrak{b})v$ and the following hold:

$$e_i v = 0 \quad (i \in I), \qquad \phi^+_{i,m} v = \Psi_{i,m} v \quad (i \in I,\ m \geq 0).$$

The ℓ-weight $\boldsymbol{\Psi} \in P_\ell$ is uniquely determined by V. It is called the highest ℓ-weight of V. The vector v is said to be a highest ℓ-weight vector of V. For any $\boldsymbol{\Psi} \in P_\ell$, there exists a simple highest ℓ-weight module $L(\boldsymbol{\Psi})$ of highest ℓ-weight $\boldsymbol{\Psi}$. This module is unique up to isomorphism.

For $\boldsymbol{\Psi}$ an ℓ-weight, we set $\widetilde{\boldsymbol{\Psi}} = (\varpi(\boldsymbol{\Psi}))^{-1}\boldsymbol{\Psi}$ and we introduce the simple $U_q(\mathfrak{b})$-module

$$\widetilde{L}(\boldsymbol{\Psi}) := L(\widetilde{\boldsymbol{\Psi}}).$$

This is the simple $U_q(\mathfrak{b})$-module obtained from $L(\boldsymbol{\Psi})$ by shifting all ℓ-weights by $\varpi(\boldsymbol{\Psi})^{-1}$ (see [Hernandez and Jimbo 2012, Remark 2.5]).

The submodule of $L(\boldsymbol{\Psi}) \otimes L(\boldsymbol{\Psi}')$ generated by the tensor product of the highest ℓ-weight vectors is of highest ℓ-weight $\boldsymbol{\Psi}\boldsymbol{\Psi}'$. In particular, $L(\boldsymbol{\Psi}\boldsymbol{\Psi}')$ is a subquotient of $L(\boldsymbol{\Psi}) \otimes L(\boldsymbol{\Psi}')$.

Definition 3.3 [Hernandez and Jimbo 2012]. For $i \in I$ and $a \in \mathbb{C}^\times$, let

$$L^\pm_{i,a} = L(\boldsymbol{\Psi}^{\pm 1}_{i,a}), \quad \text{where } (\boldsymbol{\Psi}^{\pm 1}_{i,a})_j(z) = \begin{cases} (1 - za)^{\pm 1}, & j = i, \\ 1, & j \neq i. \end{cases} \tag{3-7}$$

We call $L^+_{i,a}$ (resp. $L^-_{i,a}$) a positive (resp. negative) *prefundamental representation.*

Definition 3.4. For $\omega \in P_{\mathbb{Q}}$, let $[\omega] = L(\boldsymbol{\Psi}_\omega)$, where $(\boldsymbol{\Psi}_\omega)_i(z) = q_i^{\omega(\alpha_i^\vee)}$ $(i \in I)$.

Note that the representation $[\omega]$ is 1-dimensional with a trivial action of $e_0, \ldots, e_n$.

For $a \in \mathbb{C}^\times$, the subalgebra $U_q(\mathfrak{b})$ is stable by τ_a. Denote its restriction to $U_q(\mathfrak{b})$ by the same letter. Then the pullbacks of the $U_q(\mathfrak{b})$-modules $L^\pm_{i,b}$ by τ_a is $L^\pm_{i,ab}$.

3B. *Category* $\mathcal{O}$. For $\lambda \in P_\mathbb{Q}$, we set $D(\lambda) = \{\omega \in P_\mathbb{Q} \mid \omega \leq \lambda\}$.

Definition 3.5 [Hernandez and Jimbo 2012]. A $U_q(\mathfrak{b})$-module V is said to be in category $\mathcal{O}$ if:

(i) V is Cartan-diagonalizable;

(ii) for all $\omega \in P_\mathbb{Q}$ we have $\dim(V_\omega) < \infty$;

(iii) there exist a finite number of elements $\lambda_1, \ldots, \lambda_s \in P_\mathbb{Q}$ such that the weights of V are in $\bigcup_{j=1,\ldots,s} D(\lambda_j)$.

The category $\mathcal{O}$ is a monoidal category.

Remark 3.6. The definition of $\mathcal{O}$ is slightly different from that in [Hernandez and Jimbo 2012] as we allow only rational powers of q for the eigenvalues of k_i.

Let $P^\mathfrak{r}_\ell$ be the subgroup of P_ℓ consisting of $\mathbf{\Psi}$ such that $\Psi_i(z)$ is a rational function of z for any $i \in I$.

Theorem 3.7 [Hernandez and Jimbo 2012]. *Let $\mathbf{\Psi} \in P_\ell$. A simple object in the category $\mathcal{O}$ is of highest ℓ-weight and the simple module $L(\mathbf{\Psi})$ is in category $\mathcal{O}$ if and only if $\mathbf{\Psi} \in P^\mathfrak{r}_\ell$. Moreover, for V in category $\mathcal{O}$, $V_\mathbf{\Psi} \neq 0$ implies $\mathbf{\Psi} \in P^\mathfrak{r}_\ell$.*

Given a map $c : P^\mathfrak{r}_\ell \to \mathbb{Z}$, consider its support

$$\operatorname{supp}(c) = \{\mathbf{\Psi} \in P^\mathfrak{r}_\ell \mid c(\mathbf{\Psi}) \neq 0\}.$$

Let $\mathcal{E}_\ell$ be the additive group of maps $c : P^\mathfrak{r}_\ell \to \mathbb{Z}$ such that $\varpi(\operatorname{supp}(c))$ is contained in a finite union of sets of the form $D(\mu)$, and such that for every $\omega \in P_\mathbb{Q}$ the set $\operatorname{supp}(c) \cap \varpi^{-1}(\{\omega\})$ is finite. Similarly, let $\mathcal{E}$ be the additive group of maps $c : P_\mathbb{Q} \to \mathbb{Z}$ whose support is contained in a finite union of sets of the form $D(\mu)$. The map ϖ is naturally extended to a surjective homomorphism $\varpi : \mathcal{E}_\ell \to \mathcal{E}$.

As for the category $\mathcal{O}$ of a classical Kac–Moody Lie algebra, the multiplicity of a simple module in a module of our category $\mathcal{O}$ is well-defined (see [Kac 1990, Section 9.6]) and we have the Grothendieck ring $K_0(\mathcal{O})$. Its elements are the formal sums

$$\chi = \sum_{\mathbf{\Psi} \in P^\mathfrak{r}_\ell} \lambda_\mathbf{\Psi} [L(\mathbf{\Psi})],$$

where the $\lambda_\mathbf{\Psi} \in \mathbb{Z}$ are set so that

$$\sum_{\mathbf{\Psi} \in P^\mathfrak{r}_\ell,\, \omega \in P_\mathbb{Q}} |\lambda_\mathbf{\Psi}| \dim((L(\mathbf{\Psi}))_\omega)[\omega]$$

is in $\mathcal{E}$. For V, V' representations in $\mathcal{O}$, the product $[V].[V']$ in $K_0(\mathcal{O})$ is naturally obtained by considering the multiplicities of simple modules in $V \otimes V'$ in the sense of [Kac 1990, Section 9.6].

We naturally identify $\mathcal{E}$ with the Grothendieck ring of the category of representations of $\mathcal{O}$ with constant ℓ-weights, the simple objects of which are the $[\omega]$, $\omega \in P_{\mathbb{Q}}$. Thus as in [Kac 1990, Section 9.7] we will regard elements of $\mathcal{E}$ as formal sums

$$c = \sum_{\omega \in \mathrm{supp}(c)} c(\omega)[\omega].$$

The multiplication is given by $[\omega][\omega'] = [\omega + \omega']$ and $\mathcal{E}$ is regarded as a subring of $K_0(\mathcal{O})$. If $(c_i)_{i \in \mathbb{N}}$ is a countable family of elements of $\mathcal{E}$ such that, for any $\omega \in P_{\mathbb{Q}}$, we have $c_i(\omega) \neq 0$ for finitely many $i \in \mathbb{N}$, then the sum $\sum_{i \in \mathbb{N}} c_i$ is well-defined as a map $P_{\mathbb{Q}} \to \mathbb{Z}$. When this map is in $\mathcal{E}$ we say that $\sum_{i \in \mathbb{N}} c_i$ is a countable sum of elements in $\mathcal{E}$. Note that we have the analogous notion of a countable sum in $K_0(\mathcal{O})$, compatible with countable sums of characters in $\mathcal{E}$.

3C. *Finite-dimensional representations.* Let $\mathcal{C}$ be the category of (type-1) finite-dimensional representations of $U_q(\mathfrak{g})$.

For $i \in I$, let $P_i(z) \in \mathbb{C}[z]$ be a polynomial with constant term 1. Set

$$\boldsymbol{\Psi} = (\Psi_i(z))_{i \in I}, \quad \Psi_i(z) = q_i^{\deg(P_i)} \frac{P_i(zq_i^{-1})}{P_i(zq_i)}.$$

Then $L(\boldsymbol{\Psi})$ is finite-dimensional. Moreover the action of $U_q(\mathfrak{b})$ can be uniquely extended to an action of $U_q(\mathfrak{g})$, and any simple object in the category $\mathcal{C}$ is of this form. Hence $\mathcal{C}$ is a subcategory of $\mathcal{O}$ and the inclusion functor preserves simple objects.

For $i \in I$ and $a \in \mathbb{C}^*$, we denote by $V_{i,a}$ the simple finite-dimensional representation associated with the polynomials

$$P_i(z) = 1 - za, \quad P_j(z) = 1 \quad (j \neq i).$$

The modules $V_{i,a}$ are called the *fundamental representations*.

3D. *The categories $\mathcal{O}^+$ and $\mathcal{O}^-$.* We introduce two new subcategories $\mathcal{O}^+$ and $\mathcal{O}^-$ of the category $\mathcal{O}$.

Definition 3.8. An ℓ-weight is said to be positive (resp. negative) if it is a monomial in the following ℓ-weights:

- $Y_{i,a} = q_i \boldsymbol{\Psi}_{i,aq_i}^{-1} \boldsymbol{\Psi}_{i,aq_i^{-1}}$, where $i \in I,\ a \in \mathbb{C}^*$,
- $\boldsymbol{\Psi}_{i,a}$ (resp. $\boldsymbol{\Psi}_{i,a}^{-1}$), where $i \in I,\ a \in \mathbb{C}^*$,
- $[\omega]$, where $\omega \in P_{\mathbb{Q}}$.

Definition 3.9. $\mathcal{O}^+$ (resp. $\mathcal{O}^-$) is the category of representations in $\mathcal{O}$ whose simple constituents have a positive (resp. negative) highest ℓ-weight.

Remark 3.10. (i) By construction, $\mathcal{O}^+$, $\mathcal{O}^-$ are stable by extensions. We will prove they are also stable by tensor products (Theorem 5.17).

(ii) There are other remarkable subcategories of $\mathcal{O}$, for example, the category $\widehat{\mathcal{O}}$ of representations of $U_q(\mathfrak{g})$ which belong to $\mathcal{O}$ as representations of $U_q(\mathfrak{b})$. This category $\widehat{\mathcal{O}}$ was introduced in [Hernandez 2005] and further studied in [Mukhin and Young 2014].

(iii) One motivation of Definition 3.9 is that $\mathcal{O}^\pm$ contains $\mathcal{C}$ as well as the prefundamental representations $L^\pm_{i,a}$. We have the following inclusion diagram:

$$\begin{array}{ccc} \mathcal{O} & \supset & \mathcal{O}^+, \mathcal{O}^- \\ \cup & & \cup \\ \widehat{\mathcal{O}} & \supset & \mathcal{C} \end{array}$$

Note that $\widehat{\mathcal{O}}$ is not contained in $\mathcal{O}^+$ or $\mathcal{O}^-$, and conversely neither $\mathcal{O}^+$ nor $\mathcal{O}^-$ is contained in $\widehat{\mathcal{O}}$. For instance, for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, the representation $L((1-zq)/(q-z))$ is in the category $\widehat{\mathcal{O}}$ by [Mukhin and Young 2014, Theorem 3.6], but not in $\mathcal{O}^+$ or $\mathcal{O}^-$ because its highest ℓ-weight has no factorization as in equation (5-12) below. On the other hand, the prefundamental representations $L^\pm_{1,a}$ are in the category $\mathcal{O}^\pm$ but not in $\widehat{\mathcal{O}}$ (see [Hernandez and Jimbo 2012, Section 4.1] or [Mukhin and Young 2014, Theorem 3.6]).

(iv) All generalized Baxter's relations established in [Frenkel and Hernandez 2015] hold in the Grothendieck rings $K_0(\mathcal{O}^+)$ or $K_0(\mathcal{O}^-)$ (see Theorem 5.5 below).

(v) The factorization of real simple modules in $\mathcal{O}$ into prime representations is not unique, so the full category $\mathcal{O}$ is not a good candidate for the notion of monoidal categorification discussed in the introduction. For example, for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, it follows from [Mukhin and Young 2014, Remark 4.3, Theorem 4.6] that

$$L\left(q^{-5}\frac{1-q^4z}{1-q^{-6}z}\right)\otimes L\left(q^{-9}\frac{1-q^8z}{1-q^{-10}z}\right)\simeq L\left(q^{-7}\frac{1-q^4z}{1-q^{-10}z}\right)\otimes L\left(q^{-7}\frac{1-q^8z}{1-q^{-6}z}\right).$$

Moreover the tensor product is simple real and each simple factor has the character of a Verma module. Consequently each factor is not isomorphic to a tensor product of prefundamental representations and so is prime.

4. Cluster algebras

We state the main results of this paper. We refer the reader to [Fomin and Zelevinsky 2003] and [Gekhtman et al. 2010] for an introduction to cluster algebras, and for any undefined terminology.

4A. ***An infinite-rank cluster algebra.*** Let us recall the infinite quiver G introduced in [Hernandez and Leclerc 2016, Section 2.1.3]. Put $\widetilde{V} = I \times \mathbb{Z}$. Let $\widetilde{\Gamma}$ be the quiver with vertex set $\widetilde{V}$. The arrows of $\widetilde{\Gamma}$ are given by

$$((i,r) \to (j,s)) \iff (C_{i,j} \neq 0 \text{ and } s = r + d_i C_{i,j}).$$

By [Hernandez and Leclerc 2016], the quiver $\widetilde{\Gamma}$ has two isomorphic connected components. We pick one of the two isomorphic connected components of $\widetilde{\Gamma}$ and call it Γ. The vertex set of Γ is denoted by V. A second labeling of the vertices of Γ is deduced from the first one by means of the function ψ defined by

$$\psi(i,r) = (i, r + d_i) \quad ((i,r) \in V). \tag{4-8}$$

Let $W \subset I \times \mathbb{Z}$ be the image of V under ψ. We shall denote by G the same quiver as Γ but with vertices labeled by W.

By analogy with [Hernandez and Leclerc 2016, Section 2.2.1], consider an infinite set of indeterminates

$$z = \{z_{i,r} \mid (i,r) \in V\}$$

over $\mathbb{Q}$. Let $\mathcal{A}(\Gamma)$ be the cluster algebra defined by the initial seed (z, Γ). Thus, $\mathcal{A}(\Gamma)$ is the subring of the field of rational functions $\mathbb{Q}(z)$ generated by all the cluster variables, that is, the elements obtained from some element of z via a finite sequence of seed mutations. Each element of $\mathcal{A}(\Gamma)$ is a linear combination of finite monomials in some cluster variables. By the Laurent phenomenon [Fomin and Zelevinsky 2002], $\mathcal{A}(\Gamma)$ is contained in $\mathbb{Z}[z_{i,r}^{\pm 1}]_{(i,r)\in V}$.

For our purposes, it is always possible to work with sufficiently large finite subseeds of the seed (z, Γ), and replace $\mathcal{A}(\Gamma)$ by the genuine cluster subalgebras attached to them. On the other hand, statements become nicer if we allow ourselves to formulate them in terms of the infinite-rank cluster algebra $\mathcal{A}(\Gamma)$.

Define an $\mathcal{E}$-algebra homomorphism $\chi : \mathbb{Z}[z_{i,r}^{\pm 1}] \otimes_{\mathbb{Z}} \mathcal{E} \to \mathcal{E}$ by setting

$$\chi(z_{i,r}^{\pm 1}) = \left[\left(\frac{\mp r}{2d_i}\right)\omega_i\right] \quad ((i,r) \in V).$$

For $A \in \mathcal{A}(\Gamma) \otimes_{\mathbb{Z}} \mathcal{E}$, we write $\chi(A) = \sum_\omega A_\omega[\omega]$ and $|\chi|(A) = \sum_\omega |A_\omega|[\omega]$. We will consider the completed tensor product

$$\mathcal{A}(\Gamma) \mathbin{\hat{\otimes}_{\mathbb{Z}}} \mathcal{E},$$

that is, the algebra of countable sums $\sum_{i\in\mathbb{N}} A_i$ of elements $A_i \in \mathcal{A}(\Gamma) \otimes_{\mathbb{Z}} \mathcal{E}$ such that $\sum_{i\in\mathbb{N}} |\chi|(A_i)$ is in $\mathcal{E}$ as a countable sum (as defined in Section 3B). Note that in particular we have the analogous notion of a countable sum in $\mathcal{A}(\Gamma) \mathbin{\hat{\otimes}_{\mathbb{Z}}} \mathcal{E}$.

4B. ***Main theorem.***

Definition 4.1. Define the category $\mathcal{O}^+_{2\mathbb{Z}}$ as the subcategory of representations in $\mathcal{O}^+$ whose simple constituents have a highest ℓ-weight $\mathbf{\Psi}$ such that the roots and the poles of $\mathbf{\Psi}_i(z)$ are of the form q^r with $(i,r) \in V$.

We will write for short $\ell^+_{i,a} = [L^+_{i,a}]$.

Theorem 4.2. *The category $\mathcal{O}^+_{2\mathbb{Z}}$ is monoidal and the identification*

$$z_{i,r} \otimes \left[\frac{r}{2d_i}\omega_i\right] \equiv \ell^+_{i,q^r} \quad ((i,r) \in V) \tag{4-9}$$

defines an isomorphism of $\mathcal{E}$-algebras

$$\mathcal{A}(\Gamma)\,\hat{\otimes}_{\mathbb{Z}}\,\mathcal{E} \simeq K_0(\mathcal{O}^+_{2\mathbb{Z}})$$

compatible with countable sums.

Remark 4.3. (i) The identification (4-9) gives an isomorphism of $\mathcal{E}$-algebras

$$\mathcal{E}[z^{\pm 1}_{i,r}]_{(i,r)\in V} \xrightarrow{\sim} \mathcal{E}[(\ell^+_{i,q^r})^{\pm 1}]_{(i,r)\in V},$$

which can be extended to countable sums as above. So the main point of the proof of Theorem 4.2 will be to show that the subalgebra $\mathcal{A}(G)\,\hat{\otimes}_{\mathbb{Z}}\,\mathcal{E}$ is mapped to $K_0(\mathcal{O}^+_{2\mathbb{Z}})$ by this isomorphism.

(ii) As in the case of finite-dimensional representations, the description of the simple objects of $\mathcal{O}^+$ essentially reduces to the description of the simple objects of $\mathcal{O}^+_{2\mathbb{Z}}$ (the decomposition explained in [Hernandez and Leclerc 2010, Section 3.7] can be extended to our more general situation by using the asymptotic approach of Section 7B below). Hence the Grothendieck ring $K_0(\mathcal{O}^+_{2\mathbb{Z}})$ contains all the interesting information on $K_0(\mathcal{O}^+)$.

The proof of Theorem 4.2 will be given in Section 6, using material presented in Section 5.

5. Properties of the category $\mathcal{O}^+$

5A. ***q*-*characters.*** The q-character morphism was first considered in [Frenkel and Reshetikhin 1999] and is a very useful tool for our proofs.

Recall from Section 3B the notation $\mathcal{E}$ and $\mathcal{E}_\ell$. Because of the support condition, we can endow $\mathcal{E}$ with a ring structure defined by

$$(c \cdot d)(\omega) = \sum_{\omega'+\omega''=\omega} c(\omega')d(\omega'') \quad (c, d \in \mathcal{E},\ \omega \in P_{\mathbb{Q}}).$$

Similarly, $\mathcal{E}_\ell$ also has a ring structure given by

$$(c \cdot d)(\mathbf{\Psi}) = \sum_{\mathbf{\Psi}'\mathbf{\Psi}''=\mathbf{\Psi}} c(\mathbf{\Psi}')d(\mathbf{\Psi}'') \quad (c, d \in \mathcal{E}_\ell,\ \mathbf{\Psi} \in P_\ell^{\mathfrak{r}}),$$

and such that ϖ becomes a ring homomorphism.

For $\mathbf{\Psi} \in P_\ell^{\mathfrak{r}}$ and $\omega \in P_{\mathbb{Q}}$, we define the delta functions $[\mathbf{\Psi}] = \delta_{\mathbf{\Psi},\cdot} \in \mathcal{E}_\ell$ and $[\omega] = \delta_{\omega,\cdot} \in \mathcal{E}$, where as usual δ denotes the Kronecker symbol. Note that the above multiplications give

$$[\mathbf{\Psi}'] \cdot [\mathbf{\Psi}''] = [\mathbf{\Psi}'\mathbf{\Psi}''], \quad [\omega'] \cdot [\omega''] = [\omega' + \omega''].$$

Let V be a $U_q(\mathfrak{b})$-module in category $\mathcal{O}$. We define [Frenkel and Reshetikhin 1999; Hernandez and Jimbo 2012] the q-character and the character of V:

$$\chi_q(V) := \sum_{\mathbf{\Psi} \in P_\ell^{\mathfrak{r}}} \dim(V_{\mathbf{\Psi}})[\mathbf{\Psi}] \in \mathcal{E}_\ell, \quad \chi(V) := \varpi(\chi_q(V)) = \sum_{\omega \in P_{\mathbb{Q}}} \dim(V_\omega)[\omega] \in \mathcal{E}.$$

If $V \in \mathcal{O}$ has a unique ℓ-weight $\mathbf{\Psi}$ whose weight $\varpi(\mathbf{\Psi})$ is maximal, we also consider its normalized q-character $\widetilde{\chi}_q(V)$ and normalized character $\widetilde{\chi}(V)$ defined by

$$\widetilde{\chi}_q(V) := [\mathbf{\Psi}^{-1}] \cdot \chi_q(V), \quad \widetilde{\chi}(V) := \varpi(\widetilde{\chi}_q(V)).$$

Note that

$$\chi_q(\widetilde{L}(\mathbf{\Psi})) = [\widetilde{\mathbf{\Psi}}] \cdot \widetilde{\chi}_q(L(\mathbf{\Psi})) \neq \widetilde{\chi}_q(L(\mathbf{\Psi})).$$

Proposition 5.1 [Hernandez and Jimbo 2012]. *The q-character morphism*

$$\chi_q : K_0(\mathcal{O}) \to \mathcal{E}_\ell, \quad [V] \mapsto \chi_q(V),$$

is an injective ring morphism.

Following [Frenkel and Reshetikhin 1999], consider the ring of Laurent polynomials $\mathcal{Y} = \mathbb{Z}[Y_{i,a}^{\pm 1}]_{i \in I, a \in \mathbb{C}^*}$ in the indeterminates $\{Y_{i,a}\}_{i \in I, a \in \mathbb{C}^*}$. Let $\mathcal{M}$ be the multiplicative group of Laurent monomials in $\mathcal{Y}$. For example, for $i \in I$ and $a \in \mathbb{C}^*$ define $A_{i,a} \in \mathcal{M}$ by

$$A_{i,a} = Y_{i,aq_i^{-1}} Y_{i,aq_i} \Biggl(\prod_{j : C_{j,i}=-1} Y_{j,a} \prod_{j : C_{j,i}=-2} Y_{j,aq^{-1}} Y_{j,aq} \prod_{j : C_{j,i}=-3} Y_{j,aq^{-2}} Y_{j,a} Y_{j,aq^2} \Biggr)^{-1}.$$

For a monomial $m = \prod_{i \in I, a \in \mathbb{C}^*} Y_{i,a}^{u_{i,a}} \in \mathcal{M}$, we consider its "evaluation at $\phi^+(z)$". By definition it is the element $m(\phi(z)) \in P_\ell^{\mathfrak{r}}$ given by

$$m(\phi(z)) = \prod_{i \in I, a \in \mathbb{C}^*} (Y_{i,a}(\phi(z)))^{u_{i,a}}, \quad \text{where } (Y_{i,a}(\phi(z)))_j = \begin{cases} q_i \frac{1 - aq_i^{-1} z}{1 - aq_i z}, & j = i, \\ 1, & j \neq i. \end{cases}$$

This defines an injective group morphism $\mathcal{M} \to P_\ell^{\mathfrak{r}}$. We identify a monomial $m \in \mathcal{M}$ with its image in $P_\ell^{\mathfrak{r}}$. This is compatible with the notation $Y_{i,a}$ used in Definition 3.8. Note that $\varpi(Y_{i,a}) = \omega_i$ and $\varpi(A_{i,a}) = \alpha_i$.

It is proved in [Frenkel and Reshetikhin 1999] that a finite-dimensional $U_q(\mathfrak{g})$-module V satisfies $V = \bigoplus_{m\in\mathcal{M}} V_{m(\phi(z))}$. In particular, $\chi_q(V)$ can be viewed as an element of $\mathcal{Y}$.

A monomial $M \in \mathcal{M}$ is said to be dominant if $M \in \mathbb{Z}[Y_{i,a}]_{i\in I, a\in\mathbb{C}^*}$. Given a finite-dimensional simple $U_q(\mathfrak{g})$-module $L(\mathbf{\Psi})$, there exists a dominant monomial $M \in \mathcal{M}$ such that $\mathbf{\Psi} = M(\phi(z))$. We will also set $L(\mathbf{\Psi}) = L(M)$.

For example, for $i \in I$, $a \in \mathbb{C}^*$, $k \geq 0$, we have the Kirillov–Reshetikhin module

$$W_{k,a}^{(i)} = L(Y_{i,a} Y_{i,aq_i^2} \cdots Y_{i,aq_i^{2(k-1)}}). \tag{5-10}$$

Example 5.2. If $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, we have ($k \geq 0$, $a \in \mathbb{C}^*$) [Frenkel and Reshetikhin 1999]

$$\begin{aligned}\chi_q(W_{k,aq^{1-2k}}^{(1)}) \\ = Y_{aq^{-1}} Y_{aq^{-3}} \cdots Y_{aq^{-2k+1}} (1 + A_{1,a}^{-1} + A_{1,a}^{-1} A_{1,aq^{-2}}^{-1} + \cdots + A_{1,a}^{-1} \cdots A_{1,aq^{-2(k-1)}}^{-1}).\end{aligned}$$

Theorem 5.3. (i) [Hernandez and Jimbo 2012; Frenkel and Hernandez 2015]. *For any $a \in \mathbb{C}^*$, $i \in I$ we have*

$$\chi_q(L_{i,a}^+) = [\mathbf{\Psi}_{i,a}]\chi(L_{i,a}^+) = [\mathbf{\Psi}_{i,a}]\chi(L_{i,a}^-),$$

where $\chi(L_{i,a}^+) = \chi(L_{i,a}^-)$ does not depend on a.

(ii) [Hernandez and Jimbo 2012]. *For any $a \in \mathbb{C}^*$, $i \in I$ we have*

$$\chi_q(L_{i,a}^-) \in [\mathbf{\Psi}_{i,a}^{-1}](1 + A_{i,a}^{-1}\mathbb{Z}[\![A_{j,b}^{-1}]\!]_{j\in I, b\in\mathbb{C}^*}).$$

Example 5.4. In the case $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, we have

$$\chi_q(L_{1,a}^+) = [(1 - za)] \sum_{r\geq 0} [-2r\omega_1],$$

$$\chi_q(L_{1,a}^-) = \left[\frac{1}{(1 - za)}\right] \sum_{r\geq 0} A_{1,a}^{-1} A_{1,aq^{-2}}^{-1} \cdots A_{1,aq^{-2(r-1)}}^{-1}.$$

Thus, although positive and negative prefundamental representations have the same character,

$$\chi(L_{1,a}^+) = \chi(L_{1,a}^-) = \sum_{r\geq 0} [-2r\omega_1],$$

their q-characters are very different: the normalized q-character $\widetilde{\chi}_q(L_{1,a}^+)$ is independent of the spectral parameter a, whereas $\widetilde{\chi}_q(L_{1,a}^-)$ does depend on a.

5B. *Baxter relations.* We can now state the *generalized Baxter relations*:

Theorem 5.5 [Frenkel and Hernandez 2015, Theorem 4.8, Remark 4.10].[2] *Let V be a finite-dimensional representation of $U_q(\mathfrak{g})$. Replace in $\chi_q(V)$ each variable $Y_{i,a}$ by $[\omega_i][L^+_{i,aq_i^{-1}}]/[L^+_{i,aq_i}]$ (resp. $[-\omega_i][L^-_{i,aq_i^{-1}}]/[L^-_{i,aq_i}]$) and $\chi_q(V)$ by $[V]$ (resp. $[V^*]$). Then multiplying by a common denominator we get a relation in the Grothendieck ring $K_0(\mathcal{O})$.*

Example 5.6. Taking $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ and $V = L(Y_{1,a})$, we obtain relation (1-2), namely,

$$[L(Y_{1,a})][L^+_{1,aq}] = [L^+_{1,aq^{-1}}][\omega_1] + [L^+_{1,aq^3}][-\omega_1],$$
$$[L(Y_{1,aq^2})][L^-_{1,aq}] = [L^-_{1,aq^{-1}}][-\omega_1] + [L^-_{1,aq^3}][\omega_1].$$

Remark 5.7. By our main result, Theorem 4.2, (and its dual version; see Remark 5.18), these generalized Baxter relations are interpreted as relations in the cluster algebra $\mathcal{A}(G)$. Moreover, when $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, the original Baxter relation (1-2) gets interpreted as a Fomin–Zelevinsky mutation relation.

The right-hand side of a generalized Baxter relation is an $\mathcal{E}$-linear combination of classes of tensor products of prefundamental representations. As shown in [Frenkel and Hernandez 2015], this can be seen as a tensor product decomposition in $K(\mathcal{O})$. Indeed, we have:

Theorem 5.8 [Frenkel and Hernandez 2015]. *Any tensor product of positive (resp. negative) prefundamental representations $L^+_{i,a}$ (resp. $L^-_{i,a}$) is simple.*

5C. *Prefundamental characters of representations and duality.* Let $K_0^\pm$ be the $\mathcal{E}$-subalgebra of $K_0(\mathcal{O}^\pm)$ generated by the $[V_{i,a}]$, $[L^\pm_{i,a}]$ ($i \in I$, $a \in \mathbb{C}^*$).

It follows from Theorem 5.5 that the fraction field of $\mathcal{E}[\ell^+_{i,a}]_{i\in I, a\in\mathbb{C}^*}$ contains $K_0^\pm$. More precisely, each element of $K_0^\pm$ is a Laurent polynomial in $\mathcal{E}[(\ell^+_{i,a})^{\pm 1}]_{i\in I, a\in\mathbb{C}^*}$. Note that the $\ell^+_{i,a}$ are algebraically independent (this is for example a consequence of Theorem 5.8, which implies that the monomials in the $\ell^+_{i,a}$ are linearly independent in $K_0(\mathcal{O}^+)$). In particular the expansion in $\mathcal{E}[(\ell^+_{i,a})^{\pm 1}]_{i\in I, a\in\mathbb{C}^*}$ is unique and we can define the injective ring morphism

$$\chi^+ : K_0^+ \to \mathcal{E}[(\ell^+_{i,a})^{\pm 1}]_{i\in I, a\in\mathbb{C}^*},$$

which is called the positive prefundamental character morphism.

In the same way, by using the relations in Theorem 5.5 in terms of the prefundamental representations $[L^-_{i,a}]$ and by setting $\ell^-_{i,a} = [L^-_{i,a^{-1}}]$, we get the negative

[2] The result in [Frenkel and Hernandez 2015] is stated in terms of the $L^+_{i,a}$, and the $R^+_{i,a}$ such that $(R^+_{i,a})^* \simeq L^-_{i,a}$. In the last case, the variable $Y_{i,a}$ has to be replaced by $[\omega_i][R_{i,aq_i^{-1}}]^+/[R^+_{i,aq_i}]$ to get $[V]$. That is why in the statement of Theorem 5.5 in terms of the $L^-_{i,a}$ the representations $[-\omega_i] \simeq [\omega_i]^*$ and V^* appear.

prefundamental character morphism, which is an injective ring morphism

$$\chi^- : K_0^- \to \mathcal{E}[(\ell^-_{i,a})^{\pm 1}]_{i\in I,\, a\in\mathbb{C}^*}.$$

Example 5.9. We can reformulate Theorem 5.5 as follows. For a finite-dimensional representation V of $U_q(\mathfrak{g})$, the positive (resp. negative) prefundamental character $\chi^+(V)$ (resp. $\chi^-(V^*)$) is obtained from the q-character $\chi_q(V)$ by replacing each variable $Y_{i,a}$ by $[\omega_i]\ell^+_{i,aq_i^{-1}}(\ell^+_{i,aq_i})^{-1}$ (resp. $[-\omega_i]\ell^-_{i,a^{-1}q_i}(\ell^-_{i,a^{-1}q_i^{-1}})^{-1}$). In fact, this can also be seen as a change of variables analogous to that of [Hernandez and Leclerc 2016, Section 5.2.2]. For example, for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, using Example 5.6 we have

$$\begin{aligned}\chi^+([L(Y_{1,a})]) &= \frac{[\omega_1]\ell^+_{1,aq^{-1}} + [-\omega_1]\ell^+_{1,aq^3}}{\ell^+_{1,aq}},\\ \chi^-([L(Y_{1,a})]) &= \frac{[-\omega_1]\ell^-_{1,a^{-1}q^3} + [\omega_1]\ell^-_{1,a^{-1}q^{-1}}}{\ell^-_{i,a^{-1}q}}.\end{aligned} \tag{5-11}$$

Example 5.10. Set $\mathfrak{g} = \widehat{\mathfrak{sl}}_3$. Let

$$\mathbf{\Psi} = ((1-zq^{-2})/(1-z),\, (1-qz)) = [-\omega_1]Y_{1,q^{-1}}\mathbf{\Psi}_{2,q}.$$

Let us compute $\chi^+(L(\mathbf{\Psi}))$. For $k \geq 0$, let $W_k = \widetilde{L}(M_k)$, where

$$M_k = (Y_{2,q^{2k}} \cdots Y_{2,q^4} Y_{2,q^2}) Y_{1,q^{-1}}.$$

We can prove as in [Hernandez and Jimbo 2012, Section 7.2] that $(W_k)_{k\geq 0}$ gives rise to a limiting $U_q(\mathfrak{b})$-module W_∞ whose q-character is

$$\chi_q(W_\infty) = \mathbf{\Psi}(1+A^{-1}_{1,1})\chi(L^+_{2,q}) = \chi_q(L^+_{2,q})[-\omega_1]Y_{1,q^{-1}} + [-\omega_1+\omega_2]Y^{-1}_{1,q}\chi_q(L^+_{2,q^{-1}}).$$

But it follows from Theorem 7.6(ii) that $L(\mathbf{\Psi})$ and W_∞ have the same character, $\lim_{k\to+\infty}\chi(\widetilde{L}(M_k)) = \lim_{k\to+\infty}\chi(\widetilde{L}(\overline{M_k}^{-1}))$. As $L(\mathbf{\Psi})$ is a subquotient of W_∞, they are isomorphic. Consequently

$$\chi^+(L(\mathbf{\Psi})) = \frac{\ell^+_{2,q}\ell^+_{1,q^{-2}} + [-\alpha_1]\ell^+_{1,q^2}\ell^+_{2,q^{-1}}}{\ell^+_{1,1}}.$$

Proposition 5.11. *A representation V in $\mathcal{O}^\pm$ satisfying $[V] \in K_0^\pm$ has finite length.*

Remark 5.12. An object in the category $\mathcal{O}$ does not necessarily have finite length. The subcategory of objects of finite length is not stable by tensor product [Boos et al. 2009, Lemma C.1].

Proof. Suppose $[V] \in K_0^+$. Then there is a monomial M in the $\ell^+_{i,a}$ such that $M\chi^+(V)$ is a polynomial in the $\ell^+_{i,a}$. There is a tensor product L of positive prefundamental representations such that $\chi^+(L) = M$. Then $L \otimes V$ has finite length, hence the result. □

Proposition 5.13. *The identification of the variables $\ell^+_{i,a}$ and $\ell^-_{i,a}$ induces a unique isomorphism of $\mathcal{E}$-algebras*

$$D : K_0^+ \to K_0^- .$$

Proof. The identification gives a well-defined injective ring morphism

$$D' : K_0^+ \to \mathcal{E}[(\ell^-_{i,a})^{\pm 1}]_{i\in I, a\in\mathbb{C}^*}.$$

It suffices to prove that its image is $\chi^-(K_0^-)$. For $V = L^+_{i,b}$ a prefundamental, its image by D' is $\ell^-_{i,b} = \chi^-([L^-_{i,b^{-1}}])$. For $V = L(Y_{i,b})$ a fundamental, its image by D' is obtained from its q-character $\chi_q(V)$ by replacing each variable $Y_{i,a}$ by $[\omega_i][L^-_{i,aq_i^{-1}}]/[L^-_{i,aq_i}]$, that is, by $[\omega_i]\ell^-_{i,a^{-1}q_i}/\ell^-_{i,a^{-1}q_i^{-1}}$. In the construction of χ^-, this corresponds to $Y^{-1}_{i,a^{-1}}$ (see the formula in Theorem 5.5). By [Hernandez 2007, Lemma 4.10], there is a finite-dimensional representation V' whose q-character is obtained from $\chi_q(V)$ by replacing each $Y_{i,a}$ by $Y^{-1}_{i,a^{-1}}$. Hence $D'(V) = \chi^-((V')^*) \in \chi^-(K_0^-)$. We have proved $\mathrm{Im}(D') \subset \chi^-(K_0(\mathcal{O}^-_{f,\mathbb{Z}}))$. Similarly, for W such that $V = W^*$, we have $D'(W') = \chi^-(V)$ and we get the other inclusion. □

Example 5.14. (i) For any $(i, a) \in I \times \mathbb{C}^*$, we have

$$D([L^+_{i,a}]) = [L^-_{i,a^{-1}}].$$

(ii) For any dominant monomial m, we have

$$D([L(m)]) = [L(m_1)],$$

where m_1 is obtained from m by replacing each $Y_{i,a}$ by $Y_{i,a^{-1}}$. Indeed $L(m_1)$ is isomorphic to $((L(m))')^*$, whose highest weight monomial is the inverse of the lowest weight monomial $(m_1)^{-1}$ of $L(m)'$. For example, for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$, we have $D([L(Y_{1,a})] = [L(Y_{1,a^{-1}})]$, as this can be observed by comparing the two formulas in (5-11).

Remark 5.15. The duality D is compatible with characters by (i) in Theorem 5.3. However it is not compatible with q-characters (for example, negative and positive prefundamental representations have very different q-characters, as explained above).

Proposition 5.16. *An element in the Grothendieck group $K_0(\mathcal{O}^\pm)$ is a* (*possibly countable*) *sum of elements in $K_0^\pm$.*

Proof. Let us prove it for $K_0(\mathcal{O}^+)$ (the proof is analogous for $K_0(\mathcal{O}^-)$). By definition, the positive ℓ-weights label the simple modules in $\mathcal{O}^+$. Moreover, an ℓ-weight is positive if and only if it is a product of highest ℓ-weights of representations $L^+_{i,a}$, $V_i(a)$ and $[\omega]$. This implies that, for each positive ℓ-weight $\boldsymbol{\Psi}$, we can choose (and fix) a monomial in the $[L^+_{i,a}]$, $[V_i(a)]$, $[\omega]$ such that the corresponding representation has highest ℓ-weight equal to $\boldsymbol{\Psi}$. Hence the positive ℓ-weights also label the linearly independent family of these monomials in K_0^+. Expanding these monomials we get finite sums of classes of simple modules by Proposition 5.11. We get an (infinite) transition matrix from the classes of simple objects in $\mathcal{O}^+$ to such products in K_0^+, and this matrix is unitriangular (for the standard partial ordering with respect to weight). Hence the result. □

This implies immediately the following.

Theorem 5.17. *$\mathcal{O}^+$ and $\mathcal{O}^-$ are monoidal and the morphism D extends uniquely to an isomorphism of $\mathcal{E}$-algebras*

$$D : K_0(\mathcal{O}^+) \to K_0(\mathcal{O}^-).$$

Remark 5.18. Consequently, our main result in Theorem 4.2 may also be written in terms of the subcategory $\mathcal{O}^-_{2\mathbb{Z}}$ of $\mathcal{O}^-_{\mathbb{Z}}$ whose Grothendieck ring is $D(K_0(\mathcal{O}^+_{2\mathbb{Z}}))$.

Proof. For L, L' simple in $\mathcal{O}^+$, we may consider a decomposition of $[L]$, $[L']$ as a countable sum of elements in K_0^+ as in Proposition 5.16. Then $[L][L']$ is also such a countable sum and is in $K_0(\mathcal{O}^+)$. Hence $\mathcal{O}^+$ is monoidal. This is analogous for $\mathcal{O}^-$.

The isomorphism of Proposition 5.13 is extended by linearity to $K_0(\mathcal{O}^+)$ by using Proposition 5.16. This map $D : K_0(\mathcal{O}^+) \to K_0(\mathcal{O}^-)$ is an injective ring morphism. The ring morphism $D^{-1} : K_0(\mathcal{O}^-) \to K_0(\mathcal{O}^+)$ is constructed in the same way and so D is a ring isomorphism. □

Proposition 5.19. *A simple object in $\mathcal{O}^+$ (resp. in $\mathcal{O}^-$) is a subquotient of a tensor product of two simple representations $V \otimes L$, where V is finite-dimensional and L is a tensor product of positive (resp. negative) prefundamental representations.*

Proof. Let $L(\mathbf{\Psi})$ be simple in $\mathcal{O}^\pm$. By definition, its highest ℓ-weight is a product of highest ℓ-weights of representations $[\omega]$, $L^\pm_{i,a}$, $V_i(a)$, where $\omega \in P_\mathbb{Q}$, $i \in I$, $a \in \mathbb{C}^*$. So $\mathbf{\Psi}$ can be factorized as

$$\mathbf{\Psi} = [\omega] \times m \times \prod_{i \in I, a \in \mathbb{C}^*} \mathbf{\Psi}_{i,a}^{u_{i,a}}, \tag{5-12}$$

where $\omega \in P_\mathbb{Q}$, $\pm u_{i,a} \geq 0$ and $m \in \mathcal{M}$ is a dominant monomial. The result follows by taking $V = [\omega] \otimes L(m)$ and $L = \bigotimes_{i \in I, a \in \mathbb{C}^*} (L^\pm_{i,a})^{\otimes |u_{i,a}|}$, which is simple by Theorem 5.8. □

Proposition 5.20. *The normalized q-character of a simple object in $\mathcal{O}^-$ belongs to the ring $\mathbb{Z}[\![A^{-1}_{i,a}]\!]_{i \in I, a \in \mathbb{C}^*}$.*

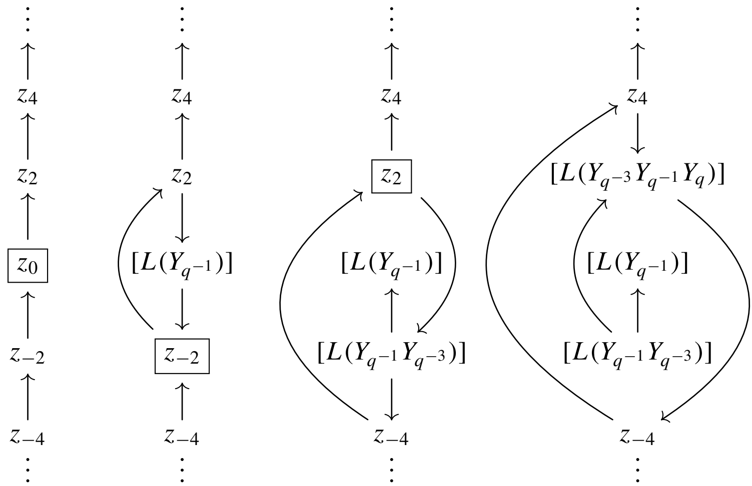
Proof. The result is known for the category $\mathcal{C}$ by [Frenkel and Mukhin 2001]. For negative prefundamental representations, the result is known by [Hernandez and Jimbo 2012, Theorem 6.1]. The general result follows from Proposition 5.19. □

Note that this property is not satisfied in $\mathcal{O}^+$; see Example 5.4.

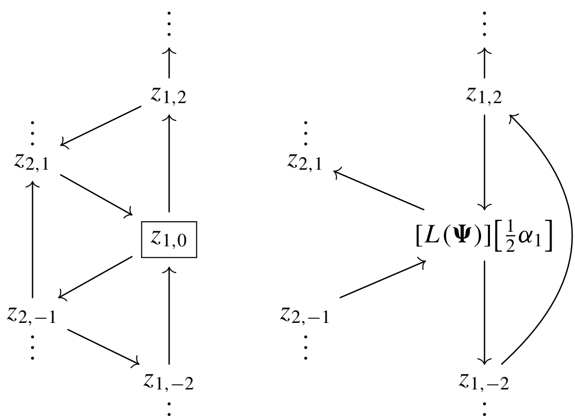
6. Proof of the main theorem

6A. *Examples of mutations.*

6A1. Let $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$. We display a sequence of 3 mutations starting from the initial seed of $\mathcal{A}(G)$. The mutated cluster variables are indicated by a framebox.



6A2. Let $\mathfrak{g} = \widehat{\mathfrak{sl}}_3$. Example 5.10 can be reformulated as a one-step mutation from the initial seed, as follows:



Recall that here $\mathbf{\Psi} = [-\omega_1]Y_{1,q^{-1}}\mathbf{\Psi}_{2,q}$, as in Example 5.10.

6A3. For an arbitrary $\mathfrak{g}$, let us calculate the first mutation relation for each cluster variable $z_{i,r}$ of the initial seed, generalizing Section 6A2. We denote by $z_{i,r}^*$ the new cluster variable obtained by mutating $z_{i,r}$. Then we have

$$z_{i,r}^* z_{i,r} = \prod_{j:C_{j,i}\neq 0} z_{j,r-d_jC_{j,i}} + \prod_{j:C_{j,i}\neq 0} z_{j,r+d_jC_{j,i}}.$$

We claim that $z^*_{i,r} = [\lambda][L(\mathbf{\Psi})]$, where

$$\mathbf{\Psi} = [-\omega_i] Y_{i,q^{r-d_i}} \prod_{j:C_{j,i}<0} \mathbf{\Psi}_{j,q^{r-d_j C_{j,i}}} \quad \text{and} \quad \lambda = \frac{\alpha_i}{2} - r \sum_{j:C_{j,i}<0} \frac{\omega_j}{2d_j}.$$

As in Section 6A2, this is derived from the explicit q-character formula

$$\chi_q(L(\mathbf{\Psi})) = [\mathbf{\Psi}](1 + A^{-1}_{i,q^r}) \prod_{j:C_{j,i}<0} \chi_j, \tag{6-13}$$

where $\chi_j = \chi(L^+_{j,a})$ does not depend on a. (By considering $L(\mathbf{\Psi}) \otimes L(Y_{i,q^{r-d_i}} \mathbf{\Psi}^{-1})$ and $L(\mathbf{\Psi}) \otimes L^+_{i,q^r}$ we prove that the multiplicities in $\chi_q(L(\mathbf{\Psi}))$ are larger than in the right-hand side of (6-13). The reverse inequality is established by considering $L(\overline{M_R}^{-1}) \otimes L(\mathbf{\Psi} \overline{M_R})$, where the monomials M_R are defined for $\overline{\mathbf{\Psi}}^{-1}$ as in Theorem 7.6.) The mutation relation thus becomes the following relation in the Grothendieck ring $K_0(\mathcal{O}^+)$:

$$[L(\mathbf{\Psi}) \otimes L^+_{i,q^r}] = \Big[\bigotimes_{j:C_{j,i}\neq 0} L^+_{j,q^{r-d_j C_{j,i}}} \Big] + [-\alpha_i] \Big[\bigotimes_{j:C_{j,i}\neq 0} L^+_{j,q^{r+d_j C_{j,i}}} \Big]. \tag{6-14}$$

By Theorem 5.8, the two terms on the right-hand side are simple. Hence this is the decomposition of the class of the tensor product into simple modules.

6B. ***Proof of Theorem 4.2.*** We identify the $\mathcal{E}$-algebras

$$\mathcal{E}[z^{\pm 1}_{i,r}]_{(i,r)\in V} \quad \text{and} \quad \mathcal{E}[(\ell^+_{i,q^r})^{\pm 1}]_{(i,r)\in V}$$

as in Remark 4.3.

Proposition 6.1. *We have* $K_0^+ \cap K_0(\mathcal{O}^+_{2\mathbb{Z}}) \subseteq \mathcal{A}(G) \otimes_{\mathbb{Z}} \mathcal{E}$.

Note that by Proposition 5.16 this implies $K_0(\mathcal{O}^+_{2\mathbb{Z}}) \subseteq \mathcal{A}(G) \mathbin{\hat{\otimes}_{\mathbb{Z}}} \mathcal{E}$.

Proof. Clearly, $\ell^+_{i,q^s} = z_{i,s}[(s/2d_i)\omega_i]$ belongs to $\mathcal{A}(G)_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathcal{E}$ for any $(i,s) \in V$. By Proposition 5.16, it remains to show that $[V_{i,q^r}]$ belongs to $\mathcal{A}(G)_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathcal{E}$ for any $(i,r) \in W$.

Remember from Section 4A that we denote by G the same quiver as Γ with vertices labeled by W instead of V. In the next discussion, we divide the vertices of G and Γ into *columns*, as in [Hernandez and Leclerc 2016, Example 2.3], and we denote by k the number of columns. As in [Hernandez and Leclerc 2016, Section 2.1.3], consider the full subquiver G^- of G whose vertex set is $W^- = \{(i,r) \in W \mid r \leq 0\}$. The definition of G shows that there is only one vertex of $G \setminus G^-$ in each column which is connected to G^- by some arrow. Let H^- denote the ice quiver obtained from G^- by adding these k vertices together with their connecting arrows, and by declaring the new vertices frozen.

Consider the cluster algebras $\mathcal{A}(H^-)$ (with k frozen variables) and $\mathcal{A}(G^-)$ (with no frozen variable). It follows from the definitions that $\mathcal{A}(H^-)$ can be regarded as a subalgebra of $\mathcal{A}(G)$, and $\mathcal{A}(G^-)$ is the coefficient-free counterpart of $\mathcal{A}(H^-)$, studied in [Hernandez and Leclerc 2016]. Let f_j $(1 \le j \le k)$ be the frozen variable of $\mathcal{A}(H^-)$ sitting in column j. The cluster of the initial seed of $\mathcal{A}(H^-)$ thus consists of the frozen variables f_j $(1 \le j \le k)$ and the ordinary cluster variables $z_{i,s}$ $((i,s) \in V^-)$, where $V^- = \{(i,s) \in V \mid s + d_i \le 0\}$. Let us denote by $u_{i,s}$ $((i,s) \in V^-)$ the cluster variables of the initial seed of $\mathcal{A}(G^-)$. We can use a similar change of variables as in [Hernandez and Leclerc 2016, Section 2.2.2]:

$$y_{i,r} = u_{i,r-d_i} \quad \text{if } r + d_i > 0, \qquad y_{i,r} = \frac{u_{i,r-d_i}}{u_{i,r+d_i}} \quad \text{otherwise.}$$

Let

$$F : \mathbb{Z}[u_{i,s}^{\pm 1} \mid (i,s) \in V^-] \to \mathbb{Z}[f_j^{\pm 1}, z_{i,s}^{\pm 1} \mid 1 \le j \le k,\ (i,s) \in V^-]$$

be the ring homomorphism defined by

$$F(u_{i,s}) = \frac{z_{i,s}}{f_j} \quad \text{if } (i,s) \text{ sits in column } j.$$

Thus

$$F(y_{i,r}) = \begin{cases} \dfrac{z_{i,r-d_i}}{f_j} & \text{if } r + d_i > 0 \text{ and } (i,r) \text{ sits in column } j, \\ \dfrac{z_{i,r-d_i}}{z_{i,r+d_i}} & \text{otherwise.} \end{cases}$$

We introduce a $\mathbb{Z}^k$-grading on $\mathbb{Z}[f_j^{\pm 1}, z_{i,s}^{\pm 1} \mid 1 \le j \le k,\ (i,s) \in V^-]$ by declaring that

$$\deg(f_j) = e_j, \quad \deg(z_{i,s}) = e_j \quad \text{if } (i,s) \text{ sits in column } j,$$

where e_j $(1 \le j \le k)$ denotes the canonical basis of $\mathbb{Z}^k$.

Let x be the cluster variable of $\mathcal{A}(G^-)$ obtained from the initial seed $(\{u_{i,s}\}, G^-)$ via a sequence of mutations σ, and let y be the cluster variable of $\mathcal{A}(H^-)$ obtained from the initial seed $(\{z_{i,s}, f_j\}, H^-)$ via the *same* sequence of mutations σ. We want to compare the Laurent polynomials y and $F(x)$. Since $\deg(F(u_{i,s})) = (0, \dots, 0)$ for every (i,s), we see that $F(x)$ is multihomogeneous of degree $(0, \dots, 0)$ for the above grading. On the other hand, it is easy to check that for every nonfrozen vertex (i,s) of the ice quiver H^- the sum of the multidegrees of the initial cluster variables and frozen variables sitting at the targets of the arrows going out of (i,s) is equal to the sum of the multidegrees of the initial cluster variables and frozen variables sitting at the sources of the arrows going into (i,s). Therefore, $\mathcal{A}(H^-)$ is a multigraded cluster algebra, in the sense of [Grabowski and Launois 2014]. It follows that y is also multihomogeneous, of degree $(a_1, \dots, a_k)$. Now, by construction, we have

$$F(x)|_{f_1=1,\dots,f_k=1} = y|_{f_1=1,\dots,f_k=1}.$$

Therefore,

$$y = F(x) \prod_{j=1}^{k} f_j^{a_j}.$$

Taking a cluster expansion with respect to the initial cluster of $\mathcal{A}(H^-)$, we write $y = N/D$ where D is a monomial in the nonfrozen cluster variables and N is a polynomial in the nonfrozen and frozen variables. Moreover N is not divisible by any of the f_j. It follows that $\prod_{j=1}^{k} f_j^{a_j}$ is the smallest monomial such that $F(x)\prod_{j=1}^{k} f_j^{a_j}$ contains only nonnegative powers of the variables f_j.

Now we can conclude using [Hernandez and Leclerc 2016, Theorem 3.1], which implies that, for all $(i,r) \in W^-$ with $r \ll 0$, the q-character of V_{i,q^r} (expressed in terms of the variables $y_{i,s} \equiv Y_{i,q^s}$) is a cluster variable x of $\mathcal{A}(G^-)$. By [Frenkel and Mukhin 2001, Corollary 6.14], for $r \ll 0$ this cluster variable does not contain any variable $y_{i,s}$ with $s + 2d_i > 0$, hence $F(x)$ does not contain any frozen variable f_j. Therefore $y = F(x)$, and the q-character of V_{i,q^r} (expressed in terms of the variables $z_{i,s}$) is a cluster variable of $\mathcal{A}(H^-)$, that is, a cluster variable of $\mathcal{A}(\Gamma)$. This proves the claim for every fundamental module V_{i,q^r} with $r \ll 0$. But by definition of the cluster algebra $\mathcal{A}(\Gamma)$, the set of cluster variables is invariant under the change of variables $z_{i,s} \mapsto z_{i,s+2d_i}$. Thus we are done. □

Proposition 6.2. *We have* $\mathcal{A}(G) \hat{\otimes}_{\mathbb{Z}} \mathcal{E} \subseteq K_0(\mathcal{O}^+_{2\mathbb{Z}})$.

To prove this proposition, we need to establish some preliminary results. We first prove in Lemma 6.4 that, for $\chi \in \mathcal{A}(G)$, at least one negative ℓ-weight occurs in $\chi_q(D(\chi))$. Then in Lemma 6.5 we construct a family of distinguished elements

$$F(\mathbf{\Psi}) \in \chi_q(D(K_0(\mathcal{O}^+))).$$

The first result allows us to write each element of $\chi_q(D(\mathcal{A}(G)))$ as a linear combination of the $F(\mathbf{\Psi})$. This implies the inclusion in Proposition 6.2.

So consider an element χ in $\mathcal{A}(G)$. By the Laurent phenomenon [Fomin and Zelevinsky 2002], χ is a Laurent polynomial in the initial cluster variables:

$$\chi = P(\{z_{i,r}\}_{(i,r)\in V}).$$

Hence $\mathcal{A}(G)$ is a subalgebra of the fraction field of $K_0(\mathcal{O}^+_{2\mathbb{Z}})$, and the duality D of Proposition 5.13 can be algebraically extended to $\mathcal{A}(G)$. In particular we have

$$D(\chi) = P(\{D(z_{i,r})\}_{(i,r)\in V}) \in \mathrm{Frac}(K_0(\mathcal{O}^-_{2\mathbb{Z}}))$$

in the fraction field of $K_0(\mathcal{O}^-_{2\mathbb{Z}})$. The q-character morphism can also be algebraically extended to $\mathrm{Frac}(K_0(\mathcal{O}^-_{2\mathbb{Z}}))$. Then $\chi_q(D(\chi))$ is obtained by replacing each $z_{i,r}$ by the corresponding q-character

$$\chi_q(D(z_{i,r})) = \left[\binom{-r}{2d_i}\omega_i\right]\chi_q(L^-_{i,q^{-r}}) = \left[\binom{-r}{2d_i}\omega_i\right]\boldsymbol{\Psi}^{-1}_{i,q^{-r}}(1+\mathcal{A}_{i,r}), \tag{6-15}$$

where, by Theorem 7.6, $\mathcal{A}_{i,r}$ is a formal power series in the $A^{-1}_{j,a}$ without constant term. In particular, we have an analogous formula for the inverse,

$$(\chi_q(D(z_{i,r})))^{-1} = \left[\binom{r}{2d_i}\omega_i\right]\boldsymbol{\Psi}_{i,q^{-r}}(1+\mathcal{B}_{i,r}),$$

where

$$\mathcal{B}_{i,r} = \sum_{k\geq 1}(-\mathcal{A}_{i,r})^k$$

is a formal power series in the $A^{-1}_{j,a}$ without constant term. In particular $\chi_q(D(\chi))$ is in $\mathcal{E}_\ell$ and we get a sum of the form

$$\chi_q(D(\chi)) = \sum_{1\leq\alpha\leq R}\lambda_\alpha[\omega_\alpha]m_\alpha(1+\mathcal{A}_\alpha) \in \mathcal{E}_\ell, \tag{6-16}$$

where ω_α is a weight, m_α is a Laurent monomial in the $\boldsymbol{\Psi}_{i,q^{-r}}$, $\mathcal{A}_\alpha$ is a formal power series in the $A^{-1}_{j,a}$ without constant term and $\lambda_\alpha \in \mathbb{Z}$.

We recall the notion of negative ℓ-weight introduced in Definition 3.8.

Remark 6.3. We say that a sequence $(\boldsymbol{\Psi}^{(m)})_{m\geq 0}$ of ℓ-weights converges pointwise as a rational fraction to an ℓ-weight $\boldsymbol{\Psi}$ if, for every $i \in I$ and $z \in \mathbb{C}$, the ratio $\boldsymbol{\Psi}_i^{(m)}(z)/\boldsymbol{\Psi}_i(z)$ converges to 1 when $N \to +\infty$ and $|q| > 1$. For example, defining the monomials $M_{i,r,N}$ as in equation (6-17) below, the sequence $(\widetilde{M_{i,r,N}})_{N\geq 0}$ converges pointwise as a rational fraction to $\boldsymbol{\Psi}^{-1}_{i,q^{-r}}$.

Lemma 6.4. *Let $\chi \in \mathcal{A}(G)$ be nonzero. Then at least one negative ℓ-weight occurs in $\chi_q(D(\chi))$.*

Proof. We will use the following partial ordering $\preceq$ on the set of ℓ-weights $\boldsymbol{\Psi}$ satisfying $\varpi(\boldsymbol{\Psi}) = 1$: for such ℓ-weights $\boldsymbol{\Psi}$, $\boldsymbol{\Psi}'$, we set $\boldsymbol{\Psi} \succeq \boldsymbol{\Psi}'$ if

$$\boldsymbol{\Psi}'(\boldsymbol{\Psi})^{-1} = \prod_{i,r\geq -M}\widetilde{A_{i,q^r}}^{-v_{i,r}}$$

is a possibly infinite product (that is, pointwise the limit of the partial products) with the $v_{i,r} \geq 0$. If $\boldsymbol{\Psi} = \widetilde{m}$ and $\boldsymbol{\Psi}' = \widetilde{m'}$ with m, m' monomials in $\mathcal{M}$, then $\boldsymbol{\Psi} \preceq \boldsymbol{\Psi}'$ is equivalent to $m \preceq m'$ for the partial ordering considered in [Nakajima 2004].

As the sum (6-16) is finite, there is α_0 such that m_{α_0} is maximal for $\preceq$. We prove that m_{α_0} is a negative ℓ-weight.

Let N be such that all the cluster variables $z_{i,r}$ of the initial seed occurring in the Laurent monomials of equation (6-16) satisfy $r > -2d(N+2)$, where $d = \mathrm{Max}_{i\in I}(d_i)$ is the lacing number of $\dot{\mathfrak{g}}$. We consider as above the semi-infinite cluster algebra $\mathcal{A}(H_N^+)$ obtained from $\mathcal{A}(G)$, where the cluster variables sitting at $(i,r) \in V$, $r \le -2d(N+2)$ have been removed. As explained in the proof of Proposition 6.1, $\mathcal{A}(H_N^+)$ can be regarded as a subalgebra of $\mathcal{A}(G)$. We replace every cluster variable $z_{i,r}$ of the initial seed by the class of the Kirillov–Reshetikhin module $W_{i,r,N}$ of highest monomial

$$M_{i,r,N} = \prod_{\substack{k\ge 0 \\ r+2kd_i \le 2dN}} Y_{i,q^{-r-d_i-2kd_i}}. \tag{6-17}$$

(Here $M_{i,r,N}$ is set to be 1 if $r > 2dN$.) We obtain

$$\phi_N(\chi) \in \mathrm{Frac}(K_0(\mathcal{C})),$$

the image of χ in the fraction field of $K_0(\mathcal{C})$. By using the duality D, we reverse all spectral parameters (by (ii) in Example 5.14 illustrating Proposition 5.13, or by [Hernandez and Leclerc 2010, Section 3.4]). We obtain the same[3] as in [Hernandez and Leclerc 2016, Section 2.2.2]. Then by [Hernandez and Leclerc 2016, Theorem 5.1], $D(\phi_N(\chi))$ belongs to the Grothendieck ring of $\mathcal{C}$. So by applying D again, $\phi_N(\chi)$ is in $K_0(\mathcal{C})$.

Now we get, as for equation (6-16),

$$\chi_q(\phi_N(\chi)) = \sum_{1\le\alpha\le R} \lambda_\alpha m_\alpha^{(N)}(1+\mathcal{A}_\alpha^{(N)}), \tag{6-18}$$

where $m_\alpha^{(N)}$ is a monomial and $\mathcal{A}_\alpha^{(N)}$ is a formal power series in the $A_{i,a}^{-1}$. Note that, when $N \to +\infty$, $\mathcal{A}_\alpha^{(N)}$ converges to $\mathcal{A}_\alpha$ as a formal power series in the $A_{i,a}^{-1}$. Let $\omega_{i,r,N}$ be the highest weight of $W_{i,r,N}$. Now if the initial cluster variable $z_{i,r}$ is replaced by

$$\left[\left(\frac{r}{2d_i}\right)\omega_i - \omega_{i,r,N}\right][W_{i,r,N}] \tag{6-19}$$

instead of $[W_{i,r,N}]$, we just have to replace in (6-18) each $m_\alpha^{(N)}$ by $[\omega_\alpha](m_\alpha^{(N)})^\sim$ (the $\mathcal{A}_\alpha^{(N)}$ are unchanged). Then $(m_\alpha^{(N)})^\sim$ converges to m_α when $N \to +\infty$ pointwise as a rational fraction.

Let us show that there are infinitely many N such that $(m_{\alpha_0}^{(N)})^\sim$ is maximal among the $(m_\alpha^{(N)})^\sim$ for $\preceq$. Otherwise, since equation (6-16) has finitely many summands, there is α such that $(m_{\alpha_0}^{(N)})^\sim \prec (m_\alpha^{(N)})^\sim$ for infinitely many N. In the limit, we get that $m_{\alpha_0} \prec m_\alpha$, contradiction.

For N such that $(m_{\alpha_0}^{(N)})^\sim$ is maximal for $\preceq$, $m_{\alpha_0}^{(N)}$ is necessarily dominant, as $\phi_N(\chi) \in K_0(\mathcal{C})$ (see [Frenkel and Mukhin 2001, Section 5.4]). Then the limit m_{α_0} of the $(m_{\alpha_0}^{(N)})^\sim$ is negative, as it is easy to check that a limit of dominant monomials

[3] We have a term $-d_i$ which does not occur in [Hernandez and Leclerc 2016, Section 2.2.2], as we use the labeling by V and not by W.

is a negative ℓ-weight. Finally, since m_{α_0} is maximal for $\preceq$, it necessarily occurs with a nonzero coefficient in the expansion of $\chi_q(D(\chi))$. □

Lemma 6.5. *Let $\mathbf{\Psi}$ be a negative ℓ-weight such that the roots and the poles of $\mathbf{\Psi}_i(z)$ are of the form q^r with $(i,r) \in V$. Then there is a unique $F(\mathbf{\Psi}) \in \chi_q(D(K_0(\mathcal{O}^+_{2\mathbb{Z}})))$ such that $\mathbf{\Psi}$ is the unique negative ℓ-weight occurring in $F(\mathbf{\Psi})$ and its coefficient is* 1.

Moreover $F(\mathbf{\Psi})$ is of the form

$$F(\mathbf{\Psi}) = [\mathbf{\Psi}] + \sum_{\mathbf{\Psi}' : \varpi(\mathbf{\Psi}') \prec \varpi(\mathbf{\Psi})} \lambda_{\mathbf{\Psi}'} [\mathbf{\Psi}'], \tag{6-20}$$

for the usual partial ordering on weights and with the $\lambda_{\mathbf{\Psi}'} \in \mathbb{Z}$.

Proof. The uniqueness follows from Proposition 6.1 and Lemma 6.4. For each negative ℓ-weight $\mathbf{\Psi}$ as in the lemma, there is a representation $M(\mathbf{\Psi})$ in $\mathcal{O}^+_{2\mathbb{Z}}$ such that $\chi_q(D([M(\mathbf{\Psi})]))$ is of the form

$$\chi_q(D([M(\mathbf{\Psi})])) = [\mathbf{\Psi}] + \sum_{\mathbf{\Psi}' : \varpi(\mathbf{\Psi}') \prec \varpi(\mathbf{\Psi})} \mu_{\mathbf{\Psi}', \mathbf{\Psi}} [\mathbf{\Psi}'].$$

Indeed it suffices to consider a tensor product of fundamental and positive prefundamental representations. Now if the $F(\mathbf{\Psi})$ do exist, we have an infinite triangular transition matrix from the $(F(\mathbf{\Psi}))$ to the $(\chi_q(D([M(\mathbf{\Psi})])))$ with 1 on the diagonal and whose off-diagonal coefficients are the $\mu_{\mathbf{\Psi}', \mathbf{\Psi}}$ for $\mathbf{\Psi}$, $\mathbf{\Psi}'$ negative. So to prove the existence, it suffices to consider the inverse of this matrix (which is well-defined, as for given $\mathbf{\Psi}$, $\mathbf{\Psi}'$ there is a finite number of $\mathbf{\Psi}''$ satisfying $\varpi(\mathbf{\Psi}') \preceq \varpi(\mathbf{\Psi}'')$ and $\mu_{\mathbf{\Psi}'', \mathbf{\Psi}} \neq 0$). □

We can now finish the proof of Proposition 6.2:

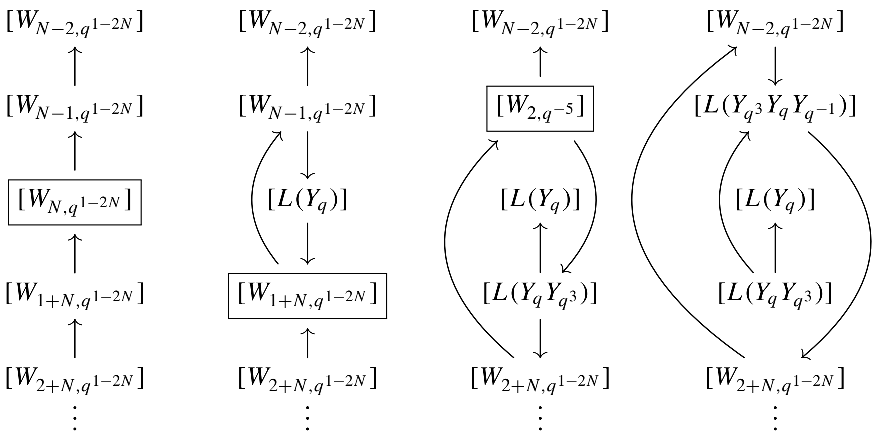
Proof. Let χ be in $\mathcal{A}(G)$. For $\mathbf{\Psi}$ a negative ℓ-weight, we denote by $\lambda_{\mathbf{\Psi}}$ the coefficient of $\mathbf{\Psi}$ in $\chi_q(D(\chi))$. Then by Lemma 6.4 we have

$$\chi_q(D(\chi)) = \sum_{\mathbf{\Psi} \text{ negative}} \lambda_{\mathbf{\Psi}} F(\mathbf{\Psi}).$$

As the $F(\mathbf{\Psi})$ are of the form (6-20), this sum is well-defined in $\chi_q(D(K_0(\mathcal{O}^+_{2\mathbb{Z}})))$ and we get $\chi \in K_0(\mathcal{O}^+_{2\mathbb{Z}})$. □

Example 6.6. Let us illustrate the proof of Lemma 6.4, which is the crucial technical point for the proof of Proposition 6.2. Consider the sequence of mutations of 6A1.

Let us write the cluster variables $\phi_N(\chi)$:



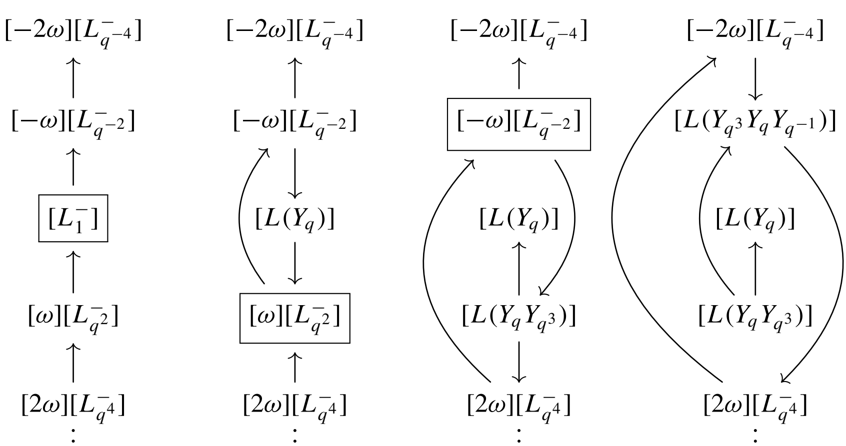
The cluster variable corresponding to $[L(Y_q)]$ has its q-character, which can be written in the form of equation (6-16):

$$\chi_q(L(Y_q)) = Y_q + Y_{q^3}^{-1} = \frac{\chi_q(W_{N-1,q^{1-2N}}) + \chi_q(W_{N+1,q^{1-2N}})}{\chi_q(W_{N,q^{1-2N}})}$$

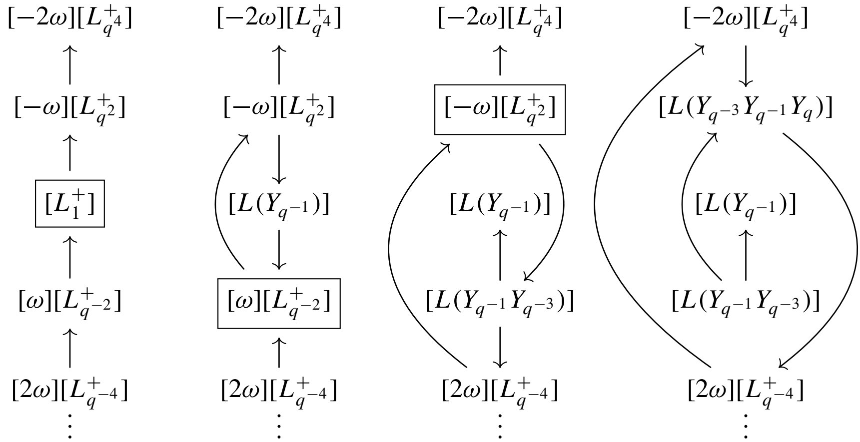
$$= Y_{q^{-1}}^{-1} \frac{1 + A_{q^{-2}}^{-1}\big(1 + A_{q^{-4}}^{-1}(1 + \cdots (1 + A_{q^{2(1-N)}}^{-1}))\cdots\big)}{1 + A_1^{-1}\big(1 + A_{q^{-2}}^{-1}(1 + \cdots (1 + A_{q^{2(1-N)}}^{-1}))\cdots\big)}$$

$$+ Y_q \frac{1 + A_{q^2}^{-1}\big(1 + A_1^{-1}(1 + \cdots (1 + A_{q^{2(1-N)}}^{-1}))\cdots\big)}{1 + A_1^{-1}\big(1 + A_{q^{-2}}^{-1}(1 + \cdots (1 + A_{q^{2(1-N)}}^{-1}))\cdots\big)}.$$

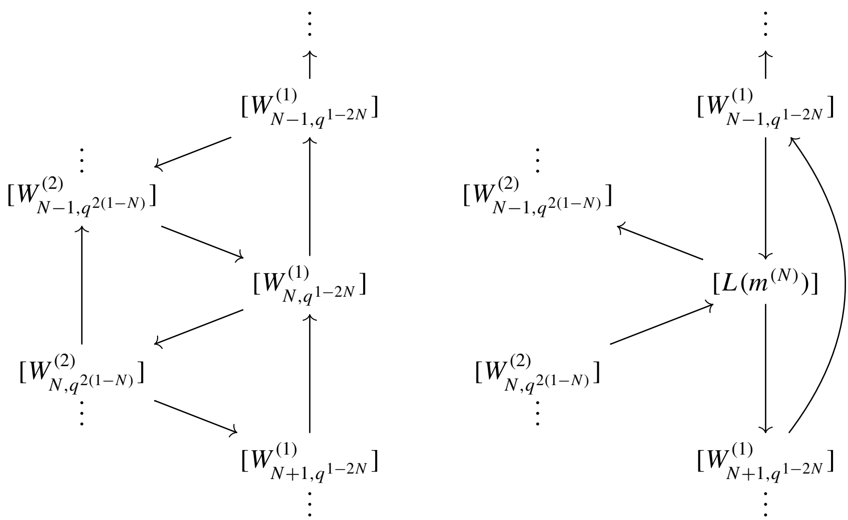
We see as in the statement of Lemma 6.4 that the monomial Y_q maximal for $\preceq$ is dominant and so negative in the sense of Definition 3.8. In the limit $N \to +\infty$ (with the renormalized weights) we get the representations in $K_0(\mathcal{O}^-)$:



The duality again gives the representations in $K_0(\mathcal{O}^+)$:



Example 6.7. We now illustrate the proof of Lemma 6.4 by means of the mutation of Section 6A2. Let us write the cluster variables $\phi_N(\chi)$:

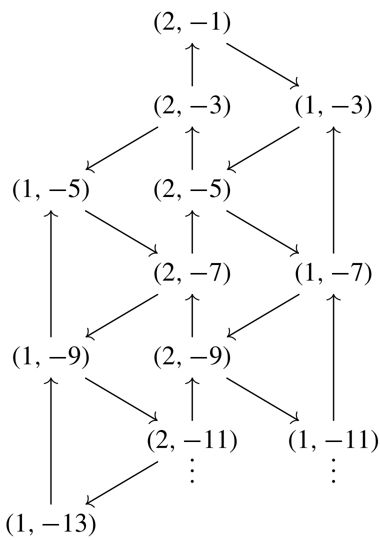


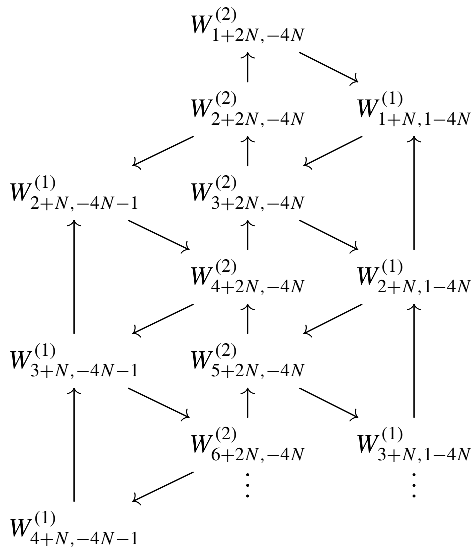
Here $m^{(N)} = Y_{1,q}(Y_{2,q^{2(1-N)}} Y_{2,q^{4-2N}} \cdots Y_{2,q^{-2}})$. The q-character corresponding to the cluster variable $[L(m^{(N)})]$ can be written in the form of equation (6-16):

$$\chi_q(L(m^{(N)})) = \frac{\chi_q(W^{(1)}_{N-1,q^{1-2N}})\chi_q(W^{(2)}_{N,q^{2(1-N)}})}{\chi_q(W^{(1)}_{N,q^{1-2N}})} + \frac{\chi_q(W^{(1)}_{N+1,q^{1-2N}})\chi_q(W^{(2)}_{N-1,q^{2(1-N)}})}{\chi_q(W^{(1)}_{N,q^{1-2N}})}$$
$$= Y^{-1}_{1,q^{-1}}(Y_{2,q^{2(1-N)}} Y_{2,q^{4-2N}} \cdots Y_{2,1})(1+\mathcal{A}^{(N)}_1)$$
$$+ Y_{1,q}(Y_{2,q^{2(1-N)}} Y_{2,q^{4-2N}} \cdots Y_{2,q^{-2}})(1+\mathcal{A}^{(N)}_2).$$

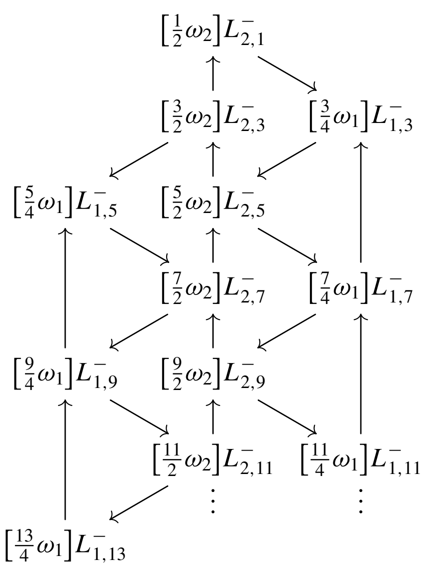
Here $\mathcal{A}_1^{(N)}$ and $\mathcal{A}_2^{(N)}$ are formal power series in the A_{1,q^a}^{-1}, A_{2,q^a}^{-1} ($a \in \mathbb{C}^*$) without constant term. The monomial $(m^{(N)})^{\sim}$ is maximal for $\preceq$. Its limit for $N \to +\infty$ is $[-\omega_1]Y_{1,q}\mathbf{\Psi}_{2,q^{-1}}^{-1}$, which is negative in the sense of Definition 3.8.

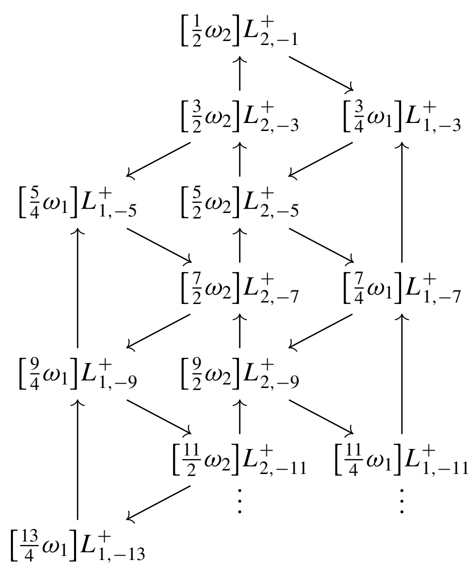
Example 6.8. In this example we check that the images of the initial cluster variables considered in the proof of Lemma 6.4 do match. Let us consider type B_2 with the following initial seed and the initial cluster variables replaced by the $W_{i,r}$:





In the limit $N \to +\infty$ (with the renormalized weights) we get the images of the initial cluster variables:





7. Conjectures and evidence

7A. ***A conjecture.*** The concept of a monoidal categorification of a cluster algebra was introduced in [Hernandez and Leclerc 2016, Definition 2.1]. We say that a simple object S of a monoidal category is *real* if $S \otimes S$ is simple. Let us recall that a cluster monomial is a monomial in the cluster variables of a single cluster.

Definition 7.1. Let $\mathcal{A}$ be a cluster algebra and let $\mathcal{M}$ be an abelian monoidal category. We say that $\mathcal{A}$ is a monoidal categorification of $\mathcal{A}$ if there is an isomorphism between $\mathcal{A}$ and the Grothendieck ring of $\mathcal{M}$ such that the cluster monomials of $\mathcal{A}$ are the classes of all the real simple objects of $\mathcal{M}$ (up to invertibles).

See [Hernandez and Leclerc 2013, Section 2] for a discussion on applications of monoidal categorifications. In view of Theorem 4.2, it is natural to formulate the following conjecture.

Conjecture 7.2. *The isomorphism of Theorem 4.2 defines a monoidal categorification*; *that is*, *the cluster monomials in* $\mathcal{A}(\Gamma)$ *get identified with real simple objects in* $\mathcal{O}^+_{2\mathbb{Z}}$ *up to invertible representations.*

Remark 7.3. By using the duality in Proposition 5.13, the statements of Theorem 4.2 and Conjecture 7.2 can also be formulated in terms of the category $\mathcal{O}^-_{2\mathbb{Z}}$.

Note that Theorem 5.8 implies that all cluster monomials of the initial seed are identified with real simple objects, more precisely, with simple tensor products of positive prefundamental representations, in agreement with Conjecture 7.2. To give other evidence supporting Conjecture 7.2, we will use the results in the next subsection.

7B. ***Limiting characters.*** We will be using the dual category $\mathcal{O}^*$ considered in [Hernandez and Jimbo 2012], whose definition we now recall.

Definition 7.4. Let $\mathcal{O}^*$ be the category of Cartan-diagonalizable $U_q(\mathfrak{b})$-modules V such that V^* is in category $\mathcal{O}$.

A $U_q(\mathfrak{b})$-module V is said to be of lowest ℓ-weight $\boldsymbol{\Psi} \in P_\ell$ if there is $v \in V$ such that $V = U_q(\mathfrak{b})v$ and the following hold:

$$U_q(\mathfrak{b})^- v = \mathbb{C}v, \quad \phi^+_{i,m} v = \Psi_{i,m} v \quad (i \in I,\ m \geq 0).$$

For $\boldsymbol{\Psi} \in P_\ell$, there exists up to isomorphism a unique simple $U_q(\mathfrak{b})$-module $L'(\boldsymbol{\Psi})$ of lowest ℓ-weight $\boldsymbol{\Psi}$. This module belongs to $\mathcal{O}^*$. More precisely, we have:

Proposition 7.5 [Hernandez and Jimbo 2012]. *For* $\boldsymbol{\Psi} \in P_\ell$, $(L'(\boldsymbol{\Psi}))^* \simeq L(\boldsymbol{\Psi}^{-1})$.

We can also define as in Section 5A notions of characters and q-characters for $\mathcal{O}^*$.

We now explain that characters (resp. q-characters) of certain simple objects in the category $\mathcal{O}^-$ can be obtained as limits of characters (resp. q-characters) of finite-dimensional representations. This is known for negative prefundamental representations [Hernandez and Jimbo 2012].

Let $L(\mathbf{\Psi})$ be a simple module whose highest ℓ-weight can be written as a finite product

$$\mathbf{\Psi} = [\omega] \times m \times \prod_{i\in I}\Big(\prod_{r\geq -R_0} \mathbf{\Psi}_{i,q^r}^{u_{i,q^r}}\Big),$$

where $\omega \in P_{\mathbb{Q}},\ R_0 \geq 0,\ u_{i,q^r} \leq 0$ and $m \in \mathcal{M}$ is a dominant monomial. For $R \geq R_0$, set

$$M_R = m \prod_{i\in I}\Big(\prod_{\substack{r\geq -R_0\\ r'\geq 0\\ r-2d_ir'\geq -R}} Y_{i,q^{r-2d_ir'-d_i}}^{-u_{i,q^r}}\Big). \tag{7-21}$$

Theorem 7.6. (1) *We have the limit as formal power series*

$$\widetilde{\chi}_q(L(M_R)) \xrightarrow[R\to+\infty]{} \widetilde{\chi}_q(L(\mathbf{\Psi})) \in \mathbb{Z}[\![A_{i,a}^{-1}]\!]_{i\in I, a\in\mathbb{C}^*}.$$

(2) *We have* $\widetilde{\chi}(L(\overline{\mathbf{\Psi}}^{-1})) = \widetilde{\chi}(L(\mathbf{\Psi}))$ *and so we have the limit as formal power series*

$$\widetilde{\chi}(L(M_R)) \xrightarrow[R\to+\infty]{} \widetilde{\chi}(L(\overline{\mathbf{\Psi}}^{-1})) \in \mathbb{Z}[\![-\alpha_i]\!]_{i\in I}.$$

The proof of Theorem 7.6 is essentially the same as that of [Hernandez and Jimbo 2012, Theorem 6.1], so we just give an outline.

Proof. First let us prove that the dimensions of weight spaces of $\widetilde{L}(\mathbf{\Psi})$ are larger than those of $\widetilde{L}(M_R)$. Consider the tensor product

$$T = \widetilde{L}(\mathbf{\Psi}) \otimes \widetilde{L}(M_R\mathbf{\Psi}^{-1}).$$

By definition of M_R, the ℓ-weight $M_R\mathbf{\Psi}^{-1}$ is a product of $\mathbf{\Psi}_{i,q^r}^{+}$ times $[\lambda]$ for some $\lambda \in P_{\mathbb{Q}}$, so by Theorem 5.8 the module $\widetilde{L}(M_R\mathbf{\Psi}^{-1})$ is a tensor product of positive fundamental representations. Moreover T and $\widetilde{L}(M_R)$ have the same highest ℓ-weight, so $\widetilde{L}(M_R)$ is a subquotient of T. By [Frenkel and Mukhin 2001, Theorem 4.1], each ℓ-weight of $\widetilde{L}(M_R)$ is the product of the highest ℓ-weight $M_R(\varpi(M_R))^{-1}$ by a product of $A_{j,b}^{-1},\ j \in i,\ b \in \mathbb{C}^*$. Hence, by Theorem 5.3, an ℓ-weight of T is an ℓ-weight of $\widetilde{L}(M_R)$ only if it is of the form

$$\mathbf{\Psi}'\varpi(\mathbf{\Psi})^{-1}(\widetilde{M_R}\widetilde{\mathbf{\Psi}}^{-1}),$$

where $\mathbf{\Psi}'$ is an ℓ-weight of $L(\mathbf{\Psi})$ and $\widetilde{M_R}\widetilde{\mathbf{\Psi}}^{-1}$ is the highest ℓ-weight of $\widetilde{L}(M_R\mathbf{\Psi}^{-1})$. We get the result for the dimensions.

Then we prove as in [Hernandez and Jimbo 2012, Section 4.2] that we can define an inductive linear system

$$L(M_0) \to L(M_1) \to \cdots \to L(M_R) \to L(M_{R+1}) \to \cdots$$

from the $L(M_R)$ so that we have the convergence of the action of the subalgebra $\widetilde{U}_q(\mathfrak{g})$ of $U_q(\mathfrak{g})$ generated by the $x_{i,r}^+$ and the $k_i^{-1}x_{i,r}^-$. We get a limiting representation of $\widetilde{U}_q(\mathfrak{g})$ from which one can construct a representation L_1 of $U_q(\mathfrak{b})$ in the category $\mathcal{O}$ and a representation L_2 of $U_q(\mathfrak{b})$ in the category $\mathcal{O}^*$ [Hernandez and Jimbo 2012, Proposition 2.4]. Moreover, L_1 (resp. L_2) is of highest (resp. lowest) ℓ-weight $\mathbf{\Psi}$ (resp. $\overline{\mathbf{\Psi}}$).

By construction, the normalized q-character of L_1 is the limit of the normalized q-characters $\widetilde{\chi}_q(L(M_R))$ as formal power series. Combining with the result of the first paragraph of this proof, the representation L_1 is necessarily simple isomorphic to $L(\mathbf{\Psi})$. We have proved the first statement in the theorem.

Now, by construction L_2^* is in the category $\mathcal{O}$ with highest ℓ-weight $\overline{\mathbf{\Psi}}^{-1}$ and satisfies $\widetilde{\chi}(L_2^*) = \widetilde{\chi}(L_1)$. To conclude, it suffices to prove that L_2^* is irreducible. This is proved as in [Hernandez and Jimbo 2012, Theorem 6.3]. □

We have the following application:

Theorem 7.7. *Let $L(\mathbf{\Psi})$ be a simple module in the category $\mathcal{O}^-$ such that $\widetilde{\mathbf{\Psi}} = \mathbf{\Psi}$. Then its image by D^{-1} in $K_0(\mathcal{O}^+)$ is simple equal to $D^{-1}([L(\mathbf{\Psi})]) = [L(\overline{\mathbf{\Psi}}^{-1})]$.*

Proof. From Example 5.14(i), the property is satisfied by negative prefundamental representations. Since these representations generate the fraction field of $K_0(\mathcal{O}^-)$, it suffices to show that the assignment $[L(\mathbf{\Psi})] \mapsto [L(\overline{\mathbf{\Psi}}^{-1})]$ for ℓ-weights $\mathbf{\Psi}$ satisfying $\mathbf{\Psi} = \widetilde{\mathbf{\Psi}}$ is multiplicative (recall that D is a morphism of $\mathcal{E}$-algebras). Let us use the same notation as in the proof of Theorem 7.6 above. For $L(\mathbf{\Psi}) \simeq L_1$ and $L(\mathbf{\Psi}') \simeq L_1'$ simple modules in $\mathcal{O}^-$ with $\widetilde{\mathbf{\Psi}} = \mathbf{\Psi}$ and $\widetilde{\mathbf{\Psi}'} = \mathbf{\Psi}'$, we have the corresponding modules L_2, L_2' in $\mathcal{O}^*$. We consider the decomposition

$$[L(\mathbf{\Psi}) \otimes L(\mathbf{\Psi}')] = \sum_{\mathbf{\Psi}'', \widetilde{\mathbf{\Psi}''} = \mathbf{\Psi}''} m_{\mathbf{\Psi}''}[L(\mathbf{\Psi}'')]$$

in $K_0(\mathcal{O}^-)$ with $m_{\mathbf{\Psi}''} \in \mathcal{E}$. Each $\chi_q(L(\mathbf{\Psi}''))$ is obtained as a limit as in Theorem 7.6, and by construction the corresponding modules $L'(\mathbf{\Psi}'')$ in the category $\mathcal{O}^*$ satisfy

$$[L_2 \otimes L_2'] = \sum_{\mathbf{\Psi}'', \widetilde{\mathbf{\Psi}''} = \mathbf{\Psi}''} m'_{\mathbf{\Psi}''}[L'(\overline{\mathbf{\Psi}''})],$$

where each $m'_{\mathbf{\Psi}''} \in \mathcal{E}$ is obtained from $m_{\mathbf{\Psi}''}$ via the substitution $[\omega] \mapsto [-\omega]$. This implies

$$[L_2^* \otimes (L_2')^*] = \sum_{\mathbf{\Psi}'', \widetilde{\mathbf{\Psi}''} = \mathbf{\Psi}''} m_{\mathbf{\Psi}''}[(L'(\overline{\mathbf{\Psi}''}))^*],$$

that is, in view of Proposition 7.5,

$$[L(\overline{\mathbf{\Psi}}^{-1})] \otimes [L(\overline{\mathbf{\Psi}'}^{-1})] = \sum_{\mathbf{\Psi}'', \widetilde{\mathbf{\Psi}''} = \mathbf{\Psi}''} m_{\mathbf{\Psi}''}[L(\overline{\mathbf{\Psi}''}^{-1})]. \qquad \square$$

Example 7.8. Applying D to (6-14), we get the relation in $K_0(\mathcal{O}^-)$

$$[L(\bar{\boldsymbol{\Psi}}^{-1})\otimes L^-_{i,q^{-r}}]=\Bigl[\bigotimes_{j:C_{j,i}\neq 0} L^-_{j,q^{-r+d_jC_{j,i}}}\Bigr]+[-\alpha_i]\Bigl[\bigotimes_{j:C_{j,i}\neq 0} L^-_{j,q^{-r-d_jC_{j,i}}}\Bigr],$$

where

$$\bar{\boldsymbol{\Psi}}^{-1}=[-\omega_i]Y_{i,q^{-r+d_i}}\prod_{j:C_{j,i}<0}\boldsymbol{\Psi}^{-1}_{j,q^{-r+d_jC_{j,i}}}.$$

7C. *Proof of Conjecture 7.2 for $\mathfrak{g}=\widehat{\mathfrak{sl}}_2$.* In this section we give an explicit description of all simple modules in $\mathcal{O}^+$ and in $\mathcal{O}^-$ for $\mathfrak{g}=\widehat{\mathfrak{sl}}_2$.

A q-set is a subset of $\mathbb{C}^*$ of the form $\{aq^{2r}\mid R_1\leq r\leq R_2\}$ for some $a\in\mathbb{C}^*$ and $R_1\leq R_2\in\mathbb{Z}\cup\{-\infty,+\infty\}$. The KR-modules $W_{k,a}$, $W_{k',b}$ are said to be in *special position* if the union of $\{a,aq^2,\dots,aq^{2(k-1)}\}$ and $\{b,bq^2,\dots,bq^{2(k'-1)}\}$ is a q-set which contains both properly. The KR-module $W_{k,a}$ and the prefundamental representation L^+_b are said to be in special position if the union of $\{a,aq^2,aq^4,\dots,aq^{2(k-1)}\}$ and $\{bq,bq^3,bq^5,\dots\}$ is a q-set which contains both properly. Two positive prefundamental representations are never in special position. Two representations are in *general position* if they are not in special position.

The invertible elements in the category $\mathcal{O}^+$ are the 1-dimensional representations $[\omega]$.

Theorem 7.9. *Suppose that $\mathfrak{g}=\widehat{\mathfrak{sl}}_2$. The prime simple objects in the category $\mathcal{O}^+$ are the positive prefundamental representations and the* KR*-modules* (*up to invertibles*). *Any simple object in $\mathcal{O}^+$ can be factorized in a unique way as a tensor product of prefundamental representations and* KR*-modules* (*up to permutation of the factors and to invertibles*). *Moreover*, *such a tensor product is simple if and only if all its factors are pairwise in general position.*

Proof. As in the classical case of finite-dimensional representations, it is easy to check that every positive ℓ-weight has a unique factorization as a product of highest ℓ-weights of KR-modules and positive prefundamental representations in pairwise general position. Hence it suffices to prove the equivalence in the last sentence. By [Chari and Pressley 1994], the result is known for finite-dimensional representations. Now, by using Section 7B, this result implies that a tensor product with factors which are in general position is simple. Conversely, it is known that a tensor product of KR-modules which are in special position is not simple. Also, it is easy to see that the tensor product of a KR-module and a positive prefundamental representation which are in special position is not simple. □

Remark 7.10. (i) This is a generalization of the factorization of simple representations in $\mathcal{C}$ when $\mathfrak{g}=\widehat{\mathfrak{sl}}_2$ [Chari and Pressley 1994].

(ii) This result for $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$ implies that all simple objects in $\mathcal{O}^+$ are real and that their factorization into prime representations is unique.

(iii) In [Mukhin and Young 2014], a factorization is proved for simple modules in $\hat{\mathcal{O}}$ when $\mathfrak{g} = \widehat{\mathfrak{sl}}_2$. But the factorization is not unique in the category $\hat{\mathcal{O}}$; see Remark 3.10.

(iv) By Proposition 5.13, our result implies a similar factorization in the category $\mathcal{O}^-$.

(v) The combinatorics of q-sets in pairwise general position is very similar to the combinatorics of triangulations of the ∞-gon studied in [Grabowski and Gratz 2014], in relation with certain cluster structures of infinite rank. However, in that paper only arcs (m, n) joining two integers m and n are considered, whereas we also allow arcs of the form $(m, +\infty)$ corresponding to positive prefundamental representations. Also, we are only interested in one mutation class, namely the mutation class of the initial triangulation $\{(m, +\infty) \mid m \in \mathbb{Z}\}$.

Theorem 7.11. *Conjecture 7.2 is true in the* $\mathfrak{sl}_2$*-case.*

Proof. Theorem 7.9 provides an explicit factorization of simple objects in $\mathcal{O}^+_{2\mathbb{Z}}$ into positive prefundamental representations and finite-dimensional KR-modules. In particular, we get an explicit q-character formula for such a simple object and so a complete explicit description of the Grothendieck ring $K_0(\mathcal{O}^+_{2\mathbb{Z}})$. The cluster algebra $\mathcal{A}(\Gamma)$ can also be explicitly described by using triangulations of the ∞-gon (see Remark 7.10). Hence we can argue as in [Hernandez and Leclerc 2010, Section 13.4]. □

7D. ***Equivalence of conjectures.*** In general, we have the following:

Theorem 7.12. *Conjecture 7.2 is equivalent to Conjecture* 5.2 *of* [Hernandez and Leclerc 2016].

Combining with the recent results in [Qin 2015], this would imply a part of Conjecture 7.2 for *ADE* types, namely that all cluster monomials are classes of real simple objects.

As recalled in the introduction, [Hernandez and Leclerc 2016, Conjecture 5.2] states that $\mathcal{C}^-_{\mathbb{Z}}$ is the monoidal categorification of a cluster algebra $\mathcal{A}(G^-)$. Note that $\mathcal{C}^-_{\mathbb{Z}}$ is a subcategory of $\mathcal{O}^-_{2\mathbb{Z}}$ and $\mathcal{O}^+_{2\mathbb{Z}}$.

Proof. For $N > 0$, let $\mathcal{C}_N$ be the category of finite-dimensional $U_q(\mathfrak{g})$-modules V satisfying

$$[V] \in \mathbb{Z}[\![V_{i,q^m}]\!]_{i\in I,\, -2dN-d_i \leq m < d(N+2)-d_i} \subset K_0(\mathcal{C}).$$

It is a monoidal category similar to the categories considered in [Hernandez and Leclerc 2010]. It contains the KR-module $W_{i,r,N}$ with highest monomial $M_{i,r,N}$ given in equation (6-17), where $i \in I$ and $-d(N+2) < r \leq 2dN$. The Grothendieck ring $K_0(\mathcal{C}_N)$ has a cluster algebra structure with an initial seed consisting of these KR-modules $W_{i,r,N}$ (here we use the initial seed as in [Hernandez and Leclerc 2016]). We have established in the proof of Theorem 4.2 that $K_0(\mathcal{C}_N) \otimes \mathcal{E}$ may be seen as a subalgebra of $K_0(\mathcal{O}^+_{2\mathbb{Z}})$ by using the identification of $z_{i,r}$ with the element defined in equation (6-19). This induces embeddings $K_0(\mathcal{C}_N) \subset K_0(\mathcal{C}_{N+1})$,

$$K_0(\mathcal{C}_1) \subset K_0(\mathcal{C}_2) \subset K_0(\mathcal{C}_3) \subset \cdots \subset K_0(\mathcal{O}^+_{2\mathbb{Z}}),$$

which are not the naive embeddings obtained from the inclusion of categories $\mathcal{C}_N \subset \mathcal{C}_{N+1}$. The cluster monomials in $K_0(\mathcal{C}_N)$ corresponds now to cluster monomials in $K_0(\mathcal{O}^+_{2\mathbb{Z}})$.

Note that by Theorem 7.7 we may consider simultaneously the statement of Conjecture 7.2 for $\mathcal{O}^+_{2\mathbb{Z}}$ or for $\mathcal{O}^-_{2\mathbb{Z}}$.

Suppose that [Hernandez and Leclerc 2016, Conjecture 5.2] is true. This implies that the cluster monomials in $K_0(\mathcal{C}_N)$ are the real simple modules for any $N > 0$. Consider a cluster monomial in $\mathcal{O}^+_{2\mathbb{Z}}$. Then for N large enough, we have a corresponding real representation V_N in $K_0(\mathcal{C}_N)$. The highest monomial of V_N is a Laurent monomial in the $m_{i,r,N}$ of the form considered in Theorem 7.6. By Theorem 7.6, $\chi_q(V_N)$ converges to the q-character of a simple module V in $\mathcal{O}^-_{2\mathbb{Z}}$ when $N \to +\infty$. Moreover V is real, as the q-character of $V \otimes V$ is obtained as a limit of simple q-character $\chi_q(V_N \otimes V_N)$ by Theorem 7.6. Conversely, every real simple module V in $\mathcal{O}^-_{2\mathbb{Z}}$ is obtained as such a limit of simple modules V_N. Moreover since V is real, V_N is real (since $\chi_q(V_N \otimes V_N)$ is an upper q-character of $V \otimes V$ in the sense of [Hernandez 2010, Corollary 5.8]). For N large enough the modules V_N correspond to the same cluster monomial, which is therefore identified with V.

Conversely, suppose that Conjecture 7.2 is true. Consider a cluster monomial χ in $K_0(\mathcal{C}^-_{\mathbb{Z}})$. The cluster variables occurring in χ are produced via sequences of mutations from a finite number of KR-modules in the initial seed. By the proof of Proposition 6.1, there is a seed in $K_0(\mathcal{O}^-_{2\mathbb{Z}})$ containing these KR-modules (and the quiver of this seed has the same arrows joining the corresponding vertices). By our hypothesis, χ is the class of a simple real module as an element of $K_0(\mathcal{O}^-_{2\mathbb{Z}}) \supset K_0(\mathcal{C}^-_{\mathbb{Z}})$. Hence it is also real simple in $K_0(\mathcal{C}^-_{\mathbb{Z}})$. Now consider a real simple module $[V]$ in $K_0(\mathcal{C}^-_{\mathbb{Z}})$. It corresponds to a cluster monomial in $K_0(\mathcal{O}^-_{2\mathbb{Z}})$ which is a cluster monomial in $K_0(\mathcal{C}^-_{\mathbb{Z}})$ by the same arguments. □

7E. *Web property theorem.* Let us prove the following generalization of the main result of [Hernandez 2010]. If Conjecture 7.2 holds, then the statement of the next theorem is a necessary condition for simple modules in the same seed.

Theorem 7.13. *Let $S_1, \dots, S_N$ be simple objects in $\mathcal{O}^+$ (resp. in $\mathcal{O}^-$). Then $S_1 \otimes \cdots \otimes S_N$ is simple if and only if the tensor products $S_i \otimes S_j$ are simple for $i \leq j$.*

Proof. By Proposition 5.13, this is equivalent to proving the statement in the category $\mathcal{O}^-$. Note that the "only if" part is clear. For the "if" part of the statement, we may assume without loss of generality that the zeros and poles of the highest ℓ-weights of the S_i are in $q^{\mathbb{Z}}$ (see (ii) in Remark 4.3). For each simple module S_i, consider a corresponding simple finite-dimensional module $L(M_{R,i})$ as in Section 7B. Since $S_i \otimes S_j$ is simple, there exists R_1 such that for $R \geq R_1$ the tensor product $L(M_{R,i}) \otimes L(M_{R,j})$ is simple. Indeed, by Theorem 7.6, $\widetilde{\chi}_q(S_i \otimes S_j)$ is the limit of the $\widetilde{\chi}_q(L(M_{R,i}M_{R,j}))$. More precisely, there is R_1 such that for $R \geq R_1$ the image of $\widetilde{\chi}_q(S_i \otimes S_j)$ in $\mathbb{Z}[A^{-1}_{i,q^r}]_{i\in I, r\geq -R_1+r_i}$ is equal to $\widetilde{\chi}_q(L(M_{R,i}M_{R,j}))$. This implies that $\widetilde{\chi}_q(L(M_{R,i}M_{R,j})) = \widetilde{\chi}_q(L(M_{R,i}))\widetilde{\chi}_q(L(M_{R,j}))$.

Now, by [Hernandez 2010], $L(M_{R,1}) \otimes \cdots \otimes L(M_{R,N})$ is simple isomorphic to $L(M_{R,1} \cdots M_{R,N})$. This implies that the character of $S_1 \otimes \cdots \otimes S_N$ is the same as the character of the simple module with the same highest ℓ-weight. Hence they are isomorphic. □

Remark 7.14. This provides an alternative proof of Theorem 7.9.

7F. *Another conjecture.* To conclude, let us state another general conjecture. Although the cluster algebra structure presented in this paper does not appear in the statement, this conjecture arises naturally if we compare Theorem 4.2 with the results of [Hernandez and Leclerc 2016].

We consider a simple finite-dimensional representation $L(m)$ whose dominant monomial m satisfies $m \in \mathbb{Z}[Y_{i,aq^r}]_{(i,r)\in W, r\leq R}$ for a given $R \in \mathbb{Z}$. We have the corresponding truncated q-character [Hernandez and Leclerc 2010]

$$\chi_q^{\leq R}(L(m)) \in m\mathbb{Z}[A^{-1}_{i,q^r}]_{i\in I, r\leq R-d_i},$$

which is the sum (with multiplicity) of the monomials m' occurring in $\chi_q(L(m))$ and satisfying $m(m')^{-1} \in \mathbb{Z}[A_{i,q^r}]_{i\in I, r\leq R-d_i}$. As in the statement of Theorem 5.5, we consider

$$(\chi^+)^{\leq R}(L(m)) \in \mathrm{Frac}(K_0(\mathcal{O}^+)),$$

obtained from $\chi_q^{\leq R}(L(m))$ by replacing each variable $Y_{i,a}$ by $[\omega_i]\ell^+_{i,aq_i^{-1}}(\ell^+_{i,aq_i})^{-1}$.

Let us set

$$W_R = \{(i,r) \in W \mid R \geq r > R - 2d_i\}.$$

For $(i,r) \in W_R$, we set $u_{i,r}$ to be the maximum of 0 and the powers $u_{i,q^r}(m')$ of Y_{i,q^r} in all monomials m' occurring in $\chi_q^{\leq R}(L(m))$. Let

$$\boldsymbol{\Psi}_R = \prod_{(i,r)\in W_R} \boldsymbol{\Psi}^{u_{i,r}}_{i,q^{r+d_i}} \quad \text{and} \quad \boldsymbol{\Psi} = m\boldsymbol{\Psi}_R.$$

The representations $L(\boldsymbol{\Psi})$, $L(\boldsymbol{\Psi}_R)$ are in the category $\mathcal{O}^+$. By Theorem 5.8, $L(\boldsymbol{\Psi}_R)$ is a simple tensor product of positive prefundamental representations.

Conjecture 7.15. *We have the relation in* $\mathrm{Frac}(K_0(\mathcal{O}^+))$

$$\chi^+(L(\boldsymbol{\Psi})) = (\chi^+)^{\leq R}(L(m)) \prod_{(i,r)\in W_R} (\ell^+_{i,q^{r+d_i}})^{u_{i,r}} = (\chi^+)^{\leq R}(L(m))\chi^+(L(\boldsymbol{\Psi}_R)).$$

Remark 7.16. By taking the q-character, the statement is equivalent to the following q-character formula:

$$\chi_q(L(\boldsymbol{\Psi})) = \chi_q^{\leq R}(L(m))\chi_q(L(\boldsymbol{\Psi}_R)).$$

In some cases the conjecture is already proved:

Example 7.17. (i) In the case $\chi_q^{\leq R}(L(m)) = \chi_q(L(m))$, we have $u_{i,r} = 0$ for $(i,r) \in W_R$. Here the conjecture reduces to the generalized Baxter relations of Theorem 5.5.

(ii) In the case $R = r + d_i$ and $m = Y_{i,q^{r-d_i}}$, we have $\chi_q(L(m))^{\geq R} = m(1 + A^{-1}_{i,q^r})$. The conjecture reduces to the relations we have established in formula (6-14).

(iii) As discussed above, it follows from Theorem 7.12 and from the main result of [Qin 2015] that, for ADE-types, all cluster monomials are classes of real simple objects. In particular, for ADE-types, Conjecture 7.15 holds for all simple modules $L(m)$ which are cluster monomials (for the cluster algebra structure defined in [Hernandez and Leclerc 2016]). Indeed we may assume that $R = 0$. It is proved in in the same paper that, for any dominant monomial m, the truncated q-character $\chi_q^{\leq 0}(L(m))$ is an element of the cluster algebra $\mathcal{A}(G^-)$ defined in the proof of Proposition 6.1. In $\chi_q^{\leq 0}(L(m))$ we perform the same substitution as in Theorem 5.5 above (that is, we apply the ring homomorphism F of the proof of Proposition 6.1). If we assume that $\chi_q^{\leq 0}(L(m))$ is a cluster variable of $\mathcal{A}(G^-)$, then it follows from the proof of Proposition 6.1 that

$$y := F(\chi_q^{\leq 0}(L(m))) \prod_{(i,r)\in W_0} z_{i,q^{r+d_i}}^{u_{i,r}}$$

is a cluster variable in $\mathcal{A}(H^-)$. Using Theorem 7.12, we deduce that y is the positive prefundamental character of a simple module. Since F is multiplicative, the argument readily extends to simple modules $L(m)$ such that $\chi_q^{\leq 0}(L(m))$ is a cluster monomial.

References

[Baxter 1972] R. J. Baxter, "Partition function of the eight-vertex lattice model", *Ann. Physics* **70** (1972), 193–228. MR 0290733 Zbl 0236.60070

[Bazhanov et al. 1997] V. V. Bazhanov, S. L. Lukyanov, and A. B. Zamolodchikov, "Integrable structure of conformal field theory, II: Q-operator and DDV equation", *Comm. Math. Phys.* **190**:2 (1997), 247–278. MR 1489571 Zbl 0908.35114

[Bazhanov et al. 2002] V. V. Bazhanov, A. N. Hibberd, and S. M. Khoroshkin, "Integrable structure of $\mathcal{W}_3$ conformal field theory, quantum Boussinesq theory and boundary affine Toda theory", *Nuclear Phys. B* **622**:3 (2002), 475–547. MR 1880098 Zbl 0983.81088

[Beck 1994] J. Beck, "Braid group action and quantum affine algebras", *Comm. Math. Phys.* **165**:3 (1994), 555–568. MR 1301623 Zbl 0807.17013

[Boos et al. 2009] H. Boos, M. Jimbo, T. Miwa, F. Smirnov, and Y. Takeyama, "Hidden Grassmann structure in the XXZ model, II: Creation operators", *Comm. Math. Phys.* **286**:3 (2009), 875–932. MR 2472021 Zbl 1173.82003

[Chari 1995] V. Chari, "Minimal affinizations of representations of quantum groups: the rank 2 case", *Publ. Res. Inst. Math. Sci.* **31**:5 (1995), 873–911. MR 1367675 Zbl 0855.17010

[Chari and Pressley 1994] V. Chari and A. Pressley, *A guide to quantum groups*, Cambridge University Press, 1994. MR 1300632 Zbl 0839.17009

[Damiani 1998] I. Damiani, "La R-matrice pour les algèbres quantiques de type affine non tordu", *Ann. Sci. École Norm. Sup.* (4) **31**:4 (1998), 493–523. MR 1634087 Zbl 0911.17005

[Drinfeld 1987] V. G. Drinfeld, "A new realization of Yangians and of quantum affine algebras", *Dokl. Akad. Nauk SSSR* **296**:1 (1987), 13–17. In Russian; translated in *Soviet Math. Dokl.* **36**:2 (1988), 212–216. MR 914215

[Fomin and Zelevinsky 2002] S. Fomin and A. Zelevinsky, "Cluster algebras, I: Foundations", *J. Amer. Math. Soc.* **15**:2 (2002), 497–529. MR 1887642 Zbl 1021.16017

[Fomin and Zelevinsky 2003] S. Fomin and A. Zelevinsky, "Cluster algebras: notes for the CDM-03 conference", pp. 1–34 in *Current developments in mathematics, 2003*, edited by B. Mazur et al., International Press, Somerville, MA, 2003. MR 2132323 Zbl 1119.05108

[Frenkel and Hernandez 2015] E. Frenkel and D. Hernandez, "Baxter's relations and spectra of quantum integrable models", *Duke Math. J.* **164**:12 (2015), 2407–2460. MR 3397389 Zbl 1332.82022

[Frenkel and Mukhin 2001] E. Frenkel and E. Mukhin, "Combinatorics of q-characters of finite-dimensional representations of quantum affine algebras", *Comm. Math. Phys.* **216**:1 (2001), 23–57. MR 1810773 Zbl 1051.17013

[Frenkel and Reshetikhin 1999] E. Frenkel and N. Reshetikhin, "The q-characters of representations of quantum affine algebras and deformations of $\mathcal{W}$-algebras", pp. 163–205 in *Recent developments in quantum affine algebras and related topics* (Raleigh, NC, 1998), edited by N. Jing and K. C. Misra, Contemp. Math. **248**, Amer. Math. Soc., Providence, RI, 1999. Corrected version available at https://arxiv.org/abs/math/9810055. MR 1745260 Zbl 0973.17015

[Gekhtman et al. 2010] M. Gekhtman, M. Shapiro, and A. Vainshtein, *Cluster algebras and Poisson geometry*, Mathematical Surveys and Monographs **167**, American Mathematical Society, Providence, RI, 2010. MR 2683456 Zbl 1217.13001

[Grabowski and Gratz 2014] J. E. Grabowski and S. Gratz, "Cluster algebras of infinite rank", *J. Lond. Math. Soc.* (2) **89**:2 (2014), 337–363. MR 3188622 Zbl 1308.13034

[Grabowski and Launois 2014] J. E. Grabowski and S. Launois, "Graded quantum cluster algebras and an application to quantum Grassmannians", *Proc. Lond. Math. Soc.* (3) **109**:3 (2014), 697–732. MR 3260291 Zbl 1315.13036

[Hernandez 2005] D. Hernandez, "Representations of quantum affinizations and fusion product", *Transform. Groups* **10**:2 (2005), 163–200. MR 2195598 Zbl 1102.17009

[Hernandez 2007] D. Hernandez, "On minimal affinizations of representations of quantum groups", *Comm. Math. Phys.* **276**:1 (2007), 221–259. MR 2342293 Zbl 1141.17011

[Hernandez 2010] D. Hernandez, "Simple tensor products", *Invent. Math.* **181**:3 (2010), 649–675. MR 2660455 Zbl 1221.17015

[Hernandez 2015] D. Hernandez, "Asymptotical representations, spectra of quantum integrable systems and cluster algebras", pp. 1399–1402 in *Enveloping Algebras and Geometric Representation Theory*, 12 **2**, 2015. Abstracts from the workshop held May 10–16, 2015, Organized by Iain Gordon, Bernard Leclerc and Wolfgang Soergel.

[Hernandez and Jimbo 2012] D. Hernandez and M. Jimbo, "Asymptotic representations and Drinfeld rational fractions", *Compos. Math.* **148**:5 (2012), 1593–1623. MR 2982441 Zbl 1266.17010

[Hernandez and Leclerc 2010] D. Hernandez and B. Leclerc, "Cluster algebras and quantum affine algebras", *Duke Math. J.* **154**:2 (2010), 265–341. MR 2682185 Zbl 1284.17010

[Hernandez and Leclerc 2013] D. Hernandez and B. Leclerc, "Monoidal categorifications of cluster algebras of type A and D", pp. 175–193 in *Symmetries, integrable systems and representations*, edited by K. Iohara et al., Springer Proc. Math. Stat. **40**, Springer, 2013. MR 3077685 Zbl 1317.13052

[Hernandez and Leclerc 2016] D. Hernandez and B. Leclerc, "A cluster algebra approach to q-characters of Kirillov–Reshetikhin modules", *J. Eur. Math. Soc.* (*JEMS*) **18**:5 (2016), 1113–1159. MR 3500832 Zbl 06578174

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, 1990. MR 1104219 Zbl 0716.17022

[Kang et al. 2014] S.-J. Kang, M. Kashiwara, M. Kim, and S.-j. Oh, "Monoidal categorification of cluster algebras", preprint, 2014. arXiv 1412.8106

[Kang et al. 2015] S.-J. Kang, M. Kashiwara, M. Kim, and S.-j. Oh, "Monoidal categorification of cluster algebras, II", preprint, 2015. arXiv 1502.06714

[Kimura and Qin 2014] Y. Kimura and F. Qin, "Graded quiver varieties, quantum cluster algebras and dual canonical basis", *Adv. Math.* **262** (2014), 261–312. MR 3228430 Zbl 1331.13016

[Kojima 2008] T. Kojima, "Baxter's Q-operator for the W-algebra W_N", *J. Phys. A* **41**:35 (2008), 355206, 16. MR 2426018 Zbl 1146.81032

[Mukhin and Young 2014] E. Mukhin and C. A. S. Young, "Affinization of category $\mathcal{O}$ for quantum groups", *Trans. Amer. Math. Soc.* **366**:9 (2014), 4815–4847. MR 3217701 Zbl 1306.17005

[Nakajima 2004] H. Nakajima, "Quiver varieties and t-analogs of q-characters of quantum affine algebras", *Ann. of Math.* (2) **160**:3 (2004), 1057–1097. MR 2144973 Zbl 1140.17015

[Nakajima 2011] H. Nakajima, "Quiver varieties and cluster algebras", *Kyoto J. Math.* **51**:1 (2011), 71–126. MR 2784748 Zbl 1223.13013

[Qin 2015] F. Qin, "Triangular bases in quantum cluster algebras and monoidal categorification conjectures", preprint, 2015. arXiv 1501.04085

Communicated by Edward Frenkel
Received 2016-04-19 Accepted 2016-08-09

david.hernandez@imj-prg.fr *Institut de Mathématiques de Jussieu-Paris Rive Gauche, Université Paris-Diderot Paris 7, Bâtiment Sophie Germain, Case 7012, 75205 Paris Cedex 13, France*

bernard.leclerc@unicaen.fr *Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse-Normandie, 14032 Caen, France*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 10 No. 9 2016