Algebra & Number Theory

Volume 20
2026
No. 8

msp

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada	Philippe Michel	École Polytechnique Fédérale de Lausanne
Olivier Benoist	Ecole Normale Supérieure, France	Martin Olsson	University of California, Berkeley, USA
Bhargav Bhatt	University of Michigan, USA	Irena Peeva	Cornell University, USA
Frank Calegari	University of Chicago, USA	Jonathan Pila	University of Oxford, UK
J-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France	Anand Pillay	University of Notre Dame, USA
Brian D. Conrad	Stanford University, USA	Bjorn Poonen	Massachusetts Institute of Technology, USA
Samit Dasgupta	Duke University, USA	Victor Reiner	University of Minnesota, USA
Hélène Esnault	Freie Universität Berlin, Germany	Peter Sarnak	Princeton University, USA
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Benjamin Schraen	Université Claude Bernard Lyon 1, France
Sergey Fomin	University of Michigan, USA	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	SUNY Buffalo, USA
Wee Teck Gan	National University of Singapore	Shunsuke Takagi	University of Tokyo, Japan
Andrew Granville	Université de Montréal, Canada	Pham Huu Tiep	Rutgers University, USA
Ben J. Green	University of Oxford, UK	Ravi Vakil	Stanford University, USA
Christopher Hacon	University of Utah, USA	Akshay Venkatesh	Institute for Advanced Study, USA
Roger Heath-Brown	Oxford University, UK	Melanie Matchett Wood	Harvard University, USA
János Kollár	Princeton University, USA	Shou-Wu Zhang	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA		

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

The subscription price for 2026 is US $590/year for the electronic version, and $865/year (+$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 2000 Allston Way #59, Berkeley, CA 94701-4004, is published continuously online.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2026 Mathematical Sciences Publishers

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1479

Supersymmetric monoidal categories

Steven V Sam and Andrew Snowden

We develop the idea of a supersymmetric monoidal supercategory, following ideas of Kapranov. Roughly, this is a monoidal category in which the objects and morphisms are $\mathbb{Z}/2$-graded, equipped with isomorphisms $X \otimes Y \to Y \otimes X$ of parity $|X||Y|$ on homogeneous objects. There are two fundamental examples: the groupoid of spin-sets, and the category of isomeric vector spaces equipped with the half tensor product; other important examples can be derived from these (such as the category of linear spin species). There are also two general constructions. The first is the exterior algebra of a supercategory (due to Ganter and Kapranov). The second is a construction we introduce, called Clifford eversion. This defines an equivalence between a certain 2-category of supersymmetric monoidal supercategories and a corresponding 2-category of symmetric monoidal supercategories. We use our theory to better understand some aspects of the isomeric superalgebra, such as certain factors of $\sqrt{2}$ in the theory of Q-symmetric functions and Schur–Sergeev duality.

1. Introduction 1479
2. Supercategories 1486
3. Monoidal structures 1489
4. Some 2-category theory 1503
5. Spin-symmetric groups 1511
6. Linear spin-species 1517
7. Isomeric vector spaces 1522
8. Clifford eversion 1526
9. Schur–Sergeev duality 1536
Appendix. Bott periodicity 1540
References 1541

1. Introduction

The categorical analog of a commutative algebra is a symmetric (or braided) linear monoidal category. These categories have played a fundamental role in category theory and representation theory, among other places. The purpose of this paper is to develop a categorical analog of supercommutative superalgebras, following ideas of Kapranov [2018, Section 3.4]; we call these supersymmetric (or superbraided) monoidal supercategories. In the rest of this introduction, we explain the basic idea behind the definition, some of

Sam was supported by NSF grant DMS-1812462. Snowden was supported by NSF grant DMS-1453893.
MSC2020: primary 18M05; secondary 05E10.
Keywords: superalgebra, supercategories.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY).
Open Access made possible by subscribing institutions via Subscribe to Open.

our results and applications, and compare our contributions to the existing discussions of supercategories in the literature.

1.1. ***The definition.*** Recall that a supercommutative superalgebra is a $\mathbb{Z}/2$-graded algebra such that $xy = (-1)^{|x||y|}yx$ holds for all homogeneous elements x and y. Here we write $|x| \in \mathbb{Z}/2$ for the degree of a homogeneous element x. Thus, we expect that, in a superbraided monoidal category, there should be even and odd homogeneous objects, and there should be an isomorphism between $X \otimes Y$ and $(-1)^{|X||Y|}$ "times" $Y \otimes X$ for homogeneous objects X and Y.

The first puzzle here is what it means to negate an object. This has been well understood for some time. Let V be a finite-dimensional super-vector space. (Recall that a super-vector space is simply a $\mathbb{Z}/2$-graded vector space.) The superdimension of V is defined to be $\mathrm{sdim}(V) = \dim(V_0) - \dim(V_1)$. Let $\Pi(V)$ be the super-vector space given by $\Pi(V)_i = V_{i+1}$, i.e., $\Pi(V)$ is obtained from V by interchanging the two graded pieces. Then $\mathrm{sdim}(\Pi(V)) = -\mathrm{sdim}(V)$. Thus, $\Pi(V)$ can be thought of as the negative of V.

The identity map $V \to \Pi(V)$ is an odd-degree isomorphism. This map characterizes $\Pi(V)$ uniquely up to even isomorphism: if W is any super-vector space equipped with an odd-degree isomorphism $V \to W$, then there is a unique even isomorphism $W \to \Pi(V)$ making the triangle commute. We can thus speak of $\Pi(V)$ for any object V in a category in which the morphisms are $\mathbb{Z}/2$-graded. Precisely, we define a *supercategory* to be a category enriched over super-vector spaces, so that all Hom spaces are super-vector spaces. Given an object V in a supercategory, we define $\Pi(V)$ to be an object equipped with an odd-degree isomorphism $V \to \Pi(V)$ if such an object and morphism exist. (We assume in what follows that they always do.) We can thus regard $\Pi(V)$ as the negative of V.

Following our initial intuition, suppose that we have a $\mathbb{Z}/2$-supercategory (i.e., one in which the objects are $\mathbb{Z}/2$-graded) equipped with a monoidal structure $\otimes$. Then we see that a superbraiding should be a natural isomorphism

$$\beta_{X,Y} : X \otimes Y \to \Pi^{|X||Y|}(Y \otimes X),$$

defined for all homogeneous objects X and Y.

Of course, as with ordinary braidings, our superbraiding should satisfy some version of the hexagon axioms. It turns out that the usual hexagon axioms are inconsistent in this setting due to the signs that arise from composition in supercategories, but one can attempt to resolve the inconsistencies by allowing the

0-category	1-category
vector space	supercategory
commutative algebra	symmetric monoidal supercategory
super-vector space	$\mathbb{Z}/q$-supercategory
commutative superalgebra	symmetric monoidal $\mathbb{Z}/q$-supercategory
supercommutative superalgebra	supersymmetric monoidal $\mathbb{Z}/q$-supercategory

Table 1. Some analogies relevant to this paper. Here q should be even.

hexagons to commute up to scalars. This ends up working if we assume the ground field has an eighth root of unity; however, the scalars that come up are not at all obvious. By working with $\mathbb{Z}/4$-supercategories (or $\mathbb{Z}/q$-supercategories for any q divisible by 4), one can get by with much nicer systems of scalars that only involve signs (and thus make sense over general fields). In both cases, we introduce a notion of factor system (Definition 3.5) to axiomatize the conditions on the scalars that appear in the hexagon axioms. One of the main contributions of this article is identifying these conditions and verifying them in important examples of interest.

We have thus arrived at the main definition: a superbraiding is a natural isomorphism β as above such that the hexagon axioms hold up to scalars determined by a choice of factor system. See Sections 3.5 and 3.6 for details.

1.2. ***The 2-categorical perspective.*** One can define the notion of supercommutative superalgebra in elementary terms, as we did above. There is also a more abstract approach: one endows the monoidal category of super-vector spaces with a new symmetry τ (namely $x \otimes y \mapsto (-1)^{|x||y|} y \otimes x$), and realizes supercommutative superalgebras as the *commutative* algebra objects in this category. The more abstract approach has several benefits; for example, any properties of commutative algebras in symmetric tensor categories can immediately be applied to supercommutative superalgebras, and it also enables one to easily consider similar structures, such as Lie superalgebras.

In our discussion of superbraided supercategories above, we have taken the elementary approach. We now explain a version of the abstract approach. Let $\mathscr{C}$ be the 2-category of $\mathbb{Z}/q$-supercategories, with q even. This 2-category has a natural monoidal structure $\boxtimes$, given by a standard product construction for supercategories. Given two $\mathbb{Z}/q$-supercategories $\mathcal{A}$ and $\mathcal{B}$, we define $\mathrm{T} : \mathcal{A} \boxtimes \mathcal{B} \to \mathcal{B} \boxtimes \mathcal{A}$ to be the equivalence given by $X \boxtimes Y \mapsto \Pi^{|X||Y|}(Y) \boxtimes X$, where X and Y are homogeneous objects of $\mathcal{A}$ and $\mathcal{B}$. One should think of the equivalence T as analogous to the isomorphism τ from the previous paragraph. We show that T (together with a choice of factor system) defines a symmetry on $\mathscr{C}$ in Section 4. One can then regard supersymmetric monoidal supercategories as the *symmetric* monoidal objects in the symmetric monoidal 2-category $(\mathscr{C}, \mathrm{T})$.

1.3. ***Spin-symmetric groups.*** The most fundamental example of a symmetric monoidal category is the groupoid of finite sets under disjoint union. A skeletal version $\mathcal{S}$ of this category can be constructed explicitly, as follows. There is one object $[n]$ for each natural number n. We put $\mathrm{End}_{\mathcal{S}}([n]) = \mathfrak{S}_n$ and $\mathrm{Hom}_{\mathcal{S}}([n], [m]) = \varnothing$ for $n \neq m$. The monoidal operation $\sqcup$ is defined on objects by $[n] \sqcup [m] = [n+m]$ and on morphisms using the natural homomorphism $\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{n+m}$. Finally, the symmetry isomorphism $[n] \sqcup [m] \to [m] \sqcup [n]$ is given by the element $\tau_{n,m} \in \mathfrak{S}_{n+m}$ defined by

$$\tau_{n,m}(i) = \begin{cases} i+m & \text{if } 1 \le i \le n, \\ i-n & \text{if } n+1 \le i \le n+m. \end{cases}$$

The hexagon axioms amount to some identities satisfied by the $\tau_{n,m}$ elements.

The above picture has an analog in the setting of supersymmetric monoidal categories in which the symmetric groups are replaced by spin-symmetric groups. Recall that the n-th spin-symmetric group $\widetilde{\mathfrak{S}}_n$ is a certain central extension of $\mathfrak{S}_n$ by $\mathbb{Z}/2$; see Section 5.2. We define a skeletal category $\tilde{\mathcal{S}}$ as follows. There is again one object $[n]$ for each natural number n. We put $\mathrm{End}_{\tilde{\mathcal{S}}}([n]) = \boldsymbol{k}[\widetilde{\mathfrak{S}}_n]/(c+1)$, where $c \in \widetilde{\mathfrak{S}}_n$ is the distinguished central element, and define $\mathrm{Hom}_{\tilde{\mathcal{S}}}([n],[m]) = 0$ if $n \neq m$. The monoidal operation $\sqcup$ on $\tilde{\mathcal{S}}$ is defined on objects by $[n] \sqcup [m] = [n+m]$, and on morphisms using the natural map $\widetilde{\mathfrak{S}}_n \times \widetilde{\mathfrak{S}}_m \to \widetilde{\mathfrak{S}}_{n+m}$. Finally, the supersymmetry is induced by a particular choice of lift $\tilde{\tau}_{n,m} \in \widetilde{\mathfrak{S}}_{n+m}$ of the element $\tau_{n,m}$ defined above. In Section 5.4, we verify that the axioms of a supersymmetric monoidal category hold in this case; this amounts to establishing some identities satisfied by the $\tilde{\tau}_{n,m}$ elements.

The definition and study of the elements $\tilde{\tau}_{n,m}$ (in Section 5.3), while totally elementary, is one of the most important parts of this paper. For a particular choice of n and m, there are two lifts of $\tau_{n,m}$ to $\widetilde{\mathfrak{S}}_{n+m}$, and neither is preferred. We find, however, that the lifts $\tilde{\tau}_{n,m}$ we define have important compatibility properties as n and m vary; in other words, there is a preferred *system* of lifts (up to the action of a single $\mathbb{Z}/2$). This strikes us as an important discovery. In fact, as discussed in Section 1.6 below, we show that essentially every supersymmetry can, in a sense, be constructed from these elements.

The category $\mathcal{S}$ is not merely the simplest example of a symmetric monoidal category: it is also universal. Indeed, suppose that $\mathcal{C}$ is a symmetric monoidal category. Given an object X of $\mathcal{C}$, the symmetric group $\mathfrak{S}_n$ naturally acts on $X^{\otimes n}$. This construction induces a symmetric monoidal functor $\mathcal{S} \to \mathcal{C}$ via $[n] \mapsto X^{\otimes n}$. This observation shows that $(\mathcal{S}, [1])$ is the universal symmetric monoidal category equipped with an object, in the sense that giving a symmetric monoidal functor $\mathcal{S} \to \mathcal{C}$ is equivalent to giving an object of $\mathcal{C}$.

An analogous result holds for $\tilde{\mathcal{S}}$. Indeed, if $\mathcal{C}$ is a supersymmetric monoidal $\mathbb{Z}$-category and X is a degree 1 object in $\mathcal{C}$, then we show that the spin-symmetric group $\widetilde{\mathfrak{S}}_n$ naturally acts on $X^{\otimes n}$. This leads to the following theorem:

Theorem 1.1. $(\tilde{\mathcal{S}}, [1])$ *is the universal supersymmetric monoidal $\mathbb{Z}$-supercategory equipped with an object of degree* 1.

Another important example of a supersymmetric category comes from the representation theory of spin-symmetric groups. A *linear spin-species* is a sequence $(V_n)_{n\geq 0}$ where V_n is a representation of $\widetilde{\mathfrak{S}}_n$ (that interacts with superstructures appropriately). The category of such sequences, denoted by $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$, admits a monoidal structure defined by inducing from Young subgroups. The $\tilde{\tau}_{n,m}$ elements can be used to construct a supersymmetry on this category. The details are carried out in Section 6.

1.4. ***The half tensor product.*** The second most fundamental example of a symmetric monoidal category is the category **Vec** of vector spaces under tensor product. The category **SVec** of super-vector spaces is *not* supersymmetric: it is a "plain" symmetric monoidal supercategory since the isomorphism $V \otimes W \to W \otimes V$ has even degree. (The same applies to any category where the tensor product and symmetry are inherited from **SVec**, such as the category of representations of a Lie superalgebra.) However, there is a supersymmetric analog to **Vec**, which we now discuss.

Let V be a super-vector space. An *isomeric structure* on V is an odd-degree automorphism $\nu : V \to V$ that squares to the identity. Suppose that U and V are super-vector spaces equipped with isomeric structures μ and ν. Then $\mu \otimes \nu$ is an even-degree automorphism of $U \otimes V$ that squares to -1. The *half tensor product* of U and V, denoted by $2^{-1}(U \otimes V)$, is defined to be the ζ_4-eigenspace of $\mu \otimes \nu$. (Here $\zeta_4 = \sqrt{-1}$.) The choice of ζ_4 does not matter much, as $\mu \otimes 1$ (or $1 \otimes \nu$) interchanges the $+\zeta_4$ and $-\zeta_4$ eigenspaces.

This construction appears frequently in the representation theory of superalgebras. Suppose that A is a $\mathbb{C}$-superalgebra and M is a finite-dimensional simple A-module. The super version of Schur's lemma states that $\mathrm{End}_A(M)$ is either 1-dimensional ("type M") or 2-dimensional ("type Q"); in the latter case, the algebra $\mathrm{End}_A(M)$ is generated over $\mathbb{C}$ by a isomeric structure. Now suppose that B is a second $\mathbb{C}$-superalgebra and N is a finite-dimensional simple B-module. If either M or N has type M, then $M \otimes N$ is a simple $(A \otimes B)$-module. If both M and N have type Q, then $2^{-1}(M \otimes N)$ is a simple module, and $M \otimes N \cong 2^{-1}(M \otimes N) \otimes \mathbb{C}^{1|1}$. (See [Cheng and Wang 2012, Section 3.1.3] for details.)

Returning to the general situation, let (U, μ) and (V, ν) be as above. A simple computation shows that the natural symmetry $\tau_{U,V} : U \otimes V \to V \otimes U$ carries the ζ_4-eigenspace of $\mu \otimes \nu$ into the $(-\zeta_4)$-eigenspace of $\nu \otimes \mu$. In particular, $\tau_{U,V}$ does *not* map $2^{-1}(U \otimes V)$ into $2^{-1}(V \otimes U)$. However, the composition $(\nu \otimes 1) \circ \tau_{U,V}$ does map $2^{-1}(U \otimes V)$ to $2^{-1}(V \otimes U)$. This is an odd-degree isomorphism, which suggests a lurking supersymmetry.

In Section 7, we define a $\mathbb{Z}/2$-supercategory **Ismr** by taking the degree 0 objects to be super-vector spaces and the degree 1 objects to be super-vector spaces equipped with an isomeric structure. We define a monoidal operation $\odot$ on this category by using the half tensor product on two degree 1 objects and the ordinary tensor product on other pairs of objects. We show that this does indeed define a monoidal structure; this is not obvious, as the associator maps are nontrivial. We further show that the construction in the previous paragraph (scaled by an appropriate root of unity) defines a supersymmetry on this monoidal structure. Thus, **Ismr** is an example of a supersymmetric monoidal $\mathbb{Z}/2$-supercategory. We believe this places the half tensor product construction in its proper theoretical framework.

1.5. ***Exterior algebras.*** Let $\mathcal{A}$ be a $\mathbb{C}$-linear category. One can then define the n-th symmetric power of $\mathcal{A}$, denoted by $\mathrm{Sym}^n(\mathcal{A})$, to be the category of $\mathfrak{S}_n$-equivariant objects in the n-th power of $\mathcal{A}$ (for an appropriate notion of power). The sum of these categories, denoted by $\mathrm{Sym}(\mathcal{A})$, is naturally a symmetric monoidal category, and is the universal such category to which $\mathcal{A}$ maps.

Now suppose that $\mathcal{A}$ is a supercategory. Ganter and Kapranov [2014, Section 3.6] defined the n-th exterior power of $\mathcal{A}$, denoted by $\bigwedge^n(\mathcal{A})$, to be the category of appropriately $\widetilde{\mathfrak{S}}_n$-equivariant objects in the n-th power of $\mathcal{A}$. We show that the sum of these categories, denoted by $\bigwedge(\mathcal{A})$, is naturally a supersymmetric monoidal supercategory, and has a universal property similar to the symmetric power. This construction provides a large source of supersymmetric categories. See Section 6.8 for details.

1.6. ***Clifford eversion.*** Let $\mathcal{A}$ be a nice $\mathbb{Z}/q$-supercategory, with q even; "nice" means that $\mathcal{A}$ is additive, Karoubian and admits a Π structure. In Section 8, we define a new nice $\mathbb{Z}/q$-supercategory $\mathrm{Cl}(\mathcal{A})$, called

the *Clifford eversion* of $\mathcal{A}$, by considering certain kinds of Clifford modules in $\mathcal{A}$. (This construction is essentially a generalization of the category **Ismr** discussed above; see Section 7.5.) If $\mathcal{A}$ has a symmetric monoidal structure, then we show that $\mathrm{Cl}(\mathcal{A})$ naturally has a supersymmetric monoidal structure; this construction again hinges on the $\tilde{\tau}_{n,m}$ elements. Moreover, we sketch a proof for:

Theorem 1.2. *Clifford eversion defines an equivalence between the* 2*-category of nice symmetric monoidal* $\mathbb{Z}/q$*-supercategories and the* 2*-category of nice supersymmetric monoidal* $\mathbb{Z}/q$*-supercategories.*

Let $\mathscr{C}$ be the 2-category of nice $\mathbb{Z}/q$-supercategories. This category has a natural monoidal structure that admits an obvious symmetry Σ; it also admits the less obvious symmetry T discussed in Section 1.2. In fact, Clifford eversion induces an equivalence of symmetric monoidal 2-categories $(\mathscr{C}, \Sigma) \to (\mathscr{C}, \mathrm{T})$. This recovers the above theorem, but is a more fundamental statement. However, we only partially prove this statement since the notion of equivalence of symmetric monoidal 2-categories is rather complicated.

Let $\mathcal{A}$ be nice symmetric monoidal $\mathbb{Z}/q$-supercategory with an exact structure, and let $\mathrm{K}_+(\mathcal{A})$ be the quotient of the Grothendieck group by the relations $[\Pi(X)] = [X]$. We show (Proposition 8.21) that there is a natural isomorphism of rings

$$\mathrm{K}_+(\mathcal{A})\left[\tfrac{1}{\sqrt{2}}\right] \cong \mathrm{K}_+(\mathrm{Cl}(\mathcal{A}))\left[\tfrac{1}{\sqrt{2}}\right].$$

This helps explain the appearance of certain factors of $\sqrt{2}$ in the theory of Q-symmetric functions.

1.7. ***Sergeev–Schur duality.*** (We take the base field to be $\mathbb{C}$ in this section.) Classical Schur–Weyl duality connects the representation theory of the general linear group and the symmetric groups. Let $\mathrm{Rep}^{\mathrm{pol}}(\mathbf{GL})$ be the category of polynomial representations of the infinite general linear group; this is the tensor category generated by the standard representation $\boldsymbol{V}$. The symmetric group $\mathfrak{S}_n$ acts on $\boldsymbol{V}^{\otimes n}$ by permuting tensor factors, and so we obtain a functor $\mathrm{Rep}(\mathfrak{S}_n) \to \mathrm{Rep}^{\mathrm{pol}}(\mathbf{GL})$ by $M \mapsto M \otimes_{\mathbb{C}[\mathfrak{S}_n]} \boldsymbol{V}^{\otimes n}$. Let $\mathrm{Rep}(\mathfrak{S}_*)$ be the category of sequences $(V_n)_{n\geq 0}$ where V_n is a representation of the symmetric group $\mathfrak{S}_n$; such sequences are sometimes called *linear species*. Then one formulation of Schur–Weyl duality states that the above functor extends to an equivalence $\mathrm{Rep}(\mathfrak{S}_*) \to \mathrm{Rep}^{\mathrm{pol}}(\mathbf{GL})$. This equivalence respects various additional structures on the two categories. In particular, it is an equivalence of symmetric monoidal categories, where $\mathrm{Rep}^{\mathrm{pol}}(\mathbf{GL})$ is endowed with the usual tensor product and $\mathrm{Rep}(\mathfrak{S}_*)$ is given the induction tensor product, defined by

$$(V \otimes W)_n = \bigoplus_{i+j=n} \mathrm{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_n}(V_i \otimes W_j).$$

This equivalence is extremely useful when studying algebra within the tensor category $\mathrm{Rep}(\mathfrak{S}_*)$, as it provides a bridge to more classical theory.

Sergeev–Schur duality is an analog of Schur–Weyl duality in which the general linear group is replaced by the isomeric Lie superalgebra. Precisely, let $\boldsymbol{V} = \mathbb{C}^{\infty|\infty}$ and fix an odd-degree isomorphism $\alpha : \boldsymbol{V} \to \boldsymbol{V}$ squaring to 2. The infinite isomeric Lie superalgebra $\mathfrak{q}$ is defined to be the subalgebra of $\mathfrak{gl}(\boldsymbol{V})$ that preserves α. The symmetric group $\mathfrak{S}_n$ acts on $\boldsymbol{V}^{\otimes n}$ by permuting tensor factors. However, there are additional natural endomorphisms that commute with $\mathfrak{q}$; namely, we can apply α to any tensor factor. To

make this more precise, let Cl_n denote the n-th Clifford algebra; this has n anticommuting generators that each square to 2. Define the Hecke–Clifford algebra $\mathcal{H}_n$ to be the semidirect product $\mathfrak{S}_n \ltimes \mathrm{Cl}_n$. Then $\mathcal{H}_n$ acts on $\boldsymbol{V}^{\otimes n}$ and commutes with $\mathfrak{q}$. As in the previous paragraph, we obtain an induced functor $\mathrm{Rep}(\mathcal{H}_*) \to \mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$. One formulation of Sergeev–Schur duality is that this is an equivalence (and respects the natural symmetric monoidal structures).

There is a close relationship between the algebra $\mathcal{H}_n$ and the spin-symmetric groups: we have an isomorphism $\mathcal{H}_n \cong \boldsymbol{k}[\widetilde{\mathfrak{S}}_n]/(c+1) \otimes \mathrm{Cl}_n$. Thus, an $\mathcal{H}_n$-module can be viewed as a Clifford module in the category $\mathrm{Rep}(\widetilde{\mathfrak{S}}_n)$ of linear spin species. We thus see that $\mathrm{Rep}(\mathcal{H}_*)$ can be identified with the Clifford eversion of the category $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ of sequences of $\widetilde{\mathfrak{S}}_n$-representations. We can thus rephrase the above discussion as follows:

Theorem 1.3. *The category* $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$ *is equivalent to the Clifford eversion of the category* $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ *of linear spin species.*

We show that this equivalence respects natural symmetric monoidal structures on the two categories. We view the above theorem as an elegant expression of Sergeev–Schur duality, as it provides a direct link between the isomeric algebra and the spin-symmetric groups. See Section 9 for details.

The Grothendieck ring of $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$ is naturally identified with the ring Γ of odd symmetric functions. This ring has a basis consisting of the Schur Q-functions Q_λ. However, they do not represent classes of simple objects in general: one must adjust them by appropriate powers of $\sqrt{2}$ (which depend on λ as well as which category is being used). One of our goals was to better understand the nature of these powers of $\sqrt{2}$, and the effect of Clifford eversion on Grothendieck rings is one piece of that puzzle. Unfortunately, there is one more piece which is not explained by our general framework: some simple modules of $\boldsymbol{k}[\widetilde{\mathfrak{S}}_n]/(c+1)$ carry an action of Cl_1, while others do not, and this also factors into the powers of $\sqrt{2}$ that appear.

1.8. ***Relation to previous work.*** Kapranov [2018, Section 3.4] proposed a sketch for the definition of a supersymmetric monoidal supercategory. A large portion of this article is devoted to providing the details for this sketch.

Versions of (monoidal) supercategories have been considered in [Brundan and Ellis 2017; Kang et al. 2016; Ellis and Lauda 2016]. However, these sources do not treat the notion of (super)braiding or (super)symmetry for a monoidal supercategory, which, as far as we know, is treated in detail for the first time here. The Clifford twist in [Kang et al. 2016, Section 2.1] is similar to our notion of Clifford eversion, but is not the same (for instance, the Clifford twist is an involution even if the coefficient field does not contain $\sqrt{-1}$).

1.9. ***Open problems.*** We mention a few open problems:

(a) In Remark 3.16, we note a connection between even factor systems and degree 3 group cohomology. However, this connection is not complete, and it does not apply to odd factor systems (which is the relevant case for supersymmetries). It would be helpful to have a better understanding of these issues.

(b) Is there a conceptual a priori explanation for why the 2-categories of symmetric and supersymmetric monoidal supercategories are equivalent?

(c) Is there a "natural" construction of the groupoid of spin-sets? That is, is there a natural category where the objects are some kind of "enhanced" finite sets whose automorphism groups are spin-symmetric groups? This might clarify a number of our constructions (such as the $\tilde{\tau}_{n,m}$ elements). Is there a nongroupoid version?

(d) We have provided explanations for most of the proof of Theorem 8.16, but it remains incomplete.

1.10. ***Notation.*** We fix a field $\boldsymbol{k}$ throughout, which we assume does not have characteristic 2. One could allow $\boldsymbol{k}$ to be a commutative $\mathbb{Z}\left[\frac{1}{2}\right]$-algebra, but we stick to fields for simplicity. We fix a compatible system of primitive 2-power roots of unity ζ_{2^k} in $\bar{\boldsymbol{k}}$; by compatible, we mean that $\zeta_{2^{k+1}}^2 = \zeta_{2^k}$.

In this article, "superalgebra" means an algebra with a $\mathbb{Z}/2$-grading and a "commutative superalgebra" refers to a commutative algebra with a $\mathbb{Z}/2$-grading (which does not affect the definition of commutative). This is sometimes used in the literature to mean a skew-commutative algebra; we reserve the phrase "supercommutative superalgebra" for this, i.e., an algebra A with a $\mathbb{Z}/2$-grading such that $xy = (-1)^{|x||y|}yx$ for homogeneous elements $x, y \in A$.

2. Supercategories

2.1. ***Super-vector spaces.*** A *super-vector space* over $\boldsymbol{k}$ is a $\mathbb{Z}/2$-graded $\boldsymbol{k}$-vector space. Given a super-vector space V, we write V_i for the subspace of homogeneous elements of degree i. For $v \in V$, we write $|v| = i$ to indicate that v is a homogeneous element of degree i. For $n \in \mathbb{Z}$, we let $V[n]$ be the super-vector space with $V[n]_i = V_{n+i}$.

Let V and W be two super-vector spaces and let $f : V \to W$ be a linear map. We say that f is *homogeneous* of parity n if $f(V_i) \subseteq W_{i+n}$. We write $\mathrm{Hom}(V, W)$ for the set of all linear maps, and $\mathrm{Hom}(V, W)_i$ for those of parity i. Every linear map decomposes uniquely as a sum of even and odd maps, and so we have a decomposition

$$\mathrm{Hom}(V, W) = \mathrm{Hom}(V, W)_0 \oplus \mathrm{Hom}(V, W)_1.$$

In this way, we regard $\mathrm{Hom}(V, W)$ itself as a super-vector space. We let **SVec** be the category of super-vector spaces and all linear maps, and $\mathbf{SVec}^\circ$ the category of super-vector spaces and even linear maps.

The tensor product of two super-vector spaces is just the tensor product of the underlying vector spaces, graded in the usual manner. Given homogeneous linear maps $f_1 : V_1 \to W_1$ and $f_2 : V_2 \to W_2$, we define $f_1 \otimes f_2 : V_1 \otimes V_2 \to W_1 \otimes W_2$ to be the linear map given on homogeneous elements by the formula

$$(f_1 \otimes f_2)(v_1 \otimes v_2) = (-1)^{|v_1||f_2|} f_1(v_1) \otimes f_2(v_2).$$

We caution that $\otimes$ is *not* a functor $\mathbf{SVec} \times \mathbf{SVec} \to \mathbf{SVec}$; this will be explained in detail in Section 3.1. However, $\otimes$ is a functor $\mathbf{SVec}^\circ \times \mathbf{SVec}^\circ \to \mathbf{SVec}^\circ$, and naturally endows $\mathbf{SVec}^\circ$ with the structure of a monoidal category.

Let V and W be super-vector spaces. We define isomorphisms $\sigma, \tau : V \otimes W \to W \otimes V$ for homogeneous elements $v \in V$ and $w \in W$ by

$$\sigma(v \otimes w) = w \otimes v, \quad \tau(v \otimes w) = (-1)^{|v||w|} w \otimes v.$$

Both σ and τ define symmetric structures on $\otimes$. However, the symmetric structure σ is uninteresting in that it does not interact with the grading. In what follows, we regard $\mathbf{SVec}^\circ$ as a symmetric monoidal category using τ, unless explicitly said otherwise.

2.2. ***Supercategories.*** A *supercategory* is a category $\mathcal{A}$ enriched in the category $\mathbf{SVec}^\circ$ of super-vector spaces. Thus, $\mathcal{A}$ is a category and $\mathrm{Hom}_{\mathcal{A}}(X, Y)$ has the structure of a super-vector space for all objects X and Y such that composition defines an even map

$$\mathrm{Hom}_{\mathcal{A}}(Y, Z) \otimes \mathrm{Hom}_{\mathcal{A}}(X, Y) \to \mathrm{Hom}_{\mathcal{A}}(X, Z), \quad f \otimes g \mapsto f \circ g.$$

Given a supercategory $\mathcal{A}$, we let $\mathcal{A}^\circ$ be the category with the same collection of objects but only the even morphisms; that is, $\mathrm{Hom}_{\mathcal{A}^\circ}(X, Y) = \mathrm{Hom}_{\mathcal{A}}(X, Y)_0$.

Let Λ be an abelian group (or, for now, any set). A *homogeneous* Λ*-supercategory* is a collection $\mathcal{A} = \{\mathcal{A}_i\}_{i \in \Lambda}$, where each $\mathcal{A}_i$ is a supercategory. We refer to $\mathcal{A}_i$ as the category of homogeneous objects of degree i. Informally, an inhomogeneous Λ-supercategory is a supercategory $\mathcal{A}$ equipped with subcategories $\mathcal{A}_i$ for $i \in \mathcal{A}$ such that $\mathcal{A}$ is a direct sum of the $\mathcal{A}_i$'s, in some sense. This concept will appear in some examples, but we do not give a rigorous definition to avoid details of what direct sum means here (there are some subtleties if, e.g., $\mathcal{A}$ is nonadditive or Λ is infinite). In the inhomogeneous case, any construction will be determined by what it does on homogeneous objects, so we can typically restrict our attention to them. In what follows, "Λ-supercategory" will by default refer to the homogeneous version, and we almost always take Λ to be $\mathbb{Z}$ or $\mathbb{Z}/q$.

Suppose that $\mathcal{A}$ is a Λ-supercategory and we have a function $f : \Lambda \to \Lambda'$. Then we can convert $\mathcal{A}$ into a Λ'-supercategory by defining $\mathcal{A}_i = \coprod_{j \in f^{-1}(i)} \mathcal{A}_j$ for $i \in \Lambda'$. Here the coproduct is taken in the sense of linear categories: the Hom space between objects in different factors is 0. In particular, we can convert a $\mathbb{Z}$-supercategory into a $\mathbb{Z}/q$-supercategory, or a $\mathbb{Z}/q$-supercategory into a $\mathbb{Z}/q'$-supercategory if q' divides q.

Example 2.1. Let G be a supergroup over $\mathbb{C}$ containing a central $\boldsymbol{G}_m$. Then $\mathcal{A} = \mathrm{Rep}(G)$ is an (inhomogeneous) $\mathbb{Z}$-supercategory, by taking $\mathcal{A}_i$ to be the subcategory of objects where $\boldsymbol{G}_m$ acts with weight i. □

2.3. ***Superfunctors.*** Let $\mathcal{A}$ and $\mathcal{B}$ be supercategories. A *superfunctor* $F : \mathcal{A} \to \mathcal{B}$ is a functor (in the usual sense) such that the maps

$$F : \mathrm{Hom}_{\mathcal{A}}(X, Y) \to \mathrm{Hom}_{\mathcal{B}}(F(X), F(Y))$$

are even maps of super-vector spaces. If $\mathcal{A}$ and $\mathcal{B}$ are Λ-supercategories, then a Λ*-superfunctor* is a collection of superfunctors $F = \{F_i : \mathcal{A}_i \to \mathcal{B}_i\}_{i \in \Lambda}$.

Let F and G be two superfunctors from $\mathcal{A}$ to $\mathcal{B}$. A *homogeneous natural transformation* $\eta : F \to G$ of parity $p \in \mathbb{Z}/2$ is a rule that assigns to every object X of $\mathcal{A}$ a homogeneous morphism $\eta_X : F(X) \to G(X)$ of degree p in $\mathcal{B}$ such that, if $f : X \to Y$ is a homogeneous morphism in $\mathcal{A}$, then $G(f) \circ \eta_X = (-1)^{p|f|}\eta_Y \circ F(f)$. A general natural transformation is a sum of an even and an odd natural transformation. The definition is the same for Λ-superfunctors: the grading does not affect the definition at all.

2.4. Π-*structures*. Let $\mathcal{A}$ be a supercategory. A Π-*structure* on an object X of $\mathcal{A}$ is an odd-degree isomorphism $\xi_X : X \to \Pi(X)$ for some object $\Pi(X)$. A Π-*structure* on $\mathcal{A}$ is a choice of Π-structure on each object of $\mathcal{A}$. Such a structure need not exist. Note that, if $\mathcal{A}$ is Λ-graded and X is homogeneous, then $\Pi(X)$ is necessarily homogeneous of the same degree.

Suppose we have a Π-structure on $\mathcal{A}$. For a homogeneous morphism $f : X \to Y$ in $\mathcal{A}$, define $\Pi(f)$ to be the unique morphism making the following diagram commute up to $(-1)^{|f|}$:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \xi_X}\downarrow & & \downarrow{\scriptstyle \xi_Y} \\ \Pi(X) & \xrightarrow{\Pi(f)} & \Pi(Y) \end{array}$$

By uniqueness, it is clear that Π is a superfunctor $\mathcal{A} \to \mathcal{A}$, and that $\xi : \mathrm{id}_{\mathcal{A}} \to \Pi$ is an odd natural isomorphism. (If $\mathcal{A}$ is a Λ-supercategory, then Π is a Λ-superfunctor.) Furthermore, taking $f = \xi_X$, we find $\Pi(\xi_X) = -\xi_{\Pi(X)}$.

For a nonnegative integer n, we let Π^n be the n-fold iterate of Π. For $n < m$, we define $\xi_X^{n,m} : \Pi^n(X) \to \Pi^m(X)$ to be the composition $\xi_{\Pi^{m-1}(X)} \circ \cdots \circ \xi_{\Pi^{n+1}(X)} \circ \xi_{\Pi^n(X)}$; for $n > m$, we define $\xi_X^{n,m} = (\xi_X^{m,n})^{-1}$; and for $n = m$, we let $\xi_X^{n,m}$ be the identity. The map $\xi_X^{n,m}$ is an isomorphism of parity $m - n$. In particular, $\xi^{0,2} : \mathrm{id}_{\mathcal{A}} \to \Pi^2$ is an even-degree isomorphism. For q even, we define Π^n for $n \in \mathbb{Z}/q$ to be the colimit of the system

$$\cdots \to \Pi^{n-q} \xrightarrow{\xi^{n-q,n}} \Pi^n \xrightarrow{\xi^{n+q,n}} \Pi^{n+q} \to \cdots.$$

This colimit exists since all the maps are even isomorphisms.

Suppose that (Π, ξ) and (Π', ξ') are two Π-structures on $\mathcal{A}$. Define $i_X : \Pi(X) \to \Pi'(X)$ to be the morphism $\xi'_X \circ \xi_X^{-1}$. Then i_X is an even-degree isomorphism, and defines an even-degree isomorphism of functors $i : \Pi \to \Pi'$ compatible with ξ and ξ'. We thus see that any two Π-structures are canonically (and, in fact, uniquely) isomorphic by an even-degree natural transformation.

Proposition 2.2. *Let $F : \mathcal{A} \to \mathcal{B}$ be a superfunctor between supercategories with Π-structures. Then there is a canonical even isomorphism $F \circ \Pi = \Pi \circ F$.*

Proof. Let X be an object of $\mathcal{A}$. Then $\xi_{F(X)} F(\xi_X)^{-1}$ gives an even-degree isomorphism $F(\Pi(X)) \cong \Pi(F(X))$. We now check that this is a natural isomorphism. Let $f : X \to Y$ be a homogeneous morphism. Then

$$\xi_Y f = (-1)^{|f|}\Pi(f)\xi_X, \quad \xi_{F(Y)}F(f) = (-1)^{|f|}\Pi(F(f))\xi_{F(X)}.$$

We apply F to the first equation and rearrange to get $F(f)F(\xi_X)^{-1} = (-1)^{|f|}F(\xi_Y)^{-1}F(\Pi(f))$. Post-compose this equation with $\xi_{F(Y)}$ and use the second equation to get

$$\xi_{F(Y)}F(\xi_Y)^{-1}F(\Pi(f)) = \Pi(F(f))\xi_{F(X)}F(\xi_X)^{-1},$$

which shows that we have defined a natural isomorphism $F \circ \Pi \cong \Pi \circ F$. □

2.5. ***Product categories.*** Let $\mathcal{A}$ and $\mathcal{B}$ be two supercategories. We let $\mathcal{A} \check{\boxtimes} \mathcal{B}$ be the product category, in the sense of enriched categories. Precisely, the objects of $\mathcal{A} \check{\boxtimes} \mathcal{B}$ are pairs (A, B) with $A \in \mathcal{A}$ and $B \in \mathcal{B}$; we write $A \boxtimes B$ in place of (A, B). The Hom spaces are defined by

$$\mathrm{Hom}_{\mathcal{A}\check{\boxtimes}\mathcal{B}}(A \boxtimes B, A' \boxtimes B') = \mathrm{Hom}_{\mathcal{A}}(A, A') \otimes \mathrm{Hom}_{\mathcal{B}}(B, B').$$

The composition of homogeneous morphisms $f \otimes g : A \boxtimes B \to A' \boxtimes B'$ and $f' \otimes g' : A' \boxtimes B' \to A'' \boxtimes B''$ is defined by

$$(f' \otimes g') \circ (f \otimes g) = (-1)^{|f||g'|}(f' \circ f) \otimes (g' \circ g).$$

The sign here comes from the symmetry of the tensor product in **SVec**. (To define the product of enriched categories in general, the enriching category needs a symmetric monoidal structure, and here we are using the symmetry τ on $\mathbf{SVec}^\circ$.)

Let $\mathcal{A}$ and $\mathcal{B}$ be two homogeneous Λ-supercategories. We define $\mathcal{A} \boxtimes \mathcal{B}$ to be the homogeneous Λ-supercategory with

$$(\mathcal{A} \boxtimes \mathcal{B})_i = \coprod_{j+k=i} \mathcal{A}_j \check{\boxtimes} \mathcal{B}_k.$$

We note that $\check{\boxtimes}$ does not distribute over direct sums of supercategories (even finite direct sums), and so does not give the expected result on inhomogeneous supercategories; this is why we have used a different symbol in the Λ-graded case.

If either $\mathcal{A}$ or $\mathcal{B}$ admits a Π-structure, then so does $\mathcal{A} \boxtimes \mathcal{B}$. Indeed, suppose $\mathcal{A}$ does, and let $A \boxtimes B$ be a given object of $\mathcal{A} \boxtimes \mathcal{B}$. If $\xi : A \to \Pi(A)$ is an odd-degree isomorphism, then $\xi \otimes \mathrm{id}_B : A \boxtimes B \to \Pi(A) \boxtimes B$ is an odd-degree isomorphism, and so the claim is verified.

2.6. ***Grothendieck groups.*** Let $\mathcal{A}$ be a supercategory equipped with an exact structure on $\mathcal{A}^\circ$. We can then consider the Grothendieck group $\mathrm{K}(\mathcal{A}^\circ)$ of $\mathcal{A}^\circ$. For $\varepsilon \in \{+, -\}$, we define $\mathrm{K}_\varepsilon(\mathcal{A})$ to be the quotient of $\mathrm{K}(\mathcal{A}^\circ)$ by the relations $[X] = \varepsilon[Y]$ whenever there is an odd-degree isomorphism $X \to Y$. If $\mathcal{A}$ admits a Π-structure, then $\mathrm{K}_\varepsilon(\mathcal{A})$ is simply the quotient of $\mathrm{K}(\mathcal{A}^\circ)$ by the relations $[\Pi(X)] = \varepsilon[X]$. If $\mathcal{A}$ admits a Λ-grading, then $\mathrm{K}(\mathcal{A}^\circ) = \bigoplus_{i \in \Lambda} \mathrm{K}(\mathcal{A}_i^\circ)$, and so $\mathrm{K}(\mathcal{A}^\circ)$ is canonically Λ-graded. This induces a Λ-grading on $\mathrm{K}_\varepsilon(\mathcal{A})$.

3. Monoidal structures

3.1. ***Monoidal structures.*** Let $\mathcal{A}$ be a Λ-supercategory. A *monoidal structure* on $\mathcal{A}$ consists of

- a Λ-superfunctor $\otimes : \mathcal{A} \boxtimes \mathcal{A} \to \mathcal{A}$;

- a unit object $\mathbb{1}$;
- an even natural isomorphism $\alpha_{X,Y,Z} : X \otimes (Y \otimes Z) \to (X \otimes Y) \otimes Z$ satisfying the pentagon axiom, i.e., the diagram

$$\begin{array}{ccc} & (A\otimes B)\otimes(C\otimes D) & \\ {\scriptstyle \alpha_{A,B,C\otimes D}}\nearrow & & \searrow{\scriptstyle \alpha_{A\otimes B,C,D}} \\ A\otimes(B\otimes(C\otimes D)) & & ((A\otimes B)\otimes C)\otimes D \\ \downarrow{\scriptstyle \mathrm{id}_A\otimes\alpha_{B,C,D}} & & \uparrow{\scriptstyle \alpha_{A,B,C}\otimes\mathrm{id}_D} \\ A\otimes((B\otimes C)\otimes D) & \xrightarrow{\alpha_{A,B\otimes C,D}} & (A\otimes(B\otimes C))\otimes D \end{array}$$

  is commutative for all objects A, B, C, D;
- even natural isomorphisms $\lambda_X : \mathbb{1} \otimes X \to X$ and $\rho_X : X \otimes \mathbb{1} \to X$ satisfying the triangle axiom, i.e., the diagram

$$\begin{array}{ccc} A\otimes(\mathbb{1}\otimes B) & \xrightarrow{\alpha_{A,\mathbb{1},B}} & (A\otimes\mathbb{1})\otimes B \\ & {\scriptstyle \mathrm{id}_A\otimes\lambda_B}\searrow \quad \swarrow{\scriptstyle \rho_A\otimes\mathrm{id}_B} & \\ & A\otimes B & \end{array}$$

  is commutative for all objects A, B.

We note that, if A and B are homogeneous objects of $\mathcal{A}$, then $A \otimes B$ is homogeneous and $|A \otimes B| = |A| + |B|$. A *monoidal* Λ*-supercategory* is a Λ-supercategory equipped with a monoidal structure.

Example 3.1. The supercategory **SVec** is monoidal under tensor product. (Note here that we are using all of **SVec**, and not just the subcategory **SVec**$^\circ$.) □

Remark 3.2. Suppose that $\mathcal{A}$ is a monoidal supercategory and we have homogeneous morphisms

$$f_1 : A_1 \to A_2, \quad f_2 : A_2 \to A_3, \quad f_1' : A_1' \to A_2', \quad f_2' : A_2' \to A_3'.$$

Then the diagram

$$\begin{array}{ccc} A_1\otimes A_1' & \xrightarrow{f_2f_1\otimes f_2'f_1'} & A_3\otimes A_3' \\ & {\scriptstyle f_1\otimes f_1'}\searrow \quad \nearrow{\scriptstyle f_2\otimes f_2'} & \\ & A_2\otimes A_2' & \end{array}$$

commutes up to $(-1)^{|f_1||f_2'|}$, that is, we have the identity

$$(f_2 \otimes f_2') \circ (f_1 \otimes f_1') = (-1)^{|f_1||f_2'|}(f_2 f_1 \otimes f_2' f_1').$$

This follows from the definition of the composition law in the category $\mathcal{A} \boxtimes \mathcal{A}$. From this, we see that $\otimes$ does *not* define a functor $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ (in general). However, it does define a functor $\mathcal{A}^\circ \times \mathcal{A}^\circ \to \mathcal{A}^\circ$, and $\mathcal{A}^\circ$ is naturally a monoidal category. □

Let $\mathcal{A}$ be a monoidal Λ-supercategory admitting a Π-structure. Let A and B be objects of $\mathcal{A}$. Then we have odd-degree isomorphisms $\xi_{A\otimes B} : A \otimes B \to \Pi(A \otimes B)$ and $\xi_A \otimes \mathrm{id}_B : A \otimes B \to \Pi(A) \otimes B$. We thus have a canonical even-degree isomorphism $\Pi(A \otimes B) \to \Pi(A) \otimes B$. Similarly, we have a canonical even-degree isomorphism $\Pi(A \otimes B) \to A \otimes \Pi(B)$. We thus have canonical even-degree isomorphisms $\Pi(A) \cong \Pi(\mathbb{1}) \otimes A \cong A \otimes \Pi(\mathbb{1})$.

Definition 3.3. Let $\mathcal{A}$ and $\mathcal{B}$ be monoidal Λ-supercategories. A *monoidal Λ-superfunctor* $F : \mathcal{A} \to \mathcal{B}$ is a Λ-superfunctor F equipped with even isomorphisms $\Phi_{X,Y} : F(X) \otimes F(Y) \to F(X \otimes Y)$ and $\varphi : F(\mathbb{1}_{\mathcal{A}}) \to \mathbb{1}_{\mathcal{B}}$ such that the following diagrams commute for all objects X, Y, Z of $\mathcal{A}$:

$$\begin{array}{ccccc}
(F(X)\otimes F(Y))\otimes F(Z) & \xrightarrow{\Phi_{X,Y}\otimes 1} & F(X\otimes Y)\otimes F(Z) & \xrightarrow{\Phi_{X\otimes Y,Z}} & F((X\otimes Y)\otimes Z) \\
\downarrow{\scriptstyle \alpha_{F(X),F(Y),F(Z)}} & & & & \downarrow{\scriptstyle F(\alpha_{X,Y,Z})} \\
F(X)\otimes(F(Y)\otimes F(Z)) & \xrightarrow{1\otimes\Phi_{Y,Z}} & F(X)\otimes F(Y\otimes Z) & \xrightarrow{\Phi_{X,Y\otimes Z}} & F(X\otimes(Y\otimes Z))
\end{array}$$

$$\begin{array}{ccc}
\mathbb{1}\otimes F(X) & \xrightarrow{\lambda_{F(X)}} & F(X) \\
\downarrow{\scriptstyle \varphi\otimes 1} & & \uparrow{\scriptstyle F(\lambda_X)} \\
F(\mathbb{1})\otimes F(X) & \xrightarrow{\Phi_{\mathbb{1},X}} & F(\mathbb{1}\otimes X)
\end{array}
\qquad
\begin{array}{ccc}
F(X)\otimes\mathbb{1} & \xrightarrow{\rho_{F(X)}} & F(X) \\
\downarrow{\scriptstyle 1\otimes\varphi} & & \uparrow{\scriptstyle F(\rho_X)} \\
F(X)\otimes F(\mathbb{1}) & \xrightarrow{\Phi_{X,\mathbb{1}}} & F(X\otimes\mathbb{1})
\end{array}$$

We will generally suppress Φ and φ from the notation. □

Definition 3.4. Given two monoidal Λ-superfunctors (F, Φ, φ) and (G, Γ, γ) between Λ-supercategories $\mathcal{A}$ and $\mathcal{B}$, a *monoidal natural transformation* $\eta : F \to G$ is an even natural transformation η such that the following two diagrams commute for all objects X, Y of $\mathcal{A}$:

$$\begin{array}{ccc}
F(X)\otimes F(Y) & \xrightarrow{\eta_X\otimes\eta_Y} & G(X)\otimes G(Y) \\
\downarrow{\scriptstyle \Phi_{X,Y}} & & \downarrow{\scriptstyle \Gamma_{X,Y}} \\
F(X\otimes Y) & \xrightarrow{\eta_{X\otimes Y}} & G(X\otimes Y)
\end{array}
\qquad
\begin{array}{ccc}
\mathbb{1} & \xrightarrow{\gamma} & G(\mathbb{1}) \\
\downarrow{\scriptstyle \varphi} & \nearrow{\scriptstyle \eta_{\mathbb{1}}} & \\
F(\mathbb{1}) & &
\end{array}$$

We thus have a 2-category of monoidal supercategories and, in particular, a notion of monoidal equivalence. □

3.2. ***The equivalences* $\boldsymbol{\Sigma}$ *and* $\mathbf{T}$.** Let $\mathcal{A}$ and $\mathcal{B}$ be $\mathbb{Z}/q$-graded supercategories, with q even. Define an equivalence

$$\Sigma_{\mathcal{A},\mathcal{B}} : \mathcal{A} \boxtimes \mathcal{B} \to \mathcal{B} \boxtimes \mathcal{A}$$

as follows. On homogeneous objects, we put $\Sigma_{\mathcal{A},\mathcal{B}}(A \boxtimes B) = B \boxtimes A$. Given homogeneous morphisms $f : A \to A'$ and $g : B \to B'$ between objects of $\mathcal{A}$ and $\mathcal{B}$, we define $\Sigma_{\mathcal{A},\mathcal{B}}(f \otimes g) = (-1)^{|f||g|} g \otimes f$.

Now suppose that $\mathcal{A}$ and $\mathcal{B}$ admit Π-structures. We define an equivalence

$$\mathrm{T}_{\mathcal{A},\mathcal{B}} : \mathcal{A} \boxtimes \mathcal{B} \to \mathcal{B} \boxtimes \mathcal{A}$$

as follows. On homogeneous objects, we put $\mathrm{T}_{\mathcal{A},\mathcal{B}}(A \boxtimes B) = \Pi^{|A||B|}(B) \boxtimes A$. Given homogeneous morphisms $f : A \to A'$ and $g : B \to B'$ of homogeneous objects, we put $\mathrm{T}_{\mathcal{A},\mathcal{B}}(f \otimes g) = (-1)^{|f||g|}\Pi^{|A||B|}(g) \otimes f$.

We regard Σ and T as analogous to the two symmetries σ and τ we defined on **SVec** in Section 2.1. They will give rise to the notion of symmetric and supersymmetric monoidal supercategories, just as σ and τ give rise to the notion of commutative and supercommutative superalgebras.

3.3. ***Factor systems.*** We next want to define braidings and symmetries, and their super analogs. Braidings will be required to satisfy a version of the hexagon axioms (see (H1) and (H2) in Section 3.4 below) but these axioms are only consistent if we allow the diagrams to commute up to certain scalars, depending on the degrees of the objects involved. These "factor systems" are defined below, but the definition is rather involved, so we first give a bit of explanation.

We will have two functions $(\mathbb{Z}/q)^3 \to \boldsymbol{k}^\times$ recording the scalars we allow for the two hexagons (H1) and (H2); we have a $(\mathbb{Z}/q)^3$ here since these diagrams involve three objects. These scalars are themselves subject to a number of constraints, coming from compatibilities of multiple instances of the hexagon axioms (explained in Sections 3.8 and 4). Moreover, the form of these compatibilities depends on whether we are working with braidings or superbraidings, which is recorded in the parity p below. In the symmetric case, the braiding will square to a scalar multiple of the identity (depending on the degrees of the objects), so we require an additional function $(\mathbb{Z}/q)^2 \to \boldsymbol{k}^\times$ recording these scalars.

With this explanation in hand, we now give the formal definition:

Definition 3.5. Let q be an even integer and let $p \in \mathbb{Z}/2$. A *B-factor system* of parity p is a pair $\omega = (\omega_1, \omega_2)$ of functions

$$\omega_1, \omega_2 : (\mathbb{Z}/q)^3 \to \boldsymbol{k}^\times$$

satisfying the following conditions for all $a, b, c, d \in \mathbb{Z}/q$ (in what follows, the semicolon in, e.g., $\omega_1(a; b, c)$, indicates that the b and c parameters play similar roles, while the a parameter is somewhat different):

$$\omega_1(a; b, c)\omega_1(a; b + c, d) = \omega_1(a; b, c + d)\omega_1(a; c, d), \tag{3.5a}$$

$$\omega_2(a + b, c; d)\omega_2(a, b; d) = \omega_2(b, c; d)\omega_2(a, b + c; d), \tag{3.5b}$$

$$\omega_1(b; c, d)\omega_1(a; c, d)\omega_1(a + b; c, d)^{-1} = (-1)^{abcdp}\omega_2(a, b; c)\omega_2(a, b; d)\omega_2(a, b; c + d)^{-1}, \tag{3.5c}$$

$$\omega_1(a; b, c)\omega_1(a; c, b)^{-1} = \omega_2(b, a; c)\omega_2(a, b; c)^{-1}. \tag{3.5d}$$

An *S-factor system* of parity p is a triple $\omega = (\omega_1, \omega_2, \omega_\sharp)$ where (ω_1, ω_2) is a B-factor system of parity p and $\omega_\sharp : (\mathbb{Z}/q)^2 \to \boldsymbol{k}^\times$ is a function satisfying

$$\omega_1(a; b, c)\omega_2(b, c; a) = \omega_\sharp(a, b)\omega_\sharp(a, c)\omega_\sharp(a, b + c)^{-1}, \tag{3.5e}$$

$$\omega_1(c; a, b)\omega_2(a, b; c) = \omega_\sharp(a, c)\omega_\sharp(b, c)\omega_\sharp(a + b, c)^{-1}. \tag{3.5f}$$

We say that ω is *symmetric* if $\omega_\sharp(a, b) = \omega_\sharp(b, a)$. □

Remark 3.6. (a) The "B" and "S" prefixes stand for "braiding" and "syllepsis" (see Section 4 for more).

(b) The parity p of the factor system only intervenes in (3.5c).

(c) Let Γ_p be the set of all S-factor systems of parity p for $\mathbb{Z}/q$. Then $\Gamma = \Gamma_0 \oplus \Gamma_1$ is a $\mathbb{Z}/2$-graded abelian group. □

Example 3.7. Let ω_1, ω_2 and $\omega_\sharp$ be the constant functions taking the value 1. Then ω is an even S-factor system. We refer to this as the *trivial factor system*. □

Example 3.8. Suppose that q is divisible by 4. Then $n \mapsto \binom{n}{2}$ is a well-defined function $\mathbb{Z}/q \to \mathbb{Z}/2$. Put

$$\mathtt{A}_1(a; b, c) = 1, \quad \mathtt{A}_2(a, b; c) = (-1)^{\binom{c}{2}ab}, \quad \mathtt{A}_\sharp(a, b) = (-1)^{\binom{a}{2}\binom{b}{2}}.$$

Then $\mathtt{A}$ is an odd symmetric S-factor system. This is perhaps the most important example for this paper. □

Example 3.9. Suppose that $\zeta_4 \in \boldsymbol{k}$. Define

$$\mathtt{B}_1(a; b, c) = 1, \qquad \mathtt{B}_2(a, b; c) = \begin{cases} \zeta_4 & \text{if } a, b \text{ and } c \text{ are all odd,} \\ 1 & \text{otherwise.} \end{cases}$$

Then $(\mathtt{B}_1, \mathtt{B}_2)$ is an odd B-factor system. Now suppose that $\zeta_8 \in \boldsymbol{k}$, and put

$$\mathtt{B}_\sharp(a, b) = \begin{cases} \zeta_8 & \text{if } a \text{ and } b \text{ are both odd,} \\ 1 & \text{otherwise.} \end{cases}$$

Then $(\mathtt{B}_1, \mathtt{B}_2, \mathtt{B}_\sharp)$ is an odd symmetric S-factor system. This factor system has the advantage over the S-factor system $\mathtt{A}$ in that it does not require q to be divisible by 4. □

Example 3.10. Define

$$\mathtt{C}_1(r; s, t) = (-1)^{rst}, \quad \mathtt{C}_2(r, s; t) = (-1)^{rst}, \quad \mathtt{C}_\sharp(r, s) = 1.$$

Then $\mathtt{C}$ is an even symmetric S-factor system. □

Example 3.11. Define

$$\mathtt{D}_1(r; s, t) = (-1)^{rst}, \quad \mathtt{D}_2(r, s; t) = (-1)^{rst}, \quad \mathtt{D}_\sharp(r, s) = (-1)^{rs}.$$

Then $\mathtt{D}$ is an even symmetric S-factor system. □

Definition 3.12. Let $\varphi : (\mathbb{Z}/q)^2 \to \boldsymbol{k}^\times$ be an arbitrary function. Put

$$\begin{aligned} \omega_1(r; s, t) &= \varphi(r, s+t)\varphi(r, s)^{-1}\varphi(r, t)^{-1}, \\ \omega_2(r, s; t) &= \varphi(r+s, t)\varphi(r, s)^{-1}\varphi(s, t)^{-1}, \\ \omega_\sharp(r, s) &= \varphi(r, s)\varphi(s, r). \end{aligned}$$

Then $(\omega_1, \omega_2, \omega_\sharp)$ is a symmetric even S-factor system, called the *coboundary* of φ. We denote this system by $\partial(\varphi)$. □

Example 3.13. Suppose $4 \mid q$. Then $\mathtt{C}$ is the coboundary of the function $\mathtt{c}$ given by

$$\mathtt{c}(r, s) = (-1)^{\binom{rs}{2}}.$$

□

Example 3.14. Suppose q is an arbitrary even integer and $\zeta_4 \in \boldsymbol{k}$. Then $\mathtt{D}$ is the coboundary of the function $\mathtt{d}$ given by

$$\mathtt{d}(x, y) = \begin{cases} 1 & \text{if } x \text{ or } y \text{ is even,} \\ \zeta_4 & \text{if } x \text{ and } y \text{ are odd.} \end{cases}$$

□

We now give a somewhat more involved example of a coboundary.

Proposition 3.15. *Set $\epsilon(n) = -1$ if $n \equiv 3 \pmod 4$ and $\epsilon(n) = 1$ otherwise.*

(a) *Suppose that $\zeta_4 \in \boldsymbol{k}$ and that q is divisible by 8. Define*

$$\mathtt{a}(n, m) = \epsilon(n)^m \zeta_4^{-\binom{n}{2}m}.$$

Then $\mathtt{A} = \mathtt{B} \cdot \partial(\mathtt{a})$ holds as B-factor systems.

(b) *Suppose that $\zeta_{16} \in \boldsymbol{k}$ and that q is divisible by 16. Define*

$$\mathtt{a}'(n, m) = \epsilon(n)^m \zeta_4^{-\binom{n}{2}m} \zeta_{16}^{nm}.$$

Then $\mathtt{A} = \mathtt{B} \cdot \partial(\mathtt{a}')$ holds as S-factor systems.

Proof. (a) The first component of the equation $\mathtt{A} = \mathtt{B} \cdot \partial(\mathtt{a})$ amounts to the identity

$$\epsilon(a)^{b+c} \zeta_4^{-\binom{a}{2}(b+c)} = \epsilon(a)^b \zeta_4^{-\binom{a}{2}b} \epsilon(a)^c \zeta_4^{-\binom{a}{2}c},$$

which is clear. The second component amounts to

$$\mathtt{B}_2(a, b; c)\epsilon(a+b)^c \zeta_4^{-\binom{a+b}{2}c} = (-1)^{\binom{c}{2}ab} \epsilon(b)^c \zeta_4^{-\binom{b}{2}c} \epsilon(a)^c \zeta_4^{-\binom{a}{2}c}.$$

The quotient of the left side by the right side is

$$(-1)^{\binom{c}{2}ab} \mathtt{B}_2(a, b; c)\epsilon(a+b)^c \epsilon(a)^c \epsilon(b)^c \zeta_4^{-abc},$$

which now only depends on a, b, c modulo 4. We will show it is 1 in all cases. If c is even, this reduces to $((-1)^{\binom{c}{2}} \zeta_4^{-c})^{ab} = 1$. If c is odd, then this simplifies to $\mathtt{B}_2(a, b; 1)\epsilon(a+b)\epsilon(a)\epsilon(b)\zeta_4^{-ab}$. If $a = 0$, we get $\epsilon(b)^2 = 1$. If $a = 1$, we get $\mathtt{B}_2(1, b; 1)\epsilon(b+1)\epsilon(b)\zeta_4^{-b}$. If $a = 2$, we get $\epsilon(b+2)\epsilon(b)(-1)^b$. Finally, if $a = 3$, we get $-\mathtt{B}_2(1, b; 1)\epsilon(b-1)\epsilon(b)\zeta_4^b$. In all cases, we see that the expressions are 1 for all choices of b by inspection.

(b) Since $\mathtt{a}'$ is equal to $\mathtt{a}$ times a quantity that is bilinear, the first two components of $\partial(\mathtt{a}')$ agree with those of $\partial(\mathtt{a})$. In particular, the equation $\mathtt{A} = \mathtt{B} \cdot \partial(\mathtt{a}')$ holds on the first two components by part (a). On the $\sharp$ component, it amounts to the identity

$$\epsilon(b)^a \zeta_4^{-\binom{b}{2}a} \zeta_{16}^{ab} \epsilon(a)^b \zeta_4^{-\binom{a}{2}b} \zeta_{16}^{ab} (-1)^{\binom{a}{2}\binom{b}{2}} = \mathtt{B}_\sharp(a, b).$$

The quotient of the left side by the right side is

$$\epsilon(b)^a \zeta_4^{-\binom{b}{2}a - \binom{a}{2}b} \zeta_8^{ab} \epsilon(a)^b (-1)^{\binom{a}{2}\binom{b}{2}} \mathtt{B}_\sharp(a, b)^{-1}.$$

We will show this is always 1. Suppose that $a = 2c$ is even (we will assume that a, b are integers). Then the above simplifies to

$$(-1)^{\binom{b}{2}c}\zeta_4^{-2bc^2+bc}\zeta_4^{bc}(-1)^{\binom{b}{2}c} = 1.$$

Since the expression is symmetric in a and b, it remains to prove that the expression is 1 when $a = 2c+1$ and $b = 2d+1$ are odd. In that case, it simplifies to

$$\begin{aligned}\epsilon(b)\zeta_4^{-(2d^2+d)a}\zeta_4^{-(2c^2+c)b}(-1)^{cd}\zeta_4^{c+d}\zeta_8\epsilon(a)(-1)^{cd}\zeta_8^{-1} &= \epsilon(b)(-1)^d\zeta_4^{-ad}(-1)^c\zeta_4^{-bc}\zeta_4^{c+d}\epsilon(a)\\ &= \epsilon(a)(-1)^c\epsilon(b)(-1)^d = 1.\end{aligned}$$

□

Remark 3.16. Let $\mathrm{C}^r(G, A)$ denote the group of r-cochains of a group G with values in the abelian group A. Note that $\mathrm{C}^0(G, A) = A$ is the set of functions from a single point to A. These groups form a cochain complex that computes the group cohomology of G. Let $\mathrm{C}^{r,s} = \mathrm{C}^r(\mathbb{Z}/q, \boldsymbol{k}^\times) \otimes \mathrm{C}^s(\mathbb{Z}/q, \mathbb{Z})$. We have a natural map

$$i : \mathrm{C}^{r,s} \to \mathrm{C}^{r+s}(\mathbb{Z}/q, \boldsymbol{k}^\times)$$

given on pure tensors by

$$i(c \otimes c')(g_1, \ldots, g_{r+s}) = c(g_1, \ldots, g_r)^{c'(g_{r+1},\ldots,g_{r+s})}.$$

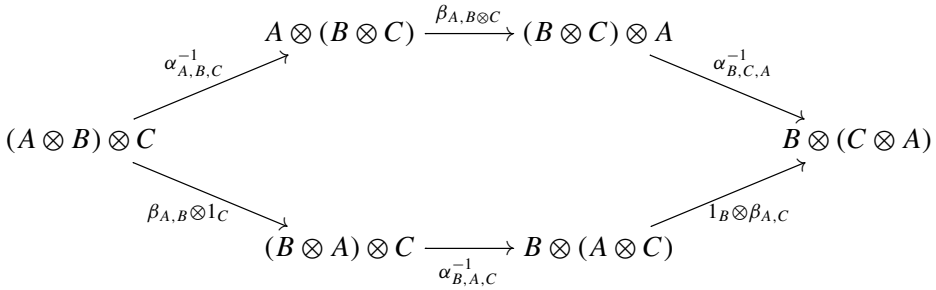
One easily sees that i is an isomorphism (e.g., by considering cochains supported at a single point). Thus, $\mathrm{C}^{1,2} \oplus \mathrm{C}^{2,1}$ can be identified with the set of pairs (ω_1, ω_2) of functions $(\mathbb{Z}/q)^3 \to \boldsymbol{k}^\times$. For such a pair to form an even B-factor system, it must be killed by the differential of the total complex (this accounts for the first three conditions in the definition of factor system). This suggests that even B-factor systems are closely related to $\mathrm{H}^3(\mathbb{Z}/q \times \mathbb{Z}/q, \boldsymbol{k}^\times)$, and that there could be a more fundamental definition including $\mathrm{C}^{0,3}$ and $\mathrm{C}^{3,0}$. We note that 3-cocycles of $\mathbb{Z}/q$ can be used to twist the associativity constraint in a monoidal $\mathbb{Z}/q$-supercategory, which might be relevant. We expect that even S-factor systems are similarly related to third abelian cohomology (as defined in [MacLane 1952]). It is not clear to us exactly how odd factor systems relate to cohomology. □

3.4. ***Braidings.*** Let $\mathcal{A}$ be a monoidal $\mathbb{Z}/q$-supercategory with q even (we allow $q = 0$). Consider an even isomorphism of superfunctors $\beta : \otimes \to \otimes \circ \Sigma$. Explicitly, β assigns to every pair (A, B) of objects an even isomorphism $\beta_{A,B} : A \otimes B \to B \otimes A$ such that, if $f : A \to A'$ and $g : B \to B'$ are homogeneous morphisms, then the diagram

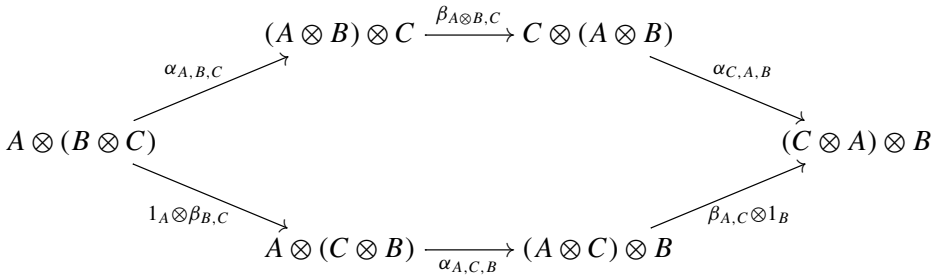
$$\begin{array}{ccc} A \otimes B & \xrightarrow{\beta_{A,B}} & B \otimes A \\ {\scriptstyle f\otimes g}\downarrow & & \downarrow{\scriptstyle g\otimes f} \\ A' \otimes B' & \xrightarrow{\beta_{A',B'}} & B' \otimes A' \end{array}$$

commutes up to $(-1)^{|f||g|}$.

Consider the two hexagons



(H1)



(H2)

Definition 3.17. We say that β is a *braiding* with respect to the even B-factor system ω if, in (H1), the bottom path is equal to $\omega_1(|A|; |B|, |C|)$ times the top path, and, in (H2), the bottom path is equal to $\omega_2(|A|, |B|; |C|)$ times the top path whenever A, B and C are homogeneous objects. □

Definition 3.18. We say that β is a *symmetry* with respect to the even S-factor system ω if it is a braiding and satisfies the condition $\beta_{B,A} \circ \beta_{A,B} = \omega_\sharp(|A|, |B|)\,\mathrm{id}_{A\otimes B}$ for all homogeneous objects A and B. □

Remark 3.19. By taking ω to be the trivial factor system (Example 3.7), one recovers the usual definition of symmetry: the diagrams (H1) and (H2) commute, and $\beta_{B,A}\beta_{A,B} = \mathrm{id}_{A\otimes B}$. □

Remark 3.20. Suppose β is a symmetry with respect to ω, and let $\varphi : (\mathbb{Z}/q)^2 \to \boldsymbol{k}^\times$ be an arbitrary function. Define β' by $\beta'_{X,Y} = \varphi(|X|, |Y|)\beta_{X,Y}$. Then β' is a symmetry with respect to $\omega \cdot \partial\varphi$, where $\partial(\varphi)$ is as in Definition 3.12. □

Remark 3.21. The reason that even factor systems are the appropriate systems of scalars will be made clear in Section 3.8. □

Definition 3.22. Let $\mathcal{A}$ and $\mathcal{B}$ be braided monoidal $\mathbb{Z}/q$-supercategories. A *braided monoidal* Λ-*superfunctor* $F : \mathcal{A} \to \mathcal{B}$ is a monoidal Λ-superfunctor (F, Φ, φ) such that the following diagram commutes for all objects X, Y of $\mathcal{A}$:

$$\begin{array}{ccc} F(X)\otimes F(Y) & \xrightarrow{\beta_{F(X),F(Y)}} & F(Y)\otimes F(X) \\ \downarrow{\scriptstyle \Phi_{X,Y}} & & \downarrow{\scriptstyle \Phi_{Y,X}} \\ F(X\otimes Y) & \xrightarrow{F(\beta_{X,Y})} & F(Y\otimes X) \end{array}$$

A natural transformation of braided monoidal superfunctors is simply one of monoidal superfunctors, i.e., the braiding is not used in the definition. □

3.5. ***Type I superbraidings.*** Suppose $\mathcal{A}$ is a monoidal $\mathbb{Z}/q$-supercategory, with q even (we allow $q=0$), admitting a Π-structure. Consider an even isomorphism of functors $\beta : \otimes \to \otimes \circ \mathrm{T}$. Explicitly, β assigns to every pair (A, B) of homogeneous objects an even isomorphism $\beta_{A,B} : A \otimes B \to \Pi^{|A||B|}(B) \otimes A$ such that, if $f : A \to A'$ and $g : B \to B'$ are nonzero homogeneous morphisms of homogeneous objects, then the diagram

$$\begin{array}{ccc} A \otimes B & \xrightarrow{\beta_{A,B}} & \Pi^{|A||B|}(B) \otimes A \\ {\scriptstyle f\otimes g}\downarrow & & \downarrow{\scriptstyle \Pi^{|A||B|}(g)\otimes f} \\ A' \otimes B' & \xrightarrow{\beta_{A',B'}} & \Pi^{|A||B|}(B') \otimes A' \end{array}$$

commutes up to $(-1)^{|f||g|}$. Note that the existence of nonzero f, g forces $|A| = |A'|$ and $|B| = |B'|$; if f or g is 0, then the power of Π may need to change, but the resulting diagram will commute for trivial reasons.

The hexagons (H1) and (H2) admit obvious modifications (H1′) and (H2′) in this setting, as follows. Let A, B and C be homogeneous objects. Then (H1′) is the diagram

$$\begin{array}{ccccc} & & A \otimes (B \otimes C) & \xrightarrow{\psi_1} & (\Pi^{b+c}(B) \otimes C) \otimes A & & \\ & {\scriptstyle\alpha^{-1}}\nearrow & & & & \searrow{\scriptstyle\alpha^{-1}} & \\ (A \otimes B) \otimes C & & & & & & \Pi^{b+c}(B) \otimes (C \otimes A) \\ & {\scriptstyle\beta\otimes\mathrm{id}}\searrow & & & & \nearrow{\scriptstyle\psi_2} & \\ & & (\Pi^b(B) \otimes A) \otimes C & \xrightarrow[\alpha^{-1}]{} & \Pi^b(B) \otimes (A \otimes C) & & \end{array} \qquad (\mathrm{H1}')$$

Here $b = |A||B|$, $c = |A||C|$, ψ_1 is the composition

$$A \otimes (B \otimes C) \xrightarrow{\beta_{A,B\otimes C}} \Pi^{b+c}(B \otimes C) \otimes A \xrightarrow{\sim} (\Pi^{b+c}(B) \otimes C) \otimes A,$$

ψ_2 is the composition

$$\Pi^b(B) \otimes (A \otimes C) \xrightarrow{\mathrm{id}\otimes\beta_{A,C}} \Pi^b(B) \otimes (\Pi^c(C) \otimes A) \xrightarrow{\xi\otimes\xi\otimes\mathrm{id}} \Pi^{b+c}(B) \otimes (C \otimes A),$$

and we have omitted indices for readability; recall $\xi = \xi^{n,m}$ is the isomorphism $\Pi^n \to \Pi^m$. The diagram (H2′) is

$$\begin{array}{ccccc} & & (A \otimes B) \otimes C & \xrightarrow{\beta} & \Pi^{a+b}(C) \otimes (A \otimes B) & & \\ & {\scriptstyle\alpha}\nearrow & & & & \searrow{\scriptstyle\alpha} & \\ A \otimes (B \otimes C) & & & & & & (\Pi^{a+b}(C) \otimes A) \otimes B \\ & {\scriptstyle\mathrm{id}\otimes\beta}\searrow & & & & \nearrow{\scriptstyle\beta\otimes\mathrm{id}} & \\ & & A \otimes (\Pi^b(C) \otimes B) & \xrightarrow[\alpha]{} & (A \otimes \Pi^b(C)) \otimes B & & \end{array} \qquad (\mathrm{H2}')$$

where $a = |A||C|$ and $b = |B||C|$, and we have again omitted indices.

It is not correct to ask for the above hexagons to commute: in fact, this leads to inconsistencies (see Section 3.8). Instead, we make the following definition:

Definition 3.23. We say that β is a *type I superbraiding* with respect to the odd B-factor system ω if, in (H1′), the bottom path is equal to $\omega_1(|A|; |B|, |C|)$ times the top path, and, in (H2′), the bottom path is equal to $\omega_2(|A|, |B|; |C|)$ times the top path. □

Definition 3.24. We say that β is a *type I supersymmetry* with respect to the odd S-factor system ω if it is a type I superbraiding with respect to ω and, for all homogeneous objects A and B, the diagram

$$\begin{array}{ccc} A \otimes B & \xrightarrow{\beta_{A,B}} & \Pi^{ab}(B) \otimes A \\ & \searrow^{\omega_\natural(a,b)\xi\otimes\xi} & \downarrow{\scriptstyle \beta_{\Pi^{ab}(B),A}} \\ & & \Pi^{ab}(A) \otimes \Pi^{ab}(B) \end{array}$$

commutes, where $a = |A|$ and $b = |B|$. □

Remark 3.25. If β satisfies the conditions to be a supersymmetry, then the two hexagon axioms are equivalent conditions: starting from either hexagon, invert all of the arrows and relabel the objects to get the other hexagon. □

Example 3.26. We define a *type Ia superbraiding* (resp. *supersymmetry*) to be a type I superbraiding (resp. supersymmetry) with respect to the factor system $\mathtt{A}$ from Example 3.8. □

Example 3.27. We define a *type Ib superbraiding* (resp. *supersymmetry*) to be a type I superbraiding (resp. supersymmetry) with respect to the factor system $\mathtt{B}$ from Example 3.9. □

Remark 3.28. One can convert between type Ia and type Ib superbraidings when $\mathtt{BA}^{-1}$ is a coboundary, in the sense of Definition 3.12 (see Remark 3.20). Proposition 3.15 gives conditions that ensure $\mathtt{BA}^{-1}$ is a coboundary. □

Definition 3.29. Let $\mathcal{A}$ and $\mathcal{B}$ be monoidal $\mathbb{Z}/q$-supercategories equipped with type I superbraidings with respect to the same factor systems. A *type I superbraided monoidal superfunctor* is a monoidal superfunctor $F : \mathcal{A} \to \mathcal{B}$ such that the diagram

$$\begin{array}{ccccc} F(X) \otimes F(Y) & & \xrightarrow{\beta_{F(X),F(Y)}} & & \Pi^{|X||Y|}(F(Y) \otimes F(X)) \\ \downarrow{\scriptstyle \Phi_{X,Y}} & & & & \downarrow{\scriptstyle \Pi^{|X||Y|}(\Phi_{Y,X})} \\ F(X \otimes Y) & \xrightarrow{F(\beta_{X,Y})} & F(\Pi^{|X||Y|}(Y \otimes X)) & = & \Pi^{|X||Y|}(F(Y \otimes X)) \end{array}$$

commutes, where the equality is the canonical isomorphism coming from Proposition 2.2. A natural transformation of such functors is simply one of monoidal superfunctors. □

3.6. ***Type II superbraidings.*** The definition of type I superbraiding followed the general procedure of replacing Σ by T, and therefore has some conceptual clarity. However, these superbraidings can be somewhat inconvenient due to the Π's that arise; for instance, if X is an odd-degree object, then a type I superbraiding does not induce an endomorphism of $X \otimes X$, but rather a map $X \otimes X \to \Pi(X) \otimes X$. Of course, one can convert an even-degree map $X \otimes Y \to \Pi^{|X||Y|}(Y) \otimes X$ into a map $X \otimes Y \to Y \otimes X$ of degree $|X||Y|$. This leads to an alternative formulation of the concept of superbraiding, in which there are no Π's, but the maps may not be even. We now give the details.

Let $\mathcal{A}$ be a monoidal $\mathbb{Z}/q$-supercategory, with q even (we allow $q = 0$); we do not require a Π-structure. Consider an isomorphism $\beta : \otimes \to \otimes \circ \Sigma$ such that, if A and B are homogeneous, then $\beta_{A,B}$ is homogeneous of degree $|A||B|$. Explicitly, β assigns to every pair of homogeneous objects (A, B) an isomorphism $\beta_{A,B} : A \otimes B \to B \otimes A$ of degree $|A||B|$ such that, if $f : A \to A'$ and $g : B \to B'$ are homogeneous morphisms of homogeneous objects, then the diagram

$$\begin{array}{ccc} A \otimes B & \xrightarrow{\beta_{A,B}} & B \otimes A \\ {\scriptstyle f\otimes g}\downarrow & & \downarrow{\scriptstyle g\otimes f} \\ A' \otimes B' & \xrightarrow{\beta_{A',B'}} & B' \otimes A' \end{array}$$

commutes up to $(-1)^{|A||B|(|f|+|g|)+|f||g|}$. The first part of the sign, namely $(-1)^{|A||B|(|f|+|g|)}$, comes from the condition that β is a natural transformation, while the second part of the sign, namely $(-1)^{|f||g|}$, comes from the definition of Σ.

Definition 3.30. We say that β is a *type II superbraiding* with respect to the B-factor system ω if, in (H1), the bottom path is equal to $\omega_1(|A|; |B|, |C|)$ times the top path, and, in (H2), the bottom path is equal to $\omega_2(|A|, |B|; |C|)$ times the top path. □

Definition 3.31. We say that β is a *type II supersymmetry* with respect to the S-factor system ω if it is a type II superbraiding with respect to ω and, for all homogeneous objects A and B, we have $\beta_{B,A}\beta_{A,B} = \omega_\sharp(|A|, |B|) \cdot \mathrm{id}_{A\otimes B}$ □

Remark 3.32. We define type IIa/IIb superbraidings/supersymmetries analogously to the type I case (see Examples 3.26 and 3.27). □

Definition 3.33. Let $\mathcal{A}$ and $\mathcal{B}$ be monoidal $\mathbb{Z}/q$-supercategories equipped with type II superbraidings with respect to the same factor systems. A *type II superbraided monoidal functor* is a monoidal superfunctor $F : \mathcal{A} \to \mathcal{B}$ such that the diagram

$$\begin{array}{ccc} F(X) \otimes F(Y) & \xrightarrow{\beta_{F(X),F(Y)}} & F(Y) \otimes F(X) \\ {\scriptstyle \Phi_{X,Y}}\downarrow & & \downarrow{\scriptstyle \Phi_{Y,X}} \\ F(X \otimes Y) & \xrightarrow{F(\beta_{X,Y})} & F(Y \otimes X) \end{array}$$

commutes for all objects X and Y. A natural transformation of such functors is simply one of monoidal superfunctors. □

3.7. ***Equivalence between type I and type II superbraidings.*** Fix a monoidal $\mathbb{Z}/q$-supercategory (with q even) $\mathcal{A}$ with a Π-structure. The following proposition shows how to convert between the two types of superbraidings:

Proposition 3.34. *Suppose that β is a type II superbraiding with respect to the odd B-factor system ω. For homogeneous objects A and B, define $\beta'_{A,B}$ to be the composition*

$$A \otimes B \xrightarrow{\beta_{A,B}} B \otimes A \xrightarrow{\xi_B^{0,|A||B|} \otimes \mathrm{id}_A} \Pi^{|A||B|}(B) \otimes A.$$

Then β' is a type I superbraiding with respect to the odd B-factor system $\mathtt{C}\omega$, where $\mathtt{C}$ is as in Example 3.10. Moreover, this construction defines a bijection between the two types of superbraidings.

Now suppose that β is a type II supersymmetry with respect to the odd S-factor system ω. Then β' is a type I supersymmetry with respect to the odd S-factor system $\mathtt{D}\omega$, where $\mathtt{D}$ is as in Example 3.11. Again, this construction is bijective.

We note that $\mathtt{C} = \mathtt{D}$ as B-factor systems, and so one could use $\mathtt{D}$ in the first part of the proposition as well.

Proof. First we check that β' is a natural transformation, i.e., that the square

$$\begin{array}{ccc} A \otimes B & \xrightarrow{\beta'_{A,B}} & \Pi^{|A||B|}(B) \otimes A \\ {\scriptstyle f\otimes g}\downarrow & & \downarrow{\scriptstyle \Pi^{|A||B|}(g)\otimes f} \\ A' \otimes B' & \xrightarrow{\beta'_{A',B'}} & \Pi^{|A||B|}(B') \otimes A' \end{array}$$

commutes up to $(-1)^{|f||g|}$ for all nonzero homogeneous morphisms $f : A \to A'$ and $g : B \to B'$ between homogeneous objects. By definition, this expands to the diagram

$$\begin{array}{ccccc} A \otimes B & \xrightarrow{\beta_{A,B}} & B \otimes A & \xrightarrow{\xi\otimes 1} & \Pi^{|A||B|}(B) \otimes A \\ {\scriptstyle f\otimes g}\downarrow & & \downarrow{\scriptstyle g\otimes f} & & \downarrow{\scriptstyle \Pi^{|A||B|}(g)\otimes f} \\ A' \otimes B' & \xrightarrow{\beta_{A',B'}} & B' \otimes A' & \xrightarrow{\xi\otimes 1} & \Pi^{|A||B|}(B') \otimes A' \end{array}$$

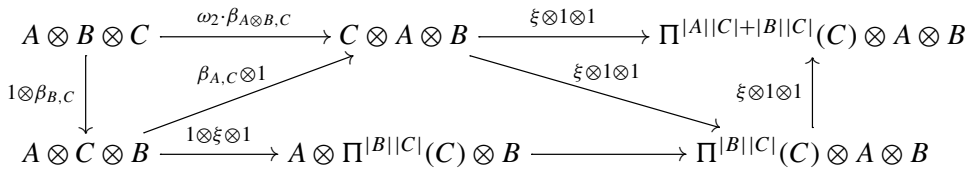
Since β is a type II superbraiding, the left square commutes up to $(-1)^{|A||B|(|f|+|g|)+|f||g|}$. By definition of how Π acts on morphisms, the right square commutes up to $(-1)^{|A||B|(|f|+|g|)}$, and so the original square commutes up to $(-1)^{|f||g|}$, as desired.

Now we check that the hexagon axiom (H1′) holds. Consider the diagram (for simplicity we ignore the associators)

$$\begin{array}{ccccc} A \otimes B \otimes C & \xrightarrow{\omega_1\cdot\beta_{A,B\otimes C}} & B \otimes C \otimes A & \xrightarrow{\xi\otimes 1} & \Pi^{|A||B|+|A||C|}(B\otimes C)\otimes A \\ {\scriptstyle \beta_{A,B}\otimes 1}\downarrow & \nearrow{\scriptstyle 1\otimes\beta_{A,C}} & & & \downarrow\cong \\ B \otimes A \otimes C & & {\scriptstyle \xi\otimes 1\otimes 1}\downarrow & \searrow{\scriptstyle \xi\otimes\xi\otimes 1} & \Pi^{|A||B|+|A||C|}(B)\otimes C\otimes A \\ {\scriptstyle \xi\otimes 1\otimes 1}\downarrow & & & & \uparrow\cong \\ \Pi^{|A||B|}(B)\otimes A\otimes C & \xrightarrow{1\otimes\beta_{A,C}} & \Pi^{|A||B|}(B)\otimes C\otimes A & \xrightarrow{1\otimes\xi\otimes 1} & \Pi^{|A||B|}(B)\otimes\Pi^{|A||C|}(C)\otimes A \end{array}$$

where the ω_1 in the top left arrow is $\omega_1(|A|;|B|,|C|)$. On the left side, the upper triangle commutes by the first hexagon axiom for β. The bottom quadrilateral commutes up to $(-1)^{|A||B||C|}$. On the right side, the bottom triangle commutes up to $(-1)^{|A||B||C|}$ by the sign rule when we compose the bottom path. The top triangle commutes: we can take $\Pi(X\otimes Y)=\Pi(X)\otimes Y$, and $\xi : X\otimes Y\to\Pi(X\otimes Y)$ to be $\xi_X\otimes \mathrm{id}_Y$, and we use this convention for the $\Pi(B\otimes C)$ above. Then the top right vertical map is the identity map; the bottom right vertical map is $\xi\otimes\xi\otimes 1$. Hence, we see that (H1′) holds for β' by following the outer paths.

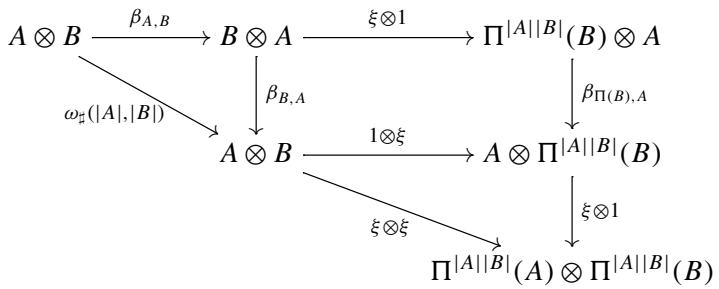
Finally, we check that the hexagon axiom (H2′) holds for β' by considering the diagram (again we ignore associators for simplicity)



where the δ in the top left arrow is $\omega_2(|A|,|B|;|C|)$, and the bottom right map is $\beta_{A,\Pi^{|B||C|}(C)}\otimes 1$. The left triangle commutes by the second hexagon axiom for β. The middle quadrilateral commutes up to $(-1)^{|A||B||C|}$ and the right triangle commutes. Hence, by considering the outer paths, we see that β' satisfies the second hexagon axiom.

Finally, the diagrams above can be used to show that β is a type II superbraiding if β' is a type I superbraiding and hence the stated construction gives a bijection between type I and type II superbraidings.

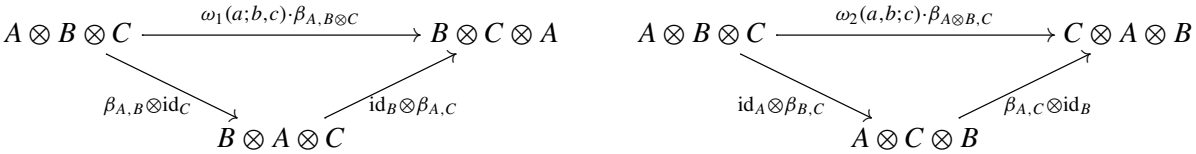
Now we handle the last statement, and suppose that β is a type I supersymmetry. Consider the diagram



The upper left triangle commutes since β is a type II supersymmetry, and the bottom triangle commutes by definition of composition. The square commutes up to $(-1)^{|A||B|}$ since β is a type II superbraiding, so we conclude that β' is a superbraiding with respect to $\mathtt{D}\omega$. □

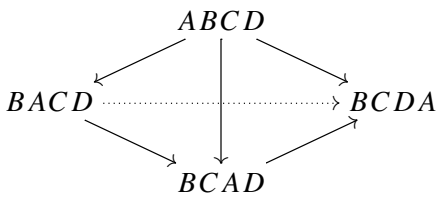
3.8. ***Explanation of factor systems.*** We now explain the rationale behind the definition of B-factor systems. We work with type II superbraidings here, but this discussion can be translated to type I superbraidings or ordinary braidings as well. Let $\mathcal{A}$ be a monoidal $\mathbb{Z}/q$-supercategory with q even. We assume, for simplicity, that the monoidal structure is strict. Suppose that we have an isomorphism $\beta_{X,Y}:\otimes\to\otimes\circ\Sigma$ such that $\beta_{X,Y}$ is homogeneous of degree $|X||Y|$ on homogeneous objects, and that (H1) and (H2) commute up to scalars. Precisely, for homogeneous objects A, B and C of degrees a, b

and c, we suppose that the diagrams



commute, where ω_1 and ω_2 are nonzero scalars. Note that these triangles are the hexagon axioms since we have assumed that the associators are the identity.

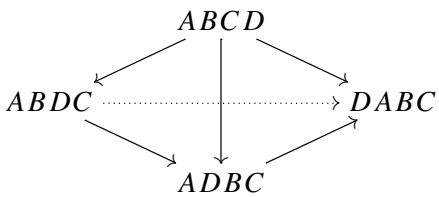
Let A, B, C and D be four homogeneous objects of degrees a, b, c and d. Consider the tetrahedron



(3.35a)

Here we have omitted the $\otimes$ symbols, and each arrow is the unique β map that makes sense. This tetrahedron has four faces that each commute up to an ω_1-factor. This yields the identity (3.5a).

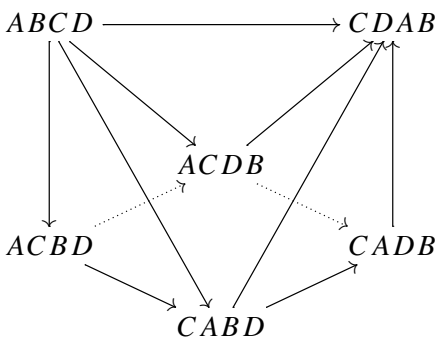
Similarly, we have the commuting tetrahedron



(3.35b)

which yields the identity (3.5b).

Now consider the diagram

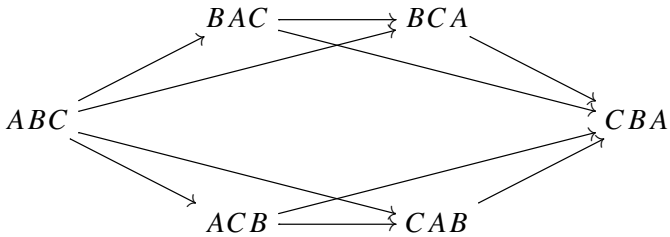


(3.35c)

This diagram has six triangles and one square (composed of the bottom four vertices). If a, b, c and d are all odd, then both $\beta_{A,C}$ and $\beta_{B,D}$ are odd, and the square anticommutes; this is the one place where the specifics of supercategories intervene. Otherwise, at least one of $\beta_{A,C}$ and $\beta_{B,D}$ are even, and the square

commutes. Thus, in general, the square commutes up to $(-1)^{abcd}$, and we obtain (3.5c) with $p = 1$. (In the case of ordinary braidings, this sign would not be present, and so we would obtain (3.5c) with $p = 0$.)

Finally, consider the diagram



(3.35d)

This diagram has four triangles: two that commute up to ω_1-factors and two that commute up to ω_2-factors. It also has two squares that commute by the naturality of β. For example, one of the squares is

$$\begin{array}{ccc} ABC & \xrightarrow{\beta_{A,BC}} & BCA \\ {\scriptstyle \mathrm{id}\otimes\beta_{B,C}}\downarrow & & \downarrow{\scriptstyle \beta_{B,C}\otimes\mathrm{id}} \\ ACB & \xrightarrow{\beta_{A,CB}} & CBA \end{array}$$

and this commutes up to $(-1)^{a(b+c)bc} = 1$. This diagram yields (3.5d).

We thus see that the four equations in the definition of a B-factor system are necessary for consistency of the hexagon axioms. Of course, one might wonder if there are additional equations that are needed that we have overlooked. In fact, this is not the case: this is best explained by the 2-categorical considerations in Section 4. (The discussion there will also explain the equations for S-factor systems.)

3.9. ***Grothendieck groups.*** Let $\mathcal{A}$ be a monoidal $\mathbb{Z}/q$-supercategory with q even. Suppose that $\mathcal{A}^\circ$ has an exact structure and $\otimes$ is exact. Then $\otimes$ induces a multiplication on $\mathrm{K}(\mathcal{A}^\circ)$ and $\mathrm{K}_\pm(\mathcal{A})$ and gives them each the structure of a superring. If the monoidal structure admits a braiding, then all three of these superrings are commutative. If the monoidal structure admits a superbraiding (of any kind), then $\mathrm{K}_+(\mathcal{A})$ is commutative while $\mathrm{K}_-(\mathcal{A})$ is supercommutative. Thus, via K_-, we can view supersymmetric monoidal supercategories as a categorification of the notion of a supercommutative superalgebra.

4. Some 2-category theory

In this section, we explain the 2-categorical underpinnings for supersymmetries and superbraidings. The main point is to show that the equivalence T (introduced in Section 3.2), together with a B-factor system, allows us to define a braiding on the 2-category of $\mathbb{Z}/q$-supercategories; the braided objects will then be our superbraided categories. There is a similar story involving S-factor systems and also the equivalence Σ.

We do not intend for this discussion to be perfectly rigorous, and we allow ourselves to take some shortcuts and make some simplifications. In Sections 4.1–4.3, we review the theory of monoidal 2-categories in general. We have mostly followed [Crans 1998], and refer there for more details; see also

[Baez and Neuchl 1996; Kapranov and Voevodsky 1994]. From Section 4.4 onwards, we specialize to our case of interest, the 2-category of $\mathbb{Z}/q$-supercategories.

4.1. ***Monoidal 2-categories.*** Let $(\mathscr{C}, \otimes)$ be a semistrict monoidal 2-category. We refer to [Crans 1998, Section 2.1] for the precise definition, and make a few remarks here. We require composition of 1-morphisms in $\mathscr{C}$ to be strictly associative (this is part of what "2-category" means). We also require $\otimes$ to be strictly associative (this is part of what "semistrict" means). We ignore the unit of the monoidal structure in our discussion. An important piece of the monoidal structure is that the square

$$\begin{array}{ccc} A \otimes B & \xrightarrow{\mathrm{id}_A \otimes g} & A \otimes B' \\ {\scriptstyle f \otimes \mathrm{id}_B}\downarrow & & \downarrow{\scriptstyle f \otimes \mathrm{id}_{B'}} \\ A' \otimes B & \xrightarrow{\mathrm{id}_{A'} \otimes g} & A' \otimes B' \end{array}$$

commutes up to a given 2-isomorphism $i_{f,g}$; precisely,

$$i_{f,g} : (f \otimes \mathrm{id}_{B'}) \circ (\mathrm{id}_A \otimes g) \xrightarrow{\sim} (\mathrm{id}_{A'} \otimes g) \circ (f \otimes \mathrm{id}_B).$$

In what follows, we omit tensor symbols (i.e., we write AB in place of $A \otimes B$), and simply write A for the identity 1-morphism on an object A of $\mathscr{C}$.

4.2. ***Braidings on monoidal 2-categories.*** A *braiding* on $\mathscr{C}$ consists of a pseudonatural equivalence $R_{A,B} : AB \to BA$ together with 2-isomorphisms

$$\rho_{A|B,C} : (BR_{A,C}) \circ (R_{A,B}C) \to R_{A,BC}, \quad \rho_{A,B|C} : (R_{A,C}B) \circ (AR_{B,C}) \to R_{AB,C}$$

satisfying the five axioms below (we have inverted the convention from [Crans 1998, Section 2.2] using ρ in place of ρ^{-1}). Note that the above two 2-isomorphisms make the diagrams

$$\begin{array}{ccc} ABC & \xrightarrow{R_{A,BC}} & BCA \\ & {\scriptstyle R_{A,B}C}\searrow \quad \nearrow{\scriptstyle BR_{A,C}} & \\ & BAC & \end{array} \qquad \begin{array}{ccc} ABC & \xrightarrow{R_{AB,C}} & CAB \\ & {\scriptstyle AR_{B,C}}\searrow \quad \nearrow{\scriptstyle R_{A,C}B} & \\ & ACB & \end{array}$$

commute up to isomorphism, and are essentially the hexagon axioms. (Remember that the associators are identity morphisms, and can thus be omitted.) We recall that, if $f : A \to A'$ and $g : B \to B'$ are 1-morphisms, then there is a given 2-isomorphism

$$R_{f,g} : (gf)R_{A,B} \xrightarrow{\sim} R_{A',B'}(fg);$$

this is part of the data of the pseudonatural equivalence R. This isomorphism makes the square

$$\begin{array}{ccc} AB & \xrightarrow{R_{A,B}} & BA \\ {\scriptstyle fg}\downarrow & & \downarrow{\scriptstyle gf} \\ A'B' & \xrightarrow{R_{A',B'}} & B'A' \end{array}$$

commute up to isomorphism.

Axiom 1. The tetrahedron in (3.35a) commutes for all objects A, B, C and D, where the edges are the obvious 1-morphisms built from R and the faces are the 2-morphisms provided by ρ. Here "commutes" is taken in the 2-categorical sense. Explicitly, this means that the diagram

$$\begin{array}{ccc}
(BCR_{A,D})(BR_{A,C}D)(R_{A,B}CD) & \xrightarrow{B\rho_{A|C,D}\circ 1} & (BR_{A,CD})(R_{A,B}CD) \\
{\scriptstyle 1\circ(\rho_{A|B,C}D)}\downarrow & & \downarrow{\scriptstyle \rho_{A|B,CD}} \\
(BCR_{A,D})(R_{A,BC}D) & \xrightarrow{\rho_{A|BC,D}} & R_{A,BCD}
\end{array}$$

commutes.

Axiom 2. The tetrahedron in (3.35b) commutes. This means that the diagram

$$\begin{array}{ccc}
(R_{A,D}BC)(AR_{B,D}C)(ABR_{C,D}) & \xrightarrow{(\rho_{A,B|D}C)\circ 1} & (R_{AB,D}C)(ABR_{C,D}) \\
{\scriptstyle 1\circ(A\rho_{B,C|D})}\downarrow & & \downarrow{\scriptstyle \rho_{AB,C|D}} \\
(R_{A,D}BC)(AR_{BC,D}) & \xrightarrow{\rho_{A,BC|D}} & R_{ABC,D}
\end{array}$$

commutes.

Axiom 3. The diagram in (3.35c) commutes; the 2-isomorphism in the square is given by i. We write this out explicitly using three diagrams. First, we require that the diagram

$$\begin{array}{ccc}
(CR_{A,D}B)(R_{A,C}DB)(ACR_{B,D})(AR_{B,C}D) & \xrightarrow{1\circ 1\circ A\rho_{B|C,D}} & (CR_{A,D}B)(R_{A,C}DB)(AR_{B,CD}) \\
{\scriptstyle \rho_{A|C,D}B\circ 1\circ 1}\downarrow & \overset{\varphi}{\searrow} & \downarrow{\scriptstyle \rho_{A|C,D}B\circ 1} \\
(R_{A,CD}B)(ACR_{B,D})(AR_{B,C}D) & \xrightarrow{1\circ A\rho_{B|C,D}} & (R_{A,CD}B)(AR_{B,CD})
\end{array}$$

commute, and define φ to be the indicated arrow. Second, we require that the diagram

$$\begin{array}{ccc}
(CR_{A,D}B)(CAR_{B,D})(R_{A,C}BD)(AR_{B,C}D) & \xrightarrow{C\rho_{A,B|D}\circ 1\circ 1} & (CR_{AB,D})(R_{A,C}BD)(AR_{B,C}D) \\
{\scriptstyle 1\circ 1\circ\rho_{A,B|C}D}\downarrow & \overset{\psi}{\searrow} & \downarrow{\scriptstyle 1\circ\rho_{A,B|C}D} \\
(CR_{A,D}B)(CAR_{B,D})(R_{AB,C}D) & \xrightarrow{C\rho_{A,B|D}\circ 1} & (CR_{AB,D})(R_{AB,C}D)
\end{array}$$

commute, and again define ψ to be the indicated arrow. Finally, we require that the diagram

$$\begin{array}{ccccc}
(CR_{A,D}B)(R_{A,C}DB)(ACR_{B,D})(AR_{B,C}D) & & \xrightarrow{1\circ i\circ 1} & & (CR_{A,D}B)(CAR_{B,D})(R_{A,C}BD)(AR_{B,C}D) \\
{\scriptstyle \varphi}\downarrow & & & & \downarrow{\scriptstyle \psi} \\
(R_{A,CD}B)(AR_{B,CD}) & \xrightarrow{\rho_{A,B|CD}} & R_{AB,CD} & \xleftarrow{\rho_{AB|C,D}} & (CR_{AB,D})(R_{AB,C}D)
\end{array}$$

commute, where $i = i_{f,g}$ with $f = R_{A,C}$ and $g = R_{B,D}$.

Axiom 4. The diagram in (3.35d) commutes; the 2-isomorphisms in the four triangles come from ρ, while the 2-isomorphisms in the two squares come from R. Explicitly, this means that the diagram

$$\begin{array}{ccccc}
(R_{BA,C})(R_{A,B}C) & \xleftarrow{\rho_{B,A|C}\circ 1} & (R_{B,C}A)(BR_{A,C})(R_{A,B}C) & \xrightarrow{1\circ\rho_{A|B,C}} & (R_{B,C}A)(R_{A,BC}) \\
\downarrow{\scriptstyle R_{R_{A,B},\mathrm{id}_C}} & & & & \downarrow{\scriptstyle R_{\mathrm{id}_A,R_{B,C}}} \\
(CR_{A,B})(R_{AB,C}) & \xrightarrow{1\circ\rho^{-1}_{A,B|C}} & (CR_{A,B})(R_{A,C}B)(AR_{B,C}) & \xleftarrow{\rho^{-1}_{A|C,B}\circ 1} & (R_{A,CB})(AR_{B,C})
\end{array}$$

commutes.

Axiom 5. A number of conditions regarding the unit object hold; see [Crans 1998, Section 2.2].

4.3. ***Symmetric and sylleptic structures.*** There are two "levels" of commutativity for monoidal 1-categories: braided and symmetric. For monoidal 2-categories, there are three levels: braided, sylleptic and symmetric. We now recall what the latter two mean.

Fix a braiding R on $\mathscr{C}$. A *syllepsis* is a 2-isomorphism (we have inverted the convention from [Crans 1998, Section 4], using v in place of v^{-1})

$$v_{A,B} : R_{B,A} \circ R_{A,B} \to \mathrm{id}_{A\otimes B}$$

such that the following diagrams commute (as before, the dotted arrows denote the definition of the common composition in the first and third diagrams):

$$\begin{array}{ccc}
(R_{B,A}C)(BR_{C,A})(BR_{A,C})(R_{A,B}C) & \xrightarrow{1\circ 1\circ\rho_{A|B,C}} & (R_{B,A}C)(BR_{C,A})(R_{A,BC}) \\
\downarrow{\scriptstyle \rho_{B,C|A}\circ 1\circ 1} & \overset{\varphi}{\cdots\cdots\to} & \downarrow{\scriptstyle \rho_{B,C|A}\circ 1} \\
(R_{BC,A})(BR_{A,C})(R_{A,B}C) & \xrightarrow{1\circ\rho_{A|B,C}} & (R_{BC,A})(R_{A,BC})
\end{array}$$

$$\begin{array}{ccc}
(R_{B,A}C)(BR_{C,A})(BR_{A,C})(R_{A,B}C) & \xrightarrow{\varphi} & (R_{BC,A})(R_{A,BC}) \\
\downarrow{\scriptstyle 1\circ Bv_{A,C}\circ 1} & & \downarrow{\scriptstyle v_{A,BC}} \\
(R_{B,A}C)(R_{A,B}C) & \xrightarrow{v_{A,B}C} & \mathrm{id}_{ABC}
\end{array}$$

$$\begin{array}{ccc}
(AR_{C,B})(R_{C,A}B)(R_{A,C}B)(AR_{B,C}) & \xrightarrow{1\circ 1\circ\rho_{A,B|C}} & (AR_{C,B})(R_{C,A}B)(R_{AB,C}) \\
\downarrow{\scriptstyle \rho_{C|A,B}\circ 1\circ 1} & \overset{\psi}{\cdots\cdots\to} & \downarrow{\scriptstyle \rho_{C|A,B}\circ 1} \\
(R_{C,AB})(R_{A,C}B)(AR_{B,C}) & \xrightarrow{1\circ\rho_{A,B|C}} & (R_{C,AB})(R_{AB,C})
\end{array}$$

$$\begin{array}{ccc}
(AR_{C,B})(R_{C,A}B)(R_{A,C}B)(AR_{B,C}) & \xrightarrow{\psi} & (R_{C,AB})(R_{AB,C}) \\
\downarrow{\scriptstyle 1\circ v_{A,C}B\circ 1} & & \downarrow{\scriptstyle v_{AB,C}} \\
(AR_{C,B})(AR_{B,C}) & \xrightarrow{Av_{B,C}} & \mathrm{id}_{ABC}
\end{array}$$

There are also conditions involving the unit object that we omit (see [Crans 1998, Section 4]). A syllepsis is a *symmetry* if $1 \circ v_{A,B} = v_{B,A} \circ 1$ agree as isomorphisms

$$R_{A,B}R_{B,A}R_{A,B} \to R_{A,B}.$$

4.4. ***The 2-category of supercategories.*** Fix an even positive integer q. For the remainder of this section, we let $\mathscr{C}$ be the 2-category whose objects are $\mathbb{Z}/q$-supercategories equipped with a Π-functor. The 1-morphisms of $\mathscr{C}$ are superfunctors, and the 2-morphisms are even natural transformations. This is a 2-category in the sense that we have been using it, as composition of 1-morphisms in $\mathscr{C}$ is strictly associative. We denote supercategories by $A, B, \ldots$ and objects of supercategories by $a, b, \ldots$, to maintain consistency with the rest of this section.

Let A and B be objects of $\mathscr{C}$. We let $AB = A \otimes B$ be the supercategory $A \boxtimes B$ defined in Section 2.5. For objects $a \in A$ and $b \in B$, we write $a \otimes b$, or simply ab, instead of $a \boxtimes b$ in this section. We equip $A \otimes B$ with the Π-structure $\Pi(a \otimes b) = \Pi(a) \otimes b$. Thus, $A \otimes B$ is again an object of $\mathscr{C}$. This defines a monoidal structure on $\mathscr{C}$. However, it is *not* semistrict, since this product is not strictly associative. We will ignore this detail, and simply treat $\otimes$ as if it were semistrict. This is the main reason for the disclaimer at the start of this section that our discussion would not be entirely rigorous.

In the remainder of this section, we investigate how Σ and T (as defined in Section 3.2) define braidings (or syllepses or symmetries) on $\mathscr{C}$. We focus on the T case since it is most relevant to this paper, but the Σ case is very similar. To maintain consistency with the rest of this section, we write R in place of T. Thus, for objects A and B of $\mathscr{C}$, we have the equivalence $R_{A,B} : A \otimes B \to B \otimes A$ defined on objects by $R_{A,B}(a \otimes b) = \Pi^{|a||b|}(b) \otimes a$ and on morphisms by $R_{A,B}(f \otimes g) = (-1)^{|f||g|}\Pi^{|a||b|}(g) \otimes f$.

4.5. ***Braiding.*** To give R the structure of a braiding, we first must define the ρ isomorphisms. Fix functions $\omega_1, \omega_2 : (\mathbb{Z}/q)^3 \to \boldsymbol{k}^\times$. Let A, B and C be objects of $\mathscr{C}$, and let a, b and c be homogeneous objects of A, B and C. We have

$$((BR_{A,C})(R_{A,B}C))(a \otimes b \otimes c) = \Pi^{|a||b|}(b) \otimes \Pi^{|a||c|}(c) \otimes a$$

and

$$R_{A,BC}(a \otimes b \otimes c) = \Pi^{|a|(|b|+|c|)}(b) \otimes c \otimes a.$$

We define

$$\rho_{A|B,C} : \Pi^{|a||b|}(b) \otimes \Pi^{|a||c|}(c) \otimes a \to \Pi^{|a|(|b|+|c|)}(b) \otimes c \otimes a$$

by

$$\rho_{A|B,C} = \omega_1(a; b, c) \cdot \xi \otimes \xi \otimes 1,$$

where we write $\omega_1(a; b, c)$ in place of $\omega_1(|a|; |b|, |c|)$. Here $\xi^{n,m} : \Pi^n \to \Pi^m$ is the isomorphism defined in Section 2.4, and we omit the indices to ξ when they are clear. Similarly, we have

$$((R_{A,C}B)(AR_{B,C}))(a \otimes b \otimes c) = \Pi^{|a||c|}(\Pi^{|b||c|}(c)) \otimes a \otimes b$$

and

$$R_{AB,C}(a \otimes b \otimes c) = \Pi^{|c|(|a|+|b|)}(c) \otimes a \otimes b.$$

We define $\rho_{A,B|C}$ to be $\omega_2(a,b;c)$ times the identity map; note that $\Pi^{|a||c|} \circ \Pi^{|b||c|} = \Pi^{|c|(|a|+|b|)}$. We have thus defined the necessary data for a braiding on $\mathscr{C}$. We now examine the axioms.

Axiom 1: The diagram becomes

$$\begin{array}{ccc} \Pi^{|a||b|}(b)\Pi^{|a||c|}(c)\Pi^{|a||d|}(d)a & \xrightarrow{\omega_1(a;c,d)1\otimes\xi\otimes\xi\otimes 1} & \Pi^{|a||b|}(b)\Pi^{|a|(|c|+|d|)}(c)da \\ \downarrow{\scriptstyle \omega_1(a;b,c)\xi\otimes\xi\otimes 1\otimes 1} & & \downarrow{\scriptstyle \omega_1(a;b,cd)\xi\otimes\xi\otimes 1\otimes 1} \\ \Pi^{|a|(|b|+|c|)}(b)c\Pi^{|a||d|}(d)a & \xrightarrow{\omega_1(a;bc,d)\xi\otimes 1\otimes\xi\otimes 1} & \Pi^{|a|(|b|+|c|+|d|)}(b)cda \end{array}$$

The top composition is $\omega_1(a;b,cd)\omega_1(a;c,d)\xi\otimes\xi\otimes\xi\otimes 1$ while the bottom composition is

$$\omega_1(a;b,c)\omega_1(a;bc,d)\xi\otimes\xi\otimes\xi\otimes 1.$$

Hence, the diagram commutes if and only if

$$\omega_1(a;b,c)\omega_1(a;bc,d) = \omega_1(a;b,cd)\omega_1(a;c,d),$$

which is equivalent to (3.5a) (note we have multiplicative notation as opposed to additive notation here since we are using the objects rather than their degrees as input).

Axiom 2: The diagram becomes

$$\begin{array}{ccc} \Pi^{|a||d|}\big(\Pi^{|b||d|}(\Pi^{|c||d|}(d))\big)abc & \xrightarrow{\omega_2(a,b;d)} & \Pi^{(|a|+|b|)|d|}(\Pi^{|c||d|}(d))abc \\ \downarrow{\scriptstyle \omega_2(b,c;d)} & & \downarrow{\scriptstyle \omega_2(ab,c;d)} \\ \Pi^{|a||d|}(\Pi^{(|b|+|c|)|d|}(d))abc & \xrightarrow{\omega_2(a,bc;d)} & \Pi^{(|a|+|b|+|c|)|d|}(d)abc \end{array}$$

where each map is a scalar times the identity. Thus, commutativity of this diagram is equivalent to

$$\omega_2(ab,c;d)\omega_2(a,b;d) = \omega_2(b,c;d)\omega_2(a,bc;d),$$

which is (3.5b).

Axiom 3: The first diagram comes out to

$$\begin{array}{ccc} \Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|a||d|}(\Pi^{|b||d|}(d))ab & \xrightarrow{①} & \Pi^{|a||c|}(\Pi^{|b|(|c|+|d|)}(c))\Pi^{|a||d|}(d)ab \\ \downarrow{\scriptstyle ②} & \overset{\varphi}{\searrow} & \downarrow{\scriptstyle ③} \\ \Pi^{|a|(|c|+|d|)}(\Pi^{|b||c|}(c))\Pi^{|b||d|}(d)ab & \xrightarrow{④} & \Pi^{|a|(|c|+|d|)}(\Pi^{|b|(|c|+|d|)}(c))dab \end{array}$$

The four numbered maps are given by

$$\begin{aligned} ① &= (-1)^{|a||b||d|(|c|+1)}\omega_1(b;c,d)\xi\otimes\xi\otimes 1\otimes 1, \\ ② &= \omega_1(a;c,d)\xi\otimes\xi\otimes 1\otimes 1, \\ ③ &= \omega_1(a;c,d)\xi\otimes\xi\otimes 1\otimes 1, \\ ④ &= (-1)^{|a||b||d|(|c|+1)}\omega_1(b;c,d)\xi\otimes\xi\otimes 1\otimes 1. \end{aligned}$$

To compute the signs, we use the identities $\Pi^k(\xi^{n,m}) = (-1)^{k(n+m)}\xi^{n+k,m+k}$ and $\xi^{n,m}_{\Pi^k} = \xi^{n+k,m+k}$. We thus see that the square commutes, and

$$\varphi = (-1)^{|a||b||c||d|}\omega_1(b;c,d)\omega_1(a;c,d)\xi\otimes\xi\otimes 1\otimes 1.$$

Note that one has to apply the sign rule when computing the compositions, and this cancels the $(-1)^{|a||b||d|}$ part of the sign. The second diagram is

$$\begin{array}{ccc}
\Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|a||d|}(\Pi^{|b||d|}(d))ab & \xrightarrow{\omega_2(a,b;d)} & \Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|d|(|a|+|b|)}(d)ab \\
\Big\downarrow{\scriptstyle \omega_2(a,b;c)} & \overset{\psi}{\searrow} & \Big\downarrow{\scriptstyle \omega_2(a,b;c)} \\
\Pi^{|c|(|a|+|b|)}(c)\Pi^{|a||d|}(\Pi^{|b||d|}(d))ab & \xrightarrow{\omega_2(a,b;d)} & \Pi^{|c|(|a|+|b|)}(c)\Pi^{|d|(|a|+|b|)}(d)ab
\end{array}$$

All maps are scalar multiples of the identity, so the diagram commutes and

$$\psi = \omega_2(a,b;d)\omega_2(a,b;c)\cdot\mathrm{id}.$$

Finally, we have the third diagram. It is

$$\begin{array}{ccc}
\Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|a||d|}(\Pi^{|b||d|}(d))ab & \xrightarrow{\mathrm{id}} & \Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|a||d|}(\Pi^{|b||d|}(d))ab \\
\Big\downarrow{\scriptstyle \varphi} & & \Big\downarrow{\scriptstyle \psi} \\
\Pi^{|a|(|c|+|d|)}(\Pi^{|b|(|c|+|d|)}(c))dab & & \Pi^{|c|(|a|+|b|)}(c)\Pi^{|d|(|a|+|b|)}(d)ab \\
& {\scriptstyle \omega_2(a,b;cd)\,\mathrm{id}}\searrow \quad \swarrow{\scriptstyle \omega_1(ab;c,d)\xi\xi 11} & \\
& \Pi^{(|a|+|b|)(|c|+|d|)}(c)dab &
\end{array}$$

(We have omitted $\otimes$ symbols in the bottom right map.) This diagram commutes if and only if (3.5c) holds.

Axiom 4: The diagram becomes

$$\begin{array}{ccccc}
\Pi^{|c|(|a|+|b|)}(c)\Pi^{|a||b|}(b)a & \xleftarrow{\omega_2(b,a;c)} & \Pi^{|b||c|}(\Pi^{|a||c|}(c))\Pi^{|a||b|}(b)a & \longrightarrow & \Pi^{|b||c|}(c)\Pi^{|a|(|b|+|c|)}(b)a \\
\| & & & & \Big\downarrow{\scriptstyle (-1)^{|a||b||c|}\xi_c\otimes\xi\otimes 1} \\
\Pi^{(|a|+|b|)|c|}(c)\Pi^{|a||b|}(b)a & \xrightarrow{\omega_2(a,b;c)^{-1}} & \Pi^{|a||c|}(\Pi^{|b||c|}(c))\Pi^{|a||b|}(b)a & \longleftarrow & \Pi^{|a|(|b|+|c|)}(\Pi^{|b||c|}(c))ba
\end{array}$$

where the top right map is $(-1)^{|a||c|(|b|+1)}\omega_1(a;b,c)\xi_c\otimes\xi\otimes 1$ and the bottom right map is

$$(-1)^{|a||b|(|c|+1)}\omega_1(a;c,b)^{-1}\xi_c^{-1}\otimes\xi^{-1}\otimes 1.$$

Taking into account signs from composing tensor products of maps, the right-side composition is $\omega_1(a;b,c)\omega_1(a;c,b)^{-1}$ times the identity, and hence the commutativity of this diagram is equivalent to (3.5d).

We thus see that R and ρ define a braiding if and only if (ω_1,ω_2) is a B-factor system. In particular, we obtain braidings using the systems defined in Examples 3.26 and 3.27.

4.6. ***Symmetry.*** Now we attempt to define a syllepsis on R. Fix a function $\omega_\sharp : (\mathbb{Z}/q)^2 \to \boldsymbol{k}^\times$. We have

$$R_{B,A}(R_{A,B}(a \otimes b)) = \Pi^{|a||b|}(a) \otimes \Pi^{|a||b|}(b).$$

We define $v_{A,B} : R_{B,A}R_{A,B} \to \mathrm{id}$ to be the isomorphism $\omega_\sharp(a,b)\xi \otimes \xi$ on an object $a \otimes b$. We check the commutativity of the four diagrams in order. The first is

$$\begin{array}{ccc}
\Pi^{|a||b|}(\Pi^{(|a||c|)}(a))\Pi^{|a||b|}(b)\Pi^{|a||c|}(c) & \xrightarrow{\omega_1(a;b,c)1\otimes\xi\otimes\xi} & \Pi^{|a||b|}(\Pi^{|a||c|}(a))\Pi^{|a|(|b|+|c|)}(b)c \\
\downarrow{\scriptstyle \omega_2(b,c;a)\,\mathrm{id}} & \overset{\varphi}{\searrow} & \downarrow{\scriptstyle \omega_2(b,c;a)\,\mathrm{id}} \\
\Pi^{|a|(|b|+|c|)}(a)\Pi^{|a||b|}(b)\Pi^{|a||c|}(c) & \xrightarrow{\omega_1(a;b,c)1\otimes\xi\otimes\xi} & \Pi^{|a|(|b|+|c|)}(a)\Pi^{|a|(|b|+|c|)}(b)c
\end{array}$$

This formally commutes, and so

$$\varphi = \omega_1(a;b,c)\omega_2(b,c;a)1 \otimes \xi \otimes \xi.$$

The second diagram is

$$\begin{array}{ccc}
\Pi^{|a||b|}(\Pi^{|a||c|}(a))\Pi^{|a||b|}(b)\Pi^{|a||c|}(c) & \xrightarrow{\varphi} & \Pi^{|a|(|b|+|c|)}(a)\Pi^{|a|(|b|+|c|)}(b)c \\
\downarrow{\scriptstyle (-1)^{|a||b||c|}\omega_\sharp(a,c)\xi\otimes 1\otimes\xi} & & \downarrow{\scriptstyle \omega_\sharp(a,bc)\xi\otimes\xi\otimes 1} \\
\Pi^{|a||b|}(a)\Pi^{|a||b|}(b)c & \xrightarrow{\omega_\sharp(a,b)\xi\otimes\xi\otimes 1} & abc
\end{array}$$

This diagram commutes if and only if

$$\omega_1(a;b,c)\omega_2(b,c;a)\omega_\sharp(a,bc) = \omega_\sharp(a,b)\omega_\sharp(a,c),$$

which is equivalent to (3.5e). The third diagram is

$$\begin{array}{ccc}
\Pi^{|a||c|}(a)\Pi^{|b||c|}(b)\Pi^{|a||c|}(\Pi^{|b||c|}(c)) & \xrightarrow{\omega_2(a,b;c)\,\mathrm{id}} & \Pi^{|a||c|}(a)\Pi^{|b||c|}(b)\Pi^{(|a|+|b|)|c|}(c) \\
\downarrow{\scriptstyle \omega_1(c;a,b)\xi\otimes\xi\otimes 1} & \overset{\psi}{\searrow} & \downarrow{\scriptstyle \omega_1(c;a,b)\xi\otimes\xi\otimes 1} \\
\Pi^{(|a|+|b|)|c|}(a)b\Pi^{|a||c|}(\Pi^{|b||c|}(c)) & \xrightarrow{\omega_2(a,b;c)\,\mathrm{id}} & \Pi^{(|a|+|b|)|c|}(a)b\Pi^{(|a|+|b|)|c|}(c)
\end{array}$$

This formally commutes, and we find that

$$\psi = \omega_1(c;a,b)\omega_2(a,b;c)\xi \otimes \xi \otimes 1.$$

The fourth diagram is

$$\begin{array}{ccc}
\Pi^{|a||c|}(a)\Pi^{|b||c|}(b)\Pi^{|a||c|}(\Pi^{|b||c|}(c)) & \xrightarrow{\psi} & \Pi^{(|a|+|b|)|c|}(a)b\Pi^{(|a|+|b|)|c|}(c) \\
\downarrow{\scriptstyle \omega_\sharp(a,c)\xi\otimes 1\otimes\xi} & & \downarrow{\scriptstyle \omega_\sharp(ab,c)\xi\otimes 1\otimes\xi} \\
a\Pi^{|b||c|}(b)\Pi^{|b||c|}(c) & \xrightarrow{\omega_\sharp(b,c)1\otimes\xi\otimes\xi} & abc
\end{array}$$

This diagram commutes if and only if

$$\omega_1(c; a, b)\omega_2(a, b; c)\omega_\sharp(ab, c) = \omega_\sharp(a, c)\omega_\sharp(b, c),$$

which is equivalent to (3.5f). Finally, the two isomorphisms

$$(R_{A,B}R_{B,A}R_{A,B})(ab) = \Pi^{2|a||b|}(b)\Pi^{|a||b|}(a) \to \Pi^{|a||b|}(b)a = R_{A,B}(ab)$$

are

$$1 \circ v_{A,B} = \omega_\sharp(a, b)\xi \otimes \xi, \quad v_{B,A} \circ 1 = \omega_\sharp(b, a)\xi \otimes \xi.$$

These agree if and only if $\omega_\sharp(a, b) = \omega_\sharp(b, a)$, i.e., γ is symmetric. We thus see that v defines a syllepsis if and only if $(\omega_1, \omega_2, \omega_\sharp)$ is an S-factor system, and it defines a symmetry if and only if, additionally, $\omega_\sharp$ is symmetric.

We thus obtain symmetries from the systems in Examples 3.26 and 3.27.

Remark 4.1. A variant of the above discussion shows that even factor systems can be used to endow Σ with a braided, sylleptic or symmetric structure. □

5. Spin-symmetric groups

5.1. ***S-sets and s-groups.*** An *s-set* is a triple $(X, |\cdot|, c)$ where X is a set, $|\cdot| : X \to \mathbb{Z}/2$ is a function (the degree function), and c is an involution of X such that $|cx| = |x|$ for all $x \in X$. We write X_i for the set of elements of parity i.

Let X and Y be s-sets. A *homogeneous map* $X \to Y$ of degree n is a function that maps X_i into Y_{i+n} and is compatible with c. We let $\mathrm{Hom}(X, Y)_i$ denote the set of maps of degree i, and let $\mathrm{Hom}(X, Y) = \mathrm{Hom}(X, Y)_0 \amalg \mathrm{Hom}(X, Y)_1$, which is naturally an s-set. We thus have a category **SSet** of s-sets. We let **SSet**$^\circ$ be the subcategory where the morphisms are even (degree 0).

Let X and Y be s-sets. We define the *product* of X and Y, denoted by $X \divideontimes Y$, as follows. The underlying set is the quotient of the usual cartesian product $X \times Y$ by the involution $c \times c$. That is, elements of $X \divideontimes Y$ are represented by ordered pairs (x, y), and the pairs (x, y) and (cx, cy) represent the same element. The degree function is given by $|(x, y)| = |x| + |y|$. The involution c is given by $c(x, y) = (cx, y) = (x, cy)$. We define an isomorphism $X \divideontimes Y \to Y \divideontimes X$ by $(x, y) \mapsto c^{|x||y|}(y, x)$. The product $\divideontimes$ gives **SSet**$^\circ$ the structure of a symmetric monoidal category.

An *s-group* is a group object in the monoidal category **SSet**$^\circ$. Explicitly, an s-group is a triple $(G, |\cdot|, c)$ where G is a group, $|\cdot| : G \to \mathbb{Z}/2$ is a group homomorphism, and $c \in G$ is a central element satisfying $c^2 = 1$ and $|c| = 0$. A *homomorphism* of s-groups $f : G \to H$ is a group homomorphism such that $f(c) = c$ and $|f(g)| = |g|$ for all $g \in G$.

If G and H are s-groups, then $G \divideontimes H$ is naturally an s-group, with multiplication given by $(g, h)(g', h') = c^{|h||g'|}(gg', hh')$. Two elements g and h of an s-group G are said to *supercommute* if $gh = c^{|h||g|}hg$. Similarly, two subsets A and B of G are said to supercommute if every element of A supercommutes with

every element of B. If H_1 and H_2 are supercommuting s-subgroups of G, then there is a homomorphism $H_1 \rtimes H_2 \to G$ given by $(h_1, h_2) \mapsto h_1 h_2$.

Example 5.1. Let X be an object in a supercategory, and let $\mathrm{Aut}^h(X)$ be the group of all homogeneous automorphisms of X. Then $\mathrm{Aut}^h(X)$ is an s-group, taking $c = -1$ and the degree map to be the usual degree map. For a super-vector space V, we write $\mathbf{GL}^h(V)$ in place of $\mathrm{Aut}^h(V)$. □

5.2. ***Spin-symmetric groups.*** Recall that the braid group $\mathfrak{B}_n$ is generated by elements $s_1, \dots, s_{n-1}$ modulo

- the braid relations $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$ for $1 \le i \le n-2$; and
- the commutativity relations $s_i s_j = s_j s_i$ for $|i-j| \ge 2$.

The symmetric group $\mathfrak{S}_n$ is isomorphic to the quotient of $\mathfrak{B}_n$ by the relations $s_i^2 = 1$; the isomorphism takes s_i to the transposition $(i\ i{+}1)$.

For $\delta \in \mathbb{Z}/2$, we define $\widetilde{\mathfrak{B}}_n^\delta$ to be the group generated by elements $\tilde{s}_1, \dots, \tilde{s}_{n-1}, c$ modulo

- the element c is central and $c^2 = 1$;
- the braid relations $\tilde{s}_i \tilde{s}_{i+1} \tilde{s}_i = \tilde{s}_{i+1} \tilde{s}_i \tilde{s}_{i+1}$ for $1 \le i \le n-2$; and
- the (super)commutativity relations $\tilde{s}_i \tilde{s}_j = c^\delta \tilde{s}_j \tilde{s}_i$ for $|i-j| \ge 2$.

For $\epsilon \in \mathbb{Z}/2$, define $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)}$ to be the quotient of $\widetilde{\mathfrak{B}}_n^\delta$ by the relations $\tilde{s}_i^2 = c^\epsilon$. It is well known that $c \ne 1$ in $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)}$ (for $\delta = 1$ and $\epsilon = 0$, this follows from Proposition 8.1, for instance), and thus $c \ne 1$ in $\widetilde{\mathfrak{B}}_n^\delta$ as well. We therefore have central extensions

$$1 \to \langle c \rangle \to \widetilde{\mathfrak{B}}_n^\delta \to \mathfrak{B}_n \to 1 \quad \text{and} \quad 1 \to \langle c \rangle \to \widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)} \to \mathfrak{S}_n \to 1,$$

where $\langle c \rangle$ is cyclic of order two.

We define the degree function on $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)}$ to be the composition $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)} \to \mathfrak{S}_n \xrightarrow{\ell} \mathbb{Z}/2$, where ℓ is the length modulo 2 of a permutation, i.e., 0 for even permutations and 1 for odd permutations. In this way, $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)}$ has the structure of an s-group. We let $\widetilde{\mathfrak{S}}_n = \widetilde{\mathfrak{S}}_n^{(1,0)}$, and refer to this as the *spin-symmetric group*.

Remark 5.2. For $n = 0, 1$, the central extension $\widetilde{\mathfrak{S}}_n^{(\delta,\epsilon)}$ is trivial for all ϵ and δ, while, for $n = 2, 3$, it is independent of δ. For $n \ge 4$, the four central extensions are nonisomorphic, and the group $\mathrm{H}^2(\mathfrak{S}_n, \mathbb{Z}/2)$ of central extensions is isomorphic to $(\mathbb{Z}/2)^2$ via (δ, ϵ).

The calculation of $\mathrm{H}^2(\mathfrak{S}_n, \mathbb{Z}/2)$ can be done as follows. First, $\mathrm{H}_1(\mathfrak{S}_n, \mathbb{Z})$ is the abelianization of $\mathfrak{S}_n$, which is just $\mathbb{Z}/2$ for $n \ge 2$ and trivial otherwise. Next, $\mathrm{H}_2(\mathfrak{S}_n, \mathbb{Z})$ is dual to $\mathrm{H}^2(\mathfrak{S}_n, \mathbb{C}^\times)$, which is $\mathbb{Z}/2$ for $n \ge 4$ and trivial otherwise [Hoffman and Humphreys 1992, Theorems 2.7 and 2.9]. Our earlier claim then follows from the universal coefficients theorem. □

5.3. ***The $\tilde{\tau}$ elements.*** Let $\tau_{n,m} \in \mathfrak{S}_{n+m}$ be the permutation defined by

$$\tau_{n,m}(i) = \begin{cases} i+m & \text{if } 1 \le i \le n, \\ i-n & \text{if } n+1 \le i \le n+m. \end{cases}$$

In a sense, $\tau_{n,m}$ induces the natural bijection of sets $[n] \sqcup [m] \to [m] \sqcup [n]$. Let $\sigma_{n,i} \in \mathfrak{S}_{n+m}$ be the $(n+1)$-cycle $(i\ i+1 \cdots i+n)$. Then

$$\sigma_{n,i} = s_i s_{i+1} \cdots s_{i+n-1}, \quad \tau_{n,m} = \sigma_{n,m} \cdots \sigma_{n,1}.$$

We define

$$\tilde{\sigma}_{n,i} = \tilde{s}_i \tilde{s}_{i+1} \cdots \tilde{s}_{i+n-1}, \quad \tilde{\tau}_{n,m} = \tilde{\sigma}_{n,m} \cdots \tilde{\sigma}_{n,1}$$

as elements of $\widetilde{\mathfrak{S}}_{n+m}$. The $\tilde{\tau}_{n,m}$ elements are very important, and are essentially responsible for all examples of supersymmetric monoidal structures that we write down.

Proposition 5.3. *We have* $\tilde{\tau}_{n,m}\tilde{\tau}_{m,n} = c^N$ *with* $N = \binom{n}{2}\binom{m}{2}$.

Proof. For any i, j, k, we have

$$\tilde{\sigma}_{i,k}\tilde{\sigma}_{j,i+k-1} = (\tilde{s}_k \tilde{s}_{k+1} \cdots \tilde{s}_{i+k-1})(\tilde{s}_{i+k-1}\tilde{s}_{i+k} \cdots \tilde{s}_{i+j+k-2}) = c^{(i-1)(j-1)}\tilde{\sigma}_{j-1,i+k}\tilde{\sigma}_{i-1,k},$$

where we have canceled the $\tilde{s}_{i+k-1}$ and commuted the resulting $\tilde{\sigma}_{i-1,k}$ past $\tilde{\sigma}_{j-1,i+k}$. By applying this relation successively, this implies that

$$\tilde{\sigma}_{i,k}\tilde{\sigma}_{j,i+k-1}\tilde{\sigma}_{j,i+k-2} \cdots \tilde{\sigma}_{j,i} = c^{\binom{i}{2}(j-1)}\tilde{\sigma}_{j-1,i+k}\tilde{\sigma}_{j-1,i+k-1} \cdots \tilde{\sigma}_{j-1,i+1}.$$

Finally, we successively use this relation to get

$$\begin{aligned}
\tilde{\tau}_{n,m}\tilde{\tau}_{m,n} &= (\tilde{\sigma}_{n,m}\tilde{\sigma}_{n,m-1} \cdots \tilde{\sigma}_{n,1})(\tilde{\sigma}_{m,n}\tilde{\sigma}_{m,n-1} \cdots \tilde{\sigma}_{m,1}) \\
&= c^{(m-1)\binom{n}{2}}(\tilde{\sigma}_{n,m}\tilde{\sigma}_{n,m-1} \cdots \tilde{\sigma}_{n,2})(\tilde{\sigma}_{m-1,n+1}\tilde{\sigma}_{m-1,n} \cdots \tilde{\sigma}_{m-1,1}) \\
&\ \ \vdots \\
&= c^{\binom{m}{2}\binom{n}{2}}.
\end{aligned}$$

□

There is a natural injective homomorphism of s-groups

$$j_{n,m} : \widetilde{\mathfrak{S}}_n \ltimes \widetilde{\mathfrak{S}}_m \to \widetilde{\mathfrak{S}}_{n+m} \tag{5.4}$$

defined by $(\tilde{s}_i, 1) \mapsto \tilde{s}_i$ and $(1, \tilde{s}_i) \mapsto \tilde{s}_{n+i}$. (To see that this map is injective, note that it is injective modulo $\langle c \rangle$, and so the kernel is contained in $\langle c \rangle$, but c itself is not in the kernel.)

Proposition 5.5. *For* $g \in \widetilde{\mathfrak{S}}_n$ *and* $h \in \widetilde{\mathfrak{S}}_m$, *we have*

$$\tilde{\tau}_{n,m}\, j_{n,m}(g,h)\, \tilde{\tau}_{n,m}^{-1} = c^{nm|g|+nm|h|+|g||h|}\, j_{m,n}(h,g).$$

Proof. First consider the case $h = 1$, $m = 1$ and $g = \tilde{s}_i$ with $1 \le i \le n-1$. Then $\tilde{\tau}_{n,1} = \tilde{\sigma}_{n,1} = \tilde{s}_1\tilde{s}_2 \cdots \tilde{s}_n$ and hence

$$\tilde{\tau}_{n,1}\tilde{s}_i\tilde{\tau}_{n,1}^{-1} = \tilde{s}_1\tilde{s}_2 \cdots \tilde{s}_n\tilde{s}_i\tilde{s}_n\tilde{s}_{n-1} \cdots \tilde{s}_1 = c^{n-i-1}\tilde{s}_1\tilde{s}_2 \cdots (\tilde{s}_{i+1}\tilde{s}_i\tilde{s}_{i+1}) \cdots \tilde{s}_1 = c^{n-2}\tilde{s}_{i+1} = c^n\tilde{s}_{i+1},$$

where in the second equality we used the commuting relations $\tilde{s}_j\tilde{s}_i = c\tilde{s}_i\tilde{s}_j$ for $j = n, n-1, \ldots, i+2$, and in the third equality we used the braid relation $\tilde{s}_{i+1}\tilde{s}_i\tilde{s}_{i+1} = \tilde{s}_i\tilde{s}_{i+1}\tilde{s}_i$ and then again the commuting relations for $j = i-1, i-2, \ldots, 1$. Since conjugation is a group homomorphism, it follows immediately

that $\tilde{\tau}_{n,1} j_{n,m}(g,1)\tilde{\tau}_{n,1}^{-1} = c^{n|g|} j_{m,n}(1,g)$ for any $g \in \widetilde{\mathfrak{S}}_n$ and hence $\tilde{\tau}_{n,m} j_{n,m}(g,1)\tilde{\tau}_{n,m}^{-1} = c^{nm|g|} j_{m,n}(1,g)$. Similarly, $\tilde{\tau}_{n,m} j_{n,m}(1,h)\tilde{\tau}_{n,m}^{-1} = c^{nm|h|} j_{m,n}(h,1)$. Finally, we have

$$\begin{aligned}\tilde{\tau}_{n,m} j_{n,m}(g,h)\tilde{\tau}_{n,m}^{-1} &= \tilde{\tau}_{n,m} j_{n,m}(g,1) j_{n,m}(1,h)\tilde{\tau}_{n,m}^{-1} \\ &= c^{nm|g|+nm|h|} j_{m,n}(1,g) j_{m,n}(h,1) \\ &= c^{nm|g|+nm|h|+|g||h|} j_{m,n}(h,g). \qquad \square\end{aligned}$$

Proposition 5.6. *For any m, n, p, we have $j_{n,m+p}(1,\tilde{\tau}_{m,p}) j_{n+m,p}(\tilde{\tau}_{m,n},1) = \tilde{\tau}_{m,n+p}$.*

Proof. By definition, we have $j_{n,m+p}(1,\tilde{\sigma}_{m,i}) = \tilde{\sigma}_{m,n+i}$, so

$$\begin{aligned}j_{n,m+p}(1,\tilde{\tau}_{m,p}) j_{n+m,p}(\tilde{\tau}_{m,n},1) &= j_{n,m+p}(1,\tilde{\sigma}_{m,p}\cdots\tilde{\sigma}_{m,1}) j_{n+m,p}(\tilde{\sigma}_{m,n}\cdots\tilde{\sigma}_{m,1},1) \\ &= \tilde{\sigma}_{m,n+p}\cdots\tilde{\sigma}_{m,n+1}\tilde{\sigma}_{m,n}\cdots\tilde{\sigma}_{m,1} = \tilde{\tau}_{m,n+p}. \qquad \square\end{aligned}$$

5.4. ***The category $\tilde{\mathbf{S}}$.*** We define a homogeneous $\mathbb{Z}$-supercategory $\tilde{\mathbf{S}}$ as follows. There is a single homogeneous object $[n]$ of degree n for each $n \geq 0$, and no homogeneous object of negative degree. We put $\mathrm{End}_{\tilde{\mathbf{S}}}([n]) = \boldsymbol{k}[\widetilde{\mathfrak{S}}_n]/(c+1)$ and $\mathrm{Hom}_{\tilde{\mathbf{S}}}([n],[m]) = 0$ for $n \neq m$. We now define a monoidal structure $\otimes$ on $\tilde{\mathbf{S}}$. On objects, we define $[n]\otimes[m] = [n+m]$. We define

$$\otimes : \mathrm{End}_{\tilde{\mathbf{S}}}([n]) \otimes \mathrm{End}_{\tilde{\mathbf{S}}}([m]) \to \mathrm{End}_{\tilde{\mathbf{S}}}([n+m])$$

to be the linear map induced by $j_{n,m}$. The unit object is $[0]$. The monoidal structure is strict (e.g., the associator maps are the identities). One readily verifies that this satisfies the requisite axioms for a monoidal $\mathbb{Z}$-supercategory.

For $n, m \in \mathbb{N}$, define $\beta_{n,m} : [n]\otimes[m] \to [m]\otimes[n]$ to be $\tilde{\tau}_{n,m}$.

Proposition 5.7. *β defines a type IIa supersymmetry on $\tilde{\mathbf{S}}$.*

Proof. Proposition 5.5 exactly shows that β is a natural transformation (see the first paragraph of Section 3.6). Proposition 5.6 exactly shows that the hexagon (H1) commutes. Proposition 5.3 shows that β is symmetric, i.e., $\beta_{m,n}\beta_{n,m} = (-1)^k$ with $k = \binom{n}{2}\binom{m}{2}$, which also implies the second hexagon axiom. $\square$

Now suppose that $\zeta_{16} \in \boldsymbol{k}$. For $n \in \mathbb{Z}$, define $\epsilon(n) = -1$ if $n \equiv 3 \pmod 4$ and $\epsilon(n) = 1$ otherwise. Define $\beta'_{n,m} : [n+m] \to [n+m]$ to be the map $\epsilon(n)^m \zeta_4^{-\binom{n}{2}m} \zeta_{16}^{nm} \tilde{\tau}_{n,m}$.

Proposition 5.8. *β' defines a type IIb supersymmetry on $\tilde{\mathbf{S}}$.*

Proof. This follows from Proposition 3.15. $\square$

Let q be a positive integer. We can then regard $\tilde{\mathbf{S}}$ as $\mathbb{Z}/q$-graded, and the monoidal structure carries over. If q is divisible by 4, then β defines a type IIa supersymmetry, while, if q is any even integer, then β' defines a type IIb supersymmetry (the key point here is that the scalar $\beta'_{n,m}\beta'_{m,n}$ only depends on the parities of n and m; this is the purpose of the somewhat complicated scalar factor in the definition of β').

Remark 5.9. In the category $\tilde{\mathbf{S}}$, the object $[n]$ admits an odd-degree automorphism if and only if $n \geq 2$. Thus, $\tilde{\mathbf{S}}$ does not admit a Π-structure, since the objects $[0]$ and $[1]$ do not. $\square$

Remark 5.10. In our discussion of supersymmetric monoidal categories, we have worked exclusively with linear categories since the examples we care about are linear. However, this is not necessary. One can define a "nonlinear supercategory" to be a category enriched in **SSet**$^\circ$, and then go on to define nonlinear analogs of many of the other notions we have discussed. The category $\tilde{\mathcal{S}}$ defined above is in fact the linearization of what we would call the groupoid of spin-sets, which is likely an object of fundamental importance. □

5.5. ***Actions of spin-symmetric groups.*** Let $\mathcal{A}$ be a type IIa supersymmetric monoidal $\mathbb{Z}/q$-supercategory with $4 \mid q$. In what follows, we assume that $\mathcal{A}$ is strict (i.e., the associativity isomorphisms are identity maps) for simplicity. (In fact, this comes at no loss of generality due to a version of Mac Lane's coherence theorem in this setting; see the discussion following Definition 1.4 in [Brundan and Ellis 2017].) Let X be a homogeneous object of degree $p \in \mathbb{Z}/q$ and fix a positive integer n. For $1 \le i \le n-1$, define $\sigma_i : X^{\otimes n} \to X^{\otimes n}$ to be the map

$$X^{\otimes n} = X^{\otimes(i-1)} \otimes X^{\otimes 2} \otimes X^{\otimes(n-i-1)} \xrightarrow{\mathrm{id}\otimes\beta_{X,X}\otimes\mathrm{id}} X^{\otimes(i-1)} \otimes X^{\otimes 2} \otimes X^{\otimes(n-i-1)} = X^{\otimes n}.$$

Note that σ_i has degree p (since $p^2 = p$ modulo 2). We then have:

Proposition 5.11. *The automorphisms σ satisfy the following relations*:

(a) $\sigma_i\sigma_j = (-1)^p\sigma_j\sigma_i$ *for* $|i-j| > 1$.

(b) $\sigma_i\sigma_{i+1}\sigma_i = (-1)^p\sigma_{i+1}\sigma_i\sigma_{i+1}$.

(c) $\sigma_i^2 = (-1)^{\binom{p}{2}}$.

Proof. Relation (a) follows immediately from how composition works in product categories and the fact that each σ_i has degree p. We now verify (b). It suffices to treat the case $n = 3$. We must show that the following diagram commutes up to $(-1)^p$:

$$\begin{array}{ccccc} X\otimes X\otimes X & \xrightarrow{\sigma_2} & X\otimes X\otimes X & \xrightarrow{\sigma_1} & X\otimes X\otimes X \\ \downarrow{\scriptstyle\sigma_1} & & & & \downarrow{\scriptstyle\sigma_2} \\ X\otimes X\otimes X & \xrightarrow{\sigma_2} & X\otimes X\otimes X & \xrightarrow{\sigma_1} & X\otimes X\otimes X \end{array}$$

By the first hexagon axiom, $\sigma_1\sigma_2 = \beta_{X\otimes X,X}$. It thus suffices to show that the following diagram commutes up to $(-1)^p$:

$$\begin{array}{ccc} X\otimes X\otimes X & \xrightarrow{\beta_{X\otimes X,X}} & X\otimes X\otimes X \\ \downarrow{\scriptstyle\beta_{X,X}\otimes\mathrm{id}_X} & & \downarrow{\scriptstyle\mathrm{id}_X\otimes\beta_{X,X}} \\ X\otimes X\otimes X & \xrightarrow{\beta_{X\otimes X,X}} & X\otimes X\otimes X \end{array}$$

This follows from the definition of β (see Section 3.6) and the fact that $|X| = |\beta_{X,X}| = p$ and $|\mathrm{id}_X| = 0$. Finally, (c) follows from the definition of supersymmetry. □

We introduce the notation

$$\widetilde{\mathfrak{S}}_n^0 = \widetilde{\mathfrak{S}}_n^{(0,0)}, \quad \widetilde{\mathfrak{S}}_n^1 = \widetilde{\mathfrak{S}}_n^{(1,0)}, \quad \widetilde{\mathfrak{S}}_n^2 = \widetilde{\mathfrak{S}}_n^{(0,1)}, \quad \widetilde{\mathfrak{S}}_n^3 = \widetilde{\mathfrak{S}}_n^{(1,1)}.$$

The index p in $\widetilde{\mathfrak{S}}_n^p$ is regarded as an element of $\mathbb{Z}/4$. The above proposition shows that $\widetilde{\mathfrak{S}}_n^p$ acts on $X^{\otimes n}$ if we let the generator $\tilde{s}_i$ act by $(-1)^{ip}$ and c act by $(-1)^p$. In fact:

Proposition 5.12. *The group $\widetilde{\mathfrak{S}}_n^p$ naturally acts on the functor $\mathcal{A}_p \to \mathcal{A}_{pn}$ given by $X \mapsto X^{\otimes n}$.*

Proof. We just need to verify naturality. Given a homogeneous morphism $f : X \to Y$ between homogeneous objects, we get a morphism $f^{\otimes n} : X^{\otimes n} \to Y^{\otimes n}$, and we check that it supercommutes with the action of σ_i. We have

$$\begin{aligned}
\sigma_i f^{\otimes n} &= (1 \otimes \beta_{Y,Y} \otimes 1)(f^{\otimes i} \otimes f^{\otimes 2} \otimes f^{\otimes (i-2)}) \\
&= (-1)^{i|f||Y|}(f^{\otimes i} \otimes \beta_{Y,Y} f^{\otimes 2} \otimes f^{\otimes (i-2)}) \\
&= (-1)^{i|f||Y|+|f|}(f^{\otimes i} \otimes f^{\otimes 2} \beta_{X,X} \otimes f^{\otimes (i-2)}) \\
&= (-1)^{|f|}(f^{\otimes i} \otimes f^{\otimes 2} \otimes f^{\otimes (i-2)})(1 \otimes \beta_{X,X} \otimes 1) \\
&= (-1)^{|f|} f^{\otimes n} \sigma_i.
\end{aligned}$$

□

Remark 5.13. There are similar results to the above in the superbraiding case and the type IIb case. □

5.6. ***Universality of*** $\tilde{\mathbf{S}}$**.** Fix an integer q divisible by 4 (we allow $q = 0$). In this section, we regard the category $\tilde{\mathbf{S}}$ as a type IIa supersymmetric monoidal $\mathbb{Z}/q$-supercategory via the supersymmetry β defined in Section 5.4. The following theorem shows that $\tilde{\mathbf{S}}$ is the universal supersymmetric monoidal $\mathbb{Z}/q$-supercategory equipped with an object of degree 1:

Theorem 5.14. *Let $\mathcal{A}$ be a type IIa supersymmetric monoidal $\mathbb{Z}/q$-supercategory. Let $\mathcal{B}$ be the category of supersymmetric monoidal $\mathbb{Z}/q$-superfunctors $\tilde{\mathbf{S}} \to \mathcal{A}$. Then the functor $\Phi : \mathcal{B} \to \mathcal{A}_1^\circ$ given by $\Phi(F) = F([1])$ is an equivalence of $\boldsymbol{k}$-linear categories.*

Proof. Suppose that X is a degree 1 object of $\mathcal{A}$. We have seen that $\widetilde{\mathfrak{S}}_n$ naturally acts on $X^{\otimes n}$, and so we can build a functor $\Psi_X : \tilde{\mathbf{S}} \to \mathcal{A}$ by $\Psi_X([n]) = X^{\otimes n}$, which is easily seen to be a supersymmetric monoidal $\mathbb{Z}/q$-functor. The assignment $X \mapsto \Psi_X$ is functorial by Proposition 5.12. The resulting functor $\Psi : \mathcal{A}_1^\circ \to \mathcal{B}$ is easily seen to be quasi-inverse to Φ. □

Remark 5.15. Given $p \in \mathbb{Z}/q$, one can build a supersymmetric monoidal $\mathbb{Z}/q$-supercategory $\tilde{\mathbf{S}}^p$ using the $\widetilde{\mathfrak{S}}^p$ groups, specializing to $\tilde{\mathbf{S}}$ when $p = 1$. The category $\tilde{\mathbf{S}}^p$ is the universal supersymmetric monoidal $\mathbb{Z}/q$-category with an object of degree p. □

5.7. ***Type IIa′ supersymmetries.*** Let $\mathcal{A}$ be a type IIa monoidal $\mathbb{Z}/4$-supercategory. We have seen that $\mathfrak{S}_n^p$ acts on $X^{\otimes n}$ when X is homogeneous of degree $p \in \mathbb{Z}/4$. This may seem strange since it identifies $\mathrm{H}^2(\mathfrak{S}_n, \mathbb{Z}/2)$ with $\mathbb{Z}/4$ (as sets). In fact, this is due to a choice we made, namely the choice of factor system.

To make this clear, define a type IIa′ supersymmetry to be a type II supersymmetry with respect to ACD, where A, C and D are as in Examples 3.8, 3.10 and 3.11. Suppose $\mathcal{A}$ is a type IIa′ supersymmetric monoidal $\mathbb{Z}/4$-category and X is an object of degree $p \in \mathbb{Z}/4$. Then one can show that $\widetilde{\mathfrak{S}}_n^{-p}$ acts on $X^{\otimes n}$.

We note that, if $\zeta_4 \in \boldsymbol{k}$, then one can pass between type IIa and type IIa′ superbraidings; indeed, if β is a type IIa superbraiding, then, putting $\beta'_{X,Y} = \zeta_4^{|X||Y|}\beta_{X,Y}$, one finds that β' is a type IIa′ superbraiding. (The ζ_4 factor in this expression realizes CD as a coboundary.)

6. Linear spin-species

6.1. ***S-representations.*** Let G be an s-group. An *s-representation* of G on a superspace V is a homomorphism of s-groups $\rho : G \to \mathbf{GL}^h(V)$. More explicitly, an s-representation ρ associates to each $g \in G$ an automorphism $\rho(g)$ of V such that

(a) $\rho(c) = -1$,

(b) $\rho(g)$ is homogeneous of degree $\deg(g)$, and

(c) $\rho(gh) = \rho(g)\rho(h)$.

Note that an s-representation is simply a representation of G on V in the usual sense satisfying certain conditions. A *morphism* of s-representations $f : V \to W$ of parity p is a map of super-vector spaces of parity p such that $f(gv) = (-1)^{p|g|}gf(v)$ for all $v \in V$. We thus have an **SVec**-category $\mathrm{Rep}(G)$ of representations of G, which is abelian and admits a Π-structure (as we see below).

Let G and H be s-groups and let V and W be s-representations of them. We define an s-representation of $G \ltimes\!\!\!\gg H$ on $V \otimes W$ by

$$(g, h)(v \otimes w) = (-1)^{|v||h|}gv \otimes hw.$$

Let $j : G \ltimes\!\!\!\gg H \to H \ltimes\!\!\!\ll G$ be the usual isomorphism of s-groups, and regard $W \otimes V$ as an s-representation of $G \ltimes\!\!\!\gg H$ via j; explicitly,

$$(g, h)(w \otimes v) = (-1)^{|g||h|+|w||g|}hw \otimes gv.$$

Then the usual isomorphism $\tau : V \otimes W \to W \otimes V$ (as defined in Section 2.1) is an isomorphism of $G \ltimes\!\!\!\gg H$ s-representations.

Remark 6.1. When V and W are two s-representations of G, there is no natural s-representation of G on $V \otimes W$, due to the lack of a diagonal morphism $G \to G \ltimes\!\!\!\gg G$. However, one can regard $V \otimes W$ as a representation of the ordinary group G (or even $G/\langle c\rangle$) by $g(v \otimes w) = gv \otimes gw$. Note that there is no sign here, even if g and v are both of odd degree. This action of $G/\langle c\rangle$ is by even transformations. □

6.2. **Π-*structures.*** Let V be a super-vector space, and let $\pi \in \boldsymbol{k}[1]$ be the element $1 \in \boldsymbol{k}$. There are (at least) three natural choices of Π-structure on V:

- Let $\Pi_n(V) = V[1]$, and let $\xi_n : V \to \Pi_n(V)$ be the map $v \mapsto v$.
- Let $\Pi_\ell(V) = \boldsymbol{k}[1] \otimes V$, and let $\xi_\ell : V \to \Pi_\ell(V)$ be the map $v \mapsto \pi \otimes v$.
- Let $\Pi_r(V) = V \otimes \boldsymbol{k}[1]$, and let $\xi_r : V \to \Pi_r(V)$ be the map $v \mapsto (-1)^{|v|}v \otimes \pi$.

We refer to Π_ℓ and Π_r as the "left" and "right" Π-structures, and Π_n as the "neutral" Π-structure (since it does not favor left or right). The neutral Π-structure may seem to be the most natural choice, but there is a practical issue with it: if v is an element of V then it is also an element of $V[1]$, and thus the symbol v becomes ambiguous. The reason for the sign in the right Π-structure is so that the following diagram commutes:

$$\begin{array}{ccc} V & = & V \otimes \boldsymbol{k} \\ {\scriptstyle \xi_r}\downarrow & & \downarrow{\scriptstyle \mathrm{id}\otimes\xi_r} \\ \Pi_r(V) & = & V \otimes \Pi_r(\boldsymbol{k}) \end{array}$$

Here $\xi_r : \boldsymbol{k} \to \Pi_r(\boldsymbol{k}) = \boldsymbol{k}[1]$ is the map $1 \mapsto \pi$. Since this is an odd map, $(\mathrm{id}\otimes\xi_r)(v\otimes 1) = (-1)^{|v|} v \otimes \pi$. The sign in ξ_r also ensures that the canonical even-degree isomorphism $\Pi_\ell(V) \to \Pi_r(V)$ is just the symmetry τ of the tensor product.

Suppose that V is an s-representation of an s-group G. We define s-representations of G on $\Pi_\ell(V) = \boldsymbol{k}[1]\otimes V$ and $\Pi_r(V) = V\otimes\boldsymbol{k}[1]$ by $g(\pi\otimes v) = (-1)^{|g|}\pi\otimes gv$ and $g(v\otimes\pi) = gv\otimes\pi$. One readily verifies that the maps $\xi_r : V \to \Pi_r(V)$ and $\xi_\ell : V \to \Pi_\ell(V)$ are odd-degree isomorphisms of s-representations. Furthermore, the canonical even-degree isomorphism $\Pi_\ell(V) \to \Pi_r(V)$ is one of s-representations. (We do not bother defining Π_n for s-representations.) We thus have two Π-structures for s-representations: Π_r and Π_ℓ.

6.3. ***Induction.*** Let $H \subseteq G$ be s-groups and let V be an s-representation of H. Consider the induced representation $W = \mathrm{Ind}_H^G(V) = \boldsymbol{k}[G] \otimes_{\boldsymbol{k}[H]} V$. We give W the structure of a super-vector space by declaring $g \otimes v$ to be homogeneous of degree $|g| + |v|$ whenever $v \in V$ is homogeneous. In this way, W is an s-representation of G. As usual, induction is the left adjoint of restriction.

Consider the following situation:

- H_1 and H_2 are s-subgroups of G.
- V_1 (resp. V_2) is an s-representation of H_1 (resp. H_2).
- τ is an element of G such that $\tau H_1 \tau^{-1} = H_2$.
- $i : V_1 \to V_2^\tau$ is an even-degree isomorphism of s-representations of H_1, where V_2^τ is the s-representation of H_1 on V_2 defined by $h \cdot v = (\tau h \tau^{-1})v$. In other words, $i(hv) = (\tau h \tau^{-1}) i(v)$ for $h \in H_1$.

Define

$$\varphi : \mathrm{Ind}_{H_1}^G(V_1) \to \Pi_r^{|\tau|} \mathrm{Ind}_{H_2}^G(V_2), \quad \varphi(g\otimes v) = g\tau^{-1} \otimes i(v) \otimes \pi^{|\tau|}.$$

We then have:

Proposition 6.2. *The map φ is an even-degree isomorphism of s-representations of G.*

Proof. We first check that φ is well defined, i.e., that $\varphi(gh\otimes v) = \varphi(g\otimes hv)$ for $h \in H_1$. Omitting the factors of π for economy, we have

$$\varphi(gh\otimes v) = gh\tau^{-1}\otimes i(v) = g\tau^{-1}\otimes \tau h\tau^{-1} i(v) = g\tau^{-1}\otimes i(hv) = \varphi(g\otimes hv).$$

Thus, φ is well defined. For $g' \in G$,

$$g'\varphi(g \otimes v) = g'(g\tau^{-1} \otimes i(v) \otimes \pi^{|\tau|}) = \varphi(g'g \otimes v),$$

and so φ is a map of s-representations. Finally, interchanging the roles of τ and τ^{-1} gives a map in the opposite direction that is the inverse of φ. Thus, φ is an isomorphism. □

6.4. ***The category of linear spin-species.*** A *representation of* $\widetilde{\mathfrak{S}}_*$ is a sequence $(V_n)_{n\geq 0}$, where V_n is an s-representation of the s-group $\widetilde{\mathfrak{S}}_n$ over $\boldsymbol{k}$. A *morphism* of $\widetilde{\mathfrak{S}}_*$-representations $V \to W$ is a sequence $f = (f_n)_{n\geq 0}$ where $f_n : V_n \to W_n$ is a (not necessarily homogeneous) morphism of $\widetilde{\mathfrak{S}}_n$-representations. We say that an $\widetilde{\mathfrak{S}}_*$-representation V is *even* (resp. *odd*) if $V_n = 0$ for all odd (resp. even) values of n. This gives $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ the structure of a supercategory. For an integer q, we can regard $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ to be an (inhomogeneous) $\mathbb{Z}/q$-supercategory by declaring an object $(V_n)_{n\geq 0}$ to have degree $d \in \mathbb{Z}/q$ if $V_n = 0$ for all $n \not\equiv d \pmod q$. Every object decomposes as a direct sum of homogeneous objects, and this decomposition is finite if $q \neq 0$.

6.5. ***Monoidal structure.*** Let n and m be nonnegative integers. Suppose that V and W are two representations of $\widetilde{\mathfrak{S}}_*$. We define $V \otimes W$ to be the representation of $\widetilde{\mathfrak{S}}_*$ given by

$$(V \otimes W)_n = \bigoplus_{i+j=n} \mathrm{Ind}_{\widetilde{\mathfrak{S}}_i \rtimes \widetilde{\mathfrak{S}}_j}^{\widetilde{\mathfrak{S}}_n} (V_i \otimes W_j),$$

where here $\widetilde{\mathfrak{S}}_i \rtimes \widetilde{\mathfrak{S}}_j$ is regarded as a subgroup of $\widetilde{\mathfrak{S}}_n$ via the map $j_{n,m}$ defined in (5.4). Given a third $\widetilde{\mathfrak{S}}_*$-representation U, we have a natural identification

$$((U \otimes V) \otimes W)_n = \bigoplus_{i+j+k=n} \mathrm{Ind}_{\widetilde{\mathfrak{S}}_i \rtimes \widetilde{\mathfrak{S}}_j \rtimes \widetilde{\mathfrak{S}}_k}^{\widetilde{\mathfrak{S}}_n} (U_i \otimes V_j \otimes W_k),$$

and similarly for $U \otimes (V \otimes W)$. This yields a natural associativity isomorphism $(U \otimes V) \otimes W \to U \otimes (V \otimes W)$ that satisfies the pentagon axiom. The $\widetilde{\mathfrak{S}}_*$-representation that is $\boldsymbol{k}$ in degree 0 and 0 in other degrees is naturally a unit object for $\otimes$. Thus, $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ has the structure of a monoidal $\mathbb{Z}/q$-supercategory.

6.6. ***Supersymmetry.*** In this section, we assume q is divisible by 4 and construct a supersymmetry on $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$. We begin with the following observation:

Proposition 6.3. *Let V and W be s-representations of $\widetilde{\mathfrak{S}}_n$ and $\widetilde{\mathfrak{S}}_m$. Regard $\widetilde{\mathfrak{S}}_n \rtimes \widetilde{\mathfrak{S}}_m$ as a subgroup of $\widetilde{\mathfrak{S}}_{n+m}$ via $j_{n,m}$, and similarly with n and m reversed. Then*

$$i_{V,W} : V \otimes W \to (W \otimes V)^{\tilde{\tau}_{n,m}}, \quad v \otimes w \mapsto (-1)^{|v||w|+nm|v|+nm|w|} w \otimes v,$$

is an isomorphism of representations of $\widetilde{\mathfrak{S}}_n \rtimes \widetilde{\mathfrak{S}}_m$, where the superscript denotes the conjugate representation. Precisely, this means

$$i_{V,W}(j_{n,m}(g,h)(v \otimes w)) = \tilde{\tau}_{n,m} j_{n,m}(g,h) \tilde{\tau}_{n,m}^{-1} i_{V,W}(v \otimes w)$$

for all $v \in V$ and $w \in W$ and all $g \in \widetilde{\mathfrak{S}}_n$ and $h \in \widetilde{\mathfrak{S}}_m$.

Proof. We have

$$\begin{aligned} i_{V,W}(j_{n,m}(g,h)(v\otimes w)) &= (-1)^{|h||v|} i_{V,W}(gv\otimes hw) \\ &= (-1)^{|h||v|+(|g|+|v|)(|h|+|w|)+nm(|g|+|v|)+nm(|h|+|w|)} hw\otimes gv \\ &= (-1)^{|g||h|+nm|g|+nm|h|+|v||w|+nm|v|+nm|w|+|g||w|} hw\otimes gv \\ &= \tilde{\tau}_{n,m} j_{n,m}(g,h)\tilde{\tau}_{n,m}^{-1} i_{V,W}(v\otimes w). \end{aligned}$$

In the last equality, we used Proposition 5.5 and that c acts by -1. □

As discussed in Section 6.3, we have an isomorphism

$$\operatorname{Ind}_{\widetilde{\mathfrak{S}}_n \rtimes \widetilde{\mathfrak{S}}_m}^{\widetilde{\mathfrak{S}}_{n+m}} (V\otimes W) \to \Pi_r^{nm} \operatorname{Ind}_{\widetilde{\mathfrak{S}}_m \rtimes \widetilde{\mathfrak{S}}_n}^{\widetilde{\mathfrak{S}}_{n+m}} (W\otimes V) \tag{6.4}$$

coming from conjugation by $\tilde{\tau}_{n,m}$, and the isomorphism $i_{V,W}$ (maintaining the notation from the proposition).

Definition 6.5. Given homogeneous objects V and W of $\operatorname{Rep}(\widetilde{\mathfrak{S}}_*)$, we define

$$\beta_{V,W} : V\otimes W \to \Pi_r^{|V||W|}(W\otimes V)$$

to be the isomorphism induced by the isomorphisms (6.4) using the elements $\tilde{\tau}_{n,m}$ and the isomorphisms i. Explicitly, for $v\in V_n$ and $w\in W_m$ and $g\in \widetilde{\mathfrak{S}}_{n+m}$, we have

$$\beta_{V,W}(g\otimes v\otimes w) = (-1)^{nm(|v|+|w|)+|v||w|} g\tilde{\tau}_{n,m}^{-1}\otimes w\otimes v\otimes \pi^{nm},$$

where here $g\otimes v\otimes w$ denotes an element of $\operatorname{Ind}_{\widetilde{\mathfrak{S}}_n \rtimes \widetilde{\mathfrak{S}}_m}^{\widetilde{\mathfrak{S}}_{n+m}} (V\otimes W)$. □

Proposition 6.6. *β defines a type Ia supersymmetry on $\otimes$.*

Proof. First, we check that β is a natural transformation. Let $f: V_n\to V_n'$ and $g: W_m\to W_m'$ be nonzero homogeneous morphisms. Pick $v\in V_n$, $w\in W_m$ and $\sigma\in\widetilde{\mathfrak{S}}_{n+m}$. Then

$$\begin{aligned} &\Pi^{|V||W|}(g\otimes f)\beta_{V_n,W_m}(\sigma\otimes v\otimes w) \\ &\qquad = (-1)^{nm(|v|+|w|)+|v||w|+|f|(|\sigma|+nm+|w|)+|g|(|\sigma|+nm)} \sigma\tilde{\tau}_{n,m}^{-1}\otimes g(w)\otimes f(v)\otimes \pi^{nm}. \end{aligned}$$

On the other hand,

$$\begin{aligned} &\beta_{V_n',W_m'}(f\otimes g)(\sigma\otimes v\otimes w) \\ &\qquad = \beta_{V_n',W_m'}((-1)^{|g|(|\sigma|+|v|)+|f||\sigma|}\sigma\otimes f(v)\otimes g(w)) \\ &\qquad = (-1)^{|g|(|\sigma|+|v|)+|f||\sigma|+nm(|f|+|v|+|g|+|w|)+(|f|+|v|)(|g|+|w|)} \sigma\tilde{\tau}_{n,m}^{-1}\otimes g(w)\otimes f(v)\otimes \pi^{nm}. \end{aligned}$$

This agrees with the first expression up to $(-1)^{|f||g|}$, which is what we wanted to show.

Next, it follows from Proposition 5.3 that $\beta_{V,W}\beta_{W,V} = (-1)^{\binom{n}{2}\binom{m}{2}}$. Hence, we only need to check that one of the hexagon axioms hold, and we check (H1′). So pick nonnegative integers m, n, p and let U, V, W be representations of $\widetilde{\mathfrak{S}}_m$, $\widetilde{\mathfrak{S}}_n$ and $\widetilde{\mathfrak{S}}_p$, respectively. We need to check that $(1_V\otimes\beta_{U,W})(\beta_{U,V}\otimes 1_W) = \beta_{U,V\otimes W}$. (Here we identify U with the $\mathfrak{S}_*$-representation that is U in degree m and 0 in other degrees, and similarly for V and W. We note that it suffices to check the axiom on such objects.)

The upper path in the hexagon is the composition

$$\begin{aligned}\mathrm{Ind}_{m+n,p}^{m+n+p}(\mathrm{Ind}_{m,n}^{m+n}(U\otimes V)\otimes W) &\xrightarrow{\alpha^{-1}_{U,V,W}} \mathrm{Ind}_{m,n+p}^{m+n+p}(U\otimes \mathrm{Ind}_{n,p}^{n+p}(V\otimes W))\\ &\xrightarrow{\beta_{U,V\otimes W}} \mathrm{Ind}_{n+p,m}^{m+n+p}(\mathrm{Ind}_{n,p}^{n+p}(V\otimes W)\otimes U)\otimes \boldsymbol{k}[m(n+p)]\\ &\xrightarrow{\alpha^{-1}_{V,W,U}} \mathrm{Ind}_{n,m+p}^{m+n+p}(V\otimes \mathrm{Ind}_{p,m}^{p+m}(W\otimes U))\otimes \boldsymbol{k}[m(n+p)].\end{aligned}$$

Here we have written n in place of $\widetilde{\mathfrak{S}}_n$ in the inductions for readability. We compute the composition explicitly. Let $g\in\widetilde{\mathfrak{S}}_{m+n+p}$ and let $u\in U$, $v\in V$ and $w\in W$ be homogeneous elements. We have

$$\begin{aligned}g\otimes(1\otimes u\otimes v)\otimes w &\mapsto g\otimes u\otimes(1\otimes v\otimes w)\\ &\mapsto (-1)^{m(n+p)(|u|+|v|+|w|)+|u|(|v|+|w|)} g\tilde{\tau}^{-1}_{m,n+p}\otimes(1\otimes v\otimes w)\otimes u\otimes\pi^{m(n+p)}\\ &\mapsto (-1)^{m(n+p)(|u|+|v|+|w|)+|u|(|v|+|w|)} g\tilde{\tau}^{-1}_{m,n+p}\otimes v\otimes(1\otimes w\otimes u)\otimes\pi^{m(n+p)}.\end{aligned}$$

The lower path in the hexagon is the composition

$$\begin{aligned}\mathrm{Ind}_{m+n,p}^{m+n+p}(\mathrm{Ind}_{m,n}^{m+n}(U\otimes V)\otimes W) &\xrightarrow{\beta_{U,V}\otimes \mathrm{id}_W} \mathrm{Ind}_{m+n,p}^{m+n+p}(\mathrm{Ind}_{n,m}^{m+n}(V\otimes W)\otimes \boldsymbol{k}[mn]\otimes U)\\ &\xrightarrow{\sim} \mathrm{Ind}_{m+n,p}^{m+n+p}(\mathrm{Ind}_{n,m}^{m+n}(V\otimes W)\otimes U)\otimes\boldsymbol{k}[mn] \xrightarrow{\alpha^{-1}_{V,W,U}} \mathrm{Ind}_{n,m+p}^{m+n+p}(V\otimes \mathrm{Ind}_{n,p}^{n+p}(U\otimes W))\otimes\boldsymbol{k}[mn]\\ &\xrightarrow{\mathrm{id}_V\otimes\beta_{U,W}} \mathrm{Ind}_{n,m+p}^{m+n+p}(V\otimes \mathrm{Ind}_{p,m}^{m+p}(W\otimes U)\otimes\boldsymbol{k}[mp])\otimes\boldsymbol{k}[mn]\\ &\xrightarrow{\sim} \mathrm{Ind}_{n,m+p}^{m+n+p}(V\otimes \mathrm{Ind}_{p,m}^{m+p}(W\otimes U))\otimes\boldsymbol{k}[m(n+p)].\end{aligned}$$

We have

$$\begin{aligned}&g\otimes(1\otimes u\otimes v)\otimes w\\ &\quad\mapsto (-1)^{mn(|u|+|v|)+|u||v|} g\otimes(\tau^{-1}_{m,n}\otimes v\otimes u\otimes\pi^{mn})\otimes w\\ &\quad\mapsto (-1)^{mn(|u|+|v|)+|u||v|+mn|w|} g\tilde{\tau}^{-1}_{m,n}\otimes(1\otimes v\otimes u)\otimes w\otimes\pi^{mn}\\ &\quad\mapsto (-1)^{mn(|u|+|v|)+|u||v|+mn|w|} g\tilde{\tau}^{-1}_{m,n}\otimes v\otimes(1\otimes u\otimes w)\otimes\pi^{mn}\\ &\quad\mapsto (-1)^{mn(|u|+|v|)+|u||v|+mn|w|+mp(|u|+|w|)+|u||w|} g\tilde{\tau}^{-1}_{m,n}\otimes v\otimes(\tilde{\tau}^{-1}_{m,p}\otimes w\otimes u\otimes\pi^{mp})\otimes\pi^{mn}\\ &\quad\mapsto (-1)^{mn(|u|+|v|)+|u||v|+mn|w|+mp(|u|+|w|)+|u||w|+mp|v|} g\tilde{\tau}^{-1}_{m,n} j(1,\tilde{\tau}^{-1}_{m,p})\otimes v\otimes(1\otimes w\otimes u)\otimes\pi^{m(n+p)},\end{aligned}$$

where here $j=j_{n,m+p}$. In the final step, we used the identity

$$v\otimes(\tilde{\tau}^{-1}_{m,p}\otimes w\otimes u) = (-1)^{mp|v|}(1,\tilde{\tau}^{-1}_{m,p})(v\otimes(1\otimes w\otimes u))$$

in $V\otimes\mathrm{Ind}_{p,m}^{m+p}(W\otimes U)$, and then moved $(1,\tilde{\tau}^{-1}_{m,p})$ into the first tensor factor via j. Since $\tilde{\tau}^{-1}_{m,n} j(1,\tilde{\tau}^{-1}_{m,p}) = \tilde{\tau}^{-1}_{m,n+p}$ (Proposition 5.6) and the signs in the two compositions agree, the result follows. □

6.7. ***Universal property.*** Let $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)^{\mathrm{pf}}$ be the full subcategory of $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ on finite-length projective objects. If $\mathrm{char}(\boldsymbol{k})=0$, then the projective condition is automatic. The category $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)^{\mathrm{pf}}$ is "nice"

in the sense defined in Section 8.1: it is additive, Karoubian and admits a Π-structure. Let $\boldsymbol{V}$ be the "standard" representation of $\widetilde{\mathfrak{S}}_*$: it is $\boldsymbol{k}$ in degree 1 and 0 in other degrees.

Theorem 6.7. *The pair* $(\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)^{\mathrm{pf}}, \boldsymbol{V})$ *is the universal nice type Ia supersymmetric monoidal* $\mathbb{Z}$*-supercategory with an object of degree* 1*, in the following sense. Suppose* $\mathcal{C}$ *is a nice supersymmetric monoidal* $\mathbb{Z}$*-supercategory and* M *is an object of* $\mathcal{C}$ *of degree* 1*. Then there exists a supersymmetric monoidal superfunctor* $\Phi : \mathrm{Rep}(\widetilde{\mathfrak{S}}_*)^{\mathrm{pf}} \to \mathcal{C}$ *and an isomorphism* $\iota : \Phi(\boldsymbol{V}) \to M$*. Furthermore,* (Φ, ι) *is unique up to isomorphism.*

Proof. One can prove this directly; however, we simply note that $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)^{\mathrm{pf}}$ is the "nice envelope" of $\tilde{\mathfrak{S}}$, and so inherits the universal property of $\tilde{\mathfrak{S}}$ (Theorem 5.14). □

6.8. ***Exterior powers of categories.*** One can generalize the above discussion to use coefficients in a supercategory $\mathcal{A}$. This leads to the notion of the exterior power of $\mathcal{A}$, as defined in [Ganter and Kapranov 2014]. The following discussion is intended as an extended comment, and so we omit many details.

Let $\mathcal{A}$ be a nice supercategory (see Section 8.1), and let $\mathcal{A}^{\hat{\boxtimes} n}$ be its n-fold self-product with respect to the tensor product $\hat{\boxtimes}$ defined in Section 8.6 below. The symmetric group $\mathfrak{S}_n$ acts in a natural way on $\mathcal{A}^{\hat{\boxtimes} n}$, and so $\widetilde{\mathfrak{S}}_n$ does too through the quotient map $\widetilde{\mathfrak{S}}_n \to \mathfrak{S}_n$. A Sym^n*-object* of $\mathcal{A}$ is an object M of $\mathcal{A}^{\hat{\boxtimes} n}$ equipped with even isomorphisms $\alpha_\sigma : \sigma^*(M) \to M$ for $\sigma \in \mathfrak{S}_n$ satisfying the usual cocycle conditions. A $\bigwedge^n$*-object* of $\mathcal{A}$ is an object M of $\mathcal{A}^{\hat{\boxtimes} n}$ equipped with isomorphisms $\beta_\sigma : \sigma^*(M) \to M$ for $\sigma \in \widetilde{\mathfrak{S}}_n$ satisfying the cocycle conditions and such that $\beta_c = -1$ and β_σ is homogeneous of the same degree as σ. We let $\mathrm{Sym}^n(\mathcal{A})$ and $\bigwedge^n(\mathcal{A})$ denote the categories of Sym^n- and $\bigwedge^n$-objects of $\mathcal{A}$. We define $\mathrm{Sym}(\mathcal{A})$ and $\bigwedge(\mathcal{A})$ to be the $\mathbb{Z}$-supercategories with these homogeneous pieces.

One can endow $\mathrm{Sym}(\mathcal{A})$ with a symmetric monoidal structure and $\bigwedge(\mathcal{A})$ with a supersymmetric monoidal structure, similar to the definitions used above for $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$. We note that $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ itself can be viewed as $\bigwedge(\mathbf{SVec})$. These categories satisfy natural universal properties; e.g., given a nice supersymmetric monoidal $\mathbb{Z}$-supercategory $\mathcal{B}$, giving a supersymmetric monoidal $\mathbb{Z}$-functor $\bigwedge(\mathcal{A}) \to \mathcal{B}$ is the same as giving a superfunctor $\mathcal{A} \to \mathcal{B}$.

7. Isomeric vector spaces

We assume $\boldsymbol{k}$ *contains* ζ_{16} *throughout this section, and put* $\sqrt{2} = \zeta_8(1 - \zeta_4)$.

7.1. ***Isomeric vector spaces.*** Let V be a super-vector space. An *isomeric structure* on V is an odd isomorphism $\nu : V \to V$ such that $\nu^2 = 1$. An *isomeric vector space* is a vector space equipped with an isomeric structure. Suppose that (U, μ) and (V, ν) are isomeric vector spaces. A homogeneous morphism of isomeric spaces $(U, \mu) \to (V, \nu)$ is a homogeneous linear map $f : U \to V$ such that $f\mu = (-1)^{|f|}\nu f$. In this way, we have a supercategory **QVec** of isomeric vector spaces.

7.2. ***The half tensor product.*** Let (U, μ) and (V, ν) be two isomeric vector spaces. Then $\mu \otimes \nu$ defines an even automorphism of $U \otimes V$ squaring to -1. We define the *half tensor product* of U and V, denoted

by $2^{-1}(U \otimes V)$, to be the ζ_4-eigenspace of $\mu \otimes \nu$. We note that the $(-\zeta_4)$-eigenspace is isomorphic to the ζ_4-eigenspace, via $\mu \otimes 1$ or $1 \otimes \nu$. The half tensor product is an ordinary super-vector space; it does not come with a preferred isomeric structure.

7.3. ***The isomeric monoidal category.*** Motivated by the above discussion, we define the *isomeric category* **Ismr** to be the $\mathbb{Z}/2$-supercategory with $\mathbf{Ismr}_0 = \mathbf{SVec}$ and $\mathbf{Ismr}_1 = \mathbf{QVec}$. We now define a monoidal operation $\odot$ on **Ismr**. We define $\odot$ on objects as follows:

- Given $U, V \in \mathbf{Ismr}_0$, we let $U \odot V = U \otimes V$.
- Let $U \in \mathbf{Ismr}_0$ and $V \in \mathbf{Ismr}_1$, and let ν be the isomeric structure on V. Then $U \otimes V$ carries the isomeric structure $1 \otimes \nu$. We define $U \odot V$ to be the isomeric vector space $(U \otimes V, 1 \otimes \nu)$. We similarly define $V \odot U$.
- Given $U, V \in \mathbf{Ismr}_1$, we define $U \odot V$ to be the half tensor product $2^{-1}(U \otimes V)$.

The definition of $\odot$ on morphisms is straightforward. In this way, we have defined a superfunctor

$$\odot : \mathbf{Ismr} \boxtimes \mathbf{Ismr} \to \mathbf{Ismr}.$$

The unit object for this monoidal structure is the unit object of $\mathbf{Ismr}_0$, i.e., the super-vector space $\boldsymbol{k}$. The isomorphisms λ and ρ are similar to **SVec**, and we omit the details.

We now discuss the associativity constraint α for $\odot$. For this, we will require the following result:

Proposition 7.1. *Let (U, μ), (V, ν) and (W, σ) be isomeric vector spaces. Then $\mu\sigma - 1$ is an automorphism of $U \otimes V \otimes W$ that maps $U \otimes 2^{-1}(V \otimes W)$ isomorphically to $2^{-1}(U \otimes V) \otimes W$.*

Proof. Since $(\mu\sigma)^2 = -1$, the inverse of $\mu\sigma - 1$ is $-\frac{1}{2}(\mu\sigma + 1)$. Now suppose that $x \in U \otimes 2^{-1}(V \otimes W)$. Thus, $\nu\sigma x = \zeta_4 x$; applying σ to each side, we find $\nu x = -\zeta_4 \sigma x$. We thus have

$$\mu\nu(\mu\sigma - 1)x = -\nu\sigma x - \mu\nu x = -\zeta_4 x + \zeta_4 \mu\sigma x = \zeta_4(\mu\sigma - 1)x,$$

which shows that $(\mu\sigma - 1)x$ belongs to $2^{-1}(U \otimes V) \otimes W$. By symmetry, $\sigma\mu - 1$ maps $2^{-1}(U \otimes V) \otimes W$ to $U \otimes 2^{-1}(V \otimes W)$. Since $\mu\sigma - 1$ and $\sigma\mu - 1$ are inverses (up to a factor of 2), these maps are isomorphisms. □

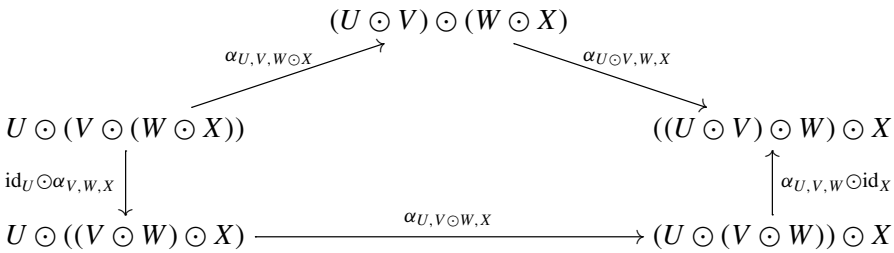
Let $U, V, W \in \mathbf{Ismr}$ be homogeneous objects. We define an isomorphism

$$\alpha_{U,V,W} : U \odot (V \odot W) \to (U \odot V) \odot W$$

as follows. We identify the domain and target above with subspaces of $U \otimes V \otimes W$. If at least one of U, V or W has degree 0, then the domain and target are identified with the same subspace, and we take α to be the identity map. Suppose now that U, V and W are each degree 1, and let μ, ν and σ be the isomeric structures on them. We then take $\alpha_{U,V,W} = \frac{1}{\sqrt{2}}(\mu\sigma - 1)$, which is an appropriate isomorphism by Proposition 7.1.

Proposition 7.2. *The pentagon axiom holds.*

Proof. Let $U, V, W, X \in \mathbf{Ismr}$ be homogeneous objects. We let μ, ν, σ, τ denote isomeric structures on these objects, respectively, whenever they have degree 1. We must show that the diagram



commutes. We identify each space above with a subspace of $U \otimes V \otimes W \otimes X$. If at most two of U, V, W and X are degree 1, then all maps in the diagram are the identity map, and so it commutes. Now suppose exactly three of the objects have degree 1. Then, in each of the two paths in the pentagon, exactly one α is not the identity, and these two α's are equal; for example, if $|U| = 0$, then $\alpha_{U\odot V,W,X} = \mathrm{id}_U \odot \alpha_{V,W,X} = \frac{1}{\sqrt{2}}(\nu\tau - 1)$. Finally, suppose that all four objects have degree 1. Then all edges in the pentagon except the bottom left and bottom right are the identity. We must show $(\mu\sigma - 1)(\nu\tau - 1) = 2$ on $U \odot (V \odot (W \odot X))$. Suppose x belongs to this space. Then $\mu\nu x = \zeta_4 x$, and so $\nu x = \zeta_4 \mu x$; similarly, $\sigma\tau x = \zeta_4 x$, and so $\tau x = \zeta_4 \sigma x$. We thus find

$$\nu\tau x = \zeta_4 \nu\sigma x = -\zeta_4 \sigma\nu x = \sigma\mu x = -\mu\sigma x.$$

Thus,

$$(\mu\sigma - 1)(\nu\tau - 1)x = (\mu\sigma - 1)(-\mu\sigma - 1)x = 2x,$$

since $(\mu\sigma)^2 = -1$. □

7.4. ***Supersymmetry.*** We now aim to define a supersymmetry on $\odot$. We let τ be the standard symmetry on **SVec**. We begin with the following observation:

Proposition 7.3. *Let (U, μ) and (V, ν) be isomeric vector spaces. Then the diagram*

$$\begin{array}{ccc} U \otimes V & \xrightarrow{\mu\otimes\nu} & U \otimes V \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau} \\ V \otimes U & \xrightarrow{\nu\otimes\mu} & V \otimes U \end{array}$$

commutes up to -1. In particular, τ maps the ζ_4-eigenspace of $\mu \otimes \nu$ on $U \otimes V$ to the $(-\zeta_4)$-eigenspace of $\nu \otimes \mu$ on $V \otimes U$.

Proof. The first statement follows from the fact that τ is a natural transformation and μ and ν are odd (see the first paragraph of Section 3.4). The second statement follows from the first. □

Let $U, V \in \mathbf{Ismr}$. We define an isomorphism

$$\beta_{U,V} : U \odot V \to V \odot U$$

as follows. If at least one of U or V has degree 0, then $\beta_{U,V} = \tau_{U,V}$. Now suppose that both U and V have degree 1, and let ν be the isomeric structure on V. Then $\beta_{U,V} = \zeta_{16}^{-1}(\nu \otimes 1)\tau_{U,V}$. This maps $U \odot V$ into $V \odot U$ by Proposition 7.3 and properties of the half tensor product (see Section 7.2).

Proposition 7.4. *β defines a type IIb supersymmetry on $\odot$.*

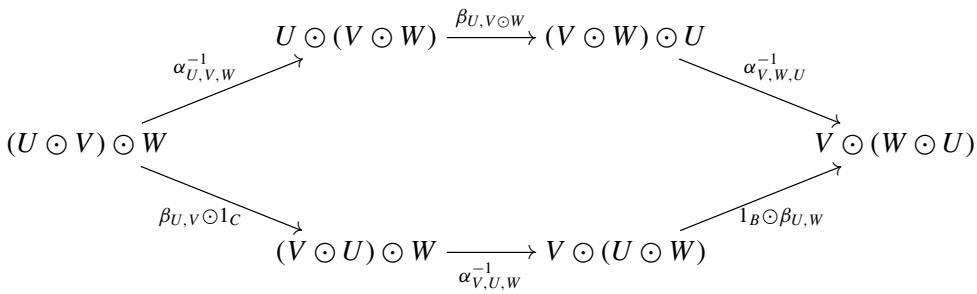
Proof. One easily sees that β is a natural transformation, and it is evident that $\beta_{U,V}$ has degree $|U||V|$ when U and V are homogeneous objects. We will use the following convention in the remainder of the proof: for an endomorphism μ of a super-vector space U, we write μ still for the induced endomorphism of $U \otimes V$ or $V \otimes U$. Note that, under this convention, μ commutes with $\tau_{U,V}$.

Now let $U, V \in \mathbf{Ismr}$ be homogeneous objects. If either U or V has degree 0, then $\beta_{V,U}\beta_{U,V}$ is the identity. Now suppose that both U and V have degree 1, and let μ and ν be the isomeric structures. Then

$$\beta_{V,U}\beta_{U,V} = \zeta_8^{-1}\mu\tau_{V,U}\nu\tau_{U,V} = \zeta_8^{-1}\mu\nu\tau_{V,U}\tau_{U,V} = \zeta_8^{-1}\mu\nu = \zeta_8.$$

In the final step, we used the fact that $\mu\nu = \zeta_4$ on $U \odot V$. We thus see that $\beta_{V,U}\beta_{U,V} = \mathtt{B}_\sharp(|U|, |V|)\,\mathrm{id}_{U\otimes V}$ in all cases, which verifies the requisite symmetry condition for β.

We now verify the hexagon axiom (H1). Let $U, V, W \in \mathbf{Ismr}$ be homogeneous objects, and let μ, ν, σ be the isomeric structures on the objects of degree 1. We must show that the diagram



commutes. (Recall that $\mathtt{B}_1 = 1$ identically, which is why the diagram must commute exactly.) If at most one of the objects has degree 1, then this is the usual hexagon in **SVec**, which commutes. Now suppose exactly two objects have degree 1. Then all the associators are the identity (or, more precisely, induced by the standard associator on the ordinary tensor products). If $|U| = 0$, then each of the β's is just a τ, and the diagram commutes. Otherwise, one of $\beta_{U,V}$ or $\beta_{U,W}$ is the usual τ, and the other involves the same isomeric structure as $\beta_{U,V\odot W}$; for instance, if $|V| = 0$, then

$$\beta_{U,V\odot W} = \zeta_{16}^{-1}\sigma\,\tau_{U,V\odot W}, \quad \beta_{U,V} = \tau_{U,V}, \quad \beta_{U,W} = \zeta_{16}^{-1}\sigma\,\tau_{U,W}.$$

Thus, the diagram commutes (the isomeric structure can be moved past all the τ's and then one has the ordinary hexagon axiom in **SVec**).

Finally, suppose that all objects have degree 1. The top path is (induced by) the map

$$\tfrac{1}{2}(\mu\nu - 1)\tau_{U,V\otimes W}(\sigma\mu - 1) = \tfrac{1}{2}\tau_{U,V\otimes W}(\mu\nu - 1)(\sigma\mu - 1).$$

The bottom path is (induced by) the map

$$\tfrac{1}{\sqrt{2}}\zeta_8^{-1}\sigma\tau_{U,W}(\sigma\nu-1)\nu\tau_{U,V} = \tfrac{1}{\sqrt{2}}\zeta_8^{-1}\tau_{U,W}\tau_{U,V}\sigma(\sigma\nu-1)\nu.$$

Since $\tau_{U,V\otimes W} = \tau_{U,W}\tau_{U,V}$, it suffices to show that

$$\tfrac{1}{2}(\mu\nu-1)(\sigma\mu-1) = \tfrac{1}{\sqrt{2}}\zeta_8^{-1}\sigma(\sigma\nu-1)\nu$$

as functions on $(U\odot V)\odot W$. This equation is equivalent to

$$\nu\sigma-\mu\nu-\sigma\mu+1 = \sqrt{2}\zeta_8^{-1}(1-\sigma\nu).$$

Since $\mu\nu=\zeta_4$ on the source, we have $\mu=-\zeta_4\nu$, and so the left side of the above equation is

$$-\sigma\nu-\zeta_4+\zeta_4\sigma\nu+1 = (1-\zeta_4)(1-\sigma\nu),$$

and so the result follows, as $1-\zeta_4=\sqrt{2}\zeta_8^{-1}$.

Since β satisfies the symmetry condition, the second hexagon axiom is a consequence of the first one. □

7.5. ***Further comments.*** The pentagon and hexagon axioms in **Ismr** are rather complicated. In fact, there is an alternative way of defining the category **Ismr** that makes these axioms more transparent. Recall that the Clifford algebra Cl_n is generated by n supercommuting odd elements that each square to 1. One formulation of Bott periodicity (discussed in the Appendix) states that Cl_n and Cl_{n+2} are Morita-equivalent. We can identify **QVec** with the category of Cl_1-modules. The half tensor product is then the composition

$$\mathrm{Mod}_{\mathrm{Cl}_1}\times\mathrm{Mod}_{\mathrm{Cl}_1}\to\mathrm{Mod}_{\mathrm{Cl}_2}\to\mathbf{SVec},$$

where the first functor takes (V, W) to $V\otimes W$, with obvious Cl_2-structure, and the second functor is the Morita equivalence provided by Bott periodicity. It is the second step that introduces the complications.

To obtain a "cleaner" version of the isomeric category, we can simply omit the Morita equivalence above. More precisely, we could redefine $\mathbf{Ismr}_i$ to be the 2-colimit of the categories $\mathrm{Mod}_{\mathrm{Cl}_n}$ taken over n of the same parity as i. The tensor product would then be more natural and the axioms would be more transparent.

This new approach also points the way to a vast generalization: one can start with an arbitrary $\mathbb{Z}/2$-supercategory $\mathcal{A}$ and construct an "isomeric version" by considering $\mathrm{Cl}_{\mathrm{even}}$-modules in $\mathcal{A}_0$ and $\mathrm{Cl}_{\mathrm{odd}}$-modules in $\mathcal{A}_1$. In fact, one can also work over more general base fields where the periodicity in Clifford algebras may be 4 or 8 instead of 2. We call this generalization "Clifford eversion", and carry out the details in the next section.

8. Clifford eversion

8.1. ***Nice supercategories.*** Let $\mathcal{C}$ be a supercategory. We say that $\mathcal{C}$ is *additive* if all finite direct sums exist. (The direct sum of two objects is an object equipped with *even* inclusion and projection maps satisfying the usual conditions.) We say that $\mathcal{C}$ is *Karoubian* if every *even* idempotent $e: A\to A$ in $\mathcal{C}$

has an image. We say that $\mathcal{C}$ is *nice* if it is additive, Karoubian and admits a Π-structure. We say that a Λ-supercategory is nice if each graded piece is. We will mostly be concerned with nice supercategories in this section. We note that, if M is an object in a nice supercategory and V is a finite-dimensional super-vector space, then we can make sense of $V \otimes M$, e.g., $\boldsymbol{k}^{r|s} \otimes M = M^{\oplus r} \oplus \Pi(M)^{\oplus s}$. (In fact, this construction does not require the Karoubian condition.)

8.2. ***The Clifford algebra.*** The *Clifford algebra* Cl_n is the $\boldsymbol{k}$-superalgebra generated by odd-degree variables $\alpha_1, \dots, \alpha_n$ subject to the relations $\alpha_i^2 = 2$ and $\alpha_i \alpha_j = -\alpha_j \alpha_i$ for $i \neq j$. If 2 is a square in $\boldsymbol{k}$, then one can renormalize α_i so that it squares to 1 (as we did in Section 7.5), which is a more common convention. We use our convention so that the following proposition works without need of $\sqrt{2}$:

Proposition 8.1. *There is a group homomorphism* $\widetilde{\mathfrak{S}}_n \to \mathrm{Cl}_n^\times$ *given by*

$$\tilde{s}_i \mapsto \tfrac{1}{2}(\alpha_{i+1} - \alpha_i), \quad c \mapsto -1.$$

Proof. Let $\sigma_i = \frac{1}{2}(\alpha_{i+1} - \alpha_i)$. We have

$$\sigma_i^2 = \tfrac{1}{4}(\alpha_{i+1}^2 - \alpha_{i+1}\alpha_i - \alpha_i\alpha_{i+1} + \alpha_i^2) = 1,$$

as the middle terms cancel and the outer terms are each equal to 2. It is clear that σ_i and σ_j anticommute if $|i - j| > 1$. Finally,

$$\sigma_i\sigma_{i+1}\sigma_i = \tfrac{1}{2}(\alpha_i - \alpha_{i+2}) = \sigma_{i+1}\sigma_i\sigma_{i+1},$$

so the braid relation holds. We thus see that the σ's satisfy the relations defining the spin-symmetric group, and the claim follows. □

8.3. ***Clifford modules.*** Let $\mathcal{A}$ be a supercategory. A Cl_n*-module* in $\mathcal{A}$ is a pair (M, ρ) consisting of an object M of $\mathcal{A}$ and a morphism of $\boldsymbol{k}$-superalgebras $\rho : \mathrm{Cl}_n \to \mathrm{End}(M)$. Explicitly, a Cl_n-module consists of an object M together with odd-degree supermorphisms $\alpha_1, \dots, \alpha_n : M \to M$ such that $\alpha_i^2 = 2$ and $\alpha_i\alpha_j = -\alpha_j\alpha_i$ for $i \neq j$. Suppose that M and N are two Cl_n-modules in $\mathcal{A}$. A *morphism* $f : M \to N$ of parity p is a morphism in $\mathcal{A}$ of parity p compatible with the Cl_n-structures, i.e., such that $f\alpha_i = (-1)^p \alpha_i f$. We let $\mathrm{Cl}_n(\mathcal{A})$ be the supercategory of Cl_n-modules in $\mathcal{A}$.

One easily sees that, if $\mathcal{A}$ is additive or Karoubian, then so is $\mathrm{Cl}_n(\mathcal{A})$. Moreover, if $\mathcal{A}$ admits a Π-structure, then so does $\mathrm{Cl}_n(\mathcal{A})$, via

$$\Pi(M; \alpha_1, \dots, \alpha_n) = (\Pi(M); \Pi(\alpha_1), \dots, \Pi(\alpha_n)).$$

We thus see that, if $\mathcal{A}$ is nice, then so is $\mathrm{Cl}_n(\mathcal{A})$. The following statement is clear, and we omit its proof:

Proposition 8.2. *We have an equivalence* $\mathrm{Cl}_m(\mathrm{Cl}_n(\mathcal{A})) \to \mathrm{Cl}_{n+m}(\mathcal{A})$ *defined by*

$$((M; \alpha_1, \dots, \alpha_n); \alpha_1', \dots, \alpha_m') \mapsto (M; \alpha_1, \dots, \alpha_n, \alpha_1', \dots, \alpha_m').$$

Given a Cl_n-module $(M; \alpha_1, \dots, \alpha_n)$ in $\mathcal{A}$ and an element $\sigma \in \mathfrak{S}_n$, put

$$\sigma_*(M; \alpha_1, \dots, \alpha_n) = (M; \alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)}).$$

This construction defines a left action of $\mathfrak{S}_n$ on the category $\mathrm{Cl}_n(\mathcal{A})$. We let $\widetilde{\mathfrak{S}}_n$ act via the homomorphism $\widetilde{\mathfrak{S}}_n \to \mathfrak{S}_n$. The following proposition shows that every object of $\mathrm{Cl}_n(\mathcal{A})$ is canonically equivariant for the $\widetilde{\mathfrak{S}}_n$-action:

Proposition 8.3. *Let $(M; \alpha_1, \dots, \alpha_n)$ be a Cl_n-module. Then $\widetilde{\mathfrak{S}}_n$ naturally acts on M, and permutes the α's via the surjection $\widetilde{\mathfrak{S}}_n \to \mathfrak{S}_n$. In particular, for $\sigma \in \widetilde{\mathfrak{S}}_n$, we have a natural isomorphism of Cl_n-modules*

$$\sigma : (M; \alpha_1, \dots, \alpha_n) \to (M; \alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)})$$

that has the same parity as σ.

Proof. The action comes from Proposition 8.1 using the homomorphism $\varphi : \widetilde{\mathfrak{S}}_n \to \mathrm{Cl}_n^\times$. We claim that $\varphi(\sigma)\alpha_i = (-1)^{|\sigma|}\alpha_{\sigma^{-1}(i)}\varphi(\sigma)$ for all $\sigma \in \widetilde{\mathfrak{S}}_n$. It is enough to check this when $\sigma = \tilde{s}_j$. Then

$$\begin{aligned}
\varphi(\tilde{s}_j)\alpha_j &= \tfrac{1}{2}(\alpha_{j+1}-\alpha_j)\alpha_j = \tfrac{1}{2}\alpha_{j+1}\alpha_j - 1 = -\tfrac{1}{2}\alpha_{j+1}(\alpha_{j+1}-\alpha_j) = -\alpha_{j+1}\varphi(\tilde{s}_j),\\
\varphi(\tilde{s}_j)\alpha_{j+1} &= \tfrac{1}{2}(\alpha_{j+1}-\alpha_j)\alpha_{j+1} = 1 - \tfrac{1}{2}\alpha_j\alpha_{j+1} = -\tfrac{1}{2}\alpha_j(\alpha_{j+1}-\alpha_j) = -\alpha_j\varphi(\tilde{s}_j),\\
\varphi(\tilde{s}_j)\alpha_k &= \tfrac{1}{2}(\alpha_{j+1}-\alpha_j)\alpha_k = -\tfrac{1}{2}\alpha_k(\alpha_{j+1}-\alpha_j) = -\alpha_k\varphi(\tilde{s}_j) \qquad (k \notin \{j, j+1\}).
\end{aligned}$$

□

8.4. ***Periodicity.*** Throughout this section, we fix a value p such that Cl_p is isomorphic as a superalgebra to a matrix superalgebra. (By Theorem A.4, we know that $p = 8$ is possible for any $\boldsymbol{k}$ and that we can take $p = 2$ if $\zeta_4 \in \boldsymbol{k}$.) Let U_p be a Cl_p-module such that the functor $\mathbf{SVec} \to \mathrm{Mod}_{\mathrm{Cl}_p}$ given by $V \mapsto V \otimes U_p$ is an equivalence, and let $\epsilon_p \in \mathrm{Cl}_p$ be an even idempotent such that the functor $\mathrm{Mod}_{\mathrm{Cl}_p} \to \mathbf{SVec}$ given by $M \mapsto \epsilon_p M$ is a quasi-inverse.

Let $\mathcal{A}$ be a nice supercategory. The following is clear:

Proposition 8.4. *The functor $\mathcal{A} \to \mathrm{Cl}_p(\mathcal{A})$ given by $M \mapsto M \otimes U_p$ is an equivalence.*

We let $\Phi_n : \mathrm{Cl}_n(\mathcal{A}) \to \mathrm{Cl}_{n+p}(\mathcal{A})$ be the equivalence induced by the identification $\mathrm{Cl}_{n+p}(\mathcal{A}) = \mathrm{Cl}_p(\mathrm{Cl}_n(\mathcal{A}))$. For a nonnegative integer i, define $\mathrm{Cl}_{(i;p)}(\mathcal{A})$ to be the 2-colimit of the system

$$\mathrm{Cl}_i(\mathcal{A}) \xrightarrow{\Phi_i} \mathrm{Cl}_{i+p}(\mathcal{A}) \xrightarrow{\Phi_{i+p}} \mathrm{Cl}_{i+2p}(\mathcal{A}) \to \cdots.$$

Explicitly, an object of $\mathrm{Cl}_{(i;p)}(\mathcal{A})$ is a Cl_n-module in $\mathcal{A}$ for some integer $n \geq 0$ such that $n \equiv i \pmod p$. If M is a Cl_m-module and N is a Cl_n-module, with $m \equiv n \equiv i \pmod p$, then

$$\mathrm{Hom}_{\mathrm{Cl}_{(i;p)}(\mathcal{A})}(M, N) = \varinjlim \mathrm{Hom}_{\mathrm{Cl}_r(\mathcal{A})}(\Phi^{(r-m)/p}(M), \Phi^{(r-n)/p}(N)),$$

where the colimit is taken as $r \to \infty$ over integers r such that $r \equiv i \pmod p$. It is clear that the natural functor $\mathrm{Cl}_i(\mathcal{A}) \to \mathrm{Cl}_{(i;p)}(\mathcal{A})$ is an equivalence, and so, in a sense, these limit categories are nothing new. However, the categories $\mathrm{Cl}_{(i;p)}(\mathcal{A})$ will sometimes be easier to work with, especially when dealing with products.

8.5. ***Clifford eversion.*** As before, let p be such that Cl_p is a matrix superalgebra, and fix an integer q divisible by p (we allow $q = 0$). Note that q is necessarily even.

Definition 8.5. The *Clifford eversion* of a nice $\mathbb{Z}/q$-supercategory $\mathcal{A}$, denoted by $\mathrm{Cl}(\mathcal{A})$, is the nice $\mathbb{Z}/q$-supercategory given by $\mathrm{Cl}(\mathcal{A})_i = \mathrm{Cl}_{(i;p)}(\mathcal{A}_i)$. □

Remark 8.6. We defined $\mathrm{Cl}(\mathcal{A})_i$ to be $\mathrm{Cl}_{(i;p)}(\mathcal{A}_i)$. In some ways, it is more natural to use $\mathrm{Cl}_{(i;q)}(\mathcal{A}_i)$, but there is one problem with this: it is not defined if $q = 0$ and $i < 0$. □

Remark 8.7. The motivation for the definition of Clifford eversion comes from the isomeric category; see Section 7.5. □

We let Cl^k be the k-fold iterate of Cl. To study this, it is convenient to introduce a slight variant. Define $\mathrm{Cl}^{(k)}(\mathcal{A})$ to be the $\mathbb{Z}/q$-supercategory given by $\mathrm{Cl}^{(k)}(\mathcal{A})_i = \mathrm{Cl}_{(ki;p)}(\mathcal{A})$. Thus, $\mathrm{Cl}^{(1)} = \mathrm{Cl}$. The equivalence in Proposition 8.2 induces an equivalence $\mathrm{Cl}_{(i;p)}(\mathrm{Cl}_{(j;p)}(\mathcal{B})) \to \mathrm{Cl}_{(i+j;p)}(\mathcal{B})$ for any nice supercategory $\mathcal{B}$. It follows that there is a natural equivalence $\mathrm{Cl}^{(k)}(\mathrm{Cl}^{(\ell)}(\mathcal{A})) \to \mathrm{Cl}^{(k+\ell)}(\mathcal{A})$. In particular, there is a natural equivalence $\mathrm{Cl}^k(\mathcal{A}) \to \mathrm{Cl}^{(k)}(\mathcal{A})$ for any $k \geq 0$.

Proposition 8.8. *We have a canonical equivalence* $\mathcal{A} \cong \mathrm{Cl}^p(\mathcal{A})$.

Proof. Combine the natural equivalences $\mathcal{A} \to \mathrm{Cl}^{(p)}(\mathcal{A})$ and $\mathrm{Cl}^p(\mathcal{A}) \to \mathrm{Cl}^{(p)}(\mathcal{A})$. □

Suppose $\mathcal{A}$ is a nice monoidal $\mathbb{Z}/q$-supercategory. We define a natural monoidal structure on $\mathrm{Cl}(\mathcal{A})$ as follows. Let $r, s \in \mathbb{Z}/q$ and let r' and s' be nonnegative integers congruent to r and s modulo p. We have a functor $\mathrm{Cl}_{r'}(\mathcal{A}_r) \boxtimes \mathrm{Cl}_{s'}(\mathcal{A}_s) \to \mathrm{Cl}_{r'+s'}(\mathcal{A}_{r+s})$ given by

$$(M; \mu_1, \dots, \mu_{r'}) \otimes (N; \nu_1, \dots, \nu_{s'}) \mapsto (M \otimes N; \mu_1, \dots, \mu_{r'}, \nu_1, \dots, \nu_{s'}).$$

Here μ_i is shorthand for $\mu_i \otimes 1$, and similarly for ν_j. Note that M is an object of $\mathcal{A}_r$ and N is an object of $\mathcal{A}_s$, so $M \otimes N$ is an object of $\mathcal{A}_{r+s}$. Thus, the above object does indeed belong to $\mathrm{Cl}(\mathcal{A})_{r+s}$. The diagram

$$\begin{array}{ccc} \mathrm{Cl}_{r'}(\mathcal{A}_r) \boxtimes \mathrm{Cl}_{s'}(\mathcal{A}_s) & \xrightarrow{1 \otimes \Phi_{s'}} & \mathrm{Cl}_{r'}(\mathcal{A}_r) \boxtimes \mathrm{Cl}_{s'+p}(\mathcal{A}_s) \\ \downarrow & & \downarrow \\ \mathrm{Cl}_{r'+s'}(\mathcal{A}_{r+s}) & \xrightarrow{\Phi_{r'+s'}} & \mathrm{Cl}_{r'+s'+p}(\mathcal{A}_{r+s}) \end{array}$$

commutes, since in both cases the extra endomorphisms for Cl_p are appended to the end of the list. The diagram

$$\begin{array}{ccc} \mathrm{Cl}_{r'}(\mathcal{A}_r) \boxtimes \mathrm{Cl}_{s'}(\mathcal{A}_s) & \xrightarrow{\Phi_{s'} \otimes 1} & \mathrm{Cl}_{r'+p}(\mathcal{A}_r) \boxtimes \mathrm{Cl}_{s'}(\mathcal{A}_s) \\ \downarrow & & \downarrow \\ \mathrm{Cl}_{r'+s'}(\mathcal{A}_{r+s}) & \xrightarrow{\Phi_{r'+s'}} & \mathrm{Cl}_{r'+s'+p}(\mathcal{A}_{r+s}) \end{array}$$

does not commute, but does commute up to canonical even isomorphism. Precisely, let $(M; \mu_\bullet)$ and $(N; \nu_\bullet)$ be objects of $\mathrm{Cl}_{r'}(\mathcal{A}_r)$ and $\mathrm{Cl}_{s'}(\mathcal{A}_s)$, and let $(U_q; \lambda_1, \dots, \lambda_p)$ be the Cl_p-module inducing periodicity. Then the top path in the above diagram yields $(M \otimes N; \mu_\bullet, \lambda_\bullet, \nu_\bullet)$, while the bottom path yields $(M \otimes N; \mu_\bullet, \nu_\bullet, \lambda_\bullet)$. These two objects are isomorphic via Proposition 8.3, using that $j_{r',p+s'}(1, \tilde{\tau}_{p,s'}) \in \widetilde{\mathfrak{S}}_{r'+s'+p}$. Note that this isomorphism is even since p is even.

It follows from the above discussion that we have a well-defined functor $\mathrm{Cl}(\mathcal{A}) \boxtimes \mathrm{Cl}(\mathcal{A}) \to \mathrm{Cl}(\mathcal{A})$. The associativity and unit constraints on $\mathrm{Cl}(\mathcal{A})$ are defined in the obvious manner, and one readily verifies that the axioms of a monoidal $\mathbb{Z}/q$-supercategory are satisfied.

We now investigate how the above construction interacts with braidings. In what follows, we assume that $4 \mid q$. Suppose we have an isomorphism $\beta : \otimes \to \otimes \circ \Sigma$. Given objects $(M; \alpha_\bullet)$ and $(N; \alpha'_\bullet)$ as above, we define

$$\beta'_{(M;\alpha_\bullet),(N;\alpha'_\bullet)} : (M; \alpha_\bullet) \otimes (N; \alpha'_\bullet) \to (N; \alpha'_\bullet) \otimes (M; \alpha_\bullet)$$

to be the composition

$$(M \otimes N; \alpha_\bullet, \alpha'_\bullet) \xrightarrow{\beta_{M,N}} (N \otimes M; \alpha_\bullet, \alpha'_\bullet) \xrightarrow{\tilde{\tau}_{r,s}^{-1}} (N \otimes M; \alpha'_\bullet, \alpha_\bullet).$$

The second map uses the action of $\widetilde{\mathfrak{S}}_{r+s}$ on $N \otimes M$, as in Proposition 8.3. We note that $|\beta'| = |\beta| + rs$. For the following, we use the notation from Section 3.3.

Proposition 8.9. *Maintain the above notation.*

(a) *If β is a symmetry (or braiding) with respect to $\omega \in \Gamma_0$, then β' is a type II supersymmetry (or superbraiding) with respect to* $\mathtt{DA}\omega$.

(b) *If β is a type II supersymmetry (or superbraiding) with respect to $\omega \in \Gamma_1$, then β' is a symmetry (or braiding) with respect to* $\mathtt{A}\omega$.

Proof. Suppose that β is a braiding with respect to $\omega = (\omega_1, \omega_2)$. We show that β' is a type II superbraiding with respect to $\mathtt{A}\omega$. We have already observed that $\beta'_{X,Y}$ is homogeneous of degree $|X||Y|$ if X and Y are homogeneous objects. We now verify that β' is a natural transformation. Suppose that $f : (M; \mu_1, \dots, \mu_r) \to (M'; \mu'_1, \dots, \mu'_r)$ and $g : (N; \nu_1, \dots, \nu_s) \to (N'; \nu'_1, \dots, \nu'_s)$ are homogeneous morphisms of homogeneous objects of $\mathrm{Cl}(\mathcal{A})$. Consider the diagram

$$\begin{array}{ccccc}
(M \otimes N; \mu_\bullet, \nu_\bullet) & \xrightarrow{\beta_{M,N}} & (N \otimes M; \mu_\bullet, \nu_\bullet) & \xrightarrow{\tilde{\tau}_{r,s}^{-1}} & (N \otimes M; \nu_\bullet, \mu_\bullet) \\
\downarrow{\scriptstyle f \otimes g} & & \downarrow{\scriptstyle g \otimes f} & & \downarrow{\scriptstyle g \otimes f} \\
(M' \otimes N'; \mu'_\bullet, \nu'_\bullet) & \xrightarrow{\beta_{M',N'}} & (N' \otimes M'; \mu'_\bullet, \nu'_\bullet) & \xrightarrow{\tilde{\tau}_{r,s}^{-1}} & (N' \otimes M'; \nu'_\bullet, \mu'_\bullet)
\end{array}$$

The first square commutes up to $(-1)^{|f||g|}$, as this is simply what it means that β is a natural transformation, while the second square commutes up to $(-1)^{(|f|+|g|)rs}$. This shows that β' is natural.

We now look at the hexagon axioms. Let $(M; \mu_\bullet)$, $(N; \nu_\bullet)$ and $(P, \pi_\bullet)$ be homogeneous objects of $\mathrm{Cl}(\mathcal{A})$ of degrees r, s and t. Consider the diagram

$$\begin{array}{ccccc}
(MNP; \mu, \nu, \pi) & \xrightarrow{\beta_{M,NP}} & (NPM; \mu, \nu, \pi) & \xrightarrow{x} & (NPM; \nu, \pi, \mu) \\
\downarrow{\scriptstyle \beta_{N,M}} & & & & \uparrow{\scriptstyle z} \\
(NMP; \mu, \nu, \pi) & & & & (NPM; \nu, \mu, \pi) \\
& \searrow^{y} & & \nearrow_{\beta_{M,P}} & \\
& & (NMP; \nu, \mu, \pi) & &
\end{array}$$

Here we have omitted $\otimes$ and $\bullet$ symbols, and x, y and z are the appropriate elements of $\widetilde{\mathfrak{S}}_*$ appearing in the definition of β'. This is the hexagon (H1) for β', where we have omitted the associators. Put

$$(\alpha_1, \dots, \alpha_{r+s+t}) = (\mu_1, \dots, \mu_r, \nu_1, \dots, \nu_s, \pi_1, \dots, \pi_t).$$

We regard each α_i as acting on MNP, or any of the other tensor products in the above diagram, in the natural manner. We let $\widetilde{\mathfrak{S}}_{r+s+t}$ act on MNP via the α's (as in Proposition 8.1). Then

$$x = \tilde{\tau}_{r,s+t}^{-1}, \quad y = j_{r+s,t}(\tilde{\tau}_{r,s}^{-1}, 1), \quad z = (-1)^{rst} \cdot y \cdot j_{s,r+t}(1, \tilde{\tau}_{r,t}^{-1}) \cdot y^{-1}.$$

The expressions for x and y are clear. By definition, z is $\mathrm{id}_N \otimes \tilde{\tau}_{r,t}^{-1}$, but where this $\tilde{\tau}_{r,t}$ element is built from the μ's and π's, i.e., the elements $\alpha_1, \dots, \alpha_r$ and $\alpha_{r+s+1}, \dots, \alpha_{r+s+t}$. This agrees with the above formula; note that

$$y\alpha_i y^{-1} = (-1)^{rs} \alpha_{\tau_{r,s}^{-1}(i)}$$

by Proposition 8.3. One readily verifies that $\beta_{M,P}$ commutes with the α_i's (in the evident sense), and thus with y as well. Since $\beta_{M,P}\beta_{N,M} = \omega_1(r; s, t)\beta_{M,NP}$ (this is the first hexagon axiom for β) and $j_{s,r+t}(1, \tilde{\tau}_{r,t}) \cdot j_{r+s,t}(\tilde{\tau}_{r,s}, 1) = \tilde{\tau}_{r,s+t}$ (Proposition 5.6), the above diagram commutes up to $(-1)^{rst}\omega_1(r; s, t)$, which verifies the first hexagon axiom for β' with respect to $\mathtt{DA}\omega$. The second hexagon axiom is similar. We thus see that β' is a type II superbraiding.

Now suppose that β is symmetric, and let us show that β' is supersymmetric. Maintaining the above notation, we have

$$\beta'_{(N;\nu),(M;\mu)}\beta'_{(M;\mu),(N;\nu)} = y\beta_{N,M}x\beta_{M,N},$$

where $x = \tilde{\tau}_{r,s}^{-1}$, and y is defined like $\tilde{\tau}_{s,r}^{-1}$ but with respect to $\alpha_{r+1}, \dots, \alpha_{r+s}, \alpha_1, \dots, \alpha_r$; in other words, $y = (-1)^{rs} x \tilde{\tau}_{s,r}^{-1} x^{-1} = (-1)^{rs} \tilde{\tau}_{s,r}^{-1}$. Since $\beta_{N,M}$ and x commute, we have

$$\beta'_{(N;\nu),(M;\mu)}\beta'_{(M;\mu),(N;\nu)} = (-1)^{rs}\tilde{\tau}_{s,r}^{-1}\tilde{\tau}_{r,s}^{-1}\beta_{N,M}\beta_{M,N} = (-1)^{\binom{r}{2}\binom{s}{2}+rs}\omega_\sharp(|M|, |N|)\,\mathrm{id}_{M\otimes N},$$

where we have used Proposition 5.3. We thus see that β' is a supersymmetry with respect to $\mathtt{DA}\omega$.

We have thus proved statement (a). Statement (b) is similar, though some additional signs come in: in the hexagon axiom we pick up an additional $(-1)^{rst}$ when commuting $\beta_{M,P}$ and y, and in the symmetry axiom we pick up an additional $(-1)^{rs}$ when commuting $\beta_{N,M}$ and $\tilde{\tau}_{r,s}$. □

Remark 8.10. If one only cares about (super)braidings, one can phrase the above proposition using B-factor systems instead of S-factor systems. □

The above proposition can be summarized pictorially as follows: for $\omega \in \Gamma_0$, Clifford eversion acts by

$$\omega \to \mathtt{DA}\omega \to \mathtt{D}\omega \to \mathtt{A}\omega \to \omega \to \cdots.$$

That is, given a (super)symmetric monoidal category for one factor system, the Clifford eversion will be a (super)symmetric monoidal category for the next factor system. The list of factor systems repeats with periodicity 4, while Clifford eversion itself repeats with period p (which is a divisor of 8).

We have seen that there is a canonical equivalence $\mathcal{A} \to \mathrm{Cl}^p(\mathcal{A})$. If $\mathcal{A}$ is (super)braided or (super-) symmetric, then so is $\mathrm{Cl}^p(\mathcal{A})$, and with respect to the same factor system (recall that q is divisible by 4). We now show that this equivalence is appropriately monoidal.

Proposition 8.11. *The equivalence $\mathcal{A} \to \mathrm{Cl}^p(\mathcal{A})$ is monoidal, (super)braided monoidal or (super-) symmetric monoidal, depending on what structure $\mathcal{A}$ has.*

Proof. The definition of the monoidal structure and (super)braiding on $\mathrm{Cl}(\mathcal{A})$ generalizes in the obvious manner to $\mathrm{Cl}^{(k)}(\mathcal{A})$ for any k. The equivalences $\mathrm{Cl}^{(k)}(\mathrm{Cl}^{(\ell)}(\mathcal{A})) \to \mathrm{Cl}^{(k+\ell)}(\mathcal{A})$ are easily seen to be compatible with these structures; it follows that the equivalence $\mathrm{Cl}^k(\mathcal{A}) \to \mathrm{Cl}^{(k)}(\mathcal{A})$ is as well. The equivalence $\mathcal{A} \to \mathrm{Cl}^{(p)}(\mathcal{A})$ induced by Clifford periodicity is also compatible with these structures. □

Combining the above results, we find:

Proposition 8.12. *Suppose q is divisible by 4 and let $\omega \in \Gamma_0$. Then the 2-category of symmetric monoidal $\mathbb{Z}/q$-supercategories with respect to ω is equivalent to the 2-category of type II supersymmetric monoidal $\mathbb{Z}/q$-supercategories with respect to $\mathtt{A}\omega$.*

Specializing to the canonical choice $\omega = 1$, we find:

Corollary 8.13. *The 2-category of symmetric monoidal $\mathbb{Z}/q$-supercategories is equivalent to the 2-category of type IIa supersymmetric monoidal $\mathbb{Z}/q$-supercategories, via Clifford eversion.*

Remark 8.14. We have assumed $4 \mid q$ throughout the above discussion. There are analogous results if $p = 2$ and $2 \mid q$, assuming one has sufficiently many roots of unity in $\boldsymbol{k}$. □

8.6. ***Envelopes.*** Let $\mathcal{C}$ be a supercategory. We define its *additive envelope* $\mathcal{C}^{\mathrm{add}}$ as follows. The objects are symbols (formal direct sums) $A_1 \oplus \cdots \oplus A_r$ where the A_i are objects of $\mathcal{C}$. A morphism $f : A_1 \oplus \cdots \oplus A_r \to B_1 \oplus \cdots \oplus B_s$ is an element $(f_{i,j}) \in \prod_{i,j} \mathrm{Hom}_{\mathcal{C}}(A_i, B_j)$, thought of as a matrix. We say that f is homogeneous of parity p if each $f_{i,j}$ is homogeneous of parity p. Given another morphism $g : B_1 \oplus \cdots \oplus B_s \to C_1 \oplus \cdots \oplus C_t$, the composition gf is defined by $(gf)_{i,j} = \sum_k g_{k,j} f_{i,k}$. One readily verifies that $\mathcal{C}^{\mathrm{add}}$ is an additive supercategory. We have a fully faithful functor $i : \mathcal{C} \to \mathcal{C}^{\mathrm{add}}$ that satisfies a universal property: given a superfunctor $j : \mathcal{C} \to \mathcal{A}$ to an additive supercategory $\mathcal{A}$, there is a unique (up to isomorphism) superfunctor $j' : \mathcal{C}^{\mathrm{add}} \to \mathcal{A}$ such that $j = j'i$. In particular, i is an equivalence if $\mathcal{C}$ is additive.

The *Karoubian envelope* of a supercategory $\mathcal{A}$, denoted by $\mathcal{A}^{\mathrm{kar}}$, consists of objects (A, e) where A is an object of $\mathcal{C}$ and e is an even idempotent of A. A morphism $(A, e) \to (B, f)$ in $\mathcal{A}^{\mathrm{kar}}$ is a morphism $\varphi : A \to B$ such that $f\varphi = \varphi = \varphi e$. (The identity of (A, e) is given by $e : A \to A$.) One easily sees that $\mathcal{A}^{\mathrm{kar}}$ is a Karoubian supercategory. We have a fully faithful superfunctor $i : \mathcal{A} \to \mathcal{A}^{\mathrm{kar}}$ defined by $i(A) = (A, 1)$ that satisfies a universal property: given an superfunctor $j : \mathcal{A} \to \mathcal{B}$ to a Karoubian supercategory $\mathcal{B}$, there is a unique (up to isomorphism) superfunctor $j' : \mathcal{A}^{\mathrm{kar}} \to \mathcal{B}$ such that $j = j'i$. In particular, i is an equivalence if $\mathcal{A}$ is Karoubian.

Lemma 8.15. *Let $F : \mathcal{A} \to \mathcal{B}$ be a fully faithful superfunctor of supercategories. Then $F^{\mathrm{add}} : \mathcal{A}^{\mathrm{add}} \to \mathcal{B}^{\mathrm{add}}$ and $F^{\mathrm{kar}} : \mathcal{A}^{\mathrm{kar}} \to \mathcal{B}^{\mathrm{kar}}$ are also fully faithful.*

Proof. This is clear for F^{add} by the explicit description for morphisms given above. Similarly, a morphism $(A, e) \to (A', e')$ in the Karoubian envelope is just a morphism $A \to A'$ satisfying some conditions which are preserved by F, so F^{kar} is also fully faithful. □

Let $\mathcal{A}$ and $\mathcal{B}$ be nice $\mathbb{Z}/q$-supercategories. The product $\mathcal{A} \boxtimes \mathcal{B}$ will typically not be additive or Karoubian, and thus not nice. We define $\mathcal{A} \hat{\boxtimes} \mathcal{B}$ to be the Karoubian envelope of the additive envelope of $\mathcal{A} \boxtimes \mathcal{B}$. This is again a nice $\mathbb{Z}/q$-supercategory, and $\hat{\boxtimes}$ endows the 2-category of nice $\mathbb{Z}/q$-supercategories with a monoidal structure. We note that, if $\mathcal{A}$ is a nice monoidal $\mathbb{Z}/q$-supercategory, then the monoidal product $\mathcal{A} \boxtimes \mathcal{A} \to \mathcal{A}$ canonically extends to a superfunctor $\mathcal{A} \hat{\boxtimes} \mathcal{A} \to \mathcal{A}$, by formal properties of envelopes.

8.7. 2-*categorical aspects.* We now explain some 2-categorical aspects of Clifford eversion. As in Section 4, we do not intend this discussion to be completely rigorous. Fix a positive integer q divisible by 4, and let $\mathscr{C}$ be the 2-category of nice $\mathbb{Z}/q$-supercategories, which we regard as monoidal under the product $\hat{\boxtimes}$. Let $\omega \in \Gamma_0$ be an even S-factor system. Then Σ and ω induce a symmetric monoidal structure on $\mathscr{C}$, just as T and an odd S-factor system do (see Section 4). We write $(\mathscr{C}, \Sigma, \omega)$ for the resulting symmetric monoidal 2-category. Of course, we similarly get a symmetric monoidal 2-category from T and an odd S-factor system. The following is the main statement we want to explain, though we do not give a complete proof:

Theorem 8.16. *Clifford eversion defines an equivalence of symmetric monoidal* 2*-categories*

$$\mathrm{Cl} : (\mathscr{C}, \Sigma, \omega) \to (\mathscr{C}, \mathrm{T}, \mathtt{A}\omega).$$

Lemma 8.17. *Let $\mathcal{A}$ and $\mathcal{B}$ be nice $\mathbb{Z}/q$-supercategories. We have a canonical equivalence*

$$i_{\mathcal{A},\mathcal{B}} : \mathrm{Cl}(\mathcal{A}) \hat{\boxtimes} \mathrm{Cl}(\mathcal{B}) \to \mathrm{Cl}(\mathcal{A} \hat{\boxtimes} \mathcal{B}).$$

Proof. Let

$$j^{n,m}_{\mathcal{A},\mathcal{B}} : \mathrm{Cl}_n(\mathcal{A}) \boxtimes \mathrm{Cl}_m(\mathcal{B}) \to \mathrm{Cl}_{n+m}(\mathcal{A} \hat{\boxtimes} \mathcal{B})$$

be the functor defined by

$$(M; \mu_1, \dots, \mu_n) \boxtimes (N; \nu_1, \dots, \nu_m) \mapsto (M \boxtimes N; \mu_1, \dots, \mu_n, \nu_1, \dots, \nu_m),$$

where here we write μ_i in place of $\mu_i \otimes \mathrm{id}$ on the right side, and similarly for ν_j.

We claim that $j^{n,m}_{\mathcal{A},\mathcal{B}}$ is fully faithful. Faithfulness is clear. Let A, A' be objects of $\mathrm{Cl}_n(\mathcal{A})$ and let B, B' be objects of $\mathrm{Cl}_m(\mathcal{B})$. A homogeneous morphism $A \otimes B \to A' \otimes B'$ in $\mathrm{Cl}_{n+m}(\mathcal{A} \hat{\boxtimes} \mathcal{B})$ is an element $f = \sum_k g_k \otimes h_k \in \mathrm{Hom}_{\mathcal{A}}(A, A') \otimes \mathrm{Hom}_{\mathcal{B}}(B, B')$, where the g_k are homogeneous and linearly independent, and similarly for the h_k, such that

$$f(\mu_i \otimes 1) = (-1)^{|f|}(\mu_i' \otimes 1) f \quad \text{and} \quad f(1 \otimes \nu_j) = (-1)^{|f|}(1 \otimes \nu_j') f$$

for all i, j. The first condition becomes $\sum_k (-1)^{|h_k|} g_k \mu_i \otimes h_k = \sum_k (-1)^{|f|} \mu_i' g_k \otimes h_k$, so, by linear independence of the h_k, we deduce that $g_k \mu_i = (-1)^{|g_k|} \mu_i' g_k$ for all i, k. (Note $|f| = |g_k| + |h_k|$.) Similarly, we deduce that $h_k \nu_j = (-1)^{|h_k|} \nu_j' h_k$ for all j, k. Hence, $j^{n,m}_{\mathcal{A},\mathcal{B}}$ is full. Thus, each g_k defines a morphism

$A \to A'$ in $\mathrm{Cl}_n(\mathcal{A})$ and each h_k defines a morphism $B \to B'$ in $\mathrm{Cl}_m(\mathcal{B})$, and so f is in the image of $j^{n,m}_{\mathcal{A},\mathcal{B}}$. This proves the claim.

One easily sees (using an argument we have already made) that the $j^{n,m}_{\mathcal{A},\mathcal{B}}$ are compatible with the periodicity equivalences. They therefore induce a superfunctor

$$j_{\mathcal{A},\mathcal{B}} : \mathrm{Cl}(\mathcal{A}) \boxtimes \mathrm{Cl}(\mathcal{B}) \to \mathrm{Cl}(\mathcal{A} \hat{\boxtimes} \mathcal{B})$$

that is also fully faithful. Let

$$i_{\mathcal{A},\mathcal{B}} : \mathrm{Cl}(\mathcal{A}) \hat{\boxtimes} \mathrm{Cl}(\mathcal{B}) \to \mathrm{Cl}(\mathcal{A} \hat{\boxtimes} \mathcal{B})$$

be the functor induced by $j_{\mathcal{A},\mathcal{B}}$ by taking envelopes; the target does not change since it is already additive and Karoubian. By Lemma 8.15, $i_{\mathcal{A},\mathcal{B}}$ is fully faithful. Iterating this construction, we obtain a fully faithful functor

$$i^k_{\mathcal{A},\mathcal{B}} : \mathrm{Cl}^k(\mathcal{A}) \hat{\boxtimes} \mathrm{Cl}^k(\mathcal{B}) \to \mathrm{Cl}^k(\mathcal{A} \hat{\boxtimes} \mathcal{B})$$

for any $k \geq 0$. The diagram

$$\begin{array}{ccc} \mathrm{Cl}^{p-1}(\mathrm{Cl}(\mathcal{A})) \hat{\boxtimes} \mathrm{Cl}^{p-1}(\mathrm{Cl}(\mathcal{B})) & \xrightarrow{\quad i^p_{\mathcal{A},\mathcal{B}} \quad} & \mathrm{Cl}^{p-1}(\mathrm{Cl}(\mathcal{A} \hat{\boxtimes} \mathcal{B})) \\ & {\scriptstyle i^{p-1}_{\mathrm{Cl}(\mathcal{A}),\mathrm{Cl}(\mathcal{B})}} \searrow \quad \nearrow {\scriptstyle \mathrm{Cl}^{p-1}(i_{\mathcal{A},\mathcal{B}})} & \\ & \mathrm{Cl}^{p-1}(\mathrm{Cl}(\mathcal{A}) \hat{\boxtimes} \mathrm{Cl}(\mathcal{B})) & \end{array}$$

commutes, essentially by definition. The two categories in the top row are each identified with $\mathcal{A} \hat{\boxtimes} \mathcal{B}$ by Proposition 8.8, and $i^p_{\mathcal{A},\mathcal{B}}$ is easily seen to be identified with the identity functor; thus, the top line is an equivalence. Since the two diagonal functors are fully faithful and their composite is an equivalence, each is an equivalence. Since Cl is an equivalence of 2-categories, again by Proposition 8.8, and $\mathrm{Cl}^{p-1}(i_{\mathcal{A},\mathcal{B}})$ is an equivalence, it follows that $i_{\mathcal{A},\mathcal{B}}$ is an equivalence. □

Lemma 8.18. *The diagram*

$$\begin{array}{ccc} \mathrm{Cl}(\mathcal{A}) \hat{\boxtimes} \mathrm{Cl}(\mathcal{B}) & \xrightarrow{i_{\mathcal{A},\mathcal{B}}} & \mathrm{Cl}(\mathcal{A} \hat{\boxtimes} \mathcal{B}) \\ {\scriptstyle \mathrm{T}_{\mathrm{Cl}(\mathcal{A}),\mathrm{Cl}(\mathcal{B})}} \downarrow & & \downarrow {\scriptstyle \mathrm{Cl}(\Sigma_{\mathcal{A},\mathcal{B}})} \\ \mathrm{Cl}(\mathcal{B}) \hat{\boxtimes} \mathrm{Cl}(\mathcal{A}) & \xrightarrow{i_{\mathcal{B},\mathcal{A}}} & \mathrm{Cl}(\mathcal{B} \hat{\boxtimes} \mathcal{A}) \end{array}$$

commutes up to a canonical even isomorphism.

Proof. Let $(M; \mu_\bullet)$ be $(N; \nu_\bullet)$ be objects of $\mathrm{Cl}_m(\mathcal{A})$ and $\mathrm{Cl}_n(\mathcal{B})$. Then

$$i_{\mathcal{A},\mathcal{B}}((M; \mu_\bullet) \otimes (N; \nu_\bullet)) = (M \otimes N; \mu_\bullet, \nu_\bullet),$$

where here we have written μ_i in place of $\mu_i \otimes 1$ and ν_j in place of $1 \otimes \nu_j$. The functor $\Sigma_{\mathcal{A},\mathcal{B}} : \mathcal{A} \hat{\boxtimes} \mathcal{B} \to \mathcal{B} \hat{\boxtimes} \mathcal{A}$ induces a functor $\mathrm{Cl}_{m+n}(\Sigma_{\mathcal{A},\mathcal{B}}) : \mathrm{Cl}_{m+n}(\mathcal{A} \hat{\boxtimes} \mathcal{B}) \to \mathrm{Cl}_{m+n}(\mathcal{B} \hat{\boxtimes} \mathcal{A})$ that takes the above object to $(N \otimes M; \mu_\bullet, \nu_\bullet)$, where now μ_i means $1 \otimes \mu_i$ and ν_j means $\nu_j \otimes 1$. Note that the order of the μ and ν

does not change, since $\mathrm{Cl}_{m+n}(\Sigma_{\mathcal{A},\mathcal{B}})$ simply applies $\Sigma_{\mathcal{A},\mathcal{B}}$ to all components of the input. On the other hand,

$$\mathrm{T}((M;\mu_\bullet)\otimes(N;\nu_\bullet)) = \Pi^{mn}(N;\nu_\bullet)\otimes(M;\mu_\bullet).$$

Applying $i_{\mathcal{B},\mathcal{A}}$ to this yields

$$\Pi^{mn}(N\otimes M;\nu_\bullet,\mu_\bullet).$$

We have a natural even-degree isomorphism

$$(N\otimes M;\mu_\bullet,\nu_\bullet)\to\Pi^{mn}(N\otimes M;\nu_\bullet,\mu_\bullet)$$

coming from Proposition 8.3 using the element $\tilde{\tau}_{m,n}^{-1}$ defined in Section 5.3; note that $\tilde{\tau}_{m,n}^{-1}$ has parity mn. □

We know that Clifford eversion is an equivalence of 2-categories $\mathscr{C}\to\mathscr{C}$. The main content of Theorem 8.16 is that it admits a symmetric monoidal structure. A definition of braided monoidal 2-functor is given in [Baez and Neuchl 1996, Section 4], though this definition is not the most general possible, and only concerns braidings (not symmetries or syllepses). The above lemmas define much of the structure required for Cl to be a braided monoidal 2-functor (though we have not discussed the modification α in [ibid., Definition 16]). To complete the proof that Cl is braided monoidal, we would need to verify the quite complicated conditions in [loc. cit.]. We leave this (together with the problem of upgrading from braided monoidal to symmetric monoidal) as an open problem. We note that the particular factor systems appearing in Theorem 8.16 have not appeared in the above discussion; however, these factor systems must be the correct ones due to Proposition 8.9: precisely, if $\mathcal{A}$ is a symmetric monoidal object in $(\mathscr{C},\Sigma,\omega)$, i.e., a symmetric monoidal supercategory with respect to ω, then, by Proposition 8.9, $\mathrm{Cl}(\mathcal{A})$ is a type II supersymmetric monoidal supercategory with respect to $\mathtt{AD}\omega$, and therefore (can be converted to) a type I supersymmetric monoidal supercategory with respect to $\mathtt{A}\omega$ (via Proposition 3.34), i.e., a symmetric monoidal object of $(\mathscr{C},\mathrm{T},\mathtt{A}\omega)$.

8.8. ***Grothendieck groups and eversion.*** Fix a Cl_1-module $(\boldsymbol{k}^{1|1},\xi)$ in **SVec**. Explicitly, let e_0 be a basis vector for $\boldsymbol{k}^{1|0}$ and let e_1 be a basis vector for $\boldsymbol{k}^{0|1}$, then let ξ be the map given by $e_0\mapsto e_1$ and $e_1\mapsto 2e_0$.

Let $\mathcal{A}$ be a nice supercategory. For an object M of $\mathcal{A}$, let $F(M)$ be the Cl_1-module $(M\otimes\boldsymbol{k}^{1|1},1\otimes\xi)$ in $\mathcal{A}$; this defines a functor $F:\mathcal{A}\to\mathrm{Cl}_1(\mathcal{A})$. Let $G:\mathrm{Cl}_1(\mathcal{A})\to\mathcal{A}$ be the forgetful functor, i.e., $G(M,\alpha)=M$.

Proposition 8.19. *We have natural even isomorphisms* $G\circ F=\mathrm{id}\oplus\Pi$ *and* $F\circ G=\mathrm{id}\oplus\Pi$.

Proof. The statement for $G\circ F$ is clear. Suppose that (M,α) is a Cl_1-module in $\mathcal{A}$. Then $M\otimes\boldsymbol{k}^{1|1}$ is naturally a Cl_2-module via $\alpha=\alpha\otimes 1$ and $\xi=1\otimes\xi$. One easily verifies that $2-\xi\alpha$ defines an even isomorphism $(M\otimes\boldsymbol{k}^{1|1},\alpha)\to(M\otimes\boldsymbol{k}^{1|1},\xi)$ of Cl_1-modules; this is a variant of Proposition 8.3. Since $(M\otimes\boldsymbol{k}^{1|1},\alpha)$ is naturally isomorphic to $(M,\alpha)\oplus\Pi(M,\alpha)$ in $\mathrm{Cl}_1(\mathcal{A})$ by an even-degree isomorphism, the result follows. □

Now suppose that $\mathcal{A}$ is equipped with an exact structure. Then $\mathrm{Cl}_n(\mathcal{A})$ has a natural exact structure as well.

Proposition 8.20. *The forgetful functor* $\mathrm{Cl}_1(\mathcal{A}) \to \mathcal{A}$ *induces an isomorphism*

$$\mathrm{K}_+(\mathrm{Cl}_1(\mathcal{A}))\left[\tfrac{1}{2}\right] \to \mathrm{K}_+(\mathcal{A})\left[\tfrac{1}{2}\right].$$

Proof. This follows from Proposition 8.19 since $\mathrm{id} \oplus \Pi$ induces multiplication by 2 on K_+. □

Now suppose that $\mathcal{A}$ is a $\mathbb{Z}/2$-supercategory with an exact structure and that $\zeta_4 \in \boldsymbol{k}$. We take $p = 2$ in Clifford periodicity, so that the Clifford eversion $\mathrm{Cl}(\mathcal{A})$ is also a $\mathbb{Z}/2$-supercategory. We define a $\mathbb{Z}\left[\frac{1}{\sqrt{2}}\right]$-linear map

$$i : \mathrm{K}_+(\mathrm{Cl}(\mathcal{A})) \otimes \mathbb{Z}\left[\tfrac{1}{\sqrt{2}}\right] \to \mathrm{K}_+(\mathcal{A}) \otimes \mathbb{Z}\left[\tfrac{1}{\sqrt{2}}\right]$$

as follows. If A is an object of $\mathrm{Cl}(\mathcal{A})_0 = \mathcal{A}_0$, then $i([A]) = [A]$. If (A, α) is an object of $\mathrm{Cl}_1(\mathcal{A}) = \mathrm{Cl}_1(\mathcal{A}_1)$, then we define $i\big([(A, \alpha)]\big) = \frac{1}{\sqrt{2}}[A]$. Thus, on the odd piece we are just scaling the map induced by the forgetful functor by $\frac{1}{\sqrt{2}}$. It follows from Proposition 8.20 that i is an isomorphism of abelian groups. The following proposition explains the reason for the $\sqrt{2}$ factor:

Proposition 8.21. *If* $\mathcal{A}$ *has a monoidal structure, then* i *is a ring isomorphism.*

Proof. Let A, A' be objects of $\mathcal{A}_0$ and let (M, α), (M', α') be objects of $\mathrm{Cl}_1(\mathcal{A}_1)$. Then $A \otimes A'$ is the same when computed in $\mathrm{Cl}(\mathcal{A})$ or $\mathcal{A}$, so $i([A][A']) = i([A])i([A'])$. Similarly, $A \otimes (M, \alpha) = (A \otimes M, 1 \otimes \alpha)$, so again $i\big([A][(M, \alpha)]\big) = i([A])i\big([(M, \alpha)]\big)$.

Finally, $(M, \alpha) \otimes (M', \alpha')$ is identified via $\mathrm{Cl}_2(\mathcal{A}) \cong \mathcal{A}$ with the ζ_4-eigenspace N of $\alpha \otimes \alpha'$. The ζ_4- and $(-\zeta_4)$-eigenspaces of $M \otimes M'$ are interchanged by $\alpha \otimes 1$, and hence $(M, \alpha) \otimes (M', \alpha') \cong N \oplus \Pi(N)$, so

$$i\big([(M, \alpha)][(M', \alpha')]\big) = [N] = \tfrac{1}{2}[M][M'] = i\big([(M, \alpha)]\big)i\big([(M', \alpha')]\big).$$ □

9. Schur–Sergeev duality

We work over $\boldsymbol{k} = \mathbb{C}$ *in this section for simplicity.*

9.1. ***Polynomial representations.*** For an integer $n \geq 0$, we let $\mathfrak{q}(n) \subset \mathfrak{gl}(n|n)$ denote the super-Lie subalgebra of matrices of the form $\left[\begin{smallmatrix} A & B \\ B & A \end{smallmatrix}\right]$ where A, B are $n \times n$ matrices. This is the *isomeric Lie superalgebra*. Alternatively, if V_n is a vector space of dimension n, $\mathfrak{q}(V_n)$ is the commutator of the Cl_1-module structure on $V_{n|n} = V_n \otimes \boldsymbol{k}^{1|1}$ in $\mathfrak{gl}(V_{n|n})$. There are natural embeddings $V_{n|n} \to V_{n+1|n+1}$ which induce embeddings $\mathfrak{q}(n) \to \mathfrak{q}(n+1)$, and we define $\mathfrak{q}(\infty)$ to be the direct limit.

Let $\boldsymbol{V} = \varinjlim_n V_{n|n}$; this is the *standard representation* of $\mathfrak{q}(\infty)$. The tensor powers $V_{n|n}^{\otimes d}$ are semisimple, and the centralizer of $\mathfrak{q}(n)$ is the Hecke–Clifford algebra $\mathcal{H}_d$ (defined below) when $n \gg d$ [Wan and Wang 2012, Theorem 4.7]. The inclusion maps $V_{n|n}^{\otimes d} \to V_{n+1|n+1}^{\otimes d}$ are $\mathcal{H}_d$-linear and hence the tensor powers $\boldsymbol{V}^{\otimes d}$ are semisimple. The simple constituents of this representation are naturally indexed by strict partitions of d; we let L_λ denote the simple corresponding to the partition λ [ibid., Theorem 4.8]. We say that a representation of $\mathfrak{q}(\infty)$ is *polynomial* if it is a direct sum of L_λ's. The category of polynomial representations is denoted by $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$. It is naturally a symmetric monoidal $\mathbb{Z}$-supercategory. The object L_λ is homogeneous of degree $|\lambda|$.

Remark 9.1. Even though $\Pi(\boldsymbol{V}) \cong \boldsymbol{V}$, we do not have $\Pi(V) \cong V$ for all objects V in $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$. More precisely, the endomorphism algebra of the irreducible L_λ, where λ is a strict partition, is Cl_1 if and only if $\ell(\lambda)$ is odd [Cheng and Wang 2012, Lemma 1.40]. □

9.2. ***The Hecke–Clifford algebra.*** The symmetric group $\mathfrak{S}_n$ acts on the algebra Cl_n by $g \cdot \alpha_i = \alpha_{g(i)}$. The *Hecke–Clifford algebra* is $\mathcal{H}_n = \mathbb{C}[\mathfrak{S}_n] \ltimes \mathrm{Cl}_n$. This is a superalgebra via the grading $\deg(g) = 0$ for $g \in \mathfrak{S}_n$ and $\deg(\alpha_i) = 1$. We define a category $\mathrm{Rep}(\mathcal{H}_*)$ whose objects are sequences $(V_n)_{n\geq 0}$ where V_n is a finite-dimensional super-vector space with a homogeneous action of $\mathcal{H}_n$ such that $V_n = 0$ for $n \gg 0$. We define $\Pi : \mathrm{Rep}(\mathcal{H}_*) \to \mathrm{Rep}(\mathcal{H}_*)$ by $(\Pi V)_n = \Pi(V_n)$. There is a natural identification of $\mathcal{H}_{n,m} = \mathcal{H}_n \otimes \mathcal{H}_m$ with a subalgebra of $\mathcal{H}_{n+m}$, and we define a monoidal structure on $\mathrm{Rep}(\mathcal{H}_*)$ via

$$(V \otimes W)_n = \bigoplus_{i=0}^{n} \mathcal{H}_n \otimes_{\mathcal{H}_{i,n-i}} (V_i \boxtimes W_{n-i}).$$

We define a symmetry β' on this tensor product by

$$a \otimes v \otimes w \mapsto (-1)^{|v||w|} a\tau_{i,n-i}^{-1} \otimes w \otimes v,$$

where $a \in \mathcal{H}_n$, $v \in V_i$ and $w \in W_{n-i}$. We omit the straightforward check that this is well defined.

The Hecke–Clifford algebra is closely related to the isomeric superalgebra. Let $\xi : \boldsymbol{V} \to \boldsymbol{V}$ be the given odd endomorphism. The algebra Cl_n has a right action on $\boldsymbol{V}^{\otimes n}$ via

$$(v_1 \otimes \cdots \otimes v_n)\alpha_i = (-1)^{|v_{i+1}|+\cdots+|v_n|}(v_1 \otimes \cdots \otimes \xi(v_i) \otimes \cdots \otimes v_n),$$

where the v_i are homogeneous elements of $\boldsymbol{V}$. The symmetric group $\mathfrak{S}_n$ also has a right action on $\boldsymbol{V}^{\otimes n}$, in the usual manner. Together, these define a right action of the Hecke–Clifford algebra $\mathcal{H}_n$ on $\boldsymbol{V}^{\otimes n}$. The following is a typical formulation of Schur–Sergeev duality:

Proposition 9.2. *The functor* $\mathrm{Rep}(\mathcal{H}_*) \to \mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$ *given by* $(N_n)_n \mapsto \bigoplus_n \boldsymbol{V}^{\otimes n} \otimes_{\mathcal{H}_n} N_n$ *is an equivalence of symmetric monoidal* $\mathbb{Z}$*-supercategories.*

Proof. Let Φ denote the functor under discussion. The explicit decomposition of $\boldsymbol{V}^{\otimes n}$ given in [Wan and Wang 2012, Theorem 4.8] shows that Φ is an equivalence of abelian categories. Let M and N be $\mathcal{H}_m$- and $\mathcal{H}_n$-modules, respectively. Then

$$\begin{aligned}
\Phi(M \otimes N) &= \boldsymbol{V}^{\otimes(m+n)} \otimes_{\mathcal{H}_{m+n}} (\mathcal{H}_{m+n} \otimes_{\mathcal{H}_{m,n}} (M \otimes N)) \\
&\cong (\boldsymbol{V}^{\otimes m} \otimes \boldsymbol{V}^{\otimes n}) \otimes_{\mathcal{H}_{m,n}} (M \otimes N) \\
&\cong (\boldsymbol{V}^{\otimes m} \otimes_{\mathcal{H}_m} M) \otimes (\boldsymbol{V}^{\otimes n} \otimes_{\mathcal{H}_n} N) \\
&= \Phi(M) \otimes \Phi(N),
\end{aligned}$$

which shows that Φ is monoidal. Finally, we verify the symmetry. Let β be the symmetry on $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$. Pick $v \otimes x \in \Phi(M)$ and $v' \otimes y \in \Phi(N)$, where $x \in M$, $y \in N$, $v \in \boldsymbol{V}^{\otimes m}$ and $v' \in \boldsymbol{V}^{\otimes n}$ are homogeneous.

Then

$$\begin{aligned}\beta((v\otimes x)\otimes(v'\otimes y)) &= (-1)^{(|v|+|x|)(|v'|+|y|)}(v'\otimes y)\otimes(v\otimes x)\\ &= (-1)^{|v||v'|+|x||v'|+|x||y|}(v'\otimes v)\otimes(y\otimes x).\end{aligned}$$

On the other hand, under the identification $\Phi(M)\otimes\Phi(N)\cong\Phi(M\otimes N)$, $(v\otimes x)\otimes(v'\otimes y)$ maps to $(-1)^{|v'||x|}(v\otimes v')\otimes(1\otimes x\otimes y)$, where 1 is the unit of $\mathcal{H}_{m+n}$. If we apply the symmetry β' in $\mathrm{Rep}(\mathcal{H}_*)$ to the second factor, then

$$(-1)^{|v'||x|+|x||y|}(v\otimes v')\otimes(\tau_{m,n}^{-1}\otimes y\otimes x) = (-1)^{|v'||x|+|x||y|+|v||v'|}(v'\otimes v)\otimes(1\otimes y\otimes x),$$

which agrees with the previous expression, so Φ is symmetric. □

9.3. ***Sergeev–Schur duality.*** Let β be the type II supersymmetry on $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ obtained from the type I supersymmetry described in Section 6.6 by the general process in Section 3.7. Thus, explicitly, if M is a representation of $\widetilde{\mathfrak{S}}_m$ and N is a representation of $\widetilde{\mathfrak{S}}_n$, then

$$\beta_{M,N}:\mathrm{Ind}_{\widetilde{\mathfrak{S}}_m\rtimes\widetilde{\mathfrak{S}}_n}^{\widetilde{\mathfrak{S}}_{n+m}}(M\otimes N)\to\mathrm{Ind}_{\widetilde{\mathfrak{S}}_n\rtimes\widetilde{\mathfrak{S}}_m}^{\widetilde{\mathfrak{S}}_{n+m}}(N\otimes M)$$

is given by

$$g\otimes v\otimes w\mapsto(-1)^{|v||w|+mn|g|+mn}g\tilde{\tau}_{m,n}^{-1}\otimes w\otimes v.$$

By results in the aforementioned sections, β is a type II supersymmetry with respect to the factor system $\mathtt{DA}$. Define β^* by

$$\beta^*_{M,N} = (-1)^{\binom{mn}{2}}\zeta_4^{mn}\beta_{M,N},$$

where $m=|M|$ and $n=|N|$. Note that $(-1)^{\binom{mn}{2}}\zeta_4^{mn}=\mathtt{d}(m,n)$, where $\mathtt{d}$ is as defined in Example 3.14, and so β^* is a type II supersymmetry with respect to $\mathtt{A}$.

The following is our reformulation of Sergeev–Schur duality, and the main result of this section:

Theorem 9.3. *We have a natural equivalence of symmetric monoidal $\mathbb{Z}$-supercategories*

$$\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})\cong\mathrm{Cl}(\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)),$$

where here $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ *is equipped with the supersymmetry* β^*.

The proof will take the rest of the section. We omit the proof of the following fact, which is a straightforward calculation:

Lemma 9.4. *We have an isomorphism of superalgebras*

$$\mathbb{C}[\widetilde{\mathfrak{S}}_n]/(c+1)\otimes\mathrm{Cl}_n\to\mathcal{H}_n$$

given by $1\otimes\alpha_i\mapsto\alpha_i$ *and* $\tilde{s}_j\otimes 1\mapsto\frac{1}{2}\zeta_4 s_j(\alpha_j-\alpha_{j+1})$.

Proposition 9.5. *We have an equivalence of symmetric monoidal $\mathbb{Z}$-supercategories*

$$\mathrm{Cl}(\mathrm{Rep}(\widetilde{\mathfrak{S}}_*))\cong\mathrm{Rep}(\mathcal{H}_*).$$

Proof. Via the isomorphism in Lemma 9.4, an $\mathcal{H}_n$-module M is the same thing as a Cl_n-module in $\mathrm{Rep}(\widetilde{\mathfrak{S}}_n)$. This gives an equivalence $\Phi : \mathrm{Rep}(\mathcal{H}_*) \to \mathrm{Cl}(\mathrm{Rep}(\widetilde{\mathfrak{S}}_*))$ which preserves the $\mathbb{Z}$-grading. We claim that this is a monoidal equivalence. Let M be an $\mathcal{H}_m$-module and let N be an $\mathcal{H}_n$-module. We have $\mathrm{Cl}_m \otimes \mathrm{Cl}_n \cong \mathrm{Cl}_{m+n}$, so $\mathcal{H}_{m+n} \otimes_{\mathcal{H}_{m,n}} (M \otimes N)$ is isomorphic, as a vector space, to $\mathrm{Ind}_{\widetilde{\mathfrak{S}}_m \ltimes \widetilde{\mathfrak{S}}_n}^{\widetilde{\mathfrak{S}}_{m+n}} (M \otimes N)$. The action of Cl_m is via M and the action of Cl_n is via N, so that the action of Cl_{m+n} is via the tensor product action, and hence Φ is a monoidal equivalence.

We now check symmetry. We have homomorphisms

$$\varphi : \widetilde{\mathfrak{S}}_{n+m} \to \mathcal{H}_{n+m}, \quad \tilde{s}_i \mapsto \tfrac{1}{2}\zeta_4 s_i(\alpha_{i+1} - \alpha_i),$$
$$\psi : \widetilde{\mathfrak{S}}_{n+m} \to \mathcal{H}_{n+m}, \quad \tilde{s}_i \mapsto \tfrac{1}{2}(\alpha_{i+1} - \alpha_i).$$

The existence of ψ comes from Proposition 8.1, while the existence of φ comes from Lemma 9.4. Observe that $\varphi(\tilde{s}_i)$ skew-commutes with α_j, for all i and j, and so $\varphi(\tilde{s}_i)$ and $\psi(\tilde{s}_j)$ skew-commute for all i and j. We also have $\varphi(\tilde{s}_i) = \zeta_4 s_i \psi(\tilde{s}_i)$, and so $\varphi(\tilde{s}_i)\psi(\tilde{s}_i) = \zeta_4 s_i$. It follows that

$$\tau_{m,n} = (-1)^{\binom{mn}{2}} \zeta_4^{mn} \varphi(\tilde{\tau}_{m,n}) \psi(\tilde{\tau}_{m,n}).$$

The vector space isomorphism

$$\mathrm{Ind}_{\widetilde{\mathfrak{S}}_m \ltimes \widetilde{\mathfrak{S}}_n}^{\widetilde{\mathfrak{S}}_{m+n}} (M \otimes N) \to \mathcal{H}_{m+n} \otimes_{\mathcal{H}_{m,n}} (M \otimes N)$$

is given by $g \otimes v \otimes w \mapsto \varphi(g) \otimes v \otimes w$. We thus have

$$\begin{aligned}
\beta'_{M,N}(\varphi(g) \otimes v \otimes w) &= (-1)^{|v||w|} \varphi(g) \tau_{m,n}^{-1} \otimes w \otimes v \\
&= (-1)^{|v||w| + \binom{mn}{2}} \zeta_4^{-mn} \varphi(g) \psi(\tilde{\tau}_{m,n})^{-1} \varphi(\tilde{\tau}_{m,n})^{-1} \otimes w \otimes v \\
&= (-1)^{|v||w| + \binom{mn}{2} + mn|g|} \zeta_4^{-mn} \psi(\tilde{\tau}_{m,n})^{-1} \varphi(g \tilde{\tau}_{m,n}^{-1}) \otimes w \otimes v \\
&= \psi(\tilde{\tau}_{m,n})^{-1} \beta^*(g \otimes v \otimes w).
\end{aligned}$$

The expression on the final line is exactly the symmetry on $\mathrm{Cl}(\mathrm{Rep}(\widetilde{\mathfrak{S}}_*))$, and so the result follows. □

Theorem 9.3 now follows from Propositions 9.5 and 9.2.

9.4. ***Symmetric functions.*** In this section, $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ denotes an integer partition, $\ell(\lambda)$ denotes the number of nonzero entries of λ, and $|\lambda| = \sum_i \lambda_i$.

Let Λ be the ring of symmetric functions in an alphabet $x_1, x_2, \ldots$. Define symmetric functions $q_k(x)$ by the identity

$$\sum_{k \geq 0} q_k(x) t^k = \prod_{i \geq 1} \frac{1 + x_i t}{1 - x_i t}.$$

Let Γ be the subring of Λ generated by $q_1, q_2, \ldots$. They satisfy the relations, for $n > 0$,

$$\sum_{i=0}^{n} (-1)^i q_i q_{n-i} = 0.$$

So $\Gamma \otimes \mathbb{Z}\left[\frac{1}{2}\right]$ is generated by $q_1, q_3, q_5, \ldots$, which are algebraically independent. For a, b, define

$$Q_{(a,b)} = q_a q_b + 2\sum_{i=1}^{b} (-1)^i q_{a+i} q_{b-i}$$

(with the convention that $q_i = 0$ if $i < 0$). Note that $Q_{(b,a)} = -Q_{(a,b)}$. Let $\lambda = (\lambda_1 > \cdots > \lambda_{2r} \geq 0)$ be a strict partition (here $\ell(\lambda) = 2r$ if $\lambda_{2r} > 0$ and otherwise $\ell(\lambda) = 2r - 1$); we define

$$Q_\lambda = \mathrm{Pf}(Q_{(\lambda_i, \lambda_j)})_{1 \leq i, j \leq 2r},$$

where Pf denotes the Pfaffian of a skew-symmetric matrix. The Q_λ form a $\mathbb{Z}$-basis for Γ [Macdonald 1995, (8.9)] and $\Gamma \otimes \mathbb{Z}\left[\frac{1}{\sqrt{2}}\right]$ is a model for the Grothendieck groups $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$, $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$ and $\mathrm{Rep}(\mathcal{H}_*)$ (in each case, the ring structure comes from the monoidal product) as we now explain. In particular, in each case, the simple objects are indexed by strict partitions.

For $\mathrm{Rep}^{\mathrm{pol}}(\mathfrak{q})$ and $\mathrm{Rep}(\mathcal{H}_*)$, the classes of the simple objects L_λ are $2^{-\lfloor \ell(\lambda)/2 \rfloor} Q_\lambda$ [Cheng and Wang 2012, Theorem 3.48], and, for $\mathrm{Rep}(\widetilde{\mathfrak{S}}_*)$, the classes of the simple objects $\mathbb{N}_\lambda$ are $2^{(\varepsilon(\lambda) - \ell(\lambda))/2} Q_\lambda$, where $\varepsilon(\lambda) \in \{0, 1\}$ is the reduction modulo 2 of $|\lambda| - \ell(\lambda)$ [Hoffman and Humphreys 1992, Theorem 8.6].

In particular, when $|\lambda|$ is even, $[L_\lambda] = [\mathbb{N}_\lambda]$, and, when $|\lambda|$ is odd,

$$[L_\lambda] = \begin{cases} \frac{1}{\sqrt{2}}[\mathbb{N}_\lambda] & \text{if } \ell(\lambda) \text{ is even,} \\ \sqrt{2}[\mathbb{N}_\lambda] & \text{if } \ell(\lambda) \text{ is odd.} \end{cases}$$

Recall from Proposition 9.5 that we have an equivalence $\mathrm{Rep}(\mathcal{H}_n) \simeq \mathrm{Cl}_1(\mathrm{Rep}(\widetilde{\mathfrak{S}}_n))$ when n is odd. Hence, from Proposition 8.21, we have an isomorphism

$$\mathrm{K}_+\big(\mathrm{Cl}_1(\mathrm{Rep}(\widetilde{\mathfrak{S}}_n))\big) \otimes \mathbb{Z}\left[\tfrac{1}{\sqrt{2}}\right] \to \mathrm{K}_+(\mathrm{Rep}(\widetilde{\mathfrak{S}}_n)) \otimes \mathbb{Z}\left[\tfrac{1}{\sqrt{2}}\right], \quad [(A, \alpha)] \mapsto \tfrac{1}{\sqrt{2}}[A].$$

To match this with the Q-functions, we note that $\mathbb{N}_\lambda$ is an isomeric object if and only if $|\lambda| - \ell(\lambda)$ is odd [Józefiak 1991, Corollary 4.19]. Hence, when $|\lambda|$ is odd, the simple objects of $\mathrm{Cl}_1(\mathrm{Rep}(\widetilde{\mathfrak{S}}_n))$ are

$$\{\mathbb{N}_\lambda \mid \ell(\lambda) \text{ is even}\} \cup \{\mathbb{N}_\lambda \otimes \boldsymbol{k}^{1|1} \mid \ell(\lambda) \text{ is odd}\}$$

and the correspondence becomes

$$[L_\lambda] \mapsto \begin{cases} \frac{1}{\sqrt{2}}[\mathbb{N}_\lambda] & \text{if } \ell(\lambda) \text{ is even,} \\ \frac{1}{\sqrt{2}}[\mathbb{N}_\lambda \otimes \boldsymbol{k}^{1|1}] & \text{if } \ell(\lambda) \text{ is odd.} \end{cases}$$

Appendix: Bott periodicity

Let $\boldsymbol{k}$ be a field of characteristic different from 2. In this paper, the Clifford algebra Cl_k is the $\boldsymbol{k}$-superalgebra generated by odd elements $\alpha_1, \ldots, \alpha_k$ subject to the relations $\alpha_i^2 = 2$ for all i and $\alpha_i \alpha_j = -\alpha_j \alpha_i$ for all $i \neq j$. Our goal here is to establish a version of Bott periodicity (Theorem A.4) for these Clifford algebras. While this is well known, we have provided the details since most references work over $\boldsymbol{k} = \mathbb{R}$ and our definition of the algebra is not standard. Here we follow the proof in [Trimble 2005].

Given a superalgebra A, the opposite superalgebra A^{op} has the same underlying vector space but multiplication of homogeneous elements is given by $a \cdot_{\mathrm{op}} b = (-1)^{|a||b|} b \cdot a$. In particular, the presentation of $\mathrm{Cl}_k^{\mathrm{op}}$ is the same as Cl_k except that the α_i square to -2 instead of 2. Also, note that $\mathrm{Cl}_k \otimes \mathrm{Cl}_\ell \cong \mathrm{Cl}_{k+\ell}$, where we emphasize that we use the tensor product of superalgebras, so that, on homogeneous elements, $(x_1 \otimes y_1)(x_2 \otimes y_2) = (-1)^{|x_2||y_1|} x_1 x_2 \otimes y_1 y_2$.

Lemma A.1. *We have an isomorphism of superalgebras* $\mathrm{Cl}_1 \otimes \mathrm{Cl}_1^{\mathrm{op}} \cong \mathrm{End}(\boldsymbol{k}^{1|1})$.

Proof. Let e_0, e_1 be a homogeneous basis for $\boldsymbol{k}^{1|1}$. We represent $\alpha_1 \otimes 1$ by the operator $e_0 \mapsto e_1$, $e_1 \mapsto 2e_0$, and $1 \otimes \alpha_1$ by the operator $e_0 \mapsto e_1$, $e_1 \mapsto -2e_0$. □

Let H denote the quaternion algebra concentrated in degree 0, i.e., H is the associative $\boldsymbol{k}$-algebra generated by even elements $\boldsymbol{i}$ and $\boldsymbol{j}$ such that $\boldsymbol{ij} = -\boldsymbol{ji}$ and $\boldsymbol{i}^2 = \boldsymbol{j}^2 = -1$.

Lemma A.2. *We have an isomorphism of superalgebras* $\mathrm{Cl}_3 \cong H \otimes \mathrm{Cl}_1^{\mathrm{op}}$.

Proof. We define $H \otimes \mathrm{Cl}_1^{\mathrm{op}} \to \mathrm{Cl}_3$ by $\boldsymbol{i} \otimes 1 \mapsto \frac{1}{2}\alpha_1\alpha_2$, $\boldsymbol{j} \otimes 1 \mapsto \frac{1}{2}\alpha_1\alpha_3$ and $1 \otimes \alpha_1 \mapsto \frac{1}{2}\alpha_1\alpha_2\alpha_3$. It is readily verified to be surjective and both algebras have dimension 8 and hence it is an isomorphism. □

Lemma A.3. *We have an isomorphism of superalgebras* $\mathrm{Cl}_4 \cong \mathrm{Cl}_4^{\mathrm{op}}$.

Proof. By Lemmas A.1 and A.2,

$$\mathrm{Cl}_4 \cong \mathrm{Cl}_3 \otimes \mathrm{Cl}_1 \cong H \otimes \mathrm{Cl}_1^{\mathrm{op}} \otimes \mathrm{Cl}_1 \cong H \otimes \mathrm{End}(\boldsymbol{k}^{1|1}).$$

Next, $H \cong H^{\mathrm{op}}$ via $\boldsymbol{i} \mapsto -\boldsymbol{i}$ and $\boldsymbol{j} \mapsto -\boldsymbol{j}$ and $\mathrm{End}(\boldsymbol{k}^{1|1}) \cong \mathrm{End}(\boldsymbol{k}^{1|1})^{\mathrm{op}}$ via $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} a & -c \\ b & d \end{smallmatrix}\right]$, where the matrices are given with respect to a homogeneous basis of $\boldsymbol{k}^{1|1}$. □

Theorem A.4 (Bott periodicity). *We have an isomorphism of superalgebras* $\mathrm{Cl}_8 \cong \mathrm{End}(\boldsymbol{k}^{8|8})$.

Furthermore, if $\boldsymbol{k}$ *contains a primitive fourth root of unity* ζ_4, *then* $\mathrm{Cl}_2 \cong \mathrm{End}(\boldsymbol{k}^{1|1})$.

Proof. By Lemmas A.1 and A.3,

$$\mathrm{Cl}_8 \cong \mathrm{Cl}_4 \otimes \mathrm{Cl}_4 \cong \mathrm{Cl}_4 \otimes \mathrm{Cl}_4^{\mathrm{op}} \cong \mathrm{End}(\boldsymbol{k}^{1|1})^{\otimes 4} \cong \mathrm{End}(\boldsymbol{k}^{8|8}).$$

The second statement follows from Lemma A.1 since $\mathrm{Cl}_1 \cong \mathrm{Cl}_1^{\mathrm{op}}$ via $\alpha_1 \mapsto \zeta_4\alpha_1$. □

The periodicity comes when we consider Morita equivalence. Since $\mathrm{End}(\boldsymbol{k}^{m|n})$ is a simple superalgebra for any m, n, its category of (super)modules is equivalent to the category of super-vector spaces, and hence $\mathrm{Mod}_{\mathrm{Cl}_{k+8}} \simeq \mathrm{Mod}_{\mathrm{Cl}_k}$ for any k and $\mathrm{Mod}_{\mathrm{Cl}_{k+2}} \simeq \mathrm{Mod}_{\mathrm{Cl}_k}$ when $\zeta_4 \in \boldsymbol{k}$.

References

[Baez and Neuchl 1996] J. C. Baez and M. Neuchl, "Higher-dimensional algebra, I: Braided monoidal 2-categories", *Adv. Math.* **121**:2 (1996), 196–244. MR

[Brundan and Ellis 2017] J. Brundan and A. P. Ellis, "Monoidal supercategories", *Comm. Math. Phys.* **351**:3 (2017), 1045–1089. MR

[Cheng and Wang 2012] S.-J. Cheng and W. Wang, *Dualities and representations of Lie superalgebras*, Graduate Studies in Math. **144**, Amer. Math. Soc., Providence, RI, 2012. MR

[Crans 1998] S. E. Crans, "Generalized centers of braided and sylleptic monoidal 2-categories", *Adv. Math.* **136**:2 (1998), 183–223. MR

[Ellis and Lauda 2016] A. P. Ellis and A. D. Lauda, "An odd categorification of $U_q(\mathfrak{sl}_2)$", *Quantum Topol.* **7**:2 (2016), 329–433. MR

[Ganter and Kapranov 2014] N. Ganter and M. Kapranov, "Symmetric and exterior powers of categories", *Transform. Groups* **19**:1 (2014), 57–103. MR

[Hoffman and Humphreys 1992] P. N. Hoffman and J. F. Humphreys, *Projective representations of the symmetric groups: Q-functions and shifted tableaux*, Clarendon, New York, 1992. MR

[Józefiak 1991] T. Józefiak, "Schur Q-functions and cohomology of isotropic Grassmannians", *Math. Proc. Cambridge Philos. Soc.* **109**:3 (1991), 471–478. MR

[Kang et al. 2016] S.-J. Kang, M. Kashiwara, and S. Tsuchioka, "Quiver Hecke superalgebras", *J. Reine Angew. Math.* **711** (2016), 1–54. MR

[Kapranov 2018] M. Kapranov, "Supergeometry in mathematics and physics", preprint, 2018. arXiv 1512.07042v2

[Kapranov and Voevodsky 1994] M. M. Kapranov and V. A. Voevodsky, "2-categories and Zamolodchikov tetrahedra equations", pp. 177–259 in *Algebraic groups and their generalizations: quantum and infinite-dimensional methods* (University Park, PA, 1991), edited by W. J. Haboush and B. J. Parshall, Proc. Sympos. Pure Math. **56.2**, Amer. Math. Soc., Providence, RI, 1994. MR

[Macdonald 1995] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Clarendon, New York, 1995. MR

[MacLane 1952] S. MacLane, "Cohomology theory of abelian groups", pp. 8–14 in *Proceedings of the International Congress of Mathematicians* (Cambridge, MA, 1950), vol. II, edited by L. M. Graves et al., Amer. Math. Soc., Providence, RI, 1952. MR

[Trimble 2005] T. Trimble, "The super Brauer group and super division algebras", webpage, 2005, available at http://math.ucr.edu/home/baez/trimble/superdivision.html.

[Wan and Wang 2012] J. Wan and W. Wang, "Lectures on spin representation theory of symmetric groups", *Bull. Inst. Math. Acad. Sin.* (*N.S.*) **7**:1 (2012), 91–164. MR

Communicated by Antoine Chambert-Loir
Received 2021-04-15 Revised 2024-04-05 Accepted 2025-08-05

ssam@ucsd.edu *Department of Mathematics, University of California, San Diego, La Jolla, CA, United States*
http://math.ucsd.edu/~ssam/

asnowden@umich.edu *Department of Mathematics, University of Michigan, Ann Arbor, MI, United States*
http://www-personal.umich.edu/~asnowden/

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1543

Rational curves and the Hilbert property on Jacobian Kummer varieties

Damián Gvirtz-Chen and Zhizhong Huang

Let K be a finitely generated field of characteristic zero. A conjecture by Corvaja and Zannier predicts that smooth, projective, simply connected varieties over K with Zariski dense set of rational points have the Hilbert property. We confirm this conjecture for all Kummer surfaces associated to the Jacobian of a genus 2 curve over K, and in general for all Jacobian Kummer varieties associated to a hyperelliptic curve of genus ≥ 3 of odd degree over K.

1. Introduction

The aim of this article is to introduce the use of rational curves into the study of the Hilbert property for nonunirational varieties.

1.1. ***Background for the Hilbert property.*** Let X be a quasiprojective integral variety of dimension at least one over a field K. Recall after [Serre 2008, §3] (see also [Colliot-Thélène 2020, §2]) that a subset $A \subset X(K)$ is *thin* if there exists a K-morphism $f : Y \to X$ from a K-variety Y to X such that

- $A \subset f(Y(K))$;
- the fibre of f over the generic point of X is finite, and f does not have any rational sections.

Definition 1.1 [Colliot-Thélène and Sansuc 1987, p. 189]. We say that X has the *Hilbert property* (or X is *of Hilbert type*) if $X(K)$ is not a thin set.

The field K is called *Hilbertian* if the projective line $\mathbb{P}^1_K$ satisfies the Hilbert property. Not every base field is Hilbertian. The field of rational numbers, and, more generally, all number fields, are classically known to be Hilbertian by the Hilbert irreducibility theorem [1892] (see [Serre 2008, Theorem 3.4.3]). We will assume, unless explicitly stated to the contrary, that K is a finitely generated field of characteristic 0. These fields are known to be Hilbertian (see, e.g., [Lang 1983, Chapter 9; Fried and Jarden 2008, §13.4]).

It should be pointed out that Definition 1.1 looks different from the one introduced later in [Corvaja and Zannier 2017, §1.1, Definition; Corvaja et al. 2022, Definition 1.1]. We shall provide a proof (see Proposition 2.1 and Corollary 2.2 below) that all these definitions are actually equivalent.

It is clear from the definition that Zariski density of $X(K)$ in X is a necessary condition for the Hilbert property, and that the Hilbert property is a birational invariant. Moreover, if X satisfies the Hilbert property,

MSC2020: 11G35, 14G05, 14J28.
Keywords: Hilbert property, thin sets, Jacobian Kummer varieties.

© 2026 The Authors, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

then it is geometrically integral (see [Serre 2008, p. 20]). The Hilbert property is a specialisation of weak weak approximation (see [Serre 2008, §3.5, Corollary 3.5.4, Definition 3.5.6] and also [Colliot-Thélène 2020, §2 (AFF), Théorème 2.3]) and provides an approach to solving the inverse Galois problem (see for example [Serre 2008], notably Theorem 3.5.9).

There is a well-known topological obstruction to the Hilbert property which goes back to the Chevalley–Weil theorem. Namely, a projective variety X cannot have the Hilbert property if it is not algebraically simply connected (or, equivalently for normal X, if its fundamental group has a proper subgroup of finite index); see [Corvaja and Zannier 2017, Theorem 1.4]. They conjecture that the converse is true:

Conjecture 1.2 [Corvaja and Zannier 2017, §2, Question–Conjecture 1]. *Let X be a smooth, projective and algebraically simply connected variety over K. If $X(K)$ is Zariski dense, then X has the Hilbert property.*[1]

While there have been quite a number of results confirming the weak weak approximation property for unirational varieties (they are simply connected by [Serre 1959]), as conjectured by Colliot-Thélène (see [Colliot-Thélène and Sansuc 1987; Serre 2008, Conjecture 3.5.8]), progress towards the Hilbert property for intermediate type varieties has been very slow. In particular, Corvaja and Zannier [2017, Appendix 1] ask for a proof that the Hilbert property holds for Kummer surfaces with a Zariski dense set of rational points. (See [Harpaz and Skorobogatov 2016, p. 817] for a stronger conjecture of weak weak approximation on K3 surfaces and [Holmes and Pannekoek 2015] for a consequence on the rank of quadratic twists of abelian varieties.)

1.2. ***Main results and discussion.*** Let H be a smooth, projective hyperelliptic curve of genus $g \geq 2$ over a finitely generated field K of characteristic 0. Let $\mathrm{Jac}(H)$ be its Jacobian variety and let $\mathrm{Kum}(\mathrm{Jac}(H))$ be the associated Kummer quotient, i.e., the quotient of $\mathrm{Jac}(H)$ by the natural involution. The minimal desingularisation of $\mathrm{Kum}(\mathrm{Jac}(H))$ is a simply connected variety of dimension g, called a *Jacobian Kummer variety* (see Section 3.1 for details).

In dimension two, all principally polarised abelian surfaces over K are isomorphic to either the Jacobian of a genus 2 curve or, up to a quadratic extension, the product of two elliptic curves (see [González et al. 2005, Theorem 3.1]). The Jacobian case is the generic one, in the sense of forming a Zariski dense open subvariety of the corresponding moduli space. Our main theorem confirms Conjecture 1.2 in the Jacobian case.

Theorem 1.3. *Let H be a smooth, projective genus 2 curve over K. Then the variety* $\mathrm{Kum}(\mathrm{Jac}(H))$ *has the Hilbert property.*[2]

The study of the Hilbert property for K3 surfaces was first addressed by Corvaja and Zannier [2017], and they verified it for the diagonal quartic surface $x^4 + y^4 = z^4 + w^4$ [ibid., Theorem 1.6] using multiple

[1] In [Corvaja and Zannier 2017, §2, Question–Conjecture 1bis], one finds an affine analogue of this conjecture concerning integral points. However, we are informed by Prof. Corvaja that in some cases of the analogue a finite field extension is necessary. Whether this is also the case for Conjecture 1.2 is unclear to the authors.

[2] Note that any smooth genus 2 curve is hyperelliptic.

elliptic fibrations. The same method was formalised and applied by Demeio [2021, Proposition 4.2.1]. He establishes the Hilbert property for the Kummer surface $\mathrm{Kum}(E_1 \times E_2)$ associated to the product of two elliptic curves E_1, E_2 over K, under the assumption that rational points on $\mathrm{Kum}(E_1 \times E_2)$ are Zariski dense. (Demeio's theorem is stated over a number field and contains the stronger assumption that the elliptic curves have positive Mordell–Weil ranks. In fact, the proof of [ibid., Theorem 1.0.2] works for any finitely generated field K of characteristic 0 and only assumes that the set of rational points is Zariski dense. This is known to hold unconditionally unless $j(E_1) = j(E_2) \in \{0, 1728\}$, thanks to [Kuwata and Wang 1993].)

Huang [2021] studied various Kummer surfaces associated to torsors under the product of two elliptic curves, though the Hilbert property for such surfaces remains still open. More recently, Gvirtz-Chen and Mezzedimi [2023] have shown that the Hilbert property holds for Kummer surfaces if one allows a finite field extension of the base field K.

The elliptic fibrations (modulo automorphisms) of a generic Jacobian Kummer surface (i.e., $\mathrm{Jac}(H)$ has no extra endomorphisms) over an algebraically closed field have been classified by Kumar [2014]. Unfortunately, in general none of these fibrations descend to K and, thus, it is not possible to apply directly the method of [Demeio 2021] over K. Instead, the novelty in our proof of Theorem 1.3 is, amongst other things, the use of rational curves on $\mathrm{Kum}(\mathrm{Jac}(H))$. To our knowledge, this is the first time that rational curves have been used to prove the Hilbert property for nonunirational varieties.

In fact, our approach allows us to settle Conjecture 1.2 for a family of higher-dimensional varieties. We are not aware of any previous results on the Hilbert property for nonunirational varieties of dimension > 2, apart from naive constructions like product varieties.

Theorem 1.4. *Let H be a smooth, projective hyperelliptic curve of genus $g \geq 3$ and of odd degree over K, i.e., H is defined by an affine model $y^2 = f(x)$ with f separable of degree $2g + 1$. Then the variety* $\mathrm{Kum}(\mathrm{Jac}(H))$ *has the Hilbert property.*

1.3. ***Strategy and outline.*** In order to reduce the a priori unknown Hilbert property of a higher-dimensional variety X to the known Hilbert property of (possibly singular) rational subcurves on X, their behaviour under pullbacks along covers of X has to be understood. This is the criterion of Proposition 5.1, which we formulate in a way tailored to our application to Jacobian Kummer varieties.

Let $\mathrm{Alb}^1(H)$ be the (first) Albanese torsor of the hyperelliptic curve H, and let $\mathrm{Kum}(\mathrm{Alb}^1(H))$ be the associated Kummer quotient, i.e., the quotient of $\mathrm{Alb}^1(H)$ by the natural involution. See Section 3.2 for details. In general, $\mathrm{Kum}(\mathrm{Alb}^1(H))$ is a nonisomorphic Galois twist of $\mathrm{Kum}(\mathrm{Jac}(H))$. In practice, it turns out that working with $\mathrm{Kum}(\mathrm{Alb}^1(H))$ instead of $\mathrm{Kum}(\mathrm{Jac}(H))$ is crucial for our purposes. The most general form of our main theorem is the following, from which Theorems 1.3 and 1.4 follow as special cases:

Theorem 1.5. *Let H be a smooth, projective hyperelliptic curve over K. Then the variety* $\mathrm{Kum}(\mathrm{Alb}^1(H))$ *has the Hilbert property.*

The variety $\mathrm{Kum}(\mathrm{Alb}^1(H))$ always possesses a countable family of rational curves whose union is Zariski dense. To show that the conditions in the criterion of Proposition 5.1 are satisfied for $\mathrm{Kum}(\mathrm{Alb}^1(H))$, we prove pullback lemmas that deal with the three occurring types of covers of $\mathrm{Kum}(\mathrm{Alb}^1(H))$:

(1) covers that factor rationally through $\mathrm{Alb}^1(H)$;

(2) covers that arise as Kummer quotients of étale covers of $\mathrm{Alb}^1(H)$;

(3) covers that pull back to ramified covers of $\mathrm{Alb}^1(H)$.

While type (1) covers are relatively easy to treat, for the other types we rely on two separate methods. Type (2) covers are dealt with through a fundamental group argument supplemented by a *coprimality trick*. As for type (3) covers, we make use of the recently established *weak Hilbert property for abelian varieties* due to Corvaja, Demeio, Javanpeykar, Lombardo and Zannier [Corvaja et al. 2022], which guarantees the existence of many rational points avoiding these covers. To control the pullbacks of rational curves on $\mathrm{Kum}(\mathrm{Alb}^1(H))$, we apply this result not to $\mathrm{Alb}^1(H)$ itself but instead to the base change of $\mathrm{Alb}^1(H)$ to the function field of H.

Finally, for any given finite collection of these covers, we need to carefully match these methods together, so as to extract a subfamily of rational curves which simultaneously have no rational sections along any one of them. Thus, a fortiori, these curves contribute to many rational points not coming from these covers, thereby achieving the Hilbert property for $\mathrm{Kum}(\mathrm{Alb}^1(H))$. The deduction of the Hilbert property for $\mathrm{Kum}(\mathrm{Jac}(H))$ in the case $g = 2$ is based on the fact (due to Cassels and Flynn; see Theorem 3.5) that $\mathrm{Kum}(\mathrm{Alb}^1(H))$ and $\mathrm{Kum}(\mathrm{Jac}(H))$ are birational to each other over K. However, in light of the lack of an analogue of Theorem 3.5 for $g > 2$, we cannot extend Theorem 1.4 to the setting where H has no K-rational Weierstrass point (or, slightly weaker, when H has no zero-cycle of degree 1 fixed under τ_H — the existence of such a zero-cycle is, for example, guaranteed if f factors into polynomials of coprime degrees). It would be interesting to remove the assumption on the degree of f in future investigations.

The article is structured as follows. After notation and background in Section 2, in Section 3 we construct and prove properties of rational curves on $\mathrm{Kum}(\mathrm{Jac}(H))$ and $\mathrm{Kum}(\mathrm{Alb}^1(H))$. In Section 4 the crucial pullback lemmas are formulated for covers of types (1)–(3). These are applied in Section 5 to prove the main theorems.

2. Notation, conventions and preliminaries

A *variety* over K is a separated reduced K-scheme of dimension > 0 and of finite type. To simplify exposition, throughout the rest of this article all varieties are assumed to be integral. We write $K(X)$ for the function field of a K-variety X.

Let $\phi : Y \to X$ be a morphism of varieties and let $X' \subseteq X$ be a subvariety. We say that X' *has a rational section along* ϕ if the morphism $\phi|_{\phi^{-1}(X')} : \phi^{-1}(X') \to X'$ has a rational section, i.e., there exists an open dense subset $U' \subset X'$ and $\psi : U' \to Y$ such that $\phi \circ \psi$ is the identity map on U'.

A K-variety birationally equivalent to $\mathbb{P}^n_K$ is called *K-rational*. By a *rational curve* on a K-variety X, we mean a closed (possibly singular) subvariety on X which is birationally equivalent to $\mathbb{P}^1_K$. By a *hyperelliptic curve* over K, we mean a smooth, projective curve over K with a finite morphism to $\mathbb{P}^1_K$ of degree 2. (Note that a curve may be geometrically hyperelliptic but not hyperelliptic over K, the image of the canonical map being a conic without a K-rational point. This does not happen when g is even.)

Assume that a K-variety V is equipped with an involution $\tau : V \to V$, usually taken to be implicit from the context. Then we define the *Kummer quotient* $\mathrm{Kum}(V)$ as the (possibly singular) quotient of V by the $\mathbb{Z}/2\mathbb{Z}$-action for which the group generator acts by τ.

Let $(G, +)$ be an abelian group. A *coset* $\mathscr{C}$ of G is of the form $g + H$, where $g \in G$ and $H \subset G$ is a subgroup.

Our next proposition says that proving the Hilbert property boils down to a specific family of maps.

Proposition 2.1. *Let X be a quasiprojective variety over a field K. Then a subset $M \subset X(K)$ is thin if and only if there exists a finite collection $(\phi_i : Y_i \to X)$, where Y_i are normal and geometrically integral varieties over K, and ϕ_i are finite surjective K-morphisms of degree $[K(Y_i) : K(X)] > 1$, such that $M \setminus \bigcup_{i=1}^r \phi_i(Y_i(K))$ is not Zariski dense.*

Proof. Upon considering the morphism from the variety Y, as the disjoint union of Y_i with a Zariski closed subset of X, to X, the "if" part follows just by definition.

We now prove the "only if" part. Namely, for every thin set $M \subset X(K)$, we want to construct such a finite collection $\phi_i : Y_i \to X$ as above such that $M \setminus \bigcup_{i=1}^r \phi_i(Y_i(K))$ is contained in a proper Zariski closed subset. Assume that $M \subset f(Z(K))$, where Z is a K-variety and the morphism f is generically finite and has no rational section. Upon considering the irreducible components of Z, it suffices to assume that $M = \phi(Y(K))$, where Y is an integral K-variety with $\dim Y = \dim X$, and $\phi : Y \to X$ is a dominant K-morphism with degree $[K(Y) : K(X)] > 1$. Let us fix a projective model $\overline{X}$ of X in what follows.

By the Nagata compactification theorem (see [Lütkebohmert 1993, Theorem 3.2]), ϕ extends to a proper morphism $\bar{\phi} : \overline{Y} \to \overline{X}$, where $\overline{Y}$ is a proper model of Y. Using the Stein factorisation [Hartshorne 1977, III, Corollary 11.5], we can decompose into $\bar{\phi} = g_2 \circ g_1 : \overline{Y} \to W \to \overline{X}$, where $g_1 : \overline{Y} \to W$ is proper with connected fibres and $g_2 : W \to \overline{X}$ is finite. It follows that g_1 is birational, and g_2 is proper by [ibid., Exercise II.4.1] and hence it is surjective (since ϕ is dominant). Let Z denote the nonnormal locus of W and let $h : \widehat{W} \to W$ be the normalisation, which is finite birational. We may assume that $\widehat{W}(K) \neq \varnothing$ and hence $\widehat{W}$ is geometrically integral (see [Serre 2008, p. 20]). Otherwise M is contained in the Zariski closed subset $g_2(Z)$. We let $\widetilde{W} = \widehat{W} \setminus (g_2 \circ h)^{-1}(\overline{X} \setminus X)$. We claim that the morphism $g_2 \circ h|_{\widetilde{W}} : \widetilde{W} \to X$ meets our need. Indeed, $g_2 \circ h$ is finite surjective of degree $[K(W) : K(X)] = [K(Y) : K(X)] > 1$, and $M \setminus \big((g_2 \circ h)(\widetilde{W}(K))\big)$ is contained in the Zariski closed subset $g_2(Z)$. □

The proposition above motivates the following convention.

A *cover* $\phi : Y \to X$ between K-varieties is a finite surjective morphism with Y being normal and geometrically integral. The *degree* of ϕ is $[K(Y) : K(X)]$.

Therefore, Proposition 2.1 may be rephrased as follows:

Corollary 2.2. *Let X be a quasiprojective variety over a Hilbertian field K. Then X has the Hilbert property over K if and only if, for any finite family of covers $\phi_i : Y_i \to X$ for $i = 1, \dots, r$ with $\deg(\phi_i) > 1$, $X(K) \setminus \bigcup_{i=1}^{r} \phi_i(Y_i(K))$ is Zariski dense in X.*

3. Rational curves on Jacobian Kummer varieties

Let H be a hyperelliptic curve of genus g over K. Recall the following classical facts about H, its Jacobian and Albanese varieties, and their Kummer quotients.

The hyperelliptic involution $\tau_H : H \to H$ is defined over K. The Kummer quotient $C = \mathrm{Kum}(H)$ is noncanonically isomorphic to $\mathbb{P}^1_K$, and we write $\pi_H : H \to C$ for the quotient morphism. The set $W \subset H(\overline{K})$ of Weierstrass points of H, i.e., the fixed points under τ_H, forms a zero-dimensional reduced K-subscheme of degree $2g + 2$.

3.1. ***The Jacobian variety and rational curves on its Kummer quotient.*** The Jacobian $J = \mathrm{Jac}(H)$ is a principally polarised abelian variety over K, even when $H(K) = \varnothing$. See, e.g., [Milne 1986b, Theorem 1.1, Summary 6.11]. For $n \in \mathbb{Z}$, write $[n_J] : J \to J$ for the multiplication-by-n morphism, and for a closed point $x \in J$, write $t_x : J \to J$ for the translation-by-x morphism. The fixed points of the involution $\tau_J = [(-1)_J]$ consist precisely of the 2-torsion $J[2]$, which is a zero-dimensional subgroup scheme of degree 2^{2g}. Let $X = \mathrm{Kum}(J)$ and $\pi : J \to X$ be the quotient morphism. Clearly the morphisms $[n_J]$ and τ_J commute, so $[n_J]$ descends to X and we denote the resulting morphism by $[n_X]$. The morphism π is étale outside the images of $J[2]$, which form the singular locus of X. Blowing up these 2^{2g} canonical singularities gives a minimal resolution $\widetilde{X} \to X$, where $\widetilde{X}$ is known as the *Kummer variety* of J. It is a smooth, projective and simply connected variety (see [Spanier 1956, Theorem 1]).

For every $x_0 \in H(K)$, there exists an Abel–Jacobi closed K-embedding

$$\iota_{x_0} : H \hookrightarrow J, \quad x \mapsto \langle x_0 - x \rangle,$$

where $\langle \cdot \rangle$ denotes the class of a divisor. By the Riemann–Roch theorem, the closed points on H generate J (see [Hartshorne 1977, IV, Remark 4.10.9] or [Milne 1986b, §5]).

We assume throughout the rest of Section 3.1 that H has a K-rational Weierstrass point ∞. We consider H as a subvariety of J via the Abel–Jacobi embedding ι_∞. Since the restriction of the map τ_J to H is just τ_H, the image of H under π is a smooth rational curve, which we denote by C. For every $n \in \mathbb{Z}_{>0}$, let $H_n = [n_J](H)$, the image of H under the map $[n_J]$, and let $C_n = \pi(H_n)$. The curves C_n are rational for all $n \in \mathbb{Z}_{>0}$. (See, e.g., [Huybrechts 2016, Chapter 13, Example 0.4].)

Lemma 3.1. *The family $(H_n)_{n\in\mathbb{N}}$ consists of pairwise distinct curves. Consequently, $(C_n)_{n\in\mathbb{N}}$ are pairwise distinct rational curves on X.*

Proof. Assume $H_m = H_n$ for some $0 < m < n$. The m^{2g} translates $t_x(H)$ for $x \in J[m]$ are pairwise distinct. Since $[m_J]$ is étale of degree m^{2g}, it follows that $[m_J]|_H : H \to H_m$ is of degree 1, i.e., a birational map.

The same holds for $[n_J]|_H$. Thus, there is an automorphism σ of H such that, as rational maps,

$$\sigma = [m_J]|_H \circ ([n_J]|_H)^{-1} = ([n_J]|_H)^{-1} \circ [m_J]_H.$$

Because σ has finite order, after replacing m and n by powers thereof, we can assume $[m_J]_H = [n_J]_H$. It follows that $H \subset \mathrm{Ker}([(n-m)_J]) \subset J$, which is impossible since the closed points of H generate J. □

Remark 3.2. There is an oversight in [Huybrechts 2016, p. 274], where it is stated that, when $g = 2$, the rational curve C_n is smooth for all n. In fact this is false for $n > 1$ and it is for this reason that we will need the coprimality argument in Proposition 4.4. If C_n were smooth, then, since it is an ample geometrically irreducible divisor of J, its pullback along $[n_J]$ would also be geometrically irreducible by [Birkenhake and Lange 2004, Corollary 4.3.4]. But this is clearly not the case since it is (geometrically) a sum of translates of C. See also Remark 4.3 below.

More concretely, let us take a look at the curve C_2. Since H contains six points of $J[2]$, $H_2 = [2_J](H)$ passes through $0 \in J$ with multiplicity at least 6. Hence, C_2 has multiplicity at least 3 at $\pi(0)$. In general, the singular locus of H_n (resp. C_n) is the closed subscheme of all points $[n_J]x$ (resp. $[n_X]\pi(x)$) for $x \in H$ such that there exists $y \in H$ different from x with $[n_J](x-y) = 0$. In other words, x is in a *torsion package* of size > 1 in the sense of [Baker and Poonen 2001]. We omit the proof since this fact will not be needed.

Lemma 3.3. *Let $\eta_H \in J_{K(H)}$ be the image of the generic point of H under the embedding $\iota_\infty : H \hookrightarrow J$. Then η_H is nondegenerate in $J_{K(H)}$, i.e., $\mathbb{Z}\eta_H$ is Zariski dense in $J_{K(H)}$.*

Proof. Assume instead that η_H is contained in a proper abelian subvariety of $J_{K(H)}$. Then, by the Poincaré irreducibility theorem, there exists a nonzero endomorphism $e \in \mathrm{End}(J_{\overline{K(H)}})$ (actually in $\mathrm{End}(J_{K(H)})$) such that $e\eta_H = 0$. Because $\mathrm{End}(J_{\overline{K(H)}}) \simeq \mathrm{End}(J_{\overline{K}})$ — see [Chai et al. 2014, Lemma 1.2.1.2] — it follows that there exists $e' \in \mathrm{End}(J_{\overline{K}})$ such that $e's(x) = e'x = 0$ for all $x \in H$. Thus, H would have to be contained in a proper abelian subvariety of J, which is impossible again since the closed points of H generate J. □

Corollary 3.4. *For every finite-index coset $\mathcal{N} \subset \mathbb{Z}$, $\bigcup_{n\in\mathcal{N}} H_n$ is Zariski dense in J and $\bigcup_{n\in\mathcal{N}} C_n$ is Zariski dense in X.*

Proof. Write $\mathcal{N} = a\mathbb{Z}+b$, where a, b are integers with $a \neq 0$. Since the morphism $t_{b\eta_H} \circ [a_{J_{K(H)}}] : J_{K(H)} \to J_{K(H)}$ is dominant, the conclusion follows directly from Lemma 3.3. □

3.2. ***The Albanese torsor $\mathbf{Alb}^1(H)$ and rational curves on its Kummer quotient.*** The zero-cycles of degree 1 (up to linear equivalence) on H fixed under τ_H form a torsor T under $J[2]$, which is trivial if H has a K-rational Weierstrass point. Let $c \in \mathrm{H}^1(K, J[2])$ be the associated cohomology class. Then, for every fixed $\overline{K}$-Weierstrass point $\infty \in W$, c is given by the cocycle $\gamma \mapsto \langle \gamma(\infty) - \infty \rangle$. The twist of J by the image of c in $\mathrm{H}^1(K, J)$, i.e., the quotient of $J \times T$ by the diagonal action of $J[2]$, is isomorphic to the Albanese torsor $J^1 = \mathrm{Alb}^1(H)$.

In general, we denote by $J^i = \mathrm{Alb}^i(H)$ the i-th Albanese torsor, which parametrises zero-cycles of degree i on H. There is an isomorphism $J^i \simeq J^{i+2}(H)$ for all $i \geq 0$, given by adding $\langle 2\infty \rangle$, and $J^0 \simeq J$. The variety J^1 is an initial object within the category of torsors V under J with a morphism

$\iota_V : H \to V$ (see [Wittenberg 2008, Appendix A]). The involution τ_H extends formally to an involution τ_{J^1} on J^1 in a way that is compatible with the involution τ_J on J and the J-torsor structure. We write $\pi^1 : J^1 \to X^1 = \mathrm{Kum}(J^1)$ for the corresponding Kummer quotient morphism.

In the case $g = 2$, the surfaces X and X^1 are projective duals of each other as quartic surfaces in $\mathbb{P}^3$. (By [Cassels and Flynn 1996, Chapter 4], they are isomorphic if and only if at least one of the translates of $C \subset X$ by $J[2](\overline{K})$ — the so-called *tropes* — is defined over K.) However, Cassels and Flynn [1996, Chapter 16] point out (somehow surprisingly) that:

Theorem 3.5 (Cassels and Flynn). *The minimal desingularisations of the surfaces X and X^1 are isomorphic over K.*

Proof. We offer a more conceptional and simplified proof here. Indeed, consider the involutions

$$\sigma_1 : H \times H \to H \times H, \quad (x_1, x_2) \mapsto (\tau_H(x_2), \tau_H(x_1)),$$
$$\sigma_2 : H \times H \to H \times H, \quad (x_1, x_2) \mapsto (x_2, x_1).$$

We define $\mathrm{Inv}^2(H)$ (resp. $\mathrm{Sym}^2(H)$) to be the quotient of $H \times H$ by σ_1 (resp. by σ_2). For any point $(x_1, x_2) \in H \times H$, we write $\{x_1, x_2\}$ for the corresponding quotient image.

We now define

$$\beta : \mathrm{Inv}^2(H) \to J^0 = J, \quad \{x_1, x_2\} \mapsto \langle x_1 - x_2 \rangle,$$
$$\beta' : \mathrm{Sym}^2(H) \to J^2, \quad \{x_1, x_2\} \mapsto \langle x_1 + x_2 \rangle.$$

Then, by the Riemann–Roch theorem (see also [Baker and Poonen 2001, Proposition 4(2)]), outside the image of the diagonal $(\mathrm{id}_H \times \mathrm{id}_H)(H)$ (resp. the antidiagonal $(\mathrm{id}_H \times \tau_H)(H)$), β (resp. β') is an isomorphism. Hence, both β, β' are birational.

Let us define the rational maps b, b' over K by

$$b : X \dashrightarrow X^1, \quad \{z, \tau_J(z)\} \mapsto \{z + \langle x_2 \rangle, z + \langle \tau_H(x_1) \rangle\},$$

for $z \in J \setminus J[2]$, where $\{x_1, x_2\}$ is the preimage of $2z$ under β, and

$$b' : X^1 \dashrightarrow X, \quad \{z^1, \tau_{J^1}(z^1)\} \mapsto \{z^1 - \langle x \rangle, z^1 - \langle x' \rangle\},$$

for $z^1 \in J^1 \setminus T$, where $\{x, x'\}$ is the preimage of $2z^1$ under β'. Then b and b' are inverse to each other, and hence are birational maps. Blowing up the singular loci of X and X' gives the desired isomorphism. □

Remark 3.6. Alternatively, in terms of cocycles, one can check that the image of c in $\mathrm{H}^1(K, \mathrm{Aut}_{\overline{K}}\, \overline{K}(X))$ is the coboundary of the automorphism of $\overline{K}(X)$ which is induced by the birational map

$$\sigma : X_{\overline{K}} \dashrightarrow X_{\overline{K}}, \quad \pi(z) \mapsto \pi(z + \langle x_2 - \infty \rangle);$$

thus, c vanishes in $\mathrm{H}^1(K, \mathrm{Aut}_{\overline{K}}\, \overline{K}(X))$.

The image H^1 under the natural embedding $\iota_{J^1} : H \to J^1$ is stable under the involution τ_{J^1} and thus yields a smooth rational subcurve $C^1 = \pi^1(H^1) \subset X^1$. For any *odd* integer $n \in \mathbb{Z}_{>0}$, we denote by $[n_{J^1}]$ the composition of the formal multiplication-by-n map (of zero-cycles of degree 1) $J^1 \to J^n$ and the

previously defined isomorphism $J^n \simeq J^1$. Because $[n_{J^1}]$ commutes with involution, it descends to a morphism $[n_{X^1}]$. We let $H_n^1 = [n_{J^1}](H^1)$, the image of H^1 under the map $[n_{J^1}]$, and let $C_n^1 = \pi^1(H_n^1)$. Since

$$C_n^1 = (\pi^1 \circ [n_{J^1}])(H^1) = ([n_{X^1}] \circ \pi^1)(H^1) = [n_{X^1}](C^1),$$

all the curves C_n^1 are rational.

If H has a K-rational Weierstrass point ∞, then the K-morphism

$$\alpha_J : J^1 \to J, \quad x \mapsto x - \langle \infty \rangle,$$

commutes with the involutions τ_J, τ_{J^1}, so it descends to a K-morphism $\alpha_X : X^1 \to X$ and yields a commutative diagram

$$\begin{array}{ccc} J^1 & \xrightarrow{\pi^1} & X^1 \\ \downarrow{\scriptstyle \alpha_J} & & \downarrow{\scriptstyle \alpha_X} \\ J & \xrightarrow{\pi} & X \end{array} \tag{$\dagger$}$$

where the vertical arrows are K-isomorphisms. It is easy to see that H_n^1 maps to H_n (as a subvariety of J) under α_J, and C_n^1 maps to C_n under α_X for all odd $n \in \mathbb{Z}_{>0}$.

4. Pullbacks along covers

Throughout this section, K denotes a finitely generated field of characteristic zero and $X = \mathrm{Kum}(J)$ denotes the Kummer quotient associated to the Jacobian J of a genus g hyperelliptic curve H over K with a K-rational Weierstrass point ∞. We continue to consider H embedded into J via ι_∞.

4.1. *Pullback along covers factoring through the Jacobian.*

Lemma 4.1. *Let $\phi : Y \to X$ be a cover of degree > 1. Suppose that the base change variety $Y \times_X J$ is not geometrically integral. Then, for every finite-index coset $\mathcal{N} \subset \mathbb{Z}$, there exists a subcoset $\mathcal{N}' \subset \mathcal{N}$ of finite index such that none of the rational curves $(C_n)_{n \in \mathcal{N}'}$ have a rational section along ϕ over $\overline{K}$.*

Proof. We may base change everything to $\overline{K}$ and we fix an algebraic closure Ω of $\overline{K}(X)$. We have that $Y_{\overline{K}} \times_{X_{\overline{K}}} J_{\overline{K}}$ is not geometrically integral if and only if $\overline{K}(Y) \otimes_{\overline{K}(X)} \overline{K}(J)$ is not a field, or, equivalently, the field extensions $\overline{K}(Y)$ and $\overline{K}(J)$ in Ω are not linearly disjoint over $\overline{K}(X)$. Since $[\overline{K}(Y) \cdot \overline{K}(J) : \overline{K}(Y)] \leq [\overline{K}(J) : \overline{K}(X)] = 2$, this happens if and only if $[\overline{K}(Y) \cdot \overline{K}(J) : \overline{K}(Y)] < 2$, or equivalently $\overline{K}(J) \subseteq \overline{K}(Y)$. It follows that, over $\overline{K}$, $\phi : Y \to X$ factors as a rational map $\phi_1 : Y \dashrightarrow J$ then to X.

By Lemma 3.3, we fix a closed point $P \in H$ which is nondegenerate in J. By [Corvaja et al. 2022, Lemma 4.6], there exists a finite-index subcoset $\mathcal{N}' \subset \mathcal{N}$ such that the set $\{nP : n \in \mathcal{N}'\}$ does not intersect with the indeterminacy locus, say $Z \subset J$, of ϕ_1. This implies that none of the curves $(H_n)_{n \in \mathcal{N}'}$ is contained in Z. If a rational curve $(C_n)_{\overline{K}} \subset X_{\overline{K}}$ for $n \in \mathcal{N}'$ has a rational section along ϕ, say $\tilde{\phi}$, then it has a rational section $\phi_1 \circ \tilde{\phi}$ along π. This is impossible since there are no nonconstant rational maps from the projective line to abelian varieties (see [Milne 1986a, Corollary 3.8]). □

4.2. *Pullbacks along unramified covers.*

Lemma 4.2. *Let D be a smooth, projective curve of genus ≥ 1 over a separably closed field. Fix a point x in D inducing the Abel–Jacobi embedding $\iota_x : D \hookrightarrow \mathrm{Jac}(D)$. Let $\pi_1(D, x)$ (resp. $\pi_1(\mathrm{Jac}(D), 0)$) be the pointed étale fundamental group of D (resp. $\mathrm{Jac}(D)$). Then the map*

$$(\iota_x)_* : \pi_1(D, x) \to \pi_1(\mathrm{Jac}(D), 0)$$

induced by ι_x is surjective. Consequently, every pullback via ι_x of a connected finite étale covering of $\mathrm{Jac}(D)$ remains connected as a finite étale covering over D.

Proof. This is explained in [Milne 1986b, Proposition 9.1]. Indeed, $(\iota_x)_*$ induces an isomorphism

$$\pi_1(D, x)^{\mathrm{ab}} \to \pi_1(\mathrm{Jac}(D), 0),$$

where $\pi_1(D, x)^{\mathrm{ab}}$ is the maximal abelian quotient of $\pi_1(D, x)$, so, in particular, $(\iota_x)_*$ is surjective. Thus, by the theory of Galois categories, all connected finite étale coverings of $\mathrm{Jac}(D)$ have abelian Galois group, and the surjectivity of $(\iota_x)_*$ implies that all connected finite abelian coverings of D are obtained precisely by pulling back the connected finite étale coverings of J. See also [Szamuely 2009, Proposition 5.5.4(2)]. □

Remark 4.3. If $g = 2$, one may also prove Lemma 4.2 by using the fact that the pullback of a geometrically irreducible, unibranch, ample divisor on an abelian variety along an isogeny is again geometrically irreducible. (See [Birkenhake and Lange 2004, Corollary 4.3.4]. Note that the statement is missing the crucial unibranch — or at the least smoothness — assumption.)

Lemma 4.2 applies to $H \hookrightarrow J$, while an analogous statement for H_n with $n > 1$ fails because of the nonsmoothness. However, one can remedy the situation by restricting to a subfamily:

Proposition 4.4 ("coprimality trick"). *Let $\phi : J' \to J$ be an unramified cover. Then $\phi^{-1}(H_n)$ are geometrically irreducible for all n coprime to* $\deg(\phi)$.

Proof. We may base change everything to $\overline{K}$, without affecting the geometrical irreducibility. By the Lang–Serre theorem [Serre 1988, Chapter I, §2], J' is an abelian variety and ϕ can be made into an isogeny. The case where $\deg(\phi) = 1$ is clear and so we assume from now on that $\deg(\phi) > 1$.

If $n = 1$, then the claim follows from Lemma 4.2. In particular, $\mathrm{Ker}(\phi) \subset \phi^{-1}(H_1)$. In general, if n is coprime to $\deg(\phi) = \deg(\mathrm{Ker}(\phi))$, then $[n_{J'}]|_{\mathrm{Ker}\,\phi}$ is an isomorphism. Thus,

$$\phi^{-1}(H_n) = \phi^{-1}([n_J](H_1)) = [n_{J'}](\phi^{-1}(H_1)) + \mathrm{Ker}(\phi) = [n_{J'}](\phi^{-1}(H_1) + \mathrm{Ker}(\phi)) = [n_{J'}](\phi^{-1}(H_1)),$$

and hence $\phi^{-1}(H_n)$ is geometrically irreducible. □

Corollary 4.5. *Let $\phi_i : Y_i \to X$ for $i \in I$ be any finite collection of covers of X with $\deg(\phi_i) > 1$ such that $Y_i \times_X J$ are all geometrically integral, and the base change morphisms $\tilde{\phi}_i : Y_i \times_X J \to J$ are all unramified. Then the $\phi_i^{-1}(C_n)$ are all geometrically irreducible for all n coprime to $\prod_i \deg(\phi_i)$.*

Proof. Since H_n is stable under the involution, $\phi_i^{-1}(C_n)$ is geometrically irreducible if and only if $\tilde{\phi}_i^{-1}(H_n)$ is. Since $Y_i \times_X J$ are all geometrically integral, we have $\deg(\tilde{\phi}_i) = \deg(\phi_i)$. We then apply Proposition 4.4 to $\tilde{\phi}_i : Y_i \times_X J \to J$. □

4.3. ***Pullbacks along ramified covers.*** The following result is essentially Theorem 1.3 of [Corvaja et al. 2022]:

Theorem 4.6 (Corvaja, Demeio, Javanpeykar, Lombardo and Zannier). *Let A be an abelian variety over a finitely generated field K of characteristic 0 with a nondegenerate point $x \in A(K)$. Let $f_i : B_i \to A$ be a finite collection of ramified covers. Then, for any finite-index coset $\mathcal{P} \subset \mathbb{Z}x$, there exists a finite-index subcoset of $\mathcal{P}' \subset \mathcal{P}$ such that $(f_i^{-1}(x'))(K) = \varnothing$ for every $x' \in \mathcal{P}'$.*

Proof. It suffices to work with a single ramified cover $f : B \to A$ by successively analysing the finite-index coset obtained at each step. Write $\mathcal{P} = \Pi + bx$, where $b \in \mathbb{Z}$ and Π is a finite-index subgroup of $\mathbb{Z}x$ and $P \in A(K)$. Define $f' = t_{-bx} \circ f$, where $t_{-bx} : A \to A$ is the translation-by-$(-bx)$ morphism. Then we apply [Corvaja et al. 2022, Theorem 1.3] to Π and get a finite-index coset $\Pi' \subset \Pi$ for which any K-fibre of f' has no K-point. Then we may take $\mathcal{P}' = \Pi' + bx$ to obtain the same conclusion for f. □

Corollary 4.7. *Let $\phi_i : Y_i \to X$ for $i \in I$ be any finite collection of covers such that all the base change morphisms $\tilde{\phi}_i : Y_i \times_X J \to J$ are ramified with $Y_i \times_X J$ geometrically integral. Then, for any finite-index coset $\mathcal{N} \subset \mathbb{Z}$, there exists a finite-index subcoset $\mathcal{N}' \subset \mathcal{N}$ such that, for any $n \in \mathcal{N}'$, the rational curve C_n does not have a rational section along ϕ_i for all $i \in I$.*

Proof. Consider the base change

$$\tilde{\phi}_{i,K(H)} : (Y_i \times_X J)_{K(H)} \to J_{K(H)}$$

to the function field $K(H)$ of H. This is still a ramified morphism with $(Y_i \times_X J)_{K(H)}$ being geometrically integral.

By Lemma 3.3, the $K(H)$-rational point $\eta_H \in J_{K(H)}$ induced by $\iota_\infty : H \hookrightarrow J$ is nondegenerate. By Theorem 4.6 applied to $\mathcal{N}\eta_H$ over the finitely generated field $K(H)$, there exists a finite-index coset $\mathcal{N}' \subset \mathcal{N}$ such that, for any $\eta'_H \in \mathcal{N}'\eta_H \subset J_{K(H)}$,

$$\eta'_H \notin \bigcup_{i \in I} \tilde{\phi}_{i,K(H)}\big((Y_i \times_X J)_{K(H)}(K(H))\big).$$

Consider the cartesian diagram

$$\begin{array}{ccccc} (Y_i \times_X J)_{K(H)} & \longrightarrow & Y_{i,K(H)} & \longrightarrow & Y_{i,K(C)} \\ \downarrow{\scriptstyle \tilde{\phi}_{i,K(H)}} & & \downarrow{\scriptstyle \phi_{i,K(H)}} & & \downarrow{\scriptstyle \phi_{i,K(C)}} \\ J_{K(H)} & \longrightarrow & X_{K(H)} & \longrightarrow & X_{K(C)} \end{array}$$

where the rightmost square is the base change from $K(C)$ to $K(H)$. Let $\eta_C \in X_{K(C)}$ be the image of η_H under $J_{K(H)} \to X_{K(C)}$. Then, for any $\eta'_C \in \mathcal{N}'\eta_C \subset X_{K(C)}$,

$$\eta'_C \notin \bigcup_{i \in I} \phi_{i,K(C)}\big(Y_{i,K(C)}(K(C))\big).$$

Equivalently, none of the curves C_n for $n \in \mathcal{N}'$ has a rational section along ϕ_i for all $i \in I$. □

5. Hilbert property for Jacobian Kummer varieties

Proposition 5.1. *Let X be a quasiprojective variety over a Hilbertian field K. Assume that there exists an infinite set $\mathscr{C}$ of (not necessarily smooth) K-rational subvarieties on X satisfying the following condition*:

For every finite family of covers $\phi_i : Y_i \to X$ for $i = 1, \dots, r$ with $\deg(\phi_i) > 1$, there exists a subset $\mathscr{C}' \subset \mathscr{C}$ such that $\bigcup_{C \in \mathscr{C}'} C$ is Zariski dense in X and, for all $C \in \mathscr{C}'$ and $i = 1, \dots, r$, C does not have a rational section along ϕ_i.

Then X has the Hilbert property.

Proof. Let us fix a finite family of covers $\phi_i : Y_i \to X$ for $i = 1, \dots, r$ with $\deg(\phi_i) > 1$. Let $\mathscr{C}' \subset \mathscr{C}$ be as stated in the proposition. Because the union of the $C \in \mathscr{C}'$ is Zariski dense, upon replacing $\mathscr{C}'$ by a subset if necessary, we may assume that, for every subvariety $C \in \mathscr{C}'$, C is not contained in the ramification locus of any ϕ_i for $i = 1, \dots, r$. In view of Proposition 2.1, to show the Hilbert property for X, it suffices to show that

$$\left(\bigcup_{C \in \mathscr{C}'} C(K) \right) \setminus \left(\bigcup_{i=1}^{r} \phi_i(Y_i(K)) \right)$$

is not contained in any proper Zariski closed subset of X.

Assume that there exists $Z \subset X$ proper Zariski closed containing all the K-points above. We fix $C_0 \in \mathscr{C}'$ such that $C_0 \not\subset Z$. For every $i \in \{1, \dots, r\}$, let C_{ij} for $j = 1, \dots, k_i$ be the irreducible components of $\phi_i^{-1}(C_0)$. The maps $\phi_i|_{C_{ij}} : C_{ij} \to C_0$ are finite morphisms because C_0 does not have a rational section along ϕ_i, and, since C_0 is not contained in the ramification locus of ϕ_i by assumption, if $\dim(C_{ij}) = \dim(C_0)$, then $\phi_i|_{C_{ij}}$ is dominant, and hence has degree > 1. It thus follows from the Hilbert property for rational varieties (see [Serre 2008, p. 30, Examples]) that $C_0(K)$ is not thin, whence

$$C_0(K) \setminus \left((Z \cap C_0)(K) \cup \left(\bigcup_{i=1}^{r} \bigcup_{j=1}^{k_i} \phi_i(C_{ij}(K)) \right) \right) \neq \varnothing,$$

which is a contradiction. □

We now turn to the proofs of the main theorems. We will apply Proposition 5.1 only in the case where $\mathscr{C}$ is a countable set of subcurves.

Proof of Theorem 1.5. Consider the rational K-subcurves $\{C_n^1\}_{n \in 1+2\mathbb{Z}}$ on X^1. Let $\phi_i : Y_i \to X^1$ for $i = 1, \dots, r$ be any fixed finite family of covers with $\deg(\phi_i) > 1$. We want to construct a finite-index subcoset $\mathcal{N}' \subset 1 + 2\mathbb{Z}$ of odd integers such that, for $\{C_n^1\}_{n \in \mathcal{N}'}$, all conditions of Proposition 5.1 are satisfied. Let us take a finite field extension L of K such that H has an L-rational Weierstrass point ∞. We base change everything to L (and we omit the subscript L). In view of the diagram (†), we shall replace X^1 by the isomorphic X and C_n^1 by C_n in the following.

Let $\tilde{\phi}_i : \widetilde{Y}_i \to J$ be the base change of ϕ_i along the Kummer quotient $\pi : J \to X$. It may be assumed that, for certain r_1, all $\widetilde{Y}_i$ for $1 \leq i < r_1$ are not geometrically integral and the remaining $\widetilde{Y}_i$ for $r_1 \leq i \leq r$ are geometrically integral. For $r_1 \leq i \leq r$, we may moreover assume that, for certain $r_2 \geq r_1$, the morphisms $\tilde{\phi}_i$ for $r_1 \leq i \leq r_2$ are all unramified, while the remaining $\tilde{\phi}_i$ for $r_2 < i \leq r$ are all ramified.

We start with

$$\mathscr{C} = \{C_n : n \in 1 + 2\mathbb{Z}\}.$$

If $r_2 > r_1$, i.e., certain $\tilde{\phi}_i$ are unramified, we consider

$$d = \operatorname{lcm}(\deg \phi_i : i = r_1, \dots, r_2), \quad \mathcal{N}^{\text{ét}} = 1 + 2d\mathbb{Z}.$$

Then $\mathcal{N}^{\text{ét}}$ is a finite-index coset of $\mathbb{Z}$, and, by Corollary 4.5, any C_n with $n \in \mathcal{N}^{\text{ét}}$ does not have a rational section along ϕ_i for all $r_1 \le i \le r_2$. Otherwise, if there are no such unramified covers, we just take $\mathcal{N}^{\text{ét}} = 1 + 2\mathbb{Z}$.

By Corollary 4.7, one obtains a finite-index subcoset $\mathcal{N}^{\text{geo}} \subset \mathcal{N}^{\text{ét}}$ such that any C_n with $n \in \mathcal{N}^{\text{geo}}$ does not have a rational section along ϕ_i for all $i = r_2 + 1, \dots, r$. Then, on applying Lemma 4.1, one obtains a further finite-index subcoset $\mathcal{N}' \subset \mathcal{N}^{\text{geo}}$ such that any C_n with $n \in \mathcal{N}'$ does not have a rational section along ϕ_i for all $i = 1, \dots, r_1 - 1$ (still over the field extension L of K).

Any such C_n^1, $n \in \mathcal{N}'$ cannot have a rational section along any ϕ_i over the base field K as well. Thus, by using again Corollary 3.4, all conditions of Proposition 5.1 are satisfied for the family of rational curves

$$\mathscr{C}' = \{C_n : n \in \mathcal{N}'\},$$

from which we conclude the Hilbert property. □

Proof of Theorem 1.3. This follows from Theorem 1.5 together with Theorem 3.5 using the birational invariance of the Hilbert property. □

Proof of Theorem 1.4. This follows immediately from Theorem 1.5 since H has, by assumption, a K-rational Weierstrass point (at infinity), and thus $X \cong X^1$ over K in view of the diagram (†). □

Acknowledgements

We are in debt to J. Demeio for generously sharing many key insights into the Hilbert property which simplify and generalise our previous arguments considerably. Gvirtz-Chen was supported by the Engineering and Physical Sciences Research Council (grant EP/R513143/1) and UKRI award UKRI094. Huang was supported by the DFG grant DE1646/4-2 at Leibniz Universität Hannover, and by the FWF grant P32428-N35 at IST Austria. This work was partly undertaken at the workshop *Arithmetic statistics and local–global principles* hosted by the Erwin Schrödinger International Institute for Mathematics and Physics, Vienna and the workshop *Rational points on higher-dimensional varieties* hosted by the International Centre for Mathematical Sciences, Edinburgh. We thank the organisers and the institutes for the hospitality. Last but not least, we thank the referee for helpful remarks.

References

[Baker and Poonen 2001] M. Baker and B. Poonen, "Torsion packets on curves", *Compositio Math.* **127**:1 (2001), 109–116. MR

[Birkenhake and Lange 2004] C. Birkenhake and H. Lange, *Complex abelian varieties*, 2nd ed., Grundl. Math. Wissen. **302**, Springer, 2004. MR

[Cassels and Flynn 1996] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus 2*, London Mathematical Society Lecture Note Series **230**, Cambridge Univ. Press, 1996. MR

[Chai et al. 2014] C.-L. Chai, B. Conrad, and F. Oort, *Complex multiplication and lifting problems*, Mathematical Surveys and Monographs **195**, Amer. Math. Soc., Providence, RI, 2014. MR

[Colliot-Thélène 2020] J.-L. Colliot-Thélène, "Point générique et saut du rang du groupe de Mordell–Weil", *Acta Arith.* **196**:1 (2020), 93–108. MR

[Colliot-Thélène and Sansuc 1987] J.-L. Colliot-Thélène and J.-J. Sansuc, "Principal homogeneous spaces under flasque tori: applications", *J. Algebra* **106**:1 (1987), 148–205. MR

[Corvaja and Zannier 2017] P. Corvaja and U. Zannier, "On the Hilbert property and the fundamental group of algebraic varieties", *Math. Z.* **286**:1-2 (2017), 579–602. MR

[Corvaja et al. 2022] P. Corvaja, J. L. Demeio, A. Javanpeykar, D. Lombardo, and U. Zannier, "On the distribution of rational points on ramified covers of abelian varieties", *Compos. Math.* **158**:11 (2022), 2109–2155. MR

[Demeio 2021] J. L. Demeio, "Elliptic fibrations and the Hilbert property", *Int. Math. Res. Not.* **2021**:13 (2021), 10260–10277. MR

[Fried and Jarden 2008] M. D. Fried and M. Jarden, *Field arithmetic*, 3rd ed., Ergebnisse der Math. (3) **11**, Springer, 2008. MR

[González et al. 2005] J. González, J. Guàrdia, and V. Rotger, "Abelian surfaces of GL_2-type as Jacobians of curves", *Acta Arith.* **116**:3 (2005), 263–287. MR

[Gvirtz-Chen and Mezzedimi 2023] D. Gvirtz-Chen and G. Mezzedimi, "A Hilbert irreducibility theorem for Enriques surfaces", *Trans. Amer. Math. Soc.* **376**:6 (2023), 3867–3890. MR

[Harpaz and Skorobogatov 2016] Y. Harpaz and A. N. Skorobogatov, "Hasse principle for Kummer varieties", *Algebra Number Theory* **10**:4 (2016), 813–841. MR

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **No. 52**, Springer, 1977. MR

[Hilbert 1892] D. Hilbert, "Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten", *J. Reine Angew. Math.* **110** (1892), 104–129. MR

[Holmes and Pannekoek 2015] D. Holmes and R. Pannekoek, "The Brauer–Manin obstruction on Kummer varieties and ranks of twists of abelian varieties", *Bull. Lond. Math. Soc.* **47**:4 (2015), 565–574. MR

[Huang 2021] Z. Huang, "Rational points on elliptic K3 surfaces of quadratic twist type", *Q. J. Math.* **72**:3 (2021), 755–772. MR

[Huybrechts 2016] D. Huybrechts, *Lectures on K3 surfaces*, Cambridge Studies in Advanced Mathematics **158**, Cambridge Univ. Press, 2016. MR

[Kumar 2014] A. Kumar, "Elliptic fibrations on a generic Jacobian Kummer surface", *J. Algebraic Geom.* **23**:4 (2014), 599–667. MR

[Kuwata and Wang 1993] M. Kuwata and L. Wang, "Topology of rational points on isotrivial elliptic surfaces", *Int. Math. Res. Not.* **1993**:4 (1993), 113–123. MR

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, 1983. MR

[Lütkebohmert 1993] W. Lütkebohmert, "On compactification of schemes", *Manuscripta Math.* **80**:1 (1993), 95–111. MR

[Milne 1986a] J. S. Milne, "Abelian varieties", pp. 103–150 in *Arithmetic geometry* (Storrs, CO, 1984), edited by G. Cornell and J. H. Silverman, Springer, 1986. MR

[Milne 1986b] J. S. Milne, "Jacobian varieties", pp. 167–212 in *Arithmetic geometry* (Storrs, CO, 1984), edited by G. Cornell and J. H. Silverman, Springer, 1986. MR

[Serre 1959] J.-P. Serre, "On the fundamental group of a unirational variety", *J. London Math. Soc.* **34** (1959), 481–484. MR

[Serre 1988] J.-P. Serre, *Algebraic groups and class fields*, Graduate Texts in Mathematics **117**, Springer, 1988. MR

[Serre 2008] J.-P. Serre, *Topics in Galois theory*, 2nd ed., Research Notes in Mathematics **1**, A K Peters, Wellesley, MA, 2008. MR

[Spanier 1956] E. Spanier, "The homology of Kummer manifolds", *Proc. Amer. Math. Soc.* **7** (1956), 155–160. MR

[Szamuely 2009] T. Szamuely, *Galois groups and fundamental groups*, Cambridge Studies in Advanced Mathematics **117**, Cambridge Univ. Press, 2009. MR

[Wittenberg 2008] O. Wittenberg, "On Albanese torsors and the elementary obstruction", *Math. Ann.* **340**:4 (2008), 805–838. MR

Communicated by Antoine Chambert-Loir
Received 2023-05-14 Revised 2025-02-07 Accepted 2025-08-05

damian.gvirtz@glasgow.ac.uk *School of Mathematics and Statistics, University of Glasgow, Glasgow, United Kingdom*

zhizhong.huang@yahoo.com *Institute of Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1559

Homological regularities and concavities

Ellen Kirkman, Robert Won and James J. Zhang

This paper concerns homological notions of regularity for noncommutative algebras. Properties of an algebra A are reflected in the regularities of certain (complexes of) A-modules. We study the classical Tor regularity and Castelnuovo–Mumford regularity, which were generalized from the commutative setting to the noncommutative setting by Jørgensen, Dong and Wu. We also introduce two new numerical homological invariants: concavity and Artin–Schelter regularity. Artin–Schelter regular algebras occupy a central position in noncommutative algebra and noncommutative algebraic geometry, and we use these invariants to establish criteria which can be used to determine whether a noetherian connected graded algebra is Artin–Schelter regular.

0. Introduction

Let $\Bbbk$ be a base field. An $\mathbb{N}$-graded algebra $A := \bigoplus_{i=0}^{\infty} A_i$ is called *connected graded* if $A_0 = \Bbbk$. Throughout let A denote a connected $\mathbb{N}$-graded $\Bbbk$-algebra unless otherwise stated. The Hilbert series $h_A(t)$ of A will be recalled in Definition 1.1.

A seminal result due to Stanley [1978, Theorem 4.4] states:

Theorem 0.1. *Let A be a commutative finitely generated connected graded Cohen–Macaulay domain. Then A is Gorenstein if and only if its Hilbert series $h_A(t)$ satisfies*

$$h_A(t^{-1}) = \pm t^{\ell} h_A(t) \tag{0.1.1}$$

for some integer ℓ.

The above theorem provides a surprising and effective criterion which is equivalent to the Gorenstein property. For example, if G is a finite group acting linearly on the polynomial ring $B := \Bbbk[x_1, \ldots, x_n]$, then the Hilbert series of the fixed subring B^G can be calculated using Molien's theorem [Kirkman 2016, Proposition 1.6]. One can then use (0.1.1) to determine whether or not B^G is Gorenstein. In general, verifying (0.1.1) is much easier than verifying the Gorenstein property directly. A noncommutative version of Stanley's criterion was proved in [Jørgensen and Zhang 2000, Theorems 6.1 and 6.2].

Note that a commutative connected graded algebra A is regular (i.e., has finite global dimension) if and only if A is isomorphic to a polynomial ring $\Bbbk[x_1, \ldots, x_n]$ with $\deg x_i > 0$ for all i. Hence, in the commutative case, checking whether or not A is regular is relatively easy. However, there are many more noncommutative connected graded regular algebras, namely the so-called Artin–Schelter regular

MSC2020: 16E10, 16E65, 20J99.
Keywords: Artin–Schelter regular algebra, Castelnuovo–Mumford regularity, Tor regularity, Koszul algebra, concavity.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

algebras, which we now define. For a connected graded algebra A, let $\mathfrak{m} = A_{\geq 1}$. The trivial graded A-bimodule $A/\mathfrak{m}$ is also denoted by $\Bbbk$.

Definition 0.2 [Artin and Schelter 1987, p. 171]. A connected graded algebra T is called *Artin–Schelter Gorenstein* (or *AS Gorenstein*, for short) if the following conditions hold:

(a) T has injective dimension $d < \infty$ on the left and on the right.

(b) $\operatorname{Ext}^i_T({}_T\Bbbk, {}_TT) = \operatorname{Ext}^i_T(\Bbbk_T, T_T) = 0$ for all $i \neq d$.

(c) $\operatorname{Ext}^d_T({}_T\Bbbk, {}_TT) \cong \operatorname{Ext}^d_T(\Bbbk_T, T_T) \cong \Bbbk(\mathfrak{l})$ for some integer $\mathfrak{l}$. Here $\mathfrak{l}$ is called the *AS index* of T.

In this case, we say T is of type $(d, \mathfrak{l})$. If, in addition,

(d) T has finite global dimension, and

(e) T has finite Gelfand–Kirillov dimension,

then T is called *Artin–Schelter regular* (or *AS regular*, for short) of dimension d.

In this paper we generally reserve the letters S and T for AS regular algebras. Artin–Schelter regular algebras play an important role in noncommutative algebraic geometry and many other subjects [Artin and Schelter 1987; Artin et al. 1990; 1991; Reyes et al. 2014]. Therefore, it would be very useful to have reasonably verifiable criteria for the AS regular property.

Recall that, for a graded left A-module M, we say that M *has a linear resolution* (or simply say that M *is linear*) if M has a minimal free A-resolution of the form

$$\cdots \to A(-i)^{\beta_i} \to A(-i+1)^{\beta_{i-1}} \to \cdots \to A(-1)^{\beta_1} \to A^{\beta_0} \to M \to 0 \tag{0.2.1}$$

for some integers β_i, or, equivalently, $\operatorname{Tor}^A_i(\Bbbk, M)_j = 0$ for all $j \neq i$. In this paper it is necessary to deal with (cochain) complexes of graded A-modules instead of graded A-modules. A complex X of graded left A-modules is naturally graded by homological and internal degrees. By examining certain complexes, properties of A can be reflected in the relationships between these degrees. For example, A is *Koszul* if the trivial graded A-module $\Bbbk$ has a linear resolution, or, equivalently if the *Tor regularity* (Definition 0.4) of $\Bbbk$ is 0. Here is one of our main results:

Theorem 0.3 (Theorem 4.6). *Let A be a noetherian connected graded s-Cohen–Macaulay algebra. Suppose that there is a graded algebra map $f : T \to A$, where T is a Koszul Artin–Schelter regular algebra, such that the induced modules A_T and ${}_TA$ are both finitely generated. Then the following are equivalent*:

(a) *A is Artin–Schelter regular.*

(b) *The Hilbert series $h_A(t)$ of A satisfies* $\deg h_A(t) = -s$.

In the commutative case, the existence of the desired map $f : T \to A$ is automatic (assuming A is generated in degree 1). Similar to Theorem 0.1, the combinatorial property in Theorem 0.3(b) is much easier to check than the definition of AS regularity and so Theorem 0.3 provides an important criterion

which can be used to determine whether or not A is AS regular. In the rest of paper, we will prove other criteria for determining AS regularity by using more sophisticated numerical homological invariants.

We now recall or introduce the following homological invariants, which will be studied throughout the paper:

- Tor regularity,
- Castelnuovo–Mumford regularity,
- (numerical) Artin–Schelter regularity,
- concavities associated to the Tor and Castelnuovo–Mumford regularities.

Definition 0.4 [Jørgensen 2004; 1999; Dong and Wu 2009]. The *Tor regularity* of a nonzero complex X of graded left A-modules is defined to be

$$\operatorname{Torreg}(X) = \sup_{i,j\in\mathbb{Z}} \{j - i \mid \operatorname{Tor}_i^A(\Bbbk, X)_j \neq 0\}. \tag{0.4.1}$$

It is clear that $\operatorname{Torreg}(X)$ provides a measure of the growth of the degrees of generators of the free modules in a minimal free resolution of X.

Since $\operatorname{Torreg}({}_A\Bbbk) \geq 0$ and the equality holds if and only if A is Koszul, $\operatorname{Torreg}({}_A\Bbbk)$ can be regarded as an invariant that measures how far A is from being Koszul, at least in the noncommutative setting. When A is a commutative algebra generated in degree one, $\operatorname{Torreg}({}_A\Bbbk)$ takes on only the values of 0 or ∞ [Avramov and Peeva 2001], so it indicates only whether or not A is Koszul. However, in the noncommutative case, $\operatorname{Torreg}({}_A\Bbbk)$ can be any value in $\mathbb{N} \cup \{+\infty\}$ (see Example 2.4(4) and Lemma 5.6).

Next we recall the definition of Castelnuovo–Mumford regularity, which was first introduced and studied in the noncommutative setting in [Jørgensen 2004; 1999; Dong and Wu 2009]. Recall that the *i-th local cohomology* of a complex X of graded left A-modules is defined to be

$$H^i_{\mathfrak{m}}(X) = \lim_{n\to\infty} \operatorname{Ext}^i_A(A/\mathfrak{m}^n, X). \tag{0.4.2}$$

See [Jørgensen 1997; 2004] and Section 1 for more details.

Definition 0.5 [Jørgensen 2004; 1999; Dong and Wu 2009]. Let A be a noetherian connected graded algebra and let X be a nonzero complex of graded left A-modules. The *Castelnuovo–Mumford regularity* (or *CM regularity*, for short) of X is defined to be

$$\operatorname{CMreg}(X) = \sup_{i,j\in\mathbb{Z}} \{j + i \mid H^i_{\mathfrak{m}}(X)_j \neq 0\}.$$

The notions of regularity defined in Definitions 0.4 and 0.5 are natural generalizations of the classical Tor and Castelnuovo–Mumford regularities in the commutative case. For a finitely generated graded A-module M, the relationship between the regularities $\operatorname{Torreg}(M)$ and $\operatorname{CMreg}(M)$ has been studied. When A is a polynomial ring generated in degree 1, $\operatorname{Torreg}(M) = \operatorname{CMreg}(M)$ for any finitely generated graded A-module M [Eisenbud and Goto 1984]. Other relationships between these invariants were established in the commutative case [Römer 2008] and were extended to the noncommutative case in [Jørgensen 2004;

1999; Dong and Wu 2009]. In this paper we provide further relationships between these invariants in the noncommutative case.

Now we introduce a new numerical homological invariant associated to every noetherian connected graded algebra A.

Definition 0.6. Let A be a noetherian connected graded algebra. The *Artin–Schelter regularity* (or *AS regularity*) of A is defined to be

$$\mathrm{ASreg}(A) = \mathrm{Torreg}(\Bbbk) + \mathrm{CMreg}(A).$$

We remark that, in the literature, the phrase "AS regularity of A" often refers to A possessing the Artin–Schelter regular property given in Definition 0.2. Here we use the phrase "AS regularity of A" to refer to the numerical invariant of A defined in Definition 0.6, which is related to the AS regular property. Indeed, by [Jørgensen 2004, Theorems 2.5 and 2.6] or Corollary 2.6(2), $\mathrm{ASreg}(A) \geq 0$; and, by Theorem 0.8, equality holds if and only if A is AS regular. Hence, $\mathrm{ASreg}(A)$ can be considered as an invariant that measures how far A is from being AS regular.

By Example 2.4(3), if T is AS regular of type (d, ℓ), then

$$\mathrm{CMreg}(T) = d - \ell = -\mathrm{Torreg}(\Bbbk), \tag{0.6.1}$$

which will appear in several places in this paper. As a consequence, $\mathrm{ASreg}(T) = 0$. Some further computations of regularities in the non-Koszul case are provided in Example 2.4(3).

The next theorem is another main result of the paper, which was announced in [Dong and Wu 2009, Proposition 5.6] without proof. It is also an extension of [Römer 2008, Theorem 4.2]. See Corollary 3.4 for a related result.

Theorem 0.7 (Theorem 2.8). *Let A be a noetherian connected graded algebra with balanced dualizing complex. Let X be a nonzero object in $\mathrm{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$ with finite projective dimension. Then*

$$\mathrm{CMreg}(X) = \mathrm{Torreg}(X) + \mathrm{CMreg}(A).$$

As a companion to Theorem 0.7, we generalize two very nice results of [Dong and Wu 2009, Theorems 4.10 and 5.4] to the not-necessarily Koszul setting.

Theorem 0.8. *Let A be a noetherian connected graded algebra with balanced dualizing complex. Then the following are equivalent*:

(a) *A is AS regular.*

(b) $\mathrm{ASreg}(A) = 0$.

Another main goal of this paper is to introduce a new notion of regularity, which we call the *concavity* of a graded algebra. If $f : A \to B$ is a graded algebra homomorphism between two $\mathbb{N}$-graded algebras. We say that f is a *finite map* if the modules ${}_AB$ and B_A are finitely generated.

Definition 0.9. Let $\mathcal{P}$ be a numerical invariant that is defined on all AS regular algebras. Let A be a locally finite $\mathbb{N}$-graded algebra (that is not necessarily connected graded).

(1) The $\mathcal{P}$-*concavity* of A is defined to be

$$c_{\mathcal{P}}(A) := \inf\{\mathcal{P}(T) \mid T \text{ is AS regular and there is a finite map } f : T \to A\}.$$

If no such T exists, we define $c_{\mathcal{P}}(A) = \infty$. A similar convention is used in the other parts of this definition.

If $\mathcal{P}$ is a numerical invariant that is defined on locally finite $\mathbb{N}$-graded algebras and if $\mathcal{P}(A)$ is finite, then we define the *normalized* $\mathcal{P}$-*concavity* of A to be

$$c_{\mathcal{P},-}(A) := c_{\mathcal{P}}(A) - \mathcal{P}(A).$$

Below are some special cases.

(2) If we take $\mathcal{P} = \mathrm{Torreg}$, then the Torreg-*concavity* of A is defined to be

$$c_{\mathrm{Torreg}}(A) := \inf\{\mathrm{Torreg}(T) \mid T \text{ is AS regular and there is a finite map } f : T \to A\}.$$

(3) If we take $\mathcal{P} = -\mathrm{CMreg}$, then the *concavity* of A is defined to

$$c(A) := \inf\{-\mathrm{CMreg}(T) \mid T \text{ is AS regular and there is a finite map } f : T \to A\}.$$

The *normalized concavity* of A is defined to be

$$c_{-}(A) := c(A) + \mathrm{CMreg}(A).$$

In this paper we focus primarily on the concavity and the normalized concavity of the invariant $\mathcal{P} = -\mathrm{CMreg}$. By (0.6.1), for any AS regular algebra T, $\mathrm{CMreg}(T) \leq 0$, and so, by definition, for any locally finite $\mathbb{N}$-graded algebra A, we have $c(A) \geq 0$. If S is a noetherian connected graded AS regular algebra, then $c(S) = 0$ if and only if S is Koszul (Theorem 0.10(2)). Let us explain the motivation behind the terminology "concavity". By analogy with the commutative case, for noncommutative T, we can imagine $\mathrm{Spec}\, T$ as a noncommutative space associated to T. If T is Koszul, then we should consider $\mathrm{Spec}\, T$ to be a flat space, as the minimal free resolution of the trivial module is linear (0.2.1). If $T_1 \to T_2$ is a finite map between two noetherian AS regular algebras (or, by analogy, if there is a finite map $\mathrm{Spec}\, T_2 \to \mathrm{Spec}\, T_1$), then we can show that $c(T_2) \leq c(T_1)$ and so the concavity of T_2 is bounded by the concavity of T_1. As a consequence, if T_1 is Koszul, then so is T_2 (Theorem 0.10(3)). Hence, in some sense, "concavity" measures how far away a noncommutative space is from being "flat".

By definition, the concavity of A depends on all finite maps f from AS regular algebras T to A. However, when A itself is AS regular, this invariant can be calculated by taking f to be the identity map, as shown in the following theorem, which can be considered as an extension of Theorem 0.3 in the case that $A = S$ is AS regular:

Theorem 0.10. *Let T and S be noetherian AS regular algebras.*

(1) $c(S) = -\mathrm{CMreg}(S)$.

(2) *S is Koszul if and only if $c(S) = 0$.*

(3) *Suppose $f : T \to S$ is a finite map. Then $c(T) \geq c(S) \geq 0$. If further $c(T) = c(S)$, then f is surjective and ${}_TS$ and S_T are linear Cohen–Macaulay modules over T.*

As a corollary to the above results, we see that, if there is a finite map between AS regular algebras $f : T \to S$, then the Koszul property of S is controlled by the Koszul property of T.

Corollary 0.11 (Corollary 4.4). *Let T and S be noetherian AS regular algebras and let $f : T \to S$ be a finite map.*

(1) *If T is Koszul (or, equivalently, $\mathrm{Torreg}({}_T\Bbbk) = 0$), then S is Koszul. Further, f is surjective and ${}_TS$ and S_T are linear Cohen–Macaulay modules over T.*

(2) *Suppose $\mathrm{Torreg}({}_T\Bbbk) \leq \deg(\Bbbk \otimes_T S)$. Then S is Koszul.*

(3) *Suppose $\mathrm{Torreg}({}_T\Bbbk) = 1$. If f is not surjective, then S is Koszul. As a consequence, if T is a proper subalgebra of S, then S is Koszul.*

We also show that (normalized) concavity is related to the AS regular property.

Theorem 0.12. *Let A be a noetherian connected graded algebra with balanced dualizing complex. Let $c_-(A)$ be the normalized concavity defined in Definition 0.9(3). Then the following hold*:

(1) $c_-(A) \geq 0$.

(2) $c_-(A) = 0$ *if and only if A is AS regular.*

By Theorem 0.12, $c_-(A)$ can be viewed as an invariant, similar to $\mathrm{ASreg}(A)$, that measures how far away A is from being AS regular. Here is a list of numerical invariants that qualify as indicators of the AS regular property for different special classes of algebras:

(i) $\deg_t h_A(t)$ (Theorem 0.3),

(ii) $\mathrm{ASreg}(A)$ (Theorem 0.8),

(iii) $c_-(A)$ (Theorem 0.12),

(iv) $\mathrm{CMreg}(A)$ (Theorem 4.5), and

(v) $\mathrm{Torreg}({}_TA)$ (Remark 5.10).

Remark 0.13. As noted in [Stanley 1979, p. 506] (or [Kirkman et al. 2015, p. 256]) we have the following hierarchy of homological properties for commutative graded algebras:

$$\text{regular} \Longrightarrow \text{hypersurface} \Longrightarrow \text{complete intersection} \Longrightarrow \text{Gorenstein} \Longrightarrow \text{Cohen–Macaulay}.$$

In this paper, we mainly study the "regular" property. It would be very interesting if new numerical (and hopefully computable) invariants could be found to detect or characterize the other properties in the above

diagram. Of course, Stanley's theorem (Theorem 0.1) for the Gorenstein property is our model. Some necessary conditions to be a "noncommutative complete intersection" are given in [Kirkman et al. 2015].

To conclude the introduction, we mention one application of concavity to noncommutative invariant theory. Kirkman et al. [2022] used CMreg to bound the degrees of generators of invariant subrings T^H when a semisimple Hopf algebra H acts on an AS regular algebra T homogeneously. The following proposition is an easy consequence of [ibid., Theorem 0.8]. For any connected graded algebra A, let

$$\beta_i(A) = \deg \operatorname{Tor}_i^A(\Bbbk, \Bbbk).$$

Then $\beta_1(A)$ is the largest degree of an element in a minimal generating set of A and $\beta_2(A)$ is the largest degree of an element in a minimal relation set of A. We have the following lower bounds for $c(T^H)$ in terms of the degrees of the generators and the relations of T^H:

Proposition 0.14. *Let H be a semisimple Hopf algebra acting on a noetherian AS regular algebra T homogeneously. Let $R = T^H$ denote the invariant subring of this action. Then the following hold*:

(1) $c(R) \geq \beta_1(R) - 1$.

(2) $c(R) \geq \min\{\frac{1}{2}\beta_2(R) - \operatorname{CMreg}(T), \frac{1}{2}(\beta_2(R) - \operatorname{CMreg}(T) - 1), \beta_2(R) - 2\}$.

By [Kirkman et al. 2009, Lemma 3.2(b)], T^H is Cohen–Macaulay with balanced dualizing complex. Now, if we are given a noetherian Cohen–Macaulay domain A with balanced dualizing complex (or even a noetherian AS Gorenstein domain), it is generally very difficult to determine if there exist (T, H) such that A is isomorphic to T^H. The simple inequality in Proposition 0.14(1) provides an easy way of showing that a noetherian connected graded algebra A cannot be isomorphic to any invariant subring T^H of an AS regular algebra T under a semisimple Hopf algebra H action (if $c(A) < \beta_1(A) - 1$). For example, let

$$A = \Bbbk[x_1, \ldots, x_n][t]/(t^2 = f(x_1, \ldots, x_n)),$$

where $\deg x_i = 1$, $\deg t \geq 2$ and f is an irreducible homogeneous polynomial in x_i of degree $2 \deg t$. It follows from the definition that

$$0 = c(A) < 1 \leq \deg t - 1 = \beta_1(A) - 1.$$

Therefore, A cannot be isomorphic to T^H by Proposition 0.14(1). See Example 4.2 for further examples.

The paper is organized as follows. Section 1 recalls some basic definitions and properties from homological algebra (including local cohomology). Section 2 gives the definitions of invariants such as Castelnuovo–Mumford regularity, Ext or Tor regularity, and basic inequalities and equalities relating these regularities. Theorem 0.7 is proved in Section 2. In Section 3 we define AS regularity and prove Theorem 0.8. Section 4 concerns concavity associated to the Castelnuovo–Mumford regularity. We prove Theorems 0.10 and 0.12 and Proposition 0.14 on concavity in Section 4. Finally, Section 5 contains some examples, questions and remarks.

1. Preliminaries

For an $\mathbb{N}$-graded $\Bbbk$-algebra A, we let A-Gr denote the category of $\mathbb{Z}$-graded left A-modules. When convenient, we identify the graded right A-modules with graded left A^{op}-modules and denote the category by A^{op}-Gr. The derived category of complexes of graded A-modules is denoted by $\mathsf{D}(A\text{-Gr})$. We use the standard notation $\mathsf{D}^+(A\text{-Gr})$, $\mathsf{D}^-(A\text{-Gr})$ and $\mathsf{D}^{\mathrm{b}}(A\text{-Gr})$ for the full subcategories of complexes which are bounded below, bounded above and bounded, respectively. We use the subscript fg to denote the full subcategories consisting of complexes with finitely generated cohomology, e.g., $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. We use the standard convention that a left A-module M can be viewed as complex concentrated in position 0.

Let ℓ be an integer. For a graded A-module M, the shifted A-module $M(\ell)$ is defined by

$$M(\ell)_m = M_{m+\ell}$$

for all $m \in \mathbb{Z}$. For a cochain complex $X = (X^n, d^n : X^n \to X^{n+1})$, we define two notions of shifting: $X(\ell)$ shifts the degrees of each graded vector space $X^i(\ell)_m = X^i_{m+\ell}$ for all $i, m \in \mathbb{Z}$; and $X[\ell]$ shifts the complex $X^i[\ell] = X^{i+\ell}$ for all $i \in \mathbb{Z}$.

Definition 1.1. Let $A := \bigoplus_{i\geq 0} A_i$ be an $\mathbb{N}$-graded locally finite algebra. The *Hilbert series of* A is defined to be

$$h_A(t) = \sum_{i\in\mathbb{N}} (\dim_{\Bbbk} A_i) t^i.$$

Similarly, if $M = \bigoplus_{i\in\mathbb{Z}} M_i$ is a $\mathbb{Z}$-graded A-module (or $\mathbb{Z}$-graded vector space), the *Hilbert series of* M is defined to be

$$h_M(t) = \sum_{i\in\mathbb{Z}} (\dim_{\Bbbk} M_i) t^i.$$

We say that M is *locally finite* if $\dim_{\Bbbk} M_d < \infty$ for all $d \in \mathbb{Z}$. Define the *degree* of M to be the maximal degree of the nonzero homogeneous elements in M, namely

$$\deg(M) = \inf\{d \mid (M)_{\geq d} = 0\} - 1 = \sup\{d \mid (M)_d \neq 0\} \in \mathbb{Z} \cup \{\pm\infty\}. \tag{1.1.1}$$

By convention, we define $\deg(0) = -\infty$. Similarly, we define

$$\operatorname{ged}(M) = \sup\{d \mid (M)_{\leq d} = 0\} + 1 = \inf\{d \mid (M)_d \neq 0\} \in \mathbb{Z} \cup \{\pm\infty\}. \tag{1.1.2}$$

By convention, we define $\operatorname{ged}(0) = \infty$.

For a nonzero cochain complex X in $\mathsf{D}(A\text{-Gr})$, the *degree* of X is defined to be

$$\deg(X) = \sup_{m,n\in\mathbb{Z}} \{m + n \mid H^n(X)_m \neq 0\} = \sup_{n\in\mathbb{Z}} \{\deg H^n(X) + n\}, \tag{1.1.3}$$

where $H^d(X)$ is the d-th homology of the complex X. Similarly, the ged of X is defined to be

$$\operatorname{ged}(X) = \inf_{m,n\in\mathbb{Z}} \{m + n \mid H^n(X)_m \neq 0\} = \inf_{n\in\mathbb{Z}} \{\operatorname{ged} H^n(X) + n\}. \tag{1.1.4}$$

We also define

$$\sup(X) = \sup\{d \mid H^d(X) \neq 0\}$$

and

$$\inf(X) = \inf\{d \mid H^d(X) \neq 0\}.$$

Let A be a connected graded algebra. Recall that $\mathfrak{m}$ denotes the graded Jacobson radical (or maximal graded ideal) $A_{\geq 1}$ and that $\Bbbk$ denotes the graded A-bimodule $A/\mathfrak{m}$. For a graded left A-module M, let

$$t_i^A({}_AM) = \deg \operatorname{Tor}_i^A(\Bbbk, M). \tag{1.1.5}$$

If M is a graded right A-module, let

$$t_i^A(M_A) = \deg \operatorname{Tor}_i^A(M, \Bbbk). \tag{1.1.6}$$

It is clear that $t_i^A({}_A\Bbbk) = t_i^A(\Bbbk_A)$. If the context is clear, we will use $t_i^A(M)$ instead of $t_i^A({}_AM)$ (or $t_i^A(M_A)$).

For each graded left A-module M, we define

$$\Gamma_{\mathfrak{m}}(M) = \{x \in M \mid A_{\geq n}x = 0 \text{ for some } n \geq 1\} = \lim_{n\to\infty} \operatorname{Hom}_A(A/A_{\geq n}, M)$$

and call this the $\mathfrak{m}$-*torsion submodule* of M. It is standard that the functor $\Gamma_{\mathfrak{m}}(-)$ is a left exact functor $A\text{-Gr} \to A\text{-Gr}$. Since this category has enough injectives, the i-th right derived functors, denoted by $H^i_{\mathfrak{m}}$ or $\mathrm{R}^i\Gamma_{\mathfrak{m}}$, are defined and called the *local cohomology functors*; see [Artin and Zhang 1994; Van den Bergh 1997; Jørgensen 1997; 2004] for more details. For a complex X of graded left A-modules, the i-th local cohomology group of X is given by (0.4.2). For example, if M is a graded left A-module, then

$$H^i_{\mathfrak{m}}(M) = \mathrm{R}^i\Gamma_{\mathfrak{m}}(M) := \lim_{n\to\infty} \operatorname{Ext}^i_A(A/A_{\geq n}, M). \tag{1.1.7}$$

Definition 1.2. Let A be a connected graded noetherian graded algebra. Let M be a finitely generated graded left A-module. We call M *s-Cohen–Macaulay* or simply *Cohen–Macaulay* if $H^i_{\mathfrak{m}}(M) = 0$ for all $i \neq s$ and $H^s_{\mathfrak{m}}(M) \neq 0$. We say A is *Cohen–Macaulay* if ${}_AA$ is Cohen–Macaulay.

Throughout the rest of this paper, when we need a dualizing complex we assume the following hypothesis; we refer the reader to [Yekutieli 1992] for the definitions of a dualizing complex and a balanced dualizing complex.

Hypothesis 1.3. Let A be a noetherian connected graded algebra with balanced dualizing complex. In this case, by [Van den Bergh 1997, Theorem 6.3], the balanced dualizing complex will be given by $\mathrm{R}\Gamma_{\mathfrak{m}}(A)'$, where $'$ denotes the graded vector space dual.

The local cohomological dimension of a graded A-module M is defined to be

$$\operatorname{lcd}(M) := \sup\{i \in \mathbb{Z} \mid H^i_{\mathfrak{m}}(M) \neq 0\}$$

and the cohomological dimension of $\Gamma_{\mathfrak{m}}$ is defined to be

$$\operatorname{cd}(\Gamma_{\mathfrak{m}}) = \sup_{M\in A\text{-Gr}} \{\operatorname{lcd}(M)\}.$$

We will use the following *local duality theorem* of Van den Bergh several times.

Theorem 1.4 [Van den Bergh 1997, Theorem 5.1]. *Let A be a noetherian connected graded $\Bbbk$-algebra with $\operatorname{cd}(\Gamma_{\mathfrak{m}}) < \infty$ and let C be a connected graded $\Bbbk$-algebra. Then, for any $X \in \mathsf{D}((A \otimes C^{\mathrm{op}})\text{-Gr})$, there is an isomorphism*

$$\mathrm{R}\Gamma_{\mathfrak{m}}(X)' \cong \mathrm{RHom}_A(X, \mathrm{R}\Gamma_{\mathfrak{m}}(A)')$$

in $\mathsf{D}((C \otimes A^{\mathrm{op}})\text{-Gr})$.

2. Equalities and inequalities

In this section we study the relationships between the regularities defined in the previous sections, recalling and generalizing results of [Jørgensen 2004; 1999; Dong and Wu 2009]. Throughout this section, we assume Hypothesis 1.3.

Recall, from Definition 0.5, that the Castelnuovo–Mumford regularity of cochain complex X of left A-modules is defined to be

$$\mathrm{CMreg}(X) = \deg(\mathrm{R}\Gamma_{\mathfrak{m}}(X)) = \sup_{i,j \in \mathbb{Z}} \{j + i \mid H^i_{\mathfrak{m}}(X)_j \neq 0\}.$$

As noted in [Jørgensen 2004, Observation 2.3], if $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$, then $\mathrm{R}\Gamma_{\mathfrak{m}}(X)' \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A^{\mathrm{op}}\text{-Gr})$ and $\mathrm{R}\Gamma_{\mathfrak{m}}(X)' \ncong 0$. It follows that $\mathrm{CMreg}(X)$ is finite. In particular, if A is a finitely generated commutative algebra, then $\mathrm{CMreg}(X)$ is finite for all $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. In Example 5.1, we show that there exists a noetherian connected domain which does not satisfy Hypothesis 1.3 with $\mathrm{GKdim}\, A = 2$ and $\mathrm{CMreg}(A) = \infty$.

Example 2.1. Assume Hypothesis 1.3.

(1) If M is a finite-dimensional nonzero graded left A-module, then $H^i_{\mathfrak{m}}(M) = 0$ for all $i \neq 0$ and $H^0_{\mathfrak{m}}(M) = M$, so

$$\mathrm{CMreg}(M) = \deg(M). \tag{2.1.1}$$

A more general case is considered in part (4), as a finite-dimensional module is 0-Cohen–Macaulay.

(2) Let A be an AS Gorenstein algebra of type (d, ℓ). Then $\mathrm{CMreg}(A) = d - \ell$. This is a well-known fact, which is a consequence of [Artin and Zhang 1994, Theorem 8.1(3)].

(3) Let A be an AS regular algebra of type (d, ℓ). Recall that, when regarded as a rational function, $\deg_t h_A(t) = -\ell$ [Stephenson and Zhang 1997, Proposition 3.1(4)]. Hence,

$$\mathrm{CMreg}(A) = d - \ell = \mathrm{gldim}\, A + \deg_t h_A(t).$$

By the second statement in [Stephenson and Zhang 1997, Proposition 3.1(4)], $d \leq \ell$. As a consequence,

$$\mathrm{CMreg}(A) = d - \ell \leq 0. \tag{2.1.2}$$

(4) If M is s-Cohen–Macaulay, then, by definition,

$$\mathrm{CMreg}(M) = s + \deg(H^s_{\mathfrak{m}}(M)). \tag{2.1.3}$$

Recall that the *Tor regularity* of a nonzero complex X of graded left A-modules was defined in Definition 0.4, namely

$$\operatorname{Torreg}(X) = \deg(\Bbbk \otimes_A^L X) = \sup_{i,j\in\mathbb{Z}} \{j - i \mid \operatorname{Tor}_i^A(\Bbbk, X)_j \neq 0\}.$$

Definition 2.2. Let X be a nonzero object in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. The Ext *regularity* of X is defined to be

$$\operatorname{Extreg}(X) = -\operatorname{ged}(\mathrm{R}\operatorname{Hom}_A(X, \Bbbk)) = -\inf_{i\in\mathbb{Z}}\{\operatorname{ged}(\operatorname{Ext}_A^i(X, \Bbbk)) + i\}.$$

By [Dong and Wu 2009, Remark 4.5], if X has a finitely generated minimal free resolution over A, then $\operatorname{Extreg}(X) = \operatorname{Torreg}(X)$, and we will not distinguish between $\operatorname{Extreg}(X)$ and $\operatorname{Torreg}(X)$ in this case. The following easy lemma is useful for computing Torreg of modules over tensor products. Throughout this paper, $\otimes$ means $\otimes_{\Bbbk}$.

Lemma 2.3. *Let A and B be connected graded algebras. Let P (resp. Q) be a nonzero object in $\mathsf{D}^-_{\mathrm{fg}}(A\text{-Gr})$ (resp. $\mathsf{D}^-_{\mathrm{fg}}(B\text{-Gr})$). Then* $\operatorname{Torreg}({}_{A\otimes B}P \otimes Q) = \operatorname{Torreg}({}_A P) + \operatorname{Torreg}({}_B Q)$.

Proof. Replacing P by its (minimal) free resolution, we can assume that each term P^i is a finitely generated free A-module. The same applies to Q. Let X and Y denote the complexes given by tensoring P and Q with $\Bbbk_A$ and $\Bbbk_B$, respectively, on the left. Then $\operatorname{Tor}_i^A(\Bbbk, P)$ and $\operatorname{Tor}_i^B(\Bbbk, Q)$ can be computed by taking homology of X and Y, respectively. Further, $\operatorname{Tor}_i^{A\otimes B}(\Bbbk, P \otimes Q)$ can be computed by taking homology of the complex $X \otimes Y$. By the Künneth formula (see, e.g., [Rotman 1979, Theorem 10.8.1]), we have

$$\bigoplus_{p+q=n} \operatorname{Tor}_p^A(\Bbbk, P) \otimes \operatorname{Tor}_q^B(\Bbbk, Q) \cong \bigoplus_{p+q=n} H_p(X) \otimes H_q(Y) \cong H_n(X \otimes Y) \cong \operatorname{Tor}_n^{A\otimes B}(\Bbbk, P \otimes Q).$$

Therefore, using the convention that $\deg(0) = -\infty$,

$$\begin{aligned}
\operatorname{Torreg}({}_{A\otimes B}P \otimes Q) &= \sup_{n\in\mathbb{Z}}\{\deg(\operatorname{Tor}_n^{A\otimes B}(\Bbbk, P \otimes Q)) - n\} \\
&= \sup_{p,q\in\mathbb{Z}} \{\deg(\operatorname{Tor}_p^A(\Bbbk, P)) + \deg(\operatorname{Tor}_q^B(\Bbbk, Q)) - (p+q)\} \\
&= \sup_{p\in\mathbb{Z}}\{\deg(\operatorname{Tor}_p^A(\Bbbk, P)) - p\} + \sup_{q\in\mathbb{Z}}\{\deg(\operatorname{Tor}_q^A(\Bbbk, Q)) - q\} \\
&= \operatorname{Torreg}({}_A P) + \operatorname{Torreg}({}_B Q),
\end{aligned}$$

as desired. □

Example 2.4. Assume Hypothesis 1.3.

(1) If $M \in A\text{-Gr}$ and $r = \operatorname{Torreg}(M)$, then

$$t_i^A({}_A M) := \deg(\operatorname{Tor}_i^A(\Bbbk, M)) \leq (r + i) \tag{2.4.1}$$

for all i.

(2) $\operatorname{Extreg}(A) = \operatorname{Torreg}(A) = 0$.

(3) Let T be any noetherian AS regular algebra of type (d, ℓ). It is well known that

$$\operatorname{Torreg}({}_T\Bbbk) = \ell - d \quad (\text{which is equal to } -\operatorname{CMreg}(T)). \tag{2.4.2}$$

This assertion follows from [Stephenson and Zhang 1997, Proposition 3.1(3) and equation (3.2)].

Next we give an explicit example. Let T be a non-Koszul AS regular algebra of global dimension 3 that is generated in degree 1. By [Artin and Schelter 1987], T is generated by two elements that satisfy cubic relations. So the minimal free resolution of ${}_T\Bbbk$ has the form

$$0 \to T(-4) \to T(-3)^{\oplus 2} \to T(-1)^{\oplus 2} \to T \to \Bbbk \to 0.$$

Therefore, T is of type $(3, 4)$ and

$$t_i^T(\Bbbk) = \begin{cases} 0 & \text{if } i = 0, \\ 1 & \text{if } i = 1, \\ 3 & \text{if } i = 2, \\ 4 & \text{if } i = 3, \\ -\infty & \text{if } i > 3. \end{cases}$$

By Example 2.1(3), $\operatorname{CMreg}(T) = -4 + 3 = -1$ and it is easy to check that

$$\operatorname{Torreg}(\Bbbk) = \max\{0, 1-1, 3-2, 4-3, -\infty\} = 1.$$

As a consequence, $\operatorname{Torreg}(\Bbbk) = -\operatorname{CMreg}(T)$, or, equivalently,

$$\operatorname{ASreg}(T) = 0.$$

Of course, this equation always holds for all AS regular algebras by (2.4.2).

(4) Let n be a fixed positive integer. Let A be any finitely generated commutative Koszul algebra (but not regular) and let T be the algebra in part (3). Let $B = A \otimes T^{\otimes n}$. By a similar argument to [Kirkman et al. 2023, Lemma 2.7], using the Künneth formula, one can easily check that $\operatorname{Torreg}({}_B\Bbbk) = n$ (Lemma 2.3) and that B is neither AS regular nor Koszul.

The following result of Jørgensen plays an important role in this paper:

Theorem 2.5 [Jørgensen 2004, Theorems 2.5 and 2.6]. *Let A be a noetherian connected graded algebra with a balanced dualizing complex, and let X be a nonzero object in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.*

(1) $\operatorname{Torreg}(X) \le \operatorname{CMreg}(X) + \operatorname{Torreg}(\Bbbk)$.

(2) $\operatorname{CMreg}(X) \le \operatorname{Torreg}(X) + \operatorname{CMreg}(A)$.

We have the following immediate consequences:

Corollary 2.6. *Let A be a noetherian connected graded algebra with a balanced dualizing complex and let X be a nonzero object in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.*

(1) *If* $\operatorname{Torreg}(\Bbbk)$ *is finite, then so is* $\operatorname{Torreg}(X)$.

(2) $\operatorname{ASreg}(A) \ge 0$.

(3) *If* $\mathrm{ASreg}(A) = 0$, *then equality holds in both parts of Theorem 2.5. Namely*,

$$\mathrm{Torreg}(X) = \mathrm{CMreg}(X) + \mathrm{Torreg}(\Bbbk).$$

Proof. (1) By Theorem 2.5(1),

$$\mathrm{Torreg}(X) \leq \mathrm{CMreg}(X) + \mathrm{Torreg}(\Bbbk),$$

and, since $\mathrm{CMreg}(X)$ is finite, the result follows.

(2) The statement follows by taking the sum of two inequalities in Theorem 2.5.

(3) If $\mathrm{ASreg}(A) = 0$, then $\mathrm{Torreg}(\Bbbk) = -\mathrm{CMreg}(A)$. Hence, using both inequalities in Theorem 2.5, we have

$$\mathrm{Torreg}(X) \leq \mathrm{CMreg}(X) + \mathrm{Torreg}(\Bbbk) = \mathrm{CMreg}(X) - \mathrm{CMreg}(A) \leq \mathrm{Torreg}(X)$$

and so equality holds throughout. □

We will show that if X has finite projective dimension, then Theorem 2.5(2) becomes an equality. We will need the following straightforward lemma. Recall that, for a cochain complex X and $\ell \in \mathbb{Z}$, the complexes $X(\ell)$ and $X[\ell]$ were defined at the beginning of Section 1.

Lemma 2.7. *Let A be a noetherian connected graded algebra and X be a nonzero complex of graded left A-modules. Suppose* $\deg(X)$ *is finite. Then*

$$\deg(X[1]) = \deg(X) - 1 \quad \textit{and} \quad \deg(X(1)) = \deg(X) - 1.$$

Similar equations hold for $\mathrm{ged}(X)$, $\mathrm{CMreg}(X)$, $\mathrm{Extreg}(X)$ *and* $\mathrm{Torreg}(X)$.

For a cochain complex

$$X = (\cdots \to X^{s-1} \to X^s \to X^{s+1} \to \cdots),$$

we denote the *brutal truncations* of X by

$$X^{\geq s} := (\cdots \to 0 \to \cdots \to 0 \to X^s \to X^{s+1} \to \cdots)$$

and

$$X^{\leq s} := (\cdots \to X^{s-1} \to X^s \to 0 \to \cdots \to 0 \to \cdots).$$

(We remark that the notation $X^{\geq s}$ and $X^{\leq s}$ may mean different truncations in other papers.)

Theorem 2.8. *Let A be a noetherian connected graded algebra with a balanced dualizing complex. Let X be a nonzero object in* $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$ *with finite projective dimension. Then*

$$\mathrm{CMreg}(X) = \mathrm{Torreg}(X) + \mathrm{CMreg}(A).$$

Proof. By Lemma 2.7, we may assume that $X^n = 0$ for all $n \geq 1$. Let F be a minimal free resolution of X, which we write as

$$F := (\cdots \to 0 \to F^{-s} \xrightarrow{d^{-s}} \cdots \to F^{-1} \xrightarrow{d^{-1}} F^0 \to 0 \to \cdots)$$

for some $s \geq 0$. We will prove the assertion by induction on s, which is (an upper bound on) the projective dimension of X.

For the initial step, we assume that $s = 0$, or $X = F^0 = \bigoplus_i A(-a_i)$ for some integers a_i. In this case, it is clear that

$$\operatorname{Torreg}(X) = \operatorname{Torreg}(F^0) = \max_i\{a_i\} =: a.$$

By Lemma 2.7,

$$\begin{aligned}\operatorname{CMreg}(X) &= \operatorname{CMreg}\left(\bigoplus_i A(-a_i)\right) = \max_i\{\operatorname{CMreg}(A(-a_i))\}\\ &= \operatorname{CMreg}(A) + \max_i\{a_i\} = \operatorname{CMreg}(A) + \operatorname{Torreg}(X),\end{aligned}$$

so the assertion holds for $X = F^0$, as required.

For the inductive step, assume that $s > 0$. Let $F^{\leq -1}$ be the brutal truncation of the complex F,

$$F^{\leq -1} := (\cdots \to 0 \to F^{-s} \to \cdots \to F^{-1} \to 0 \to 0 \to \cdots),$$

which is obtained by replacing F^0 by 0. We have a distinguished triangle in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$,

$$F^0 \xrightarrow{f} F \to F^{\leq -1} \to F^0[1], \tag{2.8.1}$$

where F^0 is viewed as a complex concentrated at position 0 and f is the inclusion. Let G be the complex $F^{\leq -1}[-1]$, which is a minimal free complex concentrated in position $\{-(s-1), \ldots, 0\}$. Then we have a distinguished triangle in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$,

$$G \xrightarrow{\phi_2} F^0 \xrightarrow{f} F \to G[1], \tag{2.8.2}$$

obtained by rotating (2.8.1). By the induction hypothesis, the assertion holds for both G and F^0. We need to show that the assertion holds for X, or, equivalently, for F, as $F \cong X$ in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. By Theorem 2.5(2), the assertion is equivalent to

$$\operatorname{CMreg}(F) \geq \operatorname{Torreg}(F) + \operatorname{CMreg}(A). \tag{2.8.3}$$

We fix the temporary notation

$$a = \operatorname{Torreg}(F^0), \quad b = \operatorname{Torreg}(G), \quad c = \operatorname{Torreg}(F) = \operatorname{Torreg}(X),$$

and

$$\alpha = \operatorname{CMreg}(F^0), \quad \beta = \operatorname{CMreg}(G), \quad \gamma = \operatorname{CMreg}(F) = \operatorname{CMreg}(X).$$

By definition and the minimality of F, we have

$$c = \max\{\operatorname{Torreg}(F^0), \operatorname{Torreg}(F^{\leq -1})\} = \max\{a, b-1\}.$$

By the above equation, we have

$$a \leq c \quad \text{and} \quad b - 1 \leq c, \tag{2.8.4}$$

and c must equal to either a or $b-1$. There are two cases:

(1) $c = a$ and $a \geq b$.

(2) $c = b-1$,

Case 1: Suppose that $c = a$ and $a \geq b$. By the definition of a, we have $F^0 = A(-a) \oplus C^0$, where C^0 is a graded free left A-module. Let $\phi_1 : F^0 \to A(-a)$ be the corresponding projection. By the definition of $\alpha := \mathrm{CMreg}(F^0)$, there is an integer $j \in \mathbb{Z}$ such that $H^j_{\mathfrak{m}}(F^0)_{\alpha-j} \neq 0$ and the induced projection

$$\tau_1 := H^j_{\mathfrak{m}}(\phi_1)_{\alpha-j} : H^j_{\mathfrak{m}}(F^0)_{\alpha-j} \to H^j_{\mathfrak{m}}(A(-a))_{\alpha-j}$$

is nonzero. The triangle (2.8.2) gives rise to a long exact sequence

$$\cdots \to H^j_{\mathfrak{m}}(G)_{\alpha-j} \to H^j_{\mathfrak{m}}(F^0)_{\alpha-j} \to H^j_{\mathfrak{m}}(F)_{\alpha-j} \to H^{j+1}_{\mathfrak{m}}(G)_{\alpha-j} \to \cdots . \tag{2.8.5}$$

If

$$\tau_2 := H^j_{\mathfrak{m}}(\phi_2)_{\alpha-j} : H^j_{\mathfrak{m}}(G)_{\alpha-j} \to H^j_{\mathfrak{m}}(F^0)_{\alpha-j}$$

is not surjective, then (2.8.5) implies that $H^j_{\mathfrak{m}}(F)_{\alpha-j} \neq 0$. By definition, the assumption that $a = c$, and the induction hypothesis, we have

$$\begin{aligned}
\mathrm{CMreg}(F) &\geq \alpha - j + j = \alpha \\
&= a + \mathrm{CMreg}(A) && \text{(induction hypothesis)} \\
&= c + \mathrm{CMreg}(A) && \text{(assumption in Case 1)} \\
&= \mathrm{Torreg}(F) + \mathrm{CMreg}(A),
\end{aligned}$$

as desired.

It remains to show the claim that τ_2 is not surjective. Assume to the contrary that τ_2 is surjective. Then so is the composed map

$$\tau_3 := \tau_1 \circ \tau_2 : H^j_{\mathfrak{m}}(G)_{\alpha-j} \to H^j_{\mathfrak{m}}(A(-a))_{\alpha-j}.$$

In particular, τ_3 is not a zero map. Note that

$$\tau_3 = \tau_1 \circ \tau_2 = H^j_{\mathfrak{m}}(\phi_1)_{\alpha-j} \circ H^j_{\mathfrak{m}}(\phi_2)_{\alpha-j} = H^j_{\mathfrak{m}}(\phi_1 \circ \phi_2)_{\alpha-j},$$

which implies that $\phi_3 := \phi_1 \circ \phi_2$ is nonzero in $\mathrm{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. Consider F as the cone of the map $\phi_2 : G \to F^0$; it is clear that ϕ_2 is the map from the top row G to the middle row F^0 in the diagram

$$\begin{array}{ccccccccc}
F^{-s} & \longrightarrow & \cdots & \longrightarrow & F^{-2} & \longrightarrow & F^{-1} & \longrightarrow & 0 \\
\downarrow{\scriptstyle 0} & & & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle d^{-1}=\phi_2} & & \\
0 & \longrightarrow & \cdots & \longrightarrow & 0 & \longrightarrow & F^0 & \longrightarrow & 0 \\
\downarrow{\scriptstyle 0} & & & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle \phi_1} & & \\
0 & \longrightarrow & \cdots & \longrightarrow & 0 & \longrightarrow & A(-a) & \longrightarrow & 0
\end{array}$$

Since $b \leq a$, F^{-1}, which is the zeroth term in the minimal free resolution of G, is generated in degree $\leq a$. Since F is a minimal free resolution, $\operatorname{im}\phi_2 = \operatorname{im} d^{-1} \subseteq \mathfrak{m}F^0$, and, consequently, $\operatorname{im}\phi_3 \subseteq \mathfrak{m}A(-a)$. For every generator x in F^{-1} which has degree $\leq a$, the image $\phi_3(x)$ lies in $\mathfrak{m}A(-a)$, which has degree at least $a+1$. Therefore, $\phi_3(x) = 0$. This implies that $\phi_3(F^{-1}) = 0$, yielding a contradiction. So we have proved the claim and finished the proof in Case 1.

Case 2: Suppose $c = b - 1$. By the definition of $\beta := \operatorname{CMreg}(G)$, there is an integer $j \in \mathbb{Z}$ such that $H^j_{\mathfrak{m}}(G)_{\beta-j} \neq 0$. The triangle (2.8.2) gives rise to a long exact sequence

$$\cdots \to H^{j-1}_{\mathfrak{m}}(F)_{\beta-j} \to H^j_{\mathfrak{m}}(G)_{\beta-j} \to H^j_{\mathfrak{m}}(F^0)_{\beta-j} \to \cdots. \tag{2.8.6}$$

By the induction hypothesis, the assumption that $c < b$, (2.8.4), and the definitions of α, β, a and b, we have

$$\begin{aligned}
\beta &= \operatorname{CMreg}(G) \\
&= \operatorname{Torreg}(G) + \operatorname{CMreg}(A) && \text{(induction hypothesis)} \\
&= b + \operatorname{CMreg}(A) > c + \operatorname{CMreg}(A) \\
&\geq a + \operatorname{CMreg}(A) = \operatorname{CMreg}(F^0) && \text{(induction hypothesis)} \\
&= \alpha,
\end{aligned}$$

and thus $H^j_{\mathfrak{m}}(F^0)_{\beta-j} = 0$. Since $H^j_{\mathfrak{m}}(G)_{\beta-j} \neq 0$ by definition, (2.8.6) implies that $H^{j-1}_{\mathfrak{m}}(F)_{\beta-j} \neq 0$. By definition, $\operatorname{CMreg}(F) \geq \beta - j + (j-1) = \beta - 1$. This inequality implies that

$$\begin{aligned}
\operatorname{CMreg}(F) \geq \beta - 1 &= \operatorname{CMreg}(G) - 1 \\
&= \operatorname{Torreg}(G) + \operatorname{CMreg}(A) - 1 && \text{(induction hypothesis)} \\
&= b + \operatorname{CMreg}(A) - 1 = (b-1) + \operatorname{CMreg}(A) \\
&= c + \operatorname{CMreg}(A) \\
&= \operatorname{Torreg}(F) + \operatorname{CMreg}(A),
\end{aligned}$$

as desired; see (2.8.3).

Combining these two cases completes the proof. □

We conclude this section with the following remark:

Remark 2.9. Let A and B be noetherian connected graded algebras with balanced dualizing complexes. Let P (resp. Q) be a nonzero object in $\mathsf{D}^b_{\mathrm{fg}}(A\text{-Gr})$ (resp. $\mathsf{D}^b_{\mathrm{fg}}(B\text{-Gr})$). Then

$$\operatorname{CMreg}({}_{(A\otimes B)}P \otimes Q) = \operatorname{CMreg}({}_A P) + \operatorname{CMreg}({}_B Q).$$

The proof of the above equality is similar to the proof of Lemma 2.3 using a version of the Künneth formula for local cohomology, so it is omitted.

3. Artin–Schelter regularities

In this section we prove results that are related to AS regular algebras. Recall from Definition 0.6 that *the AS regularity* of A is defined to be

$$\mathrm{ASreg}(A) = \mathrm{Torreg}(\Bbbk) + \mathrm{CMreg}(A).$$

In general, $\mathrm{ASreg}(A)$ can be any positive integer, as the next example shows.

Example 3.1. Let $d \geq 2$ be an integer and let $B = \Bbbk[x]/(x^d)$ with $\deg x = 1$. By Example 2.1(2), $\mathrm{CMreg}(B) = d - 1$. By an easy computation,

$$\deg \mathrm{Tor}_n^B(\Bbbk, \Bbbk) = \begin{cases} 0 & \text{if } n = 0, \\ \lfloor \frac{1}{2} n \rfloor d & \text{if } n > 0 \text{ is even}, \\ 1 + \lfloor \frac{1}{2} n \rfloor d & \text{if } n > 0 \text{ is odd}. \end{cases}$$

As a consequence, if $d > 2$, then $\mathrm{Torreg}({}_B\Bbbk) = \infty$ and $\mathrm{ASreg}(B) = \infty$. If $d = 2$, then $\mathrm{Torreg}({}_B\Bbbk) = 0$ and $\mathrm{ASreg}(B) = d - 1 = 1$.

Now let $d = 2$ so $B = \Bbbk[x]/(x^2)$ and let C be the algebra $B^{\otimes m}$ for a positive integer m. Then, by Example 2.1(2), $\mathrm{CMreg}(C) = m$. Since C is Koszul, $\mathrm{Torreg}({}_C\Bbbk) = 0$. Therefore, $\mathrm{ASreg}(C) = m$.

The goal of this section is to prove Theorem 0.8. We begin with a generalization of a nice result of Dong and Wu [2009, Theorem 4.10] that provides the first step towards to the proof of Theorem 0.8.

Theorem 3.2. *Let A be a noetherian connected graded algebra with balanced dualizing complex. Then the following are equivalent*:

(i) *A is AS regular.*

(ii) *A is Cohen–Macaulay and* $\mathrm{ASreg}(A) = 0$.

When A is Koszul, [Dong and Wu 2009, Theorem 4.10] can be recovered from Theorem 3.2 since standard AS Gorenstein algebras satisfy (ii) in the above theorem.

Proof. We first prove that (i) implies (ii). Suppose that A is AS regular of type $(d, \mathfrak{l})$. It is well known that A is Cohen–Macaulay. By (2.4.2), $\mathrm{ASreg}(A) = 0$.

We now show that (ii) implies (i). Let A be noetherian connected graded with balanced dualizing complex. If $\mathrm{projdim}\, {}_A\Bbbk < \infty$, then A has finite global dimension. Since A is noetherian, if it has finite global dimension, then it has finite GK dimension. By [Zhang 1997, Theorem 0.3], A is AS Gorenstein and so A is AS regular by definition. Hence, it suffices to show that $\mathrm{projdim}\, \Bbbk < \infty$.

Let

$$F := (\cdots \to F_i \to \cdots \to F_0 \to \Bbbk \to 0)$$

be a minimal free resolution of the trivial left A-module $\Bbbk$. Since A is Cohen–Macaulay, by [Van den Bergh 1997, Theorem 6.3], the balanced dualizing complex over A is

$$R := \mathrm{R}\Gamma_{\mathfrak{m}}(A)' \cong \omega[d],$$

where ω is a dualizing A-bimodule and $d := \operatorname{lcd}(A)$. By the local duality theorem (Theorem 1.4), for every complex X of graded left A-modules,

$$\mathrm{R}\Gamma_{\mathfrak{m}}(X)' \cong \mathrm{R}\operatorname{Hom}_A(X, R) \cong \mathrm{R}\operatorname{Hom}_A(X, \omega[d]). \tag{3.2.1}$$

Since the dualizing complex has finite injective dimension, we obtain that

$$\operatorname{injdim}(\omega) = d < \infty.$$

By the consequence of the local duality theorem (3.2.1), $\Gamma_{\mathfrak{m}}$ has cohomological dimension d.

For each $j \geq 0$, let $Z_j(F)$ denote the j-th syzygy of the complex F. We will show that $Z_j(F) = 0$ for $j \gg 0$, which implies that $\operatorname{projdim} \Bbbk < \infty$, as desired. Assume to the contrary that $Z_j(F) \neq 0$ for all $j \geq 0$. Note that

$$\cdots \to F_{j+2} \to F_{j+1} \to Z_j(F) \to 0$$

is a minimal free resolution of $Z_j(F)$.

Claim. *For all $j \geq d$, $t^A_{j+1}(\Bbbk) \leq t^A_j(\Bbbk)$.*

Proof. By the balanced condition, $\operatorname{Ext}^i_A(\Bbbk, \omega) = 0$ for all $i \neq d = \operatorname{injdim} \omega$. By induction on syzygies, we have $\operatorname{Ext}^i_A(Z_{d-1}(F), \omega) = 0$ for all $i \neq 0$. Further, by induction, one sees that $\operatorname{Ext}^i_A(Z_{j-1}(F), \omega) = 0$ for all $i \neq 0$ and all $j \geq d$. From now on, we fix $j \geq d$. By local duality (3.2.1), we obtain that $H^i_{\mathfrak{m}}(Z_{j-1}(F)) = 0$ for all $i \neq d$. Since A is Cohen–Macaulay, $H^i_{\mathfrak{m}}(F_j) = 0$ for all $i \neq d$. Applying $\mathrm{R}\Gamma_{\mathfrak{m}}(-)$ to the short exact sequence

$$0 \to Z_j(F) \to F_j \to Z_{j-1}(F) \to 0,$$

we obtain a long exact sequence, which has only three nonzero terms, yielding a short exact sequence

$$0 \to H^d_{\mathfrak{m}}(Z_j(F)) \to H^d_{\mathfrak{m}}(F_j) \to H^d_{\mathfrak{m}}(Z_{j-1}(F)) \to 0.$$

This short exact sequence implies that $\deg H^d_{\mathfrak{m}}(Z_j(F)) \leq \deg H^d_{\mathfrak{m}}(F_j)$. By definition, $\operatorname{CMreg}(Z_j(F)) \leq \operatorname{CMreg}(F_j)$. Since $\operatorname{ASreg}(A) = 0$, Corollary 2.6(3), for any $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$,

$$\operatorname{Torreg}(X) = \operatorname{CMreg}(X) + c,$$

where $c = \operatorname{Torreg}({}_A\Bbbk) = -\operatorname{CMreg}(A)$. Then

$$\begin{aligned} t^A_{j+1}(\Bbbk) &= t^A_0(F_{j+1}) = t^A_0(Z_j(F)) = t^A_0(Z_j(F)) - 0 \\ &\leq \operatorname{Torreg}(Z_j(F)) = \operatorname{CMreg}(Z_j(F)) + c \\ &\leq \operatorname{CMreg}(F_j) + c = \operatorname{Torreg}(F_j) = \sup\{t^A_i(F_j) - i \mid i \in \mathbb{Z}\} \\ &= t^A_0(F_j) - 0 = t^A_0(F_j) \\ &= t^A_j(\Bbbk), \end{aligned}$$

as desired. This finishes the proof of the claim.

Since $t_j^A(\Bbbk) \geq j$ when $F_j \neq 0$, for $j \gg 0$ the claim contradicts the fact that $t_j^A(\Bbbk) \geq j$. Therefore, we obtain a contradiction, and hence $\operatorname{projdim} \Bbbk < \infty$, as required. □

Next we work to complete the proof of Theorem 0.8, which is a generalization of [Dong and Wu 2009, Theorem 5.4]. We need the following lemma [ibid., Lemma 5.3]. Let R be a nonzero object in $\mathsf{D}^{\mathrm{b}}(A\text{-Gr})$, which will be a balanced dualizing complex over A in the proof of Theorem 0.8. Let $f : F \to R$ be a minimal free resolution of R. Since $R \in \mathsf{D}^{\mathrm{b}}(A\text{-Gr})$, each term in F is a finitely generated free graded left A-module. Set $s = \inf R$. Then f naturally induces a morphism from the truncated complex $F^{\geq s} \to R$, denoted by $\bar{f}$.

Lemma 3.3. *Retain the above notation.*

(1) [Dong and Wu 2009, Lemma 5.3] *There is a quasi-isomorphism* $g : F^{\leq s-1} \to \operatorname{cone}(\bar{f})$.

(2) $H^i(\operatorname{cone}(\bar{f})) = \begin{cases} F^{s-1}/\operatorname{im}(d_X^{s-2}) & \text{if } i = s-1, \\ 0 & \text{if } i \neq s-1. \end{cases}$

(3) $F^{\leq s-1}$ *is a minimal free resolution of* $\operatorname{cone}(\bar{f})$.

Proof. Part (1) is a consequence of [Dong and Wu 2009, Lemma 5.3] as f is a quasi-isomorphism.

Parts (2) and (3) follow immediately from part (1). □

Proof of Theorem 0.8. By (2.4.2), (i) implies (ii).

We now show that (ii) implies (i). Let $c = -\operatorname{CMreg}(A) = \operatorname{Torreg}(\Bbbk)$. By Corollary 2.6(3),

$$\operatorname{Extreg}(X) = \operatorname{Torreg}(X) = \operatorname{CMreg}(X) + c$$

for any nonzero object X in $\mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.

Let R be a balanced dualizing complex over A. Then, by [Van den Bergh 1997, Theorem 6.3], $R \cong \mathrm{R}\Gamma_{\mathfrak{m}}(A)'$. By assumption, $\operatorname{CMreg}(A) = -c$, so, for all $i \in \mathbb{Z}$, we have $H^i_{\mathfrak{m}}(A)_{>-(i+c)} = 0$. Hence, for all integers i, $H^i(R)_{<-i+c} = 0$, or $H^i(R(c))_{<-i} = 0$, and so, by [Dong and Wu 2009, Lemma 5.2], $\operatorname{Ext}^i_A(R(c), \Bbbk)_{>-i} = 0$. By a degree shift, we obtain that $\operatorname{Ext}^i_A(R, \Bbbk)_{>-i-c} = 0$ for all $i \in \mathbb{Z}$.

Since R is a balanced dualizing complex over A, by definition, $\mathrm{R}\Gamma_{\mathfrak{m}}(R) \cong A'$, which implies that $\operatorname{CMreg}(R) = 0$. Hence, by the first paragraph, $\operatorname{Extreg}(R) = \operatorname{Torreg}(R) = \operatorname{CMreg}(R) + c = c$. By definition of $\operatorname{Extreg}(R)$, for all integers i, we have $\operatorname{Ext}^i_A(R, \Bbbk)_{<-i-c} = 0$. Combining this with the last paragraph, we obtain that

$$\operatorname{Ext}^i_A(R, \Bbbk)_j \neq 0 \quad \text{if and only if} \quad j = -i - c. \tag{3.3.1}$$

Since $R \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$, R has a minimal free resolution $F \xrightarrow{\cong} R$ such that each term in F is a finitely generated free graded left A-module. Set

$$s = \inf R = \inf\{j \mid H^j(R) \neq 0\}.$$

By [Van den Bergh 1997, Theorem 6.3], $R \cong \mathrm{R}\Gamma_{\mathfrak{m}}(A)'$, which implies that

$$s = -\sup\{j \mid H^j_{\mathfrak{m}}(A) \neq 0\}.$$

Since both brutal truncations $F^{\geq s}$ and $F^{\leq s-1}$ of F are minimal free complexes, (3.3.1) implies that

$$\operatorname{Extreg}(F^{\geq s}) = \begin{cases} c & \text{if } F^{\geq s} \text{ is not acyclic,} \\ -\infty & \text{if } F^{\geq s} \text{ is acyclic,} \end{cases} \tag{3.3.2}$$

and

$$\operatorname{Extreg}(F^{\leq s-1}) = \begin{cases} c & \text{if } F^{\leq s-1} \text{ is not acyclic,} \\ -\infty & \text{if } F^{\leq s-1} \text{ is acyclic.} \end{cases} \tag{3.3.3}$$

By the choice of s, $F^s \neq 0$ and $\operatorname{Ext}_A^{-s}(F^{\geq s}, \Bbbk) = \operatorname{Hom}_A(F^s, \Bbbk) \neq 0$. Hence, $F^{\geq s}$ is not acyclic, and it follows that $\operatorname{Extreg}(F^{\geq s}) = c$ by (3.3.2). Let f be the quasi-isomorphism from $F \to R$. Then f naturally induces a morphism from $F^{\geq s}$ to R, denoted by $\bar{f}$. Hence, we are in the setting of Lemma 3.3.

We claim that $\bar{f}$ is a quasi-isomorphism. Suppose on the contrary that $\bar{f}$ is not a quasi-isomorphism. Then $Y := \operatorname{cone}(\bar{f}) \ncong 0$ in $\mathrm{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. Consider the distinguished triangle

$$F^{\geq s} \xrightarrow{\bar{f}} R \to Y (:= \operatorname{cone}(\bar{f})) \to F^{\geq s}[1].$$

Since $\mathrm{R}\Gamma_{\mathfrak{m}}(R) \cong A'$ in $\mathrm{D}(A^e\text{-Gr})$, we have the exact sequences

$$0 \to H^{-1}_{\mathfrak{m}}(Y) \to H^0_{\mathfrak{m}}(F^{\geq s}) \to H^0_{\mathfrak{m}}(R) \to H^0_{\mathfrak{m}}(Y) \to H^1_{\mathfrak{m}}(F^{\geq s}) \to 0, \tag{3.3.4}$$

and the isomorphisms

$$H^{j-1}_{\mathfrak{m}}(Y) \cong H^j_{\mathfrak{m}}(F^{\geq s}), \quad j \neq 0, 1. \tag{3.3.5}$$

By Lemma 3.3(2), $Y \cong H^{s-1}(Y)[1-s]$ in $\mathrm{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$. By the local duality theorem (Theorem 1.4), we have isomorphisms

$$\mathrm{R}\Gamma_{\mathfrak{m}}(Y)' \cong \mathrm{RHom}_A(Y, R) \cong \mathrm{RHom}_A(H^{s-1}(Y)[1-s], R) \cong \mathrm{RHom}_A(H^{s-1}(Y), R)[s-1].$$

By taking the zeroth cohomology, we obtain the isomorphism

$$H^0_{\mathfrak{m}}(Y)' \cong \operatorname{Ext}^{s-1}(H^{s-1}(Y), R).$$

Since $H^j(R) = 0$ for all $j < s$, one sees that

$$H^0_{\mathfrak{m}}(Y) \cong \operatorname{Ext}^{s-1}(H^{s-1}(Y), R)' = 0.$$

Since $\operatorname{CMreg}(F^{\geq s}) = \operatorname{Extreg}(F^{\geq s}) - c = c - c = 0$, $H^{j-1}_{\mathfrak{m}}(Y)_{>-j} = 0$ for $j \neq 1$ by (3.3.4) and (3.3.5). By definition, $\operatorname{CMreg}(Y) \leq -1$. However, Lemma 3.3(3) says that $F^{\leq s-1}$ is a minimal free resolution of Y. Hence, (3.3.3) implies that $\operatorname{Extreg}(Y) = c$, and consequently, by the first paragraph, $\operatorname{CMreg}(Y) = c - c = 0$. This yields a contradiction. Therefore, $\bar{f}$ is a quasi-isomorphism and $Y \cong 0$.

Since $F^{\leq s-1}$ is a minimal free resolution of Y, $F^j = 0$ for all $j \leq s-1$, which means that the projective dimension of R is finite. By [Dong and Wu 2009, Theorem 3.6], A is AS Gorenstein. As a consequence, A is Cohen–Macaulay. It follows from Theorem 3.2 that A is AS regular. □

Corollary 3.4. *Let A be a noetherian connected graded algebra with balanced dualizing complex. Then the following are equivalent*:

(a) *A is AS regular.*

(b) $\mathrm{Torreg}(X) = \mathrm{CMreg}(X) + \mathrm{Torreg}(\Bbbk)$ *for every nonzero* $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.

(c) $\mathrm{CMreg}(X) = \mathrm{Torreg}(X) + \mathrm{CMreg}(A)$ *for every nonzero* $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.

(d) *There is a constant c such that* $\mathrm{CMreg}(X) = \mathrm{Torreg}(X) + c$ *for every nonzero* $X \in \mathsf{D}^{\mathrm{b}}_{\mathrm{fg}}(A\text{-Gr})$.

(e) *There is a constant c such that* $\mathrm{CMreg}(M) = \mathrm{Torreg}(M) + c$ *for every nonzero finitely generated graded left A-module M.*

Proof. By Theorem 0.8 and Corollary 2.6(3), (a) implies both (b) and (c). It is clear that either (b) or (c) implies (d), and also that (d) implies (e). Hence, we need only show that (e) implies (a).

Suppose there is a constant c such that $\mathrm{CMreg}(M) = \mathrm{Torreg}(M) + c$ for every nonzero finitely generated graded left A-module M. Since $\mathrm{Torreg}(A) = 0$ (Example 2.4(2)), it follows by setting $M = A$ that $c = \mathrm{CMreg}(A)$. Since $\mathrm{CMreg}(\Bbbk) = 0$ (Example 2.1(1)), setting $M = \Bbbk$ implies that $c = -\mathrm{Torreg}(\Bbbk)$. Therefore,

$$\mathrm{ASreg}(A) = \mathrm{CMreg}(A) + \mathrm{Torreg}(\Bbbk) = c - c = 0.$$

Hence, by Theorem 0.8, A is AS regular. □

We remark that Corollary 3.4 is a generalization of [Römer 2008, Theorem 4.1].

4. Concavity

In this section we use the letters A and B for connected graded noetherian algebras, S and T for connected graded noetherian AS regular algebras, and F and G for general locally finite $\mathbb{N}$-graded noetherian algebras.

For a locally finite $\mathbb{N}$-graded noetherian algebra F, let

$$\Phi(F) := \{T \mid \text{there is a finite map } \phi : T \to F\}, \tag{4.0.1}$$

where T ranges over all connected graded noetherian AS regular algebras.

Now we recall Definition 0.9. Let $\mathcal{P}$ be any numerical invariant that is defined for locally finite $\mathbb{N}$-graded noetherian rings (or a subclass of such algebras).

(1) The *$\mathcal{P}$-concavity* of F is defined to be

$$c_{\mathcal{P}}(F) := \inf_{T \in \Phi(F)} \{\mathcal{P}(T)\}.$$

If no such T exists, we write $c_{\mathcal{P}}(F) = \infty$. The *normalized $\mathcal{P}$-concavity* of F is defined to be

$$c_{\mathcal{P},-}(F) := c_{\mathcal{P}}(F) - \mathcal{P}(F).$$

(2) The *concavity* of F is defined to

$$c(F) := \inf_{T \in \Phi(F)} \{-\mathrm{CMreg}(T)\}$$

and the *normalized concavity* of F is defined to be

$$c_-(F) := c(F) + \mathrm{CMreg}(F).$$

Proposition 4.1. *Let $f : F \to G$ be a finite map of locally finite graded noetherian algebras. Then $c_{\mathcal{P}}(F) \geq c_{\mathcal{P}}(G)$. As a consequence, the following hold*:

(1) *If F is a subalgebra of G such that ${}_FG$ and G_F are finitely generated, then $c_{\mathcal{P}}(F) \geq c_{\mathcal{P}}(G)$.*

(2) *If H is a finite dimensional semisimple Hopf algebra acting on F homogeneously, then $c_{\mathcal{P}}(F^H) \geq c_{\mathcal{P}}(F) \geq c_{\mathcal{P}}(F \# H)$.*

(3) $c_{\mathcal{P}}(F) \geq c_{\mathcal{P}}(M_n(F))$.

(4) *Let t be a commutative indeterminate of degree 1. Then $c(F[t]) = c(F)$.*

Proof. By definition, $c_{\mathcal{P}}(F) = \inf_{T \in \Phi(F)}\{\mathcal{P}(T)\}$. If $T \in \Phi(F)$, then, by definition, there is a finite map $\phi : T \to F$. But, since $f : F \to G$ is a finite map, the composition $f \circ \phi : T \to G$ is a finite map. Hence, $T \in \Phi(G)$, so $\Phi(F) \subseteq \Phi(G)$, which implies that $c_{\mathcal{P}}(F) \geq c_{\mathcal{P}}(G)$.

Parts (1)–(3) are immediate consequences of the main assertion.

For part (4), note that there is a finite map $F[t] \to F$ given by sending t to 0. Hence, taking $\mathcal{P} = -\mathrm{CMreg}$ in the main assertion, we have $c(F[t]) \geq c(F)$.

Fix a real number $\epsilon > 0$. By definition of $c(F)$, there is a noetherian AS regular algebra T of type $(d, \mathfrak{l})$ and a finite map $\phi : T \to F$ such that $-\mathrm{CMreg}(T) \leq c(F) + \epsilon$. Then $T[t] \to F[t]$ is a finite map. Hence,

$$\begin{aligned} c(F[t]) &\leq -\mathrm{CMreg}(T[t]) = -((d+1) - (\mathfrak{l} + \deg t)) \\ &= -(d - \mathfrak{l}) - (1 - \deg t) = -(d - \mathfrak{l}) \\ &= -\mathrm{CMreg}(T) \leq c(F) + \epsilon. \end{aligned}$$

Since ϵ was arbitrary, we obtain that $c(F[t]) \leq c(F)$. Combined with the previous paragraph, we conclude that $c(F[t]) = c(F)$. □

We do not have any examples with strict inequalities $c_{\mathcal{P}}(F) > c_{\mathcal{P}}(F \# H)$ and $c_{\mathcal{P}}(F) > c_{\mathcal{P}}(M_n(F))$. Proposition 0.14 can be used to provide many examples with $c_{\mathcal{P}}(F^H) > c_{\mathcal{P}}(F)$. We now prove Proposition 0.14. For the rest of the paper we consider only connected graded noetherian algebras.

Proof of Proposition 0.14. Let H be a semisimple Hopf algebra acting on a noetherian AS regular algebra T homogeneously and let $R = T^H$ denote the invariant subring.

(1) By [Kirkman et al. 2022, Theorem 0.8(1)], if $S \in \Phi(R)$, then

$$\beta_1(R) \leq \max\{\beta_1(S), \mathrm{CMreg}(T) - \mathrm{CMreg}(S)\}.$$

It is clear that $\mathrm{CMreg}(T) - \mathrm{CMreg}(S) \leq -\mathrm{CMreg}(S) < -\mathrm{CMreg}(S) + 1$. By definition and Theorem 0.7,

$$\beta_1(S) = t_1^S(\Bbbk) \leq \mathrm{Torreg}({}_S\Bbbk) + 1 = \mathrm{CMreg}({}_S\Bbbk) - \mathrm{CMreg}(S) + 1 = -\mathrm{CMreg}(S) + 1,$$

which implies that $\beta_1(R) \leq -\mathrm{CMreg}(S)+1$, or, equivalently, $-\mathrm{CMreg}(S) \geq \beta_1(R)-1$. Since $S \in \Phi(R)$ was arbitrary, we obtain that $c(R) \geq \beta_1(R)-1$.

(2) By [Kirkman et al. 2022, Theorem 0.8(2)], if $S \in \Phi(R)$, then

$$\beta_2(R) \leq \max\{2(\mathrm{CMreg}(T)-\mathrm{CMreg}(S)), \mathrm{CMreg}(T)-\mathrm{CMreg}(S)+\beta_1(S), \beta_2(S)\}.$$

If $\beta_2(R) \leq 2(\mathrm{CMreg}(T)-\mathrm{CMreg}(S))$, then $-\mathrm{CMreg}(S) \geq \frac{1}{2}\beta_2(R)-\mathrm{CMreg}(T)$. Since $S \in \Phi(R)$ was arbitrary, we obtain that

$$c(R) \geq \tfrac{1}{2}\beta_2(R)-\mathrm{CMreg}(T).$$

If $\beta_2(R) \leq \mathrm{CMreg}(T)-\mathrm{CMreg}(S)+\beta_1(S)$, then $\beta_2(R) \leq \mathrm{CMreg}(T)-\mathrm{CMreg}(S)+(-\mathrm{CMreg}(S)+1)$, which implies that

$$c(R) \geq \tfrac{1}{2}(\beta_2(R)-\mathrm{CMreg}(T)-1).$$

Finally, if $\beta_2(R) \leq \beta_2(S)$, then

$$\beta_2(R) \leq \beta_2(S) \leq \mathrm{Torreg}(\Bbbk)+2 \leq \mathrm{CMreg}(\Bbbk)-\mathrm{CMreg}(S)+2 = -\mathrm{CMreg}(S)+2,$$

which implies that $c(R) \geq \beta_2(R)-2$.

Combining these three cases, we obtain that

$$c(R) \geq \min\left\{\tfrac{1}{2}\beta_2(R)-\mathrm{CMreg}(T), \tfrac{1}{2}(\beta_2(R)-\mathrm{CMreg}(T)-1), \beta_2(R)-2\right\},$$

as desired. □

The inequalities in Proposition 0.14 can be used to test if an algebra is isomorphic to the invariant subring of a semisimple Hopf action on an AS regular algebra.

Example 4.2. Let A be a connected graded algebra that is not generated in degree 1 (or, equivalently, $\beta_1(A) \geq 2$). Suppose there is a noetherian Koszul AS regular algebra $S \in \Phi(A)$. Then $c(A)=0$ by definition, and so, by Proposition 0.14(1), A cannot be isomorphic to an invariant subring T^H.

A specific example is the graded algebra

$$A = \Bbbk[x_1, \ldots, x_n, t_1, \ldots, t_m]/(t_j^2 - f_j(x_1, \ldots, x_n) \mid j = 1, \ldots, m),$$

where $\deg x_i = 1$ for all $i = 1, \ldots, n$, $\deg t_{j_0} > 1$ for some j_0, and each f_j is a homogeneous polynomial in the x_i of degree equal to $2 \deg t_j$. Then $c(A)=0$, as the polynomial ring $\Bbbk[x_1, \ldots, x_n]$ is in $\Phi(A)$. It is clear that $\beta_1(A) = \max_{1 \leq j \leq m}\{\deg t_j\} > 1$.

Our next aim is to prove Theorems 0.10 and 0.12. We begin by proving the following lemma. As usual, let T denote a noetherian AS regular algebra.

Lemma 4.3. *Let A be a connected graded noetherian algebra, let $\phi : T \to A$ be a finite map, and let M be a finitely generated graded left A-module. For parts* (2)–(5)*, we further assume that ${}_TA$ is linear.*

(1) $\mathrm{Torreg}({}_TM) \leq \mathrm{Torreg}({}_TA) + \mathrm{Torreg}({}_AM)$.

(2) *For all* $s \geq 0$, *we have* $t_s^T(\Bbbk) - s \leq \max_{0\leq i\leq s}\{t_i^A(\Bbbk) - i\}$.

(3) *For all* $0 \leq s \leq \operatorname{gldim} T$, *we have* $\max_{0\leq i\leq s}\{t_i^A(\Bbbk) - i\} = \max_{0\leq i\leq s}\{t_i^T(\Bbbk) - i\}$.

(4) *For all* $j > \operatorname{gldim} T$, *we have*

$$t_j^A(\Bbbk) - j \leq \operatorname{Torreg}({}_T\Bbbk) = \max_{0\leq i\leq \operatorname{gldim} T}\{t_i^T(\Bbbk) - i\}.$$

(5) $\operatorname{Torreg}({}_A\Bbbk) = \operatorname{Torreg}({}_T\Bbbk)$. *As a consequence*, A *is Koszul if and only if* T *is Koszul.*

Proof. We will use the change of rings spectral sequence given in [Rotman 1979, Theorem 10.60], namely

$$E^2_{p,q} := \operatorname{Tor}^A_p(\operatorname{Tor}^T_q(\Bbbk_T, A), {}_AM) \Rightarrow \operatorname{Tor}^T_{p+q}(\Bbbk_T, {}_TM). \tag{4.3.1}$$

The E^2-page of the spectral sequence is similar to the one given after [Kirkman et al. 2022, Lemma 5.1].

(1) We have

$$\begin{aligned}
\operatorname{Torreg}({}_TM) &= \operatorname{CMreg}({}_TM) + \operatorname{Torreg}({}_T\Bbbk) && \text{(Corollary 2.6)}\\
&= \operatorname{CMreg}({}_TM) - \operatorname{CMreg}(T) && \text{(Example 2.4(3))}\\
&= \operatorname{CMreg}({}_TM) + \operatorname{Torreg}({}_TA) - \operatorname{CMreg}({}_TA) && \text{(Corollary 3.4(c))}\\
&= \operatorname{CMreg}({}_TM) + \operatorname{Torreg}({}_TA) - \operatorname{CMreg}(A)\\
&= \operatorname{Torreg}({}_TA) + \operatorname{CMreg}({}_AM) - \operatorname{CMreg}(A)\\
&\leq \operatorname{Torreg}({}_TA) + \operatorname{Torreg}({}_AM) && \text{(Theorem 2.5(2))}.
\end{aligned}$$

(2) For the remainder of the proof, we assume that ${}_TA$ is linear; that is, we assume for all q that there exists some $n_q \geq 0$ such that $\operatorname{Tor}^T_q(\Bbbk_T, A) = \Bbbk(-q)^{n_q}$. We claim that the map $f : T \to A$ is surjective. Since ${}_TA$ is linear, $\Bbbk \otimes_T A = \Bbbk^{\oplus n}$ for some positive integer n. So ${}_TA$ is generated by elements of degree 0. But A has only one element in degree 0 (up to a scalar). So ${}_TA$ is generated by a single element of degree 0. This implies that the map $\phi : T \to A$ is surjective. Note that $n_q = 0$ for all $q > \operatorname{projdim}{}_TA$. By taking $M = \Bbbk$ in (4.3.1), we have

$$\begin{aligned}
t_n^T(\Bbbk) &= \deg \operatorname{Tor}^T_n(\Bbbk_T, {}_T\Bbbk)\\
&\leq \max_{p+q=n}\{\deg E^2_{p,q}\} = \max_{p+q=n}\{\deg \operatorname{Tor}^A_p(\operatorname{Tor}^T_q(\Bbbk_T, A), {}_A\Bbbk)\}\\
&\leq \max_{p+q=n}\{\deg \operatorname{Tor}^A_p(\Bbbk_A, {}_A\Bbbk) + q\} = \max_{p+q=n}\{t_p^A(\Bbbk) + q\}.
\end{aligned} \tag{4.3.2}$$

Therefore, the assertion in (2) follows.

(3) We use induction on s from 0 to $g := \operatorname{gldim} T$. Since $t_0^T(\Bbbk) = t_0^A(\Bbbk) = 0$, the assertion holds for $s = 0$. Now assume that $s > 0$ and assume the assertion holds for $s - 1$. Part (2) above implies that $\max_{0\leq i\leq s}\{t_i^A(\Bbbk) - i\} \geq \max_{0\leq i\leq s}\{t_i^T(\Bbbk) - i\}$. Hence, we need only show that $\max_{0\leq i\leq s}\{t_i^A(\Bbbk) - i\} \leq \max_{0\leq i\leq s}\{t_i^T(\Bbbk) - i\}$. By the induction hypothesis, it suffices to show that

$$t_s^A(\Bbbk) - s \leq \max_{0\leq i\leq s}\{t_i^T(\Bbbk) - i\}. \tag{4.3.3}$$

By [Stephenson and Zhang 1997, (3-4), p. 1600], $\{t_0^T(\Bbbk), \dots, t_g^T(\Bbbk)\}$ is strictly increasing, or, equivalently,

$$\{t_0^T(\Bbbk), t_1^T(\Bbbk) - 1, \dots, t_{g-1}^T(\Bbbk) - (g-1), t_g^T(\Bbbk) - g\} \quad \text{is nondecreasing.} \tag{4.3.4}$$

So (4.3.3) is equivalent to

$$t_s^A(\Bbbk) - s \le t_s^T(\Bbbk) - s, \quad \text{or} \quad t_s^A(\Bbbk) \le t_s^T(\Bbbk). \tag{4.3.5}$$

Let $\alpha = t_s^T(\Bbbk)$ and $\beta = t_s^A(\Bbbk)$. Using the spectral sequence (4.3.1) (and an inequality in (4.3.2)), we see that, for each $p < s$,

$$\begin{aligned}\deg E^\infty_{p,s-p} \le \deg E^2_{p,s-p} &\le \max_{p+q=s, p<s}\{t_p^A(\Bbbk) + q\} \\ &= \max_{0\le p\le s-1}\{t_p^A(\Bbbk) - p\} + s = \max_{0\le i\le s-1}\{t_i^T(\Bbbk) - i\} + s \\ &\le t_s^T(\Bbbk) - s + s = t_s^T(\Bbbk) = \alpha.\end{aligned}$$

Hence, if $\beta > \alpha$, then $(E^\infty_{p,s-p})_\beta = 0$ for all $p < s$. Similarly, for all $r \ge 2$ and for all $p + q < s$, we can show $\deg E^r_{p,q} \le t^T_{p+q}(\Bbbk) \le \alpha$. Observe that, if $r \ge 2$, the incoming differentials to $E^r_{s,0}$ are all zero, and the outgoing differentials map to $E^r_{s-r,r-1}$ with $\deg E^r_{s-r,r-1} \le \alpha$. Hence, $0 \ne (E^2_{s,0})_\beta$ is in the kernel of all outgoing differentials and so $(E^2_{s,0})_\beta$ survives on the ∞-page. Thus, by (4.3.1), $\operatorname{Tor}^T_s(\Bbbk, \Bbbk)_\beta \ne 0$, which contradicts that $\deg \operatorname{Tor}^T_s(\Bbbk, \Bbbk) = \alpha < \beta$. Therefore, $\beta \le \alpha$, which is (4.3.5). This completes the inductive step and the proof.

(4) We prove this by induction on $j > \operatorname{gldim} T$. The initial step and the inductive step are similar, so we treat them together. Let $\gamma = \operatorname{Torreg}({}_T\Bbbk)$. Similar to the proof of part (3), one sees that, for every $r \ge 2$,

$$\deg E^r_{j-r,r-1} \le \max_{p\le j-1}\{t_p^A(\Bbbk) - p + (j-1)\} \le \gamma + (j-1),$$

where the second inequality is the inductive step when $j > \operatorname{gldim} T + 1$, and is part (3) when $j = \operatorname{gldim} T + 1$. Since $E^\infty_{j,0} = 0$ as $j > \operatorname{gldim} T$, the "outgoing differential" argument in the proof of part (3) shows that

$$\deg E^2_{j,0} \le \max_{r\ge 2}\{\deg E^r_{j-r,r-1}\} \le \gamma + j - 1.$$

This is equivalent to

$$t_j^A(\Bbbk) \le \gamma + j - 1 < \gamma + j,$$

and therefore the assertion holds.

(5) The equation follows from parts (3) and (4). The consequence follows from the fact that A is Koszul if and only if $\operatorname{Torreg}({}_A\Bbbk) = 0$. □

Now we are ready to prove Theorems 0.10 and 0.12.

Proof of Theorem 0.10. Fix a noetherian AS regular algebra S. We first remark that, by [Artin and Zhang 1994, Theorem 8.3(3)], if there is a finite map $T \to S$, then, for any finitely generated graded S-module M, we have $\operatorname{CMreg}({}_T M) = \operatorname{CMreg}({}_S M)$. In particular, $\operatorname{CMreg}({}_T S) = \operatorname{CMreg}({}_S S)$.

(1) Suppose T is any noetherian AS regular algebra and suppose $T \to S$ is a finite map. By Theorem 0.7 and the fact that $\mathrm{Torreg}({}_T S) \geq 0$, we obtain that

$$\mathrm{CMreg}(T) = \mathrm{CMreg}({}_T S) - \mathrm{Torreg}({}_T S) \leq \mathrm{CMreg}(S),$$

and hence $-\mathrm{CMreg}(T) \geq -\mathrm{CMreg}(S)$. Therefore, $c(S) \geq -\mathrm{CMreg}(S)$. By definition, it is clear that $c(S) \leq -\mathrm{CMreg}(S)$, and so we have equality, as desired.

(2) If S is Koszul, then $c(S) = -\mathrm{CMreg}(S) = \mathrm{Torreg}({}_S\Bbbk) = 0$. Conversely, if $c(S) = 0$, then $\mathrm{Torreg}({}_S\Bbbk) = -\mathrm{CMreg}(S) = c(S) = 0$, which implies that S is Koszul.

(3) Suppose T is a noetherian AS regular algebra and $f : T \to S$ is a finite map. By Theorem 0.7, we have

$$\mathrm{CMreg}(S)(= \mathrm{CMreg}({}_S S)) = \mathrm{CMreg}({}_T S) = \mathrm{Torreg}({}_T S) + \mathrm{CMreg}(T).$$

This implies that $-\mathrm{CMreg}(T) \geq -\mathrm{CMreg}(S)$, as $\mathrm{Torreg}({}_T S) \geq 0$ by definition. The first assertion now follows from part (1).

If $c(T) = c(S)$, then, by part (1), we have $\mathrm{CMreg}(T) = \mathrm{CMreg}(S) = \mathrm{CMreg}({}_T S)$. Hence, by Theorem 0.7, $\mathrm{Torreg}({}_T S) = 0$, which means that ${}_T S$ is linear. Similarly, S_T is linear. By the proof of Lemma 4.3(2), the map $f : T \to S$ is surjective.

Since the local cohomology of S can be computed as a left S-module or as a left T-module, $H^i_{\mathfrak{m}}({}_T S)$ is zero unless $i = \mathrm{gldim}\, S$. Hence, ${}_T S$ is Cohen–Macaulay. □

Proof of Theorem 0.12. (1) We have

$$\begin{aligned} c_-(A) &= c(A) + \mathrm{CMreg}(A) = \inf_{T \in \Phi(A)} \{-\mathrm{CMreg}(T) + \mathrm{CMreg}({}_T A)\} \\ &= \inf_{T \in \phi(A)} \{\mathrm{Torreg}({}_T A)\} \geq 0. \end{aligned}$$

(2) If A is AS regular, by letting $T = A$, we obtain that $c(A) \leq -\mathrm{CMreg}(A)$. As a consequence, $c_-(A) \leq 0$. By part (1), $c_-(A) = 0$.

Conversely, suppose $c_-(A) = 0$, or, equivalently, $c(A) = -\mathrm{CMreg}(A)$. By definition and the fact that $c(A)$ is an integer, there is a noetherian AS regular algebra T with a finite map $\phi : T \to A$ such that $-\mathrm{CMreg}(T) = c(A) = -\mathrm{CMreg}(A)$. By Theorem 0.7,

$$\mathrm{Torreg}({}_T A) = \mathrm{CMreg}(A) - \mathrm{CMreg}(T) = 0.$$

Then ${}_T A$ is a linear left T-module (i.e., $t_i^T({}_T A) = i$ if $\mathrm{Tor}_i^T(\Bbbk, {}_T A) \neq 0$).

By Lemma 4.3(5), $\mathrm{Torreg}({}_A\Bbbk) = \mathrm{Torreg}({}_T\Bbbk)$. We know that $\mathrm{CMreg}(A) = \mathrm{CMreg}(T)$ and, since T is AS regular, we have $\mathrm{CMreg}(T) = -\mathrm{Torreg}({}_T\Bbbk)$. Combining these equations, we obtain that

$$\mathrm{ASreg}(A) = \mathrm{CMreg}(A) + \mathrm{Torreg}({}_A\Bbbk) = \mathrm{CMreg}(T) + \mathrm{Torreg}({}_T\Bbbk) = \mathrm{ASreg}(T) = 0.$$

Finally, the assertion follows from Theorem 0.8. □

We collect some criteria for Koszulness of AS regular algebras that follow from our previous results.

Corollary 4.4. *Let T and S be noetherian AS regular algebras and $f : T \to S$ be a finite map.*

(1) *If T is Koszul (or, equivalently,* $\operatorname{Torreg}({}_T\Bbbk) = 0$*), then so is S. Further, f is surjective, and ${}_TS$ and S_T are linear Cohen–Macaulay modules over T.*

(2) *Suppose* $\operatorname{Torreg}({}_T\Bbbk) \leq \deg(\Bbbk \otimes_T S)$. *Then S is Koszul.*

(3) *Suppose* $\operatorname{Torreg}({}_T\Bbbk) = 1$. *If f is not surjective, then S is Koszul. As a consequence, if T is a proper subalgebra of S, then S is Koszul.*

Proof. (1) Suppose T is a Koszul noetherian AS regular algebra and suppose $f : T \to S$ is a finite map. By Theorem 0.10(3) and the fact that T is Koszul, we have

$$0 = -\operatorname{CMreg}(T) \geq -\operatorname{CMreg}(S) \geq 0.$$

Hence, $\operatorname{CMreg}(S) = \operatorname{CMreg}(T) = 0$. The result follows from parts (2) and (3) of Theorem 0.10 (the Koszulness of S also follows from Lemma 4.3(5)).

(2) Suppose S and T are AS regular. By Theorem 0.7, we have

$$\operatorname{CMreg}(S) = \operatorname{CMreg}({}_TS) = \operatorname{Torreg}({}_TS) + \operatorname{CMreg}(T).$$

By Theorem 0.8, $\operatorname{CMreg}(S) = -\operatorname{Torreg}({}_S\Bbbk)$ and $\operatorname{CMreg}(T) = -\operatorname{Torreg}({}_T\Bbbk)$, so

$$\begin{aligned} 0 \leq \operatorname{Torreg}({}_S\Bbbk) &= -\operatorname{CMreg}(S) = -\operatorname{CMreg}(T) - \operatorname{Torreg}({}_TS) \\ &= \operatorname{Torreg}({}_T\Bbbk) - \operatorname{Torreg}({}_TS) \leq \operatorname{Torreg}({}_T\Bbbk) - \deg(\Bbbk \otimes_T S). \end{aligned}$$

Hence, if $\operatorname{Torreg}({}_T\Bbbk) \leq \deg(\Bbbk \otimes_T S)$, then $\operatorname{Torreg}({}_S\Bbbk) = 0$, whence S is Koszul.

(3) If f is not surjective, then $\deg(\Bbbk \otimes_T S) \geq 1$ and so, by part (2), S is Koszul. □

Theorem 4.5. *Let A be a noetherian connected graded algebra. Suppose that there is a finite map $f : T \to A$, where T is a noetherian Koszul AS regular algebra. Then the following hold*:

(1) $\operatorname{CMreg}(A) \geq 0$ *and* $\operatorname{CMreg}(A) = 0$ *if and only if A is AS regular (and Koszul).*

(2) *If A is commutative and generated in degree* 1, *then* $\operatorname{CMreg}(A) \geq 0$, *and A is a polynomial ring if and only if* $\operatorname{CMreg}(A) = 0$.

Part (2) is an improvement of [Römer 2008, Theorem 4.1(iv)$\Longleftrightarrow$(v)].

Proof of Theorem 4.5. (1) Since the map $\phi : T \to A$ is finite and T is Koszul, by Theorem 0.7,

$$\operatorname{CMreg}(A) = \operatorname{CMreg}({}_TA) = \operatorname{Torreg}({}_TA) + \operatorname{CMreg}(T) = \operatorname{Torreg}({}_TA) \geq 0.$$

Since $T \in \Phi(A)$ and T is Koszul, by definition,

$$0 \leq c(A) \leq -\operatorname{CMreg}(T) = 0,$$

which implies that $c(A) = 0$. Consequently, by the definition of $c_-(A)$, we have $c_-(A) = \operatorname{CMreg}(A)$. Now, by Theorem 0.12(2), A is AS regular if and only if $\operatorname{CMreg}(A) = 0$. If this happens, $\operatorname{CMreg}(A) = 0$ implies that $\operatorname{Torreg}({}_A\Bbbk) = 0$ by Theorem 0.8, and hence A is Koszul.

(2) Let $T = \Bbbk[A_1]$. Then $A = T/J$ for some ideal J. Since T is noetherian, Koszul and AS regular, the assertion follows from part (1). □

Note that part (2) of the above theorem fails when A is not commutative (Example 5.4(ii)). Now we are ready to prove Theorem 0.3. Recall that the Hilbert series of a graded A-module M, denoted by $h_M(t)$, is given in Definition 1.1. If $h_M(t)$ is a rational function, then the *a-invariant of* M, denoted by $a(M)$, is defined to be the t-degree of the rational function $h_M(t)$.

Theorem 4.6. *Let A be a noetherian connected graded s-Cohen–Macaulay algebra. Suppose that there is a finite map $f : T \to A$, where T is a noetherian Koszul AS regular algebra. Then the following are equivalent*:

(a) *A is AS regular.*

(b) *A is AS regular and Koszul.*

(c) $\deg h_A(t) = -s$.

Proof. Since there is a finite map to A from an AS regular algebra T, by taking the minimal free resolution of the graded T-module A, it follows that the Hilbert series $h_A(t)$ is rational. Then, by the proof of [Kirkman et al. 2022, Theorem 4.7(2)], we have

$$\mathrm{CMreg}(A) = s + a(A) = s + \deg h_A(t).$$

Now the assertion follows from Theorem 4.5(1). □

The following simple corollary is useful in practice:

Corollary 4.7. *Let A and B be algebras satisfying Hypothesis 1.3. Assume that there is a finite map $f : A \to B$ such that ${}_AB$ is linear of finite projective dimension. Then A is AS regular if and only if B is AS regular.*

Proof. Suppose B is AS regular. By the change of rings spectral sequence given in [Rotman 1979, Theorem 10.60] (or see (4.3.1)), we have

$$\mathrm{gldim}\, A = \mathrm{projdim}\, {}_A\Bbbk \leq \mathrm{projdim}\, {}_B\Bbbk + \mathrm{projdim}\, {}_AB < \infty.$$

By [Zhang 1997, Theorem 0.3], A is AS regular.

Now we assume A is AS regular. By Lemma 4.3(5), $\mathrm{Torreg}({}_A\Bbbk) = \mathrm{Torreg}({}_B\Bbbk)$. By Theorem 0.7,

$$\mathrm{CMreg}(B) = \mathrm{CMreg}({}_AB) = \mathrm{Torreg}({}_AB) + \mathrm{CMreg}(A) = \mathrm{CMreg}(A).$$

Then

$$\mathrm{ASreg}(B) = \mathrm{Torreg}({}_B\Bbbk) + \mathrm{CMreg}(B) = \mathrm{Torreg}({}_A\Bbbk) + \mathrm{CMreg}(A) = \mathrm{ASreg}(A) = 0.$$

The assertion follows from Theorem 0.8. □

5. Examples and remarks

This section contains some examples and remarks. To save space we will omit some nonessential details. Our first example shows that it is necessary to assume Hypothesis 1.3 for most of the results in this paper.

Example 5.1. We refer to [Artin and Zhang 1994; Stafford and Zhang 1994] for the definition of the χ-*condition*. Let A be the noetherian connected graded domain of GK-dimension 2 given in [Stafford and Zhang 1994, Theorem 2.3], which does not satisfy the χ-condition. By [Van den Bergh 1997, Theorem 6.3], A does not admit a balanced dualizing complex. Further, by [Stafford and Zhang 1994, Theorem 2.3], $\mathrm{Ext}^1_A(\Bbbk, A)$ is not bounded above. Since $\mathrm{Hom}_A(\Bbbk, A) = 0$, $H^1_{\mathfrak{m}}(A)$ contains $\mathrm{Ext}^1_A(\Bbbk, A)$ as a graded $\Bbbk$-subspace. Therefore, $\deg H^1_{\mathfrak{m}}(A) = +\infty$, and, consequently, $\mathrm{CMreg}(A) = +\infty$.

If there is a finite map $T \to A$ for some noetherian AS regular algebra T, then [Artin and Zhang 1994, Theorem 8.1(1) and Lemma 8.2(4)] implies that A satisfies the χ-condition. Since A does not satisfy the χ-condition, we conclude that there does not exist a finite map from a noetherian AS regular algebra to A. As a consequence, $c(A) = +\infty$ by definition.

Rogalski and Sierra [2012] provide a family of examples of noetherian Koszul algebras of global dimension 4 for which the χ-condition fails. By similar arguments, these algebras have infinite CM regularity and concavity.

Before we give more examples, we prove some well-known lemmas.

Lemma 5.2. *Let A be a noetherian connected graded algebra and let Ω be a regular normal element of degree* 1 *or* 2. *Let $B = A/(\Omega)$. Then* $\mathrm{Torreg}({}_B\Bbbk) \leq \mathrm{Torreg}({}_A\Bbbk)$. *Consequently, the following hold*:

(1) *If A has finite global dimension, then* $\mathrm{Torreg}({}_B\Bbbk) < \infty$.

(2) [Polishchuk 2006, Theorem 1.2] *If A is Koszul, then so is B.*

(3) *Suppose* $\deg \Omega = 1$. *If A has finite global dimension, then so does B.*

Proof. The first part of the proof is copied from the proof of [Kirkman et al. 2015, Theorem 1.11]. Since B is a factor ring of A, there is a graded version of the change of rings spectral sequence given in [Rotman 1979, Theorem 10.71],

$$E^2_{p,q} := \mathrm{Tor}^B_p(\Bbbk, \mathrm{Tor}^A_q(B, \Bbbk)) \Rightarrow \mathrm{Tor}^A_n(\Bbbk, \Bbbk). \tag{5.2.1}$$

Let $a = \deg \Omega$. Since $B = A/(\Omega)$ and Ω is a regular normal element, we have $\mathrm{Tor}^A_0(B, \Bbbk) = \Bbbk$, $\mathrm{Tor}^A_1(B, \Bbbk) = \Bbbk(-a)$ and $\mathrm{Tor}^A_i(B, \Bbbk) = 0$ for $i > 1$. Hence, the E^2-page of the spectral sequence (5.2.1) has only two possibly nonzero rows, namely

$$\begin{aligned} q = 0: &\quad \mathrm{Tor}^B_p(\Bbbk, \Bbbk) &&\text{for } p = 0, 1, 2, \ldots, \\ q = 1: &\quad \mathrm{Tor}^B_p(\Bbbk, \Bbbk(-a)) &&\text{for } p = 0, 1, 2, \ldots. \end{aligned}$$

Since (5.2.1) converges, we have $\mathrm{Tor}_0^B(\Bbbk,\Bbbk)=\mathrm{Tor}_0^A(\Bbbk,\Bbbk)=\Bbbk$ and a long exact sequence

$$\begin{aligned}\cdots\to \mathrm{Tor}_4^B(\Bbbk,\Bbbk)&\to \mathrm{Tor}_2^B(\Bbbk,\Bbbk)(-a)\\ \to \mathrm{Tor}_3^A(\Bbbk,\Bbbk)&\to \mathrm{Tor}_3^B(\Bbbk,\Bbbk)\to \mathrm{Tor}_1^B(\Bbbk,\Bbbk)(-a)\\ \to \mathrm{Tor}_2^A(\Bbbk,\Bbbk)&\to \mathrm{Tor}_2^B(\Bbbk,\Bbbk)\to \mathrm{Tor}_0^B(\Bbbk,\Bbbk)(-a)\\ \to \mathrm{Tor}_1^A(\Bbbk,\Bbbk)&\to \mathrm{Tor}_1^B(\Bbbk,\Bbbk)\to 0.\end{aligned}$$

By the k-th row from the bottom in the above exact sequence, we have

$$\begin{aligned}\deg(\mathrm{Tor}_k^B(\Bbbk,\Bbbk))-k&\le \max\{\deg(\mathrm{Tor}_k^A(\Bbbk,\Bbbk))-k,\deg(\mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk)(-a))-k\}\\ &=\max\{\deg(\mathrm{Tor}_k^A(\Bbbk,\Bbbk))-k,\deg(\mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk))-(k-2)-(2-a)\}\\ &\le \max\{\mathrm{Torreg}({}_A\Bbbk),\mathrm{Torreg}({}_A\Bbbk)-(2-a)\}\\ &\le \mathrm{Torreg}({}_A\Bbbk),\end{aligned}$$

where we use $\mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk)-(k-2)\le\mathrm{Torreg}({}_A\Bbbk)$ by the induction hypothesis and $2-a\ge 0$ as $a=1$ or 2. The assertion follows from the definition.

Parts (1) and (2) of this lemma follow immediately from the main assertion.

(3) For each $k>\mathrm{gldim}\,A+1$, we have an exact sequence

$$(\mathrm{Tor}_k^A(\Bbbk,\Bbbk)=)\ \ 0\to \mathrm{Tor}_k^B(\Bbbk,\Bbbk)\to \mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk)(-1)\to 0\ \ (=\mathrm{Tor}_{k-1}^A(\Bbbk,\Bbbk)).$$

So $\mathrm{ged}\,\mathrm{Tor}_k^B(\Bbbk,\Bbbk)=\mathrm{ged}\,\mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk)+1$.

If $\mathrm{Tor}_k^B(\Bbbk,\Bbbk)\ne 0$, B being connected graded implies that $\mathrm{ged}\,\mathrm{Tor}_k^B(\Bbbk,\Bbbk)\ge\mathrm{ged}\,\mathrm{Tor}_{k-2}^B(\Bbbk,\Bbbk)+2$ by using the minimal free resolution of $\Bbbk$. This yields a contradiction, and thus $\mathrm{Tor}_k^B(\Bbbk,\Bbbk)=0$, or $\mathrm{gldim}\,B<k<\infty$. □

See Lemma 4.3 and Corollary 4.7 for related results. Note that the inequality $\mathrm{Torreg}({}_B\Bbbk)\le\mathrm{Torreg}({}_A\Bbbk)$ in Lemma 5.2 can be strict; see Example 5.4(iv). Part (3) of Lemma 5.2 is a special case of Corollary 4.7 if A satisfies Hypothesis 1.3.

Lemma 5.3. *Let A be a noetherian connected graded AS Gorenstein algebra of type (d,ℓ). Let Ω be a regular normal element of degree a.*

(1) *If $a\ge\ell-d+2$, then $A/(\Omega)$ is not AS regular.*

(2) *Suppose A is generated in degree 1 and is not Koszul. If $a\ge\max\{3,\ell-d+1\}$, then $A/(\Omega)$ is not AS regular.*

Proof. By the Rees lemma, $B:=A/(\Omega)$ is also AS Gorenstein (of injective dimension $d-1$). Applying $H^i_{\mathfrak{m}}(-)$ to the short exact sequence of left A-modules

$$0\to A(-a)\to A\to B\to 0,$$

we obtain that $H^i_{\mathfrak{m}}(B) = 0$ for all $i \neq d-1$ and

$$0 \to H^{d-1}_{\mathfrak{m}}(B) \to H^d_{\mathfrak{m}}(A(-a)) \to H^d_{\mathfrak{m}}(A) \to 0$$

is an exact sequence. Since $H^d_{\mathfrak{m}}(A)$ is nonzero and bounded above,

$$\deg H^{d-1}_{\mathfrak{m}}(B) = \deg H^d_{\mathfrak{m}}(A(-a)) = \deg H^d_{\mathfrak{m}}(A) + a = -\mathfrak{l} + a.$$

This implies that

$$\mathrm{CMreg}(B) = (d-1) + (-\mathfrak{l} + a) = (d - \mathfrak{l}) + (a-1). \tag{5.3.1}$$

As a consequence, B is of type $(d-1, \mathfrak{l}-a)$.

(1) If $a \geq \mathfrak{l} - d + 2$, then, by (5.3.1), $\mathrm{CMreg}(B) \geq 1$. By (2.1.2), B is not AS regular.

(2) Part (1) takes care of the case when $a \geq \mathfrak{l} - d + 2$. It remains to consider the case $a = \mathfrak{l} - d + 1 \geq 3$. By (5.3.1), $\mathrm{CMreg}(B) = a - (\mathfrak{l} - d + 1) = 0$. If B is AS regular, then $\mathrm{Torreg}({}_B\Bbbk) = -\mathrm{CMreg}(B) = 0$ or B is Koszul. Since A is isomorphic to $\Bbbk\langle x_1, \ldots, x_n\rangle/(R)$ for some x_i of degree 1, B is isomorphic to $\Bbbk\langle x_1, \ldots, x_n\rangle/(R, \Omega)$. Since $\deg \Omega \geq 3$, B is not quadratic. So B is not Koszul, yielding a contradiction. Therefore, B is not AS regular. □

Example 5.4(iv) shows that, if $a = \max\{2, \mathfrak{l} - d + 1\}$, then $A/(\Omega)$ may be AS regular.

Example 5.4. Let T be a noetherian AS regular algebra. Then $\mathrm{Torreg}({}_T\Bbbk) < \infty$. Let B be a factor ring $T/(\Omega)$, where Ω is a regular normal element of degree a. The algebra B is s-Cohen–Macaulay, where $s = \mathrm{gldim}\, T - 1 = d - 1$ and, by (5.3.1),

$$\mathrm{CMreg}(B) = \mathrm{CMreg}(T) + (a-1). \tag{5.4.1}$$

If a is 1 or 2, by Lemma 5.2(1), $\mathrm{Torreg}({}_B\Bbbk) < \infty$.

Now we give some explicit examples:

(i) Let $A = \Bbbk[x]/(x^a)$ with $a \geq 2$. If $a \geq 3$, $\mathrm{Torreg}({}_A\Bbbk) = \infty$ by Example 3.1. This means that the assertion in Lemma 5.2(1) fails if $\deg \Omega > 2$.

(ii) Let T be the algebra $\Bbbk\langle x, y\rangle/(x^2y - yx^2, xy^2 - y^2x)$, which is a noetherian AS regular algebra of type $(3, 4)$ (a special case of the algebra given in Example 2.4(3)). By Example 2.4(3), $\mathrm{Torreg}({}_T\Bbbk) = 1$.

It is easy to see that $\Omega := x^2$ is a normal regular element of T. Let $B = T/(\Omega)$. Then, by Lemma 5.2(1), $\mathrm{Torreg}({}_B\Bbbk) \leq \mathrm{Torreg}({}_T\Bbbk) = 1$. Since B is isomorphic to $\Bbbk\langle x, y\rangle/(x^2, xy^2 - y^2x)$, which is not Koszul, $\mathrm{Torreg}({}_B\Bbbk) \geq 1$. Therefore, $\mathrm{Torreg}({}_B\Bbbk) = 1$. By (5.4.1),

$$\mathrm{CMreg}(B) = \mathrm{CMreg}(T) + (2-1) = -\mathrm{Torreg}({}_T\Bbbk) + 1 = -1 + 1 = 0.$$

Hence, $\mathrm{ASreg}(B) = 1$ and B is not AS regular. This example shows that Theorem 4.5(2) fails without the commutativity assumption.

Let $A = B^{\otimes n}$ for any $n \geq 1$. Then it is easy to check that $\mathrm{CMreg}(A) = 0$ and $\mathrm{Torreg}({}_A\Bbbk) = \mathrm{ASreg}(A) = n$. As a consequence, A is not AS regular. By Theorem 4.5(1), there does not exist a finite map from a Koszul AS regular algebra to A (for any $n \geq 1$).

(iii) Let B be the algebra in part (ii). Since $T \to B$ is a surjective map, $c(B) \leq -\mathrm{CMreg}(T) = \mathrm{Torreg}({}_T\Bbbk) = 1$. We claim that $c(B) = 1$. First, by definition, $c(B) \geq 0$, and hence $c(B)$ is either 0 or 1. Now, since $\mathrm{CMreg}(B) = 0$, we have $c_-(B) = c(B) + \mathrm{CMreg}(B) = c(B)$, which implies that $c_-(B)$ is either 0 or 1. Since B is not AS regular by part (ii), by Theorem 0.12, $c_-(B) = 1$. Therefore, $c(B) = c_-(B) = 1$ and we have proved the claim.

It is easy to see that

$$1 \leq c(B^{\otimes n}) = c_-(B^{\otimes n}) \leq n.$$

It would be interesting to work out the exact value of $c(B^{\otimes n})$.

(iv) Let T be as in part (ii) and let Ω be $xy - yx$. Then Ω is a regular normal element of T such that $B := T/(\Omega)$ is the commutative polynomial ring $\Bbbk[x, y]$. So

$$\mathrm{Torreg}({}_B\Bbbk) = 0 < 1 = \mathrm{Torreg}({}_T\Bbbk),$$

where the last equation is given in part (ii).

Since T is of type $(3, 4)$, $a = 2 = \max\{2, \ell - d + 1\}$. Therefore, Lemma 5.3(2) fails if the hypothesis $a \geq \max\{3, \ell - d + 1\}$ is replaced by $a \geq \max\{2, \ell - d + 1\}$.

Throughout this paper, we were mainly interested in five numerical homological invariants:

$$\mathrm{Torreg}({}_A\Bbbk), \quad \mathrm{CMreg}(A), \quad c(A), \quad \mathrm{ASreg}(A) \quad \text{and} \quad c_-(A).$$

By definition (Definitions 0.6 and 0.9(3)), the last two invariants are dependent on the first three. We now seek to understand what values are possible for the first three invariants.

Definition 5.5. Let A be a noetherian connected graded algebra with balanced dualizing complex.

(1) The $\mathfrak{ta}$-*pair associated to* A is defined to be

$$\mathfrak{ta}(A) := (\mathrm{Torreg}({}_A\Bbbk), \mathrm{ASreg}(A)) \in (\mathbb{N} \times \mathbb{N}) \cup \{(+\infty, +\infty)\}.$$

(2) The $\mathfrak{tc}$-*pair* associated to A is defined to be

$$\mathfrak{tc}(A) := (\mathrm{Torreg}({}_A\Bbbk), \mathrm{CMreg}(A)) \in (\mathbb{N} \cup \{+\infty\}) \times \mathbb{Z}.$$

Note that $\mathrm{CMreg}(A) \geq -\mathrm{Torreg}({}_A\Bbbk)$ for every A.

In the next lemma, let D denote a noetherian connected graded algebra with balanced dualizing complex.

Lemma 5.6. (1) *For every* $t \in \mathbb{N} \cup \{+\infty\}$ *and* $c \in \mathbb{Z}$ *with* $c \geq -t$, *there is an algebra* D *such that* $\mathfrak{tc}(D) = (t, c)$.

(2) *For every* $t \in \mathbb{N}$ *and* $c \in \mathbb{Z}$ *with* $c \geq -t$, *there is an algebra* D *such that* $\mathfrak{tc}(D) = (t, c)$ *and* $c(D) = t$.

(3) *For every pair* $(t, a) \in (\mathbb{N} \times \mathbb{N}) \cup \{(+\infty, +\infty)\}$, *there is an algebra* D *such that* $\mathfrak{ta}(D) = (t, a)$.

(4) *For every pair* $(t, b) \in \mathbb{N} \times \mathbb{N}$, *there is an algebra* D *such that*

$$(\operatorname{Torreg}({}_A\Bbbk), c_-(D)) = (c(D), c_-(D)) = (t, b).$$

Proof. (1) First assume $t = \infty$. Let T be the algebra in Example 5.4(ii). Then $\operatorname{CMreg}(T) = -1$. Let $A = \Bbbk[x]/(x^3)$. Then $\operatorname{CMreg}(A) = 2$. For every integer c, there are nonnegative integers p, q such that $c = (-1)p + 2q$. Let $D = T^{\otimes p} \otimes A^{\otimes q}$. One can calculate

$$\operatorname{CMreg}(D) = p \operatorname{CMreg}(T) + q \operatorname{CMreg}(A) = p(-1) + q(2) = c,$$

as desired.

Now we assume that $t \in \mathbb{N}$. Let T be the algebra in Example 5.4(ii). Then $\operatorname{Torreg}({}_T\Bbbk) = 1$ and $\operatorname{Torreg}({}_{T^{\otimes t}}\Bbbk) = t$ by Lemma 2.3. Let $a = t + c$, which is a nonnegative integer by the hypothesis. Let E be the algebra $\Bbbk[x]/(x^2)$, which has $\operatorname{CMreg}(E) = 1$ and $\operatorname{Torreg}({}_E\Bbbk) = 0$. Let $D = T^{\otimes t} \otimes E^{\otimes a}$. Then

$$\operatorname{CMreg}(D) = t \operatorname{CMreg}(T) + a \operatorname{CMreg}(E) = -t + a = c$$

and

$$\operatorname{Torreg}(D) = t \operatorname{Torreg}(T) + a \operatorname{Torreg}(E) = t + a0 = t,$$

as desired.

(2) We use the second half of the proof of (1) and note that

$$c(D) = c(T^{\otimes t} \otimes E^{\otimes a}) = c(T^{\otimes t}) = t,$$

as desired.

(3) Since $\operatorname{ASreg}(A) = \operatorname{Torreg}({}_A\Bbbk) + \operatorname{CMreg}(A)$, the assertion follows easily from part (1).

(4) Since $c_-(A) = c(A) + \operatorname{CMreg}(A)$, the assertion follows easily from part (2). □

Note that the algebras A in the above proof are generated in degree 1 and are module-finite over their centers. We conclude the paper with the following remarks:

Remark 5.7. Assume Hypothesis 1.3.

(1) It would be interesting to work out the range of $c(A)$ (resp. $c_-(A)$) for every fixed $\mathfrak{tc}$-pair (t, c). When $c = -t$, by Theorems 0.8 and 0.12(1), $c(A)$ must be $-c$. However, computing the concavity $c(A)$ and normalized concavity $c_-(A)$ is generally much harder than computing $\operatorname{Torreg}({}_A\Bbbk)$ and $\operatorname{CMreg}(A)$.

(2) Further it would be interesting to study algebras with $\mathfrak{ta}$ near $(0, 0)$, for example $\mathfrak{ta}(A) = (1, 0)$ or $\mathfrak{ta}(A) = (0, 1)$. It is not clear to us whether an algebra A with $\mathfrak{ta}(A) = (0, 1)$ is AS Gorenstein, although we do not have any counterexample to this.

Remark 5.8. It is an open question whether the tensor product of two noetherian AS regular algebras (or two algebras satisfying Hypothesis 1.3) is always noetherian. If we ignore this issue, we can consider the behavior of regularities with respect to the tensor product.

(1) (Lemma 2.3) $\mathrm{Torreg}({}_A\Bbbk)$ is additive in the following sense:

$$\mathrm{Torreg}({}_{A\otimes B}\Bbbk) = \mathrm{Torreg}({}_A\Bbbk) + \mathrm{Torreg}({}_B\Bbbk).$$

(2) (Remark 2.9) $\mathrm{CMreg}(A)$ is additive in the following sense:

$$\mathrm{CMreg}(A \otimes B) = \mathrm{CMreg}(A) + \mathrm{CMreg}(B).$$

(3) It follows from parts (1) and (2), $\mathrm{ASreg}(A)$ is additive in the following sense:

$$\mathrm{ASreg}(A \otimes B) = \mathrm{ASreg}(A) + \mathrm{ASreg}(B).$$

(4) We do not know if $c(A)$ (resp. $c_-(A)$) is additive. It follows from the definition that

$$c(A \otimes B) \leq c(A) + c(B).$$

The same is true for $c_-(A)$.

Remark 5.9. It follows immediately from Theorem 0.10(3) that, if S and T are noetherian AS regular algebras and $f : T \to S$ and $g : S \to T$ are finite maps, then both f and g are isomorphisms of graded algebras.

The above statement fails if we assume that S and T are only AS Gorenstein. Let $S = T = \Bbbk[x]/(x^2)$, which is noetherian AS Gorenstein. Let $f(= g) : T \to T$ be defined by $a + bx \mapsto a$ for all $a, b \in \Bbbk$. Then f is a finite map that is not an isomorphism of graded algebras.

Remark 5.10. Let T be a noetherian AS regular Koszul algebra and let A be a connected graded algebra. Suppose $f : T \to A$ is a finite map. It is trivial that:

(1) $\mathrm{Torreg}({}_T A) \geq 0$.

By Theorem 0.7,

$$\mathrm{Torreg}({}_T A) = \mathrm{Torreg}({}_T A) + \mathrm{CMreg}(T) = \mathrm{CMreg}({}_T A) = \mathrm{CMreg}(A).$$

Now it follows from Theorem 4.5(1) that:

(2) $\mathrm{Torreg}({}_T A) = 0$, or, equivalently, ${}_T A$ is linear, if and only if A is AS regular (and Koszul).

Therefore, $\mathrm{Torreg}({}_T A)$ qualifies as an indicator of the AS regular property in this case.

We can say more in terms of the Hilbert series of the algebras.

Remark 5.11. We continue Remark 5.10 and consider a special case when $h_T(t) = (1-t)^{-d}$, where $d = \mathrm{gldim}\, T$. Let $f : T \to A$ be a finite map.

(1) Then A is AS regular if and only if there is an integer $0 \leq d' \leq d$ such that the i-th term in the minimal free resolution of the T-module ${}_T A$ is isomorphic to $T(-i)^{\oplus \binom{d'}{i}}$ for all i. If this happens, A is Koszul and $h_A(t) = (1-t)^{-d+d'}$.

(2) If A is known to be Cohen–Macaulay, then A is AS regular if and only if $h_A(t) = (1-t)^{-d+d'}$ for an integer $0 \leq d' \leq d$. If this happens, all properties in part (1) hold.

Remark 5.12. Let A and B be two connected graded algebras. We say that A and B are *finite-equivalent* if there are finite graded algebra homomorphisms $f : A \to B$ and $g : B \to A$. It is easy to see that being finite-equivalent is an equivalence relation. For example, $\Bbbk$ and $\Bbbk[x]/(x^a)$ with $a \geq 2$ are finite-equivalent. It follows from Proposition 4.1 that:

(1) If A and B are finite equivalent, then $c(A) = c(B)$.

The concavity $c(A)$ also has the following properties:

(2) If A is commutative and generated in degree 1, then $c(A) = 0$.

(3) (Proposition 4.1(4)) If x is a commutative indeterminate of degree 1, then $c(A[x]) = c(A)$.

By Remark 5.9, if S and T are finite-equivalent noetherian AS regular algebras, then $S \cong T$. Properties (1) and (2) fail for other invariants that we have defined in this paper. See Remark 5.8 for property (3).

Remark 5.13. The relationship between $\mathrm{Torreg}({}_A\Bbbk)$ and $\mathrm{CMreg}(A)$ (resp. $c(A)$ and $\mathrm{CMreg}(A)$) is given in Lemma 5.6(1,2). It seems that there is a mysterious (maybe closer) connection between $\mathrm{Torreg}({}_A\Bbbk)$ and $c(A)$. By Jørgensen's theorem (Theorem 2.5),

$$\mathrm{Torreg}({}_A\Bbbk) \geq -\mathrm{CMreg}(A),$$

and, by Theorem 0.12(1) and the definition,

$$c(A) = c_-(A) - \mathrm{CMreg}(A) \geq -\mathrm{CMreg}(A).$$

By Lemma 5.6(2), for each $t \geq 0$, there is an algebra D such that $\mathrm{Torreg}({}_A\Bbbk) = c(A) = t$.

Based on this very limited evidence, we ask the following two questions:

Question 5.14. Suppose both $c(A)$ and $\mathrm{Torreg}({}_A\Bbbk)$ are finite. Then is $\mathrm{Torreg}({}_A\Bbbk)$ uniformly bounded by a function of $c(A)$ (or vice versa)?

A special case of the above question is related to a nice result in [Avramov and Peeva 2001] (also see [Römer 2008, Theorem 2.3]). Is there a noncommutative version of this result?

Question 5.15. Suppose $c(A) = 0$ and $\mathrm{Torreg}({}_A\Bbbk) < \infty$. Is then $\mathrm{Torreg}({}_A\Bbbk) = 0$ (or, equivalently, is A Koszul)?

Acknowledgements

The authors thank the referee for their careful reading of the manuscript and useful suggestions, particularly in improving the proofs of Theorems 0.10(1) and 0.12(1). Won was partially supported by an AMS–Simons Travel Grant and Simons Foundation grant #961085. Zhang was partially supported by the US National Science Foundation (DMS-1700825, DMS-2001015 and DMS-2302087).

References

[Artin and Schelter 1987] M. Artin and W. F. Schelter, "Graded algebras of global dimension 3", *Adv. in Math.* **66**:2 (1987), 171–216. MR

[Artin and Zhang 1994] M. Artin and J. J. Zhang, "Noncommutative projective schemes", *Adv. Math.* **109**:2 (1994), 228–287. MR

[Artin et al. 1990] M. Artin, J. Tate, and M. Van den Bergh, "Some algebras associated to automorphisms of elliptic curves", pp. 33–85 in *The Grothendieck Festschrift*, *I*, edited by P. Cartier et al., Progr. Math. **86**, Birkhäuser, Boston, MA, 1990. MR

[Artin et al. 1991] M. Artin, J. Tate, and M. Van den Bergh, "Modules over regular algebras of dimension 3", *Invent. Math.* **106**:2 (1991), 335–388. MR

[Avramov and Peeva 2001] L. L. Avramov and I. Peeva, "Finite regularity and Koszul algebras", *Amer. J. Math.* **123**:2 (2001), 275–281. MR

[Van den Bergh 1997] M. Van den Bergh, "Existence theorems for dualizing complexes over non-commutative graded and filtered rings", *J. Algebra* **195**:2 (1997), 662–679. MR

[Dong and Wu 2009] Z.-C. Dong and Q.-S. Wu, "Non-commutative Castelnuovo–Mumford regularity and AS-regular algebras", *J. Algebra* **322**:1 (2009), 122–136. MR

[Eisenbud and Goto 1984] D. Eisenbud and S. Goto, "Linear free resolutions and minimal multiplicity", *J. Algebra* **88**:1 (1984), 89–133. MR

[Jørgensen 1997] P. Jørgensen, "Local cohomology for non-commutative graded algebras", *Comm. Algebra* **25**:2 (1997), 575–591. MR

[Jørgensen 1999] P. Jørgensen, "Non-commutative Castelnuovo–Mumford regularity", *Math. Proc. Cambridge Philos. Soc.* **125**:2 (1999), 203–221. MR

[Jørgensen 2004] P. Jørgensen, "Linear free resolutions over non-commutative algebras", *Compos. Math.* **140**:4 (2004), 1053–1058. MR

[Jørgensen and Zhang 2000] P. Jørgensen and J. J. Zhang, "Gourmet's guide to Gorensteinness", *Adv. Math.* **151**:2 (2000), 313–345. MR

[Kirkman 2016] E. E. Kirkman, "Invariant theory of Artin–Schelter regular algebras: a survey", pp. 25–50 in *Recent developments in representation theory*, edited by A. Martsinkovsky et al., Contemp. Math. **673**, Amer. Math. Soc., Providence, RI, 2016. MR

[Kirkman et al. 2009] E. Kirkman, J. Kuzmanovich, and J. J. Zhang, "Gorenstein subrings of invariants under Hopf algebra actions", *J. Algebra* **322**:10 (2009), 3640–3669. MR

[Kirkman et al. 2015] E. Kirkman, J. Kuzmanovich, and J. J. Zhang, "Noncommutative complete intersections", *J. Algebra* **429** (2015), 253–286. MR

[Kirkman et al. 2022] E. Kirkman, R. Won, and J. J. Zhang, "Degree bounds for Hopf actions on Artin–Schelter regular algebras", *Adv. Math.* **397** (2022), art. id. 108197. MR

[Kirkman et al. 2023] E. Kirkman, R. Won, and J. J. Zhang, "Weighted homological regularities", *Trans. Amer. Math. Soc.* **376**:10 (2023), 7407–7445. MR

[Polishchuk 2006] A. Polishchuk, "Koszul configurations of points in projective spaces", *J. Algebra* **298**:1 (2006), 273–283. MR

[Reyes et al. 2014] M. Reyes, D. Rogalski, and J. J. Zhang, "Skew Calabi–Yau algebras and homological identities", *Adv. Math.* **264** (2014), 308–354. MR

[Rogalski and Sierra 2012] D. Rogalski and S. J. Sierra, "Some projective surfaces of GK-dimension 4", *Compos. Math.* **148**:4 (2012), 1195–1237. MR

[Römer 2008] T. Römer, "On the regularity over positively graded algebras", *J. Algebra* **319**:1 (2008), 1–15. MR

[Rotman 1979] J. J. Rotman, *An introduction to homological algebra*, Pure and Applied Mathematics **85**, Academic, New York, 1979. MR

[Stafford and Zhang 1994] J. T. Stafford and J. J. Zhang, "Examples in non-commutative projective geometry", *Math. Proc. Cambridge Philos. Soc.* **116**:3 (1994), 415–433. MR

[Stanley 1978] R. P. Stanley, "Hilbert functions of graded algebras", *Adv. Math.* **28**:1 (1978), 57–83. MR

[Stanley 1979] R. P. Stanley, "Invariants of finite groups and their applications to combinatorics", *Bull. Amer. Math. Soc.* (*N.S.*) **1**:3 (1979), 475–511. MR

[Stephenson and Zhang 1997] D. R. Stephenson and J. J. Zhang, "Growth of graded Noetherian rings", *Proc. Amer. Math. Soc.* **125**:6 (1997), 1593–1605. MR

[Yekutieli 1992] A. Yekutieli, "Dualizing complexes over noncommutative graded algebras", *J. Algebra* **153**:1 (1992), 41–84. MR

[Zhang 1997] J. J. Zhang, "Connected graded Gorenstein algebras with enough normal elements", *J. Algebra* **189**:2 (1997), 390–405. MR

Communicated by Irena Peeva
Received 2024-02-16 Revised 2025-03-21 Accepted 2025-08-05

kirkman@wfu.edu *Department of Mathematics, Wake Forest University, Winston-Salem, NC, United States*

robertwon@gwu.edu *Department of Mathematics, The George Washington University, Washington, DC, United States*

zhang@math.washington.edu *Department of Mathematics, University of Washington, Seattle, WA, United States*

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1597

Transcendence properties of the Artin–Hasse exponential modulo p

Joe Kramer-Miller

Let $E_p(x)$ denote the Artin–Hasse exponential and $\bar{E}_p(x)$ denote its reduction modulo p in $\mathbb{F}_p[\![x]\!]$. We study transcendence properties of $\bar{E}_p(x)$ over $\mathbb{F}_p[x]$. We give two proofs that $\bar{E}_p(x)$ is transcendental, affirmatively answering a question of Thakur. We also prove algebraic independence results: (i) for $f_1, \dots, f_r \in x\overline{\mathbb{F}}_p[x]$ satisfying certain linear independence properties, we show that the $\bar{E}_p(f_1), \dots, \bar{E}_p(f_r)$ are algebraically independent over $\mathbb{F}_p[x]$; and (ii) we determine the algebraic relations between the $\bar{E}_p(\xi x)$ for $\xi \in \mathbb{F}_p^\times$. Our proof studies the higher derivatives of $\bar{E}_p(x)$ and makes use of iterative differential Galois theory.

1. Introduction

The Artin–Hasse exponential is defined by

$$E_p(x) := \exp\left(\sum_{n=0}^{\infty} \frac{x^{p^n}}{p^n}\right)$$

and is known to have coefficients in $\mathbb{Z}_{(p)}$. This series plays an enormous role in p-adic analysis, showing up in the construction of Witt vectors (see, e.g., [Dieudonné 1957]) and the p-adic approach to exponential sums (see, e.g., [Adolphson and Sperber 1989; Wan 1993]). From the relation $E_p(x^p)/E_p(x)^p = \exp(-xp)$, it is immediate that $E_p(x)$ is a transcendental function over $\mathbb{C}(x)$. However, we may also consider the reduction of $E_p(x)$ modulo p, which we denote by $\bar{E}_p(x)$. Then $\bar{E}_p(x)$ is a power series in $\mathbb{F}_p[\![x]\!]$ that in some sense plays the role of the exponential function in characteristic p. For example, one can use $\bar{E}_p(x)$ to translate Witt vector addition into power series multiplication, as explained in [Dieudonné 1957]. Here we investigate the transcendence properties of $\bar{E}_p(x)$. We give two proofs that $\bar{E}_p(x)$ is transcendental over $\mathbb{F}_p(x)$. This answers a question posed by Dinesh Thakur (see, e.g., [Thakur 2020] or the open questions in [Allouche and Shallit 2003, Chapter 12]). We also prove algebraic independence results for the evaluations of $\bar{E}_p(x)$ at different values in $x\overline{\mathbb{F}}_p[x]$.

1.1. *Main results and outline.*

Our first result is the following.

Theorem 1.1. *The series $\bar{E}_p(x)$ is transcendental over $\mathbb{F}_p(x)$.*

We give two proofs of Theorem 1.1. The first proof is given in Section 2. It uses Christol's automatic criterion for algebraicity of power series together with a functional equation: for $\xi \in \mathbb{F}_p^\times$, we have

MSC2020: 11B85, 11J91.
Keywords: transcendence in positive characteristic, differential Galois theory.

© 2026 The Author, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

$\bar{E}_p(\xi x) = \bar{E}_p(x)^{[\xi]}$, where $[\xi]$ denotes the Teichmüller lift of ξ (i.e., $[\xi]$ is the unique $p-1$-st root of unity whose reduction modulo p is ξ). This proof is relatively elementary and short, but has the disadvantage of requiring p to be at least 5. The second proof uses iterative differential Galois theory and iterative Frobenius equations. These are the same methods we use for the stronger algebraic independence results, which we describe in detail below.

To state our first algebraic independence result, we need the following definition.

Definition 1.2. Let $f(x) = \sum c_i x^i \in \bar{\mathbb{F}}_p[x]$. We define the *primitive* part of f to be

$$f^*(x) := \sum_{\substack{i \geq 0 \\ p \nmid i}} c_i^*(f) x^i, \; c_i^*(f) := \sum_{k \geq 0} c_{ip^k}^{1/p^k}.$$

Note that $f^*(x) = 0$ if and only if $f(x) = g(x)^p - g(x)$ for some $g(x) \in \bar{\mathbb{F}}_p[x]$.

Theorem 1.3. *Let $f_1, \dots, f_r \in x\bar{\mathbb{F}}_p[x]$ be nonzero. Assume that one of the following conditions holds*:

(1) *The elements $f_1^*, \dots, f_r^*$ are linearly independent over $\mathbb{F}_p$.*

(2) *The f_i^* are not all zero and no nontrivial power product of the f_i are contained in $\bar{\mathbb{F}}_p(x)^p$. In other words, if $f_1^{n_1} \cdots f_r^{n_r}$ is a p-th power, then $n_1, \dots, n_r \in p\mathbb{Z}$.*

Then $\bar{E}_p(f_1), \dots, \bar{E}_p(f_r)$ are algebraically independent.

Example 1.4. For $f(x) = x^k$ with $p \neq k$, we have $f^*(x) = f(x)$. In particular, we see that the elements of $\{\bar{E}_p(x^k)\}_{k \geq 1,\, p \nmid n}$ are algebraically independent using the first condition of Theorem 1.3. It is also clear that the second condition of Theorem 1.3 does not apply, as for any k_1, k_2 coprime to p we can find n_1, n_2 coprime to p so that $p \mid (n_1 k_1 + n_2 k_2)$. In particular $(x^{k_1})^{n_1} \cdot (x^{k_2})^{n_2} \in \bar{\mathbb{F}}_p(x)^p$, but $n_1, n_2 \notin p\mathbb{Z}$.

Example 1.5. Let $f_1, \dots, f_r \in x\bar{\mathbb{F}}_p[x]$ be nonzero. Assume that there exists $c_1, \dots, c_r \in \bar{\mathbb{F}}_p$ such that $f_i(c_i) = 0$ and $f_i(c_j) \neq 0$ for $i \neq j$. We furthermore assume that the zero of f_i at c_i is simple. Then, for integers $n_1, \dots, n_r$ with $p \nmid n_i$, the order of $f_1^{n_1} \cdots f_r^{n_r}$ at c_i is coprime to p. In particular, the product $f_1^{n_1} \cdots f_r^{n_r}$ cannot be in $\bar{\mathbb{F}}_p(x)^p$. Thus, the second condition of Theorem 1.3 is satisfied. On the other hand, it is clear that the first condition of Theorem 1.3 can fail. For example, take $c_1, \dots, c_r$ to be distinct elements of $\mathbb{F}_p$ and set $f_i = x(x - c_i)$. Assume that $p \geq 3$ and $r \geq 3$. Then $f_i^* = f_i$ and the $f_1^*, \dots, f_r^*$ are linearly dependent over $\mathbb{F}_p$.

Example 1.6. Consider the case where $f_2 = f_1^p$. Then $\bar{E}_p(f_1)$ and $\bar{E}_p(f_2)$ are algebraically dependent, as $\bar{E}_p(f_2) = \bar{E}_p(f_1)^p$. We know $f_2 \in \bar{\mathbb{F}}_p(x)^p$, so the second condition of Theorem 1.3 fails. It is also easy to verify that $f_2^* = f_1^*$, so the first condition of Theorem 1.3 fails.

Our second algebraic independence result is the following.

Theorem 1.7. *We have*

$$\operatorname{tr.deg}_{\mathbb{F}_p(x)} \mathbb{F}_p(x)(\bar{E}_p(\xi x))_{\xi \in \mathbb{F}_p^\times} = \phi(p-1),$$

where ϕ is the Euler totient function.

The proofs of Theorems 1.3 and 1.7 follow the same strategy. This strategy can be broken into three steps, which we outline below. For simplicity, we restrict our attention to the case where the f_i in Theorem 1.3 have coefficients in $\mathbb{F}_p$.

Iterated Frobenius equations. Let σ be the lifted Frobenius endomorphism of $\mathbb{Z}_p[\![x]\!]$ that fixes $\mathbb{Z}_p$ and sends x to x^p. For $f \in x\mathbb{Z}_p[\![x]\!]$, we define $\tau(f)$ to be the unique solution to $y^\sigma - y = f$ in $x\mathbb{Z}_p[\![x]\!]$ (explicitly, $\tau(f) = \sum_{n\geq 0} f^{\sigma^n}$). Let $b_1, \ldots, b_k$ be positive integers coprime to p and consider the power series defined by

$$y_1 = \tau(x^{b_1}), \quad y_2 = \tau(x^{b_2} y_1), \quad \ldots, \quad y_k = \tau(x^{b_k} y_{k-1}).$$

We refer to y_k as an iterated Frobenius solution (abbreviated IF-solution) of depth k. This terminology is justified by observing that the y_i satisfy the iterated Frobenius equations

$$y_1^\sigma - y_1 = x^{b_1}, \quad y_2^\sigma - y_2 = x^{b_2} y_1, \quad \ldots, \quad y_k^\sigma - y_k = x^{b_k} y_{k-1}.$$

Let M_k denote the $\mathbb{Z}_p(x)$-module generated by all IF-solutions of depth at most k and let $M = \bigcup M_k$. In Section 3, we study M_k and M. More precisely, we do the following:

(1) We prove that the IF-solutions freely generate M_k (i.e., there are no relations over $\mathbb{Z}_p(x)$ between different IF-solutions).

(2) We show that M is closed under multiplication and $M_{k_1} M_{k_2} \subset M_{k_1+k_2}$. Furthermore, we study the multiplication law in terms of the IF-solutions.

Higher derivatives. Let $\partial^{(k)}$ denote the differential operator $\frac{1}{k!} d^k/d^k x$ and let $\mathrm{dLog}^{(k)}$ be the operator sending f to $\partial^{(k)}(f)/f$. We regard $\partial^{(k)}$ and $\mathrm{dLog}^{(k)}$ as operators on $\mathbb{F}_p[\![x]\!]$ and on $\mathbb{Z}_p[\![x]\!]$. Note that, if f is contained in a finite separable extension L of $\mathbb{F}_p(x)$, then so is $\mathrm{dLog}^{(k)}(f)$ for all $k \geq 0$. This gives a differential transcendence criterion: if the $\mathrm{dLog}^{(k)}(f)$ generate an infinite extension of $\mathbb{F}_p(x)$, then f is transcendental over $\mathbb{F}_p(x)$. In Section 4, we study the iterated differential equations satisfied by $\bar{E}_p(x)$ and power products of the $\bar{E}_p(f_r)$. Define $\overline{M}_k$ to be M_k/pM_k. We prove that $\mathrm{dLog}^{(k)}(\bar{E}_p(x))$ is contained in $\overline{M}_k$, but not in $\overline{M}_{k-1}$. For example, when $k = 1$, a quick calculation shows that $\frac{d}{dx} E_p(x) = \frac{y}{x} E_p(x)$, where $y = \tau(x)$. More precisely, we determine $\mathrm{dLog}^{(k)}(\bar{E}_p(x))$ modulo $\overline{M}_{k-1}$. This is enough to show $\bar{E}_p(x)$ is transcendental. Next, using standard properties of higher derivatives and results from Section 3, we are able to determine $\mathrm{dLog}^{(k)}(\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r})$ modulo $\overline{M}_{k-1}$, where $n_i \in \mathbb{Z}$ and $f_i \in \mathbb{F}_p[x]$.

Iterative differential Galois groups. Let K be the algebraic closure of $\mathbb{F}_p(x)$ in $\overline{\mathbb{F}}_p((x))$. An ID-module over K is a vector space over K with compatible actions of $\partial^{(k)}$. For any $f \in x\mathbb{F}_p[x]$, we let N_f denote the ID-module associated to the differential equation satisfied by $\bar{E}_p(f)$ (see Section 5.1 for a precise description of this ID-module). Following Matzat and van der Put [2003], we may define an ID-Galois group $\mathrm{Gal}(N_f)$. Then $\mathrm{Gal}(N_f)$ is a closed subgroup of the multiplicative group $\mathbb{G}_m$ over $\overline{\mathbb{F}}_p$ and is equal to all of $\mathbb{G}_m$ precisely when $\bar{E}_p(f)$ is transcendental over $\mathbb{F}_p(x)$. More generally, the algebraic independence of $\bar{E}_p(f_1), \ldots, \bar{E}_p(f_r)$ is equivalent to the ID-Galois group $\mathrm{Gal}(N_{f_1} \oplus \cdots \oplus N_{f_r})$ being isomorphic to $\mathbb{G}_m^r$. Using the Tannakian interpretation of ID-Galois groups, it is enough to

prove $\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}$ is transcendental for any $n_i \in \mathbb{Z}$. This is essentially a version of Kolchin's theorem [1968] for iterated differential equations. We then address this transcendence question assuming the hypotheses of Theorems 1.7 and 1.3 using the results from Section 4.

Remark 1.8. The appearance of differential Galois theory naturally places Theorem 1.3 in a rich historical context. Ostrowski [1946] gives a simple algebraic independence criterion for integrals of elements of a differential field K of meromorphic functions. This was extended to include exponentials of elements of a differential field (as well as further results) by Kolchin [1968]. The use of algebraic groups and differential Galois theory has since played an outstanding role in the transcendence theory of functions. To the best of our knowledge, this paper is the first application of differential Galois theory to transcendence in positive characteristic. However, we mention that the use of Tannakian theory for Frobenius equations has played an exceptional role in function field arithmetic. This was pioneered by Papanikolas [2008] and lays the framework for studying transcendence of periods of Drinfeld modules and t-motives.

1.2. ***Further questions.*** It would be interesting to completely determine the algebraic relations between values $\bar{E}_p(f)$ for $f \in x\overline{\mathbb{F}}_p[x]$. More generally, it would be interesting to determine the algebraic relations between values of $\bar{E}_p(x)$ at any algebraic power series. At the moment it is even unclear to the author what the algebraic relations are between the $\bar{E}_p(\xi x)$ where ξ ranges over $\overline{\mathbb{F}}_p$. The reader may observe the influence of Ax's work [1971] on Schanuel's conjecture on the final steps of the proof of Theorem 1.3. Thus, it is natural to ask if an analogue of Ax–Schanuel holds for the Artin–Hasse exponential.

2. An automatic approach to the transcendence of $\bar{E}_p(x)$

In this section we give a short proof of the transcendence of $\bar{E}_p(x)$ using the functional equation satisfied by $\bar{E}_p(x)$ and automatic methods.

2.1. ***A transcendental criterion.*** We recall Christol's well-known automatic criterion for algebraicity of series in $\mathbb{F}_p[\![x]\!]$. See [Christol 1979] or [Christol et al. 1980] for proofs.

Theorem 2.1. *Let $f = \sum f_n x^n \in \mathbb{F}_p[\![x]\!]$. The following are equivalent*:

(1) *The series f is algebraic over $\mathbb{F}_p[x]$.*

(2) *There is a p-automaton that gives output f_n on input n.*

(3) *There are only finitely many subsequences of $(f_n)_{n\geq 0}$ of the form $(f_{p^k n+r})_{n\geq 0}$, where k is positive and r is in the range $0 \leq r < p^k$.*

This allows us to establish the following transcendence criterion.

Lemma 2.2. *Let $\lambda \in \mathbb{Z}_p$. Let $f \in 1 + x\mathbb{F}_p[\![x]\!]$ be algebraic over $\mathbb{F}_p[x]$. Then $f^\lambda \in \mathbb{F}_p[\![x]\!]$ is algebraic over $\mathbb{F}_p[x]$ if and only if λ is rational.*

Proof. This is a special case of [Allouche et al. 1988]. We give a proof in the interest of having a quick and self-contained automatic proof that $\bar{E}_p(x)$ is transcendental. We easily reduce to the case where $f = 1 + x$. If λ is rational, it is clear that $(1 + x)^\lambda$ is algebraic. For the other direction, assume λ is

irrational. Write $\lambda = \sum_{n=0}^{\infty} a_n p^n$. For $1 \le c \le p-1$, we define sets

$$S_c := \{n \mid a_n = c\} \quad \text{and} \quad S_c^k := S_c \cap \mathbb{Z}_{>k}.$$

We then set $\lambda_c = \sum_{n \in S_c} a_n p^n$. If λ_c is nonzero and rational, we may replace λ with $\lambda - \lambda_c$. In particular, we may assume each λ_c is either zero or irrational. Let c_0 be the largest number such that $\lambda_{c_0} \neq 0$. As λ_{c_0} is irrational, for each $k, m > 0$ we have

$$m + S_{c_0}^k \neq S_{c_0}^{k+m}, \tag{1}$$

as otherwise the digits of λ_{c_0} would eventually be periodic.

Next, write $(1+x)^\lambda = \sum b_n x^n$. By using the product expansion $(1+x)^\lambda = \prod (1+x^{p^n})^{a_n}$, we see that

$$b_{c_0 p^k} = \begin{cases} 1 & \text{if } k \in S_{c_0}, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the sequences $B_k = (b_{np^k})_{n \ge 1}$. By Theorem 2.1, if we prove that the B_k are all distinct, we will know $(1+x)^\lambda$ is transcendental. Note that the $c_0 p^n$-th element in B_k is 1 if $n + k \in S_{c_0}^k$ and zero otherwise. In particular, if $B_k = B_{k+m}$, then we must have $m + S_{c_0}^k = S_{c_0}^{k+m}$. However, by (1) we know this can never happen, which proves the lemma. □

2.2. ***A short proof of the transcendence of $\bar{E}_p(x)$.*** The first step is to prove that $\bar{E}_p(x)$ satisfies a functional equation.

Lemma 2.3. *For $\xi \in \mathbb{F}_p^\times$, we have $\bar{E}_p(\xi x) = \bar{E}_p(x)^{[\xi]}$, where $[\xi] \in \mathbb{Z}_p$ denotes the Teichmüller lift of ξ (i.e., $[\xi]$ is the unique $p-1$-st root of unity in $\mathbb{Z}_p$ that reduces to ξ modulo p).*

Proof. For $f_0 \in \mathbb{Q}_p[\![x]\!]$ and $f_1, \dots, f_n \in x\mathbb{Q}_p[\![x]\!]$, we may take the composition $f_0 \circ \cdots \circ f_n \in \mathbb{Q}_p[\![x]\!]$, and this operation is associative. For $c \in \mathbb{Z}_p$ we define

$$g(x) = \sum_{n=0}^{\infty} \frac{x^{p^n}}{p^n}, \quad L_c(x) = cx \quad \text{and} \quad P_c(x) = (1+x)^c = \sum_{n=0}^{\infty} \binom{c}{n} x^n.$$

Note that $E_p(x) = \exp \circ g$. Also, for $\xi \in \mathbb{F}_p^\times$, we have the relations

$$g \circ L_{[\xi]} = L_{[\xi]} \circ g \quad \text{and} \quad \exp \circ L_{[\xi]} = P_{[\xi]} \circ (\exp(x) - 1).$$

The first relation comes from the fact that $[\xi]^p = [\xi]$. Thus, we compute

$$E_p([\xi]x) = \exp \circ g \circ L_{[\xi]} = \exp \circ L_{[\xi]} \circ g = P_{[\xi]} \circ (\exp(x) - 1) \circ g = P_{[\xi]} \circ (E_p(x) - 1) = E_p(x)^{[\xi]}.$$

The lemma follows by reducing modulo p. □

Theorem 2.4. *The series $\bar{E}_p(x)$ is transcendental over $\mathbb{F}_p(x)$ if $p \ge 5$.*

Proof. Take $\xi \in \mathbb{F}_p^\times$. Since $p \ge 5$, we can pick ξ to not be 1 or -1. This means $[\xi]$ is an algebraic integer that is not an integer. Thus, $[\xi]$ is irrational. Assume $\bar{E}_p(x)$ is algebraic. This implies $\bar{E}_p(\xi x)$ is algebraic. Thus, by Lemma 2.3, $\bar{E}_p(x)^{[\xi]}$ is algebraic. However, $[\xi]$ is not rational, so Lemma 2.2 tells us that $\bar{E}_p(x)^{[\xi]}$ is transcendental. This is a contradiction. □

Remark 2.5. One can get algebraic independence results using the full result of [Allouche et al. 1988]. In particular, the argument above can be extended to show $\operatorname{tr.deg}_{\mathbb{F}_p(x)}(\bar{E}_p(\xi x))_{\xi\in\mathbb{F}_p^\times} \geq \phi(p-1)-1$. As we prove in Section 5 that $\operatorname{tr.deg}_{\mathbb{F}_p(x)}(\bar{E}_p(\xi x))_{\xi\in\mathbb{F}_p^\times} = \phi(p-1)$, we omit the proof of this weaker result.

3. Solutions to iterated Frobenius equations

Let $\mathbb{F}_q$ be the field with $q = p^f$ elements and fix $\zeta \in \mathbb{F}_q$ such that $\zeta, \zeta^p, \dots, \zeta^{p^{f-1}}$ is a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. Let $\mathbb{Z}_q$ denote the p-typical Witt vectors of $\mathbb{F}_q$. By abuse of notation, we regard ζ as an element of $\mathbb{Z}_q$ by identifying ζ^{p^j} with its Teichmüller lift. It will be convenient to define $x_j^i = \zeta^{p^j} x^i$ for $i \geq 1$ and any $j \in \mathbb{Z}$. We also define the set

$$X := \{x_j^i \mid 1 \leq j \leq f,\ i \geq 1,\ p\nmid i\}.$$

3.1. ***Iterated Frobenius equations.*** Let $\sigma : \mathbb{Z}_q((x)) \to \mathbb{Z}_q((x))$ denote the $\mathbb{Z}_p$-linear endomorphism sending x to x^p and acting as the Frobenius endomorphism on $\mathbb{Z}_q$. Let $f \in x\mathbb{Z}_q[\![x]\!]$. We define $\tau(f)$ to be the unique solution to $y^\sigma - y = f$ in $x\mathbb{Z}_q[\![x]\!]$. Concretely, we have

$$\tau(f) = \sum_{n=0}^{\infty} f^{\sigma^n}.$$

We regard τ as a $\mathbb{Z}_p$-linear map. Note that τ satisfies the two fundamental equations

$$\tau(f)\tau(g) = \tau(f\tau(g)) + \tau(g\tau(f)) - \tau(fg), \tag{2}$$

$$\tau(f^\sigma \tau(g)) = \tau(f\tau(g)) + \tau(f^\sigma g) - f\tau(g). \tag{3}$$

Definition 3.1. Let $[k]$ denote the set $\{1, \dots, k\}$ and let $[0]$ denote the empty set. An *iterated Frobenius datum* b (abbreviated as IF-datum) of depth k is a function

$$b : [k] \to X.$$

We denote the depth of b by $\operatorname{dep}(b)$. We remark that there is exactly one IF-datum of depth 0. We define D to be the set of all IF-data.

Definition 3.2. We let b^{tr} denote the restriction of b to $[k-1]$ (we think of b as a list of elements and b^{tr} as the truncation of the last element, whence the "tr" in the superscript). Also, for any $z \in X$, we let $\operatorname{cat}(b, z)$ be the IF-datum of depth $k+1$ defined by

$$\operatorname{cat}(b,z)(i) \begin{cases} b(i) & \text{if } i \in [k], \\ z & \text{if } i = k+1. \end{cases}$$

Definition 3.3. Let b be an IF-datum of depth k. We recursively define $w_b \in x\mathbb{Z}_q[\![x]\!]$ as

$$w_b = \begin{cases} 1 & \text{if } \operatorname{dep}(b) = 0, \\ \tau(b(k) w_{b^{\mathrm{tr}}}) & \text{if } \operatorname{dep}(b) = k > 0. \end{cases}$$

We refer to w_b as an iterative Frobenius solution (abbreviated as IF-solution).

Remark 3.4. Let b be an IF-datum of depth k. For $i \in [k]$, let b_i denote the IF-datum obtained by restricting b to $[i]$. Consider the iterated Frobenius equation

$$y_1^\sigma - y_1 = b(1), \quad y_2^\sigma - y_2 = b(2)y_1, \quad \ldots, \quad y_k^\sigma - y_k = b(k)y_{k-1}.$$

Note that we obtain a solution to this equation by setting $y_i = w_{b_i}$. This justifies the terminology "IF-solution".

3.2. ***Spaces generated by* IF-*solutions.*** We define

$$M_k := \left\{ \sum_{\substack{b \in D \\ \deg(b) \le k}} a_b w_b \;\middle|\; a_b \in \mathbb{Z}_q(x) \right\}, \quad \widehat{M}_k := p\text{-adic completion of } M_k \text{ in } \mathbb{Z}_q[\![x]\!]\langle x^{-1}\rangle,$$

$$A_k := \left\{ \sum_{\substack{b \in D \\ \deg(b) \le k}} a_b w_b \;\middle|\; a_b \in \mathbb{Z}_q[x] \right\}.$$

Here $\mathbb{Z}_q[\![x]\!]\langle x^{-1}\rangle$ denotes the p-adic completion of $\mathbb{Z}_q[\![x]\!][x^{-1}]$.

Lemma 3.5. *For $k \ge 0$, we have $\tau(xA_k) \subset A_{k+1}$. In particular, for any $z \in x\mathbb{Z}_q[x]$, we have $\tau(zA_k) \subset A_{k+1}$.*

Proof. We proceed by induction on k. For $k = 0$, we are reduced to showing $\tau(x_j^i) \in A_1$. Write $i = p^t h$, where h is coprime to p. Then

$$\tau(x_j^i) = \tau(x_{j-t}^h) - \sum_{m=1}^{t} x_{j-m}^{p^{t-m}h},$$

so $\tau(x_j^i) \in A_1$. For $k > 0$, it suffices to show $\tau(x_j^i w_b) \in A_{k+1}$ for any IF-datum b of depth k. Again we write $i = p^t h$. By repeatedly using (3), we obtain

$$\tau(x_j^i w_b) = \tau(x_{j-t}^h w_b) + \sum_{m=1}^{t} \tau(x_{j-m}^{p^{t-m}h} b(k) w_{b^{\mathrm{tr}}}) - \sum_{m=1}^{t} x_{j-m}^{p^{t-m}h} w_b.$$

Note that $\tau(x_{j-t}^h w_b)$ is an IF-solution of degree $k+1$, so it is contained in A_{k+1}. The second summation is clearly contained in A_k. Finally, by our inductive hypothesis, $\tau(x_{j-m}^{p^{t-m}h} b(k) w_{b^{\mathrm{tr}}})$ is contained in A_k. □

Lemma 3.6. *We have $\sigma(M_k) \subset M_k$ and $\sigma(\widehat{M}_k) \subset \widehat{M}_k$.*

Proof. Let b be an IF-datum of depth k. We have $w_b^\sigma = w_b + b(k) w_{b^{\mathrm{tr}}}$, which is in M_k. Thus, $\sigma(M_k) \subset M_k$. We see $\sigma(\widehat{M}_k) \subset \widehat{M}_k$ by taking p-adic limits. □

Lemma 3.7.

$$\frac{d}{dx}\widehat{M}_k \subset \widehat{M}_k.$$

Proof. We induct on k. When $k = 0$, the statement is immediate, since $\widehat{M}_0$ is the ring $\mathbb{Z}_q\langle x, x^{-1}\rangle$, which is preserved by $\frac{d}{dx}$. Let $k > 0$ and assume the lemma holds for $k-1$. From the definitions and the Leibnitz rule, we reduce to showing $\frac{d}{dx} w_b \in \widehat{M}_k$ whenever $\deg(b) = k$. Applying $\frac{d}{dx}$ to the equation $w_b^\sigma - w_b = b(k) w_{b^{\mathrm{tr}}}$ and using the chain rule gives

$$p x^{p-1} (w_b')^\sigma + w_b' = b(k)' w_{b^{\mathrm{tr}}} + b(k) w_{b^{\mathrm{tr}}}'.$$

We deduce that

$$w_b' = \sum_{n=0}^{\infty} p^n x^{p^{n+1}-1} (b(k)' w_{b^{\mathrm{tr}}} + b(k) w_{b^{\mathrm{tr}}}')^{\sigma^n}. \tag{4}$$

By our inductive hypothesis, we know $w_{b^{\mathrm{tr}}}' \in \widehat{M}_{k-1}$. Then, from Lemma 3.6, we see that each summand in (4) is in $\widehat{M}_{k-1}$. As the terms in the sum converge p-adically, the lemma follows. □

Remark 3.8. It is not true that $\frac{d}{dx} M_k \subset M_k$. For example, we see that $\frac{d}{dx}\tau(x) = \sum p^i x^{p^i-1}$ is not in M_1. This is the sole reason for introducing the p-adically closed spaces $\widehat{M}_k$.

Lemma 3.9. *For $k_1, k_2 \geq 0$, we have $M_{k_1} M_{k_2} \subset M_{k_1+k_2}$ and $\widehat{M}_{k_1} \widehat{M}_{k_2} \subset \widehat{M}_{k_1+k_2}$.*

Proof. We will show $A_{k_1} A_{k_2} \subset A_{k_1+k_2}$. The lemma will follow by inverting x and taking the p-adic completion. We proceed by induction on $k = k_1 + k_2$. The case $k = 0$ is immediate. Let $k > 0$ and assume the result holds for $k-1$. Let b_i be an IF-datum of degree k_i. Then, using (2), we have

$$w_{b_1} w_{b_2} = \tau(b_1(k_1) w_{b_1^{\mathrm{tr}}} w_{b_2}) + \tau(b_2(k_2) w_{b_2^{\mathrm{tr}}} w_{b_1}) - \tau(b_1(k_1) w_{b_1^{\mathrm{tr}}} b_2(k_2) w_{b_2^{\mathrm{tr}}}).$$

The lemma follows from our inductive hypothesis and Lemma 3.5. □

3.3. IF-*solutions modulo p.*

We now study the space of IF-solutions modulo p.

Definition 3.10. Let b be an IF-datum. We define $\overline{w}_b$ to be the reduction of w_b modulo p in $\mathbb{F}_q[\![x]\!]$.

We define the spaces

$$\overline{M}_k := \left\{ \sum_{\substack{b \in D \\ \mathrm{dep}(b) \leq k}} a_b \overline{w}_b \;\middle|\; a_b \in \mathbb{F}_q(x) \right\}, \quad \overline{M} = \bigcup_{n \geq 1} \overline{M}_k.$$

Our main result of this section is the following proposition.

Proposition 3.11. *The elements $\overline{w}_b$, where b ranges over $b \in D$, are linearly independent over $\mathbb{F}_q(x)$. In particular, the set $\{\overline{w}_b\}_{b \in D}$ is a basis of $\overline{M}$ over $\mathbb{F}_q(x)$.*

Proof. It will be useful for us to introduce an ordering on D. We set $b_1 > b_2$ if $\mathrm{dep}(b_1) > \mathrm{dep}(b_2)$ and, for IF-data with the same depth k, we use the lexicographical ordering on $(b(k), \ldots, b(1))$. Consider $c = \{c_b\}_{b \in D}$ with $c_b \in \mathbb{F}_q(x)$, where almost all $c_b = 0$. We define $\max(c)$ to be the largest b such that $c_b \neq 0$. Assume there exists such a c such that

$$\sum_{b \in D} c_b \overline{w}_b = 0. \tag{5}$$

We may further assume that $\max(c)$ is minimal in the sense that, for any other $c' = \{c_b'\}_{b \in D}$ with $\max(c') < \max(c)$, the corresponding sum $\sum c_b' \overline{w}_b$ is nonzero. Set $b_{\max} = \max(c)$ and $k_{\max} = \mathrm{dep}(b_{\max})$. After multiplying (5) with an element of $\mathbb{F}_q(x)$, we may assume $c_{b_{\max}} = 1$.

Let F be the p-th power Frobenius. For $b \in D$ with $\mathrm{dep}(b) = k$, we have

$$(F-1) c_b \overline{w}_b = c_b^p \overline{w}_b^p - c_b \overline{w}_b = (c_b^p - c_b) \overline{w}_b + c_b^p b(k) \overline{w}_{b^{\mathrm{tr}}}.$$

Thus, if we apply $F-1$ to (5), we get

$$\sum_{b\in D} d_b \overline{w}_b = 0,$$

where, for b of depth k,

$$d_b := (c_b^p - c_b) + \sum_{\substack{\beta\in D\\ \beta^{\mathrm{tr}}=b}} c_\beta \beta(k+1).$$

It is clear that $d_{b_{\max}} = 0$ and $\max(d) < \max(c)$. Thus, we are reduced to proving that not all the d_b are zero, as this will contradict our minimality condition on c.

First assume each $c_b \in \mathbb{F}_p$. Then

$$d_{b_{\max}^{\mathrm{tr}}} = \sum_{\substack{b\in D\\ b^{\mathrm{tr}}=b_{\max}^{\mathrm{tr}}}} c_b b(k_{\max}).$$

The terms $b_0(k_{\max})$ are all distinct elements of X. In particular, they are linearly independent over $\mathbb{F}_p$. Not all of the c_b are equal to zero, since we know $c_{b_{\max}} \neq 0$, so we see that $d_{b_{\max}^{\mathrm{tr}}} \neq 0$. Next, assume that not all the c_b are in $\mathbb{F}_p$. Let b_0 be the largest IF-datum such that $c_{b_0} \notin \mathbb{F}_p$ and let $k_0 = \mathrm{dep}(b_0)$. If $d_{b_0} = 0$, we have

$$c_{b_0}^p - c_{b_0} = -\sum_{\substack{\beta\in D\\ \beta^{\mathrm{tr}}=b_0}} c_\beta \beta(k_0+1).$$

The left side of this equation is zero, as $c_{b_0} \notin \mathbb{F}_p$. However, the right side is an $\mathbb{F}_p$-linear combination of terms of the form $x_j^i = \zeta^{p^j} x^i$, where $p \nmid i$. In particular, the right side is a polynomial $f(x) \in \mathbb{F}_q[x]$ whose degree is coprime to p. However, the equation $y^p - y = f(x)$ cannot have a solution in $\mathbb{F}_q(x)$, which gives a contradiction. □

Corollary 3.12. *The space M_k (resp. $\widehat{M}_k$) is freely generated by $\{w_b\}_{\mathrm{dep}(b)\le k}$ over $M_0 = \mathbb{Z}_q(x)$ (resp. $\widehat{M}_0 = \mathbb{Z}_q\langle x, x^{-1}\rangle$).*

Corollary 3.13. *For all $k, m \ge 0$, we have $p^m \mathbb{Z}_q((x)) \cap \widehat{M}_k = p^m \widehat{M}_k$. In particular, let $x, y \in \widehat{M}_{k+1}$. If $p^m x \equiv p^m y \bmod \widehat{M}_k$, then $x \equiv y \bmod \widehat{M}_k$.*

In light of Proposition 3.11, we make the following definitions.

Definition 3.14. Let $f \in \overline{M}$. By Proposition 3.11, we may express f uniquely as a sum $\sum a_b \bar{b}$, where $a_b \in \mathbb{F}_q(x)$ and almost all the a_b are zero. We define the depth of f, written as $\mathrm{dep}(f)$, to be the largest k such that, for some b with $\mathrm{dep}(b) = k$, we have $a_b \neq 0$. Equivalently, the depth of $f \in \overline{M}$ is the unique $k \ge 0$ such that $f \in \overline{M}_k$ and $f \notin \overline{M}_{k-1}$.

Corollary 3.15. *Let $\alpha_1, \alpha_2, \ldots$ be a sequence of elements in $\overline{M}$ with $\mathrm{dep}(\alpha_k) = k$ for each k. Then the α_k are linearly independent over $\mathbb{F}_p(x)$.*

Definition 3.16. Let $\{\dot{w}_b\}_{b\in D}$ be the basis of $\mathrm{Hom}_{\mathbb{F}_p(x)}(\overline{M}, \mathbb{F}_p(x))$ that is dual to $\{\overline{w}_b\}_{b\in D}$. That is, for $b, b' \in D$, $\dot{w}_b(\overline{w}_{b'})$ is one if $b = b'$ and zero otherwise.

3.4. ***Products of* IF-*solutions.*** Let $b_1, \dots, b_r$ be IF-data with $\mathrm{dep}(b_i) = k_i$. Let $k = k_1 + \cdots + k_r$. From Lemma 3.9, we know $w_{b_1} \cdots w_{b_r}$ is in M_k, and thus, by Corollary 3.12, can be uniquely written as a sum $\sum_{\mathrm{dep}(b) \leq k} a_b w_b$ with $a_b \in M_0$. In this subsection, we determine a_b when $\mathrm{dep}(b) = k$. In particular, in Proposition 3.21, we completely determine $w_{b_1} \cdots w_{b_r}$ modulo M_{k-1}.

Definition 3.17. Let $k_1, \dots, k_r \geq 1$ and set $k = k_1 + \cdots + k_r$. We define

$$[k_1, \dots, k_r] := \{(i, j) \mid 1 \leq i \leq r,\ j \in [k_i]\}.$$

A splicing function is a bijection $\eta : [k] \to [k_1, \dots, k_r]$ such that, for any $1 \leq j_1 < j_2 \leq k_i$, we have $\eta^{-1}(i, j_1) < \eta^{-1}(i, j_2)$. We let $C(k_1, \dots, k_r)$ denote the set of all splicing functions. Note that $C(k_1, \dots, k_r)$ has cardinality $\binom{k}{k_1, \dots, k_r}$.

Definition 3.18. Let $b_1, \dots, b_r$ be IF-data with $\mathrm{dep}(b_i) = k_i$. For $\eta \in C(k_1, \dots, k_r)$, we define the splicing of $b_1, \dots, b_r$ according to η to be the IF-datum of depth k defined by the composition

$$b_\eta : [k] \xrightarrow{\eta} [k_1, \dots, k_r] \xrightarrow{(i,j) \mapsto b_i(j)} X.$$

We let $S(b_1, \dots, b_r)$ denote the multiset of all possible splicings of $b_1, \dots, b_r$. That is,

$$S(b_1, \dots, b_r) := \{b_\eta \mid \eta \in C(k_1, \dots, k_r)\}.$$

The next lemma follows immediately from these definitions.

Lemma 3.19. *Let b_i and k_i be as in Definition 3.18. Then*

$$S(b_1, \dots, b_r) = \bigsqcup_{i=1}^{r} \{\mathrm{cat}(b, b_i(k_i)) \mid b \in S(b_1, \dots, b_{i-1}, b_i^{\mathrm{tr}}, b_{i+1}, \dots, b_r)\},$$

where the $\sqcup$ denotes a disjoint union.

Lemma 3.20. *Continuing with the same notation, we define*

$$\mathcal{B}_r := \prod_{i=1}^{r} w_{b_i} \quad \text{and} \quad \mathcal{B}_r^{(j)} := w_{b_j^{\mathrm{tr}}} \prod_{\substack{1 \leq i \leq r \\ i \neq j}} w_{b_i}.$$

Then we have

$$\mathcal{B}_r = \sum_{i=1}^{r} \tau(b_i(k_i) \mathcal{B}_r^{(i)}) \mod M_{k-1}.$$

Proof. We proceed by induction on r, with the base case $r = 1$ being immediate. Let $r > 1$ and assume the lemma holds for $r - 1$. In particular, we have

$$\mathcal{B}_{r-1} \equiv \sum_{i=1}^{r-1} \tau(b_i(k_i) \mathcal{B}_{r-1}^{(i)}) \mod M_{k-k_r-1}. \tag{6}$$

We know $M_{k-k_r-1} M_{k_r} \subset M_{k-1}$ by Lemma 3.9. Thus, multiplying (6) by w_{b_r} gives

$$\mathcal{B}_r \equiv w_{b_r} \sum_{i=1}^{r-1} \tau(b_i(k_i) \mathcal{B}_{r-1}^{(i)}) \mod M_{k-1}. \tag{7}$$

For $1 \le i \le r-1$, we know by (2) that

$$w_{b_r}\tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)}) = \tau\big(b_r(k_r)w_{b_r^{\mathrm{tr}}}\tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)})\big) + \tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)}w_{b_r}) - \tau(b_r(k_r)w_{b_r^{\mathrm{tr}}}b_i(k_i)\mathcal{B}_{r-1}^{(i)}). \quad (8)$$

We know $\mathcal{B}_{r-1}^{(i)} \in M_{k-k_r-1}$ and $w_{b_r^{\mathrm{tr}}} \in M_{k_r-1}$, so that the last term in the right side of (8) is in M_{k-1}. As $\mathcal{B}_{r-1}^{(i)}w_{b_r} = \mathcal{B}_r^{(i)}$, we have

$$w_{b_r}\tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)}) \equiv \tau\big(b_r(k_r)w_{b_r^{\mathrm{tr}}}\tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)})\big) + \tau(b_i(k_i)\mathcal{B}_r^{(i)}) \mod M_{k-1}. \quad (9)$$

Next, by our inductive hypothesis (6) and the linearity of τ, we obtain

$$\sum_{i=1}^{r-1}\tau\big(b_r(k_r)w_{b_r^{\mathrm{tr}}}\tau(b_i(k_i)\mathcal{B}_{r-1}^{(i)})\big) \equiv \tau(b_r(k_r)w_{b_r^{\mathrm{tr}}}\mathcal{B}_{r-1}) \mod M_{k-1}.$$

We can combine this with (9) and (7) to get

$$\mathcal{B}_r \equiv \tau(b_r(k_r)w_{b_r^{\mathrm{tr}}}\mathcal{B}_{r-1}) + \sum_{i=1}^{r-1}\tau(b_i(k_i)\mathcal{B}_r^{(i)}) \mod M_{k-1}.$$

The result follows by observing that $\mathcal{B}_r^{(r)} = w_{b_r^{\mathrm{tr}}}\mathcal{B}_{r-1}$. □

Proposition 3.21. *Continue with the notation from Lemma 3.20. We have*

$$\mathcal{B}_r \equiv \sum_{b\in S(b_1,\dots,b_r)} w_b \mod M_{k-1}.$$

Proof. We induct on r. When $r = 1$, the result is immediate. Let $r > 1$ and assume the result holds for $r-1$. From Lemma 3.20, we know

$$\mathcal{B}_r = \sum_{i=1}^{r}\tau(b_i(k_i)\mathcal{B}_r^{(i)}) \mod M_{k-1}.$$

By our inductive hypothesis, we have

$$\mathcal{B}_r^{(i)} \equiv \sum_{b\in S(b_1,\dots,b_{i-1},b_i^{\mathrm{tr}},b_{i+1},\dots,b_r)} w_c \mod M_{k-2}.$$

The proposition then follows from Lemmas 3.19 and 3.5. □

We will now give some corollaries of Proposition 3.21. First we need a definition.

Definition 3.22. Let $k > 0$ and $z \in X$. Let β_z^k be the IF-datum of depth k such that $\beta_z^k(i) = z$ for each $1 \le i \le k$. We define ω_z^k to be $w_{\beta_z^k}$. We let $\bar{\omega}_z^k$ be the reduction of ω_z^k modulo p and we let $\dot{\omega}_z^k$ be $\dot{w}_{\beta_z^k}$, as defined in Definition 3.16.

We readily compute that

$$\omega_z^k = \sum_{i_k\ge\cdots\ge i_1\ge 0} z^{bp^{i_1+\cdots+i_k}}.$$

The elements ω_z^k and $\bar{\omega}_z^k$ will appear when studying the higher derivatives of the Artin–Hasse exponential. We have the following corollaries of Proposition 3.21.

Corollary 3.23. *Fix $z \in X$. Let $k_1, \dots, k_r \geq 1$ and set $k = k_1 + \cdots + k_r$. Then*

$$\prod_{i=1}^{r} \omega_z^{k_i} \equiv \binom{k}{k_1, \dots, k_r} \omega_z^k \mod M_{k-1}.$$

Proof. This follows from Proposition 3.21 and the fact that there are $\binom{k}{k_1,\dots,k_r}$ splicing functions. □

Corollary 3.24. *For $z \in X$, we have $\tau(z)^k \equiv k! \omega_z^k \mod M_{k-1}$.*

Corollary 3.25. *Let $b_1, \dots, b_r \in D$, let $k_i = \operatorname{dep}(b_i)$, and set $k = k_1 + \cdots + k_r$. Then, for any $z \in X$, we have*

$$\dot{\omega}_z^k(\overline{w}_{b_1} \cdots \overline{w}_{b_r}) = \begin{cases} \binom{k}{k_1,\dots,k_r} & \text{if each } \overline{w}_{b_i} = \bar{\omega}_z^{k_i}, \\ 0 & \text{otherwise.} \end{cases}$$

4. Higher derivatives and the Artin–Hasse exponential

4.1. ***Higher derivatives and a transcendental criterion.*** We define the *higher derivatives* on $\mathbb{F}_q((x))$ and $\mathbb{Z}_q((x))$,

$$\partial^{(k)} := \frac{1}{k!} \frac{d^k}{d^k x}.$$

We define the *higher logarithmic derivatives* of $f \in \mathbb{F}_q((x))$ or $f \in \mathbb{Z}_q((x))$ by

$$\mathrm{dLog}^{(k)}(f) := \frac{\partial^{(k)}(f)}{f}.$$

Note that $\partial^{(k)}$ and $\mathrm{dLog}^{(k)}$ restrict to maps on $\mathbb{F}_q(x)$. Furthermore, if K is a finite separable extension of $\mathbb{F}_q(x)$, then each $\partial^{(k)}$ and $\mathrm{dLog}^{(k)}$ extends uniquely to K by a theorem of Schmidt (see [Matzat and van der Put 2003, Section 2]). We begin with the following general transcendence result.

Theorem 4.1. *Let $\alpha_1, \alpha_2, \dots$ be a sequence of power series in $\mathbb{F}_q[\![x]\!]$. Let $f(x) \in \mathbb{F}_q[\![x]\!]$ be a power series satisfying the iterative differential equation*

$$\mathrm{dLog}^{(k)}(f) = \alpha_k \quad \text{for all } k \geq 1.$$

If $\mathbb{F}_q(x)(\alpha_1, \alpha_2, \dots)$ is an infinite extension of $\mathbb{F}_q(x)$, then $f(x)$ is transcendental over $\mathbb{F}_q(x)$.

Proof. Assume $f(x)$ is algebraic (necessarily separable) and contained in a field K. As $\mathrm{dLog}^{(k)}$ extends to K, we see that α_n is contained in K as well. This contradicts our assumption that $\mathbb{F}_q(x)(\alpha_1, \alpha_2, \dots)$ is an infinite extension of $\mathbb{F}_q(x)$. □

Corollary 4.2. *Let $\alpha_1, \alpha_2, \dots$ be a sequence of power series contained in $\overline{M}$ with $\operatorname{dep}(\alpha_k) = k$ for each k. Let $f(x) \in \mathbb{F}_q[\![x]\!]$ be a power series satisfying $\mathrm{dLog}^{(k)}(f) = \alpha_k$ for all $k \geq 1$. Then $f(x)$ is transcendental.*

Proof. From Corollary 3.15, we know that the α_k are linearly independent over $\mathbb{F}_q(x)$, so they cannot be contained in a finite extension. Then use Theorem 4.1. □

4.2. ***Estimating*** $\mathbf{dLog}^{(k)}(E_p(x))$**.** Let L denote the differential operator $\frac{d}{dx} - \tau(x)/x$ on $\mathbb{Z}_q((x))$. Using the description $E_p(x) = \exp\left(\sum_{n=0}^{\infty} x^{p^n}/p^n\right)$, we see that, for $k \geq 1$,

$$k! \cdot \mathrm{dLog}^{(k)}(E_p(x)) = L^{k-1}\left(\frac{\tau(x)}{x}\right). \tag{10}$$

We have the following proposition.

Proposition 4.3. *For $k \geq 1$, we have*

$$k! \cdot \mathrm{dLog}^{(k)}(E_p(x)) \equiv \frac{\tau(x)^k}{x^k} \mod \widehat{M}_{k-1}.$$

Proof. We proceed by induction on k. When $k = 1$, the result is immediate. Assume the result holds for k. From (10) and our inductive hypothesis, we have

$$L^k\left(\frac{\tau(x)}{x}\right) = L\left(\frac{\tau(x)^k}{x^k} + c\right),$$

where $c \in \widehat{M}_{k-1}$. We know $\frac{d}{dx}(\tau(x)^k/x^k + c) \in \widehat{M}_k$ from Lemmas 3.7 and 3.9. We also know $\tau(x)/xc \in \widehat{M}_k$ by Lemma 3.9. Thus,

$$L\left(\frac{\tau(x)^k}{x^k} + c\right) \equiv \frac{\tau(x)^{k+1}}{x^{k+1}} \mod \widehat{M}_k.$$

The proposition follows from (10). □

Corollary 4.4. *For $k \geq 1$, we have* $\mathrm{dLog}^{(k)}(\bar{E}_p(x)) \equiv \bar{\omega}_x^k/x^k \mod \overline{M}_{k-1}$.

Proof. From Corollary 3.24 and Proposition 4.3, we know $k! \cdot \mathrm{dLog}^{(k)}(E_p(x)) \equiv k!\omega_x^k/x^k \mod \widehat{M}_{k-1}$. By Corollary 3.13, we see that $\mathrm{dLog}^{(k)}(E_p(x)) \equiv \omega_x^k/x^k \mod \widehat{M}_{k-1}$. The corollary follows by reducing modulo p. □

At this point we may establish the transcendence of $\bar{E}_p(x)$.

Theorem 4.5. *The series $\bar{E}_p(x)$ is transcendental over $\mathbb{F}_p(x)$.*

Proof. Set $\alpha_k = \mathrm{dLog}^{(k)}(\bar{E}_p(x))$. From Corollary 4.4, we see that $\alpha_k \in \overline{M}$ and $\mathrm{dep}(\alpha_k) = k$. The theorem follows from Corollary 4.2. □

4.3. ***Estimating*** $\mathbf{dLog}^{(k)}[\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}]$**.** Recall from the introduction that, for $f = \sum c_i x^i \in x\mathbb{F}_q[x]$, we define

$$f^* = \sum_{\substack{i \geq 0 \\ p \nmid i}} c_i^*(f) x^i \quad \text{with} \quad c_i^*(f) = \sum_{k \geq 0} c_{ip^k}^{1/p^k}.$$

We can express $c_i^*(f)$ uniquely as $\sum_{j=0}^{f-1} c_{x_j^i}^* \zeta^{p^j}$, where $c_{x_j^i}^* \in \mathbb{F}_p$. In particular, we have

$$f^* = \sum_{z \in X} c_z^* z.$$

In this subsection we prove the following proposition.

Proposition 4.6. *Let $k \geq 0$ and let $n_1, \dots, n_r \geq 0$. Let $f_1, \dots, f_r \in x\mathbb{F}_q[x]$. Then*

$$\Delta_k := \mathrm{dLog}^{(k)}[\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}]$$

is in $\overline{M}_k$ and, for any $z \in X$, we have

$$x^k \dot{\omega}_z^k\big(\mathrm{dLog}^{(k)}[\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}]\big) = \left(\sum_{i=1}^{r} n_i c_z^*(f_i)\, \mathrm{dLog}^{(1)}(f_i)\right)^k.$$

In particular, Δ_k has depth exactly k if there exists $z \in X$ such that $\sum_{i=1}^r n_i c_z^*(f_i)\, \mathrm{dLog}^{(1)}(f_i)$ is nonzero.

The proof of Proposition 4.6 is broken into several steps. The main ingredients are Corollary 4.4 and the following well-known identities on higher derivatives:

$$\partial^{(k)}(f_1 \cdots f_n) = \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=k}} \partial^{(k_1)}(f_1) \cdots \partial^{(k_n)}(f_n), \tag{11}$$

$$\partial^{(k)}(f(g)) = \sum_{j=1}^{k} \partial^{(j)}(f)(g) \cdot \sum_{\substack{k_1,\dots,k_j \geq 1 \\ k_1+\cdots+k_j=k}} \partial^{(k_1)}(g) \cdots \partial^{(k_j)}(g). \tag{12}$$

Lemma 4.7. *Let n be any p-adic integer. We have*

$$\mathrm{dLog}^{(k)}(\bar{E}_p(x)^n) \equiv n^k \frac{\bar{\omega}_x^k}{x^k} \mod \overline{M}_{k-1}.$$

Proof. First assume n is a positive integer. By Corollaries 4.4 and 3.23 and (11), we have

$$\begin{aligned}
\mathrm{dLog}^{(k)}(\bar{E}_p(x)^n) &\equiv \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=k}} \frac{1}{x^k} \bar{\omega}_x^{k_1} \cdots \bar{\omega}_x^{k_n} \mod \overline{M}_{k-1} \\
&\equiv \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=k}} \binom{k}{k_1, \dots, k_n} \frac{\bar{\omega}_x^k}{x^k} \mod \overline{M}_{k-1} \\
&\equiv n^k \frac{\bar{\omega}_x^k}{x^k} \mod \overline{M}_{k-1}.
\end{aligned}$$

For the general result, write $n = a + p^m b$, where a is a positive integer. If $m > v_p(k)$, we have $\mathrm{dLog}^{(k)}(\bar{E}_p(x)^n) = \mathrm{dLog}^{(k)}(\bar{E}_p(x)^a)$, since $\partial^{(k)}(f^{p^m}) = 0$. The general result follows. □

Corollary 4.8. *For $f \in x\mathbb{F}_q[x]$, we have*

$$\mathrm{dLog}^{(k)}(\bar{E}_p(f)^n) \equiv n^k \frac{\bar{\omega}_x^k(f)}{x^k} (\mathrm{dLog}^{(1)} f)^k \mod \overline{M}_{k-1}.$$

Proof. From Lemma 4.7, we see that $\mathrm{dLog}^{(i)}(\bar{E}_p(x)^n) \in \overline{M}_{k-1}$ for $1 \leq i < k$. Note that, for $g \in \overline{M}_{k-1}$ and $h \in x\mathbb{F}_q[x]$, we have $g \circ h \in \overline{M}_{k-1}$, so $\mathrm{dLog}^{(i)}(\bar{E}_p(x)^n) \circ f \in \overline{M}_{k-1}$ for $1 \leq i < k$. The corollary follows from (12) and Lemma 4.7. □

Lemma 4.9. *Let $f \in x\mathbb{F}_q[x]$. Then*

$$\bar{\omega}_x^k(f) \equiv \sum_{\substack{b\in D\\ \deg(b)=k}} c_b^*(f)\bar{w}_b \mod \overline{M}_{k-1}, \quad \text{where } c_b^*(f) = \prod_{i=1}^{k} c_{b(i)}^*(f).$$

Proof. First note that $\bar{\omega}_x^k(f) \equiv \bar{\omega}_x^k(f^*) \bmod \overline{M}_{k-1}$. This can be proven by inducting on k and using (3). We are therefore reduced to the case where $f = f^*$. The result follows from the $\mathbb{F}_p$-linearity of τ modulo p and unraveling the nested τ's. □

Corollary 4.10. *For all $z \in X$, we have $\dot{\omega}_z^k(\bar{\omega}_x^k(f)) = c_z^*(f)^k$.*

Lemma 4.11. *Let $k_1, \dots, k_r \geq 1$ with $k_1 + \cdots + k_r = k$. Then*

$$\dot{\omega}_z^k\big(\bar{\omega}_x^{k_1}(f_1(x)) \cdots \bar{\omega}_x^{k_r}(f_r(x))\big) = \binom{k}{k_1, \dots, k_r} \prod_{i=1}^{r} c_z^*(f_i)^{k_i}.$$

Proof. This follows from Corollaries 4.10 and 3.25. □

Proof of Proposition 4.6. From (11) and Corollary 4.8, we have

$$\begin{aligned} x^k \operatorname{dLog}^{(k)}[\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}] &= \sum_{\substack{k_1,\dots,k_r\geq 0\\ k_1+\cdots+k_r=k}} \prod_{i=1}^{r} \operatorname{dLog}^{(k_i)}(\bar{E}_p(f_i)^{n_i}) \\ &\equiv \sum_{\substack{k_1,\dots,k_r\geq 0\\ k_1+\cdots+k_r=k}} \prod_{i=1}^{r} n_i^{k_i} (\operatorname{dLog}^{(1)}(f_i))^{k_i} \bar{\omega}_x^{k_i}(f_i) \mod \overline{M}_{k-1}. \end{aligned}$$

Applying Lemma 4.11 to $\prod_{i=1}^{r} \bar{\omega}_x^{k_i}(f_i)$ gives

$$\begin{aligned} x^k \dot{\omega}_z^k\big(\operatorname{dLog}^{(k)}[\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}]\big) &= \sum_{\substack{k_1,\dots,k_r\geq 0\\ k_1+\cdots+k_r=k}} \binom{k}{k_1, \dots, k_r} \prod_{i=1}^{s} [n_i \operatorname{dLog}^{(1)}(f_i) c_z^*(f_i)]^{k_i} \\ &= \Big(\sum_{i=1}^{r} n_i \operatorname{dLog}^{(1)}(f_i) c_z^*(f_i) \Big)^k. \end{aligned}$$

□

5. Algebraic independence results

In this section we prove Theorems 1.3 and 1.7.

5.1. ***Iterative differential Galois theory and transcendence.*** We will make use of iterative differential (ID for short) Galois theory, as expounded on by Matzat and van der Put [2003] (in particular, Sections 3–4 therein). Let K be the algebraic closure of $\mathbb{F}_q(x)$ in $\overline{\mathbb{F}}_p((x))$. The higher derivatives $\partial^{(k)}$ make K into an ID-field. For any $f \in x\mathbb{F}_q[x]$ we recursively define a sequence $A_{f,0}, A_{f,1}, \dots \in K$ by

$$A_{f,0} = 1 \quad \text{and} \quad A_{f,k} = -\sum_{i=0}^{k-1} A_{f,i} \operatorname{dLog}^{(k-i)}(\bar{E}_p(f)).$$

We then define an ID-module $N_f = Ke_f$ over K of rank one by the rule

$$\partial^{(k)}(e_f) = A_{f,k}e_f.$$

Note that N_f trivializes over the ring $K[\bar{E}_p(f), 1/\bar{E}_p(f)]$. For $f_1, \dots, f_r \in x\mathbb{F}_q[x]$, we define the rank r ID-module

$$N := \bigoplus_{i=1}^{r} N_{f_1}.$$

Then N has a full system of solutions over the ring

$$R_0 := K\left[\bar{E}_p(f_1), \dots, \bar{E}_p(f_r), \frac{1}{\bar{E}_p(f_1)\cdots\bar{E}_p(f_r)}\right].$$

The Picard–Vessiot ring of N is $R = R_0/I$, where I is a maximal ID-ideal (i.e., I is a maximal element in the ordered set of ideals of R_0 satisfying $\partial^{(k)}(I) \subset I$). The ID-Galois group $\mathcal{G}(N)$ is the group of ID-automorphisms of R that fix K. A key fact is that $\mathcal{G}(N)$ is a reduced algebraic group over the field of constants $\bar{\mathbb{F}}_p$ and that $\operatorname{Spec}(R)$ is a $\mathcal{G}(N)$-torsor over K. In particular, we have

$$\operatorname{tr.deg}_K(\operatorname{Frac}(R_0)) \geq \operatorname{tr.deg}_K(\operatorname{Frac}(R)) = \dim_{\bar{\mathbb{F}}_p}(\mathcal{G}(N)).$$

Thus, to prove the algebraic independence of the $\bar{E}_p(f_i)$, it suffices to show $\mathcal{G}(N)$ has dimension r.

On the other hand, there is a Tannakian interpretation of the ID-Galois group. Let ID_K denote the category of ID-modules over K. Then ID_K is an $\bar{\mathbb{F}}_p$-linear tensor category. Let $[N]$ be the full subcategory of ID_K generated by tensor powers of N and its dual. Then $[N]$ is again an $\bar{\mathbb{F}}_p$-linear tensor category and thus isomorphic to the category of representations for an affine algebraic group $\mathrm{Gal}(N)$ defined over $\bar{\mathbb{F}}_p$. By standard arguments (see [van der Put and Singer 2003] for the characteristic 0 case and [Matzat and van der Put 2003] for our precise situation), we have

$$\mathcal{G}(N) \cong \mathrm{Gal}(N).$$

Note that, since N_{f_i} has rank one, we know $\mathrm{Gal}(N_{f_i})$ is a closed subgroup of $\mathbb{G}_m$. In particular, we see that $\mathrm{Gal}(N)$ is a closed subgroup of $\mathbb{G}_m^r$. To prove that $\mathrm{Gal}(N) \cong \mathbb{G}_m^r$, it is enough to show that, for any $n_1, \dots, n_r \in \mathbb{Z}$ that are not all zero, the ID-module $N_{f_1}^{\otimes n_1} \otimes \cdots \otimes N_{f_r}^{\otimes n_r}$ is nontrivial. This amounts to proving $\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}$ is transcendental over K. Furthermore, we easily reduce to the case where the n_i are not all divisible by p.

5.2. ***Proof of Theorem 1.3.*** Let $f_1, \dots, f_r$ satisfy one of the two conditions from Theorem 1.3. From the discussion in Section 5.1 we must show $\bar{E}_p(f_1)^{n_1} \cdots \bar{E}_p(f_r)^{n_r}$ is transcendental, where not all the n_i are divisible by p. Using Corollary 4.2 and Proposition 4.6, we reduce to showing that, for some $z \in X$, we have

$$\sum_{i=1}^{r} n_i c_z^*(f_i)\, \mathrm{dLog}^{(1)}(f_i) \neq 0.$$

The case where the f_i^* are not all zero and no nontrivial power product of the f_i are contained in $\mathbb{F}_q(x)^p$. We have

$$\sum_{i=1}^{r} n_i c_z^*(f_i)\,\mathrm{dLog}^{(1)}(f_i) = \mathrm{dLog}^{(1)}\Bigg[\prod_{i=1}^{r} f_i^{n_i c_z^*(f_i)}\Bigg].$$

By our assumption, the product in the $\mathrm{dLog}^{(1)}$ is not in $\mathbb{F}_q(x)^p$. The theorem follows by observing that the kernel of $\mathrm{dLog}^{(1)}$ on $\mathbb{F}_q(x)$ is $\mathbb{F}_q(x)^p$.

The case where $f_1^*, \ldots, f_r^*$ are linearly independent. Without loss of generality, assume $p\nmid n_1$. We may also assume that none of the f_i are p-th powers. Let $g \in \mathbb{F}_q[x]$ be an irreducible polynomial dividing f_1 and let v_g denote the valuation associated to g. Write $f_i = g^{b_i} h_i$, where $g\nmid h_i$. We can find g such that $p\nmid b_1$, as otherwise f_1 would be a p-th power. Since $f_1^*, \ldots, f_r^*$ are linearly dependent over $\mathbb{F}_p$,

$$n_1 b_1 f_1^* + \cdots + n_r b_r f_r^* \neq 0.$$

In particular, there exists $z \in X$ such that

$$D_z := n_1 b_1 c_z^*(f_1) + \cdots + n_r b_r c_z^*(f_r) \neq 0.$$

Then

$$\sum_{i=1}^{r} n_i c_z^*(f_i)\,\mathrm{dLog}^{(1)}(f_i) = D_z\,\mathrm{dLog}^{(1)}(g) + \sum_{i=1}^{r} n_i c_z^*(f_i)\,\mathrm{dLog}^{(1)}(h_i).$$

If α is a root of g, we see that $D_z\,\mathrm{dLog}^{(1)}(g)$ has a simple pole at α, while $\mathrm{dLog}^{(1)}(h_i)$ is regular at α. In particular, the right side of this sum has a simple pole at α and thus is nonzero.

5.3. ***Proof of Theorem 1.7.*** We now prove Theorem 1.7. Let ξ be a generator of $\mathbb{F}_p^\times$. Let $[\xi]$ denote the Teichmüller lift of ξ. Set $r = \phi(p-1)$. First, we show that $\bar{E}_p(x), \ldots, \bar{E}_p(\xi^{r-1}x)$ are algebraically independent. By the ID-Galois group argument in Section 5.1, it suffices to show

$$A(x) := \prod_{i=0}^{r-1} \bar{E}_p(\xi^i x)^{n_i}$$

is transcendental, where $n_0, \ldots, n_{r-1}$ are integers not all divisible by p. Indeed, this implies that the ID-Galois group of the ID-module $N_x \oplus N_{\xi x} \oplus \cdots \oplus N_{\xi^{r-1}x}$ is all of $\mathbb{G}_m^r$. From Lemma 2.3, we have

$$A(x) = \bar{E}_p(x)^{n_0 + n_1[\xi] + \cdots + n_{r-1}[\xi]^{r-1}}.$$

As the $1, \ldots, [\xi]^{r-1}$ are independent over $\mathbb{Z}$, we have $A(x) = \bar{E}_p(x)^n$, where n is a nonzero p-adic number. Thus, it suffices to show $\bar{E}_p(x)^n$ is transcendental for any nonzero p-adic number with $p\nmid n$. By Lemma 4.7, we see that $\mathrm{dLog}^{(k)}(\bar{E}_p(x)^n)$ has depth k. It follows from Corollary 4.2 that $\bar{E}_p(x)^n$ is transcendental over $\mathbb{F}_p(x)$. Thus, we have proven

$$\mathrm{tr.deg}_{\mathbb{F}_p(x)}\,\mathbb{F}_p(x)(\bar{E}_p(\xi^i x))_{0\le i\le r-1} = r.$$

For $j > r-1$, we can write $[\xi]^j = c_0 + c_1[\xi] + \cdots + c_{r-1}[\xi]^{r-1}$, where the c_i are integers. Then Lemma 2.3 gives

$$\bar{E}_p(\xi^j x) = \prod_{i=0}^{r-1} \bar{E}_p(\xi^i x)^{c_i},$$

so

$$\operatorname{tr.deg}_{\mathbb{F}_p(x)} \mathbb{F}_p(x)(\bar{E}_p(\xi^i x))_{0 \le i \le p-2} = r.$$

Acknowledgements

The author wishes to thank Dinesh Thakur for his enthusiasm and comments on some very preliminary versions of this article. We also thank the reviewers for helpful comments and for providing some historical context for this article.

References

[Adolphson and Sperber 1989] A. Adolphson and S. Sperber, "Exponential sums and Newton polyhedra: cohomology and estimates", *Ann. of Math.* (2) **130**:2 (1989), 367–406. MR

[Allouche and Shallit 2003] J.-P. Allouche and J. Shallit, *Automatic sequences: theory, applications, generalizations*, Cambridge Univ. Press, 2003. MR

[Allouche et al. 1988] J.-P. Allouche, M. Mendès France, and A. J. van der Poorten, "Indépendance algébrique de certaines séries formelles", *Bull. Soc. Math. France* **116**:4 (1988), 449–454. MR

[Ax 1971] J. Ax, "On Schanuel's conjectures", *Ann. of Math.* (2) **93** (1971), 252–268. MR

[Christol 1979] G. Christol, "Ensembles presque periodiques k-reconnaissables", *Theoret. Comput. Sci.* **9**:1 (1979), 141–145. MR

[Christol et al. 1980] G. Christol, T. Kamae, M. Mendès France, and G. Rauzy, "Suites algébriques, automates et substitutions", *Bull. Soc. Math. France* **108**:4 (1980), 401–419. MR

[Dieudonné 1957] J. Dieudonné, "On the Artin–Hasse exponential series", *Proc. Amer. Math. Soc.* **8** (1957), 210–214. MR

[Kolchin 1968] E. R. Kolchin, "Algebraic groups and algebraic dependence", *Amer. J. Math.* **90** (1968), 1151–1164. MR

[Matzat and van der Put 2003] B. H. Matzat and M. van der Put, "Iterative differential equations and the Abhyankar conjecture", *J. Reine Angew. Math.* **557** (2003), 1–52. MR

[Ostrowski 1946] A. Ostrowski, "Sur les relations algébriques entre les intégrales indéfinies", *Acta Math.* **78** (1946), 315–318. MR

[Papanikolas 2008] M. A. Papanikolas, "Tannakian duality for Anderson–Drinfeld motives and algebraic independence of Carlitz logarithms", *Invent. Math.* **171**:1 (2008), 123–174. MR

[van der Put and Singer 2003] M. van der Put and M. F. Singer, *Galois theory of linear differential equations*, Grundl. Math. Wissen. **328**, Springer, 2003. MR

[Thakur 2020] D. S. Thakur, "Automata methods in transcendence", pp. 351–372 in *t-motives: Hodge structures, transcendence and other motivic aspects*, edited by G. Böckle et al., Eur. Math. Soc., Berlin, 2020. MR

[Wan 1993] D. Q. Wan, "Newton polygons of zeta functions and L functions", *Ann. of Math.* (2) **137**:2 (1993), 249–293. MR

Communicated by Michael Singer
Received 2024-05-25 Revised 2025-05-05 Accepted 2025-06-27

jjk221@lehigh.edu *Department of Mathematics, Lehigh University, Bethlehem, PA, United States*

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1615

Anticyclotomic Euler system over biquadratic fields

Kim Tuan Do

We construct a new Euler system (anticyclotomic, in the sense of Jetchev–Nekovář–Skinner) for the Galois representation $V_{f,\chi}$ attached to a newform f of weight $k \geq 2$ twisted by an anticyclotomic Hecke character χ defined over an imaginary biquadratic field K_0. We then show some arithmetic applications of the constructed Euler system, including results on the Bloch–Kato conjecture and a divisibility towards the Iwasawa–Greenberg main conjecture for $V_{f,\chi}$.

1. Introduction 1615
2. Preliminaries 1618
3. The construction 1621
4. Selmer groups 1629
5. Triple product p-adic L-function and Selmer group 1631
6. Arithmetic applications 1636
Acknowledgements 1645
References 1645

1. Introduction

Let $f = \sum_{n=1}^{\infty} a_n q^n \in S_k(\Gamma_0(N_f))$ be an elliptic newform of even weight $k = 2r \geq 2$, and let $p \nmid 6N_f$ be a prime. Let $K_0/\mathbb{Q}$ be an imaginary biquadratic field in which p splits. This means that K_0 contains two distinct imaginary quadratic subfields K_1, K_2 together with one real quadratic subfield K_3. Let L be a number field containing K_0 and the Fourier coefficients of f, and let $\mathfrak{P}$ be a prime of L above p at which f is ordinary, i.e., $v_{\mathfrak{P}}(a_p) = 0$. Let χ be an anticyclotomic Hecke character of K_0 with infinity type $(-a, a, -b, b)$, where $a \geq b \geq 0$,[1] that satisfies the decomposition hypothesis, Assumption 6.1, i.e., χ can be factored as

$$\chi = \tilde{\psi}_1 \tilde{\psi}_2 \mathbf{N}^{(k_1+k_2-2)/2}.$$

Here, for $i \in \{1, 2\}$, ψ_i is a Hecke character of K_i of infinity type $(1 - k_i, 0)$ and modulus $\mathfrak{f}_i$; and $\tilde{\psi}_i$ is the Hecke character of K_0, obtained by composing $\mathbb{A}_{K_0}^{\times} \xrightarrow{\mathbb{N}_{K_0/K_i}} \mathbb{A}_{K_i}^{\times} \xrightarrow{\psi_i} \mathbb{C}$. Note that, if this happens, we must have $k_1 = a - b + 1$ and $k_2 = a + b + 1$. We then focus on the conjugate self-dual

[1] By either using $L(f/K_0, \chi, r) = L(f/K_0, \chi^{\boldsymbol{c}}, r)$, where $\chi^{\boldsymbol{c}}$ is the composition of χ with the action of complex conjugation, or swapping the order of K_1 and K_2, we would be able to cover other cases of a and b.

MSC2020: 11F67, 11R23.
Keywords: Euler system, Bloch–Kato conjecture, Iwasawa theory, Selmer groups.

© 2026 The Author, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

$G_{K_0} = \mathrm{Gal}(\overline{\mathbb{Q}}/K_0)$-representation

$$V_{f,\chi} := V_f^\vee(1-r) \otimes \chi^{-1},$$

where $V_f^\vee$ is the contragredient of Deligne's $\mathfrak{P}$-adic Galois representation associated to f.

For the remainder of this section, we assume

- f is ordinary and non-Eisenstein at $\mathfrak{P}$;
- p splits completely in K_0;
- $p \nmid h_{K_0}$, where h_{K_0} is the class number of K_0.

For every integral ideal μ_3 of $\mathcal{O}_{K_3}$, let $K_0[\mu_3]$ be the maximal p-subextension of the ring class field of K_0 of conductor μ_3. Denote by $\mathcal{N}$ the set of squarefree products of primes $\mu_3 \subset \mathcal{O}_{K_3}$, where $m = N_{K_3/\mathbb{Q}}(\mu_3)$ is squarefree, prime to p, and split in K_0.

Theorem A (Theorem 3.3.1). *There exists a collection of Iwasawa cohomology classes*

$$z_{f,\chi,\mu_3} \in H^1_{\mathrm{Iw}}(K_0[\mu_3 p^\infty], T_{f,\chi}),$$

indexed by the ideals $\mu_3 \in \mathcal{N}$ with $m = N_{K_3/\mathbb{Q}}(\mu_3)$, where $T_{f,\chi}$ is a certain G_K-stable $\mathcal{O}$-lattice inside $V_{f,\chi}$, such that, for every prime $\lambda_3 \in \mathcal{N}$ of norm ℓ with $(\ell, mp) = 1$, we have the norm relation

$$\mathrm{Norm}_{K_0[\mu_3]}^{K_0[\mu_3\lambda_3]}(z_{f,\chi,\mu_3\lambda_3}) = P_{\mathcal{L}_4}(\mathrm{Frob}_{\mathcal{L}_4})(z_{f,\chi,\mu_3}),$$

where $P_{\mathcal{L}_4}(X) = \det(1 - X \cdot \mathrm{Frob}_{\mathcal{L}_4} | (T_{f,\chi})^\vee(1))$, and $\mathrm{Frob}_{\mathcal{L}_4}$ is the geometric Frobenius.

Remark. Jetchev, Nekovář and Skinner [Jetchev et al. 2024] have developed a theory of "split" anticyclotomic Euler systems attached to conjugate self-dual representations over CM fields, where classes are defined over ring class extensions of CM fields (indexed by ideals of their totally real subfields). Our construction fits within their framework. Furthermore, the condition that $m = N_{K_3/\mathbb{Q}}(\mu_3)$ splits in K_0 does exclude the setting where m is inert in K_3 and μ_3 splits in K_0. Nevertheless, this does not affect the application of the Jetchev–Nekovář–Skinner machinery (see some details for the imaginary quadratic case in [Do 2022, Section 4.3]).

Due to its geometric origin, if we let

$$\kappa_{f,\chi} := \mathrm{Norm}_{K_0}^{K_0[1]}(z_{f,\chi,(1)}),$$

then it will land in a Selmer subgroup of $H^1(K_0, V_{f,\chi})$ with "nice" local conditions (see Section 6.3). Then, feeding Theorem A to the general Euler system machinery of [Jetchev et al. 2024], we deduce the following cases of the Bloch–Kato conjecture in analytic rank 0.

Theorem B (Theorem 6.5.1). *Let $f \in S_k(\Gamma_0(N_f))$ be a newform. Let χ be an anticyclotomic Hecke character of K_0 of infinity type $(-a, a, -b, b)$ satisfying Assumption 6.1. Assume further that*:

(1) *Either $k \geq 2a + 2$ or $2b \geq k$.*

(2) $N_f\mathcal{O}_{K_3} = \mathfrak{n}^+\mathfrak{n}^-$, *where* $\mathfrak{n}^+$ (*resp.* $\mathfrak{n}^-$) *is divisible only by primes which are split* (*resp. inert*) *in* K_0/K_3 *and* $\mathfrak{n}^-$ *is a squarefree product of an even number of primes.*

(3) $\bar{\rho}_f$ *is absolutely irreducible.*

(4) $(pN_f, \mathrm{Norm}_{K_1/\mathbb{Q}}(\mathfrak{f}_1)\mathrm{Norm}_{K_2/\mathbb{Q}}(\mathfrak{f}_2)D_{K_0}) = 1$.

Then

$$L(f/K_0, \chi, r) \neq 0 \implies \mathrm{Sel}_{\mathrm{BK}}(K_0, V_{f,\chi}) = 0,$$

and hence the Bloch–Kato conjecture for $V_{f,\chi}$ *holds in this case.*

Note that the first two conditions of Theorem B imply that the sign of the functional equation of $V_{f,\chi}$ is equal to $+1$; see also Remark 6.5.2. This puts us in an ideal situation for the nonvanishing of central L-values generically.

Let $\mathcal{O}$ be the ring of integers of $L_{\mathfrak{P}}$. We say that f has big image if, for a certain Galois stable $\mathcal{O}$-lattice $T_f^\vee \subset V_f^\vee$, the image of $G_{\mathbb{Q}}$ in $\mathrm{Aut}_{\mathcal{O}}(T_f^\vee)$ contains a conjugate of $\mathrm{SL}_2(\mathbb{Z}_p)$. Under this assumption, we also have results towards the Bloch–Kato conjecture in the analytic rank 1 case.

Theorem C (Theorem 6.7.1). *Let the hypotheses be as in Theorem B, and assume in addition that*:

(1) $\bar{\rho}_f$ *is* p*-distinguished.*

(2) f *has big image.*

(3) $p > k - 2$.

If $2a \geq k \geq 2b + 2$ (*which implies* $L(f/K_0, \chi, r) = 0$), *then*

$$\dim_{L_{\mathfrak{P}}} \mathrm{Sel}_{\mathrm{BK}}(K_0, V_{f,\chi}) \geq 1.$$

Finally, we note that these results also include the proof of a divisibility towards the anticyclotomic Iwasawa main conjecture for $V_{f,\chi}$; see Theorem 6.6.1.

1.1. ***Relation to previous works.*** When χ is an anticyclotomic Hecke character over an imaginary quadratic field K, the arithmetic of $V_{f,\chi}$ has been studied intensively via the Euler system of Heegner points pioneered in [Gross and Zagier 1986; Kolyvagin 1988] (see also [Zhang 1997; Tian 2003; Nekovář 2007]), and generalized Heegner cycles by Bertolini, Darmon and Prasanna [Bertolini et al. 2013]. In particular, these objects have direct implications towards the Bloch–Kato conjecture in analytic rank 0 for $V_{f,\chi}$ either by varying the generalized Heegner cycles in p-adic families like in [Castella and Hsieh 2018] (see also [Castella 2020]), or by the "level-raising" method like in [Bertolini and Darmon 2005] (see also [Longo and Vigni 2010; Chida and Hsieh 2015; Chida 2017]). In the same vein as [Bertolini and Darmon 2005], Nekovář [2012] and Wang [2023] proved results towards the rank 0 Bloch–Kato conjecture when f is a cuspidal Hilbert modular eigenform over a totally real field F of parallel weight 2 and higher weights, respectively, where χ is a finite-order character; see also [Tamiozzo 2021].

Outside of the Heegner realm, it is worthwhile to mention that the Euler system of Beilinson–Flach classes constructed by Lei, Loeffler and Zerbes [Lei et al. 2014; 2015] and Kings, Loeffler and Zerbes

[Kings et al. 2017; 2020] can be applied to obtain similar rank 0 results. Relying on this, Lamplugh [2016] constructed Euler systems for $\mathrm{Ind}_{K_0}^{K_1}\mathcal{O}(\chi\rho)$ over K_1 (where ρ is an auxiliary character) and used that to bound the associated Selmer group over K_0 via Rubin's machinery [2000].

The anticyclotomic Euler system over K_0 that we will describe in this paper is more comparable with the anticyclotomic diagonal Euler system [Do 2022; Castella and Do 2025] over K (an imaginary quadratic field) and comes together with application towards the Bloch–Kato conjecture in analytic rank 0. The construction of the cohomology classes, similar to [Do 2022; Castella and Do 2025], is based on a generalization of the diagonal cycles pioneered in [Gross and Kudla 1992; Gross and Schoen 1995], and improved in Bertolini–Seveso–Venerucci and Darmon–Rotger [Bertolini et al. 2022a]. Despite the fact that it was addressed later, the imaginary biquadratic case is actually a generic case (where $K_1 \neq K_2$), while the imaginary quadratic case is a degenerate situation (where $K = K_1 = K_2$).

In future work, we intend to construct a bipartite Euler system over a biquadratic field as well as investigate the case where p does not split completely in K_0.

2. Preliminaries

2.1. ***Galois representations associated to newforms.*** In this section, we follow [Castella and Do 2025, Section 1.1] and introduce some important notation and results. Let $f \in S_k(\Gamma_1(N_f), \chi_f)$ be a normalized newform of weight $k \geq 2$ and let $\sum_{n=1}^{\infty} a_n q^n$ be its q-expansion. Let $p \nmid N_f$ be a prime. Fix embeddings $i_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Let $L/\mathbb{Q}$ be the coefficient field of f, that is, L contains all values $i_\infty^{-1}(a_n)$ and $i_\infty^{-1} \circ \chi_f$. Let $\mathfrak{P}$ be the prime of L above p with respect to i_p. Let $S = \{\text{primes } \ell \mid pN_f\} \cup \{\infty\}$. Then Eichler and Shimura (for $k = 2$) and Deligne (for $k > 2$) construct a p-adic Galois representation associated to f,

$$\rho_{f,\mathfrak{P}} : G_{\mathbb{Q},S} \to \mathrm{GL}_2(L_{\mathfrak{P}}),$$

such that, for all primes $\ell \notin S$,

- $\mathrm{trace}(\rho_{f,\mathfrak{P}}(\mathrm{Frob}_\ell)) = i_p(a_\ell)$;
- $\det(\rho_{f,\mathfrak{P}}(\mathrm{Frob}_\ell)) = i_p(\chi_f(\ell)\ell^{k-1})$;
- $\rho_{f,\mathfrak{P}}$ is irreducible, and hence absolutely irreducible.

Here Frob_ℓ is the geometric Frobenius.

As in [loc. cit.], one obtains the geometric realization V_f of $\rho_{f,\mathfrak{P}}$ defined as the subspace of

$$H^1_{\text{ét}}(Y_1(N_f)_{\overline{\mathbb{Q}}}, \mathscr{S}_{k-2}) \otimes L_{\mathfrak{P}}.$$

Dually, $V_f^\vee = \mathrm{Hom}(V_f, L_{\mathfrak{P}})$ can be interpreted as the maximal quotient of

$$H^1_{\text{ét}}(Y_1(N_f)_{\overline{\mathbb{Q}}}, \mathscr{L}_{k-2}(1)) \otimes L_{\mathfrak{P}}$$

on which the dual Hecke operator T'_ℓ acts as multiplication by a_ℓ for all $\ell \nmid N_f p$ and $\langle d \rangle = \langle d \rangle^*$ acts as multiplication by $\chi_f(d)$ for all $d \in (\mathbb{Z}/N_f\mathbb{Z})^\times$.

Let $\mathcal{O}$ be the ring of integers of $L_{\mathfrak{P}}$. There exists a $G_{\mathbb{Q}}$-stable $\mathcal{O}$-lattice $T_f^\vee \subset V_f^\vee$ defined as the image of $H^1_{\text{ét}}(Y_1(N_f)_{\overline{\mathbb{Q}}}, \mathscr{L}_{k-2}(1)) \otimes \mathcal{O}$ in $V_f^\vee$.

If f is ordinary at p (which means $i_p(a_p) \in \mathcal{O}^\times$), then the restriction of V_f to $G_{\mathbb{Q}_p}$ is reducible. This leads us to an exact sequence of $L_{\mathfrak{P}}[G_{\mathbb{Q}_p}]$-modules

$$0 \to V_f^+ \to V_f \to V_f^- \to 0,$$

where $\dim_{L_{\mathfrak{P}}} V_f^\pm = 1$. Dually, we also obtain an exact sequence for the restriction of $V_f^\vee$ to $G_{\mathbb{Q}_p}$,

$$0 \to V_f^{\vee,+} \to V_f^\vee \to V_f^{\vee,-} \to 0, \tag{2-1}$$

where $V_f^{\vee,+} \simeq (V_f^-)^\vee(1-k)(\chi_f^{-1})$, and the $G_{\mathbb{Q}_p}$-action on the quotient $V_f^{\vee,-}$ is given by the unramified character sending the arithmetic Frobenius Frob_p^{-1} to α_p, which is the unit root of $x^2 - a_p x + \chi_f(p)p^{k-1}$.

2.2. ***Patched CM Hecke modules.*** Here, we recall the conventions on Hecke characters and the construction of certain patched CM Hecke modules from [Castella and Do 2025, Section 1.3; Lei et al. 2015].

2.2.1. *Hecke characters and theta series.* Let K be an imaginary quadratic field. Let $p = \mathfrak{p}\bar{\mathfrak{p}}$ be a prime that splits in K with $\mathfrak{p}$, the prime of K above p, induced by $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. We say that a Hecke character $\psi : \mathbb{A}_K^\times / K^\times \to \mathbb{C}^\times$ has infinity type (m, n), where m, n are integers, if $\psi_\infty(x_\infty) = x_\infty^m \bar{x}_\infty^n$.

Let $\mathrm{rec}_K : \mathbb{A}_K^\times \to G_K^{\mathrm{ab}}$ be the geometrically normalized Artin reciprocity map. Following [Castella and Do 2025, Section 1.3.1], given $g \in G_K$, we take $x \in \mathbb{A}_K^\times$ such that $\mathrm{rec}_K(x) = g|_{K^{\mathrm{ab}}}$ and define

$$\psi_{\mathfrak{P}}(g) = i_p \circ i_\infty^{-1}(\psi(x) x_\infty^{-m} \bar{x}_\infty^{-n}) x_{\mathfrak{p}}^m x_{\bar{\mathfrak{p}}}^n.$$

Such a $\psi_{\mathfrak{P}}$ will be called the p-adic avatar of ψ. We shall also use ψ to denote its p-adic avatar if the context makes this usage reasonable.

Attached to ψ, a Hecke character of K of infinity type $(-1, 0)$ with conductor $\mathfrak{f}$ that takes values in a finite extension L/K, is the theta series

$$\theta_\psi = \sum_{(\mathfrak{a}, \mathfrak{f})=1} \psi(\mathfrak{a}) q^{N_{K/\mathbb{Q}}(\mathfrak{a})} \in S_2(\Gamma_1(N_\psi), \chi_\psi \epsilon_K),$$

where $N_\psi = N_{K/\mathbb{Q}}(\mathfrak{f}) \operatorname{disc}(K/\mathbb{Q})$, χ_ψ is the unique Dirichlet character modulo $N_{K/\mathbb{Q}}(\mathfrak{f})$ such that $\psi((n)) = n\chi_\psi(n)$ for all $n \in \mathbb{Z}$ with $(n, N_{K/\mathbb{Q}}(\mathfrak{f})) = 1$, and ϵ_K is the quadratic Dirichlet character attached to K. The cuspform θ_ψ is new of level $N_\psi = N_{K/\mathbb{Q}}(\mathfrak{f}) \cdot \operatorname{disc}(K/\mathbb{Q})$ by [Miyake 1989]. One obtains the description of the $\mathfrak{P}$-adic representation of θ_ψ

$$V_{\theta_\psi}^\vee \cong \mathrm{Ind}_K^{\mathbb{Q}} L_{\mathfrak{P}}(\psi^{-1}),$$

where $\mathfrak{P}$ is the prime of L above p with respect to i_p.

2.2.2. *Hecke algebras and norm maps.* We keep the notation of the previous section and follow [Castella and Do 2025, Section 1.3.1]. Let $\mathfrak{n} \subset \mathcal{O}_K$ be an ideal divisible by $\mathfrak{f}$ and let $N = N_{K/\mathbb{Q}}(\mathfrak{n}) \operatorname{disc}(K/\mathbb{Q})$. Let $K_\mathfrak{n}$ be the ray class field of K with conductor $\mathfrak{n}$. Let $H_\mathfrak{n} = \mathrm{Gal}(K_\mathfrak{n}/K)$ be the ray class group of K modulo $\mathfrak{n}$. Let $K(\mathfrak{n})$ be the largest p-subextension of K contained in $K_\mathfrak{n}$, i.e., $\mathrm{Gal}(K(\mathfrak{n})/K) \cong H_\mathfrak{n}^{(p)}$ is

the largest p-power quotient of $H_{\mathfrak{n}}$. Given an ideal $\mathfrak{k}$ of K that is coprime to $\mathfrak{n}$, let $[\mathfrak{k}]$ be the class of $\mathfrak{k}$ in $H_{\mathfrak{n}}$. Let $\mathbb{T}'(N)$ be the subalgebra of $\mathrm{End}_{\mathbb{Z}}\big(H^1(Y_1(N)(\mathbb{C}),\mathbb{Z})\big)$ generated by $\langle d\rangle'$ and T'_ℓ for all primes ℓ; then one can prove that:

Proposition 2.2.1 [Lei et al. 2015, Proposition 3.2.1]. *There exists a homomorphism $\phi_{\mathfrak{n}}:\mathbb{T}'(N)\to\mathcal{O}[H_{\mathfrak{n}}]$ defined by*

$$\phi_{\mathfrak{n}}(T'_\ell)=\sum_{\substack{\mathfrak{l}\subset\mathcal{O}_K,\mathfrak{l}\nmid\mathfrak{n}\\ N_{K/\mathbb{Q}}(\mathfrak{l})=\ell}}[\mathfrak{l}]\psi(\mathfrak{l}),\quad \phi_{\mathfrak{n}}(\langle d\rangle')=\chi_\psi(d)\epsilon_K(d)[(d)].$$

For $\mathfrak{m}=\mathfrak{n}\mathfrak{l}$ with $\mathfrak{l}$ a prime ideal and $(\mathfrak{m},p)=1$, put $M=N_{K/\mathbb{Q}}(\mathfrak{m})\,\mathrm{disc}(K/\mathbb{Q})$; then one has the map

$$\mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}}:\mathcal{O}[H_{\mathfrak{m}}^{(p)}]\otimes_{\mathbb{T}'(M)\otimes\mathbb{Z}_p,\phi_{\mathfrak{m}}}H^1_{\text{ét}}(Y_1(M)_{\overline{\mathbb{Q}}},\mathbb{Z}_p(1))\to\mathcal{O}[H_{\mathfrak{n}}^{(p)}]\otimes_{\mathbb{T}'(N)\otimes\mathbb{Z}_p,\phi_{\mathfrak{n}}}H^1_{\text{ét}}(Y_1(N)_{\overline{\mathbb{Q}}},\mathbb{Z}_p(1)).$$

This norm map is defined explicitly by splitting into three cases (see [Castella and Do 2025, Section 1.1.2] for the definition of the degeneracy map):

- If $\mathfrak{l}\mid\mathfrak{n}$, then
$$\mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}}=1\otimes\mathrm{pr}_{1*}.$$
- If $\mathfrak{l}\nmid\mathfrak{n}$ is split or ramified in K and $N_{K/\mathbb{Q}}(\mathfrak{l})=\ell$, then
$$\mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}}=1\otimes\mathrm{pr}_{1*}-\frac{\psi(\mathfrak{l})[\mathfrak{l}]}{\ell}\otimes\mathrm{pr}_{\ell*}.$$
- If $\mathfrak{l}\nmid\mathfrak{n}$ is inert in K, say $\mathfrak{l}=(\ell)$, then
$$\mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}}=1\otimes\mathrm{pr}_{1*}-\frac{\psi(\mathfrak{l})[\mathfrak{l}]}{\ell^2}\otimes\mathrm{pr}_{\ell\ell*}.$$

Note that one can extend the definition of $\mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}}$ to any pair of ideals $\mathfrak{n}\mid\mathfrak{m}$ by composition.

Following [ibid., Section 1.3.2], if p splits in K and $(p,\mathfrak{f})=1$ then, for any ideal $\mathfrak{n}\subset\mathcal{O}_K$ divisible by $\mathfrak{f}$ such that $(\mathfrak{n},\bar{\mathfrak{p}})=1$, the maximal ideal of $\mathbb{T}'(N)$ defined by the kernel of the composition

$$\mathbb{T}'(N)\xrightarrow{\phi_{\mathfrak{n}}}\mathcal{O}[H_{\mathfrak{n}}]\xrightarrow{\text{aug}}\mathcal{O}\to\mathcal{O}/\mathfrak{P},$$

is non-Eisenstein, p-ordinary and p-distinguished.

We finish this section by extracting a crucial result in [Lei et al. 2015] in the case where p splits in K. This will be used later to prove the norm relation of our Euler system.

Theorem 2.2.2 [Lei et al. 2015, Corollary 5.2.6]. *Assume that $(p,\mathfrak{f})=1$. Let $\mathcal{A}$ be the set of ideals $\mathfrak{m}\subset\mathcal{O}_K$ with $(\mathfrak{m},\bar{\mathfrak{p}})=1$, and put $\mathcal{A}_{\mathfrak{f}}=\{\mathfrak{f}\mathfrak{m}:\mathfrak{m}\in\mathcal{A}\}$. Given $\mathfrak{n}\in\mathcal{A}_{\mathfrak{f}}$, there is a $G_{\mathbb{Q}}$-equivariant isomorphism of $\mathcal{O}[H_{\mathfrak{n}}^{(p)}]$-modules*

$$\nu_{\mathfrak{n}}:\mathcal{O}[H_{\mathfrak{n}}^{(p)}]\otimes_{\mathbb{T}'(N)\otimes\mathbb{Z}_p,\phi_{\mathfrak{n}}}H^1_{\text{ét}}(Y_1(N)_{\overline{\mathbb{Q}}},\mathbb{Z}_p(1))\xrightarrow{\cong}\mathrm{Ind}_{K(\mathfrak{n})}^{\mathbb{Q}}\mathcal{O}(\psi_{\mathfrak{P}}^{-1}).$$

Furthermore, for any $\mathfrak{n}, \mathfrak{m} \in \mathcal{A}_{\mathfrak{f}}$ *with* $\mathfrak{n} \mid \mathfrak{m}$, *the diagram*

$$\begin{array}{ccc} \mathcal{O}[H_{\mathfrak{m}}^{(p)}] \otimes_{\mathbb{T}'(M)\otimes\mathbb{Z}_p,\phi_{\mathfrak{m}}} H^1_{\text{ét}}(Y_1(M)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) & \xrightarrow[\cong]{\nu_{\mathfrak{m}}} & \mathrm{Ind}_{K(\mathfrak{m})}^{\mathbb{Q}} \mathcal{O}(\psi_{\mathfrak{P}}^{-1}) \\ \Big\downarrow \mathcal{N}_{\mathfrak{n}}^{\mathfrak{m}} & & \Big\downarrow \mathrm{Norm}_{\mathfrak{n}}^{\mathfrak{m}} \\ \mathcal{O}[H_{\mathfrak{n}}^{(p)}] \otimes_{\mathbb{T}'(N)\otimes\mathbb{Z}_p,\phi_{\mathfrak{n}}} H^1_{\text{ét}}(Y_1(N)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) & \xrightarrow[\cong]{\nu_{\mathfrak{n}}} & \mathrm{Ind}_{K(\mathfrak{n})}^{\mathbb{Q}} \mathcal{O}(\psi_{\mathfrak{P}}^{-1}) \end{array}$$

commutes, *where* $\mathrm{Norm}_{\mathfrak{n}}^{\mathfrak{m}}$ *is the natural norm map of the induced representations.*

3. The construction

For a newform f and two Hecke characters ψ_1, ψ_2 of two distinct imaginary quadratic fields K_1, K_2, respectively, using the results from [Castella and Do 2025; Bertolini et al. 2022b; Lei et al. 2015] recalled in the preceding section, we construct a family of cohomology classes for $f \otimes \tilde{\psi}_1\tilde{\psi}_2$ defined over ring class field extensions of K_0, which is the compositum of K_1 and K_2, and prove that they satisfy the norm relations of an anticyclotomic Euler system. Following [Castella and Do 2025, Section 2], we first give the construction and show the tame norm relations in the case where $(f, \theta_{\psi_1}, \theta_{\psi_2})$ have weights $(2, 2, 2)$. Then, by varying the diagonal cycle classes in Hida families, we extend the construction to more general weights and prove the wild norm relations.

Throughout this section we consider the following setup:

(1) Let $f \in S_k(\Gamma_0(N_f))$ be a newform of weight $k \geq 2$.

(2) Let $K_1/\mathbb{Q}$ be an imaginary quadratic field of discriminant D_1 coprime to N_f. Let ψ_1 be a Hecke character of K_1 of infinity type $(1-k_1, 0)$, with $k_1 \geq 1$, and modulus $\mathfrak{f}_1$.

(3) Let $K_2/\mathbb{Q}$ be an imaginary quadratic field of discriminant $D_2 \neq D_1$ and coprime to N_f. Let ψ_2 be a Hecke character of K_2 of infinity type $(1-k_2, 0)$, with $k_2 \geq 1$, and modulus $\mathfrak{f}_2$.

(4) Denote by ϵ_{K_i} the quadratic character attached to the quadratic field K_i for $i \in \{1, 2\}$.

(5) Let K_0 be the compositum of K_1 and K_2. Since K_0 is a biquadratic field, we can consider K_3, the unique real quadratic field inside K_0.

(6) Let $\tilde{\psi}_i$ be the Hecke character of K_0, obtained by composing $\mathbb{A}_{K_0}^{\times} \xrightarrow{\mathbb{N}_{K_0/K_i}} \mathbb{A}_{K_i}^{\times} \xrightarrow{\psi_i} \mathbb{C}$ for $i \in \{1, 2\}$.

(7) Denote by

$$\theta_{\psi_i} \in S_{k_i}(N_{\psi_i}, \chi_{\psi_i}\epsilon_{K_i})$$

the associated theta series, where $N_{\psi_i} = N_{K_i/\mathbb{Q}}(\mathfrak{f}_i) \cdot \mathrm{disc}(K_i/\mathbb{Q})$ and χ_{ψ_i} is the Dirichlet character modulo $N_{K_i/\mathbb{Q}}(\mathfrak{f}_i)$ defined by $\psi_i((n)) = n^{k_i-1}\chi_{\psi_i}(n)$ for all integers n prime to $N_{K_i/\mathbb{Q}}(\mathfrak{f}_i)$ for $i \in \{1, 2\}$.

(8) We assume the self-duality condition

$$\chi_{\psi_1}\epsilon_{K_1}\chi_{\psi_2}\epsilon_{K_2} = 1. \tag{3-1}$$

Let L/K_0 be a finite extension containing the Fourier coefficients of f, θ_{ψ_1} and θ_{ψ_2}. Let $p \geq 5$ be a prime that splits in K_0 such that $(p, N_f N_{\psi_1} N_{\psi_2}) = 1$, and let $\mathfrak{P} \mid \mathfrak{p}$ be the prime of L/K_0 above p determined by a fixed embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$. Finally, let $L_{\mathfrak{P}}$ be the completion of L at $\mathfrak{P}$, and denote by $\mathcal{O}$ the ring of integers of $L_{\mathfrak{P}}$.

3.1. ***Digression to primes decomposition and the left superscript notation.*** Let ℓ be a split prime in K_0, i.e., $(\ell) = \mathcal{L}_1\mathcal{L}_2\mathcal{L}_3\mathcal{L}_4$. We can write $(\ell) = \lambda_1\bar{\lambda}_1$ and $(\ell) = \lambda_2\bar{\lambda}_2$ in K_1 and K_2, respectively. Note that ℓ also splits in K_3 as $\lambda_3\tilde{\lambda}_3$, where the tilde corresponds to the nontrivial element generating the Galois group $\mathrm{Gal}(K_3/\mathbb{Q})$.

Let τ_i be the generator of $\mathrm{Gal}(K_0/K_i)$ for $i = \{1, 2, 3\}$; then $\tau_3 = \tau_1\tau_2$ (this is the complex conjugation on K_0/K_3). Due to the Galois group action on primes lying above ℓ, we can further assume that

$$\lambda_1 = \mathcal{L}_1\mathcal{L}_4, \quad \bar{\lambda}_1 = \mathcal{L}_3\mathcal{L}_2, \quad \lambda_2 = \mathcal{L}_1\mathcal{L}_3, \quad \bar{\lambda}_2 = \mathcal{L}_2\mathcal{L}_4,$$
$$\lambda_3 = \mathcal{L}_4\mathcal{L}_3 \quad (\text{so } \lambda_3 \mid \lambda_1\lambda_2) \quad \text{and} \quad \tilde{\lambda}_3 = \mathcal{L}_1\mathcal{L}_2,$$

where

$$\mathcal{L}_4 = \tau_1\mathcal{L}_1, \quad \mathcal{L}_3 = \tau_2\mathcal{L}_1, \quad \mathcal{L}_2 = \tau_3\mathcal{L}_1 = \tau_1\tau_2\mathcal{L}_1.$$

Denote by $\mathcal{L}$ the set of primes $\lambda_3 \subset \mathcal{O}_{K_3}$, where $\ell = N_{K_3/\mathbb{Q}}(\lambda_3)$ is prime to p and ℓ splits in K_0. Let $\mathcal{N}$ be the set of squarefree products of primes inside $\mathcal{L}$ whose norm down to $\mathbb{Q}$ is still squarefree. For such λ_3, we can choose $\lambda_1 \subset \mathcal{O}_{K_1}$ and $\lambda_2 \subset \mathcal{O}_{K_2}$ as above such that $\lambda_3 \mid \lambda_1\lambda_2$.

Let $\mu_3 \in \mathcal{N}$ and $N_{K_3/\mathbb{Q}}(\mu_3) = m$. Then its norm $m = \prod_i \ell_i$ will be a product of split primes ℓ_i in K_0. Similarly, we can decompose $(m) = \mathcal{M}_1\mathcal{M}_2\mathcal{M}_3\mathcal{M}_4$, $(m) = \mu_1\bar{\mu}_1$, $(m) = \mu_2\bar{\mu}_2$ and $(m) = \mu_3\tilde{\mu}_3$ inside K_0, K_1, K_2 and K_3, respectively, where we can assume

$$\mu_1 = \mathcal{M}_1\mathcal{M}_4, \quad \bar{\mu}_1 = \mathcal{M}_3\mathcal{M}_2, \quad \mu_2 = \mathcal{M}_1\mathcal{M}_3, \quad \bar{\mu}_2 = \mathcal{M}_2\mathcal{M}_4,$$
$$\mu_3 = \mathcal{M}_4\mathcal{M}_3 \quad (\text{so } \mu_3 \mid \mu_1\mu_2) \quad \text{and} \quad \tilde{\mu}_3 = \mathcal{M}_1\mathcal{M}_2.$$

Here, for every i, $\ell_i = \prod_j \mathcal{L}_{j,i}$, and $\mathcal{M}_j = \prod_i \mathcal{L}_{j,i}$ for $1 \leq j \leq 4$.

For each $i \in \{0, 1, 2\}$, we define ${}^iK_{\mathfrak{n}_i}$ as the ray class field of K_i with conductor $\mathfrak{n}_i$ (an integral ideal inside $\mathcal{O}_{K_i}$), and let ${}^iH_{\mathfrak{n}_i}$ be the ray class group of K_i modulo $\mathfrak{n}_i$. Let $K_i(\mathfrak{n}_i)$ be the largest p-subextension of K_i contained in ${}^iK_{\mathfrak{n}_i}$, so $\mathrm{Gal}(K_i(\mathfrak{n}_i)/K_i) \cong {}^iH_{\mathfrak{n}_i}^{(p)}$ is the largest p-power quotient of ${}^iH_{\mathfrak{n}_i}$.

3.2. ***Construction in weight* (2, 2, 2) *and tame norm relation.*** Suppose in this section that $(k, k_1, k_2) = (2, 2, 2)$. Let $N = \mathrm{lcm}(N_f, N_{\psi_1}, N_{\psi_2})$. Following [Castella and Do 2025, Section 2.1], which is based on the diagonal classes in the triple product of modular curves [Bertolini et al. 2022b, Section 3], we have cohomology classes

$$\mathcal{Z}_m^{(1)} := \tilde{\kappa}_m^{(3)} \in H^1\big(\mathbb{Q}, H^1_{\text{ét}}(Y_1(N)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes H^1_{\text{ét}}(Y_1(N_{\psi_1}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes H^1_{\text{ét}}(Y_1(N_{\psi_2}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1))(-1)\big)$$

for every positive integer m. One then chooses a test vector $\check{f} \in S_2(N)[f]$. As noted in [loc. cit.], the maps used to construct $\mathcal{Z}_m^{(1)}$ are compatible with correspondences. This allows one to tensor them with $\mathcal{O}$

and obtain

$$\mathcal{Z}_{\mu_3}^{(1)} \in H^1\big(\mathbb{Q}, T_f^\vee \otimes H^1_{\text{ét}}(Y_1(N_{\psi_1}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes_{\mathbb{T}'(N_{\psi_1}m)} \mathcal{O}[{}^1H^{(p)}_{\mathfrak{f}_1\mu_1}] \otimes H^1_{\text{ét}}(Y_1(N_{\psi_2}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes_{\mathbb{T}'(N_{\psi_2}m)} \mathcal{O}[{}^2H^{(p)}_{\mathfrak{f}_2\mu_2}](-1)\big). \quad (3\text{-}2)$$

Here, the chosen $\check{f}$ is used to take the image under the projection $H^1_{\text{ét}}(Y_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \to T_f^\vee$ in the first factor. The tensor products are taken from Proposition 2.2.1,

$$\phi_{\mathfrak{f}_1\mu_1} : \mathbb{T}'(N_{\psi_1}m) \to \mathcal{O}[{}^1H^{(p)}_{\mathfrak{f}_1\mu_1}], \quad \phi_{\mathfrak{f}_2\mu_2} : \mathbb{T}'(N_{\psi_2}m) \to \mathcal{O}[{}^2H^{(p)}_{\mathfrak{f}_2\mu_2}],$$

with respect to two distinct imaginary quadratic fields K_1 and K_2, respectively.

Via the isomorphisms from Theorem 2.2.2 with respect to two distinct imaginary quadratic fields,

$$\nu_{\mathfrak{f}_1\mu_1} : H^1_{\text{ét}}(Y_1(N_{\psi_1}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes_{\mathbb{T}'(N_{\psi_1}m)} \mathcal{O}[{}^1H^{(p)}_{\mathfrak{f}_1\mu_1}] \xrightarrow{\sim} \operatorname{Ind}^{\mathbb{Q}}_{K_1(\mathfrak{f}_1\mu_1)} \mathcal{O}(\psi_1^{-1}),$$
$$\nu_{\mathfrak{f}_2\mu_2} : H^1_{\text{ét}}(Y_1(N_{\psi_2}m)_{\overline{\mathbb{Q}}}, \mathbb{Z}_p(1)) \otimes_{\mathbb{T}'(N_{\psi_2}m)} \mathcal{O}[{}^2H^{(p)}_{\mathfrak{f}_2\mu_2}] \xrightarrow{\sim} \operatorname{Ind}^{\mathbb{Q}}_{K_2(\mathfrak{f}_2\mu_2)} \mathcal{O}(\psi_2^{-1}),$$

one then obtains a cohomology class in

$$H^1\big(\mathbb{Q}, T_f^\vee \otimes_{\mathcal{O}} \operatorname{Ind}^{\mathbb{Q}}_{K_1(\mathfrak{f}_1\mu_1)} \mathcal{O}(\psi_1^{-1}) \otimes_{\mathcal{O}} \operatorname{Ind}^{\mathbb{Q}}_{K_2(\mathfrak{f}_2\mu_2)} \mathcal{O}(\psi_2^{-1})(-1)\big),$$

which, under the maps ${}^1H_{\mathfrak{f}_1\mu_1} \twoheadrightarrow {}^1H_{\mu_1}$ and ${}^2H_{\mathfrak{f}_2\mu_2} \twoheadrightarrow {}^2H_{\mu_2}$, can be projected to a class

$$\mathcal{Z}_{\mu_3}^{(2)} \in H^1\big(\mathbb{Q}, T_f^\vee \otimes_{\mathcal{O}} \operatorname{Ind}^{\mathbb{Q}}_{K_1} \mathcal{O}_{\psi_1^{-1}}[{}^1H^{(p)}_{\mu_1}] \otimes_{\mathcal{O}} \operatorname{Ind}^{\mathbb{Q}}_{K_2} \mathcal{O}_{\psi_2^{-1}}[{}^2H^{(p)}_{\mu_2}](-1)\big). \quad (3\text{-}3)$$

Note that the group cohomology can be rewritten as

$$H^1\big(\mathbb{Q}, T_f^\vee \otimes_{\mathcal{O}} \operatorname{Ind}^{\mathbb{Q}}_{K_0} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[{}^1H^{(p)}_{\mu_1} \times {}^2H^{(p)}_{\mu_2}](-1)\big),$$

which, by Shapiro's lemma, gives us elements

$$\mathcal{Z}_{\mu_3}^{(3)} \in H^1\big(K_0, T_f^\vee \otimes_{\mathcal{O}} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[{}^1H^{(p)}_{\mu_1} \times {}^2H^{(p)}_{\mu_2}](-1)\big). \quad (3\text{-}4)$$

3.2.1. *Projection to ring class groups.* Recall the fundamental exact sequence for ray class groups

$$\mathcal{O}_{K_i}^\times \to (\mathcal{O}_{K_i}/\mu_i\mathcal{O}_{K_i})^\times \to {}^iH_{\mu_i} \to {}^iH_1 \to 1$$

for $i \in \{1, 2\}$. Assume that $p \nmid 6h_{K_0}$, where h_{K_0} is the class number of K_0. Note that, by [Fröhlich and Taylor 1993, Theorem 74], we have

$$p \nmid h_{K_i},$$

which is the class number of K_i for $i \in \{1, 2, 3\}$. Taking the p-primary parts of the above exact sequence induces two isomorphisms

$${}^1H^{(p)}_{\mu_1} \xrightarrow{\simeq} (\mathcal{O}_{K_1}/\mu_1\mathcal{O}_{K_1})^{\times,(p)} \xrightarrow{\simeq} (\mathcal{O}_{K_0}/\mathcal{M}_4\mathcal{O}_{K_0})^{\times,(p)},$$
$${}^2H^{(p)}_{\mu_2} \xrightarrow{\simeq} (\mathcal{O}_{K_2}/\mu_2\mathcal{O}_{K_2})^{\times,(p)} \xrightarrow{\simeq} (\mathcal{O}_{K_0}/\mathcal{M}_3\mathcal{O}_{K_0})^{\times,(p)},$$

and hence the projection

$$ {}^1H_{\mu_1}^{(p)} \times {}^2H_{\mu_2}^{(p)} \xrightarrow{\simeq} (\mathcal{O}_{K_0}/\mu_3\mathcal{O}_{K_0})^{\times,(p)} \twoheadrightarrow {}^0H_{\mu_3}^{(p)}. \tag{3-5} $$

Recall that K_3 is the totally real field sitting inside a CM field K_0. Given an integral ideal $\mathfrak{n}$ of K_3, let $H[\mathfrak{n}]$ be the ring class group of K_0 of conductor $\mathfrak{n}$, so $H[\mathfrak{n}] \simeq \mathrm{Pic}(\mathcal{O}_\mathfrak{n})$ under the Artin reciprocity map, where $\mathcal{O}_\mathfrak{n} = \mathcal{O}_{K_3} + \mathfrak{n}\mathcal{O}_{K_0}$ is the order of K_0 of conductor $\mathfrak{n}$. Let $H[\mathfrak{n}]^{(p)}$ be the maximal p-power quotient of $H[\mathfrak{n}]$, and denote by $K_0[\mathfrak{n}]$ the maximal p-extension inside the ring class field of K_0 of conductor $\mathfrak{n}$, i.e., $H[\mathfrak{n}]^{(p)} = \mathrm{Gal}(K_0[\mathfrak{n}]/K_0)$. Note that, for the ring class groups and fields of K_0, we drop the left superscript 0 from the notation.

Proposition 3.2.1. *Suppose $p \nmid 6h_{K_0}$ and μ_3 is a squarefree ideal of $\mathcal{O}_{K_3}$ of norm m, where m is divisible only by primes that are split in K_0. Using (3-5), we have the short exact sequence*

$$ 1 \to (\mathcal{O}_{K_3}/\mu_3\mathcal{O}_{K_3})^{\times,(p)} \xrightarrow{\Delta} {}^1H_{\mu_1}^{(p)} \times {}^2H_{\mu_2}^{(p)} \xrightarrow{e\circ w} H[\mu_3]^{(p)} \to 1. $$

Here, the map Δ uses the identifications

$$ (\mathcal{O}_{K_3}/\mu_3\mathcal{O}_{K_3})^{\times,(p)} \simeq (\mathcal{O}_{K_0}/\mathcal{M}_4\mathcal{O}_{K_0})^{\times,(p)}, \quad (\mathcal{O}_{K_3}/\mu_3\mathcal{O}_{K_3})^{\times,(p)} \simeq (\mathcal{O}_{K_0}/\mathcal{M}_3\mathcal{O}_{K_0})^{\times,(p)}. $$

Moreover, if $(\ell) = \mathcal{L}_1\mathcal{L}_2\mathcal{L}_3\mathcal{L}_4$ is a prime that splits in K_0 and is coprime to m, the projection $e \circ w$ (defined in the proof below) is given by

$$ [\lambda_1] \times [\bar{\lambda}_2] \mapsto \mathrm{Frob}_{\mathcal{L}_4/\lambda_3}, $$

where $\mathrm{Frob}_{\mathcal{L}_4/\lambda_3}$ is the Frobenius element of $\mathcal{L}_4$ in $H[\mu_3]^{(p)}$.

Proof. We have the exact diagram

$$ \begin{array}{ccccccccc} \mathcal{O}_{K_0}^\times & \longrightarrow & (\mathcal{O}_{K_0}/\mu_3\mathcal{O}_{K_0})^\times & \longrightarrow & {}^0H_{\mu_3} & \longrightarrow & {}^0H_1 & \longrightarrow & 1 \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ \mathcal{O}_{K_0}^\times/\mathcal{O}_{K_3}^\times & \longrightarrow & (\mathcal{O}_{K_0}/\mu_3\mathcal{O}_{K_0})^\times/(\mathcal{O}_{K_3}/\mu_3\mathcal{O}_{K_3})^\times & \longrightarrow & H[\mu_3] & \longrightarrow & {}^0H_1 & \longrightarrow & 1 \end{array} $$

Taking the p-part of this, using the assumption that $p \nmid 6h_{K_0}$ together with the fact that $|\mathcal{O}_{K_0}^\times/\mathcal{O}_{K_3}^\times|$ is a power of 2, we obtain the diagram

$$ \begin{array}{ccccccccc} \mathcal{O}_{K_0}^\times \otimes \mathbb{Q}_p/\mathbb{Z}_p & \longrightarrow & (\mathcal{O}_{K_0}/\mu_3\mathcal{O}_{K_0})^{\times,(p)} & \xrightarrow{w} & {}^0H_{\mu_3}^{(p)} & \longrightarrow & 1 \\ \downarrow & & \downarrow & & \downarrow e & & \\ 1 & \longrightarrow & (\mathcal{O}_{K_0}/\mu_3\mathcal{O}_{K_0})^{\times,(p)}/(\mathcal{O}_{K_3}/\mu_3\mathcal{O}_{K_3})^{\times,(p)} & \longrightarrow & H[\mu_3]^{(p)} & \longrightarrow & 1 \end{array} $$

Using the middle arrow and the identification (3-5), we can show the first exact sequence.

One can show the second part by noting that $\mathcal{L}_4$ is a prime of K_0 lying above both λ_1 (a prime of K_1) and $\bar{\lambda}_2$ (a prime of K_2). □

In the same setting as Proposition 3.2.1, we can consider the image of $\mathcal{Z}^{(3)}_{\mu_3}$ from (3-4) under the composition $e \circ w$. This results in the class

$$\mathcal{Z}^{(4)}_{\mu_3} \in H^1\big(K_0, T_f^\vee \otimes_{\mathcal{O}} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[H[\mu_3]^{(p)}](-1)\big).$$

By Shapiro's lemma, its image under the isomorphism

$$H^1\big(K_0, T_f^\vee \otimes_{\mathcal{O}} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[H[\mu_3]^{(p)}](-1)\big) \simeq H^1\big(K_0[\mu_3], T_f^\vee(\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1})(-1)\big)$$

then defines

$$\mathcal{Z}^{(5)}_{\mu_3} \in H^1(K_0[\mu_3], T_f^\vee(\psi_1^{-1}\psi_2^{-1})(-1)).$$

The next lemma is in the same vein as [Castella and Do 2025, Lemma 2.1.3]:

Lemma 3.2.2. *Let λ_3 be a split prime in K_3 of norm ℓ coprime to mp, where ℓ splits in K_0. The diagram*

$$\begin{array}{ccc}
\mathrm{Ind}_{K_1}^{\mathbb{Q}} \mathcal{O}_{\psi_1^{-1}}[{}^1H^{(p)}_{\mu_1\lambda_1}] \otimes_{\mathcal{O}} \mathrm{Ind}_{K_2}^{\mathbb{Q}} \mathcal{O}_{\psi_2^{-1}}[{}^2H^{(p)}_{\mu_2\lambda_2}] & \longrightarrow & \mathrm{Ind}_{K_0}^{K_3} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[H[\mu_3\lambda_3]^{(p)}] \\
\Big\downarrow {\scriptstyle \mathrm{Norm}_{\mu_1}^{\mu_1\lambda_1} \otimes \mathrm{Norm}_{\mu_2}^{\mu_2\lambda_2}} & & \Big\downarrow {\scriptstyle \mathrm{Norm}_{\mu_3}^{\mu_3\lambda_3}} \\
\mathrm{Ind}_{K_1}^{\mathbb{Q}} \mathcal{O}_{\psi_1^{-1}}[{}^1H^{(p)}_{\mu_1}] \otimes_{\mathcal{O}} \mathrm{Ind}_{K_2}^{\mathbb{Q}} \mathcal{O}_{\psi_2^{-1}}[{}^2H^{(p)}_{\mu_2}] & \longrightarrow & \mathrm{Ind}_{K_0}^{K_3} \mathcal{O}_{\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}}[H[\mu_3]^{(p)}]
\end{array}$$

commutes, where the norm maps are the natural ones and the horizontal arrows are given by the composition $e \circ w$ in Proposition 3.2.1.

Using Lemma 3.2.2, one can show the following proposition. Similar to [Castella and Do 2025, Proposition 2.2.1], this is the key result for the construction of our anticyclotomic Euler system for $T_f^\vee(\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1})(-1)$ over the biquadratic field K_0.

Proposition 3.2.3. *The family $\{\mathcal{Z}^{(5)}_{\mu_3}\}$ satisfies the norm relation*

$$\begin{aligned}
\mathrm{Norm}_{K_0[\mu_3]}^{K_0[\mu_3\lambda_3]}(\mathcal{Z}^{(5)}_{\mu_3\lambda_3}) = (\ell-1)\Bigg(a_\ell(f) - \frac{\psi_1(\lambda_1)\psi_2(\bar{\lambda}_2)}{\ell}([\lambda_1]\times[\bar{\lambda}_2]) - \frac{\psi_1(\bar{\lambda}_1)\psi_2(\lambda_2)}{\ell}([\bar{\lambda}_1]\times[\lambda_2]) \\
+ (1-\ell)\frac{\psi_1(\lambda_1)\psi_2(\lambda_2)}{\ell^2}([\lambda_1]\times[\lambda_2])\Bigg)(\mathcal{Z}^{(5)}_{\mu_3}).
\end{aligned}$$

Proof. As in the proof of [loc. cit.], one has

$$\begin{aligned}
&(1 \otimes \mathcal{N}_{\mu_1}^{\mu_1\lambda_1} \otimes \mathcal{N}_{\mu_2}^{\mu_2\lambda_2})(\mathcal{Z}^{(5)}_{\mu_3\lambda_3}) \\
&= (\ell-1)\Bigg((T_\ell, 1, 1) - \frac{\psi_1(\lambda_1)[\lambda_1]}{\ell}(1, 1, T'_\ell) - \frac{\psi_2(\lambda_2)[\lambda_2]}{\ell}(1, T'_\ell, 1) \\
&\qquad + \frac{\psi_1(\lambda_1)\psi_2(\lambda_2)}{\ell^2}([\lambda_1]\times[\lambda_2])(\ell+1)\Bigg)(\mathcal{Z}^{(5)}_{\mu_3}) \\
&= (\ell-1)\Bigg(a_\ell(f) - \frac{\psi_1(\lambda_1)[\lambda_1]}{\ell}(\psi_2(\lambda_2)[\lambda_2] + \psi_2(\bar{\lambda}_2)[\bar{\lambda}_2]) \\
&\qquad - (\psi_1(\lambda_1)[\lambda_1] + \psi_1(\bar{\lambda}_1)[\bar{\lambda}_1])\frac{\psi_2(\lambda_2)[\lambda_2]}{\ell} + \frac{\psi_1(\lambda_1)\psi_2(\lambda_2)}{\ell^2}([\lambda_1]\times[\lambda_2])(\ell+1)\Bigg)(\mathcal{Z}^{(5)}_{\mu_3})
\end{aligned}$$

$$= (\ell - 1)\Big(a_\ell(f) - \frac{\psi_1(\lambda_1)\psi_2(\bar{\lambda}_2)}{\ell}([\lambda_1] \times [\bar{\lambda}_2]) - \frac{\psi_1(\bar{\lambda}_1)\psi_2(\lambda_2)}{\ell}([\bar{\lambda}_1] \times [\lambda_2]) + (1-\ell)\frac{\psi_1(\lambda_1)\psi_2(\lambda_2)}{\ell^2}([\lambda_1] \times [\lambda_2])\Big)(\mathcal{Z}^{(5)}_{\mu_3}).$$

This implies the result via combining Theorem 2.2.2 and Lemma 3.2.2. □

Following [Castella and Do 2025, Section 2.2] verbatim, which borrows ideas from [Darmon and Rotger 2017, Section 1.4], we can strip out the $\ell - 1$ factor by quotienting out the diamond operators action and obtain modified classes

$$\mathcal{Z}^{(6)}_{\mu_3} \in H^1(K_0[\mu_3], T_f^\vee(\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1})(-1)).$$

Then the term in the right-hand side of Proposition 3.2.3 can be massaged to agree with the local Euler factor at $\mathcal{L}_2$ of the Galois representation $[T_f^\vee(\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1})(-1)]^\vee(1) = T_f(\tilde{\psi}_1\tilde{\psi}_2)(2)$, giving the correct norm relations:

Theorem 3.2.4. *Suppose $p \nmid 6h_{K_0}$ and f is non-Eisenstein modulo $\mathfrak{P}$. Let $\mu_3 \in \mathcal{N}$ be a squarefree ideal of $\mathcal{O}_{K_3}$. Then there exists a collection of cohomology classes*

$$\mathcal{Z}_{\mu_3} \in H^1(K_0[\mu_3], T_f^\vee(\tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1})(-1))$$

such that, for every split prime $\lambda_3 \in \mathcal{L}$ of $\mathcal{O}_{K_3}$ of norm ℓ with $(\ell, m) = 1$, we have the norm relation

$$\mathrm{Norm}_{K_0[\mu_3]}^{K_0[\mu_3\lambda_3]}(\mathcal{Z}_{\mu_3\lambda_3}) = P_{\mathcal{L}_4}(\mathrm{Frob}_{\mathcal{L}_4/\lambda_3})(\mathcal{Z}_{\mu_3}),$$

where $P_{\mathcal{L}_4}(X) = \det(1 - X \cdot \mathrm{Frob}_{\mathcal{L}_4/\lambda_3} | T_f(\tilde{\psi}_1\tilde{\psi}_2)(2))$.

Proof. The proof is parallel to the one of [Castella and Do 2025, Theorem 2.2.7]. First note that $[\lambda_1] \times [\bar{\lambda}_2]$ corresponds to $\mathrm{Frob}_{\mathcal{L}_4/\lambda_3} \in H[\mu_3]^{(p)}$ under the map $e \circ w$ of Proposition 3.2.1. One then multiplies the class $\mathcal{Z}^{(6)}_{\mu_3\lambda_3}$ with $-\psi_1(\lambda_1)\psi_2(\bar{\lambda}_2)([\lambda_1] \times [\bar{\lambda}_2])$. From $\psi_1\psi_2((\ell)) = \chi_{\psi_1}\chi_{\psi_2}(\ell)\ell^2 = \epsilon_{K_1}(\ell)^{-1}\epsilon_{K_2}(\ell)^{-1}\ell^2 = \ell^2$ since ℓ splits in K_0, $\psi_1(\lambda_1)\psi_2(\bar{\lambda}_2) = \tilde{\psi}_1(\mathcal{L}_4)\tilde{\psi}_2(\mathcal{L}_4)$, and the fact that $[\ell] \times [\ell]$ maps to the identity element inside the ring class group, together with [Rubin 2000, Lemmas 9.6.1 and 9.6.3], the result follows from the explicit formula of Proposition 3.2.3. □

3.3. ***Construction for general weights and wild norm relations.*** We now extend the above construction to other weights $(k, k_1, k_2) \in \mathbb{Z}^3_{\geq 1}$ following [Castella and Do 2025, Section 2.3]. Then we show that the constructed cohomology classes also satisfy the wild norm relations for the anticyclotomic $\mathbb{Z}_p^2$-extension of K_0.

First, we assume that $p \nmid 6h_{K_0}$. Assume further that p splits in K_0, i.e.,

$$(p) = \mathcal{P}_1\mathcal{P}_2\mathcal{P}_3\mathcal{P}_4 \text{ in } K_0,$$

and $\mathcal{P}_2 = \tau_3\mathcal{P}_1$, $\mathcal{P}_3 = \tau_2\mathcal{P}_1$, $\mathcal{P}_4 = \tau_1\mathcal{P}_1$. Hence,

$$(p) = \mathfrak{p}_1\bar{\mathfrak{p}}_1 \quad \text{in } K_1, \qquad (p) = \mathfrak{p}_2\bar{\mathfrak{p}}_2 \quad \text{in } K_2,$$

with $\mathcal{P}_1$ the prime of K_0 above p induced by our fixed embedding $i_p : \overline{\mathbb{Q}} \hookrightarrow \overline{\mathbb{Q}}_p$, and $\mathcal{P}_1$ lies above $\mathfrak{p}_i$, the prime of K_i, for $i \in \{1, 2\}$. Note that the numbering here is parallel to our convention in Section 3.1.

Let $\Gamma_{\mathfrak{p}_i}^{K_i}$ be the Galois group of the unique $\mathbb{Z}_p$-extension of K_i unramified outside $\mathfrak{p}_i$. There exists a unique Hecke character $\psi_{0,i}$ of K_i of infinity type $(-1, 0)$ and conductor $\mathfrak{p}_i$ such that its p-adic avatar factors through $\Gamma_{\mathfrak{p}_i}^{K_i}$. The character ψ_i fixed at the beginning of this section can be decomposed as

$$\psi_i = \xi_i \psi_{0,i}^{k_i - 1},$$

where ξ_i is a ray class character of K of conductor dividing $\mathfrak{f}_i \mathfrak{p}_i$. Noting that $\Gamma_{\mathfrak{p}_i}^{K_i}$ is a quotient of ${}^iH_{\mathfrak{f}_i \mathfrak{p}_i^\infty}$ allows us to view $\psi_{0,i}$ and ξ_i as characters of ${}^iH_{\mathfrak{f}_i \mathfrak{p}_i^\infty}$. The formal q-expansion

$$\boldsymbol{\theta}_{\xi_i}(q) = \sum_{(\mathfrak{a}, \mathfrak{f}_i \mathfrak{p}_i) = 1} \xi_i \psi_{0,i}(\mathfrak{a})[\mathfrak{a}] q^{N_{K/\mathbb{Q}}(\mathfrak{a})} \in \Lambda_{\mathfrak{p}_i}[\![q]\!],$$

where $\Lambda_{\mathfrak{p}_i} = \mathcal{O}[\![\Gamma_{\mathfrak{p}_i}^{K_i}]\!] \simeq \mathcal{O}[\![\Gamma]\!]$ and $\Gamma = 1 + p\mathbb{Z}_p$, is the Hida family passing through θ_{ψ_i} (the specialization of $\boldsymbol{\theta}_{\xi_i}$ at weight k_i and trivial character recovers the ordinary p-stabilization of θ_{ψ_i}). Here we identify $\Gamma_{\mathfrak{p}_i}^{K_i}$ with $1 + p\mathbb{Z}_p$ via the (geometrically normalized) local Artin map.

Let $\boldsymbol{f}$ be the Hida family associated to f. Let $\boldsymbol{g} = \boldsymbol{\theta}_{\xi_1}$ and $\boldsymbol{h} = \boldsymbol{\theta}_{\xi_2}$ be the CM Hida families associated to ψ_1 and ψ_2, respectively. Denote by κ_f, κ_g and κ_h the Dirichlet characters modulo p giving the p-part of the tame characters of $\boldsymbol{f}$, $\boldsymbol{g}$ and $\boldsymbol{h}$, respectively.

Under the assumption that $\xi_i \psi_{0,i} \not\equiv \omega \pmod{\mathfrak{P}}$ for $i \in \{1, 2\}$, following equation (2.17) from [Castella and Do 2025, Section 2.3] and its notation, we have the $G_{\mathbb{Q}}$-equivariant maps

$$\begin{aligned} H^1(\Gamma(m, p), \mathcal{D}'_{\kappa_1}) \otimes \mathcal{O}[\![{}^1H^{(p)}_{\mathfrak{f}_1 \mu_1 \mathfrak{p}_1^\infty}]\!] &\to \mathrm{Ind}_{K_1}^{\mathbb{Q}} \mathcal{O}_{(\xi_1 \psi_{0,1})^{-1}}[{}^1H^{(p)}_{\mu_1}][\![\Gamma_{\mathfrak{p}_1}^{K_1}]\!], \\ H^1(\Gamma(m, p), \mathcal{D}'_{\kappa_2}) \otimes \mathcal{O}[\![{}^2H^{(p)}_{\mathfrak{f}_2 \mu_2 \mathfrak{p}_2^\infty}]\!] &\to \mathrm{Ind}_{K_2}^{\mathbb{Q}} \mathcal{O}_{(\xi_2 \psi_{0,2})^{-1}}[{}^2H^{(p)}_{\mu_2}][\![\Gamma_{\mathfrak{p}_2}^{K_2}]\!], \end{aligned} \tag{3-6}$$

where $\Gamma(m, p) = \Gamma_1(Nm) \cap \Gamma_0(p)$ is a congruence subgroup. Focusing on the class $\boldsymbol{\kappa}_m^{(2)}$ in equation (2.15) of [loc. cit.], we first tensor it with $\mathcal{O}[{}^1H^{(p)}_{\mathfrak{f}_1 \mu_1 \mathfrak{p}_1^r}]$ and $\mathcal{O}[{}^2H^{(p)}_{\mathfrak{f}_2 \mu_2 \mathfrak{p}_2^r}]$, let $r \to \infty$, and then arrive at

$$\begin{aligned} \boldsymbol{\mathcal{Z}}_{\mu_3}^{(1)} \in H^1\Big(\mathbb{Q}, H^1(\Gamma(1, p), \mathcal{D}'_{\kappa_f}) \hat{\otimes}_{\mathcal{O}} \big(H^1(\Gamma(m, p), \mathcal{D}'_{\kappa_g}) \otimes \mathcal{O}[\![{}^1H^{(p)}_{\mathfrak{f}_1 \mu_1 \mathfrak{p}_1^\infty}]\!]\big) \\ \hat{\otimes}_{\mathcal{O}[D_m]} \big(H^1(\Gamma(m, p), \mathcal{D}'_{\kappa_h}) \otimes \mathcal{O}[\![{}^2H^{(p)}_{\mathfrak{f}_2 \mu_2 \mathfrak{p}_2^\infty}]\!]\big)(2 - \kappa^*_{\boldsymbol{fgh}})\Big). \end{aligned}$$

Now choose a level-N test vector for f, denoted by $\breve{f}$. It comes with a specialization map

$$\pi_f : H^1(\Gamma(1, p), \mathcal{D}'_{\kappa_f})(1) \to T_f^\vee. \tag{3-7}$$

Under the natural maps induced by (3-6) and (3-7), the image of $\boldsymbol{\mathcal{Z}}_{\mu_3}^{(1)}$ is then $\boldsymbol{\mathcal{Z}}_{\mu_3}^{(2)}$ in

$$H^1\big(\mathbb{Q}, T_f^\vee \otimes_{\mathcal{O}} (\mathrm{Ind}_{K_1}^{\mathbb{Q}} \mathcal{O}_{(\xi_1 \psi_{0,1})^{-1}}[{}^1H^{(p)}_{\mu_1}][\![\Gamma_{\mathfrak{p}_1}^{K_1}]\!]) \hat{\otimes}_{\mathcal{O}[D_m]} (\mathrm{Ind}_{K_2}^{\mathbb{Q}} \mathcal{O}_{(\xi_2 \psi_{0,2})^{-1}}[{}^2H^{(p)}_{\mu_2}][\![\Gamma_{\mathfrak{p}_2}^{K_2}]\!])(-1 - \kappa^*_{\boldsymbol{fgh}})\big).$$

We first follow (3-3) and then apply the diagonal map $e \circ w$ in Proposition 3.2.1. This induces the class

$$\boldsymbol{\mathcal{Z}}_{\mu_3}^{(3)} \in H^1\big(K_0, T_f^\vee\big(1 - \tfrac{1}{2}k\big) \otimes_{\mathcal{O}} \mathrm{Ind}_{K_0[\mu_3]}^{K_0} \Lambda_{\mathcal{O}}(\tilde{\psi}_1^{-1} \tilde{\psi}_2^{-1} \tilde{\kappa}_{\mathrm{ac},1}^{(k_1-2)/2} \tilde{\kappa}_{\mathrm{ac},2}^{(k_2-2)/2} \boldsymbol{\kappa}_{\mathrm{ac}}^{-1})\big(1 - \tfrac{1}{2}(k_1 + k_2)\big)\big). \tag{3-8}$$

Here, for $i \in \{1,2\}$, we identify $\Gamma_i^- = \mathrm{Gal}(K_{i,\infty}^-/K_i)$ with the antidiagonal in $(1+p\mathbb{Z}_p)\times(1+p\mathbb{Z}_p) \simeq \mathcal{O}_{K_i,\mathfrak{p}_i}^{(1)} \times \mathcal{O}_{K_i,\bar{\mathfrak{p}}_i}^{(1)}$ via the geometric normalized Artin map, and define

$$\kappa_{\mathrm{ac},i} : \Gamma_i^- \to \mathbb{Z}_p^\times, \quad ((1+p)^{-1/2}, (1+p)^{1/2}) \mapsto (1+p);$$

compare this with [Castella and Do 2025, equation (2.19)]. Note that here we identify the anticyclotomic $\mathbb{Z}_p$-extension of K_i with the unique $\mathbb{Z}_p$-extension of K_i unramified outside $\mathfrak{p}_i$ (projection to the anticyclotomic part introduces a square root, $\gamma \mapsto \gamma^{(1-c)/2}$; see also [Hida 2015, page 636]).

We then identify the anticyclotomic $\mathbb{Z}_p^2$-extension $\Gamma^- = \mathrm{Gal}(K_{0,\infty}^-/K_0)$ of K_0 with $\Gamma_{\mathfrak{p}_1}^{K_1} \times \Gamma_{\mathfrak{p}_2}^{K_2}$ via the diagram

$$\begin{array}{ccccc}
(\mathcal{O}_{K_0,\mathcal{P}_1}^{(1)} \times \mathcal{O}_{K_0,\mathcal{P}_2}^{(1)}) \times (\mathcal{O}_{K_0,\mathcal{P}_3}^{(1)} \times \mathcal{O}_{K_0,\mathcal{P}_4}^{(1)}) & \twoheadrightarrow & \dfrac{(1+p\mathbb{Z}_p)\times(1+p\mathbb{Z}_p)}{\mathrm{diag}} \times \dfrac{(1+p\mathbb{Z}_p)\times(1+p\mathbb{Z}_p)}{\mathrm{diag}} & \xrightarrow{\simeq} & \mathbb{Z}_p^2 \\
\downarrow & & \downarrow & & \downarrow \\
(\mathcal{O}_{K_1,\mathfrak{p}_1}^{(1)} \times \mathcal{O}_{K_1,\bar{\mathfrak{p}}_1}^{(1)}) \times (\mathcal{O}_{K_2,\mathfrak{p}_2}^{(1)} \times \mathcal{O}_{K_2,\bar{\mathfrak{p}}_2}^{(1)}) & \twoheadrightarrow & \dfrac{(1+p\mathbb{Z}_p)\times(1+p\mathbb{Z}_p)}{\mathrm{diag}} \times \dfrac{(1+p\mathbb{Z}_p)\times(1+p\mathbb{Z}_p)}{\mathrm{diag}} & \xrightarrow{\simeq} & \mathbb{Z}_p \times \mathbb{Z}_p
\end{array} \tag{3-9}$$

Let $\Lambda^- = \mathbb{Z}_p[\![\Gamma^-]\!]$ and define further $\boldsymbol{\kappa}_{\mathrm{ac}} : \Gamma^- \to \Lambda^\times$, where

$$((1+p)^{-1/2}, (1+p)^{1/2}, (1+p)^{-1/2}, (1+p)^{1/2}) \mapsto [(1+p), (1+p)].$$

Given an $\mathcal{O}$-lattice T inside a G_{K_0}-representation V, Shapiro's lemma allows us to write

$$H^1(K_0, T \hat{\otimes}_{\mathcal{O}} \Lambda_{\mathcal{O}}^-(\boldsymbol{\kappa}_{\mathrm{ac}}^{-1})) \simeq H^1_{\mathrm{Iw}}(K_0[p^\infty], T),$$

where $H^1_{\mathrm{Iw}}(K_0[p^\infty], T) := \varprojlim_{r,s} H^1(K_0[\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s], T)$ with limit under the corestriction maps. Then the image of $\boldsymbol{\mathcal{Z}}_{\mu_3}^{(3)}$ in (3-8) under Shapiro's lemma is an Iwasawa cohomology class

$$\boldsymbol{\mathcal{Z}}_{\mu_3} \in H^1_{\mathrm{Iw}}\big(K_0[\mu_3 p^\infty], T_f^\vee\big(1-\tfrac{1}{2}k\big) \otimes \tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}\tilde{\kappa}_{\mathrm{ac},1}^{(k_1-2)/2}\tilde{\kappa}_{\mathrm{ac},2}^{(k_2-2)/2}\big(1-\tfrac{1}{2}(k_1+k_2)\big)\big) \tag{3-10}$$

for the conjugate self-dual representation $T_f^\vee\big(1-\frac{1}{2}k\big)$ twisted by the Hecke character

$$\chi^{-1} = \tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}\mathbf{N}^{1-(k_1+k_2)/2}$$

(up to an anticyclotomic twist). Here χ is anticyclotomic and of infinity type (corresponding to the order $(\mathcal{P}_1,\mathcal{P}_2,\mathcal{P}_3,\mathcal{P}_4)$ or $(1,\tau_3,\tau_2,\tau_1)$)

$$\big(\tfrac{1}{2}(2-k_1-k_2), \tfrac{1}{2}(k_1+k_2-2), \tfrac{1}{2}(k_1-k_2), \tfrac{1}{2}(k_2-k_1)\big).$$

Let

$$T_{f,\chi} = T_f^\vee\big(1-\tfrac{1}{2}k\big) \otimes \chi^{-1}. \tag{3-11}$$

Following the proof of Theorem 3.2.4 and invoking [Rubin 2000, Theorem 6.4.1], we can obtain a collection of Iwasawa cohomology classes for anticyclotomic twists (to eliminate $\tilde{\kappa}_{\mathrm{ac},1}^{(k_1-2)/2}\tilde{\kappa}_{\mathrm{ac},2}^{(k_2-2)/2}$). We thus arrive at the proof for the wild norm relation, which is formulated inside the following theorem.

Theorem 3.3.1. *Suppose $p \nmid 6h_{K_0}$ and f is non-Eisenstein modulo $\mathfrak{P}$. Let $\mu_3 \in \mathcal{N}$ and $m = N_{K_3/\mathbb{Q}}(\mu_3)$. Then there exists a collection of Iwasawa cohomology classes*

$$z_{f,\chi,\mu_3} \in H^1_{\mathrm{Iw}}(K_0[\mu_3 p^\infty], T_{f,\chi})$$

such that, for every split prime λ_3 of $\mathcal{O}_{K_3}$ of norm ℓ, where ℓ splits in K_0, with $(\ell, mp) = 1$, we have the norm relation

$$\mathrm{Norm}_{K_0[\mu_3]}^{K_0[\mu_3\lambda_3]}(z_{f,\chi,\mu_3\lambda_3}) = P_{\mathcal{L}_4}(\mathrm{Frob}_{\mathcal{L}_4})(z_{f,\chi,\mu_3}),$$

where $P_{\mathcal{L}_4}(X) = \det(1 - X \cdot \mathrm{Frob}_{\mathcal{L}_4} | (T_{f,\chi})^\vee(1))$.

4. Selmer groups

In this section, we show that the classes constructed in Theorem 3.3.1 land in certain Selmer groups defined in [Greenberg 1994]. Keeping the setup at the start of Section 3, we further assume that f is a p-ordinary newform of even weight $k \geq 2$ with $p \nmid N_f$.

Let χ be an anticyclotomic Hecke character of K_0 of infinity type $(-a, a, -b, b)$ for some integers $a, b \geq 0$. We will focus on the conjugate self-dual G_{K_0}-representation

$$V_{f,\chi} := V_f^\vee\big(1 - \tfrac{1}{2}k\big) \otimes \chi^{-1}.$$

Definition 4.1. For each prime $\mathcal{P} \in \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4\}$ of K_0 above p, we fix a $G_{K_{0,\mathcal{P}}}$-stable subspace $\mathscr{F}_\mathcal{P}^+(V_{f,\chi}) \subset V_{f,\chi}$ and let

$$\mathscr{F}_\mathcal{P}^-(V_{f,\chi}) = V_{f,\chi}/\mathscr{F}_\mathcal{P}^+(V_{f,\chi}).$$

Let L be a finite extension of K_0. The *Greenberg Selmer group* $\mathrm{Sel}_\mathscr{F}(L, V_{f,\chi})$ attached to $\mathscr{F} = \{\mathscr{F}_\mathcal{P}^+(V_{f,\chi})\}_{\mathcal{P} \mid p}$ is defined by

$$\mathrm{Sel}_\mathscr{F}(L, V_{f,\chi}) := \ker\left\{ H^1(L, V_{f,\chi}) \to \prod_w \frac{H^1(L_w, V_{f,\chi})}{H^1_\mathscr{F}(L_w, V_{f,\chi})} \right\}, \tag{4-1}$$

where w runs over the finite primes of L and the local conditions are given by

$$H^1_\mathscr{F}(L_w, V_{f,\chi}) = \begin{cases} \ker\{H^1(L_w, V_{f,\chi}) \to H^1(L_w^{\mathrm{ur}}, V_{f,\chi})\} & \text{if } w \nmid p, \\ \ker\{H^1(L_w, V_{f,\chi}) \to H^1(L_w, \mathscr{F}_\mathcal{P}^-(V_{f,\chi}))\} & \text{if } w \mid \mathcal{P} \mid p. \end{cases}$$

We fix a lattice $T_{f,\chi} \subset V_{f,\chi}$. Let $H^1_\mathscr{F}(L_w, T_{f,\chi})$ be the inverse image of $H^1_\mathscr{F}(L_w, V_{f,\chi})$ under the natural map

$$H^1(L_w, T_{f,\chi}) \to H^1(L_w, V_{f,\chi}).$$

This then defines $\mathrm{Sel}_\mathscr{F}(L, T_{f,\chi})$ as in (4-1). For any $\mathbb{Z}_p^2$-extension $L_\infty = \bigcup_{r,s} L_{r,s}$ of L, we put

$$\mathrm{Sel}_\mathscr{F}(L_\infty, T_{f,\chi}) := \varprojlim_{r,s} \mathrm{Sel}_\mathscr{F}(L_{r,s}, T_{f,\chi}),$$

where the inverse limit is taken with respect to the corestriction map. We also put $\mathrm{Sel}_\mathscr{F}(L_\infty, V_{f,\chi}) := \mathrm{Sel}_\mathscr{F}(L_\infty, T_{f,\chi}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. Note that this group is independent of the chosen lattice $T_{f,\chi}$.

Definition 4.2. We also define the *Bloch–Kato Selmer group* $\mathrm{Sel}_{\mathrm{BK}}(L, V_{f,\chi})$, following [Bloch and Kato 1990], as

$$\mathrm{Sel}_{\mathrm{BK}}(L, V_{f,\chi}) := \ker\left\{ H^1(L, V_{f,\chi}) \to \prod_w \frac{H^1(L_w, V_{f,\chi})}{H^1_f(L_w, V_{f,\chi})} \right\},$$

where the local conditions are given by

$$H^1_f(L_w, V_{f,\chi}) = \ker\{H^1(L_w, V_{f,\chi}) \to H^1(L_w^{\mathrm{ur}}, V_{f,\chi})\}$$

at primes $w \nmid p$, and the crystalline condition at primes $w \mid p$,

$$H^1_f(L_w, V_{f,\chi}) = \ker\{H^1(L_w, V_{f,\chi}) \to H^1(L_w, V_{f,\chi} \otimes \boldsymbol{B}_{\mathrm{cris}})\},$$

with $\boldsymbol{B}_{\mathrm{cris}}$ being Fontaine's crystalline period ring. The local conditions $H^1_f(L_w, T_{f,\chi}) \subset H^1(L_w, T_{f,\chi})$ are defined by propagation similarly.

Besides the crystalline condition, there are three local conditions at primes $\mathcal{P} \mid p$ that we will be interested in:

(1) The *strict* condition
$$\mathscr{F}^+_{\mathcal{P}}(V_{f,\chi}) = 0.$$

(2) The *relaxed* condition
$$\mathscr{F}^+_{\mathcal{P}}(V_{f,\chi}) = V_{f,\chi}.$$

(3) The *ordinary* condition, corresponding to the fact that the restriction of $V_{f,\chi}$ to $G_{\mathbb{Q}_p}$ is reducible (see (2-1)),
$$\mathscr{F}^+_{\mathcal{P}}(V_{f,\chi}) = V^+_{f,\chi} := V_f^{\vee,+}\big(1 - \tfrac{1}{2}k\big) \otimes \chi^{-1}.$$

Definition 4.3. Denote by $\mathrm{Sel}_{\alpha,\beta,\gamma,\delta}(K_0, V)$ the subgroup of $H^1(K_0, V)$ where classes are unramified at all primes $v \nmid p$ and satisfy the conditions α, β, γ, δ at $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, $\mathcal{P}_4$, respectively, where $\alpha, \beta, \gamma, \delta \in$ {rel, str, ord}, which correspond to the relaxed, strict, and ordinary condition, respectively.

We will now compute the explicit local conditions for the Bloch–Kato Selmer group. Here we shall adopt the convention that the p-adic cyclotomic character has Hodge–Tate weight -1. Thus, since χ has infinity type $(-a, a, -b, b)$, the p-adic avatar of χ has Hodge–Tate weight $a, -a, b, -b$ at $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4$, respectively.

Lemma 4.4. *Assume that $a \geq b$. For any finite extension L of K_0, we have*

$$\mathrm{Sel}_{\mathrm{BK}}(L, V_{f,\chi}) = \begin{cases} \mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(L, V_{f,\chi}) & \text{if } k \geq 2a+2, \\ \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(L, V_{f,\chi}) & \text{if } 2b+2 \leq k < 2a+2, \\ \mathrm{Sel}_{\mathrm{rel,str,rel,str}}(L, V_{f,\chi}) & \text{if } k < 2b+2. \end{cases}$$

Proof. By the Panchiskin condition [Bloch and Kato 1990, Theorem 4.1(ii)] (see also [Nekovář 2000, (3.1)–(3.2); Flach 1990, Lemma 2, page 125]), for every prime $w \mid \mathcal{P} \mid p$ of $L/K_0/\mathbb{Q}$, we have

$$H^1_f(L_w, V_{f,\chi}) = \mathrm{im}\{H^1(L_w, \mathrm{Fil}^1_{\mathcal{P}}(V_{f,\chi})) \to H^1(L_w, V_{f,\chi})\},$$

where $\mathrm{Fil}^1_{\mathcal{P}}(V_{f,\chi}) \subset V_{f,\chi}$ is a $G_{K_{\mathcal{P}}}$-stable subspace (assuming it exists) such that the Hodge–Tate weights of $\mathrm{Fil}^1_{\mathcal{P}}(V_{f,\chi})$ (resp. $V_{f,\chi}/\mathrm{Fil}^1_{\mathcal{P}}(V_{f,\chi})$) are all < 0 (resp. ≥ 0).

We compute the Hodge–Tate weights table of $V^+_{f,\chi}$ and $V^-_{f,\chi} := V_{f,\chi}/V^+_{f,\chi}$ at the primes of K_0 above p:

	$V^+_{f,\chi}$	$V^-_{f,\chi}$
HT weight at $\mathcal{P}_1$	$-a-\frac{1}{2}k$	$-a-1+\frac{1}{2}k$
HT weight at $\mathcal{P}_2$	$a-\frac{1}{2}k$	$a-1+\frac{1}{2}k$
HT weight at $\mathcal{P}_3$	$-b-\frac{1}{2}k$	$-b-1+\frac{1}{2}k$
HT weight at $\mathcal{P}_4$	$b-\frac{1}{2}k$	$b-1+\frac{1}{2}k$

This obtains the equalities in the lemma. □

Fix a choice of Galois stable subgroups $\mathscr{F} = \{\mathscr{F}^+_{\mathcal{P}}(V_{f,\chi})\}_{\mathcal{P}\mid p}$ and let

$$A_{f,\chi} := \mathrm{Hom}_{\mathbb{Z}_p}(T_{f,\chi}, \mu_{p^\infty}).$$

Define the associated *dual Selmer group* $\mathrm{Sel}_{\mathscr{F}^*}(L, A_{f,\chi})$ by

$$\mathrm{Sel}_{\mathscr{F}^*}(L, A_{f,\chi}) := \ker\left\{ H^1(L, A_{f,\chi}) \to \prod_w \frac{H^1(L_w, A_{f,\chi})}{H^1_{\mathscr{F}^*}(L_w, A_{f,\chi})} \right\},$$

where $H^1_{\mathscr{F}^*}(L_w, A_{f,\chi})$ is the orthogonal complement of $H^1_{\mathscr{F}}(L_w, T_{f,\chi})$ under local Tate pairing

$$H^1(L_w, T_{f,\chi}) \times H^1(L_w, A_{f,\chi}) \to \mathbb{Q}_p/\mathbb{Z}_p.$$

One can then compute the following:

(1) The dual Selmer group of $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(L, T_{f,\chi})$ consists of classes that are unramified outside p and have the strict, relaxed, ordinary, ordinary condition at $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4$, respectively. Compatibly with Definition 4.3, this can be written as $\mathrm{Sel}_{\mathrm{str,rel,ord,ord}}(L, A_{f,\chi})$.

(2) The dual Selmer group of $\mathrm{Sel}_{\mathrm{rel,str,rel,str}}(L, T_{f,\chi})$ consists of classes that are unramified outside p and have the strict, relaxed, strict, relaxed condition at $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4$, respectively. Compatibly with Definition 4.3, this can be written as $\mathrm{Sel}_{\mathrm{str,rel,str,rel}}(L, A_{f,\chi})$.

(3) The dual Selmer group of $\mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(L, T_{f,\chi})$ consists of classes that are unramified outside p, and land in the image of the natural map

$$H^1(L_w, \mathscr{F}^+_{\mathcal{P}}(A_{f,\chi})) \to H^1(L_w, A_{f,\chi}), \quad \mathscr{F}^+_{\mathcal{P}}(A_{f,\chi}) := \mathrm{Hom}_{\mathbb{Z}_p}(\mathscr{F}^-_{\mathcal{P}}(T_{f,\chi}), \mu_{p^\infty}),$$

for $w \mid \mathcal{P} \mid p$. Compatibly with Definition 4.3, this can be written as $\mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(L, A_{f,\chi})$.

5. Triple product p-adic L-function and Selmer group

Here, we will recall some conventions on Hida families, triple product p-adic L-functions ($\boldsymbol{f}$-unbalanced), and Selmer groups (balanced and $\boldsymbol{f}$-unbalanced) following [Hsieh 2021].

5.1. ***Hida families.*** We follow the convention of [Hsieh 2021, Section 3.1]. Let $\mathcal{O}$ be the ring of integers of a finite extension of $\mathbb{Q}_p$. Let $\mathbb{I}$ be a normal domain, finite flat over the Iwasawa algebra

$$\Lambda := \mathcal{O}[\![1 + p\mathbb{Z}_p]\!].$$

Let N be a positive integer prime to p and $\chi : (\mathbb{Z}/Np\mathbb{Z})^\times \to \mathcal{O}^\times$ be a Dirichlet character. Denote by $S^o(N, \chi, \mathbb{I}) \subset \mathbb{I}[\![q]\!]$ the space of ordinary $\mathbb{I}$-adic cusp forms of tame level N and branch character χ.

Let $\mathfrak{X}_{\mathbb{I}}^+ \subset \operatorname{Spec} \mathbb{I}(\overline{\mathbb{Q}}_p)$ be the set of arithmetic points of $\mathbb{I}$, which consists of the ring homomorphisms $Q : \mathbb{I} \to \overline{\mathbb{Q}}_p$ such that, for some $k_Q \in \mathbb{Z}_{\geq 2}$, called the weight of Q, and $\epsilon_Q(z) \in \mu_{p^\infty}$,

$$Q|_{1+p\mathbb{Z}_p} : z \mapsto z^{k_Q - 1}\epsilon_Q(z).$$

We say that $\boldsymbol{f} = \sum_{n=1}^\infty a_n(\boldsymbol{f})q^n \in S^o(N, \chi, \mathbb{I})$ is a primitive Hida family if the specialization $\boldsymbol{f}_Q$ for every $Q \in \mathfrak{X}_{\mathbb{I}}^+$ gives the q-expansion of an ordinary p-stabilized newform of weight k_Q and tame conductor N. Let $\mathfrak{X}_{\mathbb{I}}^{\mathrm{cls}} \subset \operatorname{Spec} \mathbb{I}(\overline{\mathbb{Q}}_p)$ be the set of ring homomorphisms Q as above, with $k_Q \in \mathbb{Z}_{\geq 1}$, such that $\boldsymbol{f}_Q$ is the q-expansion of a classical modular form.

Given $\boldsymbol{f}$ a primitive Hida family of tame conductor N, one can associate a Galois representation

$$\rho_{\boldsymbol{f}} : G_{\mathbb{Q}} \to \mathrm{Aut}_{\mathbb{I}}(V_{\boldsymbol{f}}) \simeq \mathrm{GL}_2(\mathbb{I}),$$

where the determinant of $\rho_{\boldsymbol{f}}$ is $\chi_{\mathbb{I}} \cdot \varepsilon_{\mathrm{cyc}}$; see [Hsieh 2021, Section 3.2]. By [Wiles 1988, Theorem 2.2.2], the restriction of $V_{\boldsymbol{f}}$ to $G_{\mathbb{Q}_p}$ is reducible and one has a short exact sequence

$$0 \to V_{\boldsymbol{f}}^+ \to V_{\boldsymbol{f}} \to V_{\boldsymbol{f}}^- \to 0.$$

Here the quotient $V_{\boldsymbol{f}}^-$ is free of rank one over $\mathbb{I}$, with $G_{\mathbb{Q}_p}$ acting via the unramified character sending an arithmetic Frobenius Frob_p^{-1} to $a_p(\boldsymbol{f})$. Let $\mathbb{T}(N, \mathbb{I})$ be the Hecke algebra acting on $\bigoplus_\chi S^o(N, \chi, \mathbb{I})$, where χ runs over the characters of $(\mathbb{Z}/Np\mathbb{Z})^\times$. There is a $\mathbb{I}$-algebra homomorphism attached to $\boldsymbol{f}$,

$$\lambda_{\boldsymbol{f}} : \mathbb{T}(N, \mathbb{I}) \to \mathbb{I},$$

that factors through a local component $\mathbb{T}_{\mathfrak{m}}$, where $\mathfrak{m}$ is the maximal ideal containing $\ker \lambda_{\boldsymbol{f}}$. Following [Hida 1988], we define the congruence ideal $C(\boldsymbol{f})$ of $\boldsymbol{f}$ by

$$C(\boldsymbol{f}) := \lambda_{\boldsymbol{f}}(\mathrm{Ann}_{\mathbb{T}_{\mathfrak{m}}}(\ker \lambda_{\boldsymbol{f}})) \subset \mathbb{I}.$$

Under the assumption that the residual representation $\bar{\rho}_{\boldsymbol{f}}$ is absolutely irreducible and p-distinguished, Wiles [1995] and Hida [1988] prove that $C(\boldsymbol{f})$ is generated by a nonzero element $\eta_{\boldsymbol{f}} \in \mathbb{I}$.

5.2. ***CM Hida families revisited.*** We explicitly construct CM Hida families, following the exposition in [Hsieh 2021, Section 8.1]. Let K be an imaginary quadratic field of discriminant $-D_K < 0$, and suppose that $p = \mathfrak{p}\bar{\mathfrak{p}}$ splits in K, with $\mathfrak{p}$ the prime of K above p induced by our fixed embedding $\iota_p : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$.

Let K_∞ be the $\mathbb{Z}_p^2$-extension of K. Let $K(\mathfrak{p}^\infty)$ be the maximal subfield of K_∞ unramified outside $\mathfrak{p}$. Put

$$\Gamma_\infty := \mathrm{Gal}(K_\infty/K) \simeq \mathbb{Z}_p^2, \quad \Gamma_{\mathfrak{p}} := \mathrm{Gal}(K(\mathfrak{p}^\infty)/K) \simeq \mathbb{Z}_p.$$

For every ideal $\mathfrak{c} \subset \mathcal{O}_K$, recall that $K_\mathfrak{c}$ is the ray class field of K of conductor $\mathfrak{c}$. Using our notation, $K(\mathfrak{p}^\infty)$ is the maximal $\mathbb{Z}_p$-extension of K inside $K_{\mathfrak{p}^\infty}$. Denote by $\mathrm{Art}_\mathfrak{p}$ the restriction of the Artin map to $K_\mathfrak{p}^\times$, with geometric normalization. Then $\mathrm{Art}_\mathfrak{p}$ induces an embedding $1 + p\mathbb{Z}_p \to \Gamma_\mathfrak{p}$, where we identified $\mathbb{Z}_p^\times$ and $\mathcal{O}_{K_\mathfrak{p}}^\times$ via ι_p. Let $\gamma_\mathfrak{p}$ be the image of $1 + p$; hence, it will be a topological generator of $\Gamma_\mathfrak{p}$.

For each variable S, let $\Psi_S : \Gamma_\infty \to \mathcal{O}[\![S]\!]^\times$ be the universal character given by

$$\Psi_S(\sigma) = (1+S)^{l(\sigma)},$$

where $l(\sigma) \in \mathbb{Z}_p$ is such that $\sigma|_{K(\mathfrak{p}^\infty)} = \gamma_\mathfrak{p}^{l(\sigma)}$. Now assume that $\mathfrak{c}$ is prime to p. Given a finite-order character $\xi : G_K \to \mathcal{O}^\times$ of conductor dividing $\mathfrak{c}$, let

$$\boldsymbol{\theta}_\xi(S)(q) = \sum_{(\mathfrak{a},\mathfrak{p}\mathfrak{c})=1} \xi(\sigma_\mathfrak{a})\Psi^{-1}_{(1+S)/(1+p)-1}(\sigma_\mathfrak{a}) q^{N_{K/\mathbb{Q}}(\mathfrak{a})} \in \mathcal{O}[\![S]\!][\![q]\!],$$

where $\sigma_\mathfrak{a} \in \mathrm{Gal}(K_{\mathfrak{c}\mathfrak{p}^\infty}/K)$ is the Artin symbol of $\mathfrak{a}$. Then $\boldsymbol{\theta}_\xi(S)$ is a Hida family defined over $\mathcal{O}[\![S]\!]$ of tame level $N_{K/\mathbb{Q}}(\mathfrak{c})D_K$ and tame character $(\xi \circ \mathscr{V})\epsilon_K\omega^{-1}$, where $\mathscr{V} : G_\mathbb{Q}^{\mathrm{ab}} \to G_K^{\mathrm{ab}}$ is the transfer map and ϵ_K is the quadratic character corresponding to $K/\mathbb{Q}$.

5.3. ***Triple products of Hida families.*** Let

$$\boldsymbol{f} \in S^o(N_f, \chi_f, \mathbb{I}_f), \quad \boldsymbol{g} \in S^o(N_g, \chi_g, \mathbb{I}_g), \quad \boldsymbol{h} \in S^o(N_h, \chi_h, \mathbb{I}_h)$$

be three primitive Hida families such that

$$\chi_f\chi_g\chi_h = \omega^{2a} \quad \text{for some } a \in \mathbb{Z}, \tag{5-1}$$

where ω is the Teichmüller character. Let

$$\mathcal{R} = \mathbb{I}_f \hat{\otimes}_\mathcal{O} \mathbb{I}_g \hat{\otimes}_\mathcal{O} \mathbb{I}_h$$

be a finite extension of the three-variable Iwasawa algebra $\Lambda \hat{\otimes}_\mathcal{O} \Lambda \hat{\otimes}_\mathcal{O} \Lambda$.

Let $\mathfrak{X}_\mathcal{R}^+ \subset \mathrm{Spec}\,\mathcal{R}(\overline{\mathbb{Q}}_p)$ be the weight space of $\mathcal{R}$ given by

$$\mathfrak{X}_\mathcal{R}^+ := \{\underline{Q} = (Q_1, Q_2, Q_3) \in \mathfrak{X}_{\mathbb{I}_f}^+ \times \mathfrak{X}_{\mathbb{I}_g}^{\mathrm{cls}} \times \mathfrak{X}_{\mathbb{I}_h}^{\mathrm{cls}} : k_{Q_1} + k_{Q_2} + k_{Q_3} \equiv 0 \pmod 2\}.$$

One can then partition $\mathfrak{X}_\mathcal{R}^+ = \mathfrak{X}_\mathcal{R}^{\mathrm{bal}} \sqcup \mathfrak{X}_\mathcal{R}^{\boldsymbol{f}} \sqcup \mathfrak{X}_\mathcal{R}^{\boldsymbol{g}} \sqcup \mathfrak{X}_\mathcal{R}^{\boldsymbol{h}}$ into

(1) the set of *balanced* weights

$$\mathfrak{X}_\mathcal{R}^{\mathrm{bal}} := \{\underline{Q} \in \mathfrak{X}_\mathcal{R}^+ : k_{Q_1} + k_{Q_2} + k_{Q_3} > 2k_{Q_i} \text{ for all } i \in \{1,2,3\}\},$$

(2) the set of $\boldsymbol{f}$-unbalanced weights

$$\mathfrak{X}_\mathcal{R}^{\boldsymbol{f}} := \{\underline{Q} \in \mathfrak{X}_\mathcal{R}^+ : k_{Q_1} \geq k_{Q_2} + k_{Q_3}\},$$

(3) the set of $\boldsymbol{g}$-unbalanced weights

$$\mathfrak{X}_\mathcal{R}^{\boldsymbol{g}} := \{\underline{Q} \in \mathfrak{X}_\mathcal{R}^+ : k_{Q_2} \geq k_{Q_1} + k_{Q_3}\},$$

(4) the set of $\boldsymbol{h}$-unbalanced weights

$$\mathfrak{X}_{\mathcal{R}}^{\boldsymbol{h}} := \{\underline{Q} \in \mathfrak{X}_{\mathcal{R}}^{+} : k_{Q_3} \geq k_{Q_1} + k_{Q_2}\}.$$

Let $\boldsymbol{V} = V_{\boldsymbol{f}} \hat{\otimes}_{\mathcal{O}} V_{\boldsymbol{g}} \hat{\otimes}_{\mathcal{O}} V_{\boldsymbol{h}}$ be the triple tensor product Galois representation attached to $(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$. By (5-1), one can decompose the determinant of $\boldsymbol{V}$ as $\det \boldsymbol{V} = \mathcal{X}^2 \varepsilon_{\mathrm{cyc}}$. Put

$$\boldsymbol{V}^{\dagger} := \boldsymbol{V} \otimes \mathcal{X}^{-1}. \tag{5-2}$$

This is a self-dual twist of $\boldsymbol{V}$. For any $\underline{Q} = (Q_1, Q_2, Q_3) \in \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{f}}$, denote by $\boldsymbol{V}_{\underline{Q}}^{\dagger}$ the corresponding specialization.

For each prime ℓ, let $\varepsilon_{\ell}(\boldsymbol{V}_{\underline{Q}}^{\dagger})$ be the epsilon factor attached to the local representation $\boldsymbol{V}_{\underline{Q}}^{\dagger}|_{G_{\mathbb{Q}_\ell}}$ (see [Tate 1979, page 21]). We assume that, for some $\underline{Q} \in \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{f}}$, we have

$$\varepsilon_{\ell}(\boldsymbol{V}_{\underline{Q}}^{\dagger}) = +1 \quad \text{for all prime factors } \ell \text{ of } N_f N_g N_h. \tag{5-3}$$

Note that condition (5-3) is independent of $\underline{Q}$ (see [Hsieh 2021, Section 1.2]). Furthermore, it implies that the sign of the functional equation for the triple product L-function (with center at $s = 0$)

$$L(\boldsymbol{V}_{\underline{Q}}^{\dagger}, s)$$

is $+1$ (resp. -1) for all $\underline{Q} \in \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{f}} \cup \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{g}} \cup \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{h}}$ (resp. $\underline{Q} \in \mathfrak{X}_{\mathcal{R}}^{\mathrm{bal}}$).

Theorem 5.3.1 [Hsieh 2021, Theorem A]. *Let $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}$ be three primitive Hida families satisfying conditions (5-1) and (5-3). Assume also that* $\gcd(N_f, N_g, N_h)$ *is squarefree, and the residual representation $\bar{\rho}_{\boldsymbol{f}}$ is absolutely irreducible and p-distinguished. Fix a generator $\eta_{\boldsymbol{f}}$ of the congruence ideal of $\boldsymbol{f}$. Then there exists a unique element*

$$\mathscr{L}_p^{\boldsymbol{f}, \eta_{\boldsymbol{f}}}(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}) \in \mathcal{R}$$

such that, for all $\underline{Q} = (Q_1, Q_2, Q_3) \in \mathfrak{X}_{\mathcal{R}}^{\boldsymbol{f}}$ of weight (k_1, k_2, k_3) with $\epsilon_{Q_1} = 1$, we have

$$(\mathscr{L}_p^{\boldsymbol{f}, \eta_{\boldsymbol{f}}}(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})(\underline{Q}))^2 = \Gamma_{\boldsymbol{V}_{\underline{Q}}^{\dagger}}(0) \cdot \frac{L(\boldsymbol{V}_{\underline{Q}}^{\dagger}, 0)}{(\sqrt{-1})^{2k_1} \cdot \Omega_{\boldsymbol{f}_{Q_1}}^2} \cdot \mathcal{E}_p(\mathscr{F}_p^{\boldsymbol{f}}(\boldsymbol{V}_{\underline{Q}}^{\dagger})) \cdot \prod_{q \in \Sigma_{\mathrm{exc}}} (1 + q^{-1})^2,$$

where:

- $\Gamma_{\boldsymbol{V}_{\underline{Q}}^{\dagger}}(0) = 16(2\pi)^{-2k_1} \Gamma(w_{\underline{Q}}) \Gamma(w_{\underline{Q}} + 2 - k_2 - k_3) \Gamma(w_{\underline{Q}} + 1 - k_2) \Gamma(w_{\underline{Q}} + 1 - k_3)$ *and* $w_{\underline{Q}} = \frac{1}{2}(k_1 + k_2 + k_3 - 2)$.
- $\Omega_{\boldsymbol{f}_{Q_1}}$ *is the Hida canonical period*

$$\Omega_{\boldsymbol{f}_{Q_1}} := (-2\sqrt{-1})^{k_1+1} \cdot \frac{\|\boldsymbol{f}_{Q_1}^{\circ}\|_{\Gamma_0(N_f)}^2}{\eta_{\boldsymbol{f}_{Q_1}}} \cdot \left(1 - \frac{\chi_f'(p) p^{k_1 - 1}}{\alpha_{Q_1}^2}\right)\left(1 - \frac{\chi_f'(p) p^{k_1 - 2}}{\alpha_{Q_1}^2}\right),$$

with $\boldsymbol{f}_{Q_1}^{\circ} \in S_{k_1}(\Gamma_0(N_f))$ the newform of conductor N_f associated with $\boldsymbol{f}_{Q_1}$, χ_f' the prime-to-p part of χ_f, and α_{Q_1} the specialization of $a_p(\boldsymbol{f}) \in \mathbb{I}_f^{\times}$ at Q_1.

- $\mathcal{E}_p(\mathscr{F}_p^f(\underline{\underline{V}}_Q^\dagger))$ *is the modified* p*-Euler factor and* Σ_{exc} *is an explicitly defined subset of the prime factors of* $N_f N_g N_h$ [Hsieh 2021, page 416].

5.4. ***Triple product Selmer groups.*** Recall from (5-2) that $\boldsymbol{V}^\dagger = \boldsymbol{V} \otimes \mathcal{X}^{-1}$ is the self-dual twist of the Galois representation associated to a triple of primitive Hida families $(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ given by (5-1).

Definition 5.4.1. Let

$$\mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}^\dagger) := (V_{\boldsymbol{f}} \otimes V_{\boldsymbol{g}}^+ \otimes V_{\boldsymbol{h}}^+ + V_{\boldsymbol{f}}^+ \otimes V_{\boldsymbol{g}} \otimes V_{\boldsymbol{h}}^+ + V_{\boldsymbol{f}}^+ \otimes V_{\boldsymbol{g}}^+ \otimes V_{\boldsymbol{h}}) \otimes \mathcal{X}^{-1},$$

and define the balanced local condition $\mathrm{H}^1_{\mathrm{bal}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger)$ by

$$\mathrm{H}^1_{\mathrm{bal}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger) := \mathrm{im}\big(\mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}^\dagger)) \to \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}^\dagger)\big).$$

Similarly, let

$$\mathscr{F}_p^{\boldsymbol{f}}(\boldsymbol{V}^\dagger) = (V_{\boldsymbol{f}}^+ \otimes V_{\boldsymbol{g}} \otimes V_{\boldsymbol{h}}) \otimes \mathcal{X}^{-1}, \quad \mathscr{F}_p^{\boldsymbol{h}}(\boldsymbol{V}^\dagger) = (V_{\boldsymbol{f}} \otimes V_{\boldsymbol{g}} \otimes V_{\boldsymbol{h}}^+) \otimes \mathcal{X}^{-1}.$$

Define the $\boldsymbol{f}$-unbalanced local condition $\mathrm{H}^1_{\boldsymbol{f}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger)$ by

$$\mathrm{H}^1_{\boldsymbol{f}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger) := \mathrm{im}\big(\mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^{\boldsymbol{f}}(\boldsymbol{V}^\dagger)) \to \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}^\dagger)\big)$$

and the $\boldsymbol{h}$-unbalanced local condition $\mathrm{H}^1_{\boldsymbol{h}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger)$ by

$$\mathrm{H}^1_{\boldsymbol{h}}(\mathbb{Q}_p, \boldsymbol{V}^\dagger) := \mathrm{im}\big(\mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^{\boldsymbol{h}}(\boldsymbol{V}^\dagger)) \to \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}^\dagger)\big).$$

Note that the maps appearing in these definitions are injective, so we can identify $\mathrm{H}^1_\star(\mathbb{Q}_p, \boldsymbol{V}^\dagger)$ with $\mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^\star(\boldsymbol{V}^\dagger))$ for $\star \in \{\mathrm{bal}, \boldsymbol{f}, \boldsymbol{h}\}$.

Definition 5.4.2. Let $\star \in \{\mathrm{bal}, \boldsymbol{f}, \boldsymbol{h}\}$. Define the Selmer group $\mathrm{Sel}^\star(\mathbb{Q}, \boldsymbol{V}^\dagger)$ by

$$\mathrm{Sel}^\star(\mathbb{Q}, \boldsymbol{V}^\dagger) := \ker\left\{\mathrm{H}^1(\mathbb{Q}, \boldsymbol{V}^\dagger) \to \frac{\mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}^\dagger)}{\mathrm{H}^1_\star(\mathbb{Q}_p, \boldsymbol{V}^\dagger)} \times \prod_{v \neq p} \mathrm{H}^1(\mathbb{Q}_v^{\mathrm{nr}}, \boldsymbol{V}^\dagger)\right\}.$$

We call $\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^\dagger)$ the balanced Selmer group, $\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{V}^\dagger)$ the $\boldsymbol{f}$-unbalanced Selmer group, and $\mathrm{Sel}^{\boldsymbol{h}}(\mathbb{Q}, \boldsymbol{V}^\dagger)$ the $\boldsymbol{h}$-unbalanced Selmer group.

Definition 5.4.3. Let $\boldsymbol{A}^\dagger = \mathrm{Hom}_{\mathbb{Z}_p}(\boldsymbol{V}^\dagger, \mu_{p^\infty})$ and let $\star \in \{\mathrm{bal}, \boldsymbol{f}, \boldsymbol{h}\}$. Define $\mathrm{H}^1_\star(\mathbb{Q}_p, \boldsymbol{A}^\dagger) \subset \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{A}^\dagger)$ to be the orthogonal complement of $\mathrm{H}^1_\star(\mathbb{Q}_p, \boldsymbol{V}^\dagger)$ under the local Tate pairing

$$\mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}^\dagger) \times \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{A}^\dagger) \to \mathbb{Q}_p/\mathbb{Z}_p.$$

Similarly as above, we then define the balanced, the $\boldsymbol{f}$-unbalanced and the $\boldsymbol{h}$-unbalanced Selmer groups with coefficients in $\boldsymbol{A}^\dagger$ by

$$\mathrm{Sel}^\star(\mathbb{Q}, \boldsymbol{A}^\dagger) := \ker\left\{\mathrm{H}^1(\mathbb{Q}, \boldsymbol{A}^\dagger) \to \frac{\mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{A}^\dagger)}{\mathrm{H}^1_\star(\mathbb{Q}_p, \boldsymbol{A}^\dagger)} \times \prod_{v \neq p} \mathrm{H}^1(\mathbb{Q}_v^{\mathrm{nr}}, \boldsymbol{A}^\dagger)\right\}.$$

Let

$$X^\star(\mathbb{Q}, \boldsymbol{A}^\dagger) = \mathrm{Hom}_{\mathbb{Z}_p}(\mathrm{Sel}^\star(\mathbb{Q}, \boldsymbol{A}^\dagger), \mathbb{Q}_p/\mathbb{Z}_p)$$

be the Pontryagin dual of $\mathrm{Sel}^\star(\mathbb{Q}, \boldsymbol{A}^\dagger)$.

6. Arithmetic applications

Finally, we obtain some arithmetic results from our constructed Euler system through identifying our classes as an anticyclotomic Euler system in the sense of [Jetchev et al. 2024] and the explicit reciprocity law.

6.1. ***Reciprocity law and Greenberg–Iwasawa main conjectures.*** Let $(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ be a triple of primitive Hida families as in Section 5.1 satisfying (5-1). Let $N = \mathrm{lcm}(N_f, N_g, N_h)$. The big diagonal class constructed in [Bertolini et al. 2022b, Section 8.1],

$$\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}) \in \mathrm{H}^1(\mathbb{Q}, \boldsymbol{V}^\dagger(N)), \tag{6-1}$$

where $\boldsymbol{V}^\dagger(N)$ (this is $V(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ using the notation of [loc. cit.]) is a free $\mathcal{R}$-module isomorphic to finitely many copies of $\boldsymbol{V}^\dagger$, can be identified with classes $\tilde{\boldsymbol{\kappa}}_m^{(1)}$, $\boldsymbol{\kappa}_m^{(2)}$ in equations (2.14) and (2.15), respectively, of [Castella and Do 2025]. The definition of the Selmer groups in Section 5.4 extends to $\boldsymbol{V}^\dagger(N)$, and, by [Bertolini et al. 2022b, Corollary 8.2], we have $\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}) \in \mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^\dagger(N))$. Now we choose level-$N$ test vectors $(\check{\boldsymbol{f}}, \check{\boldsymbol{g}}, \check{\boldsymbol{h}})$, provided by [Hsieh 2021, Theorem A], to project the classes from $\mathbb{V}^\dagger(N)$ to $\mathbb{V}^\dagger$.

We define more $G_{\mathbb{Q}_p}$-invariant subspaces of $\boldsymbol{V}^\dagger$,

$$\begin{aligned}
\mathscr{F}_p^3(\boldsymbol{V}^\dagger) &= V_f^+ \hat{\otimes}_{\mathcal{O}} V_g^+ \hat{\otimes}_{\mathcal{O}} V_h^+ \otimes \mathcal{X}^{-1}, \\
\boldsymbol{V}_f^{gh} &= V_f^- \hat{\otimes}_{\mathcal{O}} V_g^+ \hat{\otimes}_{\mathcal{O}} V_h^+ \otimes \mathcal{X}^{-1}, \\
\boldsymbol{V}_g^{fh} &= V_f^+ \hat{\otimes}_{\mathcal{O}} V_g^- \hat{\otimes}_{\mathcal{O}} V_h^+ \otimes \mathcal{X}^{-1}, \\
\boldsymbol{V}_h^{fg} &= V_f^+ \hat{\otimes}_{\mathcal{O}} V_g^+ \hat{\otimes}_{\mathcal{O}} V_h^- \otimes \mathcal{X}^{-1},
\end{aligned} \tag{6-2}$$

and obtain

$$\mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}^\dagger)/\mathscr{F}_p^3(\boldsymbol{V}^\dagger) \cong \boldsymbol{V}_f^{gh} \oplus \boldsymbol{V}_g^{fh} \oplus \boldsymbol{V}_h^{fg}. \tag{6-3}$$

Assume that the congruence ideal $C(\boldsymbol{f}) \subset \mathbb{I}_f$ is principal, generated by the nonzero $\eta_{\boldsymbol{f}} \in \mathbb{I}_f$ (this will be satisfied when the residual representation $\bar{\rho}_{\boldsymbol{f}}$ is absolutely irreducible and p-distinguished). One can deduce from [Kings et al. 2017] the construction of an injective three-variable p-adic regulator map with pseudonull cokernel,

$$\mathrm{Log}^{\eta_f} : \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}_f^{gh}) \to \mathcal{R}; \tag{6-4}$$

see the explicit map in [Castella and Do 2025, Section 4.3.1] and the explanation in [Bertolini et al. 2022b, Section 7.3].

Let $\mathrm{res}_p(\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}))_f$ be the image of $\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ under the natural composition of maps

$$\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^\dagger) \xrightarrow{\mathrm{res}_p} \mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}^\dagger)) \to \mathrm{H}^1(\mathbb{Q}_p, \mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}^\dagger)/\mathscr{F}_p^3(\boldsymbol{V}^\dagger)) \to \mathrm{H}^1(\mathbb{Q}_p, \boldsymbol{V}_f^{gh}), \tag{6-5}$$

where we first restrict at p and then project onto the first direct summand in (6-3). The following result is an explicit reciprocity law that relates diagonal cycles with the triple product p-adic L-functions.

Theorem 6.1.1 [Bertolini et al. 2022b, Theorem A]. *Let* $(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ *be a triple of primitive Hida families as in Theorem 5.3.1. Then*

$$\mathrm{Log}^{\eta_f}\left(\mathrm{res}_p(\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}))_f\right) = \mathscr{L}_p^{f,\eta_f}(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}).$$

Assume that the associated ring $\mathcal{R}$ is regular. Similar to [Alonso et al. 2025, Section 7.3], the following result can be seen as the equivalence between two different formulations of the Iwasawa main conjecture in the style of [Greenberg 1994] for the p-adic deformation $\boldsymbol{V}^\dagger$, relating the $\boldsymbol{f}$-unbalanced Selmer group to the balanced one (or one with $\mathscr{L}_p^{\boldsymbol{f}}$ and another "without p-adic L-functions").

Proposition 6.1.2 [Castella and Do 2025, Proposition 4.3.3]. *The following are equivalent*:

(I) $\mathscr{L}_p^{\boldsymbol{f},\eta_f}(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ *is nonzero, the modules* $\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{V}^\dagger)$ *and* $X^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{A}^\dagger)$ *are both* $\mathcal{R}$*-torsion, and*

$$\mathrm{char}_{\mathcal{R}}(X^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{A}^\dagger)) = (\mathscr{L}_p^{\boldsymbol{f}}(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})^2)$$

in $\mathcal{R} \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$.

(II) $\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})$ *is not* $\mathcal{R}$*-torsion, the modules* $\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^\dagger)$ *and* $X^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{A}^\dagger)$ *both have* $\mathcal{R}$*-rank one, and*

$$\mathrm{char}_{\mathcal{R}}(X^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{A}^\dagger)_{\mathrm{tors}}) = \mathrm{char}_{\mathcal{R}}\left(\frac{\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^\dagger)}{\mathcal{R} \cdot \kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})}\right)^2$$

in $\mathcal{R} \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$, *where the subscript* tors *denotes the* $\mathcal{R}$*-torsion submodule.*

6.2. ***The setup.*** Let $f \in S_{2r}(pN_f)$ be a p-stabilized newform, and suppose the residual representation $\bar{\rho}_f$ is absolutely irreducible and p-distinguished. By Hida theory, f is the specialization of a unique primitive Hida family $\boldsymbol{f} \in S^o(N_f, \mathbb{I})$ at an arithmetic point $Q_1 \in \mathfrak{X}_{\mathbb{I}}^+$ of weight $2r$. For $i \in \{1, 2\}$, let $\mathfrak{f}_i \subset \mathcal{O}_{K_i}$ be an ideal coprime to pN_f, and ξ_i be ray class characters of K_i of conductors dividing $\mathfrak{f}_i$. Let χ_{ξ_i} be the central character of ξ_i. We assume that

$$\chi_{\xi_1}\epsilon_{K_1}\chi_{\xi_2}\epsilon_{K_2} = 1, \tag{6-6}$$

and let

$$\boldsymbol{g}_1 = \boldsymbol{\theta}_{\xi_1}(S_1) \in \mathcal{O}[\![S_1]\!][\![q]\!], \quad \boldsymbol{g}_2 = \boldsymbol{\theta}_{\xi_2}(S_2) \in \mathcal{O}[\![S_2]\!][\![q]\!] \tag{6-7}$$

be the CM Hida families attached to ξ_1 and ξ_2, respectively.

The triple $(\boldsymbol{f}, \boldsymbol{g}_1, \boldsymbol{g}_2)$ satisfies condition (5-1) and the associated $\boldsymbol{f}$-unbalanced triple product p-adic L-function $\mathscr{L}_p^{\boldsymbol{f},\eta_f}(\boldsymbol{f}, \boldsymbol{g}_1, \boldsymbol{g}_2)$ is an element in $\mathcal{R} = \mathbb{I} \hat{\otimes}_{\mathcal{O}} \mathcal{O}[\![S_1]\!] \hat{\otimes}_{\mathcal{O}} \mathcal{O}[\![S_2]\!] \simeq \mathbb{I}[\![S_1, S_2]\!]$. Let

$$\mathscr{L}_p^{f,\eta_f}(f, \boldsymbol{g}_1, \boldsymbol{g}_2) \in \mathcal{O}[\![S_1, S_2]\!] \tag{6-8}$$

be its image under the natural map $\mathbb{I}[\![S_1, S_2]\!] \to \mathcal{O}[\![S_1, S_2]\!]$ defined by Q_1.

Write $\boldsymbol{V}_{Q_1}^\dagger$ for the specialization of $\boldsymbol{V}^\dagger$ at Q_1. Let $V_f^\vee$ be the Galois representation associated to f, and recall that $\det(V_f^\vee) = \varepsilon_{\mathrm{cyc}}^{2r-1}$ in our conventions. Setting $T_i = \boldsymbol{v}^{-1}(1 + S_i) - 1$ for $i \in \{1, 2\}$, we have

$\det(V_{\boldsymbol{g}_{T_1}} \otimes V_{\boldsymbol{h}_{T_2}}) = \Psi_{T_1}\Psi_{T_2} \circ \mathscr{V}$, and so

$$\begin{aligned} \boldsymbol{V}_{Q_1}^{\dagger} &\simeq T_f^{\vee} \otimes (\mathrm{Ind}_K^{\mathbb{Q}}\, \xi_1^{-1}\Psi_{T_1}) \otimes (\mathrm{Ind}_K^{\mathbb{Q}}\, \xi_2^{-1}\Psi_{T_2}) \otimes \varepsilon_{\mathrm{cyc}}^{1-r}(\Psi_{T_1}^{-1/2}\Psi_{T_2}^{-1/2} \circ \mathscr{V}) \\ &\simeq T_f^{\vee}(1-r) \otimes \mathrm{Ind}_{K_0}^{\mathbb{Q}}\, \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}, \end{aligned} \tag{6-9}$$

where $T_f^{\vee}$ is a $G_{\mathbb{Q}}$-stable $\mathcal{O}$-lattice inside $V_f^{\vee}$. In particular, we get

$$\mathrm{H}^1(\mathbb{Q}, \boldsymbol{V}_{Q_1}^{\dagger}) \simeq \mathrm{H}^1(K_0, T_f^{\vee}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \tag{6-10}$$

by Shapiro's lemma.

6.3. ***Local conditions at p of the Euler system.*** Recall from Theorem 3.3.1 that we have constructed classes

$$\boldsymbol{z}_{f,\chi,\mu_3} \in H^1_{\mathrm{Iw}}(K_0[\mu_3 p^{\infty}], T_{f,\chi}),$$

where $T_{f,\chi} = T_f^{\vee}(1-r) \otimes \chi^{-1}$ and $\chi^{-1} = \tilde{\psi}_1^{-1}\tilde{\psi}_2^{-1}\mathbf{N}^{1-(k_1+k_2)/2}$.

Proposition 6.3.1. *Suppose $p \nmid 6h_{K_0}$ and f is non-Eisenstein modulo $\mathfrak{P}$. Let $\mu_3 \in \mathcal{N}$ (taken from Section 3.1) run over squarefree product of prime ideals of $\lambda_3 \in \mathcal{L}$ with $m = N_{K_3/\mathbb{Q}}(\mu_3)$ coprime to p. The class $\boldsymbol{z}_{f,\chi,\mu_3}$ of Theorem 3.3.1 satisfies*

$$\boldsymbol{z}_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3 p^{\infty}], T_{f,\chi}).$$

Proof. By [Bertolini et al. 2022b, Corollary 8.2; Castella and Do 2025, Section 3.2], the class $\boldsymbol{z}_{f,\chi,\mu_3}$ lands in the balanced Selmer group $\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}^{\dagger})$, where the balanced local condition at p when specialized to f is given by

$$\begin{aligned} \mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}_{Q_1}^{\dagger}) \simeq (T_f^{\vee}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \oplus (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-\boldsymbol{c}}\tilde{\Psi}_{T_2}^{\boldsymbol{c}}) \\ \oplus (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-\boldsymbol{c}}\tilde{\Psi}_{T_1}^{\boldsymbol{c}}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}). \end{aligned} \tag{6-11}$$

Put $\widetilde{\boldsymbol{V}}_{Q_1}^{\dagger} = T_f^{\vee}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}$; then Shapiro's lemma tells us that $H^1(\mathbb{Q}, \boldsymbol{V}_{Q_1}^{\dagger}) \simeq H^1(K_0, \widetilde{\boldsymbol{V}}_{Q_1}^{\dagger})$; see (6-10). Following [Castella and Do 2025, Section 5.3], the local condition $\mathscr{F}_p^{\mathrm{bal}}(\boldsymbol{V}_{Q_1}^{\dagger})$ cutting out the specialized balanced Selmer group at p corresponds to

$$\begin{aligned} \mathscr{F}_{\mathcal{P}_1}^{\mathrm{bal}}(\widetilde{\boldsymbol{V}}_{Q_1}^{\dagger}|_{G_{K_0}}) &= T_f^{\vee}\big(1-\tfrac{1}{2}k\big) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}, \\ \mathscr{F}_{\mathcal{P}_2}^{\mathrm{bal}}(\widetilde{\boldsymbol{V}}_{Q_1}^{\dagger}|_{G_{K_0}}) &= 0, \\ \mathscr{F}_{\mathcal{P}_3}^{\mathrm{bal}}(\widetilde{\boldsymbol{V}}_{Q_1}^{\dagger}|_{G_{K_0}}) &= T_f^{\vee,+}\big(1-\tfrac{1}{2}k\big) \otimes \tilde{\xi}_1^{-\boldsymbol{c}}\tilde{\Psi}_{T_1}^{\boldsymbol{c}}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}, \\ \mathscr{F}_{\mathcal{P}_4}^{\mathrm{bal}}(\widetilde{\boldsymbol{V}}_{Q_1}^{\dagger}|_{G_{K_0}}) &= T_f^{\vee,+}\big(1-\tfrac{1}{2}k\big) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-\boldsymbol{c}}\tilde{\Psi}_{T_2}^{\boldsymbol{c}}. \end{aligned} \tag{6-12}$$

Hence, the class $\boldsymbol{z}_{f,\chi,\mu_3}$ satisfies the relaxed-strict-ordinary-ordinary condition at the primes above p.

On the other hand, at the primes $w \nmid p$, because $V_{f,\chi}$ is conjugate self-dual and pure of weight -1, we see that

$$H^0(K_0[\mu_3\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s]_w, V_{f,\chi}) = H^2(K_0[\mu_3\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s]_w, V_{f,\chi}) = 0$$

for all r, s, and therefore

$$H^1(K_0[\mu_3\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s]_w, V_{f,\chi}) = 0$$

by Tate's local Euler characteristic formula. This implies the torsionness of $H^1(K_0[\mu_3\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s]_w, T_{f,\chi})$, and one has the inclusion

$$\mathrm{res}_w(z_{f,\chi,\mu_3}) \in \varprojlim_{r,s} H^1_f(K_0[\mu_3\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s]_w, T_{f,\chi}). \qquad \square$$

Proposition 6.3.2. *Via the isomorphism (6-10):*

(1) *The balanced Selmer group* $\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger)$ *can be rewritten as*

$$\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}).$$

(2) *The* $\boldsymbol{f}$*-unbalanced Selmer group* $\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger)$ *can be rewritten as*

$$\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(K_0, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}).$$

(3) *The* $\boldsymbol{h}$*-unbalanced Selmer group* $\mathrm{Sel}^{\boldsymbol{h}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger)$ *can be rewritten as*

$$\mathrm{Sel}^{\boldsymbol{h}}(\mathbb{Q}, \boldsymbol{V}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{rel,str,rel,str}}(K_0, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}).$$

Proof. For the balanced case, see Proposition 6.3.1. For the $\boldsymbol{f}$-unbalanced case, note that

$$\begin{aligned}\mathscr{F}_p^{\boldsymbol{f}}(\boldsymbol{V}_{Q_1}^\dagger) \simeq (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \oplus (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-\boldsymbol{c}}\tilde{\Psi}_{T_2}^{\boldsymbol{c}}) \\ \oplus (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-\boldsymbol{c}}\tilde{\Psi}_{T_1}^{\boldsymbol{c}}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \oplus (T_f^{\vee,+}(1-r) \otimes \tilde{\xi}_1^{-\boldsymbol{c}}\tilde{\Psi}_{T_1}^{\boldsymbol{c}}\tilde{\xi}_2^{-\boldsymbol{c}}\tilde{\Psi}_{T_2}^{\boldsymbol{c}}), \quad (6\text{-}13)\end{aligned}$$

and the result follows.

The $\boldsymbol{h}$-unbalanced case can be obtained in a similar manner, where

$$\mathscr{F}_p^{\boldsymbol{h}}(\boldsymbol{V}_{Q_1}^\dagger) \simeq (T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \oplus (T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-\boldsymbol{c}}\tilde{\Psi}_{T_1}^{\boldsymbol{c}}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}). \qquad \square$$

As a consequence, we also obtain the following isomorphisms for the Selmer groups with coefficients in $\boldsymbol{A}_{Q_1}^\dagger = \mathrm{Hom}_{\mathbb{Z}_p}(\boldsymbol{V}_{Q_1}^\dagger, \mu_{p^\infty})$ by local Tate duality. Let $A_f(r) = \mathrm{Hom}_{\mathbb{Z}_p}(T_f^\vee(1-r), \mu_{p^\infty})$.

Corollary 6.3.3. *We can identify the balanced Selmer group* $\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger)$ *as*

$$\mathrm{Sel}^{\mathrm{bal}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{str,rel,ord,ord}}(K_0, A_f(r) \otimes \tilde{\xi}_1\tilde{\Psi}_{T_1}^{-1}\tilde{\xi}_2\tilde{\Psi}_{T_2}^{-1}),$$

the $\boldsymbol{f}$*-unbalanced Selmer group* $\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger)$ *as*

$$\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(K_0, A_f(r) \otimes \tilde{\xi}_1\tilde{\Psi}_{T_1}^{-1}\tilde{\xi}_2\tilde{\Psi}_{T_2}^{-1}),$$

and the $\boldsymbol{h}$*-unbalanced Selmer group* $\mathrm{Sel}^{\boldsymbol{h}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger)$ *as*

$$\mathrm{Sel}^{\boldsymbol{h}}(\mathbb{Q}, \boldsymbol{A}_{Q_1}^\dagger) \simeq \mathrm{Sel}_{\mathrm{str,rel,str,rel}}(K_0, A_f(r) \otimes \tilde{\xi}_1\tilde{\Psi}_{T_1}^{-1}\tilde{\xi}_2\tilde{\Psi}_{T_2}^{-1}).$$

6.4. ***Applying the general machinery.*** We show some arithmetic applications by invoking the general Euler system machinery of [Jetchev et al. 2024]; see some details for the imaginary quadratic case in [Do 2022, Section 4.3; Alonso et al. 2025, Section 8]. These results will be used to deduce the Bloch–Kato conjecture and the anticyclotomic Iwasawa main conjecture by exploiting the relation between our Euler system classes and special values of complex and p-adic L-functions via an explicit reciprocity law.

For every ideal $\mu_3 \in \mathcal{N}$, denote by

$$z_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3], T_{f,\chi})$$

the image of z_{f,χ,μ_3} from Theorem 3.3.1 under the projection

$$\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3 p^\infty], T_{f,\chi}) \to \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3], T_{f,\chi}). \tag{6-14}$$

Define the base class

$$z_{f,\chi} := \mathrm{Norm}_{K_0}^{K_0[1]}(z_{f,\chi,1}) \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, T_{f,\chi}).$$

Note that since we assume $p \nmid h_{K_0}$, $K_0[1]$ is actually the same with K_0 (recall that $K_0[\mathfrak{n}]$ is the maximal p-extension inside the ring class field of K_0 of conductor $\mathfrak{n}$). Therefore, $z_{f,\chi} = z_{f,\chi,1}$.

Theorem 6.4.1. *Assume that f is not of CM type, non-Eisenstein at $\mathfrak{P}$, and that $p \nmid 6h_{K_0}$. Then*

$$z_{f,\chi} \text{ is nontorsion} \implies \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, V_{f,\chi}) \text{ is one-dimensional.}$$

Proof. Combining Theorem 3.3.1 and Proposition 6.3.1, the system of classes

$$\{z_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3], T_{f,\chi}) : \mu_3 \in \mathcal{N}\} \tag{6-15}$$

forms an anticyclotomic Euler system in the sense of [Jetchev et al. 2024] for the relaxed-strict-ordinary-ordinary Greenberg Selmer group.

Under the assumption that f is not of CM type, the following properties follow from Momose's big image results [1981] as in [Lei et al. 2015, Proposition 7.1.4]:

(i) $V_{f,\chi}$ is absolutely irreducible.

(ii) There is an element $\sigma \in G_{K_0}$ fixing $K_0[1]K_0(\mu_{p^\infty}, (\mathcal{O}_{K_0}^\times)^{1/p^\infty})$ such that $V_{f,\chi}/(\sigma - 1)V_{f,\chi}$ is one-dimensional.

(iii) There is an element $\gamma \in G_{K_0}$ fixing $K_0[1]K_0(\mu_{p^\infty}, (\mathcal{O}_{K_0}^\times)^{1/p^\infty})$ such that $V_{f,\chi}^{\gamma=1} = 0$.

Hence, the fact that $z_{f,\chi}$ is nontorsion implies the one-dimensionality of $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, V_{f,\chi})$ by the general machinery of [Jetchev et al. 2024]. □

Recall that $K_{0,\infty}^-$ is the anticyclotomic $\mathbb{Z}_p^2$-extension over K_0 and $\Lambda_{K_0}^- = \mathbb{Z}_p[\![\mathrm{Gal}(K_{0,\infty}^-/K_0)]\!]$. Let $z_{f,\chi,1}$ be the $\Lambda_{K_0}^-$-adic class of Theorem 3.3.1 of conductor $\mu_3 = (1)$, and define the Iwasawa-theoretic base class

$$z_{f,\chi} := \mathrm{Norm}_{K_0}^{K_0[1]}(z_{f,\chi,(1)}).$$

Again $z_{f,\chi} = z_{f,\chi,(1)}$ from the assumption on the class number of K_0. Note that, by Proposition 6.3.1, one has

$$z_{f,\chi} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-, T_{f,\chi}).$$

Definition 6.4.2. We say that f has big image at $\mathfrak{P}$ if the image of $G_{\mathbb{Q}}$ in $\mathrm{Aut}_{\mathcal{O}}(T_f^\vee)$ contains a conjugate of $\mathrm{SL}_2(\mathbb{Z}_p)$.

Remark 6.4.3. By a theorem of Ribet [1985], if f is not of CM type then it has big image for all but finitely many primes of L.

Let

$$X_{\mathrm{str,rel,ord,ord}}(K_{0,\infty}^-, A_{f,\chi}) = \mathrm{Hom}_{\mathbb{Z}_p}\big(\varinjlim \mathrm{Sel}_{\mathrm{str,rel,ord,ord}}(K_0[\mathfrak{p}_3^r\tilde{\mathfrak{p}}_3^s], A_{f,\chi}), \mathbb{Q}_p/\mathbb{Z}_p\big).$$

One then has a divisibility towards an anticyclotomic Iwasawa main conjecture "without p-adic L-functions" as follows:

Theorem 6.4.4. *Assume that f is not of CM type, has big image at $\mathfrak{P}$, and that $p \nmid 6h_{K_0}$. If $z_{f,\chi}$ is nontorsion, then*:

(1) $X_{\mathrm{str,rel,ord,ord}}(K_{0,\infty}^-, A_{f,\chi})$ *and* $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-, T_{f,\chi})$ *both have $\Lambda_{K_0}^-$-rank one.*

(2) *We have the divisibility*

$$\mathrm{char}_{\Lambda_{K_0}^-}(X_{\mathrm{str,rel,ord,ord}}(K_{0,\infty}^-, A_{f,\chi})_{\mathrm{tors}}) \supset \mathrm{char}_{\Lambda_{K_0}^-}\left(\frac{\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-, T_{f,\chi})}{\Lambda_{K_0}^- \cdot z_{f,\chi}}\right)^2$$

in $\Lambda_{K_0}^-$.

Here, the subscript tors *denotes the $\Lambda_{K_0}^-$-torsion submodule.*

Proof. Combining Theorem 3.3.1 and Proposition 6.3.1, the system of classes

$$\{z_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[\mu_3 p^\infty], T_{f,\chi}) : \mu_3 \in \mathcal{N}\} \tag{6-16}$$

forms a $\Lambda_{K_0}^-$-adic anticyclotomic Euler system in the sense of Jetchev–Nekovář–Skinner for the relaxed-strict-ordinary-ordinary Selmer group.

Under the assumption that f has big image at $\mathfrak{P}$, we have (see [Lei et al. 2015, Proposition 7.1.6]):

(i) $\overline{T}_{f,\chi} := T_{f,\chi}/\mathfrak{P}T_{f,\chi}$ is absolutely irreducible.

(ii) There is an element $\sigma \in G_K$ fixing $K_0[1]K_0(\mu_{p^\infty}, (\mathcal{O}_K^\times)^{1/p^\infty})$ such that $T_{f,\chi}/(\sigma - 1)T_{f,\chi}$ is free of rank 1 over $\mathcal{O}$.

(iii) There is an element $\gamma \in G_{K_0}$ fixing $K_0[1]K_0(\mu_{p^\infty}, (\mathcal{O}_K^\times)^{1/p^\infty})$ and acting as multiplication by a scalar $a_\gamma \neq 1$ on $\overline{T}_{f,\chi}$.

Thus, the nontorsionness of $z_{f,\chi}$ implies the conclusions by the general machinery of [Jetchev et al. 2024]. □

6.5. ***On the Bloch–Kato conjecture in rank*** **0.** Our first application is the Bloch–Kato conjecture in analytic rank zero for the conjugate self-dual G_{K_0}-representation $V_{f,\chi} = V_f^\vee(1-r)\otimes\chi^{-1}$.

Assumption 6.1. We assume that the anticyclotomic Hecke character χ over K_0 can be decomposed as

$$\chi = \tilde{\psi}_1\tilde{\psi}_2\mathbf{N}^{(k_1+k_2-2)/2},$$

where:

(1) ψ_1 is a Hecke character of K_1 of infinity type $(1-k_1, 0)$, with $k_1 \geq 1$, and modulus $\mathfrak{f}_1$.

(2) ψ_2 is a Hecke character of K_2 of infinity type $(1-k_2, 0)$, with $k_2 \geq 1$ and modulus $\mathfrak{f}_2$.

(3) $\tilde{\psi}_i$ is the Hecke character of K_0, obtained by composing $\mathbb{A}_{K_0}^\times \xrightarrow{\mathbb{N}_{K_0/K_i}} \mathbb{A}_{K_i}^\times \xrightarrow{\psi_i} \mathbb{C}$ for each $i \in \{1, 2\}$.

(4) By swapping K_1 and K_2, we may assume that $k_2 \geq k_1$.

In this scenario, the infinity type of χ (corresponding to the order $(\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4)$ or $(1, \tau_3, \tau_2, \tau_1)$) is

$$\big(\tfrac{1}{2}(2-k_1-k_2), \tfrac{1}{2}(k_1+k_2-2), \tfrac{1}{2}(k_1-k_2), \tfrac{1}{2}(k_2-k_1)\big).$$

Theorem 6.5.1. *Let $f \in S_k(\Gamma_0(pN_f))$ be a p-ordinary newform of weight $k = 2r \geq 2$. Let χ be an anticyclotomic Hecke character of K_0 as in Assumption 6.1. Assume that*:

(1) *Either $k \geq k_1 + k_2$ or $k_2 - k_1 \geq k$.*

(2) *$N_f\mathcal{O}_{K_3} = \mathfrak{n}^+\mathfrak{n}^-$, where $\mathfrak{n}^+$ (resp. $\mathfrak{n}^-$) is divisible only by primes which are split (resp. inert) in K_0/K_3 and $\mathfrak{n}^-$ is a squarefree product of an even number of primes.*

(3) *$\bar{\rho}_f$ is absolutely irreducible.*

(4) *$(pN_f, \mathrm{Norm}_{K_1/\mathbb{Q}}(\mathfrak{f}_1)\mathrm{Norm}_{K_2/\mathbb{Q}}(\mathfrak{f}_2)D_{K_0}) = 1$.*

(5) *$p \nmid 6h_{K_0}$, the class number of K_0.*

Then

$$L(f/K_0, \chi, r) \neq 0 \implies \mathrm{Sel}_{\mathrm{BK}}(K_0, V_{f,\chi}) = 0.$$

In other words, the Bloch–Kato conjecture holds in analytic rank zero for $V_{f,\chi}$.

Proof. We consider the CM Hida families

$$\boldsymbol{g} = \boldsymbol{\theta}_{\xi_1}(S_1), \quad \boldsymbol{h} = \boldsymbol{\theta}_{\xi_2}(S_2)$$

that pass through θ_{ψ_1} and θ_{ψ_2}, respectively. Note that the triple $(\boldsymbol{f}, \boldsymbol{g}_1, \boldsymbol{g}_2)$ satisfies (5-3). Then the isomorphism (6-9) of the associated $\boldsymbol{V}_{\mathcal{Q}_1}^\dagger$ together with the specialization $\mathcal{Q}_1$ corresponding to θ_{ψ_1} and θ_{ψ_2} show that

$$L(\mathbb{V}_{\mathcal{Q}_1}^\dagger, 0) = L(f/K_0, \chi, r).$$

By Theorem 6.1.1, we then have

$$L(f/K_0, \chi, r) \neq 0 \implies \mathrm{res}_p(\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}))_{\boldsymbol{f},\mathcal{Q}_1} \neq 0.$$

From our construction and Proposition 6.3.2, the class $\kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})_{\mathcal{Q}_1} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, V_{f,\chi})$ is the base class of the anticyclotomic Euler system

$$\{z_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0[m], T_{f,\chi}) : \mu_3 \in \mathcal{N}\}$$

of (6-15). Recall again that $K_0[1] = K_0$ from the assumption $p \nmid h_{K_0}$. By Theorem 6.4.1, we conclude that the Selmer group $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, V_{f,\chi})$ is one-dimensional, spanned by

$$z_{f,\chi} = \mathrm{Norm}_{K_0}^{K_0[1]}(z_{f,\chi,1}) = \kappa(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h})_{\mathcal{Q}_1}.$$

If $k \geq k_1 + k_2$, we observe that the composition of maps in (6-5) corresponds to the composition

$$\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2}) \xrightarrow{\mathrm{res}_{\mathcal{P}_1}} \mathrm{H}^1(K_{0,\mathcal{P}_1}, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2})$$
$$\to \mathrm{H}^1(K_{0,\mathcal{P}_1}, T_f^{\vee,-}(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2})$$

by (6-10). Hence, $\mathrm{res}_{\mathcal{P}_1}(z_{f,\chi}) \neq 0$ by the reciprocity law of Theorem 6.1.1. Then the vanishing of $\mathrm{Sel}_{\mathrm{ord,ord,ord,ord}}(K_0, V_{f,\chi})$ follows by a standard argument using Poitou–Tate duality (see [Do 2022, Section 5.1.1]). This yields the result by using Lemma 4.4 for $k \geq k_1 + k_2$ to identify the latter group with $\mathrm{Sel}_{\mathrm{BK}}(K_0, V_{f,\chi})$.

If $k_2 - k_1 \geq k$, similarly, by using $\mathrm{res}_{\mathcal{P}_2}(z_{f,\chi}) \neq 0$, we obtain the vanishing of $\mathrm{Sel}_{\mathrm{rel,str,rel,str}}(K_0, V_{f,\chi})$, which is again the Bloch–Kato Selmer group $\mathrm{Sel}_{\mathrm{BK}}(K_0, V_{f,\chi})$ by Lemma 4.4 for $k_2 - k_1 \geq k$. □

Remark 6.5.2. Define $\epsilon(f, \chi)$ to be the sign of the functional equation for $V_{f,\chi}$. Then $\epsilon(f, \chi) = \prod \epsilon\left(\pi_{K_{0,v}} \otimes \chi_v, \frac{1}{2}\right)$ over places v of K_0 as a product of local root numbers. If $v \mid \mathfrak{n}^+$, then $\epsilon\left(\pi_{K_{0,v}} \otimes \chi_v, \frac{1}{2}\right) = +1$ and, if $v \mid \mathfrak{n}^-$, then $\epsilon\left(\pi_{K_{0,v}} \otimes \chi_v, \frac{1}{2}\right) = -1$. Therefore, the contribution from the local places is $+1$ due to assumption (2). At the infinity places,

$$\epsilon_\infty\left(\pi_{K_0} \otimes \chi, \tfrac{1}{2}\right) = i^{|k-1+(k_1+k_2-2)|+|k-1-(k_1+k_2-2)|+|k-1+(k_2-k_1)|+|k-1-(k_2-k_1)|}$$
$$= \begin{cases} +1 & \text{if } k > (k_1+k_2-2), \\ -1 & \text{if } k_2 - k_1 < k \leq k_1 + k_2 - 2, \\ +1 & \text{if } k \leq k_2 - k_1. \end{cases}$$

Hence, conditions (1) and (2) of Theorem 6.5.1 then imply that $\epsilon(f, \chi) = 1$.

6.6. ***On the Iwasawa main conjecture.*** Our second application is evidence towards the anticyclotomic Iwasawa main conjecture for modular forms. Recall that we have an eigenform f of weight $k = 2r \geq 2$ with trivial nebentypus and an anticyclotomic character χ satisfying Assumption 6.1. Let

$$A_{f,\chi} = \mathrm{Hom}_{\mathbb{Z}_p}(T_f^\vee(1-r) \otimes \chi^{-1}, \mu_{p^\infty}).$$

Theorem 6.6.1. *Under the same assumption as in Theorem 6.5.1, we assume further that*:

(1) $\bar{\rho}_f$ *is p-distinguished.*

(2) f *has big image.*

(3) $p > k - 2$.

If $k \geq k_1 + k_2$, then $\mathrm{Sel}_{\mathrm{ord},\mathrm{ord},\mathrm{ord},\mathrm{ord}}(K_{0,\infty}^-, A_{f,\chi})$ is cotorsion over $\Lambda_{K_0}^-$. Furthermore, inside $\Lambda_{K_0}^- \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$, we have the inclusion

$$\mathrm{char}_{\Lambda_{K_0}^-}(\mathrm{Sel}_{\mathrm{ord},\mathrm{ord},\mathrm{ord},\mathrm{ord}}(K_{0,\infty}^-, A_{f,\chi})^\vee) \supset (\mathscr{L}_p^{\boldsymbol{f},\eta_f}(f, \boldsymbol{g}_1, \boldsymbol{g}_2)^2).$$

Proof. Recall from Corollary 6.3.3 that

$$\mathrm{Sel}^{\boldsymbol{f}}(\mathbb{Q}, \mathbb{A}^\dagger) \simeq \mathrm{Sel}_{\mathrm{ord},\mathrm{ord},\mathrm{ord},\mathrm{ord}}(K_{0,\infty}^-, A_f(r) \otimes \chi), \tag{6-17}$$

where $\mathbb{A}^\dagger = \mathrm{Hom}_{\mathbb{Z}_p}(\mathbb{V}^\dagger, \mu_{p^\infty})$.

Note that, from (6-8), $\mathscr{L}_p^{\boldsymbol{f},\eta_f}(f, \boldsymbol{g}_1, \boldsymbol{g}_2)$ is an element of $\mathcal{O}[\![S_1, S_2]\!]$. We then identify $\Lambda_{K_0}^- \simeq \mathcal{O}[\![S_1, S_2]\!]$ via the diagram (3-9). The p-adic L-function $\mathscr{L}_p^{\boldsymbol{f},\eta_f}(f, \boldsymbol{g}_1, \boldsymbol{g}_2)$ is nonzero by [Hung 2017, Theorem C]. Note that our assumption that $k \geq k_1 + k_2$ ensures that we are in the critical specializations, i.e., $-\frac{1}{2}k_\sigma < m_\sigma < \frac{1}{2}k_\sigma$ for all $\sigma \in \Sigma$, following the notation of [loc. cit.]. Hence, Theorem 6.1.1 and the proof of Theorem 6.5.1 implies that the class

$$\kappa(f, \boldsymbol{g}, \boldsymbol{h}) \in \mathrm{Sel}_{\mathrm{rel},\mathrm{str},\mathrm{ord},\mathrm{ord}}(K_0, T_f^\vee(1-r) \otimes \tilde{\xi}_1^{-1}\tilde{\Psi}_{T_1}\tilde{\xi}_2^{-1}\tilde{\Psi}_{T_2})$$

is nontorsion. By construction, we can treat $\kappa(f, \boldsymbol{g}, \boldsymbol{h})$ as the base class of the $\Lambda_{K_0}^-$-adic anticyclotomic Euler system

$$\{z_{f,\chi,\mu_3} \in \mathrm{Sel}_{\mathrm{rel},\mathrm{str},\mathrm{ord},\mathrm{ord}}(K_0[\mu_3 p^\infty], T_{f,\chi}) : \mu_3 \in \mathcal{N}\}$$

in (6-16). Via the isomorphism (6-10), the result follows immediately from Theorem 6.4.4 applied to

$$z_{f,\chi} := \mathrm{Norm}_{K_0}^{K_0[1]}(z_{f,\chi,(1)}) = \kappa(f, \boldsymbol{g}, \boldsymbol{h}), \tag{6-18}$$

the equivalence between two different formulation of the Iwasawa main conjecture in Proposition 6.1.2, and the Selmer group isomorphism (6-17). □

Remark 6.6.2. Within Theorem 6.6.1, the right-hand side can be compared to the p-adic L-function of [Wan 2015, Theorem 86; Hung 2017] (under the assumptions of [Fujiwara 2006, Theorems 11.1–11.2; Wan 2015, Theorem 103]). The author then expects that the full Iwasawa main conjecture, which means an equality of Theorem 6.6.1, will follow by combining Theorem 6.6.1 with the opposite inclusion of [Wan 2015] and the vanishing of the μ-invariant of the p-adic L-function [Hung 2017] (which are generalizations of [Skinner and Urban 2014; Vatsal 2003]).

Remark 6.6.3. One expects a similar result that, if $k_2 - k_1 \geq k$, then $\mathrm{Sel}_{\mathrm{str},\mathrm{rel},\mathrm{str},\mathrm{rel}}(K_\infty^-, A_{f,\chi})$ is cotorsion over $\Lambda_{K_0}^-$ together with the inclusion, inside $\Lambda_{K_0}^- \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$,

$$\mathrm{char}_{\Lambda_{K_0}^-}(\mathrm{Sel}_{\mathrm{str},\mathrm{rel},\mathrm{str},\mathrm{rel}}(K_{0,\infty}^-, A_{f,\chi})^\vee) \supset (\mathscr{L}_p^{\boldsymbol{h},\eta_h}(f, \boldsymbol{g}_1, \boldsymbol{g}_2)^2).$$

The only missing ingredient is the nonvanishing of the p-adic L-function in this region.

6.7. ***On the Bloch–Kato conjecture in rank* 1.** Our last application is extracted from the proof of Theorem 6.6.1. It provides a result towards the Bloch–Kato conjecture in rank 1.

Theorem 6.7.1. *Under the same assumption as in Theorem 6.6.1, if* $k_1+k_2-2\geq k\geq k_2-k_1+2$ *(which induces* $L(f/K_0,\chi,r)=0$*), then*

$$\dim_{L_{\mathfrak{P}}}\mathrm{Sel}_{\mathrm{BK}}(K_0,V_{f,\chi})\geq 1.$$

Proof. The class $z_{f,\chi}\in\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-,T_{f,\chi})$ is nontorsion via the proof of Theorem 6.6.1. Furthermore, $z_{f,\chi}$ is the base of a $\Lambda_{K_0}^-$-adic anticyclotomic Euler system as in (6-18) for the relaxed-strict-ordinary-ordinary Selmer group. Theorem 6.4.4 then implies that $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-,T_{f,\chi})$ has $\Lambda_{K_0}^-$-rank 1. The natural map (compare with the projection (6-14))

$$\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_{0,\infty}^-,T_{f,\chi})/(\gamma_{1,-}-1,\gamma_{2,-}-1)\to\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0,T_{f,\chi}) \tag{6-19}$$

is injective (see also [Mazur and Rubin 2004, Proposition 5.3.14; Greenberg 1999, page 72]).

Hence, the Selmer group $\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0,T_{f,\chi})$ has a positive $\mathcal{O}$-rank. The theorem then follows by Lemma 4.4, which computes the local conditions of the Bloch–Kato Selmer group explicitly. □

Remark 6.7.2. By letting $z_{f,\chi}\in\mathrm{Sel}_{\mathrm{rel,str,ord,ord}}(K_0,T_{f,\chi})$ be the image of $z_{f,\chi}$ under the map (6-19), such a class $z_{f,\chi}\in\mathrm{Sel}_{\mathrm{BK}}(K_0,V_{f,\chi})$ satisfies

$$z_{f,\chi}\neq 0\implies\dim_{L_{\mathfrak{P}}}\mathrm{Sel}_{\mathrm{BK}}(K_0,V_{f,\chi})=1$$

by Theorem 6.4.1.

Acknowledgements

The author would like to thank Francesc Castella and Christopher Skinner for many fruitful discussions and encouragements, Haruzo Hida for suggesting the problem, Chandrashekhar Khare, Romyar Sharifi and Alex Smith for helpful communication, and the referee for many helpful comments and suggestions on an earlier draft of this paper.

References

[Alonso et al. 2025] R. Alonso, F. Castella, and O. Rivero, "An anticyclotomic Euler system for adjoint modular Galois representations", *Ann. Inst. Fourier (Grenoble)* **75**:1 (2025), 291–329. MR

[Bertolini and Darmon 2005] M. Bertolini and H. Darmon, "Iwasawa's main conjecture for elliptic curves over anticyclotomic $\mathbb{Z}_p$-extensions", *Ann. of Math.* (2) **162**:1 (2005), 1–64. MR

[Bertolini et al. 2013] M. Bertolini, H. Darmon, and K. Prasanna, "Generalized Heegner cycles and p-adic Rankin L-series", *Duke Math. J.* **162**:6 (2013), 1033–1148. MR

[Bertolini et al. 2022a] M. Bertolini, H. Darmon, V. Rotger, M. A. Seveso, and R. Venerucci, *Heegner points, Stark–Heegner points, and diagonal classes*, Astérisque **434**, Soc. Math. France, Paris, 2022. MR

[Bertolini et al. 2022b] M. Bertolini, M. A. Seveso, and R. Venerucci, "Reciprocity laws for balanced diagonal classes", pp. 77–174 in *Heegner points, Stark–Heegner points, and diagonal classes*, Astérisque **434**, Soc. Math. France, Paris, 2022. MR

[Bloch and Kato 1990] S. Bloch and K. Kato, "L-functions and Tamagawa numbers of motives", pp. 333–400 in *The Grothendieck Festschrift, I*, edited by P. Cartier et al., Progr. Math. **86**, Birkhäuser, Boston, MA, 1990. MR

[Castella 2020] F. Castella, "On the p-adic variation of Heegner points", *J. Inst. Math. Jussieu* **19**:6 (2020), 2127–2164. MR

[Castella and Do 2025] F. Castella and K. T. Do, "Diagonal cycles and anticyclotomic Iwasawa theory of modular forms", *J. Eur. Math. Soc.* (online publication November 2025).

[Castella and Hsieh 2018] F. Castella and M.-L. Hsieh, "Heegner cycles and p-adic L-functions", *Math. Ann.* **370**:1-2 (2018), 567–628. MR

[Chida 2017] M. Chida, "Selmer groups and central values of L-functions for modular forms", *Ann. Inst. Fourier* (*Grenoble*) **67**:3 (2017), 1231–1276. MR

[Chida and Hsieh 2015] M. Chida and M.-L. Hsieh, "On the anticyclotomic Iwasawa main conjecture for modular forms", *Compos. Math.* **151**:5 (2015), 863–897. MR

[Darmon and Rotger 2017] H. Darmon and V. Rotger, "Diagonal cycles and Euler systems, II: The Birch and Swinnerton-Dyer conjecture for Hasse–Weil–Artin L-functions", *J. Amer. Math. Soc.* **30**:3 (2017), 601–672. MR

[Do 2022] K. T. Do, *Construction of an anticyclotomic Euler system with applications*, Ph.D. thesis, Princeton University, 2022, available at https://www.proquest.com/docview/2720964424. MR

[Flach 1990] M. Flach, "A generalisation of the Cassels–Tate pairing", *J. Reine Angew. Math.* **412** (1990), 113–127. MR

[Fröhlich and Taylor 1993] A. Fröhlich and M. J. Taylor, *Algebraic number theory*, Cambridge Studies in Advanced Mathematics **27**, Cambridge Univ. Press, 1993. MR

[Fujiwara 2006] K. Fujiwara, "Deformation rings and Hecke algebras in the totally real case", preprint, 2006. arXiv math/0602606

[Greenberg 1994] R. Greenberg, "Iwasawa theory and p-adic deformations of motives", pp. 193–223 in *Motives* (Seattle, WA, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55.2**, Amer. Math. Soc., Providence, RI, 1994. MR

[Greenberg 1999] R. Greenberg, "Iwasawa theory for elliptic curves", pp. 51–144 in *Arithmetic theory of elliptic curves* (Cetraro, 1997), edited by C. Viola, Lecture Notes in Math. **1716**, Springer, 1999. MR

[Gross and Kudla 1992] B. H. Gross and S. S. Kudla, "Heights and the central critical values of triple product L-functions", *Compositio Math.* **81**:2 (1992), 143–209. MR

[Gross and Schoen 1995] B. H. Gross and C. Schoen, "The modified diagonal cycle on the triple product of a pointed curve", *Ann. Inst. Fourier* (*Grenoble*) **45**:3 (1995), 649–679. MR

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

[Hida 1988] H. Hida, "Modules of congruence of Hecke algebras and L-functions associated with cusp forms", *Amer. J. Math.* **110**:2 (1988), 323–382. MR

[Hida 2015] H. Hida, "Big Galois representations and p-adic L-functions", *Compos. Math.* **151**:4 (2015), 603–664. MR

[Hsieh 2021] M.-L. Hsieh, "Hida families and p-adic triple product L-functions", *Amer. J. Math.* **143**:2 (2021), 411–532. MR

[Hung 2017] P.-C. Hung, "On the non-vanishing mod ℓ of central L-values with anticyclotomic twists for Hilbert modular forms", *J. Number Theory* **173** (2017), 170–209. MR

[Jetchev et al. 2024] D. Jetchev, J. Nekovář, and C. Skinner, "Split Euler systems for conjugate-dual Galois representations", preprint, 2024.

[Kings et al. 2017] G. Kings, D. Loeffler, and S. L. Zerbes, "Rankin–Eisenstein classes and explicit reciprocity laws", *Camb. J. Math.* **5**:1 (2017), 1–122. MR

[Kings et al. 2020] G. Kings, D. Loeffler, and S. L. Zerbes, "Rankin–Eisenstein classes for modular forms", *Amer. J. Math.* **142**:1 (2020), 79–138. MR

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

[Lamplugh 2016] J. Lamplugh, "An Euler system for characters over an imaginary biquadratic field", 2016. arXiv 1611.05702

[Lei et al. 2014] A. Lei, D. Loeffler, and S. L. Zerbes, "Euler systems for Rankin–Selberg convolutions of modular forms", *Ann. of Math.* (2) **180**:2 (2014), 653–771. MR

[Lei et al. 2015] A. Lei, D. Loeffler, and S. L. Zerbes, "Euler systems for modular forms over imaginary quadratic fields", *Compos. Math.* **151**:9 (2015), 1585–1625. MR

[Longo and Vigni 2010] M. Longo and S. Vigni, "On the vanishing of Selmer groups for elliptic curves over ring class fields", *J. Number Theory* **130**:1 (2010), 128–163. MR

[Mazur and Rubin 2004] B. Mazur and K. Rubin, *Kolyvagin systems*, Mem. Amer. Math. Soc. **799**, Amer. Math. Soc., Providence, RI, 2004. MR

[Miyake 1989] T. Miyake, *Modular forms*, Springer, 1989. MR

[Momose 1981] F. Momose, "On the l-adic representations attached to modular forms", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **28**:1 (1981), 89–109. MR

[Nekovář 2000] J. Nekovář, "p-adic Abel–Jacobi maps and p-adic heights", pp. 367–379 in *The arithmetic and geometry of algebraic cycles* (Banff, AB, 1998), edited by B. B. Gordon et al., CRM Proc. Lecture Notes **24**, Amer. Math. Soc., Providence, RI, 2000. MR

[Nekovář 2007] J. Nekovář, "The Euler system method for CM points on Shimura curves", pp. 471–547 in *L-functions and Galois representations*, edited by D. Burns et al., London Math. Soc. Lecture Note Ser. **320**, Cambridge Univ. Press, 2007. MR

[Nekovář 2012] J. Nekovář, "Level raising and anticyclotomic Selmer groups for Hilbert modular forms of weight two", *Canad. J. Math.* **64**:3 (2012), 588–668. MR

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

[Rubin 2000] K. Rubin, *Euler systems*, Annals of Math. Studies **147**, Princeton Univ. Press, 2000. MR

[Skinner and Urban 2014] C. Skinner and E. Urban, "The Iwasawa main conjectures for GL_2", *Invent. Math.* **195**:1 (2014), 1–277. MR

[Tamiozzo 2021] M. Tamiozzo, "On the Bloch–Kato conjecture for Hilbert modular forms", *Math. Z.* **299**:1-2 (2021), 427–458. MR

[Tate 1979] J. Tate, "Number theoretic background", pp. 3–26 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **XXXIII.2**, Amer. Math. Soc., Providence, RI, 1979. MR

[Tian 2003] Y. Tian, *Euler systems of CM points on Shimura curves*, Ph.D. thesis, Columbia University, 2003, available at https://www.proquest.com/docview/305333682. MR

[Vatsal 2003] V. Vatsal, "Special values of anticyclotomic L-functions", *Duke Math. J.* **116**:2 (2003), 219–261. MR

[Wan 2015] X. Wan, "The Iwasawa main conjecture for Hilbert modular forms", *Forum Math. Sigma* **3** (2015), art. id. e18. MR

[Wang 2023] H. Wang, "On the anticyclotomic Iwasawa main conjecture for Hilbert modular forms of parallel weights", *Ann. Math. Qué.* **47**:1 (2023), 195–248. MR

[Wiles 1988] A. Wiles, "On ordinary λ-adic representations associated to modular forms", *Invent. Math.* **94**:3 (1988), 529–573. MR

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

[Zhang 1997] S. Zhang, "Heights of Heegner cycles and derivatives of L-series", *Invent. Math.* **130**:1 (1997), 99–152. MR

Communicated by Shou-Wu Zhang
Received 2024-09-30 Revised 2025-05-05 Accepted 2025-07-01

ktdo@ucsb.edu *Department of Mathematics, University of California, Los Angeles, Los Angeles, CA, United States*

Current address: *Department of Mathematics, University of California, Santa Barbara, Santa Barbara, CA, United States*

ALGEBRA AND NUMBER THEORY 20:8 (2026)
https://doi.org/10.2140/ant.2026.20.1649

Explicit Hecke descent for special cycles

Syed Waqar Ali Shah

We derive an explicit formula for the action of a geometric Hecke correspondence on special cycles on a Shimura variety in terms of such cycles at a fixed neat level and compare it with another closely related expression that is valid for modular curves. We provide evidence that the two formulas do not agree in general.

1. Introduction 1649
2. Recollections 1653
3. Abstract pushforwards 1656
4. The formula 1662
5. Counterexamples 1668
6. Schwartz spaces 1673
Acknowledgements 1675
References 1675

1. Introduction

The study of special cycles on Shimura varieties often entails aspects of smooth representation theory. Kudla [1997] considered certain weighted linear combinations of special cycles on orthogonal Shimura varieties whose equivariance properties are neatly captured by Schwartz spaces admitting a smooth group action. Similar constructions have appeared in [Cornut 2018; Jetchev 2014; Li and Liu 2021; Lai and Skinner 2024] and several other recent works. One motivation for studying such equivariant labeling of cycles is to facilitate the study of Hecke actions on them. Although conventions differ from one author to another, it seems to us that such descriptions can give rise to two fundamentally different ways of defining the said Hecke action, and only one of these a priori agrees with the geometric action of Hecke correspondences in all situations. Unfortunately, the "nongeometric" action seems to be occasionally used in the literature and assumed to be compatible with the geometric one.

The purpose of this note is twofold. First, we derive a general formula for the action of a Hecke correspondence on a given irreducible special cycle in terms of such cycles that are all at a fixed (finite) level, assuming some mild conditions on the level and the Shimura data. This formula also generalizes to any abstract pushforward construction involving such cycles, e.g., Gysin maps in étale cohomology. Second, we address the aforementioned compatibility of the two actions on the group of equidimensional algebraic cycles and provide evidence that they only agree under very special circumstances.

MSC2020: 11G18, 14C25, 14G35.
Keywords: special cycles, Hecke correspondences, Shimura varieties, descent.

© 2026 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

1.1. ***Motivation.*** We motivate our question in the familiar setting of modular curves. Let K be a compact–open subgroup of $\mathrm{GL}_2(\mathbb{A}_f)$ that is contained in a principal congruence subgroup of level $N \geq 3$.[1] Let Y_K denote the modular curve of level K in the sense of [Deligne 1971]. It is a smooth quasiprojective algebraic curve defined over $\mathbb{Q}$ with complex points given by

$$Y_K(\mathbb{C}) = \mathrm{GL}_2(\mathbb{Q})\backslash(\mathscr{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K),$$

where $\mathscr{H}^{\pm} := \mathbb{C} \setminus \mathbb{R}$ and $\mathbb{A}_f$ denotes the ring of finite rational adeles. For $(x, g) \in \mathscr{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)$, we denote by $[x, g]_K \in Y_K(\mathbb{C})$ the class of (x, g). Let $\mathscr{C}(K) := \mathbb{Z}[Y_K(\mathbb{C})]$ denote the group of complex divisors on $Y_K(\mathbb{C})$. For $\sigma \in \mathrm{GL}_2(\mathbb{A}_f)$, we write $K_\sigma := K \cap \sigma K \sigma^{-1}$ and $K^\sigma := K \cap \sigma^{-1} K \sigma$ for brevity. Consider the diagram of $\mathbb{Q}$-varieties

$$Y_K \xleftarrow{\ p\ } Y_{K_\sigma} \xrightarrow[\sim]{[\sigma]} Y_{K^\sigma} \xrightarrow{\ q\ } Y_K$$

in which p, q are the natural degeneracy maps induced by the inclusions $K_\sigma \hookrightarrow K$, $K^\sigma \hookrightarrow K$, respectively, and $[\sigma]$ is the twisting isomorphism given on $\mathbb{C}$-points by $[x, g]_{K_\sigma} \mapsto [x, g\sigma]_{K^\sigma}$. The maps p and q are finite étale and induce pushforwards and pullbacks on divisors with the expected properties. The $\mathbb{Z}$-linear map $[K\sigma K]_* = q_* \circ [\sigma]_* \circ p^* : \mathscr{C}(K) \to \mathscr{C}(K)$ is called the *covariant Hecke correspondence*[2] induced by σ. It is easily seen that

$$[K\sigma K]_* \cdot [x, g]_K = \sum_{\gamma \in K\sigma K/K} [x, g\gamma]_K. \tag{1.1}$$

On the other hand, one can construct the inductive limit

$$\widehat{\mathscr{C}} := \varinjlim_L \mathscr{C}(L)$$

over all (sufficiently small) levels L of $\mathrm{GL}_2(\mathbb{A}_f)$, where the transition maps for the limit are given by pullbacks along the degeneracy maps $Y_{L'} \to Y_L$ induced by the inclusions $L' \hookrightarrow L$. This comes equipped with a smooth left action

$$\mathrm{GL}_2(\mathbb{A}_f) \times \widehat{\mathscr{C}} \to \widehat{\mathscr{C}}$$

given by pullbacks along the twisting isomorphisms. Explicitly, if $z \in \widehat{\mathscr{C}}$ can be represented by $[x, g]_L \in \mathscr{C}(L)$, then $h \cdot z$ for $h \in \mathrm{GL}_2(\mathbb{A}_f)$ is the class of $[x, gh^{-1}]_{hLh^{-1}}$ in $\widehat{\mathscr{C}}$. Since divisors satisfy étale descent, one has $\mathscr{C}(K) = \widehat{\mathscr{C}}^K$, obtained via the inclusion $\mathscr{C}(K) \hookrightarrow \widehat{\mathscr{C}}^K$. Under this identification, the pushforward $\mathscr{C}(L) \to \mathscr{C}(K)$ along the degeneracy map induced by the inclusion of any (not necessarily normal) subgroup L of K corresponds to the norm map $\sum_{\gamma \in K/L} \gamma$. In particular, q_* is identified with $\sum_{K/K^\sigma} \gamma$ and $[K\sigma K]_*$ with $\sum_{\gamma \in K/K^\sigma} \gamma\sigma^{-1}$. Thus,

$$[K\sigma K]_* \cdot [x, g]_K = \sum_{\delta \in K\backslash K\sigma K} [x, g\delta]_{\delta^{-1}K\delta}, \tag{1.2}$$

[1] See [Shah 2023, Section 2.4] for a treatment of the case of a general compact–open subgroup K.

[2] This is sometimes referred to as the "Albanese" convention for Hecke correspondences [Ribet 1990, p. 443]. See also [Lei et al. 2015, Remarks 4.1.1 and 4.3.2] and [Lei et al. 2018, Section 2.6] for a discussion.

where the right-hand side is interpreted as an element of $\widehat{\mathscr{C}}^K$. Note that individual points in this expression live on modular curves of *different* levels.

The right-hand side of (1.2) computing the action of $[K\sigma K]_*$ is formal in nature and remains valid, for instance, in the cohomology of any Shimura variety and, more generally, any locally symmetric space. The right-hand side of (1.1) is, however, more useful, since it explicitly gives a divisor in terms of points that all share the original level K. It is also what one usually finds in the literature on Shimura curves; see, for example, [Cornut and Vatsal 2007, Section 3.4; Zhang 2001, Section 1.4]. We say that (1.1) provides an *explicit descent formula* for the formal K-invariant expression given by (1.2).

1.2. ***Main problem.*** We may make similar considerations for special cycles on Shimura varieties. Say E is an imaginary quadratic field and $x = x_\iota$ is defined as the image of "pt" under the morphism of Shimura data

$$\iota : (\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m, \{\mathrm{pt}\}) \hookrightarrow (\mathrm{GL}_2, \mathscr{H}^{\pm}) \tag{1.3}$$

obtained by considering E as a $\mathbb{Q}$-vector space. Then the points $z_L(g) := [x_\iota, g]_L \in Y_L(\mathbb{C})$ for $g \in \mathrm{GL}_2(\mathbb{A}_f)$ are algebraic and referred to as special points (or CM points). One may equivalently describe $z_L(g)$ as the point obtained by taking the distinguished geometric connected component "[1]" of the zero-dimensional Shimura variety for $\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m$ of level $(\mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_m)(\mathbb{A}_f) \cap gLg^{-1}$, embedding it into $Y_{gLg^{-1}}$, and taking its image under the twisting isomorphism $[g] : Y_{gLg^{-1}} \to Y_L$. See [Shah 2023, Section 2] for details on this setup. In this paradigm, we may replace (1.3) by an arbitrary embedding of Shimura data

$$\iota : (\boldsymbol{H}, X_{\boldsymbol{H}}) \hookrightarrow (\boldsymbol{G}, X_{\boldsymbol{G}})$$

and construct, for any compact–open subgroup L of $\boldsymbol{G}(\mathbb{A}_f)$, *irreducible special cycles* $z_L(g)$ (defined over the algebraic closure $\overline{\mathbb{Q}} \subset \mathbb{C}$ of $\mathbb{Q}$) of a fixed codimension n on $\mathrm{Sh}_{\boldsymbol{G}}(L)$ for any $g \in \boldsymbol{G}(\mathbb{A}_f)$ in an analogous fashion. Let $\mathscr{C}^n(L)$ denote the free $\mathbb{Z}$-module on codimension n irreducible $\overline{\mathbb{Q}}$-subvarieties on $\mathrm{Sh}_{\boldsymbol{G}}(L)$. One may then ask for the analog of (1.1) for the Hecke action on the special cycle $z_L(g)$ in the space $\mathscr{C}^n(L)$.

More precisely, let K denote a fixed compact–open subgroup of $\boldsymbol{G}(\mathbb{A}_f)$. To avoid pathologies arising from degrees of degeneracy maps, we assume that K is *neat* [Pink 1989, Section 0.1]. Fix a $\sigma \in \boldsymbol{G}(\mathbb{A}_f)$ and write $K_\sigma := K \cap \sigma K \sigma^{-1}$ and $K^\sigma := K \cap \sigma^{-1} K \sigma$ as before. We define the *covariant Hecke correspondence*

$$[K\sigma K]_* : \mathscr{C}^n(K) \xrightarrow{\mathrm{pr}^*} \mathscr{C}^n(K_\sigma) \xrightarrow{[\sigma]_*} \mathscr{C}^n(K^\sigma) \xrightarrow{\mathrm{pr}_*} \mathscr{C}^n(K),$$

where pr^* (resp. pr_*) denotes the flat pullback (resp. proper pushforward) of cycles along the degeneracy map induced by the inclusion $K_\sigma \hookrightarrow K$ (resp. $K^\sigma \hookrightarrow K$) and $[\sigma]_*$ is the isomorphism induced by pushforward along the twisting map[3] $[\sigma] : \mathrm{Sh}_{\boldsymbol{G}}(K_\sigma) \to \mathrm{Sh}_{\boldsymbol{G}}(K^\sigma)$. See also [Noot 2006, Section 6.2; Esnault and Harris 2017, Section 1.2]. In analogy with (1.2), it is not hard to show that, for any $g \in \boldsymbol{G}(\mathbb{A}_f)$,

[3]Throughout, the action of $\sigma \in \boldsymbol{G}(\mathbb{A}_f)$ on the tower of Shimura varieties is denoted by $[\sigma]$ and gives a *right action* of $\boldsymbol{G}(\mathbb{A}_f)$, as in [Milne 2005, Definition 5.14].

we have

$$[K\sigma K]_* \cdot z_K(g) = \sum_{\delta \in K \backslash K\sigma K} z_{\delta^{-1}K\delta}(g\delta),$$

where the right-hand side is again viewed as an element in the inductive limit of $\mathcal{C}^n(L)$ over all (sufficiently small) levels L. A natural question that arises at this point is the following.

Question 1.4. *Is there a $\mathbb{Z}$-linear combination of irreducible special cycles $z_K(-) \in \mathcal{C}^n(K)$ that equals $[K\sigma K]_* z_K(g)$? Equivalently, is it possible to write the K-invariant limit expression $\sum_{\delta \in K\backslash K\sigma K} z_{\delta^{-1}K\delta}(g\delta)$ in terms of such cycles? If so, can one give a formula for it?*

A first guess suggested by (1.1) would be that

$$[K\sigma K]_* \cdot z_K(g) \stackrel{?}{=} \sum_{\gamma \in K\sigma K/K} z_K(g\gamma). \tag{1.5}$$

Unfortunately, this is not necessarily always the case. While the answer to Question 1.4 is affirmative, the correct expression is given by Corollary 4.20 below and involves various nontrivial coefficients. In fact, the number of irreducible special cycles in our expression seems to be far smaller in general than what is prescribed by (1.5). See Section 5 for computations. This discrepancy continues to exist if one replaces $z_K(g)$ with the full fundamental cycle of the source Shimura variety. Analogs of (1.5) seem to be used in [Boumasmoud et al. 2020, Section 3.1; Li and Liu 2021, p. 846] and various other writings.

On the other hand, we show that, for the cases of zero cycles and codimension zero cycles, our expression does agree with (1.5). This is of course in agreement with (1.1). For a simple group-theoretic explanation of this phenomenon, see Section 6.

Our approach to Question 1.4 is based on an elementary but useful idea: that explicit distribution relations among objects associated with $\mathrm{Sh}_{\boldsymbol{H}}$ (such as fundamental cycles of $\mathrm{Sh}_{\boldsymbol{H}}$) can be parlayed into corresponding relations on $\mathrm{Sh}_{\boldsymbol{G}}$ via a gadget we refer to as the *mixed Hecke correspondence* [Shah 2024, Section 2]. This approach utilizes the built-in functoriality properties of pushforwards and pullbacks of cycles on schemes. Consequently, our methods readily apply in the more general settings of pushforwards of cycle classes into cohomology with coefficients in local systems on $\mathrm{Sh}_{\boldsymbol{G}}$, such as the one studied in [Graham and Shah 2023]. This is a nontrivial extension, since cohomology with integral coefficients does not satisfy Galois descent in general. We establish this formula in Theorem 4.12 for abstract pushforwards using the language of RIC functors developed in [Shah 2024, Section 2], which we briefly recall in Section 3.2. We then derive the expression asked for in Question 1.4 in Corollary 4.20. We anticipate that our formula will find applications in establishing Euler system norm relations when one works with more general coefficient systems, as, for instance, in the approach originally envisioned by Cornut [2018].

1.3. ***Outline.*** This note is structured as follows. In Section 2, we collect relevant facts about Shimura varieties needed in the derivation of our formula. In Section 3, we study the special cycles $z_K(-)$ from the introduction and their analogs for abstract pushforwards. In Section 4, we derive our main formula and provide some simple examples. In Section 5, we furnish two sets of counterexamples to (1.5). In

Section 6, we investigate an analogous question in the setting of function spaces that provides another perspective on the issue. This section may be read independently of the rest of this note.

2. Recollections

In this section, we recall some basic facts about Shimura varieties that we will need later on. We also generalize a result from [Loeffler et al. 2022], which provides a simple criterion for checking when maps between Shimura varieties are closed immersions. It is, however, only needed in Section 5.

2.1. ***Shimura varieties.*** Suppose that

$$\iota : (\boldsymbol{H}, X_{\boldsymbol{H}}) \hookrightarrow (\boldsymbol{G}, X_{\boldsymbol{G}}) \tag{2.1}$$

is an embedding of Shimura data, i.e., $\boldsymbol{H}$, $\boldsymbol{G}$ are connected reductive algebraic groups over $\mathbb{Q}$, the pairs $(\boldsymbol{H}, X_{\boldsymbol{H}})$, $(\boldsymbol{G}, X_{\boldsymbol{G}})$ satisfy the Milne–Deligne axioms (SV1)–(SV3) [Milne 2005, Definition 5.5], and $\iota : \boldsymbol{H} \to \boldsymbol{G}$ is an embedding preserving the said datum. Denote by $\mathbb{A}_f$ the ring of finite rational adeles and set

$$H := \boldsymbol{H}(\mathbb{A}_f), \quad G := \boldsymbol{G}(\mathbb{A}_f).$$

For each neat [Pink 1989, Section 0.1] compact–open subgroup K of G, there is a smooth equidimensional quasiprojective Shimura variety $\mathrm{Sh}_{\boldsymbol{G}}(K)$ defined over a canonically defined subfield of $\mathbb{C}$, called the reflex field of $\boldsymbol{G}$, and similarly for $\boldsymbol{H}$. We will, however, exclusively consider all Shimura varieties as objects in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$, the category of schemes over the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$.

The $\mathbb{C}$-points of $\mathrm{Sh}_{\boldsymbol{G}}(K)$ are given by $\boldsymbol{G}(\mathbb{Q})\backslash(X_{\boldsymbol{G}} \times G/K)$ and a general point is denoted as $[x, g]_K$, where $x \in X_{\boldsymbol{G}}$ and $g \in G$. Similar notation and conventions will be used for corresponding objects associated with $\boldsymbol{H}$. We set

$$n := \dim \mathrm{Sh}_{\boldsymbol{G}}(K) - \dim \mathrm{Sh}_{\boldsymbol{H}}(U) \tag{2.2}$$

for some choice of compact–open subgroups $K \subset G$ and $U \subset H$. Then n is independent of K and U. In what follows, we think of H as a closed subgroup of G. Then, for any compact–open subgroup K of G, the intersection $H \cap K := H \cap \iota^{-1}(K)$ is a compact–open subgroup of H, which is neat if K is.

2.2. ***Maps.*** For any $g \in G$ and any neat compact–open subgroup K of G, there is a *twisting isomorphism*

$$[g] = [g]_{K, g^{-1}Kg} : \mathrm{Sh}_{\boldsymbol{G}}(K) \xrightarrow{\sim} \mathrm{Sh}_{\boldsymbol{G}}(g^{-1}Kg) \tag{2.3}$$

given on $\mathbb{C}$-points by $[x, \gamma]_K \mapsto [x, \gamma g]_{g^{-1}Kg}$ for all $x \in X_{\boldsymbol{G}}$ and $\gamma \in G$. Let $\boldsymbol{Z}_{\boldsymbol{G}}$ (resp. $\boldsymbol{Z}_{\boldsymbol{H}}$) denote the center of $\boldsymbol{G}$ (resp. $\boldsymbol{H}$). For any compact–open subgroup L contained in K that satisfies

$$K \cap \boldsymbol{Z}_{\boldsymbol{G}}(\mathbb{Q}) = L \cap \boldsymbol{Z}_{\boldsymbol{G}}(\mathbb{Q}), \tag{2.4}$$

the natural surjection $\mathrm{pr}_{L,K} : \mathrm{Sh}_{\boldsymbol{G}}(L) \to \mathrm{Sh}_{\boldsymbol{G}}(K)$ given by $[x, g]_L \mapsto [x, g]_K$ is finite étale of degree $[K : L]$, by [Shah 2024, Lemma 2.7.1]. When (SV5) of [Milne 2005] holds for $(\boldsymbol{G}, X_{\boldsymbol{G}})$, the condition (2.4) is

automatic by neatness of K. One may also consider the limit

$$\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C}) := \varprojlim_{L} \mathrm{Sh}_{\boldsymbol{G}}(L)(\mathbb{C})$$

taken over all neat compact–open subgroups L of $\boldsymbol{G}(\mathbb{A}_f)$ along the degeneracy maps induced by the inclusions $L' \hookrightarrow L$. Then $\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})$ inherits a continuous right action of $G = \boldsymbol{G}(\mathbb{A}_f)$ and $\mathrm{Sh}_{\boldsymbol{G}}(K)(\mathbb{C}) = \mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})/K$ for any K. If $(\boldsymbol{G}, X_{\boldsymbol{G}})$ satisfies (SV5), any neat compact–open subgroup acts freely on $\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})$.

We also have maps between Shimura varieties of $\boldsymbol{H}$ and $\boldsymbol{G}$. If K is a neat compact–open subgroup of G and U is a compact–open subgroup of $H \cap K$, there is a finite unramified morphism

$$\iota_{U,K} : \mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(K) \tag{2.5}$$

of smooth schemes in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$, which is given on $\mathbb{C}$-points by $[y, h]_U \mapsto [y, h]_K$ for all $y \in X_{\boldsymbol{H}}$ and $h \in H$. We can always choose a compact–open subgroup $L = L_U$ of G which contains U such that $\mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(L)$ is a closed immersion [Deligne 1971, Proposition 1.15]. We may assume that $L \subset K$ by replacing L with its intersection with K if necessary, since $\mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(L)$ factors as

$$\mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(L \cap K) \to \mathrm{Sh}_{\boldsymbol{G}}(L)$$

and [Stacks 2005–, Lemma 07RK] implies that $\mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(L \cap K)$ is also a closed immersion. We may moreover assume that (2.4) holds for L. Indeed, if we let $L' := L\Gamma$, where $\Gamma := \boldsymbol{Z}_{\boldsymbol{G}}(\mathbb{Q}) \cap K$, we have $L \subset L' \subset K$, and it is elementary to show that $\mathrm{Sh}_{\boldsymbol{G}}(L) = \mathrm{Sh}_{\boldsymbol{G}}(L')$. In any case, $\iota_{U,K}$ is a composition of a closed immersion followed by a finite étale surjection and one may therefore apply various natural pushforward constructions in this setting (see, e.g., [Graham and Shah 2023, Proposition A.5]). In all cases, the natural map

$$\iota_\infty : \mathrm{Sh}_{\boldsymbol{H}}(\mathbb{C}) \to \mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C}) \tag{2.6}$$

at the infinite level is injective, again by [Deligne 1971, Proposition 1.15]. The following generalization of [Loeffler et al. 2022, Proposition 5.3.1] gives a criterion for immersion at finite levels.

Lemma 2.7. *Assume that* $(\boldsymbol{G}, X_{\boldsymbol{G}})$ *satisfies* (SV5) *and that there exists a* $w \in \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q})$ *such that* $\boldsymbol{H}$ *is the centralizer of* w *in* $\boldsymbol{G}$. *Let* K *be a compact–open subgroup of* G *such that there exists a neat compact–open subgroup of* G *that contains both* K *and* wKw^{-1}. *Then* $\iota_{H\cap K,K}$ *is a closed immersion.*

Proof. Since $\iota_{H\cap K,K}$ is a finite unramified morphism, it suffices by [Stacks 2005–, Lemma 04XV] to check that $\iota_{H\cap K,K}$ is universally injective. By [ibid., Lemma 01S4], this amounts to the surjectivity of the diagonal map

$$\Delta : \mathrm{Sh}_{\boldsymbol{H}}(H \cap K) \to \mathrm{Sh}_{\boldsymbol{H}}(H \cap K) \times_{\mathrm{Sh}_{\boldsymbol{G}}(K)} \mathrm{Sh}_{\boldsymbol{H}}(H \cap K).$$

By [Demazure and Gabriel 1980, I, Section 3, 6.11],[4] one can check surjectivity on closed points and, therefore, also on $\mathbb{C}$-points. But it is easily seen that $\Delta(\mathbb{C})$ is surjective if and only if $\iota_{H\cap K,K}(\mathbb{C})$ is injective. So it suffices to check the injectivity of $\iota_{H\cap K,K}$ on $\mathbb{C}$-points.

[4] See also the notion of *ultraschemes* in [EGA I 1971, Appendice].

To this end, say there exists a $P \in \mathrm{Sh}_{\boldsymbol{H}}(\mathbb{C})$ and $\kappa \in K$ such that $P\kappa$ lies in $\mathrm{Sh}_{\boldsymbol{H}}(\mathbb{C})$, where we are viewing $\mathrm{Sh}_{\boldsymbol{H}}(\mathbb{C})$ as a subset of $\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})$ via (2.6). Since $w \in \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q})$ fixes $X_{\boldsymbol{H}}$ pointwise and commutes with all elements of $\boldsymbol{H}(\mathbb{A}_f)$, the (right) action of w on $\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})$ fixes $\mathrm{Sh}_{\boldsymbol{H}}(\mathbb{C})$ pointwise. Therefore,

$$P w\kappa = P\kappa = P\kappa w.$$

This implies that $\lambda := \kappa w\kappa^{-1}w^{-1}$ stabilizes P. If L is a compact–open subgroup of G that contains both K and wKw^{-1}, then $\lambda = \kappa(w\kappa^{-1}w^{-1}) \in L$. Recall that, if $(\boldsymbol{G}, X_{\boldsymbol{G}})$ satisfies (SV5), then the right action of any neat compact–open subgroup of G on $\mathrm{Sh}_{\boldsymbol{G}}(\mathbb{C})$ is free. Thus, if L is neat, the only element of L that can stabilize P is the identity. So our assumptions imply that λ must be the identity, i.e., $w = \kappa w\kappa^{-1}$. Since the centralizer of w in $\boldsymbol{G}$ is $\boldsymbol{H}$, we have $\kappa \in H$. Therefore, $\kappa \in H \cap K$ and the classes of P and $P\kappa$ in $\mathrm{Sh}_{\boldsymbol{H}}(H \cap K)(\mathbb{C})$ are forced to be equal. This implies that $\iota_{H\cap K,K}$ is injective on $\mathbb{C}$-points and is therefore a closed immersion. □

Remark 2.8. While we exclusively work in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$, many of our results hold true over canonical models over the reflex field of $\boldsymbol{H}$. For instance, (2.5) is defined over this field and the criterion of Lemma 2.7 applies if $\iota_{H\cap K,K}$ is viewed over this number field.

2.3. ***Connected components.*** Recall that we are viewing all Shimura varieties in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$. To simplify the description of the connected components of the Shimura varieties of $\boldsymbol{H}$, we assume for the rest of this note that the derived group $\boldsymbol{H}^{\mathrm{der}}$ is simply connected. Let $\boldsymbol{T} = \boldsymbol{H}/\boldsymbol{H}^{\mathrm{der}}$ and $\nu : \boldsymbol{H} \to \boldsymbol{T}$ denote the natural map. Let $\boldsymbol{T}(\mathbb{R})^{\dagger} \subset \boldsymbol{T}(\mathbb{R})$ denote the image of the real points of the center of $\boldsymbol{H}$ in the real points of $\boldsymbol{T}$ and let $\boldsymbol{T}(\mathbb{Q})^{\dagger} := \boldsymbol{T}(\mathbb{Q}) \cap \boldsymbol{T}(\mathbb{R})^{\dagger}$. Let U be a neat compact–open subgroup of H. Then $\nu(U)$ is a compact–open subgroup of $\boldsymbol{T}(\mathbb{A}_f)$ by [Milne 2005, Lemma 5.21] and neat by [Borel 2019, Corollary 17.3]. The set of connected components of $\mathrm{Sh}_{\boldsymbol{H}}(U)$ can be described via an explicit bijection (of sets)

$$\pi_0(\mathrm{Sh}_{\boldsymbol{H}}(U))(\mathbb{C}) \xrightarrow{\sim} \boldsymbol{T}(\mathbb{Q})^{\dagger}\backslash\boldsymbol{T}(\mathbb{A}_f)/\nu(U) \tag{2.9}$$

once a connected component $X_{\boldsymbol{H}}^{+} \subset X_{\boldsymbol{H}}$ in the analytic topology of $X_{\boldsymbol{H}}$ is fixed [Milne 2005, Theorem 5.17], which we do in what follows. Since $\pi_0(\mathrm{Sh}_{\boldsymbol{H}}(U))$ is a finite étale scheme and we are working over $\overline{\mathbb{Q}}$, there is no harm in identifying it with

$$\pi_0(\mathrm{Sh}_{\boldsymbol{H}}(U))(\overline{\mathbb{Q}}) \xrightarrow{\sim} \pi_0(\mathrm{Sh}_{\boldsymbol{H}}(U))(\mathbb{C}),$$

and we will merely view these as sets without a scheme structure. We however do view $\pi_0(\mathrm{Sh}_{\boldsymbol{H}}(U))$ as an abelian group by transport of structure via (2.9), and will denote this abelian group by $\pi_{0,U}$ for the rest of this note for simplicity.

Definition 2.10. By Y_U°, we denote the connected component of $\mathrm{Sh}_{\boldsymbol{H}}(U)$ over the identity element $[1] \in \pi_{0,U}$, which we consider as an object in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$. If V is a compact–open subgroup of U, we denote by $\varphi_{V,U} : \pi_{0,V} \to \pi_{0,U}$ the quotient homomorphism.

For convenience, we will write Y° for Y_U° when the level is apparent from context. We observe that the connected components of $\mathrm{Sh}_{\boldsymbol{H}}(U)$ are equidimensional of dimension $\dim \mathrm{Sh}_{\boldsymbol{H}}(U)$. It is easy to see

that the twisting isomorphism $[h]:\mathrm{Sh}_{\boldsymbol{H}}(U)\to\mathrm{Sh}_{\boldsymbol{H}}(h^{-1}Uh)$ sends Y_U° to the connected component of $\mathrm{Sh}_{\boldsymbol{H}}(h^{-1}Uh)$ indexed by the class of $\nu(h)$ in $\pi_{0,h^{-1}Uh}=\pi_{0,U}$.

Remark 2.11. Each component in $\pi_{0,U}$ is defined over a finite abelian extension of the reflex field, which is determined by an explicit Shimura–Deligne reciprocity law, and the eventual expression we derive for the Hecke action in Question 1.4 is defined over the compositum of the fields of definition of the cycles involved. This is, however, not relevant to the issues considered in this note.

3. Abstract pushforwards

In this section, we carefully define the cycles $z_K(-)$ from the introduction and consider their analogs in the more general setting of abstract pushforwards, which model the behavior of any suitable cohomology theory for Shimura varieties.

3.1. ***Special cycles.*** We maintain the notation and assumptions of Section 2. For each integer $m\geq 0$ and neat compact–open subgroup K of G, we denote by $\mathscr{C}^m(K)$ the free abelian group on the set of codimension m closed integral $\overline{\mathbb{Q}}$-subschemes of $\mathrm{Sh}_{\boldsymbol{G}}(K)$. We refer to elements of $\mathscr{C}^m(K)$ as *algebraic cycles of codimension m on* $\mathrm{Sh}_{\boldsymbol{G}}(K)$. Recall from (2.2) that n denotes the codimension of Shimura varieties of $\boldsymbol{H}$ in those of $\boldsymbol{G}$. Recall also that $\iota_{H\cap K,K}$ of (2.5) is finite and therefore proper. In particular, $\iota_{H\cap K,K}$ is a closed map.

Definition 3.1. For $g\in G$, we denote by $z_K(g)\in\mathscr{C}^n(K)$ the algebraic cycle on $\mathrm{Sh}_{\boldsymbol{G}}(K)$ given by the fundamental cycle of the reduced closed subscheme of $\mathrm{Sh}_{\boldsymbol{G}}(K)$ whose underlying topological space is given by the (necessarily irreducible and closed) image of $Y^\circ_{H\cap gKg^{-1}}$ under the morphism

$$\mathrm{Sh}_{\boldsymbol{H}}(H\cap gKg^{-1})\xrightarrow{\iota}\mathrm{Sh}_{\boldsymbol{G}}(gKg^{-1})\xrightarrow{[g]}\mathrm{Sh}_{\boldsymbol{G}}(K). \tag{3.2}$$

We refer to these as *irreducible special cycles*. We denote by $\mathscr{Z}_K$ the $\mathbb{Z}$-submodule of $\mathscr{C}^n(K)$ spanned by $z_K(g)$ for $g\in G$.

One may equivalently describe $z_K(g)$ as the fundamental cycle associated with the integral $\overline{\mathbb{Q}}$-subscheme of $\mathrm{Sh}_{\boldsymbol{G}}(K)$ whose $\mathbb{C}$-points are the image of $X^+_{\boldsymbol{H}}\times g$ in $\mathrm{Sh}_{\boldsymbol{G}}(K)(\mathbb{C})$. Then it is easily seen that the twisting isomorphism $[h]_{K,h^{-1}Kh}$ for any $h\in G$ sends $z_K(g)$ to $z_{h^{-1}Kh}(gh)$.

It is clear that the $z_K(g)$ are not uniquely labeled by $g\in G$. For instance, $z_K(g)=z_K(g\kappa)$ for any $\kappa\in K$. Let $\boldsymbol{H}(\mathbb{R})_+$ denote the stabilizer of $X^+_{\boldsymbol{H}}$ in $\boldsymbol{H}(\mathbb{R})$ and set

$$\boldsymbol{H}(\mathbb{Q})_+:=\boldsymbol{H}(\mathbb{R})_+\cap\boldsymbol{H}(\mathbb{Q}).$$

From the explicit description of $\mathbb{C}$-points underlying the scheme of $z_K(g)$, it is also clear that $z_K(g)=z_K(hg)$ for any $h\in\boldsymbol{H}(\mathbb{Q})_+$. Thus, the cycles $z_K(g)$ can be indexed by $\boldsymbol{H}(\mathbb{Q})_+\backslash G/K$ and we obtain a $\mathbb{Z}$-linear surjection

$$\Phi:\mathscr{S}_{\mathbb{Z}}(\boldsymbol{H}(\mathbb{Q})_+\backslash G/K)\to\mathscr{Z}_K, \tag{3.3}$$

where $\mathcal{S}_{\mathbb{Z}}(X)$ for a set X denotes the free $\mathbb{Z}$-module on the elements of X. Since G/K is discrete, we are free to replace $\boldsymbol{H}(\mathbb{Q})_+$ by its topological closure inside G in (3.3). If $\boldsymbol{H}^{\mathrm{der}}(\mathbb{R})$ is noncompact and contained in $\boldsymbol{H}(\mathbb{R})_+$, then this closure contains $\boldsymbol{H}^{\mathrm{der}}(\mathbb{A}_f)$ by strong approximation. In this case, one obtains a large set of relations among the labels g for $z_K(g)$, even locally at each prime.

Remark 3.4. A detailed study of such cycles for certain orthogonal Shimura varieties is done in [Cornut 2018, Section 5.14]. In [ibid., Proposition 5.1], it is shown using the Baire category theorem that a map closely related to Φ is in fact a bijection. A similar fact is established in [Jetchev 2014, Section 2.3]. See Remark 6.6 for an alternative approach to Question 1.4 when such a description is available. In general, it seems unclear what the full set of relations among the labels of these cycles is.

3.2. ***RIC functors.*** To address Question 1.4 at a greater level of generality and to streamline certain arguments, we will require some notions of functors on compact–open subgroups of locally profinite groups from [Shah 2024, Section 2]. These notions roughly coincide with those introduced in [Graham and Shah 2023, Section 2], except that the monoid "Σ" is taken to be the full group and axiom "(T2)" of [loc. cit.] is relaxed slightly. Note also that the terminology in the former has been slightly updated to acknowledge the prior widespread use of these functors (and their variations) [Lewis et al. 1981; Bley and Boltje 2004; Barwick 2017]. We briefly recall these notions below and explain their relevance in our context.

Fix any nonempty collection Υ_G of neat compact–open subgroups of $G = \boldsymbol{G}(\mathbb{A}_f)$ which is closed under intersections and conjugations by G, where any two subgroups have the same intersection with $\boldsymbol{Z}_{\boldsymbol{G}}(\mathbb{Q})$ and, for any $K, L \in \Upsilon_G$, there exists a third subgroup in Υ_G that is contained in $K \cap L$ and normal in K. It is elementary to see that such collections always exist [Shah 2024, Lemma 2.1.1]. We let $\mathcal{P}(G, \Upsilon_G)$ denote the small category whose set of objects is Υ_G and whose morphisms are given by $\mathrm{Hom}_{\mathcal{P}(G,\Upsilon_G)}(L, K) = \{g \in G \mid g^{-1}Lg \subset K\}$ for all $L, K \in \Upsilon_G$. A morphism will usually be written as either $(L \xrightarrow{g} K)$ or $[g]_{L,K}$, where, in the latter notation, we omit the subscripts if they are understood from context.[5] Composition in $\mathcal{P}(G, \Upsilon_G)$ is given by $[g] \circ [h] = [hg]$. If $1 = 1_G$ denotes the identity of G, the morphism $[1]_{L,K}$ is also denoted by $\mathrm{pr}_{L,K}$. Fix a commutative ring with identity R.

Definition 3.5. A *RIC functor M on (G, Υ_G) valued in R-*Mod is a pair of covariant functors $M^* : \mathcal{P}(G, \Upsilon_G)^{\mathrm{op}} \to R\text{-Mod}$ and $M_* : \mathcal{P}(G, \Upsilon_G) \to R\text{-Mod}$ satisfying the following three conditions:

(C1) $M^*(K) = M_*(K)$ for all $K \in \Upsilon_G$. This common R-module is denoted by $M(K)$.

(C2) For all $K \in \Upsilon_G$ and $g \in G$,

$$(gKg^{-1} \xrightarrow{g} K)^* = (K \xrightarrow{g^{-1}} gKg^{-1})_* \in \mathrm{Hom}_{R\text{-Mod}}(M(K), M(gKg^{-1})).$$

Here, for a morphism $\phi \in \mathcal{P}(G, \Upsilon_G)$, we let $\phi_* := M_*(\phi)$ and $\phi^* := M^*(\phi)$.

(C3) $[\gamma]_{K,K,*} : M(K) \to M(K)$ is the identity map for all $K \in \Upsilon_G$ and $\gamma \in K$.

[5] A priori, the notation $[g]$ here conflicts with (2.3). However, $\mathcal{P}(G, \Upsilon_G)$ can be canonically identified with the corresponding system of Shimura varieties and their degeneracy maps. This justifies our abuse of notation.

The pair of functors above will be denoted simply as $M : \mathcal{P}(G, \Upsilon_G) \to R\text{-Mod}$. We say that M is

(G) *Galois* if $\mathrm{pr}^*_{L,K}$ injects $M(K)$ onto $M(L)^{K/L}$ for all $L, K \in \Upsilon_G$ such that $L \triangleleft K$.

(Co) *cohomological* if $\mathrm{pr}_{L,K,*} \circ \mathrm{pr}^*_{L,K} = [K : L] \cdot \mathrm{id}$ for all $L, K \in \Upsilon_G$ with $L \subset K$.

(M) *Mackey* if, for all $K, L, L' \in \Upsilon_G$ with $L, L' \subset K$, we have a commutative diagram

$$\begin{array}{ccc} \bigoplus_\delta M(L'_\delta) & \xrightarrow{\sum [\delta]_*} & M(L) \\ {\scriptstyle \oplus \mathrm{pr}^*}\uparrow & & \uparrow{\scriptstyle \mathrm{pr}^*} \\ M(L') & \xrightarrow[\mathrm{pr}_*]{} & M(K) \end{array} \tag{3.6}$$

where the direct sum is over $\delta \in L' \backslash K / L$, and $L'_\delta := L' \cap \delta L \delta^{-1} \in \Upsilon_G$. This condition is independent of the choice of coset representatives.

When (Co) and (M) are both satisfied, we say that M is *CoMack* for brevity.

Remark 3.7. The acronym "RIC" stands for *restriction*, *induction*, *conjugation* and may be pronounced "Ric". To any smooth left G-representation π, one can attach a RIC functor M_π by setting $M_\pi(K) = \pi^K$ (the K-invariants of π) for all $K \in \Upsilon_G$. For $L \subset K$, the map $\mathrm{pr}^* : M_\pi(K) \to M_\pi(L)$ is the inclusion $\pi^K \hookrightarrow \pi^L$, the map $\mathrm{pr}_* : M_\pi(L) \to M_\pi(K)$ is the trace $\sum_{\gamma \in K/L} \gamma \cdot (-)$, and $[g]^* : M_\pi(K) \to M_\pi(gKg^{-1})$ is given by the action of g on π. Then M_π is Galois, cohomological and Mackey. However, not all RIC functors arise in this manner, e.g., the cohomology of Shimura varieties with $\mathbb{Z}_p$-coefficients generally fails to be Galois.

Remark 3.8. The role of Υ_G in the above is primarily to restrict any given cohomology theory for $\mathrm{Sh}_{\boldsymbol{G}}$ to those levels where pushforwards and pullbacks behave in the expected way (i.e., the cohomology over varying levels constitutes a CoMack functor). This is especially relevant in the case of Hilbert modular varieties, for which axiom (SV5) of [Milne 2005] fails. We observe that some claims in [Grossi 2020, Section 6.1] do not hold for arbitrary neat levels for precisely this reason, but this can be easily remedied by restricting to levels in a collection Υ_G as above and appealing to [Shah 2024, Lemma 2.7.1].

We can similarly define these notions for $H = \boldsymbol{H}(\mathbb{A}_f)$. To speak of maps between functors on H and G, we assume that the collection

$$\Upsilon_H := \{H \cap K \mid K \in \Upsilon_{\boldsymbol{G}}\} \tag{3.9}$$

is such that any two elements in Υ_H have the same intersection with $\boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q})$. This condition is automatic if either $\boldsymbol{Z}_{\boldsymbol{H}} \subset \boldsymbol{Z}_{\boldsymbol{G}}$ or $(\boldsymbol{H}, X_{\boldsymbol{H}})$ satisfies (SV5), but is otherwise a running assumption in this note. The other analogous conditions for Υ_H however always hold.

For convenience, we will call a pair $(U, K) \in \Upsilon_H \times \Upsilon_G$ *compatible* if U is contained in K. We fix for the rest of this subsection RIC functors $N : \mathcal{P}(H, \Upsilon_H) \to R\text{-Mod}$ and $M : \mathcal{P}(G, \Upsilon_G) \to R\text{-Mod}$.

Definition 3.10. A *pushforward* $\iota_* : N \to M$ is a collection of R-module homomorphisms $\iota_{U,K,*} : N(U) \to M(K)$ indexed by compatible pairs (U, K) such that, if $(V, L) \in \Upsilon_H \times \Upsilon_G$ is another such

pair and $h \in H$ labels morphisms $\phi = [h]_{V,U} \in \mathcal{P}(H, \Upsilon_H)$ and $\psi = [h]_{L,K} \in \mathcal{P}(G, \Upsilon_G)$, we have $\psi_* \circ \iota_{V,L,*} = \iota_{U,K,*} \circ \phi_*$.

Definition 3.11. Suppose $\iota_* : N \to M$ is a pushforward. For any $U \in \Upsilon_H$, $K \in \Upsilon_G$ and $\sigma \in G$, the *mixed Hecke correspondence* $[U\sigma K]_* : N(U) \to M(K)$ is defined as the composition

$$N(U) \xrightarrow{\mathrm{pr}^*} N(U \cap \sigma K \sigma^{-1}) \xrightarrow{\iota_*} M(\sigma K \sigma^{-1}) \xrightarrow{[\sigma]_*} M(K).$$

The *degree* of $[U\sigma K]_*$ is defined to be the index $[H \cap \sigma K \sigma^{-1} : U \cap \sigma K \sigma^{-1}]$ and denoted by $\deg[U\sigma K]_*$. We say that the pushforward ι_* is *Mackey* if, for all $V \in \Upsilon_H$ and $L, K \in \Upsilon_G$ satisfying $V, L \subset K$, we have $\mathrm{pr}^*_{L,K} \circ \iota_{V,K,*} = \sum_\gamma [V\gamma L]_*$, where γ runs over $V\backslash K/L$.

Note 3.12. Suppose $\boldsymbol{H} = \boldsymbol{G}$ in the above situation. Then there is a natural pushforward $M \to M$ given by ϕ_* for various $\phi \in \mathcal{P}(G, \Upsilon_G)$, and it is easy to see that M is Mackey if and only if this natural pushforward is. Given $K, K' \in \Upsilon_G$ and $\sigma \in G$, we refer to $[K\sigma K']_* : M(K) \to M(K')$ defined with respect to this pushforward as a *covariant Hecke correspondence* and $[K'\sigma K] := [K\sigma^{-1}K']_* : M(K) \to M(K')$ as a *contravariant Hecke correspondence*. Then our notion of degree applied to these recovers the usual notion of degrees of Hecke correspondences.

3.3. ***Pushforwards of cycles.*** We now specialize the notions of Section 3.2 to the situation of interest. Recall that, for $L \in \Upsilon_G$, $\mathscr{C}^n(L)$ denotes the abelian group of algebraic cycles of codimension n on $\mathrm{Sh}_{\boldsymbol{G}}(L) \in \mathbf{Sch}_{\overline{\mathbb{Q}}}$. Since the maps in the inverse system of Shimura varieties for $\boldsymbol{G}$ are all finite étale, the functoriality results established in [Stacks 2005–, Lemmas 02RD and 02R5] apply and we see that the groups $\mathscr{C}^n(L)$ for varying L assemble into a RIC functor

$$\mathscr{C}^n : \mathcal{P}(G, \Upsilon_G) \to \mathbb{Z}\text{-Mod}. \tag{3.13}$$

For $V \in \Upsilon_H$, we let $N(V)$ denote the free abelian group on the fundamental cycles indexed by $\pi_{0,V}$ from Section 2.3. For the same reasons, $N(V)$ for varying V assemble into a RIC functor

$$N : \mathcal{P}(H, \Upsilon_H) \to \mathbb{Z}\text{-Mod}. \tag{3.14}$$

Moreover, there exists a pushforward $\mathrm{cyc}_* : N \to \mathscr{C}^n$ in the sense of Definition 3.10 given by proper pushforward of algebraic cycles, again by [Stacks 2005–, Lemma 02R5].

Lemma 3.15. *The functors N and $\mathscr{C}^n$ are CoMack and Galois, and the pushforward cyc_* is Mackey.*

Proof. That N and $\mathscr{C}^n$ are Mackey follows by applying [Stacks 2005–, Lemma 02RG] to the diagram in [Shah 2024, Corollary 2.7.3], which is Cartesian in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$ by [ibid., Lemma 2.7.4]. That cyc_* is Mackey follows by similar considerations and our assumptions on Υ_H (which guarantee that the analog of the second commutative diagram in the proof of [Graham and Shah 2023, Proposition 4.12] in our context is also Cartesian). That N and $\mathscr{C}^n$ are cohomological follows by [Stacks 2005–, Lemma 02RH]. So it only remains to see that these functors are Galois. Since $\mathscr{C}^n$ specializes to N when $\boldsymbol{H} = \boldsymbol{G}$, it suffices to focus

on $\mathscr{C}^n$. So suppose that $L, K \in \Upsilon_G$ with $L \lhd K$. Since $\mathscr{C}^n$ is cohomological,

$$\mathrm{pr}^*_{L,K} : \mathscr{C}^n(K) \to \mathscr{C}^n(L)$$

is necessarily injective (as its postcomposition with $\mathrm{pr}_{L,K,*}$ is multiplication by $[K : L]$). It is also clear that the image of $\mathrm{pr}^*_{L,K}$ lands in the K/L-invariants of $\mathscr{C}^n(L)$. Finally, the surjectivity of $\mathrm{pr}^*_{L,K}$ onto $\mathscr{C}^n(L)^{K/L}$ follows by Galois descent for algebraic cycles with integral coefficients. Since we were unable to find a satisfactory reference for this fact, we provide a full proof below.

Let us momentarily write $X := \mathrm{Sh}_{\boldsymbol{G}}(L)$, $Y := \mathrm{Sh}_{\boldsymbol{G}}(K)$, $f : X \to Y$ for the degeneracy map induced by the inclusion $L \hookrightarrow K$, and $\Gamma := K/L$. Then f is a Galois covering with Galois group Γ in the sense of [Bosch et al. 1990, Section 6.2, Example B], i.e., the induced map $X \times \Gamma \to X \times_Y X$ given by $(x, \gamma) \mapsto (x, x\gamma)$ is an isomorphism (where Γ is viewed as a constant étale group scheme). Let Z_0 be a closed integral subscheme of X of codimension n, and Z be the scheme-theoretic union of the distinct Γ-conjugates of Z_0. Then Z is a closed and reduced subscheme of X and its ideal sheaf $\mathscr{I}$ is Γ-invariant under the induced action $\Gamma \times \mathbb{O}_X \to \mathbb{O}_X$. It therefore descends to a quasicoherent sheaf of ideals $\mathscr{J}$ for $\mathbb{O}_Y$ by [Bosch et al. 1990, Section 6.1, Theorem 4]. If W denotes the closed subscheme of Y corresponding to $\mathscr{J}$, we have $f^{-1}(W) = Z$ scheme-theoretically by construction. Then f^* sends the fundamental cycle $[W]$ to the fundamental cycle $[Z]$ by [Fulton 1998, Lemma 1.7.1].[6] Since cycles of the form $[Z]$ span $\mathscr{C}^n(L)^\Gamma$, the surjectivity of $f^* = \mathrm{pr}^*_{L,K}$ onto $\mathscr{C}^n(L)^\Gamma$ follows. □

Remark 3.16. The Galois descent for cycles with integral coefficients presumably follows from [Anschütz 2017]. See also [Poonen 2017, Section 7.6.2] for an argument similar to ours in the context of fields. For rational coefficients, Galois descent follows by [Shah 2024, Corollary 2.1.12].

One may also make similar considerations when $\mathscr{C}^n$ is replaced by Chow groups, or with the p-adic étale cohomology with coefficients as in [Graham and Shah 2023, Section 5.1]. The conclusions of Lemma 3.15 remain valid, except that the target functor is no longer necessarily Galois. However, it still makes sense to pose the analog of Question 1.4 in these settings. To make our results applicable to such situations, we will work with an arbitrary CoMack functor on (G, Υ_G) and an arbitrary Mackey pushforward to this CoMack functor. The situation of Question 1.4 can be recovered by specializing to $\mathscr{C}^n$ and cyc_*.

To this end, let R be any commutative ring with identity and $N_R := N \otimes_{\mathbb{Z}} R$ be the RIC functor on (H, Υ_H) obtained by base change, i.e., $N_R(V) := N(V) \otimes_{\mathbb{Z}} R$ for all $V \in \Upsilon_H$. For any $V \in \Upsilon_H$, denote by $[Y^\circ] \in N_R(V)$ the fundamental cycle associated with $Y^\circ = Y^\circ_V$ introduced in Definition 2.10.

Definition 3.17. Let $M_R : \mathcal{P}(G, \Upsilon_G) \to R\text{-Mod}$ be any RIC functor and $\iota_* : N_R \to M_R$ be any Mackey pushforward. For $g \in G$ and $K \in \Upsilon_G$, we define $y_K(g) \in M_R(K)$ to be the image of $[Y^\circ] \in N_R(H \cap gKg^{-1})$ under the composition

$$N_R(H \cap gKg^{-1}) \xrightarrow{\iota_*} M_R(gKg^{-1}) \xrightarrow{[g]_*} M_R(K). \tag{3.18}$$

[6] The scheme W can be shown to be reduced and irreducible, but we do not need this.

Similarly, we define $x_K(g) \in M_R(K)$ to be the image of the fundamental cycle $[\mathrm{Sh}_{\boldsymbol{H}}(H \cap gKg^{-1})] \in N_R(H \cap gKg^{-1})$ under (3.18).

Remark 3.19. Since the fundamental cycle of $\mathrm{Sh}_{\boldsymbol{H}}(V)$ is a formal sum of cycles indexed by $\pi_{0,V}$ for any $V \in \Upsilon_H$, it is not hard to see that each $x_K(g)$ is a formal sum of various $y_K(-)$. Indeed, say Z is a (geometric) connected component of $\mathrm{Sh}_{\boldsymbol{H}}(H \cap gKg^{-1})$ and $h \in H$ is such that Z is indexed by $[\nu(h)] \in \pi_{0,H\cap gKg^{-1}}$. Then the image of $[Z] \in N_R(H \cap gKg^{-1})$ under (3.18) is equal to $y_K(hg)$. Thus, $x_K(g)$ equals the sum of $y_K(hg)$ as h varies over representatives for $\pi_{0,H\cap gKg^{-1}}$.

As already noted, we consider a general ring R and a CoMack functor M_R in order to capture the various pushforward constructions and cohomology theories one may consider. For instance, one may take $R = \mathbb{Z}_p$ and

$$M_{\mathbb{Z}_p}(L) := \mathrm{H}^{2n}_{\text{ét}}(\mathrm{Sh}_{\boldsymbol{G}}(L), \mathbb{Z}_p(n)),$$

where the right-hand side denotes Jannsen's continuous étale cohomology [1988]. Indeed, $N_{\mathbb{Z}_p}(V)$ can be identified with $\mathrm{H}^0_{\text{ét}}(\mathrm{Sh}_{\boldsymbol{H}}(V), \mathbb{Z}_p)$ and the pushforward is obtained via the Gysin triangle in Ekedahl's "derived" category of constructible $\mathbb{Z}_p$-sheaves as in [Graham and Shah 2023, Appendix A]. The relevant properties of this construction can be established as in [ibid., Section 4], and the failure of (SV5) can be handled by [Shah 2024, Corollary 2.7.3]. Note however that $M_{\mathbb{Z}_p}$ in this case is *not* necessarily Galois.

Remark 3.20. The inductive limit $\varinjlim_L M_R(L)$ for $L \in \Upsilon_G$ over restriction maps is naturally a smooth G-representation, where an element $g \in G$ acts by $[g]^*$. The Mackey axiom for M_R implies that

$$[K\sigma K]_* \cdot x = \sum_{\delta \in K \backslash K\sigma K} \delta^{-1} \cdot x$$

for any $x \in M_R(K)$, where the equality is viewed in the inductive limit. This is the analog of (1.2) from the introduction. See [Shah 2024, Corollary 2.4.3] or [Graham and Shah 2023, Lemma 2.7(a)] for a justification.

3.4. ***A comparison.*** Assume for this subsection only that $R = \mathbb{Z}$, $M_R = \mathscr{C}^n$ and $\iota_* = \mathrm{cyc}_*$. Then $y_K(g) \in \mathscr{C}^n(K)$ (Definition 3.17) is not necessarily equal to $z_K(g)$ (Definition 3.1), but the two are very closely related.

Lemma 3.21. *There exists a unique positive integer $d_{g,K}$ such that $y_K(g) = d_{g,K} \cdot z_K(g)$. Moreover, if the morphism $\iota_{H\cap gKg^{-1}, gKg^{-1}}$ is a closed immersion, $d_{hg,K} = 1$ for all $h \in H$.*

Proof. The first part follows by [Fulton 1998, Section 1.4]. More precisely, $d_{g,K}$ is the degree of the field extension of the function fields that corresponds to the dominant morphism of integral schemes

$$Y^\circ_{H\cap gKg^{-1}} \to \iota_{g,K}(Y^\circ_{H\cap gKg^{-1}})^{\mathrm{red}}, \tag{3.22}$$

where $\iota_{g,K}$ denotes $\iota_{H\cap gKg^{-1}, gKg^{-1}}$ for simplicity and the right-hand side of (3.22) denotes the reduced induced closed subscheme of $\mathrm{Sh}_{\boldsymbol{G}}(gKg^{-1})$ on the image of Y° under $\iota_{g,K}$. It follows that $d_{g,K} = 1$ if

$\iota_{g,K}$ is a closed immersion. In that case, $\iota_{hg,K}$ is a closed immersion for any $h \in H$ as well, since we have a commutative diagram

$$\begin{array}{ccc} \mathrm{Sh}_{\boldsymbol{H}}(H \cap hgK(hg)^{-1}) & \xrightarrow{\iota} & \mathrm{Sh}_{\boldsymbol{G}}(hgK(hg)^{-1}) \\ {\scriptstyle [h]}\downarrow & & \downarrow{\scriptstyle [h]} \\ \mathrm{Sh}_{\boldsymbol{H}}(H \cap gKg^{-1}) & \xrightarrow{\iota} & \mathrm{Sh}_{\boldsymbol{G}}(gKg^{-1}) \end{array}$$

whose vertical arrows are isomorphisms. □

Remark 3.23. Since the isomorphism class of the map (3.22) depends only on the class of g in $\boldsymbol{H}(\mathbb{Q})_+\backslash G/K$, we have $d_{g,K} = d_{hg\kappa,K}$ for all $h \in \boldsymbol{H}(\mathbb{Q})_+$ and $\kappa \in K$. We also observe that the cycle $x_K(g)$ from Definition 3.17 in this case is closely related to, but not exactly the same as, the "natural cycle" defined in [Kudla 1997, Section 2], since the former is a sum of various $y_K(-)$ (see Remark 3.19), whereas the latter is a sum of various $z_K(-)$.

4. The formula

In this section, we derive our formula for Hecke action on the classes introduced in Definition 3.17 and the special cycles of Definition 3.1. We also highlight two scenarios where the resulting expression simplifies and agrees with (1.5).

4.1. ***The computation.*** The notation, conventions and assumptions introduced in Sections 2 and 3 are maintained. In particular, the derived group of $\boldsymbol{H}$ is assumed to be simply connected and the existence of a collection Υ_H in (3.9) whose elements have the same intersection with $\boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q})$ is also assumed. Recall also that all our Shimura varieties are viewed in $\mathbf{Sch}_{\overline{\mathbb{Q}}}$.

As in Section 3.3, we fix for all of this section a commutative ring R with identity, a CoMack functor $M_R : \mathcal{P}(G, \Upsilon_G) \to R\text{-Mod}$, and a Mackey pushforward $\iota_* : N_R \to M_R$, where N_R denotes the base change of N (see (3.14)) to R. We also fix a compact–open subgroup $K \in \Upsilon_G$ and two elements $g, \sigma \in G$. Our main goal in this subsection is to compute an expression for

$$[K\sigma K]_* \cdot y_K(g) \in M_R(K)$$

in terms of $y_K(-)$ from Definition 3.17. Here, $[K\sigma K]_*$ is as in Note 3.12.

Lemma 4.1. *Let $K' \in \Upsilon_G$ denote gKg^{-1}. Then $[K\sigma K]_* \cdot y_K(g) = [K'g\sigma K]_* \cdot y_{K'}(1)$.*

Proof. This follows by unraveling the definitions. □

Lemma 4.1 reduces our problem to computing $[K'\varsigma K]_* \cdot y_{K'}(1) \in M_R(K)$ for arbitrary $K' \in \Upsilon_G$ and $\varsigma \in G$, which we also fix for all of this subsection.[7] For this purpose, we introduce the following notation.

Notation 4.2. We let U denote the intersection $H \cap K' \in \Upsilon_H$ and I denote the finite double coset space $U\backslash K'\varsigma K/K$. For each $i \in I$, we let $\varsigma_i \in G$ denote a representative element for i.

[7] Later on, we will specialize to $K' = gKg^{-1}$ and $\varsigma = g\sigma$.

Lemma 4.3. *We have* $[K'\varsigma K]_* \cdot y_{K'}(1) = \sum_{i\in I}[U\varsigma_i K]_*([Y_U^\circ])$, *where* $[U\varsigma_i K]_*$ *denotes the mixed Hecke correspondence.*

Proof. It is elementary to deduce from the Mackey axiom for ι_* that $[K'\varsigma K]_* \circ \iota_{U,K',*} = \sum_{i\in I}[U\varsigma_i K]_*$ as R-linear maps $N_R(U) \to M_R(K)$ (see [Shah 2024, Lemma 2.5.7]). The claim now follows since $y_{K'}(1) = \iota_{U,K',*}([Y_U^\circ])$ by definition. □

Lemma 4.3 in turn reduces our problem to computing the effect of certain mixed Hecke correspondences on $[Y_U^\circ]$. It will be convenient to make the following notational convention.

Notation 4.4. For any subgroup X of H and any element $\tau \in G$, we denote by $X_{\tau,K}$ the intersection $X \cap \tau K \tau^{-1}$. We will write X_τ for $X_{\tau,K}$ if K is fixed in context.

Unraveling the definition of $[U\varsigma_i K]_*$, we obtain the following commutative diagram for each $i \in I$:

$$\begin{array}{ccccccc} & & & & N_R(U_{\varsigma_i}) & & \\ & & & \overset{\mathrm{pr}^*}{\nearrow} & \downarrow{\scriptstyle \mathrm{pr}_*} & \searrow^{\iota_*} & \\ N_R(U) & & & & N_R(H_{\varsigma_i}) & \xrightarrow{\iota_*} & M_R(\varsigma_i K \varsigma_i^{-1}) \xrightarrow{[\varsigma_i]_*} M_R(K) \\ & & & \searrow & & & \nearrow \\ & & & & [U\varsigma_i K]_* & & \end{array} \tag{4.5}$$

We wish to compute the effect of the individual maps in diagram (4.5). To this end, recall (Section 2.3) that, for $V \in \Upsilon_H$, $\pi_{0,V}$ is an abelian group that parametrizes the connected components of $\mathrm{Sh}_{\boldsymbol{H}}(V)$, and, for $V' \in \Upsilon_H$ contained in V, $\varphi_{V',V} : \pi_{0,V'} \to \pi_{0,V}$ denotes the quotient morphism.

Notation 4.6. For any $\tau \in G$, we let A_τ denote a set of representatives in H for the kernel of the homomorphism $\varphi_{U_\tau,U} : \pi_{0,U_\tau} \to \pi_{0,U}$.

Here the representatives are picked under the composition $H \xrightarrow{\nu} \boldsymbol{T}(\mathbb{A}_f) \to \pi_{0,U_\tau}$, which is surjective by [Milne 2005, Lemma 5.21] and simple connectedness of $\boldsymbol{H}^{\mathrm{der}}$.

Lemma 4.7. *For any* $\tau \in G$,

$$\mathrm{pr}^*_{U_\tau,U}([Y_U^\circ]) = \sum_{h\in A_\tau}[h]_*([Y^\circ_{hU_\tau h^{-1}}]) \in N_R(U_\tau),$$

where $[h]$ *on the right-hand side above denotes the morphism* $(hU_\tau h^{-1} \xrightarrow{h} U_\tau) \in \mathcal{P}(H, \Upsilon_H)$.

Proof. Let $\eta \in H$. Recall (Section 2.3) that the twisting isomorphism $[h] : \mathrm{Sh}_{\boldsymbol{H}}(hU_\tau h^{-1}) \xrightarrow{\sim} \mathrm{Sh}_{\boldsymbol{H}}(U_\tau)$ sends the component indexed by the class of η in $\pi_{0,hU_\tau h^{-1}} = \pi_{0,U_\tau}$ to the component of $\mathrm{Sh}_{\boldsymbol{H}}(U_\tau)$ indexed by the class of $h\eta$ (or that of ηh) in π_{0,U_τ}. So the right-hand side of the equality above is just the formal sum of the fundamental cycles corresponding to the connected components of $\mathrm{Sh}_{\boldsymbol{H}}(U_\tau)$ indexed by A_τ. We argue that this also equals the left-hand side. Since the morphism $\mathrm{pr}_{U_\tau,U} : \mathrm{Sh}_{\boldsymbol{H}}(U_\tau) \to \mathrm{Sh}_{\boldsymbol{H}}(U)$ is étale, so is its pullback $\mathrm{pr}^{-1}_{U_\tau,U}(Y_U^\circ) \to Y_U^\circ$ along $Y_U^\circ \to \mathrm{Sh}_{\boldsymbol{H}}(U)$. This implies that $\mathrm{pr}^{-1}_{U_\tau,U}(Y_U^\circ)$ is reduced [Stacks 2005–, Tag 03PC]. So the scheme $\mathrm{pr}^{-1}_{U_\tau,U}(Y_U^\circ)$ is equal to the disjoint union of the components of $\mathrm{Sh}_{\boldsymbol{H}}(U_\tau)$ indexed by A_τ. The claim now follows by definition of flat pullback [Fulton 1998, Sections 1.7 and 1.5]. □

Next we need a result for degrees of maps between connected components of $\mathrm{Sh}_{\boldsymbol{H}}$.

Lemma 4.8. *Let $V_1, V_2 \in \Upsilon_H$ with $V_2 \subset V_1$ and $\eta \in \boldsymbol{T}(\mathbb{A}_f)$. For $j = 1, 2$, let Y_{η,V_j} denote the component of $\mathrm{Sh}_{\boldsymbol{H}}(V_j)$ indexed by the class of η in π_{0,V_j} and let e denote the cardinality of the kernel of $\varphi_{V_2,V_1} : \pi_{0,V_2} \to \pi_{0,V_1}$. Then the natural map $Y_{\eta,V_2} \to Y_{\eta,V_1}$ is finite étale of degree $[V_1 : V_2]/e$.*

Proof. Let W be a compact–open subgroup of V_2 such that $W \triangleleft V_1$. To ease notation, we let V denote an element of $\{V_1, V_2\}$. By enlarging W if necessary, we may assume that $W \cap \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q}) = V \cap \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q})$, so that the map $\mathrm{Sh}_{\boldsymbol{H}}(W) \to \mathrm{Sh}_{\boldsymbol{H}}(V)$ is a Galois cover with Galois group V/W. Since $Y_{\eta,V} \hookrightarrow \mathrm{Sh}_{\boldsymbol{H}}(V)$ is an open immersion,

$$Z := \mathrm{pr}_{W,V}^{-1}(Y_{\eta,V}) \to Y_{\eta,V}$$

is a Galois cover of degree $[V:W]$ as well. Let e_V denote the cardinality of $\ker(\varphi_{W,V})$ and let $\upsilon_1, \ldots \upsilon_{e_V} \in V$ be representatives of $\ker(\varphi_{W,V})$. Then Z is the union of components of $\mathrm{Sh}_{\boldsymbol{H}}(W)$ indexed by the classes $[\eta\nu(\upsilon_k)] \in \pi_{0,W}$ for $k = 1, \ldots, e_V$. Since $[\upsilon_k] : \mathrm{Sh}_{\boldsymbol{H}}(W) \to \mathrm{Sh}_{\boldsymbol{H}}(W)$ are automorphisms that act transitively on the connected components contained in Z, we see that the degree of $Y_{\eta\nu(\upsilon_k),W} \to Y_{\eta,V}$ is independent of k and therefore equal to $[V : W]/e_V$. Since $e_{V_1} = e_{V_2} \cdot e$ and $[V_1 : W] = [V_1 : V_2] \cdot [V_2 : W]$, we conclude that $Y_{\eta,V_2} \to Y_{\eta,V_1}$ is finite étale with degree as claimed. □

Remark 4.9. In the proof, we do not require $W \in \Upsilon_H$.

Corollary 4.10. *For any $\tau \in G$ and $\eta \in \pi_{0,U_\tau}$,*

$$\mathrm{pr}_{U_\tau,H_\tau,*}([Y_{\eta,U_\tau}]) = e_\tau^{-1} \deg[U\tau K]_* \cdot [Y_{\eta,H_\tau}],$$

where Y_{η,U_τ} denotes the component of $\mathrm{Sh}_{\boldsymbol{H}}(U_\tau)$ indexed by η, Y_{η,H_τ} denotes the component of $\mathrm{Sh}_{\boldsymbol{H}}(H_\tau)$ indexed by $\varphi_{U_\tau,H_\tau}(\eta) \in \pi_{0,H_\tau}$ and e_τ denotes the cardinality of $\ker(\varphi_{U_\tau,H_\tau})$.

Proof. This follows by Lemma 4.8 and the definition of proper pushforward [Fulton 1998, Section 1.4]. □

We now apply these results to the maps in diagram (4.5) by specializing to $\tau = \varsigma_i$. Let us first fix some additional notation.

Notation 4.11. For each $i \in I$, we will let U_i, H_i, A_i denote U_{ς_i}, H_{ς_i}, A_{ς_i}, respectively. We let $B_i \subset A_i$ denote a set of representatives for the image of A_i under the quotient morphism $\varphi_{U_i,H_i} : \pi_{0,U_i} \to \pi_{0,H_i}$. Here we are identifying $A_i = A_{\varsigma_i}$ with the kernel of $\varphi_{U_i,U} : \pi_{0,U_i} \to \pi_{0,U}$. We let $e_i = e_{\varsigma_i}$ denote the cardinality of $\ker(\varphi_{U_i,H_i})$ and c_i denote the cardinality of $\ker(\varphi_{U_i,H_i}) \cap \ker(\varphi_{U_i,U})$.

If we identify A_i with $\ker(\varphi_{U_i,U})$ and B_i with $\varphi_{U_i,H_i}(A_i)$, then c_i is the cardinality of the kernel of the surjective homomorphism $A_i \to B_i$, and so $c_i = |A_i|/|B_i|$. Let us also emphasize that $H_i = H_{\varsigma_i} = H \cap \varsigma_i K \varsigma_i^{-1}$ is a *compact–open* subgroup of H (despite the notational similarity with H). To make the statement of our main theorem more self-contained, we recall most of the necessary notation.

Theorem 4.12 (explicit descent). *For $K', K \in \Upsilon_G$ and $\varsigma \in G$, let $U = H \cap K'$ and $I = U\backslash K'\varsigma K/K$. For each $i \in I$, let $\varsigma_i \in G$ denote a representative for the class i and write $U_i = U \cap \varsigma_i K \varsigma_i^{-1}$, $H_i = H \cap \varsigma_i K \varsigma_i^{-1}$ and $\deg[U\varsigma_i K]_* = [H_i : U_i]$. For each i, let $A_i \subset H$ denote a set of representatives for $\ker(\varphi_{U_i,U})$, let*

$B_i \subset A_i$ denote a set of representatives for $\varphi_{U_i,H_i}(\ker(\varphi_{U_i,U}))$, and set $c_i = |A_i|/|B_i|$, $e_i = |\ker(\varphi_{U_i,H_i})|$. Then

$$[K'\varsigma K]_* \cdot y_{K'}(1) = \sum_{i\in I}\sum_{h\in B_i} c_i e_i^{-1} \deg[U\varsigma_i K]_* \cdot y_K(h\varsigma_i)$$

as elements of $M_R(K)$.

Proof. By Lemma 4.3, it suffices to compute $[U\varsigma_i K]_*([Y_U^\circ])$ for each $i \in I$. Note that the integer c_i is the number of connected components of $\mathrm{Sh}_{\boldsymbol{H}}(U_i)$ contained in $\mathrm{pr}_{U_i,U}^{-1}(Y_U^\circ)$ that collapse into a single component of $\mathrm{Sh}_{\boldsymbol{H}}(H_i)$ under pr_{U_i,H_i}. Invoking Lemma 4.7 and Corollary 4.10, we see that

$$\mathrm{pr}_{U_i,H_i,*} \circ \mathrm{pr}_{U_i,U}^*([Y_U^\circ]) = \sum_{h\in B_i} (c_i e_i^{-1} \deg[U\varsigma_i K]_*) \cdot [h]_*([Y^\circ_{H_{h\varsigma_i}}]), \tag{4.13}$$

where $[h]$ above is the morphism $(H_{h\varsigma_i} \xrightarrow{h} H_{\varsigma_i}) \in \mathcal{P}(H, \Upsilon_H)$. Now, for each $i \in I$ and $h \in B_i$, we have a commutative diagram

$$\begin{array}{ccccc} N_R(H_{h\varsigma_i}) & \xrightarrow{\iota_*} & M_R(h\varsigma_i K \varsigma_i^{-1} h^{-1}) & & \\ \wr\downarrow{\scriptstyle [h]_*} & & \wr\downarrow{\scriptstyle [h]_*} & \searrow^{[h\varsigma_i]_*} & \\ N_R(H_{\varsigma_i}) & \xrightarrow{\iota_*} & M_R(\varsigma_i K \varsigma_i^{-1}) & \xrightarrow{[\varsigma_i]_*} & M_R(K) \end{array} \tag{4.14}$$

Let us momentarily write $K_i = \varsigma_i K \varsigma_i^{-1}$ to simplify notation. Using (4.13) and the commutativity of diagrams (4.5) and (4.14), we see that

$$\begin{aligned} [U\varsigma_i K]_*([Y_U^\circ]) &= [\varsigma_i]_{K_i,K,*} \circ \iota_{U_i,K_i,*} \circ \mathrm{pr}_{U_i,U}^*([Y_U^\circ]) \\ &= [\varsigma_i]_{K_i,K,*} \circ \iota_{H_i,K_i,*} \circ \mathrm{pr}_{U_i,H_i,*} \circ \mathrm{pr}_{U_i,U}^*([Y_U^\circ]) \\ &= [\varsigma_i]_{K_i,K,*} \circ \iota_{H_i,K_i,*} \Big(\sum_{h\in B_i} (c_i e_i^{-1} \deg[U\varsigma_i K]_*) \cdot [h]_{H_{h\varsigma_i},H_i,*}([Y^\circ_{H_{h\varsigma_i}}]) \Big) \\ &= \sum_{h\in B_i} (c_i e_i^{-1} \deg[U\varsigma_i K]_*) \cdot [h\varsigma_i]_{hK_ih^{-1},K,*} \circ \iota_{H_{h\varsigma_i},hK_ih^{-1},*}([Y^\circ_{H_{h\varsigma_i}}]) \\ &= \sum_{h\in B_i} c_i e_i^{-1} \deg[U\varsigma_i K]_* \cdot y_K(h\varsigma_i). \end{aligned}$$

□

In the formula above, we may take the inner sum to be over A_i (instead of B_i) after removing c_i from the expression, since $y_K(h\varsigma_i) = y_K(h'\varsigma_i)$ for $h, h' \in A_i$ if the classes of h, h' are equal in π_{0,H_i}.

Theorem 4.12 bis. *With notation as above,*

$$[K'\varsigma K]_* \cdot y_{K'}(1) = \sum_{i\in I}\sum_{h\in A_i} e_i^{-1} \deg[U\varsigma_i K]_* \cdot y_K(h\varsigma_i).$$

One reason for preferring the first version is that a simplification occurs when $\nu(U)$ contains $\nu(H_i)$ for all $i \in I$. We record it as a lemma for ease of reference in Section 5.

Lemma 4.15. *If $\nu(U)$ contains $\nu(H_i)$ for some $i \in I$, we have $c_i = e_i$. If, moreover, $\nu(U)$ equals $\nu(H_i)$, then B_i is a singleton.*

Remark 4.16. In the setting of Corollary 4.19, suppose that $K = K_\ell K^\ell$ and $\sigma \in \boldsymbol{G}(\mathbb{Q}_\ell)$ for a rational prime ℓ at which $\boldsymbol{T}$ is unramified. Write $U = U_\ell U^\ell$, and suppose that $\nu(U_\ell)$ is the maximal compact–open subgroup of $\boldsymbol{T}(\mathbb{Q}_\ell)$. Then $\nu(U) \supseteq \nu(H_i)$ for all $i \in I$. So, in this case, the coefficients in the expression of Theorem 4.12 only involve mixed degrees. See [Shah 2024, Section 5], where several techniques were developed to aid their computation and [ibid., Part II] for several concrete examples.

If we replace $y_{K'}(1)$ with $x_{K'}(1)$, the formula is much simpler and does not require as much work.

Proposition 4.17.
$$[K'\varsigma K]_* \cdot x_{K'}(1) = \sum_{i\in I} \deg[U\varsigma_i K]_* \cdot x_K(\varsigma_i).$$

Proof. Let $N_{\mathrm{triv},R} : \mathcal{P}(H, \Upsilon_H) \to R\text{-Mod}$ denote the functor associated with the trivial representation of H (see Remark 3.7), meaning $N_{\mathrm{triv},R}(V) = R$ for all V. Since fundamental cycles of $\mathrm{Sh}_{\boldsymbol{H}}(V)$ for varying $V \in \Upsilon_H$ map to themselves under pullbacks and to multiples by degree under pushforwards along degeneracy maps, they realize the trivial functor $N_{\mathrm{triv},R}$. The class $x_K(g)$ can then be defined as the image of $1_R \in R = N_{\mathrm{triv},R}(H_g)$ under the analogous twisting map in Definition 3.17. The claim easily follows by the obvious analog of Lemma 4.3 and the diagram (4.5) with N_R replaced by $N_{\mathrm{triv},R}$. □

Remark 4.18. Proposition 4.17 holds without the assumption that $\boldsymbol{H}^{\mathrm{der}}$ be simply connected, since the definition of $x_K(-)$ etc. does not rely on a description of the connected components of $\mathrm{Sh}_{\boldsymbol{H}}$. One may also drop (SV3) for $(\boldsymbol{H}, X_{\boldsymbol{H}})$ in light of [Graham and Shah 2023, Appendix A], which extends the formalism of Shimura varieties in the absence of (SV3), *assuming* that $(\boldsymbol{H}, X_{\boldsymbol{H}})$ embeds into a datum which does satisfy (SV3). In our case, this latter datum is $(\boldsymbol{G}, X_{\boldsymbol{G}})$.

If we specialize to $K' = gKg^{-1}$ and $\varsigma = g\sigma$ in Theorem 4.12 and invoke Lemma 4.1, we obtain the following.

Corollary 4.19. *For $K \in \Upsilon_G$ and $g, \sigma \in G$, let $U = H \cap gKg^{-1}$ and $I = U\backslash gK\sigma K/K$. For each $i \in I$, let $\varsigma_i \in G$ denote a representative for i and let $U_i = U \cap \varsigma_i K \varsigma_i^{-1}$, $H_i = H \cap \varsigma_i K \varsigma_i^{-1}$ and $\deg[U\varsigma_i K]_* = [H_i : U_i]$. For each i, let $A_i \subset H$ denote a set of representatives for $\ker(\varphi_{U_i,U})$, $B_i \subset A_i$ denote a set of representatives for $\varphi_{U_i,H_i}(\ker(\varphi_{U_i,U}))$, and set $c_i = |A_i|/|B_i|$, $e_i = |\ker(\varphi_{U_i,H_i})|$. Then*
$$[K\sigma K]_* \cdot y_K(g) = \sum_{i\in I} \sum_{h\in B_i} c_i e_i^{-1} \deg[U\varsigma_i K]_* \cdot y_K(h\varsigma_i).$$

We can finally answer Question 1.4 now. Recall from Lemma 3.21 that, when $M_R = \mathscr{C}^n$ (with $R = \mathbb{Z}$ and $\iota_* = \mathrm{cyc}_*$), there exists for each $\tau \in G$ a unique positive integer $d_{\tau,K}$ such that $y_K(\tau) = d_{\tau,K} \cdot z_K(\tau)$ as elements of $\mathscr{C}^n(K)$.

Corollary 4.20. *With notation as in Corollary 4.19, we have*
$$[K\sigma K]_* \cdot z_K(g) = \frac{1}{d_{g,K}} \sum_{i\in I} \sum_{h\in B_i} c_i d_{h\varsigma_i,K} e_i^{-1} \deg[U\varsigma_i K]_* \cdot z_K(h\varsigma_i). \tag{4.21}$$

In particular, $[K\sigma K]_* \cdot z_K(g)$ lies in the submodule of $\mathscr{C}^n(K)$ spanned by irreducible special cycles.

Proof. The expression follows from Corollary 4.19 after specializing to $R = \mathbb{Z}$, $M_R = \mathscr{C}^n$ and $\iota_* = \mathrm{cyc}_*$ (which we can do by Lemma 3.15). Since $[K\sigma K]_* \cdot z_K(g)$ belongs to both $\mathscr{C}^n(K)$ (by definition) and $\mathscr{Z}_K \otimes_{\mathbb{Z}} \mathbb{Q}$ by (4.21), it must lie inside $\mathscr{Z}_K$ (where $\mathscr{Z}_K$ is as in Definition 3.1). □

Remark 4.22. The coefficients of individual summands in (4.21) are not a priori integers, since we do not know if $d_{g,K}$ divides the product $c_i d_{h\varsigma_i,K} e_i^{-1} \deg[U\varsigma_i K]_*$ for all $i \in I$, $h \in B_i$. The point is that the coefficients in our expression can be made integral (if not already) by collecting together the coefficients of all $z_K(h\varsigma_i)$ that represent the same irreducible cycle in $\mathrm{Sh}_{\boldsymbol{G}}(K)$.

4.2. ***Examples.*** Below, we record two simple instances in which the right-hand side of (4.21) matches that of (1.5). The notation above is maintained.

Example 4.23. Suppose $\boldsymbol{H}$ is a torus. In this case, we are asking for a Hecke action on special points on $\mathrm{Sh}_{\boldsymbol{G}}(K)$. We have $\boldsymbol{H} = \boldsymbol{T}$ and ν is the identity map. Our assumption on Υ_H implies that A_i identifies with U/U_i and $e_i = [H_i : U_i] = \deg[U\varsigma_i K]_*$. Choose the representatives in A_i to lie in U. Moreover, $d_{\tau,K} = 1$ for all $\tau \in G$, since $\mathrm{Sh}_{\boldsymbol{H}}(H_\tau)$ is a finite set of reduced points over $\overline{\mathbb{Q}}$. Thus,

$$[K\sigma K]_* \cdot z_K(g) = \sum_{i\in I} \sum_{h\in B_i} c_i \cdot z_K(h\varsigma_i) = \sum_{i\in I} \sum_{h\in A_i} z_K(h\varsigma_i).$$

Now, for each $i \in I$, we have $U\varsigma_i K/K = \bigsqcup_{h\in A_i} h\varsigma_i K$ and therefore $gK\sigma K/K = \bigsqcup_{i\in I, h\in A_i} h\varsigma_i K$. So

$$[K\sigma K]_* \cdot z_K(g) = \sum_{\gamma \in K\sigma K/K} z_K(g\gamma),$$

which agrees with (1.5). This is of course what one gets by directly computing the result of a correspondence on a general zero-cycle on $\mathrm{Sh}_{\boldsymbol{G}}(K)$ as in (1.1).

Example 4.24. Suppose that $\boldsymbol{H} = \boldsymbol{G}$. In this case, we are asking for a Hecke action on connected components of $\mathrm{Sh}_{\boldsymbol{G}}(K)$ itself. We have $U = gKg^{-1}$, so $I = \{1\}$ is a singleton and we may take $\varsigma_1 = g\sigma$. By definition, we have $\deg[U\varsigma_1 K]_* = [\varsigma_1 K \varsigma_1^{-1} : U_{\varsigma_1}]$, which equals $[K : \varsigma_1^{-1} U_{\varsigma_1} \varsigma_1]$. Since $\varsigma_1^{-1} U_{\varsigma_1} \varsigma_1 = \sigma^{-1} K\sigma \cap K$, we see that

$$\deg[U\varsigma_1 K]_* = |K\sigma^{-1}K/K|.$$

Since $\nu(U) = \nu(K) = \nu(H_{\varsigma_1})$, we have $c_1 = e_1$ and $B_1 = \{1\}$ is a singleton by Lemma 4.15. Clearly, $d_{\tau,K} = 1$ for all $\tau \in G$. Therefore,

$$[K\sigma K]_* \cdot z_K(g) = |K\sigma^{-1}K/K| \cdot z_K(g\sigma).$$

Now $|K\sigma^{-1}K/K| = |K\sigma K/K|$ by unimodularity of G [Renard 2010, p. 58] and $z_K(g\gamma\sigma) = z_K(g\sigma)$ for any $\gamma \in K$ since $\pi_{0,K} = \pi_0(\mathrm{Sh}_{\boldsymbol{G}}(K))$ is abelian. So

$$[K\sigma K]_* \cdot z_K(g) = \sum_{\gamma \in K\sigma K/K} z_K(g\gamma),$$

which again agrees with (1.5).

5. Counterexamples

In this section, we furnish two (families of) examples where (1.5) fails to hold.

5.1. ***Counting cycles.*** Since irreducible special cycles form a $\mathbb{Z}$-module basis of $\mathscr{Z}_K$ by definition, we see that (1.5) holds only if

$$|K\sigma K/K| \stackrel{?}{=} \frac{1}{d_{g,K}} \sum_{i\in I} \sum_{h\in B_i} c_i d_{h\varsigma_i,K} e_i^{-1} \deg[U\varsigma_i K]_*.$$

Indeed, the right-hand side above is the sum of the coefficients in the right-hand side of (4.21). The strategy is therefore to compute both these integers explicitly and show they are not equal. We will pick K and σ so that the various $d_{-,K}$ appearing above are forced to be 1. For the computation of mixed degrees, we rely on the calculations done in [Shah 2024], though alternatives are also provided for the reader and the computations are mostly self-contained. In both our examples, we will have $g = 1_G$ (so $U = H \cap K$) and the letter g will be used for other purposes. For this reason, we denote the representatives ς_i for $i \in U\backslash K\sigma K/K$ by σ_i. In this notation, we are interested in checking if

$$|K\sigma K/K| \stackrel{?}{=} \frac{1}{d_{1_G,K}} \sum_{i\in I} \sum_{h\in B_i} c_i d_{h\sigma_i,K} e_i^{-1} \deg[U\sigma_i K]_*. \tag{5.1}$$

Throughout Sections 5.2 and 5.3, ℓ denotes a rational prime and $\mathbb{A}_f^\ell = \mathbb{A}_f/\mathbb{Q}_\ell$ denotes the group of finite rational adeles away from ℓ.

Remark 5.2. The counterexamples below also work if $z_K(1)$ is replaced by $x_K(1)$.

5.2. ***Symplectic groups.*** We let

$$\boldsymbol{H} = \mathrm{GL}_2 \times_{\mathbb{G}_m} \mathrm{GL}_2, \quad \boldsymbol{G} = \mathrm{GSp}_4,$$

which we consider as reductive group schemes over $\mathbb{Z}$. Here, we define $\boldsymbol{G}$ with respect to the standard symplectic matrix which has the identity matrix in the top right 2×2 block, negative identity in the bottom left 2×2 block, and zeros elsewhere. The embedding of $\boldsymbol{H}$ in $\boldsymbol{G}$ is as in [Shah 2024, Section 9.3], which gives a morphism of Shimura data (see [Milne 2005, Section 6]). Moreover, both data satisfy (SV1)–(SV6) of [Milne 2005] by the discussion in [ibid., Section 6]. In addition, the derived group of $\boldsymbol{H}$ is $\mathrm{SL}_2 \times \mathrm{SL}_2$, which is simply connected. Finally, our assumption on the existence of Υ_H is also satisfied, since the datum for $\boldsymbol{H}$ satisfies (SV5). Thus, Corollary 4.20 applies to this embedding of Shimura data. Let

$$w := \mathrm{diag}(1, -1, 1, -1) \in \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q}).$$

Over $\mathbb{Q}$, the centralizer of w in $\boldsymbol{G}$ equals $\boldsymbol{H}$. We have $\boldsymbol{T} = \boldsymbol{H}/\boldsymbol{H}^{\mathrm{der}} = \mathbb{G}_m$, and $\nu : \boldsymbol{H} \to \boldsymbol{T}$ is the map given by taking the common determinant of the two components. As above, we write $G = \boldsymbol{G}(\mathbb{A}_f)$ and $H = \boldsymbol{H}(\mathbb{A}_f)$. For $N \geq 1$, we let $K(N)$ denote the principal congruence subgroup of G at level N. Then

$K(N)$ is neat if $N \geq 3$. Now say K is a compact–open subgroup contained in $K(N)$ for some $N \geq 3$. Since $w \in K(1) = \boldsymbol{G}(\widehat{\mathbb{Z}})$ and $K(N) \trianglelefteq K(1)$, we have $wKw^{-1} \subset wK(N)w^{-1} = K(N)$. Thus,

$$\iota_{U,K} : \mathrm{Sh}_{\boldsymbol{H}}(U) \to \mathrm{Sh}_{\boldsymbol{G}}(K)$$

is a closed immersion by Lemma 2.7. We fix such a K from now on and write as above $U := H \cap K$, etc. We let ℓ be a rational prime such that K is unramified at ℓ, i.e., $K = K_\ell K^\ell$, where $K_\ell = \mathrm{GSp}_4(\mathbb{Z}_\ell)$ and $K^\ell \subset \mathrm{GSp}_4(\mathbb{A}_f^\ell)$ is a compact–open subgroup. Then $U = U_\ell U^\ell$ with $U_\ell = \boldsymbol{H}(\mathbb{Z}_\ell)$. Set

$$\tau_\ell := \begin{pmatrix} 1 & & & 1/\ell \\ & 1 & 1/\ell & \\ & & 1 & \\ & & & 1 \end{pmatrix} \in \boldsymbol{G}(\mathbb{Q}_\ell)$$

and denote by τ the image of τ_ℓ under the embedding $\boldsymbol{G}(\mathbb{Q}_\ell) \hookrightarrow \boldsymbol{G}(\mathbb{A}_f)$, so $\tau = \tau^\ell \tau_\ell$, where $\tau^\ell \in \boldsymbol{G}(\mathbb{A}_f^\ell)$ is the identity. The convention introduced in Notation 4.4 is maintained.

Lemma 5.3. *If $\ell = 2$, the morphism $\iota : \mathrm{Sh}_{\boldsymbol{H}}(H_\tau) \to \mathrm{Sh}_{\boldsymbol{G}}(\tau K \tau^{-1})$ is a closed immersion.*

Proof. Let $L = L_\ell L^\ell$ be a neat compact–open subgroup of G that contains both K and wKw (for example, take $L = K(N)$). Clearly, $L_\ell = K_\ell$ by maximality of K_ℓ. Write $w = w_\ell w^\ell$ and set $L' := \tau L \tau^{-1}$. Then L' is neat and contains $\tau K \tau^{-1}$. By Lemma 2.7, $\iota = \iota_{H_\tau, \tau K \tau^{-1}}$ is a closed immersion whenever L' contains $w\tau K \tau^{-1} w$. Note that

$$L' = L^\ell \cdot L'_\ell, \quad \text{where } L'_\ell = \tau_\ell K_\ell \tau_\ell^{-1}.$$

Now L^ℓ contains $w^\ell \tau^\ell K^\ell \tau^\ell w^\ell = w^\ell K^\ell w^\ell$ by our choice. So L' contains $w\tau K\tau^{-1}w$ if and only if $\tau_\ell K_\ell \tau_\ell^{-1}$ contains (and hence equals) $w_\ell \tau_\ell K_\ell \tau_\ell^{-1} w_\ell$, i.e., when

$$\gamma_\ell := \tau_\ell^{-1} w_\ell \tau_\ell = \begin{pmatrix} 1 & & & 2/\ell \\ & -1 & -2/\ell & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

normalizes K_ℓ. This is true for $\ell = 2$, where we even have $\gamma_\ell \in K_\ell$. □

Lemma 5.4. *For all $\eta \in \{h, h\tau \mid h \in H\}$, we have $\nu(H_\eta) = \nu(U)$.*

Proof. For any $h \in H$ and $\eta \in G$, we have $\nu(H_{h\eta}) = \nu(H_\eta)$ as $H_{h\eta} = hH_\eta h^{-1}$. So it suffices to restrict our attention to $\eta \in \{1_G, \tau\}$. The case $\eta = 1_G$ is trivial and the argument for $\eta = \tau$ is as follows. Let $\boldsymbol{A} \subset \boldsymbol{G}_{\mathbb{Q}_\ell}$ denote the maximal diagonal torus and let $A_\ell := \boldsymbol{A}(\mathbb{Q}_\ell)$, $A_{\tau_\ell}^\circ = A_\ell \cap \tau_\ell K_\ell \tau_\ell^{-1}$. Then $\mathrm{diag}(a, b, c, d) \in A_{\tau_\ell}^\circ$ if and only if

$$a, b, c, d \in \mathbb{Z}_\ell^\times, \quad ac = bd, \quad a - d, b - c \in \ell\mathbb{Z}_\ell.$$

It is then easily seen that $\nu(A_{\tau_\ell}^\circ) = \mathbb{Z}_\ell^\times$. Since H_τ contains $A_{\tau_\ell}^\circ U^\ell$, we have $\nu(H_\tau) = \mathbb{Z}_\ell^\times \nu(U^\ell) = \nu(U)$. □

Define $\sigma = \sigma^\ell \sigma_\ell \in \boldsymbol{G}(\mathbb{A}_f)$ by setting $\sigma_\ell = \mathrm{diag}(\ell, \ell, 1, 1) \in \boldsymbol{G}(\mathbb{Q}_\ell)$ and $\sigma^\ell = 1$.

Lemma 5.5. *We have* $K\sigma K = U\sigma K \sqcup U\sigma\tau K$. *Moreover*:

(a) $|K\sigma K/K| = 1+\ell+\ell^2+\ell^3$.

(b) $\deg[U\sigma K]_* = (1+\ell)^2$.

(c) $\deg[U\sigma\tau K]_* = 1$.

Proof. The first statement is [Shah 2024, Proposition 9.3.3]. Alternatively, note that, by [Taylor 1988, p. 38] or [Roberts and Schmidt 2007, p. 189], a set of representatives for $K\sigma K/K$ (in our convention) is given by

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & \ell \\ & & & \ell \end{pmatrix}, \quad \begin{pmatrix} 1 & & & \\ & \ell & & b \\ & & \ell & \\ & & & 1 \end{pmatrix}, \quad \begin{pmatrix} \ell & a & b & \\ & 1 & & \\ & & 1 & \\ & & -a & \ell \end{pmatrix}, \quad \begin{pmatrix} \ell & & b & a \\ & \ell & a & c \\ & & 1 & \\ & & & 1 \end{pmatrix},$$

where the entries a, b, c in each of the displayed matrices run over $0, 1, \ldots, \ell-1$. One now shows by applying elementary row and column operations that the classes of these matrices in $U\backslash G/K$ are represented by σ and $\sigma\tau$. That σ and $\sigma\tau$ represent distinct classes in $U\backslash G/K$ follows by noticing that $HK \neq H\tau K$ (which one can see by showing that $h\tau \notin K$ for any $h \in H$).

Part (a) follows from the decomposition above. Part (b) follows from [Shah 2024, Lemma 9.4.1(a)] by evaluating the function computed there at the zero matrix, and part (c) by [ibid., Lemma 9.3.2]. Alternatively, (b) can also be computed by relating $\deg[U\sigma K]_*$ to the degree of the T_ℓ Hecke operator for GL_2, and part (c) by observing that, for $h \in H$, the conditions on its matrix entries imposed by $(\sigma\tau)^{-1}h(\sigma\tau) \in K$ already force $h \in U$. □

Let $\sigma_1 := \sigma$ and $\sigma_2 := \sigma\tau$, so $I = \{1, 2\}$ is our indexing set. By our choice of K and Lemma 3.21, we have $d_{\sigma_1,K} = d_{1_G,K} = 1$. Lemmas 5.4 and 4.15 imply that we can take $B_i = \{1_H\}$ and that $c_i = e_i$ for $i = 1, 2$. Invoking Lemma 5.5, equation (5.1) reads

$$1+\ell+\ell^2+\ell^3 \stackrel{?}{=} (\ell+1)^2 + d_{\sigma_2,K}. \tag{5.6}$$

If $\ell = 2$ (e.g., take $K = K(N)$ for $N \geq 3$ odd), we have $d_{\tau,K} = 1$ by Lemma 5.3 and so $d_{\sigma_2,K} = 1$ by Lemma 3.21. But, in that case, the left-hand side of (5.6) is 15 while the right-hand side is 10.

5.3. ***Unitary groups.*** In Section 5.2, our eventual counterexample only worked for $\ell = 2$, which might seem a little unsatisfactory in terms of scope. In this section, we consider certain unitary Shimura varieties for which there is an abundance of elements w satisfying Lemma 2.7. This allows us to furnish a set of counterexamples for all primes ℓ that are split in an imaginary quadratic extension used to define the Shimura variety. Although the ideas are the same as before, a little more work is involved.

Let $E = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field, and let $\gamma \in \mathrm{Gal}(E/\mathbb{Q})$ denote the nontrivial element. For a positive integer p, let $\mathrm{GU}(p, p)$ be the unitary similitude group over $\mathbb{Q}$ of signature (p, p), defined with respect to the Hermitian pairing over E given by the Hermitian matrix $J = \mathrm{diag}(1, -1, 1, -1, \ldots, 1, -1)$. That is, J is the $2p \times 2p$ diagonal matrix with 1 in the odd-numbered

entries and -1 in the even-numbered entries (cf. [Graham and Shah 2023, Section 3.1]). Let $c : \mathrm{GU}(p, p) \to \mathbb{G}_m$ denote the similitude map and $\det : \mathrm{GU}(p, p) \to \mathrm{GU}(1)$ the determinant. Set

$$\boldsymbol{H} = \mathrm{GU}(1,1) \times_c \mathrm{GU}(1,1), \quad \boldsymbol{G} = \mathrm{GU}(2,2)$$

and let $\iota : \boldsymbol{H} \to \boldsymbol{G}$ be the embedding $(h_1, h_2) \mapsto \mathrm{diag}(h_1, h_2)$. Then both $\boldsymbol{H}$ and $\boldsymbol{G}$ admit standard Shimura data, given by sending $z \in \mathbb{C}^\times$ to alternating copies of z and $\bar{z}$ along the diagonal. Moreover, both of these data satisfy (SV5), and ι constitutes an embedding of Shimura data. We also observe that the derived group of $\boldsymbol{H}$ is simply connected, since $(\boldsymbol{H}^{\mathrm{der}})_{\overline{\mathbb{Q}}} \simeq \mathrm{SL}_2 \times \mathrm{SL}_2$. Therefore, Corollary 4.20 applies in this context as well.

Let U_1 denote the torus of norm one elements in $E^\times$, as in [Graham and Shah 2023, Section 3.1]. Then $\boldsymbol{T} = \boldsymbol{H}/\boldsymbol{H}^{\mathrm{der}}$ is isomorphic to $\mathrm{U}_1 \times \mathbb{G}_m \times \mathrm{U}_1$ in such a way that $\nu : \boldsymbol{H} \to \boldsymbol{T}$ is identified with the map

$$\nu : \boldsymbol{H} \to \mathrm{U}_1 \times \mathbb{G}_m \times \mathrm{U}_1, \quad h = (h_1, h_2) \mapsto \left(\frac{c(h)}{\det h_1}, c(h), \frac{c(h)}{\det h_2} \right),$$

where $c(h) := c(h_1) = c(h_2)$ is the common similitude. It is easily seen that $\boldsymbol{T}(\mathbb{Q})$ is discrete in $\boldsymbol{T}(\mathbb{A}_f)$. Let $\mathscr{N} : E^\times \to \mathbb{Q}^\times$ denote the norm map and let $S \subset E^\times \setminus \{1\}$ denote the set of elements such that $\mathscr{N}(\xi) = \xi \cdot \gamma(\xi) = 1$. For any $\xi \in S$, define

$$w_\xi := \mathrm{diag}(1, 1, \xi, \xi) \in \boldsymbol{Z}_{\boldsymbol{H}}(\mathbb{Q}).$$

For any $\mathbb{Q}$-algebra R and $g \in \boldsymbol{G}(R)$, the condition $g w_\xi = w_\xi g$ forces g to be block diagonal. Thus, the centralizer of w_ξ in $\boldsymbol{G}$ equals $\boldsymbol{H}$. As before, we will consider $w_\xi \in \boldsymbol{G}(\mathbb{Q})$ as an element of G via the diagonal embedding $\boldsymbol{G}(\mathbb{Q}) \hookrightarrow \boldsymbol{G}(\mathbb{A}_f) = G$.

Let $\ell \nmid 2d$ be a fixed rational prime that is split in E. If $\jmath : E \to \mathbb{Q}_\ell$ is an embedding and $\xi \in S$ is such that $\jmath(\xi) \in 1 + \ell\mathbb{Z}_\ell$, then $\jmath \circ \gamma(\xi) \in 1 + \ell\mathbb{Z}_\ell$ as well, since $\gamma(\xi)$ is the inverse of ξ. We will refer to such $\xi \in S$ as *ℓ-invertible*. For positive integers m,

$$\xi_m := \left(\frac{1 - dm^2\ell^2}{1 + dm^2\ell^2} \right) + \left(\frac{2m\ell}{1 + dm^2\ell^2} \right) \sqrt{-d} \in S$$

are examples of such elements.

Fix now a neat compact–open subgroup K and let ℓ be a prime split in E such that K is hyperspecial at ℓ, i.e., $K = K^\ell K_\ell$ with $K_\ell = \boldsymbol{G}(\mathbb{Z}_\ell)$. For any $\xi \in S$, define $K_\xi := K \cap w_\xi^{-1} K w_\xi$. Then K_ξ is a neat compact–open subgroup of G, and both K_ξ and $w_\xi K_\xi w_\xi^{-1}$ are contained in K. If ξ is also ℓ-invertible, then $\xi \in E \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \simeq \mathbb{Q}_\ell \oplus \mathbb{Q}_\ell$ lies in $\mathbb{Z}_\ell^\times \times \mathbb{Z}_\ell^\times$ and the ℓ-component $w_{\xi,\ell} \in \boldsymbol{G}(\mathbb{Q}_\ell)$ of w_ξ lies in K_ℓ. Thus, for an ℓ-invertible element $\xi \in S$, which we fix in what follows, K_ξ is hyperspecial at ℓ. The upshot of this discussion is that, by replacing K with K_ξ, we may assume that:

- K and $w_\xi K w_\xi^{-1}$ are hyperspecial (and equal) at ℓ.
- K and $w_\xi K w_\xi^{-1}$ are contained in a common neat compact–open subgroup.

Set $U := K \cap \boldsymbol{H}(\mathbb{A}_f)$. Then $\iota_{U,K}$ is a closed immersion by Lemma 2.7 and our assumptions on K.

Since ℓ is split in E, we can fix isomorphisms $\boldsymbol{H}_{\mathbb{Q}_\ell} \simeq \mathbb{G}_m \times \mathrm{GL}_2 \times \mathrm{GL}_2$ and $\boldsymbol{G}_{\mathbb{Q}_\ell} \simeq \mathbb{G}_m \times \mathrm{GL}_4$ such that the local embedding $\iota : \boldsymbol{H}_{\mathbb{Q}_\ell} \hookrightarrow \boldsymbol{G}_{\mathbb{Q}_\ell}$ is identified with the embedding

$$\mathbb{G}_m \times \mathrm{GL}_2 \times \mathrm{GL}_2 \hookrightarrow \mathbb{G}_m \times \mathrm{GL}_4, \quad (c, h_1, h_2) \mapsto (c, \mathrm{diag}(h_1, h_2)),$$

using the embedding $j : E \hookrightarrow \mathbb{Q}_\ell$. In the 4×4 matrices below, we suppress the similitude coordinate, which is 1, and write ξ for $j(\xi)$.

Let

$$\tau_\ell := \begin{pmatrix} 1 & & 1/\ell & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \in \boldsymbol{G}(\mathbb{Q}_\ell)$$

and let $\tau = \tau_\ell \tau^\ell \in \boldsymbol{G}(\mathbb{A}_f)$, where τ^ℓ is the identity.

Lemma 5.7. *The morphism $\iota : \mathrm{Sh}_{\boldsymbol{H}}(H_\tau) \to \mathrm{Sh}_{\boldsymbol{G}}(\tau K \tau^{-1})$ is a closed immersion.*

Proof. As in the proof of Lemma 5.3, it suffices to show that

$$\tau_\ell^{-1} w_{\xi,\ell} \tau_\ell = \begin{pmatrix} 1 & & (1-\xi)/\ell & \\ & 1 & & \\ & & \xi & \\ & & & \xi \end{pmatrix} \in \boldsymbol{G}(\mathbb{Q}_\ell)$$

lies in K_ℓ, which it does by ℓ-invertibility of ξ. □

Lemma 5.8. *For all $\eta \in \{h, h\tau \mid h \in H\}$, $\nu(H_\eta) = \nu(U)$.*

Proof. As in Lemma 5.4, it suffices to restrict to $\eta \in \{1_H, \tau\}$, and the claim for $\eta = 1_H$ is again trivial. At a split prime, the choice of an $\alpha \in \mathbb{Q}_\ell$ such that $\alpha^2 = -d$ determines compatible isomorphisms $\boldsymbol{T}_{\mathbb{Q}_\ell} \simeq \mathbb{G}_m^3$ and $\boldsymbol{H}_{\mathbb{Q}_\ell} \simeq \mathbb{G}_m \times \mathrm{GL}_2 \times \mathrm{GL}_2$ such that $\nu : \boldsymbol{H}_{\mathbb{Q}_\ell} \to \boldsymbol{T}_{\mathbb{Q}_\ell}$ is given by $(c, h_1, h_2) \mapsto (c \det h_1^{-1}, c, c \det h_2^{-1})$. Then one can show that $H_{\tau_\ell} := \boldsymbol{H}(\mathbb{Q}_\ell) \cap \tau_\ell K_\ell \tau_\ell^{-1}$ contains the subgroup

$$\{(u, \mathrm{diag}(a, b), \mathrm{diag}(a, c)) \in \mathbb{G}_m \times \mathrm{GL}_2 \times \mathrm{GL}_2 \mid a, b, c, u \in \mathbb{Z}_\ell^\times\}.$$

From this, one deduces that $\nu(H_{\tau_\ell}) = \mathbb{Z}_\ell^\times \times \mathbb{Z}_\ell^\times \times \mathbb{Z}_\ell^\times = \nu(U_\ell)$. □

Let $\sigma, \sigma' \in G$ be defined so that their components away from ℓ are the identity and at ℓ are given by $\sigma_\ell := (1, \mathrm{diag}(\ell, 1, 1, 1))$, $\sigma'_\ell = (1, \mathrm{diag}(1, 1, \ell, 1))$.

Lemma 5.9. *We have $K\sigma K = U\sigma K \sqcup U\sigma' K \sqcup U\sigma\tau K$. Moreover*:

(a) $|K\sigma K/K| = 1 + \ell + \ell^2 + \ell^3$.

(b) $\deg[U\sigma K]_* = \deg[U\sigma' K]_* = 1 + \ell$.

(c) $\deg[U\sigma\tau K]_* = 1$.

Proof. The first claim is a special case of [Shah 2024, Proposition 7.4.5]. Alternatively, note that the stratification of the projective space $\mathbb{P}^3$ over the field $\mathbb{Z}/\ell\mathbb{Z}$ implies that a set of representatives for

$K\sigma K/K$ is given by

$$\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&\ell\end{pmatrix},\quad \begin{pmatrix}1&&&\\&1&&\\&&\ell&a\\&&&1\end{pmatrix},\quad \begin{pmatrix}1&&&\\&\ell&a&b\\&&1&\\&&&1\end{pmatrix},\quad \begin{pmatrix}\ell&a&b&c\\&1&&\\&&1&\\&&&1\end{pmatrix},$$

where the $\mathbb{Q}_\ell^\times$-component (in $\boldsymbol{G}(\mathbb{Q}_\ell)\simeq\mathbb{Q}_\ell^\times\times\mathrm{GL}_4(\mathbb{Q}_\ell)$) is taken to be 1, and where the entries a, b, c in each of the displayed matrices run over $0,1,\dots,\ell-1$. One then reduces each of these matrices by appropriate row and column operations to show that the classes of these matrices in $U\backslash G/K$ are represented by σ, σ' and $\sigma\tau$. It is easily checked that $HK\neq H\tau K$, which distinguishes $U\sigma\tau K$ from $U\sigma K$ and $U\sigma' K$. To distinguish $U\sigma K$ from $U\sigma' K$, we use the Cartan decomposition for the double quotient $\boldsymbol{H}(\mathbb{Z}_\ell)\backslash\boldsymbol{H}(\mathbb{Q}_\ell)/\boldsymbol{H}(\mathbb{Z}_\ell)$ and an elementary trick established in [ibid., Lemma 5.9.2].

Part (a) follows from the decomposition above. Parts (b) and (c) can be deduced along the lines outlined in the proof of Lemma 5.5. □

Set $\sigma_1:=\sigma$, $\sigma_2:=\sigma'$, $\sigma_3:=\sigma\tau$, so $I=\{1,2,3\}$ is our indexing set. We have $d_{1_G,K}=d_{\sigma_i,K}=1$ for $i=1,2,3$ by our choice of K and Lemmas 5.7 and 3.21. Lemma 5.8 (in conjunction with Lemma 4.15) implies that we may take $B_i=\{1_H\}$ and that $c_i=e_i$ for $i=1,2,3$. Equation (5.1) then reads

$$1+\ell+\ell^2+\ell^3\overset{?}{=}(1+\ell)+(1+\ell)+1,$$

which is false for all ℓ.

6. Schwartz spaces

To have another — and a technically simpler — perspective on the failure of (1.5) in general, we investigate an auxiliary question in the setting of Schwartz spaces motivated by the discussion in Section 3.1. This section can be read independently of the rest of this note.

Let G be a locally profinite group and J be a closed subgroup of G. Then $X:=J\backslash G$ is a locally compact Hausdorff and totally disconnected topological space with a continuous right action $X\times G\to X$ given by $(J\gamma,g)\mapsto J\gamma g$. Let $\mathcal{S}_{\mathbb{Z}}(X)$ denote the $\mathbb{Z}$-module of all $\mathbb{Z}$-valued locally constant compactly supported functions on X. We have an induced smooth left action

$$G\times\mathcal{S}_{\mathbb{Z}}(X)\to\mathcal{S}_{\mathbb{Z}}(X),\quad (g,\xi)\mapsto\xi((-)g). \tag{6.1}$$

For each compact–open subgroup K of G, we let $M(K):=\mathcal{S}_{\mathbb{Z}}(X/K)=\mathcal{S}_{\mathbb{Z}}(X)^K$ denote the $\mathbb{Z}$-module of all functions in $\mathcal{S}_{\mathbb{Z}}(X)$ that are K-invariant under the natural G-action. For $\sigma\in G$, we define $\mathbb{Z}$-linear maps

$$\mathcal{T}(\sigma):M(K)\to M(K),\qquad \mathrm{ch}(JgK)\mapsto\sum_{\gamma\in K\sigma K/K}\mathrm{ch}(Jg\gamma K),$$

$$[K\sigma K]_*:M(K)\to M(K),\qquad \mathrm{ch}(JgK)\mapsto\sum_{\delta\in K\backslash K\sigma K}\mathrm{ch}(JgK\delta),$$

where ch(−) denotes the characteristic function. Then $\mathcal{T}(\sigma)$ and $[K\sigma K]_*$ are respectively the analogs of (1.1) and (1.2) in this setting.

Question 6.2. *Is $\mathcal{T}(\sigma) = [K\sigma K]_*$ for all $\sigma \in G$?*

Of course, this is the case when J is the trivial subgroup, since both $\mathcal{T}(\sigma)$ and $[K\sigma K]_*$ send $\mathrm{ch}(gK)$ to the characteristic function of $gK\sigma K$. More generally, we have the following.

Proposition 6.3. *Suppose that, for any $\gamma \in G$, $J \cap \gamma K \gamma^{-1}$ equals a fixed subgroup of J. Then, for all $\sigma \in G$, $\mathcal{T}(\sigma) = [K\sigma K]_*$.*

Proof. Let Σ be the collection of all compact–open subgroups of G that are equal to a finite intersection of conjugates of K. Our assumption on J implies that, for any $g \in G$ and $L \in \Sigma$ satisfying $L \subset K$, $JgK = \bigsqcup_{\mu \in K/L} Jg\mu L$. If we take L to be the intersection of K and $\bigcap_{\delta \in K\backslash K\sigma K} \delta^{-1} K \delta$, then L is normal in K and its conjugates by $\delta \in K\sigma K$ are contained in K. These properties imply that, for any $\gamma, \delta \in K\sigma K$ and $g \in G$,

$$\mathrm{ch}(Jg\gamma K) = \sum_{\mu \in K/L} \mathrm{ch}(Jg\gamma\mu L), \quad \mathrm{ch}(JgK\delta) = \sum_{\nu \in \delta^{-1}K\delta/L} \mathrm{ch}(Jg\delta\nu L).$$

We wish to show that

$$\sum_{\gamma \in K\sigma K/K} \sum_{\mu \in K/L} \mathrm{ch}(Jg\gamma\mu L) = \sum_{\delta \in K\backslash K\sigma K} \sum_{\nu \in \delta^{-1}K\delta/L} \mathrm{ch}(Jg\delta\nu L).$$

Both sides are sums of characteristic functions on cosets $J\backslash G/L$, possibly with repetitions. But both the lists

- $\gamma\mu L$ for $\gamma \in K\sigma K/K$ and $\mu \in K/L$,
- $\delta\nu L$ for $\delta \in K\backslash K\sigma K$ and $\nu \in \delta^{-1}K\delta/L$

of cosets in G/L enumerate each element of $K\sigma K/L$ exactly once. □

An example where the condition of Proposition 6.3 is satisfied is when $G = \boldsymbol{G}(\mathbb{A}_f)$ for some reductive group $\boldsymbol{G}$ over $\mathbb{Q}$, K is a neat compact–open subgroup of G, and $J = \boldsymbol{T}(\mathbb{Q})$, where $\boldsymbol{T} \hookrightarrow \boldsymbol{G}$ is a torus in $\boldsymbol{G}$ such that $\boldsymbol{T}(\mathbb{Q})$ is discrete in $\boldsymbol{T}(\mathbb{A}_f)$. Indeed, $\boldsymbol{T}(\mathbb{A}_f) \cap \gamma K \gamma^{-1}$ is a compact–open subgroup of $\boldsymbol{T}(\mathbb{A}_f)$ for any $\gamma \in G$. So the discreteness of $J = \boldsymbol{T}(\mathbb{Q})$ in $\boldsymbol{T}(\mathbb{A}_f)$ implies that $J \cap \gamma K \gamma^{-1}$ is finite and the neatness of K forces $J \cap \gamma K \gamma^{-1}$ to be torsion free (see [Graham and Shah 2023, Definition B.6]). See Example 4.23. It is however also easy to find situations where $\mathcal{T}(\sigma) \neq [K\sigma K]_*$.

Example 6.4. Suppose $J = K$. Then $M(K) = \mathbb{Z}[K\backslash G/K]$ is the free $\mathbb{Z}$-module on the characteristic functions of double cosets in $K\backslash G/K$. Now, for any $g \in G$, $\mathcal{T}(\sigma) \cdot \mathrm{ch}(KgK) = \sum_{\gamma \in K\sigma K/K} \mathrm{ch}(Kg\gamma K)$ by definition, whereas

$$[K\sigma K]_* \cdot \mathrm{ch}(KgK) = \mathrm{ch}(KgK) * \mathrm{ch}(K\sigma K),$$

where $*$ denotes the convolution product on $M(K)$ with respect to a Haar measure μ on G such that $\mu(K) = 1$. Note that the map $\mathrm{ind} : M(K) \to \mathbb{Z}$ given by $\mathrm{ch}(K\gamma K) \mapsto |K\gamma K/K|$ is a $\mathbb{Z}$-algebra

homomorphism (see, e.g., [Shah 2024, Section 2.3]), so $\mathcal{T}(\sigma)\cdot \mathrm{ch}(KgK) = [K\sigma K]_* \cdot \mathrm{ch}(KgK)$ would imply that

$$\sum_{\gamma \in K\sigma K/K} |Kg\gamma K/K| = |KgK/K| \cdot |K\sigma K/K|. \tag{6.5}$$

This is clearly false in general. For instance, if $g = 1_G$, the left-hand side of (6.5) is $|K\sigma K/K|^2$ while the right-hand side is $|K\sigma K/K|$, and these are not equal as soon as $|K\sigma K/K| \neq 1$.

Remark 6.6. We may consider Question 6.2 as a problem of "explicit descent" in the following sense. We know that $\sum_{\delta \in K\backslash K\sigma K} \mathrm{ch}(JgK\delta)$ is an element of $M(K)$ as it is K-invariant (a descent phenomenon) and we even know that its support is $JgK\sigma K$. We however want an *explicit* linear combination of the basis $\{\mathrm{ch}(J\gamma K) \mid \gamma \in G\}$ of $M(K)$ that equals this element. This amounts to computing certain volumes of subsets of J. If one establishes that (3.3) or an appropriate variant of it is a bijection and the maps are equivariant for varying K, then one has an alternative strategy for deriving Corollary 4.20. This approach, however, does not work in the generality of Theorem 4.12.

Acknowledgements

The author is grateful to Christophe Cornut for several discussions and to Barry Mazur for valuable feedback on an earlier draft. The author also thanks the referees for their diligent reading of the manuscript and for numerous helpful suggestions that significantly improved the exposition of this note.

References

[Anschütz 2017] J. Anschütz, "Descent of algebraic cycles", *Homology Homotopy Appl.* **19**:1 (2017), 135–154. MR

[Barwick 2017] C. Barwick, "Spectral Mackey functors and equivariant algebraic K-theory, I", *Adv. Math.* **304** (2017), 646–727. MR

[Bley and Boltje 2004] W. Bley and R. Boltje, "Cohomological Mackey functors in number theory", *J. Number Theory* **105**:1 (2004), 1–37. MR

[Borel 2019] A. Borel, *Introduction to arithmetic groups*, University Lecture Series **73**, Amer. Math. Soc., Providence, RI, 2019. MR

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Math. (3) **21**, Springer, 1990. MR

[Boumasmoud et al. 2020] R. Boumasmoud, E. H. Brooks, and D. P. Jetchev, "Vertical distribution relations for special cycles on unitary Shimura varieties", *Int. Math. Res. Not.* **2020**:13 (2020), 3902–3926. MR

[Cornut 2018] C. Cornut, "An Euler system of Heegner type", preprint, 2018, https://webusers.imj-prg.fr/~christophe.cornut/papers/ESHT.pdf.

[Cornut and Vatsal 2007] C. Cornut and V. Vatsal, "Nontriviality of Rankin–Selberg L-functions and CM points", pp. 121–186 in *L-functions and Galois representations*, edited by D. Burns et al., London Math. Soc. Lecture Note Ser. **320**, Cambridge Univ. Press, 2007. MR

[Deligne 1971] P. Deligne, "Travaux de Shimura", Exposé 389, pp. 123–165 in *Séminaire Bourbaki* 1970/1971, Lecture Notes in Math. **244**, Springer, 1971. MR

[Demazure and Gabriel 1980] M. Demazure and P. Gabriel, *Introduction to algebraic geometry and algebraic groups*, North-Holland Mathematics Studies **39**, North-Holland, Amsterdam, 1980. MR

[EGA I 1971] A. Grothendieck and J. A. Dieudonné, *Éléments de géométrie algébrique*, *I*, Grundl. Math. Wissen. **166**, Springer, 1971. MR

[Esnault and Harris 2017] H. Esnault and M. Harris, "Chern classes of automorphic vector bundles", *Pure Appl. Math. Q.* **13**:2 (2017), 193–213. MR

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Math. (3) **2**, Springer, 1998. MR

[Graham and Shah 2023] A. Graham and S. W. A. Shah, "Anticyclotomic Euler systems for unitary groups", *Proc. Lond. Math. Soc.* (3) **127**:6 (2023), 1577–1680. MR

[Grossi 2020] G. Grossi, "On norm relations for Asai–Flach classes", *Int. J. Number Theory* **16**:10 (2020), 2311–2377. MR

[Jannsen 1988] U. Jannsen, "Continuous étale cohomology", *Math. Ann.* **280**:2 (1988), 207–245. MR

[Jetchev 2014] D. P. Jetchev, "Hecke and Galois properties of special cycles on unitary Shimura varieties", preprint, 2014. arXiv 1410.6692

[Kudla 1997] S. S. Kudla, "Algebraic cycles on Shimura varieties of orthogonal type", *Duke Math. J.* **86**:1 (1997), 39–78. MR

[Lai and Skinner 2024] S. Lai and C. Skinner, "Anti-cyclotomic Euler system of diagonal cycles", preprint, 2024. arXiv 2408.01219

[Lei et al. 2015] A. Lei, D. Loeffler, and S. L. Zerbes, "Euler systems for modular forms over imaginary quadratic fields", *Compos. Math.* **151**:9 (2015), 1585–1625. MR

[Lei et al. 2018] A. Lei, D. Loeffler, and S. L. Zerbes, "Euler systems for Hilbert modular surfaces", *Forum Math. Sigma* **6** (2018), art. id. e23. MR

[Lewis et al. 1981] G. Lewis, J. P. May, and J. McClure, "Ordinary $RO(G)$-graded cohomology", *Bull. Amer. Math. Soc.* (*N.S.*) **4**:2 (1981), 208–212. MR

[Li and Liu 2021] C. Li and Y. Liu, "Chow groups and L-derivatives of automorphic motives for unitary groups", *Ann. of Math.* (2) **194**:3 (2021), 817–901. MR

[Loeffler et al. 2022] D. Loeffler, C. Skinner, and S. L. Zerbes, "Euler systems for GSp(4)", *J. Eur. Math. Soc.* **24**:2 (2022), 669–733. MR

[Milne 2005] J. S. Milne, "Introduction to Shimura varieties", pp. 265–378 in *Harmonic analysis, the trace formula, and Shimura varieties*, edited by J. Arthur et al., Clay Math. Proc. **4**, Amer. Math. Soc., Providence, RI, 2005. MR

[Noot 2006] R. Noot, "Correspondances de Hecke, action de Galois et la conjecture d'André-Oort (d'après Edixhoven et Yafaev)", Exposé 942, pp. 165–197 in *Séminaire Bourbaki*, 2004/2005, Astérisque **307**, Soc. Math. France, Paris, 2006. MR

[Pink 1989] R. Pink, *Arithmetical compactification of mixed Shimura varieties*, Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 1989, https://people.math.ethz.ch/~pink/dissertation.html. MR

[Poonen 2017] B. Poonen, *Rational points on varieties*, Graduate Studies in Mathematics **186**, Amer. Math. Soc., Providence, RI, 2017. MR

[Renard 2010] D. Renard, *Représentations des groupes réductifs p-adiques*, Cours Spécialisés **17**, Soc. Math. France, Paris, 2010. MR

[Ribet 1990] K. A. Ribet, "On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms", *Invent. Math.* **100**:2 (1990), 431–476. MR

[Roberts and Schmidt 2007] B. Roberts and R. Schmidt, *Local newforms for* GSp(4), Lecture Notes in Math. **1918**, Springer, 2007. MR

[Shah 2023] S. W. A. Shah, "Norm relations for CM points on modular curves", preprint, 2023. arXiv 2310.14543

[Shah 2024] S. W. A. Shah, "On constructing zeta elements for Shimura varieties", preprint, 2024. arXiv 2409.03517

[Stacks 2005–] "The Stacks project", electronic reference, 2005–, http://stacks.math.columbia.edu.

[Taylor 1988] R. L. Taylor, *On congruences between modular forms*, Ph.D. thesis, Princeton University, 1988, https://www.proquest.com/docview/303728133. MR

[Zhang 2001] S. Zhang, "Heights of Heegner points on Shimura curves", *Ann. of Math.* (2) **153**:1 (2001), 27–147. MR

Communicated by Shou-Wu Zhang
Received 2024-10-22 Revised 2025-04-05 Accepted 2025-06-27

swshah@ucsb.edu *Department of Mathematics, University of California, Santa Barbara, Santa Barbara, CA, United States*

Current address: *Department of Mathematics, Bilkent University, Ankara, Turkey*

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 20 No. 8 2026

Supersymmetric monoidal categories 1479
Steven V Sam and Andrew Snowden

Rational curves and the Hilbert property on Jacobian Kummer varieties 1543
Damián Gvirtz-Chen and Zhizhong Huang

Homological regularities and concavities 1559
Ellen Kirkman, Robert Won and James J. Zhang

Transcendence properties of the Artin–Hasse exponential modulo p 1597
Joe Kramer-Miller

Anticyclotomic Euler system over biquadratic fields 1615
Kim Tuan Do

Explicit Hecke descent for special cycles 1649
Syed Waqar Ali Shah