Pacific Journal of Mathematics

Volume 291 No. 2 December 2017

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

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

The subscription price for 2017 is US $450/year for the electronic version, and $625/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.257

TORSION PAIRS IN SILTING THEORY

LIDIA ANGELERI HÜGEL, FREDERIK MARKS AND JORGE VITÓRIA

In the setting of compactly generated triangulated categories, we show that the heart of a (co)silting t-structure is a Grothendieck category if and only if the (co)silting object satisfies a purity assumption. Moreover, in the cosilting case the previous conditions are related to the coaisle of the t-structure being a definable subcategory. If we further assume our triangulated category to be algebraic, it follows that the heart of any nondegenerate compactly generated t-structure is a Grothendieck category.

1. Introduction

Silting and cosilting objects in triangulated categories are useful generalisations of tilting and cotilting objects. While (co)tilting objects have been a source of many interactions with torsion and localisation theory, it is in the setting of (co)silting objects that classification results occur more naturally. This paper strengthens this claim by showing that, in the setting of compactly generated triangulated categories, relevant torsion-theoretic structures are parametrised by suitable classes of (co)silting objects.

The concept of a silting object, first introduced in [Keller and Vossieck 1988] in the context of derived module categories over finite dimensional hereditary algebras, has recently been extended to the setting of abstract triangulated categories [Aihara and Iyama 2012; Mendoza Hernández et al. 2013; Nicolás et al. 2015; Psaroudakis and Vitória 2015]. In this paper, our focus is on t-structures and co-t-structures arising from (co)silting objects. For this purpose, we use the vast theory of *purity* in compactly generated triangulated categories, where a central role is played by the category of contravariant functors on the compact objects. We show that a fundamental property of the t-structure associated to a cosilting object C — namely, its heart being a Grothendieck abelian category — is related to the pure-injectivity of C. An analogous result holds true for silting objects. Moreover, it turns out that in the cosilting case the pure-injectivity of C is further related to the definability (in terms of coherent functors) of the coaisle of the associated t-structure. We can summarise our results as follows.

MSC2010: 18E15, 18E30, 18E40.
Keywords: torsion pair, silting, cosilting, t-structure, Grothendieck category.

Theorem [Theorems 3.6 and 4.9, Corollary 4.10]. *Let* $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ *be a triple in a compactly generated triangulated category* $\mathcal{T}$ *such that* $(\mathcal{U}, \mathcal{V})$ *is a nondegenerate t-structure and* $(\mathcal{V}, \mathcal{W})$ *is a co-t-structure. Then the following are equivalent*:

(1) $\mathcal{V}$ *is definable in* $\mathcal{T}$;

(2) $\mathcal{V} = {}^{\perp_{>0}}C$ *for a pure-injective cosilting object* C *in* $\mathcal{T}$;

(3) $\mathcal{H} := \mathcal{U}[-1] \cap \mathcal{V}$ *is a Grothendieck category.*

In particular, *if we further assume* $\mathcal{T}$ *to be algebraic*, *it follows that any nondegenerate compactly generated t-structure in* $\mathcal{T}$ *has a Grothendieck heart.*

For partial results in this direction we refer to [Nicolás et al. 2015, Proposition 4.2; Bravo and Parra 2016, Corollary 2.5]. In a forthcoming paper ([Marks and Vitória 2017]), it will be proved that cosilting complexes in derived module categories are always pure-injective and give rise to definable subcategories as above. We do not know, however, if the same holds true for arbitrary cosilting objects in compactly generated triangulated categories. Moreover, it will be shown in [Marks and Vitória 2017] that there are cosilting complexes (in fact, cosilting modules) inducing triples $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ as above such that the t-structure has a Grothendieck heart, although it is not compactly generated. This will answer [Bravo and Parra 2016, Question 3.5].

The structure of the paper is as follows. In Section 2, we present our setup and provide the reader with some preliminaries on torsion pairs and (co)silting objects. In Section 3, we briefly recall the key concepts of pure-projectivity and pure-injectivity and we establish the connection between (co)silting objects having such properties and t-structures with Grothendieck hearts. Finally, in Section 4, we discuss definable subcategories and we prove the above mentioned relation between pure-injective cosilting objects and certain definable subcategories of the underlying triangulated category.

2. Preliminaries

Setup and notation. Throughout, we denote by $\mathcal{T}$ a *compactly generated* triangulated category, i.e., a triangulated category with coproducts for which the subcategory of compact objects, denoted by $\mathcal{T}^c$, has only a set of isomorphism classes and such that for any Y in $\mathcal{T}$ with $\mathrm{Hom}_{\mathcal{T}}(X, Y) = 0$ for all X in $\mathcal{T}^c$, we have $Y = 0$. Since $\mathcal{T}$ admits arbitrary set-indexed coproducts, it is idempotent complete (see [Neeman 2001, Proposition 1.6.8]). It is also well known (see [Neeman 2001, Proposition 8.4.6 and Theorem 8.3.3]) that such triangulated categories admit products. In some places, we will further assume $\mathcal{T}$ to be *algebraic*, i.e., $\mathcal{T}$ can be constructed as the stable category of a Frobenius exact category (see [Happel 1988]). Note that algebraic and compactly generated triangulated categories are essentially derived categories of small differential graded categories [Keller 1994].

All subcategories considered are strict and full. For a set of integers I (which is often expressed by symbols such as $> n$, $< n$, $\geq n$, $\leq n$, $\neq n$, or just n, with the obvious associated meaning) we define the following orthogonal classes:

$$
\begin{aligned}
{}^{\perp_I}X &:= \big\{Y \in \mathcal{T} : \mathrm{Hom}_{\mathcal{T}}(Y, X[i]) = 0, \text{ for all } i \in I\} \\
X^{\perp_I} &:= \{Y \in \mathcal{T} : \mathrm{Hom}_{\mathcal{T}}(X, Y[i]) = 0, \text{ for all } i \in I\big\}.
\end{aligned}
$$

If $\mathscr{C}$ is a subcategory of $\mathcal{T}$, then we denote by $\mathsf{Add}(\mathscr{C})$ (respectively, $\mathsf{Prod}(\mathscr{C})$) the smallest subcategory of $\mathcal{T}$ containing $\mathscr{C}$ and closed under coproducts (respectively, products) and summands. If $\mathscr{C}$ consists of a single object M, we write $\mathsf{Add}(M)$ and $\mathsf{Prod}(M)$ for the respective subcategories. For a ring A, we denote by $\mathsf{Mod}(A)$ the category of right A-modules and by $\mathsf{D}(A)$ the unbounded derived category of $\mathsf{Mod}(A)$. The subcategories of injective and of projective A-modules are denoted, respectively, by $\mathsf{Inj}(A)$ and $\mathsf{Proj}(A)$, and their bounded homotopy categories by $\mathsf{K}^b(\mathsf{Inj(A)})$ and $\mathsf{K}^b(\mathsf{Proj}(A))$, respectively.

Torsion pairs. We consider the notion of a torsion pair in a triangulated category (see, for example, [Iyama and Yoshino 2008]), which gives rise to the notions of a t-structure [Beĭlinson et al. 1982] and a co-t-structure [Bondarko 2010; Pauksztello 2008].

Definition 2.1. A pair of subcategories $(\mathcal{U}, \mathcal{V})$ in $\mathcal{T}$ is said to be a *torsion pair* if

(1) $\mathcal{U}$ and $\mathcal{V}$ are closed under summands;

(2) $\mathrm{Hom}_{\mathcal{T}}(\mathcal{U}, \mathcal{V}) = 0$;

(3) For every object X of $\mathcal{T}$, there are U in $\mathcal{U}$, V in $\mathcal{V}$ and a triangle

$$U \to X \to V \to U[1].$$

In a torsion pair $(\mathcal{U}, \mathcal{V})$, the class $\mathcal{U}$ is called the *aisle*, the class $\mathcal{V}$ the *coaisle*, and $(\mathcal{U}, \mathcal{V})$ is said to be

- *nondegenerate* if $\bigcap_{n\in\mathbb{Z}} \mathcal{U}[n] = 0 = \bigcap_{n\in\mathbb{Z}} \mathcal{V}[n]$;
- a *t-structure* if $\mathcal{U}[1] \subseteq \mathcal{U}$, in which case we say that $\mathcal{U}[-1] \cap \mathcal{V}$ is the *heart* of $(\mathcal{U}, \mathcal{V})$;
- a *co-t-structure* if $\mathcal{U}[-1] \subseteq \mathcal{U}$, in which case we say that $\mathcal{U} \cap \mathcal{V}[-1]$ is the *coheart* of $(\mathcal{U}, \mathcal{V})$.

It follows from [Beĭlinson et al. 1982] that the heart $\mathcal{H}_{\mathbb{T}}$ of a t-structure $\mathbb{T} := (\mathcal{U}, \mathcal{V})$ in $\mathcal{T}$ is an abelian category with the exact structure induced by the triangles of $\mathcal{T}$ lying in $\mathcal{H}_{\mathbb{T}}$. Furthermore, the triangle in Definition 2.1(3) can be expressed functorially as

$$u(X) \xrightarrow{\;f\;} X \xrightarrow{\;g\;} v(X) \longrightarrow u(X)[1],$$

where $u : \mathcal{T} \to \mathcal{U}$ is the right adjoint of the inclusion of $\mathcal{U}$ in $\mathcal{T}$ and $v : \mathcal{T} \to \mathcal{V}$ is the left adjoint of the inclusion of $\mathcal{V}$ in $\mathcal{T}$. The existence of one of these adjoints, usually called *truncation functors*, is in fact equivalent to the fact that $(\mathcal{U}, \mathcal{V})$ is a t-structure ([Keller and Vossieck 1988, Proposition 1.1]). Observe that the maps f and g in the triangle are, respectively, the counit and unit map of the relevant adjunction. In particular, it follows that if $f = 0$ (respectively, $g = 0$), then $u(X) = 0$ (respectively, $v(X) = 0$). Furthermore, u and v give rise to a cohomological functor defined by

$$H^0_{\mathbb{T}} : \mathcal{T} \to \mathcal{H}_{\mathbb{T}},\ X \mapsto H^0_{\mathbb{T}}(X) := v(u(X[1])[-1]) = u(v(X)[1])[-1].$$

Recall that an additive covariant functor from $\mathcal{T}$ to an abelian category $\mathcal{A}$ is said to be *cohomological* if it sends triangles in $\mathcal{T}$ to long exact sequences in $\mathcal{A}$.

We will also be interested in the properties of torsion pairs generated or cogenerated by certain subcategories of $\mathcal{T}$, which are defined as follows.

Definition 2.2. Let $(\mathcal{U}, \mathcal{V})$ be a torsion pair in $\mathcal{T}$ and $\mathcal{A}$ a subcategory of $\mathcal{T}$. We say that $(\mathcal{U}, \mathcal{V})$ is

- *generated by* $\mathcal{A}$ if $(\mathcal{U}, \mathcal{V}) = ({}^{\perp_0}(\mathcal{A}^{\perp_0}), \mathcal{A}^{\perp_0})$;
- *cogenerated by* $\mathcal{A}$ if $(\mathcal{U}, \mathcal{V}) = ({}^{\perp_0}\mathcal{A}, ({}^{\perp_0}\mathcal{A})^{\perp_0})$;
- *compactly generated* if $(\mathcal{U}, \mathcal{V})$ is generated by a set of compact objects.

Moreover, we say that $\mathcal{A}$ *generates* $\mathcal{T}$ if the subcategory $\bigcup_{n\in\mathbb{Z}} \mathcal{A}[n]$ generates the torsion pair $(\mathcal{T}, 0)$. Dually, we say that $\mathcal{A}$ *cogenerates* $\mathcal{T}$ if the subcategory $\bigcup_{n\in\mathbb{Z}} \mathcal{A}[n]$ cogenerates the torsion pair $(0, \mathcal{T})$.

Recall that a subcategory $\mathcal{U}$ of $\mathcal{T}$ is said to be *suspended* (respectively, *cosuspended*) if it is closed under extensions and positive (respectively, negative) shifts. For example, a torsion pair $(\mathcal{U}, \mathcal{V})$ is a t-structure if and only if $\mathcal{U}$ is suspended (or, equivalently, $\mathcal{V}$ is cosuspended). In particular, a t-structure generated (respectively, cogenerated) by a subcategory $\mathcal{A}$ is also generated (respectively, cogenerated) by the smallest suspended (respectively, cosuspended) subcategory containing $\mathcal{A}$. A dual statement holds for co-t-structures.

Definition 2.3. Two torsion pairs of the form $(\mathcal{U}, \mathcal{V})$ and $(\mathcal{V}, \mathcal{W})$ are said to be *adjacent*. More precisely, we say that $(\mathcal{U}, \mathcal{V})$ is *left adjacent* to $(\mathcal{V}, \mathcal{W})$ and that $(\mathcal{V}, \mathcal{W})$ is *right adjacent* to $(\mathcal{U}, \mathcal{V})$. Such $\mathcal{V}$ is then called a TTF *(torsion-torsion-free) class* and the triple $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ is said to be a TTF *triple*. Moreover, a TTF triple $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ is said to be *suspended* (respectively, *cosuspended*) if the corresponding TTF class is a suspended (respectively, cosuspended) subcategory of $\mathcal{T}$.

Note that, in a TTF triple, one of the torsion pairs is a t-structure if and only if the adjacent one is a co-t-structure.

Example 2.4. (1) Let A be a ring and consider its derived category $\mathsf{D}(A)$. Denote by $D^{\leq -1}$ (respectively, $D^{\geq 0}$) the subcategory of $\mathsf{D}(A)$ formed by the complexes whose usual complex cohomology vanishes in all nonnegative degrees (respectively, in all negative degrees). The pair $(D^{\leq -1}, D^{\geq 0})$ is a nondegenerate t-structure in $\mathsf{D}(A)$, called the *standard t-structure*. We note that the standard t-structure admits both a left and a right adjacent co-t-structure. We refer to [Angeleri Hügel et al. 2016, Example 2.9 (2)] for details on the left adjacent co-t-structure. Analogously, the right adjacent co-t-structure is the pair $(D^{\geq 0}, K_{\leq -1})$ where $K_{\leq -1}$ stands for the subcategory of objects in $\mathsf{D}(A)$ which are isomorphic to a complex $X^\bullet$ of injective A-modules such that $X^i = 0$ for all $i \geq 0$. The triple $(D^{\leq -1}, D^{\geq 0}, K_{\leq -1})$ is then a cosuspended TTF triple. Clearly, the heart of $(D^{\leq -1}, D^{\geq 0})$ is $\mathsf{Mod}(A)$ and the coheart of $(D^{\geq 0}, K_{\leq -1})$ coincides with $\mathsf{Inj}(A)$.

(2) It follows from [Aihara and Iyama 2012, Theorem 4.3] that if $\mathcal{A}$ is a set of compact objects, then the pair $({}^{\perp_0}(\mathcal{A}^{\perp_0}), \mathcal{A}^{\perp_0})$ is a torsion pair. If $\mathcal{T}$ is moreover an algebraic triangulated category, then such a pair admits a right adjacent torsion pair, as shown in [Šťovíček and Pospíšil 2016, Theorem 3.11]. In this case, if $\mathcal{A}$ is a suspended (respectively, cosuspended) subcategory of $\mathcal{T}^c$, then the triple $({}^{\perp_0}(\mathcal{A}^{\perp_0}), \mathcal{A}^{\perp_0}, (\mathcal{A}^{\perp_0})^{\perp_0})$ is a cosuspended (respectively, suspended) TTF triple. We investigate some properties of the heart of compactly generated cosuspended TTF triples in Section 4.

(3) Following the arguments in [Neeman 2010, Proposition 1.4], we have that if $\mathcal{V}$ is a cosuspended and preenveloping (respectively, suspended and precovering) subcategory of $\mathcal{T}$, then the inclusion of $\mathcal{V}$ in $\mathcal{T}$ has a left (respectively, right) adjoint. In particular, there is a t-structure $(\mathcal{U}, \mathcal{V})$ (respectively, a t-structure $(\mathcal{V}, \mathcal{W})$) in $\mathcal{T}$. In our context, this shows that a co-t-structure $(\mathcal{V}, \mathcal{W})$ has a left (respectively, right) adjacent t-structure if and only if $\mathcal{V}$ is preenveloping (respectively, $\mathcal{W}$ is precovering).

(Co)silting. Recall the definition of silting and cosilting objects in a triangulated category (see [Psaroudakis and Vitória 2015]):

Definition 2.5. An object M in $\mathcal{T}$ is called

- *silting* if $(M^{\perp_{>0}}, M^{\perp_{\leq 0}})$ is a t-structure in $\mathcal{T}$ and $M \in M^{\perp_{>0}}$;
- *cosilting* if $({}^{\perp_{\leq 0}}M, {}^{\perp_{>0}}M)$ is a t-structure in $\mathcal{T}$ and $M \in {}^{\perp_{>0}}M$.

We say that two silting (respectively, cosilting) objects are *equivalent*, if they give rise to the same t-structure in $\mathcal{T}$ and we call such a t-structure *silting* (respectively, *cosilting*). The heart of the t-structure associated to a silting or cosilting object M is denoted by $\mathcal{H}_M$ and the cohomological functor $\mathcal{T} \to \mathcal{H}_M$ by H^0_M.

It follows from the definition that silting and cosilting t-structures are nondegenerate and that a silting (respectively, cosilting) object generates (respectively, cogenerates) the triangulated category $\mathcal{T}$ (see [Psaroudakis and Vitória 2015]).

Example 2.6. Let A be a ring and $\mathsf{D}(A)$ its derived category.

(1) Let E be an injective cogenerator of $\mathsf{Mod}(A)$. Regarded as an object in $\mathsf{D}(A)$, E is a cosilting object and the associated cosilting t-structure is the standard one. As discussed in Example 2.4 (1), there is also a right adjacent co-t-structure with coheart $\mathsf{Prod}(E) = \mathsf{Inj}(A)$.

(2) It follows from [Angeleri Hügel et al. 2016, Theorem 4.6] that a silting object T of $\mathsf{D}(A)$ lying in $\mathsf{K}^b(\mathsf{Proj}(A))$ gives rise to a suspended TTF triple, that is, the silting t-structure $(T^{\perp_{>0}}, T^{\perp_{\leq 0}})$ admits a left adjacent co-t-structure with coheart $\mathsf{Add}(T)$ (see also [Wei 2013]). Dually, a cosilting object C of $\mathsf{D}(A)$ lying in $\mathsf{K}^b(\mathsf{Inj}(A))$ gives rise to a cosuspended TTF triple, that is, the cosilting t-structure $({}^{\perp_{\leq 0}}C, {}^{\perp_{>0}}C)$ admits a right adjacent co-t-structure with coheart $\mathsf{Prod}(C)$. For this dual statement, we refer to forthcoming work in [Marks and Vitória 2017].

Silting and cosilting objects produce hearts with particularly interesting homological properties. First, recall from [Parra and Saorín 2015] that hearts of t-structures in a triangulated category with products and coproducts also have products and coproducts. Indeed, the (co)product of a family of objects in the heart is obtained by applying the functor $H^0_{\mathbb{T}}$ to the corresponding (co)product of the same family in $\mathcal{T}$. Of course, this (co)product in the heart may differ from the (co)product formed in $\mathcal{T}$.

Lemma 2.7 [Psaroudakis and Vitória 2015, Proposition 4.3]. *Let M be a silting* (*respectively*, *cosilting*) *object in $\mathcal{T}$. Then the heart $\mathcal{H}_M$ is an abelian category with a projective generator* (*respectively*, *an injective cogenerator*) *given by $H^0_M(M)$.*

The following lemma establishes a particularly nice behaviour of the cohomological functors arising from (co)silting t-structures with respect to products and coproducts.

Lemma 2.8. *If T is a silting object in $\mathcal{T}$, then the functor H^0_T induces an equivalence between $\mathsf{Add}_{\mathcal{T}}(T)$ and $\mathsf{Add}_{\mathcal{H}_T}(H^0_T(T)) = \mathsf{Proj}(\mathcal{H}_T)$. Dually*, *if C is a cosilting object in $\mathcal{T}$, then the functor H^0_C induces an equivalence between $\mathsf{Prod}_{\mathcal{T}}(C)$ and $\mathsf{Prod}_{\mathcal{H}_C}(H^0_C(C)) = \mathsf{Inj}(\mathcal{H}_C)$.*

Proof. We prove the statement for a cosilting object C in $\mathcal{T}$ (the silting case is shown dually). Let the truncation functors of the associated cosilting t-structure $({}^{\perp_{\leq 0}}C, {}^{\perp_{>0}}C)$ be denoted by $u : \mathcal{T} \to {}^{\perp_{\leq 0}}C$ and $v : \mathcal{T} \to {}^{\perp_{>0}}C$. Recall that $\mathsf{Prod}_{\mathcal{T}}(C) = {}^{\perp_{>0}}C \cap ({}^{\perp_{>0}}C[-1])^{\perp_0}$ [Psaroudakis and Vitória 2015, Lemma 4.5(iii)].

We first show that H^0_C is fully faithful on $\mathsf{Prod}_{\mathcal{T}}(C)$ (compare with [Keller and Vossieck 1988, Lemma 5.1(d); Assem et al. 2008, Lemma 1.3; Nicolás et al. 2015, Lemma 3.2]). Let X_1 and X_2 be objects in $\mathsf{Prod}_{\mathcal{T}}(C)$. Suppose that $f : X_1 \to X_2$ is a map in $\mathcal{T}$ such that $H^0_C(f) = 0$. Since, by assumption, X_i $(i = 1, 2)$ lies in ${}^{\perp_{>0}}C$, there is a truncation triangle of the form

$$v(X_i[1])[-2] \xrightarrow{\kappa_i} H^0_C(X_i) \xrightarrow{\mu_i} X_i \longrightarrow v(X_i[1])[-1].$$

Now f induces a morphism of triangles and, in particular, we have that $0 = \mu_2 H^0_C(f) = f\mu_1$. Thus, f factors through $v(X_1[1])[-1]$. However, since X_2 lies in $({}^{\perp_{>0}}C[-1])^{\perp_0}$, we have that $\mathrm{Hom}_{\mathcal{T}}(v(X_1[1])[-1], X_2) = 0$ and, therefore, $f = 0$. Now let us show that H^0_C is also full on $\mathsf{Prod}_{\mathcal{T}}(C)$. Suppose that g is a map in $\mathrm{Hom}_{\mathcal{T}}(H^0_C(X_1), H^0_C(X_2))$. Since X_2 lies in $({}^{\perp_{>0}}C[-1])^{\perp_0}$, the composition $\mu_2 g \kappa_1$ vanishes and, therefore, there is a map $\tilde{g} : X_1 \to X_2$ such that $\tilde{g}\mu_1 = \mu_2 g$. Therefore, g extends to a morphism of triangles and, as a consequence, $g = H^0_C(\tilde{g})$.

It remains to show that the essential image of H^0_C restricted to $\mathsf{Prod}_{\mathcal{T}}(C)$ coincides with $\mathsf{Prod}_{\mathcal{H}_C}(H^0_C(C))$. Observe first that $H^0_C\big(\prod_{i\in I} X_i\big) = \prod_{i\in I} H^0_C(X_i)$ for every family $(X_i)_{i\in I}$ of objects in $\mathsf{Prod}_{\mathcal{T}}(C)$, where the product of the family $(H^0_C(X_i))_{i\in I}$ is taken in $\mathcal{H}_C$. The proof is dual to the argument for silting objects in [Nicolás et al. 2015, Lemma 3.2.2(a)]. Take an object M in $\mathsf{Prod}_{\mathcal{H}_C}(H^0_C(C))$ and let N be an object in $\mathcal{H}_C$ such that $M \oplus N = H^0_C(C)^I$ for some set I. Then there is an idempotent element e_M in $\mathsf{End}_{\mathcal{H}_C}(H^0_C(C)^I) = \mathsf{End}_{\mathcal{H}_C}(H^0_C(C^I))$ whose image is the summand M. Since H^0_C is fully faithful on $\mathsf{Prod}_{\mathcal{T}}(C)$, it follows that there is an idempotent element e in $\mathsf{End}_{\mathcal{T}}(C^I)$ such that $H^0_C(e) = e_M$. Given that $\mathcal{T}$ is idempotent complete, the map e factors as $C^I \xrightarrow{f} X \xrightarrow{g} C^I$ such that $fg = \mathrm{id}_X$, and it then follows that $H^0_C(X) = M$. ☐

We finish this section with a general observation on abelian categories that will be useful later.

Lemma 2.9. *Let $\mathcal{A}$ and $\mathcal{B}$ be abelian categories with enough injective* (*respectively*, *projective*) *objects and let $F : \mathcal{A} \to \mathcal{B}$ be a left* (*respectively*, *right*) *exact functor yielding an equivalence* $\mathsf{Inj}(\mathcal{A}) \to \mathsf{Inj}(\mathcal{B})$ (*respectively*, $\mathsf{Proj}(\mathcal{A}) \to \mathsf{Proj}(\mathcal{B})$). *Then F is an equivalence of abelian categories.*

Proof. Suppose that $\mathcal{A}$ and $\mathcal{B}$ have enough injective objects. Then both categories can be recovered as factor categories of the corresponding categories $\mathsf{Mor}(\mathsf{Inj}(\mathcal{A}))$ and $\mathsf{Mor}(\mathsf{Inj}(\mathcal{B}))$ of morphisms between injectives. Indeed, the kernel functors induce equivalences $\mathrm{Ker}_{\mathcal{A}} : \mathsf{Mor}(\mathsf{Inj}(\mathcal{A}))/\mathcal{R}_{\mathcal{A}} \to \mathcal{A}$ and $\mathrm{Ker}_{\mathcal{B}} : \mathsf{Mor}(\mathsf{Inj}(\mathcal{A}))/\mathcal{R}_{\mathcal{B}} \to \mathcal{B}$, where the relations $\mathcal{R}_{\mathcal{A}}$ and $\mathcal{R}_{\mathcal{B}}$ are the obvious ones (compare with [Auslander et al. 1995, Proposition IV.1.2] for the case of projectives). Since F induces an equivalence between $\mathsf{Inj}(\mathcal{A})$ and $\mathsf{Inj}(\mathcal{B})$, it clearly also induces an equivalence between the corresponding morphism categories and, moreover, since F is left exact, it indeed defines an equivalence $\widetilde{F} : \mathsf{Mor}(\mathsf{Inj}(\mathcal{A}))/\mathcal{R}_{\mathcal{A}} \to \mathsf{Mor}(\mathsf{Inj}(\mathcal{B}))/\mathcal{R}_{\mathcal{B}}$ such that $\mathrm{Ker}_{\mathcal{B}} \circ \widetilde{F} = F \circ \mathrm{Ker}_{\mathcal{A}}$. Hence, F is an equivalence. The dual statement follows analogously. □

3. Grothendieck hearts in compactly generated triangulated categories

Recall that a Grothendieck category is an abelian category with coproducts, exact direct limits and a generator. It is well known that Grothendieck categories have enough injective objects and every object admits an injective envelope. This section is dedicated to the question of determining when hearts of silting and cosilting t-structures are Grothendieck categories. We answer this question using a suitable category of functors and a corresponding theory of purity. We begin this section with a quick reminder of the relevant concepts.

Functors and purity. We consider the category $\mathsf{Mod}\text{-}\mathcal{T}^c$ of contravariant additive functors from $\mathcal{T}^c$ to $\mathsf{Mod}(\mathbb{Z})$, which is known to be a locally coherent Grothendieck category (see [Krause 1997; 2000, Subsection 1.2]).

Consider the restricted Yoneda functor

$$\boldsymbol{y} : \mathcal{T} \to \mathsf{Mod}\text{-}\mathcal{T}^c, \qquad \boldsymbol{y}X = \mathrm{Hom}_{\mathcal{T}}(-, X)_{|\mathcal{T}^c}, \quad \text{for all } X \in \mathcal{T}.$$

It is well known that $\boldsymbol{y}$ is not, in general, fully faithful. A triangle

$$\Delta : \; X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow X[1]$$

in $\mathcal{T}$ is said to be a *pure triangle* if $\boldsymbol{y}\Delta$ is a short exact sequence. In other words, the triangle Δ is pure, if for any compact object K in $\mathcal{T}$, the sequence

$$0 \longrightarrow \mathrm{Hom}_{\mathcal{T}}(K, X) \xrightarrow{\mathrm{Hom}_{\mathcal{T}}(K,f)} \mathrm{Hom}_{\mathcal{T}}(K, Y) \xrightarrow{\mathrm{Hom}_{\mathcal{T}}(K,g)} \mathrm{Hom}_{\mathcal{T}}(K, Z) \longrightarrow 0$$

is exact. We say that a morphism $f : X \to Y$ in $\mathcal{T}$ is a *pure monomorphism* (respectively, a *pure epimorphism*) if $\boldsymbol{y}f$ is a monomorphism (respectively, an epimorphism) in $\mathsf{Mod}\text{-}\mathcal{T}^c$. An object E of $\mathcal{T}$ is said to be *pure-injective* if any pure monomorphism $f : E \to Y$ in $\mathcal{T}$ splits. Similarly, an object P is said to be *pure-projective* in $\mathcal{T}$ if any pure epimorphism $g : X \to P$ splits.

The following theorem collects useful properties of pure-injective and pure-projective objects.

Theorem 3.1 [Krause 2000, Theorem 1.8, Corollary 1.9; Beligiannis 2000, §11]. *The following statements are equivalent for an object E in $\mathcal{T}$:*

(1) *E is pure-injective.*

(2) *$\boldsymbol{y}E$ is an injective object in $\mathsf{Mod}\text{-}\mathcal{T}^c$.*

(3) *The map $\mathrm{Hom}_{\mathcal{T}}(X, E) \to \mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}X, \boldsymbol{y}E)$, $\phi \mapsto \boldsymbol{y}\phi$ is an isomorphism for any object X in $\mathcal{T}$.*

(4) *For every set I, the summation map $E^{(I)} \to E$ factors through the canonical map $E^{(I)} \to E^I$.*

Dually, the following are equivalent for an object P in $\mathcal{T}$:

(1) *P is pure-projective.*

(2) *$\boldsymbol{y}P$ is a projective object in $\mathsf{Mod}\text{-}\mathcal{T}^c$.*

(3) *The map $\mathrm{Hom}_{\mathcal{T}}(P, Y) \to \mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}P, \boldsymbol{y}Y)$, $\phi \mapsto \boldsymbol{y}\phi$ is an isomorphism for any object Y in $\mathcal{T}$.*

(4) *P lies in $\mathsf{Add}(\mathcal{T}^c)$.*

Moreover, any projective (respectively, injective) object in $\mathsf{Mod}\text{-}\mathcal{T}^c$ is of the form $\boldsymbol{y}P$ (respectively, $\boldsymbol{y}E$), for a pure-projective object P (respectively, a pure-injective object E), uniquely determined up to isomorphism.

It follows from above that $\mathcal{T}$ has enough pure-injective objects and that every object X in $\mathcal{T}$ admits a pure-injective envelope. The following theorem collects two results that will become essential later on.

Theorem 3.2. *Let $H : \mathcal{T} \to \mathcal{A}$ be a cohomological functor from $\mathcal{T}$ to an abelian category $\mathcal{A}$.*

(1) [Beligiannis 2000, Theorem 3.4] *If H sends pure triangles in $\mathcal{T}$ to short exact sequences in $\mathcal{A}$, then there is a unique exact functor $\overline{H} : \mathsf{Mod}\text{-}\mathcal{T}^c \to \mathcal{A}$ such that $\overline{H} \circ \boldsymbol{y} = H$.*

(2) [Krause 2000, Corollary 2.5] *If $\mathcal{A}$ has exact direct limits and H preserves coproducts, then H sends pure triangles in $\mathcal{T}$ to short exact sequences in $\mathcal{A}$.*

We recall from [Beligiannis 2000] how to construct $\overline{H}$. Given F in $\mathsf{Mod}\text{-}\mathcal{T}^c$, consider an injective copresentation

$$0 \longrightarrow F \longrightarrow \boldsymbol{y}E_0 \xrightarrow{\boldsymbol{y}\alpha} \boldsymbol{y}E_1,$$

where E_0 and E_1 are pure-injective in $\mathcal{T}$ and α is a map in $\mathrm{Hom}_{\mathcal{T}}(E_0, E_1)$. Then we define $\overline{H}(F) := \mathrm{Ker}\, H(\alpha)$, and it can be checked that $\overline{H}$ is indeed well defined (that is, it does not depend on the choice of the injective copresentation of F). This functor can also be obtained in a dual way by taking a projective presentation of F.

Grothendieck hearts and purity. Note that, in general, the cohomological functor associated to a t-structure does not commute with products and coproducts in $\mathcal{T}$. The following lemma provides necessary and sufficient conditions for this to happen.

Lemma 3.3. *Let $\mathbb{T} = (\mathcal{U}, \mathcal{V})$ be a nondegenerate t-structure in $\mathcal{T}$ with heart $\mathcal{H}_{\mathbb{T}}$ and associated cohomological functor $H^0_{\mathbb{T}} : \mathcal{T} \to \mathcal{H}_{\mathbb{T}}$. Then the functor $H^0_{\mathbb{T}}$ preserves $\mathcal{T}$-coproducts (respectively, $\mathcal{T}$-products) if and only if $\mathcal{V}$ is closed under coproducts (respectively, $\mathcal{U}$ is closed under products).*

If these conditions are satisfied, we say $\mathbb{T}$ is smashing (respectively, cosmashing).

Proof. We prove the statement for coproducts; the statement for products follows dually. Notice that aisles are always closed under coproducts. If also the coaisle $\mathcal{V}$ is closed under coproducts, then both truncation functors $u : \mathcal{T} \to \mathcal{U}$ and $v : \mathcal{T} \to \mathcal{V}$ commute with $\mathcal{T}$-coproducts and, hence, so does $H^0_{\mathbb{T}}$. In particular, coproducts in $\mathcal{H}_{\mathbb{T}}$ coincide with coproducts in $\mathcal{T}$. For the converse, it is easy to check that nondegenerate t-structures can be cohomologically described, i.e., $\mathcal{V}$ can be described as the subcategory formed by objects X such that $H^0_{\mathbb{T}}(X[k]) = 0$ for all $k < 0$. Consequently, since $H^0_{\mathbb{T}}$ commutes with $\mathcal{T}$-coproducts, this description shows that $\mathcal{V}$ is closed under coproducts. □

Example 3.4. (1) By definition, every silting t-structure is cosmashing and every cosilting t-structure is smashing.

(2) If a silting object T is pure-projective, then the associated t-structure is smashing. Indeed, let $(X_i)_{i\in I}$ be a family of objects in $T^{\perp_{<0}}$ and let X be their coproduct in $\mathcal{T}$. Since T is pure-projective,

$$\mathrm{Hom}_{\mathcal{T}}(T, X[n]) \cong \mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}T, \boldsymbol{y}X[n])$$

for all n in $\mathbb{Z}$. The statement then follows from the fact that $\boldsymbol{y}$ commutes with coproducts and $\mathrm{Ker}\,\mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}T, -)$ is coproduct-closed.

(3) If a cosilting object C is pure-injective, in general, it does not follow that the associated t-structure is cosmashing. Indeed, let A be the Kronecker algebra and let C be the Reiten–Ringel cotilting module from [Reiten and Ringel 2006, Proposition 10.1] with associated torsion pair $(\mathcal{Q}, \mathsf{Cogen}(C))$ in $\mathsf{Mod}(A)$, where $\mathcal{Q}$ is the class of all modules generated by preinjective A-modules. The object C is cosilting in $D(A)$ (see [Šťovíček 2014, Theorem 4.5]). Note that, since C is pure-injective in $\mathsf{Mod}(A)$ by [Bazzoni 2003], it follows from Theorem 3.1 that C is also pure-injective when viewed as an object in $D(A)$. It turns out that the aisle of the associated cosilting t-structure consists precisely of those complexes whose zeroth cohomology belongs to $\mathcal{Q}$ and for which all positive cohomologies vanish (compare with [Happel et al. 1996]). In particular, the cosilting t-structure is cosmashing if and only if $\mathcal{Q}$ is closed under products in $\mathsf{Mod}(A)$. But the latter cannot be true due to [Angeleri Hügel 2003, Theorem 5.2 and Example 5.4].

For a compactly generated triangulated category $\mathcal{T}$, (co)silting t-structures can be obtained in a rather abstract way. First, recall that $\mathcal{T}$ satisfies a *Brown representability theorem* (i.e., every cohomological functor $H : \mathcal{T}^{op} \to \mathsf{Mod}(\mathbb{Z})$ which sends coproducts to products is representable) and a *dual Brown representability theorem* (i.e., every cohomological functor $H : \mathcal{T} \to \mathsf{Mod}(\mathbb{Z})$ which sends products to products is representable); see [Krause 2002a] for details. We can now state the following result:

Theorem 3.5 [Nicolás et al. 2015, §4]. *There is a bijection between*

- *cosmashing nondegenerate t-structures whose heart has a projective generator*;
- *equivalence classes of silting objects.*

Dually, *there is a bijection between*

- *smashing nondegenerate t-structures whose heart has an injective cogenerator*;
- *equivalence classes of cosilting objects.*

The first statement is proven in [Nicolás et al. 2015]. For the reader's convenience, we briefly sketch an argument for the second bijection. First recall that cosilting t-structures are smashing, nondegenerate and their hearts have injective cogenerators (see Lemma 2.7). Hence, there is an injective assignment from equivalence classes of cosilting objects to the t-structures with the assigned properties. To see that the assignment is surjective, we use the fact that $\mathcal{T}$ satisfies Brown representability. Indeed, given a smashing nondegenerate t-structure $\mathbb{T}$ whose heart has injective cogenerator E, the corresponding cosilting object C can be obtained as the (unique) representative of the cohomological functor $\mathrm{Hom}_{\mathcal{T}}(H^0_{\mathbb{T}}(-), E) \cong \mathrm{Hom}_{\mathcal{T}}(-, C)$. Note that $\mathrm{Hom}_{\mathcal{T}}(H^0_{\mathbb{T}}(-), E)$ sends coproducts to products by the smashing assumption. The dual arguments were used in [Nicolás et al. 2015] to show the silting case.

We can now prove the main result of this section by building on Theorem 3.5 and identifying which (co)silting t-structures have Grothendieck hearts. A similar result was obtained independently in [Nicolás et al. 2015, Proposition 4.2] with the additional assumption that all t-structures considered are cosmashing.

Theorem 3.6. *Let* $\mathbb{T} = (\mathcal{U}, \mathcal{V})$ *be a smashing nondegenerate t-structure in* $\mathcal{T}$ *with heart* $\mathcal{H}_{\mathbb{T}}$. *Denote by* $H^0_{\mathbb{T}} : \mathcal{T} \to \mathcal{H}_{\mathbb{T}}$ *the associated cohomological functor. The following statements are equivalent.*

(1) $\mathcal{H}_{\mathbb{T}}$ *is a Grothendieck category*;

(2) *There is a pure-injective cosilting object* C *in* $\mathcal{T}$ *such that* $\mathbb{T} = ({}^{\perp_{\leq 0}}C, {}^{\perp_{>0}}C)$.

If the above conditions are satisfied, *there is a* (*unique*) *exact functor* $\overline{H}^0_{\mathbb{T}} : \mathsf{Mod}\text{-}\mathcal{T}^c \to \mathcal{H}_{\mathbb{T}}$ *with a right adjoint* j_* *such that* $\overline{H}^0_{\mathbb{T}} \circ \boldsymbol{y} = H^0_{\mathbb{T}}$ *and* $j_* H^0_{\mathbb{T}}(C) \cong \boldsymbol{y}C$. *Moreover*, *there is a localisation sequence of the form*

$$\begin{tikzcd} \mathrm{Ker}\, \overline{H}^0_{\mathbb{T}} = {}^{\perp_0}\boldsymbol{y}C \arrow[r, "i_*"] & \mathsf{Mod}\text{-}\mathcal{T}^c \arrow[l, bend left, "i^!"] \arrow[r, "\overline{H}^0_{\mathbb{T}}"] & \mathcal{H}_{\mathbb{T}} \arrow[l, bend left, "j_*"] \end{tikzcd}$$

Proof. Suppose that $\mathcal{H}_{\mathbb{T}}$ is a Grothendieck category. By Theorem 3.5, $\mathbb{T}$ is a cosilting t-structure for a cosilting object C, such that $\mathrm{Hom}_{\mathcal{T}}(H^0_{\mathbb{T}}(-), E) \cong \mathrm{Hom}_{\mathcal{T}}(-, C)$ for some injective cogenerator E in $\mathcal{H}_{\mathbb{T}}$. It remains to show that C is pure-injective. Since, by Lemma 3.3, $H^0_{\mathbb{T}}$ commutes with $\mathcal{T}$-coproducts, Theorem 3.2(2) shows that

$H^0_{\mathbb{T}}$ sends pure triangles to short exact sequences. In particular, $\mathrm{Hom}_{\mathcal{T}}(-, C)$ sends pure triangles to short exact sequences, showing that C is indeed pure-injective.

Conversely, let C be a pure-injective cosilting object in $\mathcal{T}$ with associated t-structure $\mathbb{T} = ({}^{\perp_{\leq 0}}C, {}^{\perp_{>0}}C)$. It follows that the functor $\mathrm{Hom}_{\mathcal{T}}(-, C)$ is naturally equivalent to $\mathrm{Hom}_{\mathcal{T}}(H^0_{\mathbb{T}}(-), H^0_{\mathbb{T}}(C))$ and, therefore, also the functor $H^0_{\mathbb{T}}$ sends pure triangles to short exact sequences. Consequently, by Theorem 3.2(1), there is a (unique) exact functor $\overline{H}^0_{\mathbb{T}} : \mathsf{Mod}\text{-}\mathcal{T}^c \to \mathcal{H}_{\mathbb{T}}$ such that $\overline{H}^0_{\mathbb{T}} \circ \boldsymbol{y} = H^0_{\mathbb{T}}$. The following argument is inspired by the proof of [Šťovíček 2014, Theorem 6.2]. Consider the hereditary torsion pair in $\mathsf{Mod}\text{-}\mathcal{T}^c$ cogenerated by the injective object $\boldsymbol{y}C$, i.e., the pair $({}^{\perp_0}\boldsymbol{y}C, \mathsf{Cogen}(\boldsymbol{y}C))$. The quotient category $\mathcal{G}_C := \mathsf{Mod}\text{-}\mathcal{T}^c/{}^{\perp_0}\boldsymbol{y}C$ is a Grothendieck category (see [Gabriel 1962, Proposition III.9]) and the quotient functor $\pi : \mathsf{Mod}\text{-}\mathcal{T}^c \to \mathcal{G}_C$ admits a fully faithful right adjoint functor $\rho : \mathcal{G}_C \to \mathsf{Mod}\text{-}\mathcal{T}^c$, with essential image

$$\mathsf{Cogen}(\boldsymbol{y}C) \cap \mathrm{Ker}\,\mathrm{Ext}^1_{\mathsf{Mod}\text{-}\mathcal{T}^c}({}^{\perp_0}\boldsymbol{y}C, -)$$

(see [Gabriel 1962, Corollary of Proposition III.3; Prest 2009, §11.1.1]). In particular, as in the proof of [Šťovíček 2014, Theorem 6.2], it follows that an object X of $\mathcal{G}_C$ is injective if and only if $\rho(X)$ lies in $\mathsf{Prod}(\boldsymbol{y}C)$, i.e., the full subcategory of injective objects in $\mathcal{G}_C$ is equivalent to $\mathsf{Prod}(\boldsymbol{y}C)$ which, by Theorem 3.1, is further equivalent to $\mathsf{Prod}(C)$. Thus, using Lemma 2.8, we get the following commutative diagram of equivalences:

$$\begin{tikzcd}
\mathsf{Inj}(\mathcal{G}_C) \arrow[r, "\rho"] & \mathsf{Prod}(\boldsymbol{y}C) \arrow[d, "\overline{H}^0_{\mathbb{T}}"'] & \mathsf{Prod}(C) \arrow[l, "\boldsymbol{y}"'] \arrow[dl, "H^0_{\mathbb{T}}"] \\
& \mathsf{Prod}(H^0_{\mathbb{T}}(C)) &
\end{tikzcd}$$

Hence, the functor $\overline{H}^0_{\mathbb{T}} \circ \rho$ yields an equivalence between the category of injective objects in $\mathcal{G}_C$ and the category of injective objects in $\mathcal{H}_{\mathbb{T}}$. Since the functor $\overline{H}^0_{\mathbb{T}} \circ \rho$ is clearly left exact, by Lemma 2.9, it extends to an equivalence of categories $\mathcal{G}_C \cong \mathcal{H}_{\mathbb{T}}$ showing, in particular, that $\mathcal{H}_{\mathbb{T}}$ is a Grothendieck category.

Assume now that $\mathbb{T}$ satisfies (1) and (2). We first show that $\mathrm{Ker}\,\overline{H}^0_{\mathbb{T}} = {}^{\perp_0}\boldsymbol{y}C$. Indeed, if F is an object in ${}^{\perp_0}\boldsymbol{y}C$, then $\mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}\alpha, \boldsymbol{y}C)$ is an epimorphism for any map $\boldsymbol{y}\alpha : \boldsymbol{y}E_0 \to \boldsymbol{y}E_1$ between injective objects in $\mathsf{Mod}\text{-}\mathcal{T}^c$ with $\mathrm{Ker}(\boldsymbol{y}\alpha) = F$. Using the pure-injectivity of C, we get that $\mathrm{Hom}_{\mathcal{T}}(\alpha, C)$ is an epimorphism and, thus, so is $\mathrm{Hom}_{\mathcal{H}_{\mathbb{T}}}(H^0_{\mathbb{T}}(\alpha), H^0_{\mathbb{T}}(C))$. Since $H^0_{\mathbb{T}}(C)$ is an injective cogenerator of $\mathcal{H}_{\mathbb{T}}$, it follows that $H^0_{\mathbb{T}}(\alpha)$ is a monomorphism and, thus, $\overline{H}^0_{\mathbb{T}}(F) = 0$, by the construction of $\overline{H}^0_{\mathbb{T}}$. Finally, since this argument is reversible the desired equality holds.

Now, in order to show the existence of the localisation sequence above, it is enough to prove that the functor $\overline{H}^0_{\mathbb{T}}$ admits a right adjoint. To this end, since $\overline{H}^0_{\mathbb{T}} \circ \rho$ is an equivalence and π has a right adjoint, it suffices to check that $\overline{H}^0_{\mathbb{T}} \cong \overline{H}^0_{\mathbb{T}} \circ \rho \circ \pi$.

By using the unit of the adjunction (π, ρ), we get a natural transformation of functors $\overline{H}^0_{\mathbb{T}} \to \overline{H}^0_{\mathbb{T}} \circ \rho \circ \pi$. We need to see that it induces an isomorphism on objects. But this follows from the fact that the kernel and cokernel of the natural map $X \to \rho\pi(X)$, for X in $\mathsf{Mod}\text{-}\mathcal{T}^c$, are torsion, that is, they belong to ${}^{\perp_0}\boldsymbol{y}C = \operatorname{Ker} \overline{H}^0_{\mathbb{T}}$. Finally, by using the adjunction $(\overline{H}^0_{\mathbb{T}}, j_*)$, we get $j_* H^0_{\mathbb{T}}(C) \cong j_* \overline{H}^0_{\mathbb{T}}(\boldsymbol{y}C) \cong \boldsymbol{y}C$. □

One can state a somewhat dual result for silting objects.

Theorem 3.7. *Let $\mathbb{T} = (\mathbb{T}^{\leq 0}, \mathbb{T}^{\geq 0})$ be a smashing and cosmashing nondegenerate t-structure in $\mathcal{T}$ with heart $\mathcal{H}_{\mathbb{T}}$. Denote by $H^0_{\mathbb{T}} : \mathcal{T} \to \mathcal{H}_{\mathbb{T}}$ the associated cohomological functor. The following are equivalent.*

(1) *$\mathcal{H}_{\mathbb{T}}$ is a Grothendieck category with a projective generator*;

(2) *There is a pure-projective silting object T in $\mathcal{T}$ such that $\mathbb{T} = (T^{\perp_{>0}}, T^{\perp_{\leq 0}})$.*

If the above conditions are satisfied, there is a (unique) exact functor $\overline{H}^0_{\mathbb{T}} : \mathsf{Mod}\text{-}\mathcal{T}^c \to \mathcal{H}_{\mathbb{T}}$ with a left adjoint $j_!$ such that $\overline{H}^0_{\mathbb{T}} \circ \boldsymbol{y} = H^0_{\mathbb{T}}$ and $j_! H^0_{\mathbb{T}}(T) \cong \boldsymbol{y}T$. Moreover, there is a recollement of the form

$$\begin{tikzcd}[column sep=huge] \operatorname{Ker} \overline{H}^0_{\mathbb{T}} \arrow[r, "i_*"] & \mathsf{Mod}\text{-}\mathcal{T}^c \arrow[l, bend right=40, "i^*"'] \arrow[l, bend left=40, "i^!"] \arrow[r, "\overline{H}^0_{\mathbb{T}}"] & \mathcal{H}_{\mathbb{T}}. \arrow[l, bend right=40, "j_!"'] \arrow[l, bend left=40, "j_*"] \end{tikzcd}$$

Proof. The arguments are dual to those in the proof of Theorem 3.6. Note that the additional assumption of the t-structure being smashing comes into play through the use of Theorem 3.2(2), which is needed in an essential way to prove the pure-projectivity of the associated silting object. On the other hand, we have seen in Example 3.4(2) that the t-structure is smashing whenever T is a pure-projective silting object. Finally, observe that we get a recollement rather than just a localisation sequence like in Theorem 3.6, since, in the given context, $\operatorname{Ker} \overline{H}^0_{\mathbb{T}}$ is closed under products and coproducts in $\mathsf{Mod}\text{-}\mathcal{T}^c$ (see also [Psaroudakis and Vitória 2014, Corollary 4.4]). □

As an immediate consequence of these results, we can identify the t-structures with Grothendieck hearts within the bijections of Theorem 3.5.

Corollary 3.8. *There is a bijection between*

- *smashing nondegenerate t-structures of $\mathcal{T}$ whose heart is a Grothendieck category*;
- *equivalence classes of pure-injective cosilting objects.*

Dually, there is a bijection between

- *smashing and cosmashing nondegenerate t-structures in $\mathcal{T}$ whose heart is a Grothendieck category with a projective generator*;
- *equivalence classes of pure-projective silting objects.*

4. Cosuspended TTF classes

In this section, we focus on cosuspended TTF classes in a compactly generated triangulated category $\mathcal{T}$. We relate the properties of the previous section (namely, Grothendieck hearts and the pure-injectivity of the associated cosilting objects) with the definability of the cosuspended TTF class. As a consequence, if $\mathcal{T}$ is algebraic, nondegenerate compactly generated t-structures have Grothendieck hearts.

Coherent functors and definability. We begin with a short reminder on coherent functors and definable subcategories of $\mathcal{T}$, and we obtain an easy (but useful) criterion to check whether a certain subcategory of $\mathcal{T}$ is definable or not. We also prove that a definable subcategory $\mathcal{V}$ of $\mathcal{T}$ is preenveloping, i.e., for any object X in $\mathcal{T}$ there is a map $\phi : X \to V$ with V in $\mathcal{V}$ such that $\mathrm{Hom}_{\mathcal{T}}(\phi, V')$ is surjective for all V' in $\mathcal{V}$.

Recall from [Krause 2002b, Proposition 5.1] that a covariant additive functor $F : \mathcal{T} \to \mathsf{Mod}(\mathbb{Z})$ is said to be *coherent* if the following equivalent conditions are satisfied:

(1) There are compact objects K and L and a presentation

$$\mathrm{Hom}_{\mathcal{T}}(K, -) \to \mathrm{Hom}_{\mathcal{T}}(L, -) \to F \to 0.$$

(2) F preserves products and coproducts and sends pure triangles to short exact sequences.

The category of coherent functors is denoted by $\mathsf{Coh}\text{-}\mathcal{T}$. For a locally coherent Grothendieck category $\mathcal{G}$ or, more generally, a locally finitely presented additive category with products, coherent functors are defined analogously, replacing in (1) the compactness of K and L by the property of being finitely presented. The analogue of (2) then states that coherent functors are precisely the functors $\mathcal{G} \to \mathsf{Mod}(\mathbb{Z})$ preserving products and direct limits [Krause 2003, Proposition 3.2].

Definition 4.1. A subcategory $\mathcal{V}$ of $\mathcal{T}$ is said to be *definable* if there is a set of coherent functors $(F_i)_{i\in I}$ from $\mathcal{T}$ to $\mathsf{Mod}(\mathbb{Z})$ such that X lies in $\mathcal{V}$ if and only if $F_i(X) = 0$ for all i in I.

Definable subcategories of a locally finitely presented additive category $\mathcal{G}$ with products are defined as above: they are zero-sets of families of coherent functors. Recall that a subcategory of $\mathcal{G}$ is definable if and only if it is closed under products, direct limits and pure subobjects [Krause 2003, Theorem 2]. Moreover, definable subcategories of $\mathcal{G}$ are closed under pure-injective envelopes (see [Prest 2009, §16.1.2]). Note that, by definition, definable subcategories of $\mathcal{T}$ are also closed under products, coproducts, pure subobjects and pure quotients, but we do not know whether they are characterised by such closure conditions (unless stronger assumptions are imposed, see [Krause 2002b, Theorem 7.5]). A useful criterion for definability in $\mathcal{T}$ will be provided in Corollary 4.4 below.

For a subcategory $\mathcal{V}$ of a compactly generated triangulated (respectively, a locally coherent Grothendieck) category, we denote by $\mathsf{Def}(\mathcal{V})$ the smallest definable subcategory containing $\mathcal{V}$.

Example 4.2. A notion of flatness in $\mathsf{Mod}\text{-}\mathcal{T}^c$ is developed in [Krause 2000, §2.3] and [Beligiannis 2000, §8.1]. The subcategory $\mathsf{Flat}\text{-}\mathcal{T}^c$ of flat objects in $\mathsf{Mod}\text{-}\mathcal{T}^c$ is locally finitely presented and contains precisely the functors F that send triangles to exact sequences or, equivalently, that satisfy $\mathrm{Ext}^1(G, F) = 0$ for all finitely presented functors G in $\mathsf{Mod}\text{-}\mathcal{T}^c$. Moreover, $\mathsf{Flat}\text{-}\mathcal{T}^c$ is a definable subcategory of $\mathsf{Mod}\text{-}\mathcal{T}^c$ by [Prest 2009, Theorem 16.1.12]. Note that all objects of the form $\boldsymbol{y}X$, for X in $\mathcal{T}$, are flat.

The definable closure $\mathsf{Def}(\mathcal{V})$ in $\mathsf{Mod}\text{-}\mathcal{T}^c$ of a set $\mathcal{V}$ of objects contained in $\mathsf{Flat}\text{-}\mathcal{T}^c$ consists of pure subobjects of direct limits in $\mathsf{Mod}\text{-}\mathcal{T}^c$ of directed systems whose terms are products of objects in $\mathcal{V}$. To prove this, one uses the notion of a reduced product from [Krause 1998, p. 465]. Since $\mathsf{Flat}\text{-}\mathcal{T}^c$ is a definable subcategory of $\mathsf{Mod}\text{-}\mathcal{T}^c$, it suffices to show that the pure subobjects of reduced products of objects in $\mathcal{V}$ form a definable subcategory of $\mathsf{Flat}\text{-}\mathcal{T}^c$. But the latter statement follows from [Krause 1998, Corollary 4.10] combined with [Krause 1998, Proposition 2.2].

We have this the following useful fact (compare with [Arnesen et al. 2016, Theorem 1.9]):

Proposition 4.3. *Let $\mathsf{fun}(\mathsf{Flat}\text{-}\mathcal{T}^c)$ denote the category of coherent functors from the locally finitely presented category $\mathsf{Flat}\text{-}\mathcal{T}^c$ to $\mathsf{Mod}(\mathbb{Z})$. Then the assignment $\mathsf{fun}(\mathsf{Flat}\text{-}\mathcal{T}^c) \to \mathsf{Coh}\text{-}\mathcal{T}$ that sends a functor F to $F \circ \boldsymbol{y}$ is an equivalence of categories.*

Proof. First, we observe that the assignment is well defined. It is clear that given F in $\mathsf{fun}(\mathsf{Flat}\text{-}\mathcal{T}^c)$, the composition $F \circ \boldsymbol{y}$ preserves products and coproducts. Now, given a pure triangle Δ in $\mathcal{T}$, we have that $\boldsymbol{y}(\Delta)$ is a short exact sequence in $\mathsf{Flat}\text{-}\mathcal{T}^c$. Since short exact sequences in $\mathsf{Flat}\text{-}\mathcal{T}^c$ are pure or, equivalently, direct limits of split exact sequences (see [Prest 2009, Theorem 16.1.15]), we see that $F(\boldsymbol{y}(\Delta))$ is a short exact sequence of abelian groups. It then follows that $F \circ \boldsymbol{y}$ is coherent by the description (2) of coherent functors above.

In order to see that this assignment yields an equivalence of categories we show that it admits a quasi-inverse. By [Krause 2002b, Proposition 4.1], each functor G in $\mathsf{Coh}\text{-}\mathcal{T}$ gives rise to a unique functor $\overline{G}$ in $\mathsf{fun}(\mathsf{Flat}\text{-}\mathcal{T}^c)$ such that $\overline{G} \circ \boldsymbol{y} = G$. The uniqueness guarantees the functoriality of this assignment and it is clear that the assignments are inverse to each other. □

As a corollary of the proposition above, we deduce the following statement.

Corollary 4.4. *Let $\mathcal{V}$ be a class of objects in $\mathcal{T}$. The smallest definable subcategory of $\mathcal{T}$ containing $\mathcal{V}$ is*

$$\mathsf{Def}(\mathcal{V}) = \{X \in \mathcal{T} : \boldsymbol{y}X \in \mathsf{Def}(\boldsymbol{y}\mathcal{V})\}.$$

As a consequence, any definable subcategory of $\mathcal{T}$ is closed under pure-injective envelopes.

Recall from [Crivei et al. 2010, Theorem 4.1] that any definable subcategory of a locally finitely presented additive category $\mathcal{G}$ with products is preenveloping. The following proposition establishes a triangulated analogue. Its proof is inspired by the proof of [Aihara and Iyama 2012, Theorem 4.3].

Proposition 4.5. *Let $\mathcal{V}$ be a definable subcategory of $\mathcal{T}$. Then $\mathcal{V}$ is preenveloping. In particular, if $\mathcal{V}$ is cosuspended, then $({}^{\perp_0}\mathcal{V}, \mathcal{V})$ is a t-structure.*

Proof. Since $\mathcal{V}$ is definable, by definition, there is a set of maps $\{\phi_i : X_i \to Y_i \mid i \in I\}$ in $\mathcal{T}^c$ such that an object V in $\mathcal{T}$ lies in $\mathcal{V}$ if and only if $\mathrm{Hom}_{\mathcal{T}}(\phi_i, V)$ is surjective for all i in I. We need to build a $\mathcal{V}$-preenvelope for a given object $Z = Z_0$ in $\mathcal{T}$. First, setting $K_{i,0} := \mathrm{Hom}_{\mathcal{T}}(X_i, Z_0)$, we define the map

$$\overline{X}_0 := \bigoplus_{i \in I} X_i^{(K_{i,0})} \xrightarrow{\bar{\phi}_0 := \bigoplus_{i\in I} \phi_i^{(K_{i,0})}} \overline{Y}_0 := \bigoplus_{i \in I} Y_i^{(K_{i,0})}$$

and consider the canonical universal map $a_0 : \overline{X}_0 \to Z_0$. Let $z_0 : Z_0 \to Z_1$ denote the corresponding component of the cone of the map $(\bar{\phi}_0, -a_0)^T : \overline{X}_0 \to \overline{Y}_0 \oplus Z_0$, and proceed inductively to define objects Z_n and maps $z_n : Z_n \to Z_{n+1}$. We prove that the Milnor colimit V_Z of the inductive system $(Z_n, z_n)_{n\in\mathbb{N}_0}$ yields a $\mathcal{V}$-approximation of Z. Let us first observe that V_Z indeed lies in $\mathcal{V}$. Since both X_i and Y_i are compact for any i in I, it follows that

$$\mathrm{Hom}_{\mathcal{T}}(\phi_i, V_Z) \cong \varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(\phi_i, Z_n) : \varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(Y_i, Z_n) \to \varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(X_i, Z_n).$$

In order to see that this map is surjective, it suffices to show that any element in $\varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(X_i, Z_n)$ which is represented by a map g in $\mathrm{Hom}_{\mathcal{T}}(X_i, Z_m)$ for some m in $\mathbb{N}_0$ lies in the image of $\mathrm{Hom}_{\mathcal{T}}(\phi_i, V_Z)$. By construction of the inductive system, there clearly is a map h in $\mathrm{Hom}_{\mathcal{T}}(Y_i, Z_{m+1})$ such that $h\phi_i = z_m g$. As a consequence, the element in $\varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(Y_i, Z_n)$ represented by the map h is a preimage via $\varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(\phi_i, Z_n)$ of the element in $\varinjlim_{n\in\mathbb{N}_0} \mathrm{Hom}_{\mathcal{T}}(X_i, Z_n)$ that we started with. This proves that $\mathrm{Hom}_{\mathcal{T}}(\phi_i, V_Z)$ is surjective for all i in I and, thus, V_Z lies in $\mathcal{V}$.

We proceed to prove that the induced map $v : Z \to V_Z$ is a left $\mathcal{V}$-approximation. Given a morphism $f : Z \to V$ with V in $\mathcal{V}$, the composition fa_0 factors through $\bar{\phi}_0$. Let $\bar{f} : \overline{Y}_0 \to V$ be such that $\bar{f}\bar{\phi}_0 = fa_0$. By construction of Z_1 as the cone of $(\bar{\phi}_0, -a_0)^T$, it follows that there is a map $f_1 : Z_1 \to V$ such that $f = f_1 z_0$.

Inductively, one can then see that the map f indeed factors successively through any Z_n and, therefore, through the Milnor colimit V_Z, as wanted.

The final statement follows from Example 2.4(3). □

Cosuspended TTF triples. Before we discuss how definability arises in the context of cosuspended TTF triples, we first prove some auxiliary statements.

Lemma 4.6. *Let $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ be a cosuspended TTF triple in $\mathcal{T}$. Then $(\mathcal{U}, \mathcal{V})$ is a nondegenerate t-structure if and only if the coheart $\mathscr{C} := \mathcal{V} \cap \mathcal{W}[-1]$ cogenerates $\mathcal{T}$. In this case, we have $\mathcal{V} = {}^{\perp_{>0}}\mathscr{C}$.*

Proof. Let $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ be a cosuspended TTF triple in $\mathcal{T}$. Suppose that $(\mathcal{U}, \mathcal{V})$ is a nondegenerate t-structure and let X be an object of $\mathcal{T}$ such that $\mathrm{Hom}_{\mathcal{T}}(X, \mathscr{C}[k]) = 0$, for all k in $\mathbb{Z}$. Given an integer k in $\mathbb{Z}$, let us denote by $u^k : \mathcal{T} \to \mathcal{U}[k]$ and $v^k : \mathcal{T} \to \mathcal{V}[k]$ the truncation functors corresponding to the t-structure $(\mathcal{U}[k], \mathcal{V}[k])$. Consider a truncation triangle of the object $v^k(X)$ for the co-t-structure $(\mathcal{V}[k-1], \mathcal{W}[k-1])$, yielding a diagram of the form

$$\begin{tikzcd} u^k(X) \arrow[r] & X \arrow[r, "f^k"] & v^k(X) \arrow[r] \arrow[d, equal] & u^k(X)[1] & \\ & V_{k-1} \arrow[r, "g"] & v^k(X) \arrow[r, "h"] & W_{k-1} \arrow[r] & V_{k-1}[1] \end{tikzcd}$$

with V_{k-1} in $\mathcal{V}[k-1]$ and W_{k-1} in $\mathcal{W}[k-1]$. We can easily deduce that W_{k-1} lies in $\mathscr{C}[k]$ and, thus, $hf^k = 0$ by assumption on X. Then there is a morphism $\alpha : X \to V_{k-1}$ such that $g\alpha = f^k$. Now, since V_{k-1} lies in $\mathcal{V}[k-1] \subseteq \mathcal{V}[k]$, it follows that α factors through the truncation $f^{k-1} : X \to v^{k-1}(X)$. This then yields a map $v^{k-1}(X) \to v^k(X)$. Considering the two compositions of this map with the canonical morphism $v^k(X) \to v^{k-1}(X)$ and using the minimality of the maps f^k and f^{k-1}, we conclude that both maps are isomorphisms. Since this holds for arbitrary k, the nondegeneracy of $(\mathcal{U}, \mathcal{V})$ implies that $v^k(X) = 0$ for all k in $\mathbb{Z}$. Thus, X must lie in $\bigcap_{n\in\mathbb{Z}} \mathcal{U}[n]$ and, again by the nondegeneracy of $(\mathcal{U}, \mathcal{V})$ it must, therefore, be zero.

Conversely, suppose that $\mathscr{C}$ cogenerates $\mathcal{T}$ and let X lie in $\bigcap_{n\in\mathbb{Z}} \mathcal{U}[n]$. Consider a morphism $f : X \to C[k]$ for k in $\mathbb{Z}$ and C in $\mathscr{C}$. Now, since $C[k]$ lies in $\mathcal{V}[k]$ and X lies in $\mathcal{U}[k]$, it follows that $f = 0$ and, thus, by assumption, also $X = 0$. Similarly, if X is in $\bigcap_{n\in\mathbb{Z}} \mathcal{V}[n]$, since $C[k]$ lies in $\mathcal{W}[k-1]$, it must follow that $X = 0$.

Finally, assuming that $\mathscr{C}$ cogenerates $\mathcal{T}$, we show that $\mathcal{V} = {}^{\perp_{>0}}\mathscr{C}$. It is always the case that $\mathcal{V} \subseteq {}^{\perp_{>0}}\mathscr{C}$. For the reverse inclusion, let X be an object in ${}^{\perp_{>0}}\mathscr{C}$ and consider the truncation triangle

$$v(X)[-1] \to u(X) \to X \to v(X).$$

Given C in $\mathscr{C}$ and applying the functor $\mathrm{Hom}_{\mathcal{T}}(-, C[k])$ to the triangle, we see that $\mathrm{Hom}_{\mathcal{T}}(v(X), C[k+1]) = \mathrm{Hom}_{\mathcal{T}}(X, C[k]) = 0$ for all $k > 0$ and, thus, we have $\mathrm{Hom}_{\mathcal{T}}(u(X), C[k]) = 0$ for all $k > 0$. Moreover, since $\mathscr{C}[k] \subset \mathcal{V}$ for all $k \leq 0$, we see that $\mathrm{Hom}_{\mathcal{T}}(u(X), C[k]) = 0$ for all $k \leq 0$. Since $\mathscr{C}$ cogenerates $\mathcal{T}$, we have that $u(X) = 0$ and X belongs to $\mathcal{V}$, as wanted. □

Lemma 4.7. *Let $\mathscr{C}$ be subcategory of $\mathcal{T}$. Suppose that $\mathscr{C}$ is closed under products and summands, and that all objects in $\mathscr{C}$ are pure-injective. Then there is an object C in $\mathscr{C}$ such that $\mathscr{C} = \mathsf{Prod}(C)$.*

Proof. Consider the hereditary torsion pair $({}^{\perp_0}(\boldsymbol{y}\mathscr{C}), \mathcal{F} := \mathsf{Cogen}(\boldsymbol{y}\mathscr{C}))$ in $\mathsf{Mod}\text{-}\mathcal{T}^c$. It is well known that there is an injective object $\boldsymbol{y}C$ in $\mathsf{Mod}\text{-}\mathcal{T}^c$ such that $\mathcal{F} = \mathsf{Cogen}(\boldsymbol{y}C)$ (see [Stenström 1975, VI, Proposition 3.7]). It follows that $\mathsf{Prod}(\boldsymbol{y}\mathscr{C}) = \mathsf{Prod}(\boldsymbol{y}C)$. Since $\boldsymbol{y}$ commutes with products and is fully faithful on pure-injectives, we get $\mathscr{C} = \mathsf{Prod}(\mathscr{C}) = \mathsf{Prod}(C)$. □

Lemma 4.8. *Let $\mathscr{C}$ be an additive subcategory of $\mathcal{T}$ and $\mathcal{V} = {}^{\perp_{>0}}\mathscr{C}$. Then the following statements are equivalent*:

(1) *$\mathcal{V}$ is product-closed and every object in $\mathscr{C}$ is pure-injective.*

(2) *$\mathcal{V}$ is definable.*

Moreover, if the above conditions are satisfied, then there is a t-structure $(\mathcal{U}, \mathcal{V})$.

Proof. Suppose that (1) holds. We have to show that every object X in $\mathsf{Def}(\mathcal{V})$ lies in $\mathcal{V} = {}^{\perp_{>0}}\mathscr{C}$. By Corollary 4.4, the object $\boldsymbol{y}X$ lies in the definable closure $\mathsf{Def}(\boldsymbol{y}\mathcal{V})$ in $\mathsf{Mod}\text{-}\mathcal{T}^c$ of all objects of the form $\boldsymbol{y}V$ with V in $\mathcal{V}$. By the description of the definable closure given in Example 4.2, $\boldsymbol{y}X$ is a pure subobject of a direct limit in $\mathsf{Mod}\text{-}\mathcal{T}^c$ of a directed system whose objects are products of the form $\prod_{i\in I} \boldsymbol{y}X_i$, with X_i in $\mathcal{V}$. Note that, since $\boldsymbol{y}$ commutes with products, we have $\prod_{i\in I} \boldsymbol{y}X_i = \boldsymbol{y}X_I$, where $X_I = \prod_{i\in I} X_i$. Now, since $\mathcal{V}$ is closed under products, X_I lies in $\mathcal{V}$. Applying the functor $\mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(-, \boldsymbol{y}C[k])$, with $k > 0$ and C in $\mathscr{C}$ to the embedding $\boldsymbol{y}X \to \varinjlim_I \boldsymbol{y}X_I$, and using the pure-injectivity of C (and its shifts), we get an epimorphism

$$\begin{aligned}\varprojlim_I \mathrm{Hom}_{\mathcal{T}}(X_I, C[k]) \cong \mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\varinjlim_I \boldsymbol{y}X_I, \boldsymbol{y}C[k]) &\twoheadrightarrow \mathrm{Hom}_{\mathsf{Mod}\text{-}\mathcal{T}^c}(\boldsymbol{y}X, \boldsymbol{y}C[k]) \\ &\cong \mathrm{Hom}_{\mathcal{T}}(X, C[k]).\end{aligned}$$

Since X_I lies in $\mathcal{V}$, the domain of this map vanishes and, hence, so does the codomain, as wanted.

Conversely, suppose that $\mathcal{V}$ is definable. First, the subcategory $\mathcal{V}$ is closed under products. Let X be an object in $\mathscr{C}$ and $f : X \to I(X)$ its pure-injective envelope in $\mathcal{T}$. Since definable subcategories are closed under pure-injective envelopes and pure quotients (see Corollary 4.4), it follows that both $I(X)$ and $Z := \mathsf{cone}(f)$ lie in $\mathcal{V}$. Since $\mathrm{Hom}_{\mathcal{T}}(\mathcal{V}, \mathscr{C}[1]) = 0$ it follows that $\mathrm{Hom}_{\mathcal{T}}(Z, X[1]) = 0$, the triangle induced by f splits and X is a summand of $I(X)$, i.e., X is pure-injective.

The last statement of the lemma follows from Proposition 4.5, since $\mathcal{V}$ is clearly cosuspended. □

Finally, we can now use the rather technical statements above to prove the main theorem of this section.

Theorem 4.9. *Let $(\mathcal{U}, \mathcal{V}, \mathcal{W})$ be a cosuspended TTF triple in $\mathcal{T}$ such that the t-structure $(\mathcal{U}, \mathcal{V})$ is nondegenerate. Then the following are equivalent*:

(1) *$\mathcal{V}$ is definable in $\mathcal{T}$*;

(2) *$\mathcal{V} = {}^{\perp_{>0}}C$ for a pure-injective cosilting object C in $\mathcal{T}$.*

Proof. First observe that by Lemma 4.6, the coheart $\mathscr{C} = \mathcal{V} \cap \mathcal{W}[-1]$ cogenerates $\mathcal{T}$ and $\mathcal{V} = {}^{\perp_{>0}}\mathscr{C}$. Since $\mathcal{V}$ is a TTF class, it is closed under products and, therefore, by Lemma 4.8, $\mathcal{V}$ is definable if and only if every object in $\mathscr{C}$ is pure-injective. In that case, since both $\mathcal{V}$ and $\mathcal{W}$ (and, thus, $\mathscr{C}$) are closed under products and summands, by Lemma 4.7, there is C in $\mathcal{T}$ such that $\mathscr{C} = \mathsf{Prod}(C)$.

(1)⇒ (2): Suppose now that $\mathcal{V}$ is definable and let C be as above. As observed, we have that $\mathcal{V} = {}^{\perp_{>0}}C$ and we only need to show that $\mathcal{U} = {}^{\perp_{\leq 0}}C$. Since $C[k]$ lies in $\mathcal{V}$ for all $k \leq 0$, it is clear that $\mathcal{U} \subseteq {}^{\perp_{\leq 0}}C$. Now let X lie in ${}^{\perp_{\leq 0}}C$ and consider a truncation triangle

$$u(X) \to X \to v(X) \to u(X)[1]$$

with $u(X)$ in $\mathcal{U}$ and $v(X)$ in $\mathcal{V}$. Since both X and $u(X)[1]$ lie in ${}^{\perp_{\leq 0}}C$, so does $v(X)$. However, $v(X)$ also lies in ${}^{\perp_{>0}}C$, showing that $v(X) = 0$, since C is a cogenerator of $\mathcal{T}$. Hence, we have that $\mathcal{U} = {}^{\perp_{\leq 0}}C$.

(2)⇒ (1): In order to show that $\mathcal{V}$ is definable, it is enough to show that the coheart $\mathscr{C}$ coincides with $\mathsf{Prod}(C)$ (thus proving that every object in $\mathscr{C}$ is pure-injective). The argument is dual to the one used in [Angeleri Hügel et al. 2016, Lemma 4.5]. Indeed, let X be an object in $\mathscr{C}$, let I denote the set $\mathrm{Hom}_{\mathcal{T}}(X, C)$ and consider the induced universal map $\phi : X \to C^I$. If Z denotes the cone of the map ϕ, then it is easy to check that Z lies in ${}^{\perp_{>0}}C$ and, thus, the map $Z \to X[1]$ of the induced triangle is zero, by the assumption on X. Hence, the triangle splits and X lies in $\mathsf{Prod}(C)$. This shows that $\mathscr{C} \subseteq \mathsf{Prod}(C)$ and the reverse inclusion is clear. □

Corollary 4.10. *Let $\mathcal{T}$ be an algebraic, compactly generated triangulated category. Every nondegenerate compactly generated t-structure has a Grothendieck heart.*

Proof. From Example 2.4(2), every compactly generated t-structure $(\mathcal{U}, \mathcal{V})$ in $\mathcal{T}$ admits a right adjacent co-t-structure $(\mathcal{V}, \mathcal{W})$. It is clear that $\mathcal{V}$ is definable as it is the subcategory obtained as the intersection of the kernels of the coherent functors

$\mathrm{Hom}_{\mathcal{T}}(X,-)$ with X in $\mathcal{U}\cap\mathcal{T}^c$. Now, Theorem 4.9 combined with Theorem 3.6 finishes the proof. □

The corollary above extends [Bravo and Parra 2016, Corollary 2.5], which treats the special case when $\mathcal{T}$ is a derived module category and the compactly generated t-structure arises as an HRS-tilt of a torsion pair in the underlying module category.

Acknowledgements

The authors would like to thank Mike Prest for discussions leading to Corollary 4.4. Marks was supported by a grant within the DAAD P.R.I.M.E. program. Vitória acknowledges support from the Department of Computer Sciences of the University of Verona in the earlier part of this project, as well as from the Engineering and Physical Sciences Research Council of the United Kingdom, grant number EP/N016505/1, in the later part of this project. Finally, the authors acknowledge funding from the Project "Ricerca di Base 2015" of the University of Verona.

References

[Aihara and Iyama 2012] T. Aihara and O. Iyama, "Silting mutation in triangulated categories", *J. Lond. Math. Soc.* (2) **85**:3 (2012), 633–668. MR Zbl

[Angeleri Hügel 2003] L. Angeleri Hügel, "Covers and envelopes via endoproperties of modules", *Proc. London Math. Soc.* (3) **86**:3 (2003), 649–665. MR Zbl

[Angeleri Hügel et al. 2016] L. Angeleri Hügel, F. Marks, and J. Vitória, "Silting modules", *Int. Math. Res. Not.* **2016**:4 (2016), 1251–1284. MR Zbl

[Arnesen et al. 2016] K. K. Arnesen, R. Laking, D. Pauksztello, and M. Prest, "The Ziegler spectrum for derived-discrete algebras", 2016. To appear in *Adv. Math.* arXiv

[Assem et al. 2008] I. Assem, M. J. Souto Salorio, and S. Trepode, "Ext-projectives in suspended subcategories", *J. Pure Appl. Algebra* **212**:2 (2008), 423–434. MR Zbl

[Auslander et al. 1995] M. Auslander, I. Reiten, and S. O. Smalø, *Representation theory of Artin algebras*, Cambridge Studies in Advanced Mathematics **36**, Cambridge Univ. Press, 1995. MR Zbl

[Bazzoni 2003] S. Bazzoni, "Cotilting modules are pure-injective", *Proc. Amer. Math. Soc.* **131**:12 (2003), 3665–3672. MR Zbl

[Beligiannis 2000] A. Beligiannis, "Relative homological algebra and purity in triangulated categories", *J. Algebra* **227**:1 (2000), 268–361. MR Zbl

[Beĭlinson et al. 1982] A. A. Beĭlinson, J. Bernstein, and P. Deligne, "Faisceaux pervers", pp. 5–171 in *Analysis and topology on singular spaces, I* (Luminy, 1981), Astérisque **100**, Société Mathématique de France, Paris, 1982. MR Zbl

[Bondarko 2010] M. V. Bondarko, "Weight structures vs. t-structures: weight filtrations, spectral sequences, and complexes (for motives and in general)", *J. K-Theory* **6**:3 (2010), 387–504. MR Zbl

[Bravo and Parra 2016] D. Bravo and C. E. Parra, "tCG torsion pairs", preprint, 2016. arXiv

[Crivei et al. 2010] S. Crivei, M. Prest, and B. Torrecillas, "Covers in finitely accessible categories", *Proc. Amer. Math. Soc.* **138**:4 (2010), 1213–1221. MR Zbl

[Gabriel 1962] P. Gabriel, "Des catégories abéliennes", *Bull. Soc. Math. France* **90** (1962), 323–448. MR Zbl

[Happel 1988] D. Happel, *Triangulated categories in the representation theory of finite-dimensional algebras*, London Mathematical Society Lecture Note Series **119**, Cambridge Univ. Press, 1988. MR Zbl

[Happel et al. 1996] D. Happel, I. Reiten, and S. O. Smalø, *Tilting in abelian categories and quasitilted algebras*, Mem. Amer. Math. Soc. **575**, 1996. MR Zbl

[Iyama and Yoshino 2008] O. Iyama and Y. Yoshino, "Mutation in triangulated categories and rigid Cohen–Macaulay modules", *Invent. Math.* **172**:1 (2008), 117–168. MR Zbl

[Keller 1994] B. Keller, "Deriving DG categories", *Ann. Sci. École Norm. Sup.* (4) **27**:1 (1994), 63–102. MR Zbl

[Keller and Vossieck 1988] B. Keller and D. Vossieck, "Aisles in derived categories", *Bull. Soc. Math. Belg. Sér. A* **40**:2 (1988), 239–253. MR Zbl

[Krause 1997] H. Krause, "The spectrum of a locally coherent category", *J. Pure Appl. Algebra* **114**:3 (1997), 259–271. MR Zbl

[Krause 1998] H. Krause, "Exactly definable categories", *J. Algebra* **201**:2 (1998), 456–492. MR Zbl

[Krause 2000] H. Krause, "Smashing subcategories and the telescope conjecture: an algebraic approach", *Invent. Math.* **139**:1 (2000), 99–133. MR Zbl

[Krause 2002a] H. Krause, "A Brown representability theorem via coherent functors", *Topology* **41**:4 (2002), 853–861. MR Zbl

[Krause 2002b] H. Krause, "Coherent functors in stable homotopy theory", *Fund. Math.* **173**:1 (2002), 33–56. MR Zbl

[Krause 2003] H. Krause, "Coherent functors and covariantly finite subcategories", *Algebr. Represent. Theory* **6**:5 (2003), 475–499. MR Zbl

[Marks and Vitória 2017] F. Marks and J. Vitória, "Silting and cosilting classes in derived categories", preprint, 2017. arXiv

[Mendoza Hernández et al. 2013] O. Mendoza Hernández, E. C. Sáenz Valadez, V. Santiago Vargas, and M. J. Souto Salorio, "Auslander–Buchweitz context and co-t-structures", *Appl. Categ. Structures* **21**:5 (2013), 417–440. MR Zbl

[Neeman 2001] A. Neeman, *Triangulated categories*, Annals of Mathematics Studies **148**, Princeton Univ. Press, 2001. MR Zbl

[Neeman 2010] A. Neeman, "Some adjoints in homotopy categories", *Ann. of Math.* (2) **171**:3 (2010), 2143–2155. MR Zbl

[Nicolás et al. 2015] P. Nicolás, M. Saorín, and A. Zvonareva, "Silting theory in triangulated categories with coproducts", preprint, 2015. arXiv

[Parra and Saorín 2015] C. E. Parra and M. Saorín, "Direct limits in the heart of a t-structure: the case of a torsion pair", *J. Pure Appl. Algebra* **219**:9 (2015), 4117–4143. MR Zbl

[Pauksztello 2008] D. Pauksztello, "Compact corigid objects in triangulated categories and co-t-structures", *Cent. Eur. J. Math.* **6**:1 (2008), 25–42. MR Zbl

[Prest 2009] M. Prest, *Purity, spectra and localisation*, Encyclopedia of Mathematics and its Applications **121**, Cambridge Univ. Press, 2009. MR Zbl

[Psaroudakis and Vitória 2014] C. Psaroudakis and J. Vitória, "Recollements of module categories", *Appl. Categ. Structures* **22**:4 (2014), 579–593. MR Zbl

[Psaroudakis and Vitória 2015] C. Psaroudakis and J. Vitória, "Realisation functors in tilting theory", 2015. To appear in *Math. Zeit.* arXiv

[Reiten and Ringel 2006] I. Reiten and C. M. Ringel, "Infinite dimensional representations of canonical algebras", *Canad. J. Math.* **58**:1 (2006), 180–224. MR Zbl

[Stenström 1975] B. Stenström, *Rings of quotients: an introduction to methods of ring theory*, Die Grundlehren Math. Wissenschaften **217**, Springer, New York, 1975. MR Zbl

[Šťovíček 2014] J. Šťovíček, "Derived equivalences induced by big cotilting modules", *Adv. Math.* **263** (2014), 45–87. MR Zbl

[Šťovíček and Pospíšil 2016] J. Šťovíček and D. Pospíšil, "On compactly generated torsion pairs and the classification of co-t-structures for commutative Noetherian rings", *Trans. Amer. Math. Soc.* **368**:9 (2016), 6325–6361. MR Zbl

[Wei 2013] J. Wei, "Semi-tilting complexes", *Israel J. Math.* **194**:2 (2013), 871–893. MR Zbl

Received November 24, 2016. Revised March 22, 2017.

LIDIA ANGELERI HÜGEL
DIPARTIMENTO DI INFORMATICA, SETTORE DI MATEMATICA
UNIVERSITÀ DEGLI STUDI DI VERONA
STRADA LE GRAZIE 15
CA' VIGNAL 2
I-37134 VERONA
ITALY
lidia.angeleri@univr.it

FREDERIK MARKS
INSTITUT FÜR ALGEBRA UND ZAHLENTHEORIE
UNIVERSITÄT STUTTGART
PFAFFENWALDRING 57
D-70569 STUTTGART
GERMANY
marks@mathematik.uni-stuttgart.de

JORGE VITÓRIA
DEPARTMENT OF MATHEMATICS
CITY UNIVERSITY OF LONDON
NORTHAMPTON SQUARE
EC1V 0HB
UNITED KINGDOM
jorge.vitoria@city.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.279

TRANSFINITE DIAMETER ON COMPLEX ALGEBRAIC VARIETIES

DAVID A. COX AND SIONE MA'U

We use methods from computational algebraic geometry to study Chebyshev constants and the transfinite diameter of a pure m-dimensional affine algebraic variety in $\mathbb{C}^n$ ($m \le n$). The main result is a generalization of Zaharyuta's integral formula for the Fekete–Leja transfinite diameter.

1. Introduction

This paper studies a notion of transfinite diameter on a pure m-dimensional algebraic subvariety of $\mathbb{C}^n$, $1 \le m \le n$. This is a natural generalization of the Fekete–Leja transfinite diameter in $\mathbb{C}^n$, which is an important quantity in pluripotential theory and polynomial approximation. In the study of the Fekete–Leja transfinite diameter in $\mathbb{C}^n$ ($n > 1$), an important paper is that of Zaharyuta [1975]. Given a compact set $K \subseteq \mathbb{C}^n$, Zaharyuta showed that its Fekete–Leja transfinite diameter, denoted $d(K)$, was given by a well-defined limiting process analogous to the one-dimensional case. The main result of that paper is an integral formula that realizes $d(K)$ as a "geometric average" of so-called *directional Chebyshev constants* associated to K; these constants measure (in an asymptotic sense) the minimum size on K of polynomials with prescribed leading terms.

Further developments and generalizations make use of the essential techniques in that paper. In [Jędrzejowski 1991] the notion of *homogeneous* transfinite diameter was studied and a Zaharyuta-type formula proved. In [Rumely and Lau 1994], and later in [Rumely et al. 2000], Lau, Rumely and Varley developed Zaharyuta's techniques in the setting of arithmetic geometry to study the notion of *sectional capacity*. More recently, Bloom and Levenberg [2003; 2010] studied a notion of *weighted* transfinite diameter in $\mathbb{C}^n$.

In [Baleikorocau and Ma'u 2015] a notion of transfinite diameter was defined and studied on an algebraic curve $V \subseteq \mathbb{C}^n$. It was shown that Zaharyuta's arguments, which exploit standard algebraic properties of polynomials, may be adapted to handle algebraic computations in the coordinate ring of V. Well-developed methods

MSC2010: primary 32U20; secondary 14Q15.
Keywords: Chebyshev constant, transfinite diameter, Vandermonde determinant, affine variety, Noether normalization, monomial order.

exist to carry out such computations, using Gröbner bases. In this paper we will apply these methods to higher dimensional algebraic varieties.

We should mention here that the notion of transfinite diameter on algebraic varieties may be studied as a by-product of Berman and Boucksom's [2010] general theory of Monge–Ampère energy on compact complex manifolds. A generalization of Zaharyuta's result to this setting has been proved in [Witt Nyström 2014]. The Berman–Boucksom methods are quite different to those of this paper. The connection between their setting and ours is explored in [Ma‘u 2017].

Before we describe the contents of the paper more specifically, we briefly recall the definition of the Fekete–Leja transfinite diameter.

Let $\{z^{\alpha_j}\}_{j=1}^{\infty}$ be the monomials in n variables listed according to a *graded order* (i.e., $|\alpha_j| \le |\alpha_k|$ whenever $j < k$). Here we are using standard multi-index notation: if $\alpha_j = (\alpha_{j1}, \dots, \alpha_{jn}) \subseteq \mathbb{Z}^n_{\ge 0}$, then $z^{\alpha_j} = z_1^{\alpha_{j1}} z_2^{\alpha_{j2}} \cdots z_n^{\alpha_{jn}}$ and $|\alpha_j| = \alpha_{j1} + \cdots + \alpha_{jn}$ denotes the total degree. Write $\boldsymbol{e}_j = z^{\alpha_j}$; so for $a = (a_1, \dots, a_n) \in \mathbb{C}^n$ we have $\boldsymbol{e}_j(a) = a_1^{\alpha_{j1}} \cdots a_n^{\alpha_{jn}}$. Given a positive integer M and points $\{\zeta_1, \dots, \zeta_M\} \subseteq \mathbb{C}^n$, the $M \times M$ determinant

$$\text{(1-1)} \quad \operatorname{Van}(\zeta_1, \dots, \zeta_M) = \det\bigl(\boldsymbol{e}_j(\zeta_i)\bigr)_{i,j=1}^{M} = \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{e}_2(\zeta_1) & \boldsymbol{e}_2(\zeta_2) & \cdots & \boldsymbol{e}_2(\zeta_M) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{e}_M(\zeta_1) & \boldsymbol{e}_M(\zeta_2) & \cdots & \boldsymbol{e}_M(\zeta_M) \end{pmatrix}$$

is called a *Vandermonde determinant* of order M. (Note that $\boldsymbol{e}_1 = 1$.)

Let $K \subseteq \mathbb{C}^n$ be compact and s a positive integer. Let m_s be the number of monomials of degree at most s in n variables, and let $l_s = \sum_{j=1}^{m_s} |\alpha_j|$ be the sum of the degrees. Define the s-th order diameter of K by

$$\text{(1-2)} \qquad d_s(K) := \sup\{|\operatorname{Van}(\zeta_1, \dots, \zeta_{m_s})|^{1/l_s} : \{\zeta_1, \dots, \zeta_{m_s}\} \subseteq K\}.$$

The Fekete–Leja transfinite diameter of K is defined as $d(K) := \limsup_{s\to\infty} d_s(K)$.

In this paper, we construct a basis $\mathcal{C}$ of polynomials for the coordinate ring $\mathbb{C}[V]$ of a pure m-dimensional algebraic variety $V \subseteq \mathbb{C}^n$ $(1 \le m \le n)$ of degree d, as long as the ring satisfies certain algebraic conditions (see (3-1)). Write $\mathcal{C} = \{\boldsymbol{e}_j\}_{j=1}^{\infty}$ for this basis which we assume is listed in a graded ordering: $\deg(\boldsymbol{e}_j) \le \deg(\boldsymbol{e}_k)$ if $j < k$. We define $\operatorname{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_M)$ to be the Vandermonde determinant with respect to $\mathcal{C}$ using the formula (1-1).

Define $m_s = m_s(V)$ to be the number of elements of $\mathcal{C}$ of degree at most s, and let $l_s = l_s(V) = \sum_{j=1}^{m_s} \deg(\boldsymbol{e}_j)$ be the sum of the degrees. The s-th order diameter of a compact set $K \subseteq V$ is defined as in (1-2) with $\operatorname{Van}_{\mathcal{C}}(\cdot)$ replacing $\operatorname{Van}(\cdot)$ on the right-hand side. Our main theorem (Theorem 6.2) says the following.

Theorem. *The limit* $d(K) := \lim_{s\to\infty} d_s(K)$ *exists and*

$$d(K) = \left(\prod_{j=1}^{d} T(K, \lambda_j)\right)^{1/d}.$$

Following Zaharyuta's terminology, the quantities $T(K, \lambda_j)$ on the right-hand side are called *principal Chebyshev constants* and are defined in Section 5 as integral averages of so-called *directional Chebyshev constants*. Here d is the *degree* of V and the λ_j are the d points of intersection of the projective closure of V in $\mathbb{P}^n$ with a certain subspace of the hyperplane at infinity. When V is a curve the above result is in [Baleikorocau and Ma'u 2015].* When $\deg(V) = 1$ then there is only one principal Chebyshev constant, and one recovers Zaharyuta's formula, up to a normalization.

In Section 2 we give some of the background needed for subsequent sections, including Noether normalization, the grevlex monomial ordering, normal forms and Hilbert functions.

In Section 3 we construct a basis (denoted by $\mathcal{C}$) of polynomials on the variety. The basis $\mathcal{C}$ consists of d groups of polynomials associated to the Noether normalization (elements of the form $(**)$, see Proposition 3.9), together with a "smaller" collection of monomials (elements of the form $(*)$). When V is a hypersurface, the basis $\mathcal{C}$ can be computed rather explicitly.

Section 4 is a general study of *weakly submultiplicative functions*. In [Bloom and Levenberg 2003] it was observed that Zaharyuta's computations with polynomials can be reformulated abstractly as properties of submultiplicative functions. We verify here that the relevant calculations go through with small modifications under slightly weaker conditions.

In Section 5, directional and principal Chebyshev constants are defined and studied. The main point is to construct weakly submultiplicative functions using computational properties of the basis $\mathcal{C}$ (Corollary 5.4). The results of Sections 3 and 4 can then be applied to this setting.

In Section 6 we prove the main theorem relating transfinite diameter to Chebyshev constants. The standard argument, based on estimating ratios of Vandermonde determinants with directional Chebyshev constants, goes through in its entirety.

In Section 7, we show in Theorem 7.2 that the transfinite diameter may be computed using the standard basis of monomials on the variety (i.e., those monomials that give normal forms). This uses the fact that, up to a geometric factor in some finite set—the collection of the $\boldsymbol{v}_i$ in Proposition 3.9—each polynomial in the basis $\mathcal{C}$ is a monomial.

*The principal Chebyshev constants in this paper are called directional Chebyshev constants in [Baleikorocau and Ma'u 2015]; for a one-dimensional curve, the λ_j may be interpreted as the directions of its linear asymptotes.

In the Appendix we compare our method to that of Rumely, Lau and Varley [Rumely et al. 2000], whose so-called *monic basis* is constructed by generating basis elements multiplicatively from a finite collection of polynomials with prescribed behavior. We compare both methods concretely in the case of the complexified sphere in $\mathbb{C}^3$.

2. Background material

We begin with Noether normalization. Consider an ideal $I \subseteq \mathbb{C}[z_1, \dots, z_n]$ with the following properties:

(1) $\mathbb{C}[z_1, \dots, z_m] \cap I = \{0\}$;

(2) For each $i = m+1, \dots, n$ there is a $g_i \in I$ which can be written in the form

$$g_i = z_i^{d_i} + \sum_{j=0}^{d_i - 1} h_{ij}(z_1, \dots, z_{i-1}) z_i^j, \quad \text{with } \deg(h_{ij}) + j \le d_i \text{ for all } i. \tag{2-1}$$

Property (1) is equivalent to saying that the map $\mathbb{C}[z_1, \dots, z_m] \to \mathbb{C}[z_1, \dots, z_n]/I$, induced by the inclusion into $\mathbb{C}[z_1, \dots, z_n]$, is injective, and property (2) implies that the quotient is finite over $\mathbb{C}[z_1, \dots, z_m]$. The Noether normalization theorem says that one can always make a change of variables so that the above properties hold. We state a specialized version of this theorem (cf., [Greuel and Pfister 2002, Theorem 3.4.1]).

Theorem 2.1 (Noether Normalization). *Let $J \subseteq \mathbb{C}[x_1, \dots, x_n]$ be an ideal. Then there is a positive integer $m \le n$ and a complex linear change of coordinates $z = T(x)$, $z_i = \sum_{j=1}^n T_{ij} x_j$, such that the following properties hold* (*write $I = T(J)$*):

(1) *The map $\mathbb{C}[z_1, \dots, z_m] \to \mathbb{C}[z_1, \dots, z_n]/I$ induced by inclusion is injective, and exhibits $\mathbb{C}[z_1, \dots, z_n]/I$ as a finite $\mathbb{C}$-algebra over $\mathbb{C}[z_1, \dots, z_m]$.*

(2) *For $i = m+1, \dots, n$, we can find polynomials $g_i \in I$ that satisfy* (2-1).

When property (1) of the theorem holds, we write $\mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[z_1, \dots, z_n]/I$. This inclusion is called a *Noether normalization*. All Noether normalizations used in this paper will be assumed to satisfy the additional condition (2) of the theorem since the degree condition in (2-1) will be important.

The grevlex ordering, which we will denote here by $<_{gr}$, is the ordering defined on $\mathbb{Z}^n_{\ge 0}$ by $\alpha <_{gr} \beta$ if:

(1) $|\alpha| < |\beta|$, or,

(2) $|\alpha| = |\beta|$, and for some $i \in \{1, \dots, n\}$ we have $\alpha_i < \beta_i$ and $\alpha_j = \beta_j$, for all $j < i$.

Define grevlex on monomials by putting $z^\alpha <_{gr} z^\beta$ if $\alpha <_{gr} \beta$. More precisely, this gives the grevlex ordering with $z_1 <_{gr} z_2 <_{gr} \cdots <_{gr} z_n$. Note that $|\alpha| < |\beta|$

implies $z^{\alpha} <_{gr} z^{\beta}$. A monomial ordering that satisfies this property is called a *graded ordering*.

Denote by $\mathrm{LT}(p)$ the leading term of a polynomial with respect to grevlex, and for an ideal I put $\mathrm{LT}(I) := \{\mathrm{LT}(p) : p \in I\}$. It is well known that for each element of $\mathbb{C}[z_1, \dots, z_n]/I$ there is a unique polynomial representative, the *normal form (with respect to grevlex)*, which contains no monomials in the ideal $\langle \mathrm{LT}(I) \rangle$. If an element of $\mathbb{C}[z_1, \dots, z_n]/I$ contains the polynomial p, then the normal form r may be computed in practice as the remainder on dividing p by a Gröbner basis of I; see [Cox et al. 1997, §5.3].

Write $\mathbb{C}[z]_I = \mathbb{C}[z_1, \dots, z_n]_I$ for the collection of normal forms of elements of $\mathbb{C}[z]/I = \mathbb{C}[z_1, \dots, z_n]/I$. As a vector space, $\mathbb{C}[z]_I$ has a basis consisting of all monomials $z^{\gamma} \notin \langle \mathrm{LT}(I) \rangle$. We can give $\mathbb{C}[z]_I$ the structure of an algebra over $\mathbb{C}$ with multiplication operation given by

$$(r_1, r_2) \mapsto \text{“the normal form of } r_1 r_2\text{”}.$$

We will usually denote this by $r_1 r_2$, though we will write $r_1 {*} r_2$ when we want to emphasize that this is the normal form of the ordinary product. Note that $\mathbb{C}[z]_I$ and $\mathbb{C}[z]/I$ are isomorphic as $\mathbb{C}$-algebras, where the isomorphism is given by identifying normal forms with their polynomial classes.

Hilbert functions play an important role in some of our proofs. We begin with $\mathbb{C}[z]_{\le s} = \mathbb{C}[z_1, \dots, z_n]_{\le s}$, which consists of polynomials of degree $\le s$. Recall that

$$\dim \mathbb{C}[z_1, \dots, z_n]_{\le s} = \binom{s+n}{n} = \frac{(s+n)\cdots(s+1)}{n!} = \frac{1}{n!} s^n + O(s^{n-1}). \tag{2-2}$$

Then $(\mathbb{C}[z]/I)_{\le s}$ consists of all classes represented by a polynomial of degree $\le s$. The dimension $\dim(\mathbb{C}[z]/I)_{\le s}$ gives the Hilbert function of I. We also define $\mathbb{C}[z]_{I \le s}$ to consist of all normal forms of degree $\le s$. Since $<_{gr}$ is a graded order, the isomorphism $\mathbb{C}[z]_I \simeq \mathbb{C}[z]/I$ induces an isomorphism

$$\mathbb{C}[z]_{I \le s} \simeq (\mathbb{C}[z]/I)_{\le s};$$

see [Cox et al. 1997, §9.3]. This has two useful consequences:

- The Hilbert function $\dim(\mathbb{C}[z]/I)_{\le s}$ is given by the number of monomials $z^{\gamma} \notin \langle \mathrm{LT}(I) \rangle$ of degree $\le s$.
- If $r_1 \in \mathbb{C}[z]_{I \le s}$ and $r_2 \in \mathbb{C}[z]_{I \le t}$, then $r_1 {*} r_2 \in \mathbb{C}[z]_{I \le s+t}$.

A Noether normalization $\mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[z]/I$ has the following properties.

Proposition 2.2. *Every element of* $\mathbb{C}[z_1, \dots, z_m]$ *is a normal form, so that*

$$\mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[z]_I .$$

Furthermore, for $i = m+1, \dots, n$, *we have* $z_i^{d_i} \in \langle \mathrm{LT}(I) \rangle$, *where* d_i *is as in* (2-1).

Proof. For the second assertion of the proposition, suppose $i \in \{m+1, \dots, n\}$ and $g_i \in I$ is as in (2-1). Then the definition of grevlex and the degree condition in (2-1) makes it easy to see that $\text{LT}(g_i) = z_i^{d_i}$, which implies $z_i^{d_i} \in \langle \text{LT}(I) \rangle$.

Since normal forms are known to form a subspace, it suffices to show that every monomial in $\mathbb{C}[z_1, \dots, z_m]$ is a normal form. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m, 0, \dots, 0)$, so that $z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_m^{\alpha_m}$. We want to show that $z^\alpha \notin \langle \text{LT}(I) \rangle$.

Suppose not, i.e., $z^\alpha \in \langle \text{LT}(I) \rangle$. We will obtain a contradiction by studying the Hilbert function. Take $z^\gamma \notin \langle \text{LT}(I) \rangle$, where $\gamma = (\gamma_1, \dots, \gamma_n)$. If $i \geq m+1$, then $z_i^{d_i} \in \langle \text{LT}(I) \rangle$, so $z_i^{d_i}$ cannot divide z^γ. Hence

(2-3) $$\gamma_i < d_i, \quad \text{for } \textit{all } i = m+1, \dots, n.$$

Furthermore, $z^\alpha \in \langle \text{LT}(I) \rangle$, so z^α cannot divide z^γ. Then

(2-4) $$\gamma_i < \alpha_i, \quad \text{for } \textit{some } i = 1, \dots, m.$$

Now let

$$L(s) := \{\gamma : z^\gamma \notin \langle \text{LT}(I) \rangle,\ |\gamma| \leq s\},$$

so that $|L(s)| = \dim(\mathbb{C}[z]/I)_{\leq s}$ is the Hilbert function. Also, for $i = 1, \dots, m$, let

$$L_i(s) = \{\gamma \in L(s) : \gamma_i < \alpha_i \text{ and } \gamma_{m+1} < d_{m+1}, \dots, \gamma_n < d_n\}.$$

Then (2-3) and (2-4) imply that

(2-5) $$L(s) \subseteq L_1(s) \cup \cdots \cup L_m(s).$$

Observe that

$$|L_i(s)| \leq \alpha_i \cdot d_{m+1} \cdots d_n \cdot \dim \mathbb{C}[z_1, \dots, \hat{z}_i, \dots, z_m]_{\leq s}.$$

Combining this with (2-2) and (2-5), we obtain $|L(s)| = O(s^{m-1})$. It follows that

(2-6) $$\dim(\mathbb{C}[z]/I)_{\leq s} = O(s^{m-1}).$$

On the other hand, the inclusion $\mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[z]/I$ gives an inclusion

$$\mathbb{C}[z_1, \dots, z_m]_{\leq s} \subseteq (\mathbb{C}[z]/I)_{\leq s},$$

and then (2-2) implies $\dim(\mathbb{C}[z]/I)_{\leq s} \geq (1/m!)s^m + O(s^{m-1})$. This contradicts (2-6) and completes the proof. □

3. Constructing an ordered basis

In what follows we will use the following standard notation.

Notation 3.1. Given a set of polynomials $I \subseteq \mathbb{C}[z_1, \dots, z_n] = \mathbb{C}[z]$, write

$$\boldsymbol{V}(I) := \{(a_1, \dots, a_n) \in \mathbb{C}^n : p(a_1, \dots, a_n) = 0 \text{ for all } p \in I\},$$

and given a set $V \subseteq \mathbb{C}^n$, write

$$\boldsymbol{I}(V) := \{p \in \mathbb{C}[z] : p(a_1, \dots, a_n) = 0 \text{ for all } (a_1, \dots, a_n) \in V\}.$$

Let $V \subseteq \mathbb{C}^n$ be an affine algebraic variety of pure dimension m ($m \leq n$). Here, "pure" means that all irreducible components of V have dimension m. If we set $I := \boldsymbol{I}(V) \subseteq \mathbb{C}[z_1, \dots, z_n]$, then the coordinate ring $\mathbb{C}[V]$ of polynomial functions on V satisfies

$$\mathbb{C}[V] \simeq \mathbb{C}[z]/I \simeq \mathbb{C}[z]_I.$$

In what follows, we will use these isomorphisms to identify $\mathbb{C}[V]$ with $\mathbb{C}[z]_I$ and write $\mathbb{C}[V] = \mathbb{C}[z]_I$.

We will construct a special basis of $\mathbb{C}[V]$ by doing interpolation at infinity. Identify $(a_1, \dots, a_n) \in \mathbb{C}^n$ with $[1 : a_1 : \cdots : a_n] \in \mathbb{P}^n$; the hyperplane at infinity is then

$$H_\infty := \{[a_0 : a_1 : \cdots : a_n] \in \mathbb{P}^n : a_0 = 0\}$$

and we write $\mathbb{C}^n \cup H_\infty = \mathbb{P}^n$. Denote by $\bar{V} \subseteq \mathbb{P}^n$ the projective closure of V, which may be computed as follows. If $I = \boldsymbol{I}(V) \subseteq \mathbb{C}[z] = \mathbb{C}[z_1, \dots, z_n]$, let

$$I^h := \{p^h \in \mathbb{C}[z_0, \dots, z_n] : p \in I\},$$

where $p(z) = \sum_{|\alpha| \leq d} c_\alpha z^\alpha \in \mathbb{C}[z]$ of degree d homogenizes to

$$p^h(z_0, z) := \sum_{|\alpha| \leq d} c_\alpha z_0^{d-|\alpha|} z^\alpha \in \mathbb{C}[z_0, z] = \mathbb{C}[z_0, z_1, \dots, z_n].$$

Then the projective closure $\bar{V} \subseteq \mathbb{P}^n$ is given by

$$\bar{V} = \boldsymbol{V}(I^h) = \{[a_0 : \cdots : a_n] \in \mathbb{P}^n : p(a_0, \dots, a_n) = 0 \text{ for all } p \in I^h\}.$$

Note that I^h is a homogeneous ideal (i.e., it is generated by homogeneous polynomials). For a homogeneous ideal $J \subseteq \mathbb{C}[z_0, \dots, z_n]$ we will write

$$J_t = \{p \in J : p \text{ is homogeneous, } \deg p = t\},$$

and

$$(\mathbb{C}[z_0, \dots, z_n]/J)_t = \mathbb{C}[z_0, \dots, z_n]_t / J_t.$$

We will assume that V has the following properties:

(3-1)

(0) V is pure of dimension m and has degree d.

(1) $R := \mathbb{C}[z_1, \ldots, z_m] \subseteq \mathbb{C}[V]$ is a Noether normalization as above.

(2) $\bar{V} \cap P$ consists of d distinct points, where $\bar{V}$ is the projective closure of V in $\mathbb{P}^n$ and $P = \boldsymbol{V}(z_0, \ldots, z_{m-1}) \subseteq \mathbb{P}^n$.

(3) If $\bar{V} \cap P = \{p_1, \ldots, p_d\}$, with $p_i = [0 : \cdots : 0 : p_{im} : \cdots : p_{in}]$, then for each i, $p_{im} \neq 0$.

Note that $\bar{V} \subseteq \mathbb{P}^n$ is pure of dimension m and has degree d, while $P \subseteq \mathbb{P}^n$ is a linear space of dimension $n-m$ and has degree 1. Since $\bar{V} \cap P$ is finite by property (3), Bézout's theorem implies that $\bar{V} \cap P$ consists of $d \cdot 1 = d$ points counted with multiplicity. Property (3) then implies that the multiplicities of the p_i are all one. In algebraic geometry, we express this by saying that

$$\boldsymbol{V}(I^h + \langle z_0, \ldots, z_{m-1} \rangle) = \{p_1, \ldots, p_d\}$$

as subschemes of $\mathbb{P}^n$. In other words, the variety of an ideal equals a finite collection of points as a subscheme exactly when all of the points have multiplicity one.

It follows that the homogeneous ideals $I^h + \langle z_0, \ldots, z_{m-1} \rangle$ and $\boldsymbol{I}(\{p_1, \ldots, p_d\})$ define the same subscheme of $\mathbb{P}^n$. Hence there is an integer $t_0 \geq 0$ such that

$$\begin{aligned}(I^h + \langle z_0, \ldots, z_{m-1} \rangle)_t &= (\boldsymbol{I}(\{p_1, \ldots, p_d\}))_t \\ &= \{f \in \mathbb{C}[z_0, \ldots, z_n]_t : f(p_i) = 0, \text{ for all } i = 1, \ldots, d\}\end{aligned}$$

when $t \geq t_0$; see [Hartshorne 1977, II.5].

A polynomial $f \in \mathbb{C}[z_0, \ldots, z_n]_t$ gives a function on

$$U_m = \{z = [z_0 : z_1 : \cdots : z_n] \in \mathbb{P}^n : z_m \neq 0\}, \tag{3-2}$$

via $[a_0 : \cdots : a_n] \mapsto a_m^{-t} f(a_0, \ldots, a_n)$. It is easy to see that the computation is independent of homogeneous coordinates. For convenience this local evaluation will be denoted $f(a)$.

Lemma 3.2. *The map* $\mathbb{C}[z_0, \ldots, z_n]_t \to \mathbb{C}^d$ *given by* $f \mapsto (f(p_1), \ldots, f(p_d))$ *is onto for* $t \gg 0$.

Proof. By property (3) of (3-1), the points $p_1, \ldots, p_d$ are in the affine chart U_m given by (3-2). For each $i = 1, \ldots, d$ and $p_i = [0 : \cdots : 0 : 1 : u_{i(m+1)} : \cdots : u_{in}]$, put $q_i := (0, \ldots, 0, u_{i(m+1)}, \ldots, u_{in}) \in \mathbb{C}^n_m$, where $\mathbb{C}^n_m$ denotes affine space with coordinates $(z_0, \ldots, z_{m-1}, z_{m+1}, \ldots, z_n)$. It is standard that one can find interpolating polynomials $w_1, \ldots, w_d$ in $\mathbb{C}[z_0, \ldots, z_{m-1}, z_{m+1}, \ldots, z_n]$ such that $w_i(q_j) = \delta_{ij}$.

Pick any $t \geq \max(\deg w_1, \ldots, \deg w_d)$ and set

$$v_i := z_m^t w_i(z_0/z_m, \ldots, z_{m-1}/z_m, z_{m+1}/z_m, \ldots, z_n/z_m). \tag{3-3}$$

This is a homogeneous polynomial of degree t in $z_0, \ldots, z_n$ and its evaluation on U_m satisfies $v_i(p_j) = \delta_{ij}$. For each i, the polynomial $v_i \in \mathbb{C}[z_0, \ldots, z_n]$ evaluates to the standard basis vector $(0, \ldots, 0, 1, 0, \ldots, 0) = e_i \in \mathbb{C}^d$ (the 1 is in the i-th slot), so the map is onto. □

Corollary 3.3. *For $t \gg 0$, we have an exact sequence*

$$0 \to (I^h + \langle z_0, \ldots, z_{m-1}\rangle)_t \to \mathbb{C}[z_0, \ldots, z_n]_t \to \mathbb{C}^d \to 0.$$

Thus there are polynomials $v_1, \ldots, v_d \in \mathbb{C}[z_0, \ldots, z_n]_t$, unique up to elements of $(I^h + \langle z_0, \ldots, z_{m-1}\rangle)_t$, such that $v_i(p_j) = \delta_{ij}$.

Now fix such a t and let $\mathcal{S} := \mathbb{C}[z_0, \ldots, z_n]/(I^h + \langle z_0, \ldots, z_{m-1}\rangle)$. If we regard the polynomials $v_1, \ldots, v_d$ in the above corollary as elements of $\mathcal{S}_t$, then they have the following properties:

$$v_i^2 = z_m^t v_i \text{ for all } i = 1, \ldots, d; \quad \text{and } v_i v_j = 0 \text{ whenever } i \neq j. \tag{3-4}$$

Lemma 3.4. *For any $\tau \geq t$, the polynomials $\{z_m^{\tau-t} v_i\}_{i=1}^d$ form a basis of $\mathcal{S}_\tau$.*

Proof. The construction (3-3) applied to τ (in place of t) gives the additional powers of z_m. □

When we consider the v_i as polynomials in $\mathbb{C}[z_0, \ldots, z_n]/(I^h + \langle z_0\rangle)$, we have

$$v_i^2 = z_m^t v_i + \sum_{k=1}^{m-1} z_k H_k(z_1, \ldots, z_n), \tag{3-5}$$

$$v_i v_j = \sum_{k=1}^{m-1} z_k Q_k(z_1, \ldots, z_n), \tag{3-6}$$

where for each k, $H_k(z_1, \ldots, z_n)$ and $Q_k(z_1, \ldots, z_n)$ are homogeneous polynomials of degree $2t - 1$.

The next step is to translate the v_i into polynomials $\boldsymbol{v}_i$ in $\mathbb{C}[V]$, paying careful attention to their degrees and the analogs of (3-5) and (3-6). Let $\mathbb{C}[V]_{\leq t} = \mathbb{C}[z]_{I \leq t}$ be the collection of normal forms of degree $\leq t$, and let $\mathbb{C}[V]_{=t}$ be those that are homogeneous of degree t.

Lemma 3.5. *We have $\mathbb{C}[V]_{=t} \simeq \mathbb{C}[V]_{\leq t}/\mathbb{C}[V]_{\leq t-1} \simeq (\mathbb{C}[z_0, \ldots, z_n]/(I^h + \langle z_0\rangle))_t$.*

Proof. Writing a normal form as a sum of homogeneous components gives the direct sum decomposition $\mathbb{C}[V]_{\leq t} = \mathbb{C}[V]_{=t} \oplus \mathbb{C}[V]_{\leq t-1}$, and the first isomorphism follows immediately.

For the second, the map $p \mapsto z_0^t p(z_1/z_0, \dots, z_n/z_0)$ induces an isomorphism

$$\mathbb{C}[V]_{\le t} \simeq (\mathbb{C}[z]/I)_{\le t} \simeq (\mathbb{C}[z_0, z]/I^h)_t;$$

see [Cox et al. 1997, §9.3]. This isomorphism sends $\mathbb{C}[V]_{\le t-1} \subseteq \mathbb{C}[V]_{\le t}$ to $z_0(\mathbb{C}[z_0, z]/I^h)_{t-1}$, so that we get an isomorphism

$$\begin{aligned}\mathbb{C}[V]_{\le t}/\mathbb{C}[V]_{\le t-1} &\simeq (\mathbb{C}[z_0, z]/I^h)_t/z_0(\mathbb{C}[z_0, z]/I^h)_{t-1}\\ &\simeq (\mathbb{C}[z_0, z]/(I^h + \langle z_0\rangle))_t.\end{aligned}$$ □

Remark 3.6. Note that multiplication in $\mathbb{C}[z_0, \dots, z_n]/(I^h + \langle z_0\rangle)$ corresponds to linear maps $\hat{*} : \mathbb{C}[V]_{=t} \times \mathbb{C}[V]_{=s} \to \mathbb{C}[V]_{=s+t}$, where to get $p\hat{*}q$, we compute $p{*}q$ (the normal form of pq) and then take the homogeneous part of degree $s+t$.

Lemma 3.7. *For each $i = 1, \dots, d$, there is a polynomial $\boldsymbol{v}_i \in \mathbb{C}[V]_{=t}$ that satisfies the following equations in $\mathbb{C}[V]$:*

(1) $\boldsymbol{v}_i{*}\boldsymbol{v}_i = z_m^t{*}\boldsymbol{v}_i + \sum_{k=1}^{m-1} z_k{*}h_k + h_0$ *with* $\deg(h_k) \le 2t-1$ *for each* $k = 0, \dots, m-1$.

(2) $\boldsymbol{v}_i{*}\boldsymbol{v}_j = \sum_{k=1}^{m-1} z_k{*}q_k + q_0$ *if* $i \ne j$ *with* $\deg(q_k) < 2t-1$ *for each* k.

Remark 3.8. Since $\mathbb{C}[V]$ is identified with the space $\mathbb{C}[z]_I$ of normal forms, the products involving $*$ in Lemma 3.7 represent multiplication of polynomials followed by reduction to normal form.

Proof. Given $v_i \in (\mathbb{C}[z_0, \dots, z_n]/(I^h + \langle z_0\rangle))_t$, let $\boldsymbol{v}_i$ be the element of $\mathbb{C}[V]_{=t}$ given by the isomorphism in Lemma 3.5. For each $k = 1, \dots, m-1$, let $h_k \in \mathbb{C}[V]_{=2t-1}$ be the element corresponding to $H_k \in (\mathbb{C}[z_0, \dots, z_n]/(I^h + \langle z_0\rangle))_{2t-1}$ in (3-5). Then by (3-5), the polynomial

$$\boldsymbol{v}_i\hat{*}\boldsymbol{v}_i - z_m^t\hat{*}\boldsymbol{v}_i - \sum_{k=1}^{m-1} z_k\hat{*}h_k \in \mathbb{C}[V]_{=2t}$$

corresponds to the zero polynomial in $(\mathbb{C}[z_0, \dots, z_n]/(I^h + \langle z_0\rangle))_{2t}$, so it must be zero in $\mathbb{C}[V]_{=2t}$. (Here, $\hat{*}$ is as in Remark 3.6.) Thus the polynomial $h_0 := \boldsymbol{v}_i{*}\boldsymbol{v}_i - z_m^t{*}\boldsymbol{v}_i - \sum_{k=1}^{m-1} z_k{*}h_k$ is in $\mathbb{C}[V]_{\le 2t-1}$. This proves (1).

A similar argument applied to (3-6) proves (2). □

In what follows, we use the notation

$$z^\alpha = z_1^{a_1} \cdots z_{m-1}^{a_{m-1}}, \quad z^\beta = z_{m+1}^{b_{m+1}} \cdots z_n^{b_n}.$$

Define the finite set of monomials

$$\mathcal{B} := \{z_m^l z^\beta \notin \langle \mathrm{LT}(I)\rangle,\ l + |\beta| \le t-1\} \subseteq \mathbb{C}[V]. \tag{3-7}$$

Proposition 3.9. $\mathbb{C}[V]$ *is spanned over* $\mathbb{C}$ *by the homogeneous polynomials*

$$(*) \qquad z^\alpha z_m^l * z^\beta : \quad \alpha \in \mathbb{Z}_{\geq 0}^{m-1},\ z_m^l z^\beta \in \mathcal{B},$$

$$(**) \qquad z^\alpha z_m^l * \boldsymbol{v}_i : \quad \alpha \in \mathbb{Z}_{\geq 0}^{m-1},\ l \geq 0,\ i = 1, \ldots, d.$$

Remark 3.10. Note that $z^\alpha z_m^l$ is a normal form by Proposition 2.2, while the products $z^\alpha z_m^l z^\beta$ and $z^\alpha z_m^l \boldsymbol{v}_i$ may fail to be normal forms. This explains why the proposition uses $z^\alpha z_m^l * z^\beta$ and $z^\alpha z_m^l * \boldsymbol{v}_i$.

Proof. To simplify the proof, we will omit the $*$ when multiplying normal forms. It suffices to show that any monomial $z^\alpha z_m^l z^\beta \notin \langle \mathrm{LT}(I) \rangle$ can be expressed as a linear combination of elements of $(*)$ and $(**)$.

We will prove this by induction on $s = |\alpha| + l + |\beta|$. Suppose $z^\alpha z_m^l z^\beta \notin \langle \mathrm{LT}(I) \rangle$ with $s \leq t - 1$. Then $|\alpha| + l + |\beta| \leq t - 1$, so that $z_m^l z^\beta \in \mathcal{B}$. Hence the monomial is in $(*)$, which proves the base case.

Next, assume $s \geq t$ and that $\mathbb{C}[V]_{\leq s-1}$ is spanned by the polynomials $(*)$ and $(**)$ of degree $\leq s - 1$. Take $z^\alpha z_m^l z^\beta \notin \langle \mathrm{LT}(I) \rangle$ of degree s. No factor of this monomial is in the ideal either; in particular, $z_m^l z^\beta \notin \langle \mathrm{LT}(I) \rangle$. If $l + |\beta| \leq t - 1$, then $z_m^l z^\beta \in \mathcal{B}$ and therefore $z^\alpha z_m^l z^\beta$ is an element of the form $(*)$.

Otherwise, $\tau := l + |\beta| \geq t$. By Lemma 3.4, we have an equation

$$z_m^l z^\beta = \sum_{i=1}^{d} a_i z_m^{\tau - t} v_i + \sum_{j=0}^{m-1} z_j H_j(z_0, z) + H(z_0, z),$$

in $\mathbb{C}[z_0, z]$, where $a_i \in \mathbb{C}$, $\deg H_j = \tau - 1$ and $H \in I^h$. If we dehomogenize by setting $z_0 = 1$, we obtain

$$z_m^l z^\beta = \sum_{i=1}^{d} a_i z_m^{\tau - t} v_i + \sum_{j=1}^{m-1} z_j h_j(z) + h_0(z)$$

in $\mathbb{C}[z]/I$, where $a_i \in \mathbb{C}$ and $\deg h_j \leq \tau - 1$. We can multiply by z^α to obtain

$$z^\alpha z_m^l z^\beta = \sum_{i=1}^{d} a_i z^\alpha z_m^{\tau - t} v_i + \sum_{j=1}^{m-1} z_j (z^\alpha h_j(z)) + z^\alpha h_0(z)$$

in $\mathbb{C}[z]/I$. Using the isomorphism $\mathbb{C}[V] = \mathbb{C}[z]_I \simeq \mathbb{C}[z]/I$, this becomes

$$z^\alpha z_m^l z^\beta = \sum_{i=1}^{d} a_i z^\alpha z_m^{\tau - t} \boldsymbol{v}_i + \sum_{j=1}^{m-1} z_j (z^\alpha h_j(z)) + z^\alpha h_0(z).$$

in $\mathbb{C}[V]$. The first sum is a linear combination of elements of the form $(**)$. For the second sum, note that $\deg(z^\alpha h_j) \leq s - 1$ for each $j = 1, \ldots, m - 1$. By the inductive hypothesis, this means that $z^\alpha h_j$ is a linear combination of terms in $(*)$

and $(**)$, and therefore $z_j z^\alpha h_j$ is too, by definition. Finally, $\deg(z^\alpha h_0) \le s-1$, and again by induction, $z^\alpha h_0$ is a linear combination of terms in $(*)$ and $(**)$. $\square$

The following is an immediate corollary of the above proof:

Corollary 3.11. *$\mathbb{C}[V]_{\le s}$ is spanned over $\mathbb{C}$ by the polynomials in $(*)$ and $(**)$ of degree $\le s$.*

Now that we have a spanning set, the next step in constructing the desired basis for $\mathbb{C}[V]$ is to show that the elements of the form $(**)$ are linearly independent over $\mathbb{C}$. These elements are monomials in $z_1, \dots, z_m$ multiplied by one of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$. Since the inclusion $\mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[V]$ makes $\mathbb{C}[V]$ into a module over $R = \mathbb{C}[z_1, \dots, z_m]$, we can verify linear independence by showing the following:

Theorem 3.12. *The polynomials $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ generate a free R-submodule of $\mathbb{C}[V]$.*

Proof. We first observe that since V has dimension m and degree d, we have

$$\dim \mathbb{C}[V]_{\le s} = \frac{d}{m!} s^m + O(s^{m-1}); \tag{3-8}$$

see, e.g., [Cox et al. 1997, §9.3]. Now let

$$M := \sum_{i=1}^{d} R\boldsymbol{v}_i \quad \text{and} \quad N := \sum_{\mathcal{B}} \mathbb{C}[z_1, \dots, z_{m-1}] z_m^l z^\beta,$$

and for $s \ge t$ define

$$M_{\le s} := \sum_{i=1}^{d} R_{\le s-t} \boldsymbol{v}_i, \quad N_{\le s} := \sum_{\mathcal{B}} \mathbb{C}[z_1, \dots, z_{m-1}]_{\le s-l-|\beta|} z_m^l z^\beta.$$

The corollary implies $\mathbb{C}[V]_{\le s} = M_{\le s} + N_{\le s}$. Using (2-2), one easily obtains

$$\dim N_{\le s} \le \frac{|\mathcal{B}|}{(m-1)!} s^{m-1} + O(s^{m-2}).$$

Combining this with (3-8) and $\mathbb{C}[V]_{\le s} = M_{\le s} + N_{\le s}$ yields

$$\dim M_{\le s} = \frac{d}{m!} s^m + O(s^{m-1}).$$

Suppose there is a nontrivial relation

$$f_1 \boldsymbol{v}_1 + \cdots + f_d \boldsymbol{v}_d = 0, \quad f_i \in R, \text{ not all } f_i = 0. \tag{3-9}$$

Let $D = \max\{\deg f_1, \dots, \deg f_d\}$ and take a large integer $s \ge t + D$. There is an exact sequence

$$0 \to K_s \to R^d_{\le s-t} \xrightarrow{\varphi_s} M_{\le s} \to 0,$$

where $\varphi_s : R^d_{\leq s-t} \to M_{\leq s-t}$ is given by $\varphi(g_1, \dots, g_d) = \sum_i g_i \boldsymbol{v}_i$ and $K_s := \ker \varphi_s$. We have $R_{s-t-D} \cdot (f_1, \dots, f_d) \subseteq R^d_{\leq s-t}$, so by (3-9),

$$R_{s-t-D}(f_1, \dots, f_d) \subseteq K_s.$$

Since $(f_1, \dots, f_d) \neq (0, \dots, 0)$, we have $K_s \neq 0$ and so $\dim K_s \geq \dim R_{s-t-D}$. Thus

$$\dim R^d_{\leq s-t} = \dim M_{\leq s} + \dim K_s \geq \dim M_{\leq s} + \dim R_{s-t-D}.$$

A Hilbert function calculation then gives the inequality

$$\frac{d}{m!} s^m + O(s^{m-1}) \geq \Big(\frac{d}{m!} s^m + O(s^{m-1})\Big) + \Big(\frac{1}{m!}(s-t-D)^m + O(s^{m-1})\Big),$$

so that $(1/m!)s^m \leq O(s^{m-1})$, a contradiction. This says that no equation of the form (3-9) can hold, and so $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ are free over R. □

We now construct the sought-after ordered basis for $\mathbb{C}[V]$.

Definition 3.13. By Proposition 3.9, the polynomials given by $(*)$ and $(**)$ span $\mathbb{C}[V]$ and by Theorem 3.12, those from $(**)$ are linearly independent. We first create a basis of $\mathbb{C}[V]$ by adjoining a sufficient number of elements of the form $(*)$ to those of the form $(**)$. List those of the form $(*)$ in grevlex order and discard any monomial that is linearly dependent with respect to elements of the form $(**)$ together with previous elements of $(*)$; otherwise keep it. This yields the basis $\mathcal{C}$ of $\mathbb{C}[V]$. We define an ordering $\prec$ on $\mathcal{C}$ as follows. First, order the elements by total degree; then for a fixed degree s,

- let elements of $(*)$ precede elements of $(**)$;
- let $z^\alpha z_m^l * \boldsymbol{v}_i \prec z^{\hat{\alpha}} z_m^{\hat{l}} * \boldsymbol{v}_j$ if $z^\alpha z_m^l$ precedes $z^{\hat{\alpha}} z_m^{\hat{l}}$ according to grevlex;
- let $z^\alpha z_m^l * \boldsymbol{v}_i \prec z^\alpha z_m^l * \boldsymbol{v}_j$ if $i < j$; and
- let elements of the form $(*)$ be ordered according to grevlex.

It is easy to see that the elements of $\mathcal{C}$ of degree $\leq s$ form a basis of $\mathbb{C}[V]_{\leq s}$. The Chebyshev constants defined in Section 5 will use the ordered basis of $\mathbb{C}[V]$ given in Definition 3.13.

We conclude this section by computing some examples of $\mathcal{C}$ and $\prec$.

Example 3.14. Let $V = \{z = (z_1, \dots, z_n) \in \mathbb{C}^n : z_{m+1} = z_{m+2} = \dots = z_n = 0\}$. The Noether normalization is the identity, $\mathbb{C}[z_1, \dots, z_m] = \mathbb{C}[V]$, and in the notation of (3-1), $\overline{V} \cap P = \{[0 : \dots : 0 : 1 : 0 : \dots : 0]\}$, where the 1 is in the m-th slot. We take $v_1 = \boldsymbol{v}_1 = 1$ (so $t = 0$). The basis $\mathcal{C}$ consists of the monomials in $\mathbb{C}[z_1, \dots, z_m]$, which are elements of the form $(**)$, ordered by grevlex. There are no elements of the form $(*)$ in this case.

Example 3.15. Let V be the complexified sphere in $\mathbb{C}^3$, i.e., the algebraic surface given by the equation $z_1^2+z_2^2+z_3^2=1$. A basis of $\mathbb{C}[V]$ is given by all monomials not in $\langle z_3^2\rangle$, i.e.,

$$1, z_1, z_2, z_3, z_1^2, z_1z_2, z_1z_3, z_2^2, z_2z_3, z_1^3, \dots .$$

The Noether normalization is $\mathbb{C}[z_1, z_2] \subseteq \mathbb{C}[V]$.

In $\mathbb{P}^3$, $\bar{V}$ is given by all points $[z_0 : z_1 : z_2 : z_3]$ satisfying $z_1^2+z_2^2+z_3^2=z_0^2$, and $P=\{z_0=z_1=0\}$. The points of $\bar{V}\cap P$ are then $p_1=[0:0:1:-i]$ and $p_2=[0:0:1:i]$. Thus (3-1) is satisfied.

Interpolating polynomials are $v_1=\frac{1}{2}(z_2+iz_3)$ and $v_2=\frac{1}{2}(z_2-iz_3)$. In this case $t=1$ so that $\boldsymbol{v}_1=v_1$ and $\boldsymbol{v}_2=v_2$. The first few elements of the basis $\mathcal{C}$, ordered by $\prec$, are

$$1, z_1, \boldsymbol{v}_1, \boldsymbol{v}_2, z_1^2, z_1\boldsymbol{v}_1, z_1\boldsymbol{v}_2, z_2\boldsymbol{v}_1, z_2\boldsymbol{v}_2, z_1^3, \dots .$$

Basis elements of the form $(*)$ are z_1^k while those of the form $(**)$ are $z_1^{\alpha_1}z_2^{\alpha_2}\boldsymbol{v}_i$. (Note that since $l<t=1$, no factors of the form z_2^l appear in $(*)$).

Example 3.16. When $V=\boldsymbol{V}(f)\subseteq\mathbb{C}^n$ is a hypersurface given by $f\in\mathbb{C}[z_1,\dots,z_n]$, we can generalize Example 3.15 by computing the basis $\mathcal{C}$ rather explicitly. We assume that f is a product of distinct irreducible polynomials, so that $I=\boldsymbol{I}(V)=\langle f\rangle$. We also assume that $\mathrm{LT}(f)=z_n^d$ where $d=\deg(f)$. This ensures that $\mathbb{C}[z_1,\dots,z_{n-1}]\subseteq\mathbb{C}[V]$ is a Noether normalization.

Let $F:=f^h\in\mathbb{C}[z_0,\dots,z_n]$ be the homogenization of f; then in $\mathbb{P}^n$, $\bar{V}=\boldsymbol{V}(F)$ and $I^h=\langle F\rangle$. If the properties (3-1) hold, then $\boldsymbol{V}(F,z_0,\dots,z_{n-2})\subseteq\mathbb{P}^n$ consists of d distinct points, all with $z_{n-1}\neq 0$, given by $[0:\cdots:1:\beta_i]$ for $i=1,\dots,d$.

Separating the terms of F containing only the variables z_{n-1}, z_n from the others,

$$F(z)=G(z_{n-1},z_n)+\sum_{l=0}^{n-2} z_l H_l(z_0,\dots,z_n), \tag{3-10}$$

where $\deg(G)=d$ and $\deg H_l=d-1$ for each $l=0,\dots,n-2$. Thus $G(1,\beta_i)=0$ for $i=1,\dots,d$.

In the notation of earlier in the section, we have

$$\begin{aligned}\mathcal{S}&=\mathbb{C}[z_0,\dots,z_n]/(I^h+\langle z_0,\dots,z_{n-2}\rangle)=\mathbb{C}[z_0,\dots,z_n]/\langle F(z),z_0,\dots,z_{n-2}\rangle\\&=\mathbb{C}[z_0,\dots,z_n]/\langle G(z_{n-1},z_n),z_0,\dots,z_{n-2}\rangle\\&\simeq\mathbb{C}[z_{n-1},z_n]/\langle G(z_{n-1},z_n)\rangle,\end{aligned}$$

where the second line uses (3-10) and the third uses the map

$$p(z_0,z_1,\dots,z_n)\mapsto p(0,\dots,0,z_{n-1},z_n).$$

We factor $G(z_{n-1}, z_n) = \prod_{i=1}^{d}(z_n - \beta_i z_{n-1}) = \prod_{i=1}^{d} l_i(z_{n-1}, z_n)$. Note that $\beta_i \neq \beta_j$ if $i \neq j$. For each $i = 1, \ldots, d$, define

$$v_i(z_{n-1}, z_n) = \prod_{j \neq i} \frac{l_j(z_{n-1}, z_n)}{l_j(1, \beta_i)}. \tag{3-11}$$

Then $\deg(v_i) = d - 1$ for each i, and clearly

$$v_i(1, \beta_j) = \begin{cases} 0 & \text{if } j \neq i, \\ 1 & \text{if } j = i. \end{cases} \tag{3-12}$$

Note that when $f = z_1^2 + z_2^2 + z_3^2 - 1$ as in Example 3.15, we have the points $[0:0:1:-i]$ and $[0:0:1:i]$. Then $G = z_2^2 + z_3^2 = (z_3 + iz_2)(z_3 - iz_2) = l_1 l_2$ and the formula for v_1 reduces to

$$v_1 = \frac{l_2(z_2, z_3)}{l_2(1, -i)} = \frac{z_3 - iz_2}{-2i} = \frac{1}{2}(z_2 + iz_3),$$

in agreement with Example 3.15. The formula for v_2 works similarly.

By (3-12), $v_1, \ldots, v_d$ satisfy Lemma 3.4 with $t = d-1$. Since the v_i only involve z_{n-1}, z_n and are normal forms with respect to grevlex (having degree $\leq d-1$ in z_n), we can take $\boldsymbol{v}_i = v_i$ in Lemma 3.7. Thus $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ are defined by (3-11) and have degree $d - 1$.

The next step is to identify the set $\mathcal{B}$ from (3-7). Since $m = n-1$, the monomials z^α and z^β from Proposition 3.9 are

$$z^\alpha = z_1^{a_1} \cdots z_{n-2}^{a_{n-2}}, \quad z^\beta = z_n^b.$$

In this notation, a monomial in $z_1, \ldots, z_n$ is written $z^\alpha z_{n-1}^l z_n^b$. Since the $\boldsymbol{v}_i$ have degree $t = d - 1$ and $\langle \mathrm{LT}(I) \rangle = \langle \mathrm{LT}(f) \rangle = \langle z_n^d \rangle$, it follows that (3-7) becomes

$$\mathcal{B} = \{z_{n-1}^l z_n^b \notin \langle z_n^d \rangle : l + b \leq d-2\} = \{z_{n-1}^l z_n^b : l + b \leq d-2\}.$$

Hence the collections $(*)$ and $(**)$ from Proposition 3.9 are

$$\begin{array}{lll} (*) & z^\alpha z_{n-1}^l z_n^b : & \alpha \in \mathbb{Z}_{\geq 0}^{n-2},\ l + b \leq d-2, \\ (**) & z^\alpha z_{n-1}^l \boldsymbol{v}_i : & \alpha \in \mathbb{Z}_{\geq 0}^{n-2},\ l \geq 0,\ i = 1, \ldots, d. \end{array} \tag{3-13}$$

These products are all normal forms, so no $*$ is needed in the multiplications.

The nicest feature of the hypersurface case is that the basis $\mathcal{C}$ consists *precisely* of the polynomials in (3-13). They span by Proposition 3.9, so we only need to prove linear independence. The polynomials in $(**)$ are linearly independent by Theorem 3.12, and those in $(*)$ are linearly independent since they are normal-form monomials. Hence it remains to study an equation of the form

$$\text{linear combination of } z^\alpha z_{n-1}^l z_n^b = \text{linear combination of } z^\alpha z_{n-1}^l \boldsymbol{v}_i.$$

The left-hand side has degree $\leq d-2$ in z_{n-1}, z_n and the right-hand side has degree $\geq d-1$. This forces the linear combinations to be trivial, and linear independence follows.

To summarize: when $V = \boldsymbol{V}(f)$ is a hypersurface of degree d, the $\boldsymbol{v}_i$ are polynomials of degree $d-1$ that we can compute explicitly in terms of f, and the elements of $(*)$ consist of all monomials $z_1^{\alpha_1} \cdots z_{n-1}^{\alpha_{n-1}} z_n^{\alpha_n}$ with $\alpha_{n-1} + \alpha_n \leq d-2$.

4. Weakly submultiplicative functions

Bloom and Levenberg [2003] observed that the main properties of Zaharyuta's directional Chebyshev constants followed from the submultiplicative property of sup norms of Chebyshev polynomials, and could be recast rather abstractly as properties of submultiplicative functions on integer tuples. We verify here that these properties still hold under slightly weaker conditions. The arguments are those of Zaharyuta's [1975] paper with minor adjustments. We will apply these results concretely in the next section.

Definition 4.1. Let m be a positive integer. A nonnegative function $Y : \mathbb{Z}_{\geq 0}^m \to \mathbb{R}_{\geq 0}$ is said to be *weakly submultiplicative* if there is a finite subset $\mathcal{F}$ of $\mathbb{Z}_{\geq 0}^m$ such that:

For all $\alpha, \beta \in \mathbb{Z}_{\geq 0}^m$ *there exists* $\gamma \in \mathcal{F}$ *such that* $Y(\alpha+\beta+\gamma) \leq Y(\alpha)Y(\beta)$.

Y has *subexponential growth* if for some $C, r > 1$ we have $Y(\alpha) \leq Cr^{|\alpha|}$ for all α.

Remark 4.2 (cf., [Bloom and Levenberg 2003]). When $Y(\alpha+\beta) \leq Y(\alpha)Y(\beta)$, i.e., $\mathcal{F} = \{(0, \ldots, 0)\}$, Y is called *submultiplicative*. A submultiplicative function automatically has subexponential growth: if $\alpha = (\alpha_1, \ldots, \alpha_m)$ then

$$Y(\alpha) = Y\Big(\sum_{k=1}^{m} \alpha_k e_k\Big) \leq \prod_{k=1}^{m} Y(e_k)^{\alpha_k} \leq r^{|\alpha|},$$

where e_k is the k-th coordinate vector and $r = \max_k Y(e_k)$. It seems that weak submultiplicativity should also imply subexponential growth, but the above argument runs into some technical difficulties.

Let

$$\Sigma_m := \big\{\theta = (\theta_1, \ldots, \theta_m) \in \mathbb{R}^m : \theta_i \geq 0 \text{ for all } i,\ \textstyle\sum_i \theta_i = 1\big\}$$

denote the simplex in $\mathbb{R}^m$, and let $\Sigma_m^\circ := \{\theta \in \Sigma_m : \theta_i > 0 \text{ for all } i\}$ be its interior.

Lemma 4.3. *Let* $Y : \mathbb{Z}_{\geq 0}^m \to \mathbb{R}_{\geq 0}$ *be weakly submultiplicative with subexponential growth. For all* $\theta \in \Sigma_m^\circ$, *the limit* $T(\theta) := \lim\limits_{\substack{|\alpha|\to\infty \\ \alpha/|\alpha|\to\theta}} Y(\alpha)^{1/|\alpha|}$ *exists.*

Proof. Let $\{\alpha_{(j)}\}$ and $\{\tilde{\alpha}_{(j)}\}$ be sequences in $\mathbb{Z}^m_{\geq 0}$ such that $\alpha_{(j)}/|\alpha_{(j)}|, \tilde{\alpha}_{(j)}/|\tilde{\alpha}_{(j)}| \to \theta$ as $j \to \infty$ and

$$\lim_{j\to\infty} Y(\alpha_{(j)})^{1/|\alpha_{(j)}|} = \liminf_{|\alpha|\to\infty, \alpha/|\alpha|\to\theta} Y(\alpha)^{1/|\alpha|} := L_1,$$
$$\lim_{j\to\infty} Y(\tilde{\alpha}_{(j)})^{1/|\tilde{\alpha}_{(j)}|} = \limsup_{|\alpha|\to\infty, \alpha/|\alpha|\to\theta} Y(\alpha)^{1/|\alpha|} := L_2.$$

To prove the lemma it is sufficient to show that $L_2 \leq L_1$. By passing to subsequences we may assume that $|\tilde{\alpha}_{(j)}|/|\alpha_{(j)}| \to \infty$ as $j \to \infty$.

Let q_j denote the largest nonnegative integer for which all the components of $r_{(j)} := \tilde{\alpha}_{(j)} - q_j\alpha_{(j)}$ are nonnegative. We claim that

$$\frac{q_j|\alpha_{(j)}|}{|\tilde{\alpha}_{(j)}|} \to 1, \qquad \frac{|r_{(j)}|}{|\tilde{\alpha}_{(j)}|} \to 0 \quad \text{as } j \to \infty. \tag{4-1}$$

Write $\alpha_{(j)} = |\alpha_{(j)}|(\theta + \epsilon_{(j)})$ and $\tilde{\alpha}_{(j)} = |\tilde{\alpha}_{(j)}|(\theta + \tilde{\epsilon}_{(j)})$ where $\epsilon_{(j)}, \tilde{\epsilon}_{(j)} \to 0$ as $j \to \infty$. A calculation in components shows that

$$\tilde{\alpha}_{(j)\nu} = \frac{|\tilde{\alpha}_{(j)}|}{|\alpha_{(j)}|}\Big(1 + \frac{|\alpha_{(j)}|}{\alpha_{(j)\nu}}(\tilde{\epsilon}_{(j)\nu} - \epsilon_{(j)\nu})\Big)\alpha_{(j)\nu} \quad \text{for each } \nu = 1, \ldots, m, \tag{4-2}$$

where we write $\alpha_{(j)} = (\alpha_{(j)1}, \ldots, \alpha_{(j)m})$, etc. For any ν, we have

$$\frac{|\alpha_{(j)}|}{\alpha_{(j)\nu}}(\tilde{\epsilon}_{(j)\nu} - \epsilon_{(j)\nu}) \to \frac{1}{\theta_\nu}(0-0) = 0 \quad \text{as } j \to \infty.$$

(Here we use the fact that $\theta \in \Sigma_m^\circ$, so $\theta_\nu \neq 0$.) This says that given $\epsilon > 0$, the quantity in parentheses on the right-hand side of (4-2) exceeds $1-\epsilon$ for all ν when j is sufficiently large. The definition of q_j then implies that

$$q_j \geq \frac{|\tilde{\alpha}_{(j)}|}{|\alpha_{(j)}|}(1-\epsilon) - 1,$$

and hence $q_j|\alpha_{(j)}|/|\tilde{\alpha}_{(j)}| \geq 1 - \epsilon - |\alpha_{(j)}|/|\tilde{\alpha}_{(j)}| \to 1-\epsilon$ as $j \to \infty$. On the other hand, $q_j|\alpha_{(j)}|/|\tilde{\alpha}_{(j)}| \leq 1$ for all j. Since ϵ is arbitrary, (4-1) follows.

Let $c := \max\{\gamma_\nu : \nu \in \{1, \ldots, m\}, (\gamma_1, \ldots, \gamma_m) \in \mathcal{F}\}$, and let s_j be the largest nonnegative integer such that

$$s_j(\alpha_{(j)\nu} + c) \leq q_j\alpha_{(j)\nu} \quad \text{for all } \nu = 1, \ldots, m.$$

Using this, there exists $\tilde{r}_{(j)} \in \mathbb{Z}^m_{\geq 0}$ such that

$$Y(\tilde{\alpha}_{(j)}) = Y(q_j\alpha_{(j)} + r_{(j)}) = Y(s_j\alpha_{(j)} + s_j\gamma_{(j)} + \tilde{r}_{(j)}),$$

where $\gamma_{(j)} \in \mathcal{F}$ satisfies $Y(2\alpha_{(j)}+\gamma_{(j)}) \le Y(\alpha_{(j)})^2$. It is easy to see that $|q_j|/|s_j| \to 1$, and hence (4-1) holds with $q_j, r_{(j)}$ replaced by $s_j, \tilde{r}_{(j)}$. Finally,

$$\begin{aligned} Y(\tilde{\alpha}_{(j)})^{1/|\tilde{\alpha}_{(j)}|} &= Y(s_j\alpha_{(j)} + s_j\gamma_{(j)} + \tilde{r}_{(j)})^{1/|\tilde{\alpha}_{(j)}|} \\ &\le (Y(\alpha_{(j)})^{s_j} Y(\tilde{r}_{(j)}))^{1/|\tilde{\alpha}_{(j)}|} \\ &\le (Y(\alpha_{(j)})^{1/|\alpha_{(j)}|})^{s_j|\alpha_{(j)}|/|\tilde{\alpha}_{(j)}|} C^{1/|\tilde{\alpha}_{(j)}|} r^{|\tilde{r}_{(j)}|/|\tilde{\alpha}_{(j)}|}, \end{aligned}$$

where C, r are as in Definition 4.1. Taking the limit as $j \to \infty$ of the first and last expressions yields $L_2 \le L_1$. This completes the proof. □

Recall that a positive real-valued function f on a convex set $C \subseteq \mathbb{R}^n$ is said to be logarithmically convex if $f((1-t)a+tb) \le f(a)^{1-t} f(b)^t$ for all $a, b \in C$; equivalently, $\log(f)$ is convex.

Lemma 4.4. *The function $\theta \mapsto T(\theta)$, defined as in the previous lemma, is uniformly bounded and logarithmically convex on Σ_m° (and hence continuous).*

Proof. Boundedness follows easily from subexponential growth: if $Y(\alpha) \le Cr^{|\alpha|}$ for all $\alpha \in \mathbb{Z}_{\ge 0}^m$ then $T(\theta) \le r$ for all $\theta \in \Sigma_m^\circ$.

To prove logarithmic convexity, fix $\theta, \tilde{\theta} \in \Sigma_m^\circ$ and $t \in (0, 1)$. Let $\alpha_{(j)}, \alpha_{(j)}$ satisfy $\alpha_{(j)}/|\alpha_{(j)}| \to \theta$, $\tilde{\alpha}_{(j)}/|\tilde{\alpha}_{(j)}| \to \tilde{\theta}$ as $j \to \infty$ and $|\alpha_{(j)}| = |\tilde{\alpha}_{(j)}| =: a_j$ for each j. Let $q_j, \tilde{q}_j$ be positive integers such that $q_j/(q_j+\tilde{q}_j) \to t$ as $j \to \infty$.

For each j there exist $\beta_{(j)}, \gamma_{(j)}, \tilde{\gamma}_{(j)} \in \mathcal{F}$ such that

$$\begin{aligned} &Y\big(q_j\alpha_{(j)} + \tilde{q}_j\tilde{\alpha}_{(j)} + \beta_{(j)} + (q_j-1)\gamma_{(j)} + (\tilde{q}_j-1)\tilde{\gamma}_{(j)}\big) \\ &\qquad \le Y\big(q_j\alpha_{(j)} + (q_j-1)\gamma_{(j)}\big) Y\big(\tilde{q}_j\tilde{\alpha}_{(j)} + (\tilde{q}_j-1)\tilde{\gamma}_{(j)}\big) \le Y(\alpha_{(j)})^{q_j} Y(\tilde{\alpha}_{(j)})^{\tilde{q}_j}. \end{aligned}$$

Let $\zeta_{(j)} := q_j\alpha_{(j)} + \tilde{q}_j\tilde{\alpha}_{(j)} + \beta_{(j)} + (q_j-1)\gamma_{(j)} + (\tilde{q}_j-1)\tilde{\gamma}_{(j)}$. Since $\mathcal{F}$ is bounded, it is easy to see that $|\zeta_{(j)}|/|q_j\alpha_{(j)} + \tilde{q}_j\tilde{\alpha}_{(j)}| \to 1$ as $j \to \infty$ and

$$\begin{aligned} \lim_{j\to\infty} \frac{\zeta_{(j)}}{|\zeta_{(j)}|} &= \lim_{j\to\infty} \frac{q_j\alpha_{(j)} + \tilde{q}_j\tilde{\alpha}_{(j)}}{|q_j\alpha_{(j)} + \tilde{q}_j\tilde{\alpha}_{(j)}|} \\ &= \lim_{j\to\infty} \frac{q_j\alpha_{(j)}}{(q_j+\tilde{q}_j)a_j} + \frac{\tilde{q}_j\tilde{\alpha}_{(j)}}{(q_j+\tilde{q}_j)a_j} = t\theta + (1-t)\tilde{\theta}. \end{aligned}$$

Hence

$$\begin{aligned} T(t\theta + (1-t)\tilde{\theta}) &= \lim_{j\to\infty} Y(\zeta_{(j)})^{1/|\zeta_{(j)}|} \\ &= \lim_{j\to\infty} Y(\zeta_{(j)})^{1/|q_j\alpha_{(j)}+\tilde{q}_j\tilde{\alpha}_{(j)}|} \\ &\le \lim_{j\to\infty} (Y(\alpha_{(j)})^{1/|\alpha_{(j)}|})^{q_j/q_j+\tilde{q}_j} (Y(\tilde{\alpha}_{(j)})^{1/|\tilde{\alpha}_{(j)}|})^{\tilde{q}_j/q_j+\tilde{q}_j} \\ &= T(\theta)^t T(\tilde{\theta})^{1-t}, \end{aligned}$$

which concludes the proof. □

Given $b \in \partial\Sigma_m = \Sigma_m \setminus \Sigma_m^\circ$, define

$$T^-(b) := \liminf_{|\alpha|\to\infty, \alpha/|\alpha|\to b} Y(\alpha)^{1/|\alpha|}. \tag{4-3}$$

Lemma 4.5. *Let $b \in \partial\Sigma_m$. Then*

$$T^-(b) = \liminf_{\theta\to b,\, \theta\in\Sigma_m^\circ} T(\theta).$$

Proof. Let $\{\theta_{(j)}\}_{j\geq 1}$ be a sequence of points in Σ_m^0 with $\theta_{(j)} \to b$ as $j\to\infty$, and for each j choose $\alpha_{(j)}$ such that

$$\left|\frac{\alpha_{(j)}}{|\alpha_{(j)}|} - \theta_{(j)}\right| < \frac{1}{j}, \quad |Y(\alpha_{(j)})^{\frac{1}{|\alpha_{(j)}|}} - T(\theta_{(j)})| < \frac{1}{j}.$$

Then $\alpha_{(j)}/|\alpha_{(j)}| \to b$ as $j\to\infty$, so

$$T^-(b) \leq \liminf_{j\to\infty} Y(\alpha_{(j)})^{1/|\alpha_{(j)}|} \leq \liminf_{j\to\infty}\Bigl(T(\theta_{(j)}) + \frac{1}{j}\Bigr) = \liminf_{j\to\infty} T(\theta_{(j)}).$$

Hence $T^-(b) \leq \liminf_{\theta\to b,\, \theta\in\Sigma_m^\circ} T(\theta)$ since the sequence $\theta_{(j)}$ was arbitrary.

It remains to prove the reverse inequality. Let $\sigma = (\sigma_1, \dots, \sigma_m)$ satisfy $\sigma_\nu > 0$ for each ν; then $(b+\sigma)/(1+|\sigma|) \in \Sigma_m^\circ$. We will show that

$$T\Bigl(\frac{b+\sigma}{1+|\sigma|}\Bigr) \leq r^{\frac{|\sigma|}{1+|\sigma|}} T^-(b)^{\frac{1}{1+|\sigma|}}. \tag{4-4}$$

(Here r is as in Definition 4.1.)

Choose sequences $\alpha_{(j)}, \ell_{(j)}$ in $\mathbb{Z}^m_{\geq 0}$ such that $|\alpha_{(j)}| \to \infty$ and

$$\frac{\alpha_{(j)}}{|\alpha_{(j)}|} \to b \text{ with } Y(\alpha_{(j)})^{\frac{1}{|\alpha_{(j)}|}} \to T^-(b), \quad \text{and } \frac{\ell_{(j)}}{|\alpha_{(j)}|} \to \sigma.$$

Since Y is weakly submultiplicative with subexponential growth,

$$Y(\ell_{(j)} + \alpha_{(j)} + \gamma_{(j)}) \leq Y(\ell_{(j)})Y(\alpha_{(j)}) \leq Cr^{|\ell_{(j)}|}Y(\alpha_{(j)}) \tag{4-5}$$

for appropriate $\gamma_{(j)} \in \mathcal{F}$.

We compute $\ell_{(j)}/|\alpha_{(j)} + \ell_{(j)}| \to \sigma/(1+|\sigma|)$ and $\alpha_{(j)}/|\alpha_{(j)} + \ell_{(j)}| \to b/(1+|\sigma|)$ as $j\to\infty$. Since $\mathcal{F}$ is bounded we also have $\gamma_{(j)}/|\ell_{(j)} + \alpha_{(j)} + \gamma_{(j)}| \to (0,\dots,0)$ and $|\ell_{(j)} + \alpha_{(j)}|/|\ell_{(j)} + \alpha_{(j)} + \gamma_{(j)}| \to 1$. The inequality (4-5) then yields (4-4) by a similar limiting process as detailed in the previous lemmas. Finally, using (4-4),

$$\liminf_{\theta\to b,\, \theta\in\Sigma_m^\circ} T(\theta) \leq \liminf_{\substack{|\sigma|\to 0\\ \sigma_i>0 \text{ for all } i}} T\Bigl(\frac{b+\sigma}{1+|\sigma|}\Bigr) \leq \lim_{|\sigma|\to 0} r^{\frac{|\sigma|}{1+|\sigma|}} T^-(b)^{\frac{1}{1+|\sigma|}} = T^-(b),$$

which is the desired inequality. □

An immediate consequence of Lemma 4.4 and equation (4-4) is the following.

Corollary 4.6. *Suppose $T(\phi) \neq 0$ for some $\phi \in \Sigma_m^\circ$. Then $T(\theta) \neq 0$ for all $\theta \in \Sigma_m^\circ$ and $T^-(b) \neq 0$ for all $b \in \partial\Sigma_m$. The same conclusion holds if $T^-(c) \neq 0$ for some $c \in \partial\Sigma_m$.*

Lemma 4.7. *Let Q be a compact subset of Σ_m°. Then*

$$\limsup_{|\alpha|\to\infty} \left\{ |Y(\alpha)^{1/|\alpha|} - T(\theta(\alpha))| : \frac{\alpha}{|\alpha|} =: \theta(\alpha) \in Q \right\} = 0.$$

If T is as in the previous corollary, then also

$$\limsup_{|\alpha|\to\infty} \left\{ |\log Y(\alpha)^{1/|\alpha|} - \log T(\theta(\alpha))| : \frac{\alpha}{|\alpha|} =: \theta(\alpha) \in Q \right\} = 0.$$

Proof. Let L denote the first lim sup, and let $\{\alpha_{(j)}\}$ be a sequence for which

$$\lim_{j\to\infty} |Y(\alpha_{(j)})^{1/|\alpha_{(j)}|} - T(\theta_{(j)})| = L,$$

where $\theta_{(j)} = \alpha_{(j)}/|\alpha_{(j)}|$. We may assume that $\theta_{(j)} \to \theta \in Q$ by passing perhaps to a subsequence. Then

$$|Y(\alpha_{(j)})^{1/|\alpha_{(j)}|} - T(\theta_{(j)})| \leq |Y(\alpha_{(j)})^{1/|\alpha_{(j)}|} - T(\theta)| + |T(\theta) - T(\theta_{(j)})|$$

and as $j \to \infty$, the first expression on the right-hand side goes to zero by Lemma 4.3 and the second by continuity of T (Lemma 4.4). So $L = 0$ as required.

If T is as in the previous corollary, then all quantities inside the second lim sup are finite. To prove this second statement, one does a similar argument as above, writing $\log Y(\alpha_{(j)})^{1/|\alpha_{(j)}|}$, $\log T(\theta_{(j)})$, etc. in place of $Y(\alpha_{(j)})^{1/|\alpha_{(j)}|}$, $T(\theta_{(j)})$. $\square$

For a positive integer s, let $h_m(s)$ denote the number of elements in the set $\{\alpha \in \mathbb{Z}_{\geq 0}^m : |\alpha| = s\}$; we have $h_m(s) = \binom{s+m-1}{s} = \frac{(s+m-1)!}{s!(m-1)!}$.

Lemma 4.8. *We have*

$$\frac{1}{h_m(s)} \sum_{|\alpha|=s} \log Y(\alpha)^{1/|\alpha|} \to \frac{1}{\mathrm{vol}(\Sigma_m)} \int_{\Sigma_m^\circ} \log T(\theta)\, d\theta \quad \textit{as } s \to \infty, \tag{4-6}$$

where on the right-hand side we integrate over θ with respect to the usual m-dimensional volume on $\mathbb{R}^m$, with $\mathrm{vol}(\Sigma_m) = \int_{\Sigma_m} d\theta$.

Proof. By Corollary 4.6 we have two cases: either T is never zero on Σ_m° or $T \equiv 0$.

We consider the first case. For convenience write $\theta(\alpha) = \alpha/|\alpha|$. The set $\Sigma(s) := \{\theta(\alpha) : |\alpha| = s\}$ is a uniformly distributed grid of points on Σ_m such that the discrete probability measure $(1/h_m(s)) \sum_{|\alpha|=s} \delta_{\theta(\alpha)}$ supported on $\Sigma(s)$ converges weak* to $(1/\mathrm{vol}(\Sigma_m))d\theta$ as $s \to \infty$. Since $\theta \to T(\theta)$ is a bounded continuous function on Σ_m° and $\mathrm{vol}(\partial\Sigma) = 0$,

$$\frac{1}{h_m(s)} \sum_{|\alpha|=s} \log T(\theta(\alpha)) \to \frac{1}{\mathrm{vol}(\Sigma_m)} \int_{\Sigma_m^\circ} \log T(\theta)\, d\theta \quad \text{as } s \to \infty.$$

(To see this, note that the formula holds by weak* convergence when $\log T(\theta)$ is replaced by $(1-\chi)\log T(\theta)$ with χ an arbitrary smooth cutoff function supported in a neighborhood of $\partial\Sigma$; now shrink the support of χ.)

Hence to prove (4-6), it is sufficient to show that

$$\Big(\frac{1}{h_m(s)}\sum_{|\alpha|=s}|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|\Big)\to 0 \quad \text{as } s\to\infty. \tag{4-7}$$

Fix $\delta>0$ and define the compact set $Q_\delta:=\{\theta=(\theta_1,\ldots,\theta_m)\in\Sigma_m^\circ:\theta_\nu\geq\delta \text{ for all } \nu\}$. For a positive integer s, let

$$L_1(s):=\Big\{\alpha=(\alpha_1,\ldots,\alpha_m)\in\mathbb{Z}^m_{\geq 0}:|\alpha|=s,\ \frac{\alpha}{|\alpha|}\in Q_\delta\Big\}$$

and let $L_2(s):=\Big\{\alpha=(\alpha_1,\ldots,\alpha_m)\in\mathbb{Z}^m_{\geq 0}:|\alpha|=s,\ \dfrac{\alpha}{|\alpha|}\notin Q_\delta\Big\}$; write

$$L_2(s)=\bigcup_{\nu=1}^m\Big\{\alpha\in L_2(s):\frac{\alpha_\nu}{s}<\delta\Big\}=:\bigcup_{\nu=1}^m L_{2,\nu}(s).$$

Using $\alpha_\nu<\delta s$ and $\sum_{\eta\neq\nu}\alpha_\eta\leq s$, we can estimate the size of $L_{2,\nu}(s)$ for each ν as $|L_{2,\nu}(s)|\leq\delta s\binom{s+m-2}{s}$. A calculation then gives

$$\frac{|L_2(s)|}{h_m(s)}=\sum_{\nu=1}^m\frac{|L_{2,\nu}(s)|}{h_m(s)}\leq m\cdot\frac{\delta s\binom{s+m-2}{s}}{\binom{s+m-1}{s}}\leq\delta m^2.$$

Hence

$$\begin{aligned}
&\frac{1}{h_m(s)}\sum_{|\alpha|=s}|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|\\
&\qquad=\frac{1}{h_m(s)}\sum_{\alpha\in L_1(s)}|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|\\
&\qquad\quad+\frac{1}{h_m(s)}\sum_{\alpha\in L_2(s)}|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|\\
&\qquad\leq\frac{|L_1(s)|}{h_m(s)}\sup\{|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|:|\alpha|=s,\ \theta(\alpha)\in Q_\delta\}\\
&\qquad\quad+\frac{|L_2(s)|}{h_m(s)}(\log(C^{\frac{1}{s}}r)+\log r)\\
&\qquad\leq\sup\{|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|:|\alpha|=s,\ \theta(\alpha)\in Q_\delta\}\\
&\qquad\quad+\delta m^2(\log(C^{1/s}r)+\log r),
\end{aligned}$$

with C, r as in Definition 4.1. By Lemma 4.7 the sup in the above line goes to zero as $s \to \infty$, so

$$\limsup_{s\to\infty}\left(\frac{1}{h_m(s)}\sum_{|\alpha|=s}|\log Y(\alpha)^{1/|\alpha|}-\log T(\theta(\alpha))|\right)\le \delta m^2(\log(C^{1/s}r)+\log r).$$

Since $\delta > 0$ was arbitrary, (4-7) follows.

For the case $T \equiv 0$, we need to show that the left-hand side of (4-6) goes to $-\infty$ as $s \to \infty$. Fix a compact set $Q \subseteq \Sigma_m^\circ$. The first part of the previous lemma yields

$$\limsup\{Y(\alpha)^{1/|\alpha|} : |\alpha| \to \infty,\ \alpha/|\alpha| \in Q\} = 0.$$

Hence given $\epsilon > 0$,

$$\sup\{Y(\alpha)^{1/|\alpha|} : |\alpha| > N,\ \alpha/|\alpha| \in Q\} < \epsilon$$

for sufficiently large N. Using the notation $L_1(s)$, $L_2(s)$ from the proof of the first case (with Q in place of Q_δ), we have

$$\frac{1}{h_m(s)}\sum_{\substack{|\alpha|=s\\ \alpha/|\alpha|\in Q}}\log Y(\alpha)^{1/|\alpha|}\le 1/h_m(s)\sum_{\substack{|\alpha|=s\\ \alpha/|\alpha|\in Q}}\log\epsilon=\frac{|L_1(s)|}{h_m(s)}\log\epsilon\le\log\epsilon$$

for $s > N$. Finally, note that $Y(\alpha)^{1/|\alpha|}$ is uniformly bounded above for all α (say by some constant M) since Y has subexponential growth. For all s,

$$\frac{1}{h_m(s)}\sum_{\substack{|\alpha|=s\\ \alpha/|\alpha|\notin Q}}\log Y(\alpha)^{1/|\alpha|}=\frac{|L_2(s)|}{h_m(s)}M\le M.$$

Altogether, $(1/h_m(s))\sum_{|\alpha|=s}\log Y(\alpha)^{1/|\alpha|} \le M + \log\epsilon$ when $s > N$. Since ϵ is arbitrary, the left-hand side of (4-6) goes to $-\infty$ as required. □

5. Chebyshev constants

In this section we construct Chebyshev constants on an algebraic variety $V \subseteq \mathbb{C}^n$. Suppose that V satisfies the properties (3-1). As before, $R := \mathbb{C}[z_1, \dots, z_m] \subseteq \mathbb{C}[V]$ is a Noether normalization, and $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ are the polynomials of Section 2. We will write $\lambda_1, \dots, \lambda_d$ for the interpolating points denoted by $p_1, \dots, p_d$ earlier, so that we can use the letter "p" to denote polynomials. We also introduce some additional notation.

Notation 5.1. Recall that the basis $\mathcal{C}$ of $\mathbb{C}[V]$ was constructed in Definition 3.13, ordered by $\prec$. Denote by $\{\boldsymbol{e}_j\}_{j=1}^\infty$ the enumeration of $\mathcal{C}$ according to $\prec$. For $f = \sum_j a_j\boldsymbol{e}_j \in \mathbb{C}[V]$ we write $\mathrm{LT}_\prec(f) = a_k\boldsymbol{e}_k$ for the leading term, i.e., $a_k \neq 0$ and $a_j = 0$ for all $j > k$. For $f, g \in \mathbb{C}[V]$, write $f \prec g$ if $\mathrm{LT}_\prec(f) \prec \mathrm{LT}_\prec(g)$.

In what follows, α will always denote a multi-index in $\mathbb{Z}^m_{\geq 0}$, and we write $\alpha = (\alpha', \alpha_m)$ where $\alpha' \in \mathbb{Z}^{m-1}_{\geq 0}$ and $\alpha_m \in \mathbb{Z}_{\geq 0}$. For convenience, we will also identify α and α' with $(\alpha_1, \dots, \alpha_m, 0, \dots, 0)$ and $(\alpha_1, \dots, \alpha_{m-1}, 0, \dots, 0)$ in $\mathbb{Z}^n_{\geq 0}$ when using multi-index notation (i.e., in expressions such as z^α).

Definition 5.2. Let $\alpha \in \mathbb{Z}^m_{\geq 0}$ be a multi-index. Define for $i = 1, \dots, d$ the collection of polynomials

$$\mathcal{M}_i(\alpha) := \{p(z) \in \mathbb{C}[V] : p(z) = z^\alpha \boldsymbol{v}_i + g(z),\ g(z) \prec z^\alpha \boldsymbol{v}_i\}.$$

Fix a compact set $K \subseteq V$. We define the function $Y_i : \mathbb{Z}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ by

$$Y_i(\alpha) := \inf\{\|p\|_K : p \in \mathcal{M}_i(\alpha)\}.$$

For a fixed $i \in \{1, \dots, d\}$, we will write $\ell_i(z^\alpha)$ to denote an arbitrary $g \in \mathbb{C}[V]$ with $g \prec z^\alpha \boldsymbol{v}_i$. An immediate consequence of Lemma 3.7 is the following.

Lemma 5.3. *We have $\boldsymbol{v}_i^2 = z_m^t \boldsymbol{v}_i + \ell_i(z_m^t)$ and $\boldsymbol{v}_i \boldsymbol{v}_j = \ell_i(z_m^t)$. Hence if $p \in \mathcal{M}_i(\alpha)$, $q \in \mathcal{M}_i(\tilde{\alpha})$, then $pq \in \mathcal{M}_i(\alpha + \tilde{\alpha} + \gamma_m)$, where $\gamma_m = (0, \dots, 0, t, 0, \dots, 0)$, where the t is in the m-th slot.*

In the above lemma, t is as in (3-4). As a consequence, we obtain a weakly submultiplicative function on $\mathbb{Z}^m_{\geq 0}$, where the set $\mathcal{F}$ in Definition 4.1 may be taken to be the singleton $\{\gamma_m\}$.

Corollary 5.4. *The function Y_i is weakly submultiplicative with subexponential growth. In particular*,

$$Y_i(\alpha + \tilde{\alpha} + \gamma_m) \leq Y_i(\alpha) Y_i(\tilde{\alpha}), \quad \alpha, \tilde{\alpha} \in \mathbb{Z}^m_{\geq 0}.$$

Proof. Fix indices $\alpha, \tilde{\alpha} \in \mathbb{Z}^m_{\geq 0}$. Choose $p \in \mathcal{M}_i(\alpha)$ such that $\|p\|_K = Y_i(\alpha)$ and $q \in \mathcal{M}_i(\tilde{\alpha})$ such that $\|q\|_K = Y_i(\tilde{\alpha})$. By the previous lemma, $pq \in \mathcal{M}_i(\alpha + \tilde{\alpha} + \gamma_m)$, so that $Y_i(\alpha + \tilde{\alpha} + \gamma_m) \leq \|pq\|_K \leq \|p\|_K \|q\|_K = Y_i(\alpha) Y_i(\tilde{\alpha})$.

Choose $r > 1$ such that $K \subseteq B(0, r) = \{z \in \mathbb{C}^n : |z| \leq r\}$. Then $Y_i(\alpha) \leq r^{|\alpha|} \|\boldsymbol{v}_i\|_K$, so Y_i has subexponential growth (choose $C > \max\{1, \|\boldsymbol{v}_i\|_K\}$). □

As a consequence of the results in the previous section, we have the following:

Proposition 5.5. *The limit*

$$T(K, \lambda_i, \theta) := \lim_{\substack{|\alpha| \to \infty \\ \alpha/|\alpha| \to \theta}} Y_i(\alpha)^{1/|\alpha|}$$

exists for each $\theta \in \Sigma_m^\circ$, and $\theta \mapsto T(K, \lambda_i, \theta)$ defines a logarithmically homogeneous function on Σ_m°. Moreover, we have the convergence

$$\frac{1}{h_m(s)} \sum_{|\alpha| = s} \log Y_i(\alpha)^{1/|\alpha|} \to \frac{1}{\mathrm{vol}(\Sigma_m)} \int_{\Sigma_m^\circ} \log T(K, \lambda_i, \theta)\, d\theta \quad \text{as } s \to \infty.$$

Definition 5.6. We call $T(K, \lambda_i, \theta)$ the *directional Chebyshev constant of K associated to λ_i and θ.*

We call

$$T(K, \lambda_i) := \exp\left(\frac{1}{\mathrm{vol}(\Sigma_m)} \int_{\Sigma_m^\circ} \log T(K, \lambda_i, \theta)\, d\theta\right)$$

the *principal Chebyshev constant of K associated to λ_i.*

As in (4-3), we also define $T^-(K, \lambda_i, b) := \liminf_{|\alpha|\to\infty,\, \alpha/|\alpha|\to b} Y_i(\alpha)^{1/|\alpha|}$ for $b \in \partial\Sigma_m$.

In the proof of the main theorem on transfinite diameter, we will need to account for polynomials whose leading terms in $\mathcal{C}$ are of the form $(*)$. For $\alpha' \in \mathbb{Z}_{\geq 0}^{m-1}$ define

$$\tilde{\mathcal{M}}(\alpha') := \{p \in \mathbb{C}[V] : \mathrm{LT}_{\prec}(p) = z^{\alpha'} z_m^l z^\beta \text{ with } z_m^l z^\beta \in \mathcal{B}\}.$$

Recall that this means that $l + |\beta| < t$. Set $\tilde{Y}(\alpha') := \inf\{\|p\|_K : p \in \tilde{\mathcal{M}}(\alpha')\}$. If $K \subseteq B(0, r)$ it is easy to see that

$$\tilde{Y}(\alpha') \leq r^{|\alpha'|}. \tag{5-1}$$

Also, set

$$\tilde{T}(\alpha') := \inf\{\|p\|_K^{1/\deg p} : p \in \tilde{\mathcal{M}}(\alpha')\} \tag{5-2}$$

and define the function

$$\tilde{T}^-(K, \theta') := \liminf_{\substack{|\alpha'|\to\infty \\ \alpha'/|\alpha'|\to\theta'}} \tilde{T}(\alpha')$$

on $\Sigma_{m-1} := \{\theta' = (\theta_1, \ldots, \theta_{m-1}) \in \mathbb{R}^{m-1} : \sum_k \theta_k = 1\}$. We want to get a lower estimate for this quantity. First we make the following observation. Since the monomial $z_m^{t-|\beta|} z^\beta$ is not in $\mathcal{B}$ it must be expressed in $\mathbb{C}[V]$ with respect to the basis $\mathcal{C}$ as

$$z_m^{t-|\beta|} z^\beta = \sum_{i=1}^{d} C_{\beta i} \boldsymbol{v}_i + q(z), \tag{5-3}$$

where $\deg q \leq t$, $\mathrm{LT}_{\prec}(q) \prec \boldsymbol{v}_1$, and not all $C_{\beta i}$ are zero.

Lemma 5.7. *Suppose $C_{\beta i} \neq 0$ for some $i \in \{1, \ldots, d\}$. Then for each $\theta' \in \Sigma_{m-1}$,*

$$T^-(K, \lambda_i, \theta) \leq \tilde{T}^-(K, \theta'), \tag{5-4}$$

where $\theta = (\theta', 0) = (\theta_1, \ldots, \theta_{m-1}, 0) \in \partial\Sigma_m$.

Proof. Fix $\theta' \in \Sigma_{m-1}$ and let $\epsilon > 0$. Let $\{\alpha'_{(j)}\}$ be a sequence in $\mathbb{Z}_{\geq 0}^{m-1}$ with $|\alpha'_{(j)}| \to \infty$, $\alpha'_{(j)}/|\alpha'_{(j)}| \to \theta'$, and $\tilde{T}(\alpha'_{(j)}) \to \tilde{T}^-(K, \theta')$ as $j \to \infty$.

Next, choose a sequence of polynomials $\{p_j\} \subseteq \mathbb{C}[V]$ such that $p_j \in \tilde{\mathcal{M}}(\alpha'_{(j)})$ and $\|p\|_K^{1/\deg p_j} \leq \tilde{T}(\alpha'_{(j)}) + \epsilon$. Since $\mathcal{B}$ is finite, we can assume, by passing perhaps to a subsequence, that $\mathrm{LT}_{\prec}(p_j) = z^{\alpha'_{(j)}} z_m^l z^\beta$, where l and β are the same for all j.

Let $Q := C_{\beta i}^{-1} z_m^{t-l-|\beta|} \boldsymbol{v}_i$ and define $\{P_j\} \subseteq \mathbb{C}[V]$ by $P_j := Qp_j$ for each j. Then a calculation using equation (5-3) and Lemma 5.3 shows that $P_j \in \mathcal{M}_i(\alpha_{(j)})$ where $\alpha_{(j)} = (\alpha'_{(j)}, t - |\beta|)$. Clearly $\alpha_{(j)}/|\alpha_{(j)}| \to \theta$ as $j \to \infty$ since l and $|\beta|$ are bounded from above by t. Now

$$Y_i(\alpha_{(j)})^{1/|\alpha_{(j)}|} \leq \|Q\|_K^{1/|\alpha_{(j)}|} \|p_j\|_K^{1/|\alpha_{(j)}|} \leq \|Q\|_K^{1/|\alpha_{(j)}|} (\tilde{T}(\alpha'_{(j)}) + \epsilon)^{\deg p_j/|\alpha_{(j)}|}.$$

We take the lim inf as $j \to \infty$. We have $T^-(K, \lambda_i, \theta) \leq \tilde{T}^-(K, \theta') + \epsilon$ since $\deg p_j/|\alpha_{(j)}| \to 1$, and (5-4) follows since ϵ was arbitrary. $\square$

Corollary 5.8. *We have*

$$\liminf_{|\alpha'|\to\infty} \tilde{Y}(\alpha')^{1/|\alpha'|} = \liminf_{|\alpha'|\to\infty} \tilde{T}(\alpha) \geq \min\big\{T^-(K, \lambda_i, \theta) : i \in \{1, \dots, d\},\ \theta \in \partial\Sigma_m\big\}.$$

6. The transfinite diameter

Recall that $\{\boldsymbol{e}_j\}_{j=1}^\infty$ denotes the enumeration of the basis $\mathcal{C}$ according to the ordering $\prec$. For a finite set $\{\zeta_1, \dots, \zeta_s\} \subseteq V$, define

$$\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_s) := \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{e}_2(\zeta_1) & \boldsymbol{e}_2(\zeta_2) & \cdots & \boldsymbol{e}_2(\zeta_s) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{e}_s(\zeta_1) & \boldsymbol{e}_s(\zeta_2) & \cdots & \boldsymbol{e}_s(\zeta_s) \end{pmatrix}. \tag{6-1}$$

As in the previous section, fix a compact set $K \subseteq V$. We have $K \subseteq B(0, r) = \{|z| < r\}$ for some $r > 0$.

Notation 6.1. For a positive integer s,

$$V_s := \sup\{|\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_s)| : \{\zeta_1, \dots, \zeta_s\} \subseteq K\}.$$

Also, given any positive integer s, let h_s denote the dimension of $\mathbb{C}[V]_{=s}$, let $m_s := \sum_{\nu=0}^s h_\nu$ denote the dimension of $\mathbb{C}[V]_{\leq s}$, and let $l_s := \sum_{\nu=0}^s \nu h_\nu$ denote the sum of the degrees of the basis elements $\mathcal{C} \cap \mathbb{C}[V]_{\leq s}$.

We now state our main theorem:

Theorem 6.2. *The limit $d(K) = \lim_{s\to\infty} V_{m_s}^{1/l_s}$ exists and we have the formula*

$$d(K) = \left(\prod_{i=1}^d T(K, \lambda_i)\right)^{1/d}.$$

To prove the theorem we will need some lemmas. Recall that $\mathcal{B}$ is the collection of monomials given by (3-7).

Lemma 6.3. *Let s be a positive integer. If $\boldsymbol{e}_s = z^{\alpha}\boldsymbol{v}_i$ for some $i \in \{1, \dots, d\}$, then*

$$Y_i(\alpha) \le \frac{V_s}{V_{s-1}} \le sY_i(\alpha). \tag{6-2}$$

If $\boldsymbol{e}_s = z^{\alpha'} z_m^l z^{\beta}$ with $z_m^l z^{\beta} \in \mathcal{B} \cap \mathcal{C}$, then

$$\tilde{Y}(\alpha') \le \frac{V_s}{V_{s-1}} \le s\tilde{Y}_i(\alpha'). \tag{6-3}$$

Proof. Choose points $\zeta_1, \dots, \zeta_{s-1}$ in K such that $\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_{s-1}) = V_{i-1}$. It is easy to see that the polynomial $P(z) := \mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_{s-1}, z)/\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_{s-1})$ is in $\mathcal{M}(\alpha)$ by expanding the determinant, and hence

$$Y_i(\alpha) \le \|P\|_K \le \frac{V_s}{V_{s-1}},$$

which gives the first inequality of (6-2).

Now choose points $\zeta_1, \dots, \zeta_s$ in K such that $\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_s) = V_i$ and let $t(z) = \boldsymbol{e}_s + \sum_{\nu<s} c_\nu \boldsymbol{e}_\nu$ be a polynomial in $\mathcal{M}(\alpha)$ such that $\|t\|_K = Y_i(\alpha)$. Then by properties of determinants,

$$\begin{aligned} V_i &= \left| \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{e}_2(\zeta_1) & \boldsymbol{e}_2(\zeta_2) & \cdots & \boldsymbol{e}_2(\zeta_s) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{e}_{s-1}(\zeta_1) & \boldsymbol{e}_{s-1}(\zeta_2) & \cdots & \boldsymbol{e}_{s-1}(\zeta_s) \\ t(\zeta_1) & t(\zeta_2) & \cdots & t(\zeta_s) \end{pmatrix} \right| \\ &\le \sum_{\nu=1}^{s} |t(\zeta_\nu)||V(\zeta_1, \dots, \hat{\zeta}_\nu, \dots, \zeta_s)| \le \sum_{\nu=1}^{s} Y_i(\alpha)V_{s-1} = sY_i(\alpha)V_{s-1}, \end{aligned}$$

where we expand along the bottom row. This gives the second inequality of (6-2).

The proof of (6-3) is similar, so we omit it. □

We need to keep track of exponents. Let t be as in Section 3 (see the paragraph following Corollary 3.3). Fix an integer $s > t$. For an element $z^{\alpha}\boldsymbol{v}_i$ there are d choices for i and $h_m(s-t) = \binom{s-t+m-1}{m-1} = \frac{(s-t+m-1)!}{(s-t)!(m-1)!}$ choices for α when $|\alpha| = s-t$. Hence the number of basis elements of degree s of the form $(**)$ is $dh_m(s-t)$.

Let $a_s := h_s - dh_m(s-t)$ be the number of remaining basis elements, of the form $(*)$, i.e., $z^{\alpha'} a z_m^l z^{\beta}$ with $\alpha' \in \mathbb{Z}_{\ge 0}^{m-1}$ and $z_m^l z^{\beta} \in \mathcal{B}$. We then have the estimate $a_s \le |\mathcal{B}|\binom{s+m-2}{m-2}$, where $|\mathcal{B}|$ denotes the size of the set $\mathcal{B}$. Hence

$$\frac{a_s}{h_s} \le \frac{|\mathcal{B}|\binom{s+m-2}{m-2}}{d\binom{s-t+m-1}{m-1}} \to 0 \quad \text{as } s \to \infty, \quad \text{and so} \quad \frac{dh_m(s-t)}{h_s} \to 1. \tag{6-4}$$

Let $\tilde{T}_s := \inf\{\tilde{T}(\alpha') : s-t \le |\alpha'| \le s\}$, where $\tilde{T}(\alpha')$ is as in (5-2). A straightforward corollary of the previous lemma is the following:

Corollary 6.4. *For a positive integer $s > t$, we have*

$$\tilde{T}_s^{sa_s}\Big(\prod_{|\alpha|=s-t}\prod_{i=1}^{d} Y_i(\alpha)\Big) \le \frac{V_{m_s}}{V_{m_{s-1}}} \le \Big(\frac{m_s!}{m_{s-1}!}\Big)^2 r^{sa_s} \prod_{|\alpha|=s-t}\prod_{i=1}^{d} Y_i(\alpha). \tag{6-5}$$

Proof. We apply Lemma 6.3 to the product

$$\frac{V_{m_s}}{V_{m_{s-1}}} = \frac{V_{m_s}}{V_{m_s-1}}\frac{V_{m_s-1}}{V_{m_s-2}}\cdots\frac{V_{m_{s-1}+1}}{V_{m_{s-1}}}.$$

For the upper estimate, we have

$$\begin{aligned}
\frac{V_{m_s}}{V_{m_{s-1}}} &= \frac{V_{m_s}}{V_{m_s-1}}\frac{V_{m_s-1}}{V_{m_s-2}}\cdots\frac{V_{m_{s-1}+1}}{V_{m_{s-1}}}\\
&= \Big(\frac{V_{m_s}}{V_{m_s-1}}\cdots\frac{V_{m_{s-1}+a_s+1}}{V_{m_{s-1}+a_s}}\Big)\Big(\frac{V_{m_{s-1}+a_s}}{V_{m_{s-1}+a_s-1}}\cdots\frac{V_{m_{s-1}+1}}{V_{m_{s-1}}}\Big)\\
&\le \Big(m_s m_{s-1}\cdots(m_{s-1}+a_s+1)\prod_{|\alpha|=s-t}\prod_{i=1}^{d} Y_i(\alpha)\Big)\\
&\qquad\qquad \times\Big((m_{s-1}+a_s)\cdots(m_{s-1}+1)\prod_{\nu=m_{s-1}+1}^{m_{s-1}+a_s}\tilde{Y}(\alpha'(\boldsymbol{e}_\nu))\Big),
\end{aligned}$$

where in the last two lines the first large parentheses apply (6-2) to those fractions V_k/V_{k-1} for which $\boldsymbol{e}_\nu$ is of the form $(**)$[†] while the second large parentheses apply (6-3) to those fractions for which $\boldsymbol{e}_\nu$ is of the form $(*)$. We have also written $\alpha'(\boldsymbol{e}_\nu)$ to denote the multi-index $\alpha' \in \mathbb{Z}_{\ge 0}^{m-1}$ for which $\boldsymbol{e}_\nu = z^{\alpha'} z_m^l z^\beta$. We have

$$\begin{aligned}
&\Big(m_s m_{s-1}\cdots(m_{s-1}+a_s+1)\prod_{|\alpha|=s-t}\prod_{i=1}^{d} Y_i(\alpha)\Big)\\
&\qquad\times\Big((m_{s-1}+a_s)\cdots(m_{s-1}+1)\prod_{\nu=m_{s-1}+1}^{m_{s-1}+a_s}\tilde{Y}(\alpha'(\boldsymbol{e}_\nu))\Big)\\
&\qquad\qquad\le \Big(\frac{m_s!}{m_{s-1}!}\prod_{|\alpha|=s-t}\prod_{i=1}^{d} Y_i(\alpha)\Big)\Big(\frac{m_s!}{m_{s-1}!}\prod_{\nu=m_{s-1}+1}^{m_{s-1}+a_s} r^s\Big)
\end{aligned}$$

where we use (5-1) in the last line. This last expression is the upper estimate in (6-5). The lower estimate follows similarly, using the fact that $s-t \le |\alpha'(\boldsymbol{e}_\nu)| \le s$ for all $\nu = m_{s-1}+1, \dots, m_{s-1}+a_s$, so that $\tilde{Y}(\alpha'(\boldsymbol{e}_\nu)) \ge \tilde{T}_s^s$ for all ν. □

[†] Recall that $\deg(z^\alpha \boldsymbol{v}_i) = s$ when $|\alpha| = s-t$

Similar reasoning as in the paragraphs before the above corollary give

$$m_s \le d\binom{s-t+m}{m} + |\mathcal{B}|\binom{s-t+m-1}{m-1},$$

and when $s > t$,

$$l_s = \sum_{\nu=1}^{s} \nu h_\nu \ge \sum_{\nu=t}^{s} \nu h_\nu \ge \sum_{\nu=1}^{s-t} \nu h_{\nu+t} \ge \sum_{\nu=1}^{s-t} \nu \cdot d\binom{\nu+m-1}{m-1} = dm\binom{s-t+m}{m+1}.$$

Then

$$\frac{m_s}{l_s} \le \frac{m+1}{m(s-t)} + \frac{|\mathcal{B}|(m+1)}{d(s-t)(s-t+m)},$$

in particular $m_s/l_s \to 0$, and

$$1 \le (m_s!)^{1/l_s} \le m_s^{m_s/l_s} \to 1 \quad \text{as } s \to \infty. \tag{6-6}$$

Set $T_s(\lambda_i) := \left(\prod_{|\alpha|=s-t} Y_i(\alpha)\right)^{1/sh_s}$; then (6-5) becomes

$$\tilde{T}_s^{sa_s} \prod_{i=1}^{d} T_s(\lambda_i)^{sh_s} \le \frac{V_{m_s}}{V_{m_{s-1}}} \le r^{sa_s}\left(\frac{m_s!}{m_{s-1}!}\right)^2 \prod_{i=1}^{d} T_s(\lambda_i)^{sh_s}. \tag{6-7}$$

Write $V_{m_s} = (V_{m_s}/V_{m_{s-1}}) \cdots (V_{m_{t+1}}/V_{m_t})V_{m_t}$. Then the above calculation yields the following:

Corollary 6.5.

$$\prod_{\nu=t+1}^{s}\left(\tilde{T}_\nu^{\nu a_\nu} \prod_{i=1}^{d} T_\nu(\lambda_i)^{\nu h_\nu}\right) V_{m_t} \le V_{m_s} \le (m_s!)^2 \prod_{\nu=t+1}^{s}\left(r^{\nu a_\nu} \prod_{i=1}^{d} T_\nu(\lambda_i)^{\nu h_\nu}\right) V_{m_t}.$$

To prove Theorem 6.2 we take l_s-th roots in the above inequality and show that the upper and lower estimates have the desired limit as $s \to \infty$.

Lemma 6.6. *As $s \to \infty$, we have*

$$(m_s!)^{2/l_s} \to 1, \qquad \frac{\sum_{\nu=t+1}^{s} \nu a_\nu}{l_s} \to 0, \quad \textit{and} \quad \frac{sh_s}{(s-t)h_m(s-t)} \to d. \tag{6-8}$$

Proof. The first limit follows immediately from (6-6). Writing the left-hand side of the second limit as $\sum_{\nu=t+1}^{s} \nu a_\nu / \sum_{\nu=1}^{s} \nu h_\nu$, convergence of this limit to zero follows easily from $a_s/h_s \to 0$ (the first limit in (6-4)). The third limit (to d) follows easily from the second limit in (6-4). $\square$

Proof of Theorem 6.2. We first verify that

$$T_s(\lambda_i) \to T(K, \lambda_i)^{1/d} \quad \text{as } s \to \infty. \tag{6-9}$$

By Proposition 5.5,

$$\left(\prod_{|\alpha|=s-t} Y_i(\alpha)\right)^{\frac{1}{(s-t)h_m(s-t)}} = \exp\left(\frac{1}{h_m(s-t)} \sum_{|\alpha|=s-t} \log Y_i(\alpha)^{\frac{1}{|\alpha|}}\right) \to T(K,\lambda_i).$$

Together with the third limit of (6-8) and the definition of $T_s(\lambda_i)$, we get (6-9). In turn, writing $\tilde{l}_s = \sum_{\nu=t+1}^{s} \nu h_\nu$, this gives the convergence

$$\left(\prod_{\nu=t+1}^{s} T_\nu(\lambda_i)^{\nu h_\nu}\right)^{1/\tilde{l}_s} \to T(K,\lambda_i)^{1/d} \quad \text{as } s\to\infty$$

of weighted geometric means. Note that $\tilde{l}_s/l_s \to 1$ as $s\to\infty$, so we may replace $\tilde{l}_s$-th roots with l_s-th roots in what follows. We have

$$\begin{aligned}(m_s!)^{2/l_s} \prod_{\nu=t+1}^{s} &\left(r^{\nu a_\nu} \prod_{i=1}^{d} T_\nu(\lambda_i)^{\nu h_\nu}\right)^{1/l_s} V_{m_t}^{1/l_s} \\ &= (m_s!)^{2/l_s} r^{\sum \nu a_\nu/l_s} \prod_{i=1}^{d}\left(\prod_{\nu=t+1}^{s} T_\nu(\lambda_i)^{\nu h_\nu}\right)^{1/l_s} V_{m_t}^{1/l_s} \to \left(\prod_{i=1}^{d} T(K,\lambda_i)\right)^{1/d}\end{aligned}$$

as $s\to\infty$, which shows that $\limsup_{s\to\infty} V_{m_s}^{1/l_s} \le \left(\prod_{i=1}^{d} T(K,\lambda_i)\right)^{1/d}$.

If $T(K,\lambda_i)=0$ for some i then the theorem is proved, with $d(K)=0$. Otherwise, $T(K,\lambda_i)>0$ for all i; using Corollary 4.6 it is easy to see that $T^-(K,\lambda_i,b)>0$ for all $i=1,\dots,d$ and $b\in\partial\Sigma_m$; and since $\partial\Sigma_m$ is compact, there exists $c>0$ such that $T^-(K,\lambda_i,b)\ge c$ for all i and b. By Lemma 5.7,

$$\liminf_{s\to\infty} \tilde{T}_s \ge \liminf_{|\alpha'|\to\infty} \tilde{T}(\alpha') \ge \min_{\theta'\in\Sigma_{m-1}} \tilde{T}^-(\theta') \ge \min_{i,b} T^-(K,\lambda_i,b) \ge c,$$

so there is some uniform constant $\epsilon\in(0,c)$ such that $T_s>\epsilon$ for all $s>t$, which gives

$$\prod_{\nu=t+1}^{s}\left(\epsilon^{\nu a_\nu} \prod_{i=1}^{d} T_\nu(\lambda_i)^{\nu h_\nu}\right) V_{m_t} \le V_{m_s}.$$

Now the l_s-th root of the left-hand side of the above goes to $\left(\prod_{i=1}^{d} T(K,\lambda_i)\right)^{1/d}$ as $s\to\infty$ by a similar argument as before. This concludes the proof. □

7. Transfinite diameter using the standard basis

In this section we verify that the transfinite diameter of the previous section may be computed in terms of the standard (grevlex) basis of monomials in $\mathbb{C}[V]$. Recall that the basis for normal forms $\mathbb{C}[z]_I$ (where $I = \boldsymbol{I}(V)$) is given by the collection of monomials

$$\{z^\gamma : \gamma \in \mathbb{Z}_{\geq 0},\ z^\gamma \notin \langle \mathrm{LT}(I)\rangle\}.$$

Writing $\{\tilde{\boldsymbol{e}}_j\}_{j=1}^\infty$ for the enumeration of these monomials according to grevlex, define $\mathrm{Van}(\zeta_1, \dots, \zeta_M)$ as in the right-hand side of (6-1) for a finite set $\{\zeta_1, \dots, \zeta_M\} \subseteq V$, replacing the $\boldsymbol{e}_j$ with the $\tilde{\boldsymbol{e}}_j$. Put

$$W_{m_s} := \sup\{|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_s})| : \{\zeta_1, \dots, \zeta_{m_s}\} \subseteq K\}.$$

Later in this section we will need to consider Vandermonde determinants formed from other graded polynomial bases. The Vandermonde determinant associated to a basis $\mathcal{F}$ will be denoted $\mathrm{Van}_{\mathcal{F}}(\,\cdot\,)$.

Lemma 7.1. *Let $\mathcal{F}_1 = \{\tilde{\mathbf{f}}_j\}_{j=1}^\infty$ and $\mathcal{F}_2 = \{\mathbf{f}_j\}_{j=1}^\infty$ be bases of polynomials for $\mathbb{C}[V]$, enumerated according to a graded ordering, and suppose that for some positive integer M, $\tilde{\mathbf{f}}_\tau = \mathbf{f}_\tau$ whenever $\tau > M$. Then there exists a uniform constant $\kappa \neq 0$ such that for any integer $\tau \geq M$ and finite set $\{\zeta_1, \dots, \zeta_\tau\}$,*

$$\mathrm{Van}_{\mathcal{F}_1}(\zeta_1, \dots, \zeta_\tau) = \kappa\, \mathrm{Van}_{\mathcal{F}_2}(\zeta_1, \dots, \zeta_\tau).$$

Proof. Fix the set $\{\zeta_1, \dots, \zeta_\tau\}$ where $\tau \geq M$. Let $E_l = [\tilde{\mathbf{f}}_j(\zeta_k)]_{j,k=1}^l$ and $F_l = [\mathbf{f}_j(\zeta_k)]_{j,k=1}^l$ denote the Vandermonde matrices at the l-th stage for $l = 1, \dots, \tau$. With this notation, we have $E_M = P_M F_M$, where P_M is the change of basis matrix from $\{\tilde{\mathbf{f}}_j\}_{j=1}^M$ to $\{\mathbf{f}_j\}_{j=1}^M$ over the linear space spanned by these polynomials. In particular, $\det P_M \neq 0$. Taking determinants, $\mathrm{Van}_{\mathcal{F}_1}(\zeta_1, \dots, \zeta_M) = \det(P_M)\mathrm{Van}_{\mathcal{F}_2}(\zeta_1, \dots, \zeta_M)$.

Similarly, write $E_\tau = P_\tau F_\tau$; then E_τ and F_τ are of the form

$$E_\tau = \left[\begin{array}{c} E_M \mid * \\ \hline E' \end{array}\right], \quad F_\tau = \left[\begin{array}{c} F_M \mid * \\ \hline E' \end{array}\right],$$

the last rows (denoted by E') being the same since $\boldsymbol{e}_l = \mathbf{f}_l$ when $l > M$. It follows that P_τ must be of the form

$$P_\tau = \left[\begin{array}{c|c} P_M & * \\ \hline 0 & I \end{array}\right]$$

where I denotes the identity matrix, so that $\det P_\tau = \det P_M$.

Taking $\kappa := \det P_M$, the lemma follows immediately. $\square$

Recall that the basis $\mathcal{C}$ of Definition 3.13 is made up of the *normal forms* of two types of polynomials:[‡]

$$(*)\quad z^\alpha z_m^l z^\beta : \alpha \in \mathbb{Z}_{\geq 0}^{m-1}, \quad l + |\beta| \leq t-1$$
$$(**)\quad z^\alpha z_m^l \boldsymbol{v}_i : \alpha \in \mathbb{Z}_{\geq 0}^{m-1}, \quad l \geq 0,\ i = 1, \ldots, d.$$

When these polynomials are already normal forms, as in the examples of Section 3, we have the following theorem:

Theorem 7.2. *Suppose the polynomials* $(*)$ *and* $(**)$ *are already in normal form. Then* $\lim_{s\to\infty} W_{m_s}^{1/l_s} = d(K)$. (*Here* l_s, m_s *are as in Notation 6.1.*)

The idea is to show that $(V_{m_s}^{1/l_s}/W_{m_s}^{1/l_s}) \to 1$ as $s \to \infty$, where V_{m_s} is as in the notation of the previous section. To this end, we analyze the Vandermonde determinants that give these quantities in more detail.

Write

$$\boldsymbol{v}_j(z) = \sum_{\beta \in \mathcal{D}} A_{j\beta} z^\beta, \qquad j = 1, \ldots, d$$

where $\mathcal{D}$ is the collection of all basis monomials that appear in the polynomials $\boldsymbol{v}_j$ for all $j = 1 \ldots, d$. Choose constants $c, C > 0$ such that for any positive integer $k \leq d$,

$$c \leq |\det A| \leq C \tag{7-1}$$

whenever A is a $k \times k$ nonsingular square matrix obtained by deleting sufficiently many rows and columns of the $d \times |\mathcal{D}|$ matrix $[A_{j\beta}]_{j,\beta}$.[§] There are finitely many possible values for $|\det A|$, so we may take the maximum and minimum of these as our constants.

We are interested in $|\mathrm{Van}(\zeta_1, \ldots, \zeta_{m_\tau})|$ for a finite set $\{\zeta_1, \ldots, \zeta_{m_\tau}\}$. The value is the same for any graded ordering of the monomials of $\mathbb{C}[V]_{\leq\tau}$, so let us construct yet another graded ordering that will be convenient for calculation.

Fix the usual grevlex ordering on monomials of degree $< t$. For $\tau \geq t$, and supposing that monomials of degree $< \tau$ have already been ordered, we order the monomials of degree τ as follows. First, list the monomials of the form $(*)$ according to the ordering on $\mathcal{C}$. We set up some convenient notation before continuing.

Notation 7.3. Let $\mathcal{W}_0$ be the set consisting of the monomial basis of $\mathbb{C}[V]_{\leq\tau-1}$ together with the monomials of the form $(*)$ of degree τ. Let $\boldsymbol{W}_0$ denote this same set with our ordering imposed. (With this notation, the matrices given below are uniquely determined.) Also, $\boldsymbol{W}_k$ will have the same meaning when $\mathcal{W}_k$, $k = 1, 2, \ldots$ is defined later in the section.

[‡]cf., Remark 3.10.

[§]Since only the absolute value of the determinant appears, the order of the columns (indexed by β) is not important.

Having listed the monomials in $\mathcal{W}_0$, we will use the elements of $(**)$ to order the remaining monomials in $\mathbb{C}[V]_{\le\tau}$. Before we do this, observe that for $\alpha\in\mathbb{Z}^m_{\ge 0}$,

$$z^\alpha \boldsymbol{v}_j = \sum_{\beta\in\mathcal{D}} A_{j\beta} z^{\alpha+\beta},$$

and since $z^\alpha\boldsymbol{v}_j$ is a normal form, each of the monomials in the sum on the right-hand side is a basis monomial.

Returning to the construction of our ordering, let us enumerate the multi-indices $\alpha\in\mathbb{Z}^m_{\ge 0}$ of total degree $\tau-t$ as $\alpha(1),\alpha(2),\dots$, according to their order of appearance in the elements of the form $(**)$ in $\mathcal{C}$.

The polynomials $\{z^{\alpha(1)}\boldsymbol{v}_j\}_{j=1}^d$ are linearly independent by Theorem 3.12. This allows us to choose, for each $j=1,\dots,d$, a term $z^{\alpha(1)+\beta(j)}$ of $z^{\alpha(1)}\boldsymbol{v}_j$ that is not a term of $z^{\alpha(1)}\boldsymbol{v}_i$ whenever $i<j$. We can also arrange that none of these terms be in $\mathcal{W}_0$ either, since by the construction of $\mathcal{C}$ in Section 2, none of the polynomials $z^{\alpha(1)}\boldsymbol{v}_j$ are in the span of $\mathcal{W}_0$. The set of monomials defined by

$$\mathcal{W}_1 := \{z^\gamma : z^\gamma\in\mathcal{W}_0 \text{ or } z^\gamma = z^{\alpha(1)+\beta(j)}\}$$

is therefore a linearly independent subset of basis monomials in $\mathbb{C}[V]_{\le\tau}$.

Remark 7.4. When $k>1$, note that $z^{\alpha(k)}\boldsymbol{v}_j$ is not in the span of $\mathcal{W}_1$. If it were, then all its monomials would be in $\mathcal{W}_1$, and, irrespective of how one orders the remaining monomials that are not in $\mathcal{W}_1$, the change of basis matrix on $\mathbb{C}[V]_{\le\tau}$ from $\mathcal{C}$ to the monomial basis would not have full rank. This contradicts the fact that a change of basis matrix must be invertible.

Now, write

$$\left[\begin{array}{c} \boldsymbol{W}_0 \\ \hline z^{\alpha(1)}\boldsymbol{v}_1 \\ \vdots \\ z^{\alpha(1)}\boldsymbol{v}_d \\ \hline \text{rest of } \mathcal{C} \\ (\deg\le\tau) \end{array}\right] = \left[\begin{array}{c} \boldsymbol{W}_0 \\ \hline \sum_\beta A_{1\beta} z^{\alpha(1)+\beta} \\ \vdots \\ \sum_\beta A_{d\beta} z^{\alpha(1)+\beta} \\ \hline \text{rest of } \mathcal{C} \\ (\deg\le\tau) \end{array}\right] = \left[\begin{array}{c|c|c} I & 0 & 0 \\ \hline * & A_{(1)} & * \\ \hline 0 & 0 & I \end{array}\right] \left[\begin{array}{c} \boldsymbol{W}_0 \\ \hline z^{\alpha(1)+\beta(1)} \\ \vdots \\ z^{\alpha(1)+\beta(d)} \\ \hline \text{rest of } \mathcal{C} \\ (\deg\le\tau) \end{array}\right],$$

where the (j,k)-th entry in the block $A_{(1)}$ is given by $A_{j\beta}$ with $\beta=\beta(k)$. (The "$*$" in the blocks adjacent to $A_{(1)}$ also consist of entries of the form $A_{j\beta}$ but do not enter into subsequent calculations.) Clearly $c\le\det A_{(1)}\le C$ as in (7-1).

Let us write this more compactly as

$$\left[\begin{array}{c} \boldsymbol{W}_0 \\ \hline \text{rest} \\ \text{of } \mathcal{C} \end{array}\right] = \left[\begin{array}{c|c|c} I & 0 & 0 \\ \hline * & A_{(1)} & * \\ \hline 0 & 0 & I \end{array}\right] \left[\begin{array}{c} \boldsymbol{W}_1 \\ \hline \text{rest} \\ \text{of } \mathcal{C} \end{array}\right].$$

The ordering of the remaining monomials is done by repeating the same process as above with the polynomials $z^{\alpha(2)}, z^{\alpha(3)}, \dots$, in turn, to form $\mathcal{W}_2, \mathcal{W}_3, \dots$, etc. Assuming that $\mathcal{W}_{\nu-1}$ has already been constructed, consider the polynomials $\{z^{\alpha(\nu)}\boldsymbol{v}_j\}_{j=1}^d$. They are linearly independent, and by similar reasoning as in Remark 7.4, none of them are in the span of $\mathcal{W}_{\nu-1}$. Hence they yield d additional basis monomials which, adjoined to $\mathcal{W}_{\nu-1}$, form the set $\mathcal{W}_\nu$. We also have an equation of the form

$$\left[\begin{array}{c} \boldsymbol{W}_{\nu-1} \\ \hline \text{rest} \\ \text{of } \mathcal{C} \end{array}\right] = \left[\begin{array}{c|c|c} I & 0 & 0 \\ \hline * & A_{(\nu)} & * \\ \hline 0 & 0 & I \end{array}\right] \left[\begin{array}{c} \boldsymbol{W}_{\nu} \\ \hline \text{rest} \\ \text{of } \mathcal{C} \end{array}\right], \tag{7-2}$$

with $c \le |\det A_{(\nu)}| \le C$ as in (7-1). This is the main formula needed for the proposition below.

Example 7.5. For the complexified sphere $\boldsymbol{V}(z_1^2+z_2^2+z_3^2-1)$ in $\mathbb{C}^3$, the elements of degree τ in the basis $\mathcal{C}$ are

$$z_1^\tau,\ z_1^{\tau-1}\boldsymbol{v}_1,\ z_1^{\tau-1}\boldsymbol{v}_2,\ z_1^{\tau-2}z_2\boldsymbol{v}_1,\ z_1^{\tau-2}z_2\boldsymbol{v}_2,\ \dots,$$

where $\boldsymbol{v}_1 = \frac{1}{2}(z_2+iz_3)$ and $\boldsymbol{v}_2 = \frac{1}{2}(z_2-iz_3)$. Then

$$\mathcal{W}_0 = \{\dots, z_1^\tau\}, \quad \mathcal{W}_1 = \{\dots, z_1^\tau, z_1^{\tau-1}z_2, z_1^{\tau-1}z_3\}, \quad \mathcal{W}_2 = \mathcal{W}_1 \cup \{z_1^{\tau-2}z_2^2, z_1^{\tau-2}z_2z_3\}.$$

Recall that for a positive integer $\tau \ge t$, $h_m(\tau-t)$ coincides with the number of multi-indices α for which $z^\alpha\boldsymbol{v}_j$ is an element in the basis $\mathcal{C}$ of degree τ, where $j \in \{1, \dots, d\}$. Introduce the notation

$$b_\tau := \sum_{s=t}^{\tau} h_m(s-t).$$

A straightforward calculation shows that

$$b_\tau/l_\tau \to 0 \quad \text{as } \tau \to \infty. \tag{7-3}$$

Proposition 7.6. *For any collection of points* $\{\zeta_1, \dots, \zeta_{m_\tau}\}$, *with* $\tau \ge t$, *we have*

$$c^{b_\tau}|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})| \le |\mathrm{Van}_{\mathcal{C}}(\zeta_1, \dots, \zeta_{m_\tau})| \le C^{b_\tau}|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})|,$$

where c, C *are as in* (7-1).

Proof. The proof is by induction on τ. We concentrate on the upper inequality involving C, and note that the same proof works for the lower inequality. When

$\tau = t$, we have

$$\left[\begin{array}{c} \text{monomials in } (*) \\ \text{of deg} \le t \\ \hline \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_d \end{array}\right] = \left[\begin{array}{c} \boldsymbol{W}_0 \\ \hline \sum_\beta A_{1\beta} z^\beta \\ \vdots \\ \sum_\beta A_{d\beta} z^\beta \end{array}\right] = \left[\begin{array}{c} I \mid 0 \\ \hline A \end{array}\right] [\boldsymbol{W}_1],$$

and note that in this case, $[\boldsymbol{W}_1]$ uses *all* monomials of degree $\le t$. Forming Vandermonde determinants, we have

$$|\mathrm{Van}_C(\zeta_1, \dots, \zeta_{m_t})| = \left|\det\left[\begin{array}{c} I \mid 0 \\ \hline A \end{array}\right] \mathrm{Van}(\zeta_1, \dots, \zeta_{m_t})\right| \le C|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_t})|,$$

where we apply (7-1) and the fact that the determinant in the middle term is the determinant of a $d \times d$ minor of A. This proves the base case.

Suppose the inequality holds when τ is replaced by $\tau - 1$. For $j = 0, \dots, b_\tau$, let us introduce the convenient notation $\mathrm{Van}_j(\zeta_1, \dots, \zeta_{m_\tau})$ for the "intermediate" Vandermonde determinants:

$$\mathrm{Van}_j(\zeta_1, \dots, \zeta_{m_\tau}) = \det\left[\begin{array}{ccc} \boldsymbol{W}_j(\zeta_1) & \cdots & \boldsymbol{W}_j(\zeta_{m_\tau}) \\ \hline z^{\alpha(j+1)}\boldsymbol{v}_1(\zeta_1) & \cdots & z^{\alpha(j+1)}\boldsymbol{v}_1(\zeta_{m_\tau}) \\ \vdots & \ddots & \vdots \\ z^{\alpha(b_\tau)}\boldsymbol{v}_1(\zeta_1) & \cdots & z^{\alpha(b_\tau)}\boldsymbol{v}_d(\zeta_{m_\tau}) \end{array}\right].$$

In particular, $|\mathrm{Van}_{h_m(\tau-t)}(\zeta_1, \dots, \zeta_{m_\tau})| = |\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})|$.

Using equation (7-2),

$$|\mathrm{Van}_{\nu-1}(\zeta_1, \dots, \zeta_{m_\tau})| = |\det(A_{(\nu)})| \cdot |\mathrm{Van}_\nu(\zeta_1, \dots, \zeta_{m_\tau})| \le C|\mathrm{Van}_\nu(\zeta_1, \dots, \zeta_{m_\tau})|$$

for all $\nu = 1, \dots, b_\tau$, and hence by repeated application of the above,

$$|\mathrm{Van}_0(\zeta_1, \dots, \zeta_{m_\tau})| \le C^{h_m(\tau-t)}|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})|.$$

If we define κ by the equation $\mathrm{Van}_C(\zeta_1, \dots, \zeta_{m_\tau-1}) = \kappa\,\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau-1})$, then by Lemma 7.1,

$$\mathrm{Van}_C(\zeta_1, \dots, \zeta_{m_\tau}) = \kappa\,\mathrm{Van}_0(\zeta_1, \dots, \zeta_{m_\tau}),$$

as both determinants use the same elements $\{\boldsymbol{e}_{m_{\tau-1}+1}, \dots, \boldsymbol{e}_{m_\tau}\}$ of degree τ. Also, note that by the inductive hypothesis, we have $|\kappa| \le C^{b_{\tau-1}}$.

Putting everything together,

$$\begin{aligned} |\mathrm{Van}_C(\zeta_1, \dots, \zeta_{m_\tau})| &\le C^{b_{\tau-1}}|\mathrm{Van}_0(\zeta_1, \dots, \zeta_{m_\tau})| \\ &\le C^{b_{\tau-1}+h_m(\tau-t)}|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})| = C^{b_\tau}|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})|, \end{aligned}$$

and the induction is complete. □

Theorem 7.2 is now an easy corollary.

Proof of Theorem 7.2. Let $K \subset V$ be compact. If $W_{m_\tau} = 0$ for some τ, then by a similar argument as in Lemma 7.1, $W_{m_s} = V_{m_s} = 0$ for all $s \geq \tau$, and the theorem follows.

Otherwise, suppose $W_{m_\tau} > 0$ for all τ. It follows from the above proposition that

$$c^{b_\tau} W_{m_\tau} \leq V_{m_\tau} \leq C^{b_\tau} W_{m_\tau}. \tag{7-4}$$

Using (7-3), we have $c^{b_\tau/l_\tau}, C^{b_\tau/l_\tau} \to 1$ as $\tau \to \infty$. Hence dividing by W_{m_τ} and taking l_τ-th roots in (7-4), we have $(V_{m_\tau})^{1/l_\tau}/(W_{m_\tau})^{1/l_\tau} \to 1$ as $\tau \to \infty$. The theorem is proved. □

We close the section by sketching an argument that shows how to get rid of the assumption that the products $z^\alpha z_m^l z^\beta$ and $z^\alpha z_m^l \boldsymbol{v}_j$ used in Theorem 7.2 are normal forms. In general, the methods of this section can be used to construct a basis $\mathcal{W}$ of linearly independent (but not necessarily normal form) monomials on the variety V, made up of the terms in these products. The same proofs also show that transfinite diameter defined in terms of $\mathrm{Van}_{\mathcal{W}}(\,\cdot\,)$ gives the same value as that defined in terms of $\mathrm{Van}_{\mathcal{C}}(\,\cdot\,)$.

Now all monomials in $\mathcal{W}$ are of the form

$$z^\alpha z^\beta = z_1^{\alpha_1} \cdots z_m^{\alpha_m} z_{m+1}^{\beta_{m+1}} \cdots z_n^{\beta_n}$$

with $|\beta| \leq t$, since $\deg \boldsymbol{v}_i = t$ for all i. Given $z^\alpha z^\beta$ as above, consider a monomial $z^\alpha z^{\tilde{\beta}}$ with $|\tilde{\beta}| \leq s$ for some $s \geq t$. Then for any compact set $K \subset V$ that avoids the coordinate axes in $\mathbb{C}^n$,¶ one can find constants m and M such that, upon evaluating these monomials at any point $\zeta \in K$,

$$m^s \leq \frac{|z^\alpha z^{\tilde{\beta}}(\zeta)|}{|z^\alpha z^\beta(\zeta)|} \leq M^s. \tag{7-5}$$

For example, choose an $M > 1$ such that

$$M \geq \frac{\max\{|z| : z \in K\}}{\min\{|z_i| : z = (z_1, \dots, z_n) \in K\})}.$$

All elements of the (grevlex) monomial basis for $\mathbb{C}[V]$ have their total degree in the variables $z_{m+1}, \dots, z_n$ uniformly bounded above (say by $s \geq t$), as a consequence of our hypotheses in Section 3 on Noether normalization. We can therefore compare these basis monomials to those in $\mathcal{W}$ using (7-5).

For an integer $\tau \geq t$ and collection of points $\{\zeta_1, \dots, \zeta_{m_\tau}\} \subset K$, it follows that one can estimate the ratio $|\mathrm{Van}_{\mathcal{W}}(\zeta_1, \dots, \zeta_{m_\tau})|/|\mathrm{Van}(\zeta_1, \dots, \zeta_{m_\tau})|$ with powers of m and M, by repeatedly applying (7-5) to compare rows of the associated Vandermonde matrices. One can verify that the growth of these powers is strictly smaller, as a function of τ, than the growth of l_τ. Finally, a similar argument as carried out in the above proof (forming an equation similar to (7-4), taking l_τ-th roots, etc.) shows

¶Further analysis can be carried out at the end to remove this condition on the axes.

that transfinite diameter defined in terms of $\mathrm{Van}(\cdot)$ gives the same value as that defined in terms of $\mathrm{Van}_{\mathcal{W}}(\cdot)$.

Appendix: The monic basis

In [Rumely et al. 2000], Rumely, Lau and Varley construct the *sectional capacity* of an algebraic variety. As in our case above, Zaharyuta's method plays an essential role. A so-called *monic basis* is constructed on the variety with good multiplicative properties, similar to those of the basis $\mathcal{C}$ from Definition 3.13. Using the monic basis, Chebyshev constants are then defined in terms of normalized polynomial classes, and products of Chebyshev constants give the sectional capacity.

The monic basis of [Rumely et al. 2000, §4] is defined in a very general, abstract setting. For simplicity, let $X \subseteq \mathbb{P}^n$ be an irreducible variety of dimension m and degree d over $\mathbb{C}$. As before, homogeneous coordinates in $\mathbb{P}^n$ are denoted by $z = [z_0 : z_1 : \cdots : z_n]$. Then X gives the graded ring $\mathbb{C}[X] = \mathbb{C}[z]/\mathbf{I}(X)$. The monic basis is a vector space basis of $\mathbb{C}[X]$ consisting of homogeneous elements $\eta_\gamma \in \mathbb{C}[X]_s$. Here is a brief sketch of how the monic basis is constructed:

(1) Write $X = X^{(0)} \supseteq X^{(1)} \supseteq X^{(2)} \supseteq \cdots \supseteq X^{(m-1)}$, where for $\ell = 1, \ldots, m-1$ we have $X^{(\ell)} = \{z \in X^{(\ell-1)} : z_\ell = 0\}$. We assume $X^{(\ell)}$ to be an irreducible variety of dimension $m - \ell$, and that the curve $X^{(m-1)}$ intersects $z_0 = 0$ in distinct smooth points of points of $X^{(m-1)}$; say on the set $D = \{q_1, \ldots, q_d\}$.

(2) Fix a sufficiently large positive integer j_0, so that for $j \geq j_0$,
   (a) For each $i = 1, \ldots, d$ there exists a rational function on $X^{(m-1)}$ with a pole of order j at q_i and no other poles.
   (b) The collection of rational functions on $X^{(m-1)}$ with poles of order at most j on D is isomorphic to the collection of homogeneous polynomials on $X^{(m-1)}$ of degree j.

(3) For each i, j as above, choose a rational function $\eta_{i,j}$ (normalized appropriately) that satisfies part (a) of the previous step. Choose these functions so that the collection $\{\eta_{i,j}\}$ is multiplicatively finitely generated.[‖]

(4) Use these rational functions to construct, for each j, a basis for the homogeneous polynomials of degree j on $X^{(m-1)}$. (Note that these are polynomials in the variables $z_0, z_m, z_{m+1}, \ldots, z_n$ only.)

(5) Construct a basis for homogeneous polynomials on the spaces $X^{(m-2)}, \ldots, X^{(1)}$, X in turn by inductively adjoining monomials in the remaining variables.

The properties of the monic basis and a justification of the above steps is given in [Rumely et al. 2000, §§4 and 5]. See especially their Theorem 4.1.

[‖] This will ensure that the monic basis has good multiplicative properties, as can be seen in Example A.1 below.

Note in particular that the monic basis gives a basis of $\mathbb{C}[X]_s$ for every s. This differs from our setting, where $V \subseteq \mathbb{C}^n$ is an affine variety with coordinate ring $\mathbb{C}[V] = \mathbb{C}[z_1, \ldots, z_n]/I(V)$. The basis $\mathcal{C}$ we construct in Definition 3.13 consists of polynomials that restrict to a basis of $\mathbb{C}[V]_{\leq s}$ for every s. Thus our basis is compatible with a *filtration*, while the monic basis in [Rumely et al. 2000] is compatible with a *grading*.

We illustrate how the two bases are related by examining the monic basis for the complexified sphere considered in Example 3.15.

Example A.1. Let

$$X = \{[z_0 : z_1 : z_2 : z_3] \in \mathbb{P}^3 : z_1^2 + z_2^2 + z_3^2 = z_0^2\} \subseteq \mathbb{P}^3,$$

and $\mathbb{C}[X] = \mathbb{C}[z]/\langle z_1^2 + z_2^2 + z_3^2 - z_0^2\rangle$. Then $X^{(1)}$ is the quadratic curve given by $z_1 = z_2^2 + z_3^2 - z_0^2 = 0$ that intersects $z_0 = 0$ in $[0 : 0 : 1 : \pm i]$.

For each $j = 1, 2, \ldots$, it is easy to see that

$$\eta_{1,j}(z_0, z_2, z_3) := \left(\frac{z_2 + i z_3}{2z_0}\right)^j = \left(\frac{\boldsymbol{v}_1}{z_0}\right)^j$$

defines a rational function on $X^{(1)}$ with a pole of order j at $[0 : 0 : 1 : -i]$ and no other poles. The function defined by

$$\eta_{2,j}(z_0, z_2, z_3) := \left(\frac{z_2 - i z_3}{2z_0}\right)^j = \left(\frac{\boldsymbol{v}_2}{z_0}\right)^j$$

has the same property in relation to $[0 : 0 : 1 : i]$. The rational functions with at most poles of order j at $[0 : 0 : 1 : \pm i]$ are then spanned by

$$\{1, \eta_{1,1}, \eta_{2,1}, \eta_{1,2}, \eta_{2,2}, \ldots, \eta_{1,j}, \eta_{2,j}\}.$$

A multiplicative generating set is $\{1, \eta_{1,1}, \eta_{2,1}\}$.

Clearing denominators (i.e., multiplying by z_0^j) gives the corresponding basis of homogeneous polynomials of degree j on $X^{(1)}$. For example, when $j = 2$ we obtain the polynomials

$$z_0^2, z_0\boldsymbol{v}_1, z_0\boldsymbol{v}_2, \boldsymbol{v}_1^2, \boldsymbol{v}_2^2.$$

To get the basis for the variety X, we adjoin powers of z_1 to basis elements for $X^{(1)}$ using the decomposition $\mathbb{C}[X]_j = z_1\mathbb{C}[X]_{j-1} \oplus \mathbb{C}[X^{(1)}]_j$. When $j = 2$, for example, we compute that

$$\begin{aligned}
\text{(A-1)} \quad \mathbb{C}[X]_2 &= z_1\mathbb{C}[X]_1 \oplus \mathbb{C}[X^{(1)}]_2 \\
&= z_1(z_1\mathbb{C}[X]_0 \oplus \mathbb{C}[X^{(1)}]_1) \oplus \mathbb{C}[X^{(1)}]_2 \\
&= z_1^2\mathbb{C}[X]_0 \oplus z_1\mathbb{C}[X^{(1)}]_1 \oplus \mathbb{C}[X^{(1)}]_2 \\
&= z_1^2 \operatorname{span}\{1\} \oplus z_1 \operatorname{span}\{z_0, \boldsymbol{v}_1, \boldsymbol{v}_2\} \oplus \operatorname{span}\{z_0^2, z_0\boldsymbol{v}_1, z_0\boldsymbol{v}_2, \boldsymbol{v}_1^2, \boldsymbol{v}_2^2\} \\
&= \operatorname{span}\{z_0^2, z_0z_1, z_1^2, z_0\boldsymbol{v}_1, z_1\boldsymbol{v}_1, \boldsymbol{v}_1^2, z_0\boldsymbol{v}_2, z_1\boldsymbol{v}_2, \boldsymbol{v}_2^2\}.
\end{aligned}$$

The last line gives the monic basis for $j=2$, where the basis elements are listed according to the ordering used in [Rumely et al. 2000].

For arbitrary j, monic basis elements $\mathbb{C}[X]_j$ are either monomials in z_0 and z_1 of degree j, or are homogeneous polynomials of the form $z_0^{\alpha_0}z_1^{\alpha_1}\boldsymbol{v}_i^{\alpha_2}$ with $\alpha_0+\alpha_1+\alpha_2=j$. Monomials in z_0, z_1 are listed first in lexicographic order (with z_0 preceding z_1), followed by elements of the form $z_0^{\alpha_0}z_1^{\alpha_1}\boldsymbol{v}_i^{\alpha_2}$. The latter are listed in increasing order on i, then lexicographically by $\alpha=(\alpha_0,\alpha_1,\alpha_2)\in\mathbb{Z}^3_{\geq 0}$. This completes the construction of the monic basis for X.

The monic basis constructed in Example A.1 involves arbitrarily large powers of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. This is related to the multiplicative properties of the monic basis described in [Rumely et al. 2000, Theorem 4.1].

It is interesting to compare the monic basis of Example A.1 to the basis constructed in Example 3.15. There, we worked with

$$V=\mathbf{V}(z_1^2+z_2^2+z_3^2-1)\subseteq\mathbb{C}^3.$$

Since the Zariski closure of V is $\bar{V}=X=\mathbf{V}(z_1^2+z_2^2+z_3^2-z_0^2)\subseteq\mathbb{P}^3$, homogenization with respect to z_0 induces an isomorphism

$$\mathbb{C}[V]_{\leq j}\simeq\mathbb{C}[X]_j$$

for all j. It follows that the basis of Example 3.15, when restricted to elements of degree $\leq j$, gives a basis of $\mathbb{C}[X]_j$. However, this basis differs from the monic basis in degree j. For example, when $j=2$, homogenizing the basis of Example 3.15 in degree ≤ 2 gives the homogeneous polynomials

$$z_0^2,\ z_0z_1,\ z_0\boldsymbol{v}_1,\ z_0\boldsymbol{v}_2,\ z_1^2,\ z_1\boldsymbol{v}_1,\ z_1\boldsymbol{v}_2,\ z_2\boldsymbol{v}_1,\ z_2\boldsymbol{v}_2.$$

Comparing this to the last line of (A-1), we see that in degree 2, the monic basis uses $\boldsymbol{v}_1^2$ and $\boldsymbol{v}_2^2$, while our basis uses $z_1\boldsymbol{v}_1$ and $z_1\boldsymbol{v}_2$. These are related by

$$\boldsymbol{v}_1^2=z_1\boldsymbol{v}_1+\tfrac{1}{4}z_1^2-\tfrac{1}{4}z_0^2,\qquad \boldsymbol{v}_2^2=z_1\boldsymbol{v}_2+\tfrac{1}{4}z_1^2-\tfrac{1}{4}z_0^2.$$

At the conceptual level, the basis $\mathcal{C}$ constructed in Definition 3.13 focuses on the *module* properties of the basis, as highlighted in Theorem 3.12. In contrast, the monic basis constructed in [Rumely et al. 2000] focuses on the *multiplicative* properties of the basis. In our treatment, the multiplicative properties of $\mathcal{C}$ follow from Lemma 3.7. Our construction is more direct (we avoid the inductive approach needed in that work) but less general.

Acknowledgement

Sione Ma‘u was partially supported by University of Auckland grant no. 3712302.

References

[Baleikorocau and Ma‘u 2015] W. Baleikorocau and S. Ma‘u, “Chebyshev constants, transfinite diameter, and computation on complex algebraic curves”, *Comput. Methods Funct. Theory* **15**:2 (2015), 291–322. MR Zbl

[Berman and Boucksom 2010] R. Berman and S. Boucksom, “Growth of balls of holomorphic sections and energy at equilibrium”, *Invent. Math.* **181**:2 (2010), 337–394. MR Zbl

[Bloom and Levenberg 2003] T. Bloom and N. Levenberg, “Weighted pluripotential theory in $\mathbb{C}^N$”, *Amer. J. Math.* **125**:1 (2003), 57–103. MR Zbl

[Bloom and Levenberg 2010] T. Bloom and N. Levenberg, “Transfinite diameter notions in $\mathbb{C}^N$ and integrals of Vandermonde determinants”, *Ark. Mat.* **48**:1 (2010), 17–40. MR Zbl

[Cox et al. 1997] D. Cox, J. Little, and D. O’Shea, *Ideals, varieties, and algorithms*, 2nd ed., Springer, New York, 1997. MR Zbl

[Greuel and Pfister 2002] G.-M. Greuel and G. Pfister, *A Singular introduction to commutative algebra*, Springer, Berlin, 2002. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR Zbl

[Jędrzejowski 1991] M. Jędrzejowski, “The homogeneous transfinite diameter of a compact subset of $\mathbb{C}^N$”, *Ann. Polon. Math.* **55** (1991), 191–205. MR Zbl

[Ma‘u 2017] S. Ma‘u, “Okounkov bodies and transfinite diameter”, preprint, 2017. arXiv

[Rumely and Lau 1994] R. Rumely and C. F. Lau, “Arithmetic capacities on $\mathbb{P}^N$”, *Math. Z.* **215**:4 (1994), 533–560. MR Zbl

[Rumely et al. 2000] R. Rumely, C. F. Lau, and R. Varley, *Existence of the sectional capacity*, Mem. Amer. Math. Soc. **690**, American Mathematical Society, Providence, RI, 2000. MR Zbl

[Witt Nyström 2014] D. Witt Nyström, “Transforming metrics on a line bundle to the Okounkov body”, *Ann. Sci. Éc. Norm. Supér.* (4) **47**:6 (2014), 1111–1161. MR Zbl

[Zaharyuta 1975] V. P. Zaharyuta, “Transfinite diameter, Chebyshev constants and capacity for a compactum in $\mathbb{C}^n$”, *Mat. Sb.* (*N.S.*) **96**:3 (1975), 374–389. In Russian; translated in *Mat. USSR Sb.* **25**:3 (1975), 350–364. MR Zbl

Received March 16, 2016. Revised February 11, 2017.

DAVID A. COX
DEPARTMENT OF MATHEMATICS
AMHERST COLLEGE
AMHERST, MA 01002
UNITED STATES
dacox@amherst.edu

SIONE MA‘U
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF AUCKLAND
AUCKLAND 1142
NEW ZEALAND
s.mau@auckland.ac.nz

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.319

A UNIVERSAL CONSTRUCTION OF UNIVERSAL DEFORMATION FORMULAS, DRINFELD TWISTS AND THEIR POSITIVITY

CHIARA ESPOSITO, JONAS SCHNITZER AND STEFAN WALDMANN

We provide an explicit construction of star products on $\mathcal{U}(\mathfrak{g})$-module algebras by using the Fedosov approach. This allows us to give a constructive proof to Drinfeld's theorem and to obtain a concrete formula for Drinfeld twists. We prove that the equivalence classes of twists are in one-to-one correspondence with the second Chevalley–Eilenberg cohomology of the Lie algebra $\mathfrak{g}$. Finally, we show that for Lie algebras with Kähler structure we obtain a strongly positive universal deformation of *-algebras by using a Wick-type deformation. This results in a positive Drinfeld twist.

1. Introduction

The concept of deformation quantization was defined by Bayen, Flato, Fronsdal, Lichnerowicz and Sternheimer in [Bayen et al. 1978a; 1978b] based on Gerstenhaber's theory [1964] of associative deformations of algebra. A formal star product on a symplectic (or Poisson) manifold M is defined as a formal associative deformation $\star$ of the algebra of smooth functions $\mathscr{C}^\infty(M)$ on M. The existence as well as the classification of star products has been studied in many different settings, e.g., in [De Wilde and Lecomte 1983; Fedosov 1986; 1994; 1996; Kontsevich 2003; Nest and Tsygan 1995; Bertelson et al. 1997]; see also the textbooks [Esposito 2015;

MSC2010: primary 17B37; secondary 53D55, 81R50.
Keywords: Star Product, Drinfeld twist, universal deformation formula.

Waldmann 2007] for more details in deformation quantization. Quite parallel to this, Drinfeld introduced the notion of quantum groups and started the deformation of Hopf algebra; see, e.g., the textbooks [Kassel 1995; Chari and Pressley 1994; Etingof and Schiffmann 1998] for a detailed discussion.

It turned out that under certain circumstances one can give simple and fairly explicit formulas for associative deformations of algebras: whenever a Lie algebra $\mathfrak{g}$ acts on an associative algebra $\mathscr{A}$ by derivations, the choice of a *formal Drinfeld twist* $\mathcal{F} \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ allows one to deform $\mathscr{A}$ by means of a *universal deformation formula*

$$a \star_{\mathcal{F}} b = \mu_{\mathscr{A}}(\mathcal{F} \triangleright (a \otimes b)) \tag{1-1}$$

for $a, b \in \mathscr{A}[[t]]$. Here

$$\mu_{\mathscr{A}} : \mathscr{A} \otimes \mathscr{A} \to \mathscr{A}$$

is the algebra multiplication and $\triangleright$ is the action of $\mathfrak{g}$ extended to the universal enveloping algebra $\mathscr{U}(\mathfrak{g})$ and then to $\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g})$ acting on $\mathscr{A} \otimes \mathscr{A}$. Finally, all operations are extended $\mathsf{R}[[t]]$-multilinearly to formal power series. Recall that a formal Drinfeld twist [Drinfeld 1983; 1986] is an invertible element

$$\mathcal{F} \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$$

satisfying

$$(\Delta \otimes \mathsf{id})(\mathcal{F})(\mathcal{F} \otimes 1) = (\mathsf{id} \otimes \Delta)(\mathcal{F})(1 \otimes \mathcal{F}), \tag{1-2}$$

$$(\epsilon \otimes 1)\mathcal{F} = 1 = (1 \otimes \epsilon)\mathcal{F}, \tag{1-3}$$

$$\mathcal{F} = 1 \otimes 1 + \mathcal{O}(t). \tag{1-4}$$

The properties of a twist are now easily seen to guarantee that (1-1) is indeed an associative deformation.

Yielding the explicit formula for the deformation universally in the algebra $\mathscr{A}$, Drinfeld twists are considered to be of great importance in deformation theory in general, and in fact, are used at many different places. We just mention a few recent developments, certainly not exhaustive: Giaquinto and Zhang studied the relevance of universal deformation formulas like (1-1) in great detail in the seminal paper [Giaquinto and Zhang 1998]. Bieliavsky and Gayral [2015] used universal deformation formulas also in a nonformal setting by replacing the notion of a Drinfeld twist with a certain integral kernel. This sophisticated construction leads to a wealth of new strict deformations having the above formal deformations as asymptotic expansions. But also beyond pure mathematics the universal deformation formulas found applications, e.g., in the construction of quantum field theories on noncommutative spacetimes; see, e.g., [Aschieri and Schenkel 2014].

In characteristic zero, there is one fundamental example of a Drinfeld twist in the case of an abelian Lie algebra $\mathfrak{g}$. Here one chooses any bivector $\pi \in \mathfrak{g} \otimes \mathfrak{g}$ and considers the formal exponential

$$\mathcal{F}_{\text{Weyl–Moyal}} = \exp(t\pi), \tag{1-5}$$

viewed as element in $(\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$. An easy verification shows that this is indeed a twist. The corresponding universal deformation formula goes back at least till [Gerstenhaber 1968, Theorem 8] under the name of *deformation by commuting derivations*. In deformation quantization the corresponding star product is the famous Weyl–Moyal star product if one takes π to be antisymmetric.

While this is an important example, it is not at all easy to find explicit formulas for twists in the general nonabelian case. A starting point is the observation that the antisymmetric part of the first order of a twist, $\mathcal{F}_1 - \mathsf{T}(\mathcal{F}_1)$, where T is the usual flip isomorphism, is first an element in $\Lambda^2\mathfrak{g}$ instead of $\Lambda^2\mathscr{U}(\mathfrak{g})$, and second a *classical r-matrix*. This raises the question whether one can go the opposite direction of a quantization: does every classical r-matrix $r \in \Lambda^2\mathfrak{g}$ on a Lie algebra $\mathfrak{g}$ arise as the first order term of a formal Drinfeld twist? It is now a celebrated theorem of Drinfeld [1983, Theorem 6] that this is true.

But even more can be said: given a twist $\mathcal{F}$ one can construct a new twist by conjugating with an invertible element $S \in \mathscr{U}(\mathfrak{g})[[t]]$ starting with $S = 1 + \mathcal{O}(t)$ and satisfying $\epsilon(S) = 1$. More precisely,

$$\mathcal{F}' = \Delta(S)^{-1}\mathcal{F}(S \otimes S) \tag{1-6}$$

turns out to be again a twist. In fact, this defines an equivalence relation on the set of twists, preserving the semiclassical limit, i.e., the induced r-matrix. In the spirit of Kontsevich's formality theorem, and in fact building on its techniques, Halbout [2006] showed that the equivalence classes of twists quantizing a given classical r-matrix are in bijection with the equivalence classes of formal deformations of the r-matrix in the sense of r-matrices. In fact, this follows from Halbout's more profound result on formality for general Lie bialgebras; the quantization of r-matrices into twists is just a special case thereof. His theorem holds in a purely algebraic setting (in characteristic zero) but relies heavily on the fairly inexplicit formality theorems of Kontsevich [2003] and Tamarkin [1998] which in turn require a rational Drinfeld associator.

On the other hand, there is a simpler approach to the existence of twists in the case of real Lie algebras: in seminal work of Drinfeld [1983] he showed that a twist is essentially the same as a left G-invariant star product on a Lie group G with Lie algebra $\mathfrak{g}$, by identifying the G-invariant bidifferential operators on G with elements in $\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g})$. The associativity of the star product gives then immediately the properties necessary for a twist and vice versa. Moreover, an r-matrix is nothing else as a left G-invariant Poisson structure; see his Theorem 1. In that paper, Drinfeld also gives an existence proof of such G-invariant star products and therefore of twists; see Theorem 6. His argument uses the canonical star product on the dual of a central extension of the Lie algebra by the cocycle defined by the (inverse of the) r-matrix, suitably pulled back to the Lie group; see also Remark 5.8 for further details.

The equivalence of twists translates into the usual G-invariant equivalence of star products as discussed in [Bertelson et al. 1998]. Hence one can use the existence (and classification) theorems for invariant star products to yield the corresponding theorems for twists, a fact we learned from personal communication with Beliavsky. This is also the point of view taken by Dolgushev et al. in [Dolgushev et al. 2002], where the star product is constructed in a way inspired by Fedosov's construction of star products on symplectic manifolds.

A significant simplification concerning the existence comes from the observation that for every r-matrix $r \in \Lambda^2\mathfrak{g}$ there is a Lie subalgebra of $\mathfrak{g}$, namely

$$\mathfrak{g}_r = \{(\alpha \otimes \mathsf{id})(r) \mid \alpha \in \mathfrak{g}^*\}, \tag{1-7}$$

such that $r \in \Lambda^2\mathfrak{g}_r$ and r becomes *nondegenerate* as an r-matrix on this Lie subalgebra [Etingof and Schiffmann 1998, Propositions 3.2–3.3]. Thus it will always be sufficient to consider nondegenerate classical r-matrices when interested in the existence of twists. For the classification this is of course not true since a possibly degenerate r-matrix might be deformed into a nondegenerate one only in higher orders: here one needs Halbout's results for possibly degenerate r-matrices. However, starting with a nondegenerate r-matrix, one will have a much simpler classification scheme as well.

The aim of this paper is now twofold: On the one hand, we want to give a direct construction to obtain the universal deformation formulas for algebras acted upon by a Lie algebra with nondegenerate r-matrix. This will be obtained in a purely algebraic fashion for sufficiently nice Lie algebras and algebras over a commutative ring R containing the rationals. Our approach is based on a certain adaptation of the Fedosov construction of symplectic star products, which is in some sense closer to the original Fedosov construction compared to the approach of [Dolgushev et al. 2002] but yet completely algebraic. More precisely, the construction will not involve a twist at all but just the classical r-matrix. Moreover, it will be important to note that we can allow for a nontrivial symmetric part of the r-matrix, provided a certain technical condition on it is satisfied. This will produce deformations with more specific features: as in usual deformation quantization one is not only interested in the Weyl–Moyal like star products, but certain geometric circumstances require more particular star products like Wick-type star products on Kähler manifolds [Karabegov 1996; 2013; Bordemann and Waldmann 1997] or standard-ordered star products on cotangent bundles [Bordemann et al. 1998; 2003].

On the other hand, we give an alternative construction of Drinfeld twists, again in the purely algebraic setting, based on the above correspondence to star products but avoiding the techniques from differential geometry completely in order to be able to work over a general field of characteristic zero. We also obtain a classification of the above restricted situation where the r-matrix is nondegenerate.

In fact, both questions turn out to be intimately linked since applying our universal deformation formula to the tensor algebra of $\mathscr{U}(\mathfrak{g})$ will yield a deformation of the tensor product which easily allows one to construct the twist. This is remarkable insofar as the tensor algebra is of course rigid, the deformation is equivalent to the undeformed tensor product, but the deformation is not the identity, allowing one therefore to consider nontrivial products of elements in $\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$.

We show that the universal deformation formula we construct in fact coincides with (1-1) for the twist we construct. However, it is important to note the detour via the twist is not needed to obtain the universal deformation of an associative algebra.

Finally, we add the notion of positivity: this seems to be new in the whole discussion of Drinfeld twists and universal deformation formulas so far. To this end we consider now an ordered ring R containing $\mathbb{Q}$ and its complex version $\mathsf{C} = \mathsf{R}(\mathrm{i})$ with $\mathrm{i}^2 = -1$, and *-algebras over C with a *-action of the Lie algebra $\mathfrak{g}$, which is assumed to be a Lie algebra over R admitting a Kähler structure. Together with the nondegenerate r-matrix we can define a Wick-type universal deformation which we show to be *strongly positive*: every undeformed positive linear functional stays positive also for the deformation. Applied to the twist we conclude that the Wick-type twist is a convex series of positive elements.

The paper is organized as follows. In Section 2 we explain the elements of the (much more general) Fedosov construction which we will need. Section 3 contains the construction of the universal deformation formula. Here not only the deformation formula will be universal for all algebras $\mathscr{A}$ but also the construction itself will be universal for all Lie algebras $\mathfrak{g}$. In Section 4 we construct the Drinfeld twist while Section 5 contains the classification in the nondegenerate case. Finally, Section 6 discusses the positivity of the Wick-type universal deformation formula. In two Appendices we collect some more technical arguments and proofs. The results of this paper are partially based on the master thesis [Schnitzer 2016].

For symplectic manifolds with suitable polarizations one can define various types of star products with separation of variables [Karabegov 1996; 2013; Bordemann and Waldmann 1997; Donin 2003; Bordemann et al. 1998; 1999; 2003] which have specific properties adapted to the polarization. The general way to construct (and classify) them is to modify the Fedosov construction by adding suitable symmetric terms to the fiberwise symplectic Poisson tensor. We have outlined that this can be done for twists as well in the Kähler case, but there remain many interesting situations. In particular a more cotangent bundle-like polarization might be useful. We plan to come back to these questions in a future project.

2. The Fedosov Setup

In the following we present the Fedosov approach in the particular case of a Lie algebra $\mathfrak{g}$ with a nondegenerate r-matrix r. We follow the presentation of the Fedosov approach given in [Waldmann 2007] but replace differential geometric concepts by algebraic versions in order to be able to treat not only the real case. The setting for this work will be to assume that $\mathfrak{g}$ is a Lie algebra over a commutative ring R containing the rationals $\mathbb{Q} \subseteq \mathsf{R}$ such that $\mathfrak{g}$ is a finite-dimensional free module.

We denote by $\{e_1, \dots, e_n\}$ a basis of $\mathfrak{g}$ and by $\{e^1, \dots, e^n\}$ its dual basis of $\mathfrak{g}^*$. We also assume the r-matrix $r \in \Lambda^2\mathfrak{g}$ to be nondegenerate in the strong sense from the beginning, since, at least in the case of R being a field, we can replace $\mathfrak{g}$ by $\mathfrak{g}_r$ from (1-7) if necessary. Hence r induces the *musical isomorphism*

$$\sharp : \mathfrak{g}^* \to \mathfrak{g} \tag{2-1}$$

by pairing with r, the inverse of which we denote by $\flat$ as usual. Then the defining property of an r-matrix is $[\![r, r]\!] = 0$, where $[\![\,\cdot\,,\,\cdot\,]\!]$ is the unique extension of the Lie bracket to $\Lambda^\bullet\mathfrak{g}$ turning the Grassmann algebra into a Gerstenhaber algebra. Since we assume r to be (strongly) nondegenerate we have the inverse $\omega \in \Lambda^2\mathfrak{g}^*$ of r and $[\![r, r]\!] = 0$ becomes equivalent to the linear condition $\delta_{\mathrm{CE}}\omega = 0$, where δ_{CE} is the usual Chevalley–Eilenberg differential. Moreover, the musical isomorphisms intertwine δ_{CE} on $\Lambda^\bullet\mathfrak{g}^*$ with the differential $[\![r, \cdot\,]\!]$ on $\Lambda^\bullet\mathfrak{g}$. We refer to ω as the induced symplectic form.

Remark 2.1. For the Lie algebra $\mathfrak{g}$ there seems to be little gain in allowing a ring R instead of a field $\mathbb{K}$ of characteristic zero, as we have to require $\mathfrak{g}$ to be a free module and (2-1) to be an isomorphism. However, for the algebras which we would like to deform there will be no such restrictions later on. Hence allowing for algebras over rings in the beginning seems to be the cleaner way to do it, since after the deformation we will arrive at an algebra over a ring, namely $\mathsf{R}[\![t]\!]$ anyway.

Definition 2.2 (Formal Weyl algebra). The algebra $\left(\prod_{k=0}^\infty \mathrm{S}^k\mathfrak{g}^* \otimes \Lambda^\bullet\mathfrak{g}^*\right)[\![t]\!]$ is called the formal Weyl algebra where the product μ is defined by

$$(f \otimes \alpha) \cdot (g \otimes \beta) = \mu(f \otimes \alpha, g \otimes \beta) = f \vee g \otimes \alpha \wedge \beta. \tag{2-2}$$

for any factorizing tensors $f \otimes \alpha, g \otimes \beta \in \mathcal{W} \otimes \Lambda^\bullet$ and extended $\mathsf{R}[\![t]\!]$-bilinearly. We write $\mathcal{W} = \prod_{k=0}^\infty \mathrm{S}^k\mathfrak{g}^*[\![t]\!]$ and $\Lambda^\bullet = \Lambda^\bullet\mathfrak{g}^*[\![t]\!]$.

Since $\mathfrak{g}$ is assumed to be finite-dimensional we have

$$\mathcal{W} \otimes \Lambda^\bullet = \left(\prod_{k=0}^\infty \mathrm{S}^k\mathfrak{g}^* \otimes \Lambda^\bullet\mathfrak{g}^*\right)[\![t]\!]. \tag{2-3}$$

Since we will deform this product μ we shall refer to μ also as the *undeformed product* of $\mathcal{W} \otimes \Lambda^\bullet$. It is clear that μ is associative and graded commutative with respect to the antisymmetric degree. In order to handle this and various other degrees, it is useful to introduce the degree maps

$$\deg_{\mathrm{s}}, \deg_{\mathrm{a}}, \deg_t : \mathcal{W} \otimes \Lambda^\bullet \to \mathcal{W} \otimes \Lambda^\bullet, \tag{2-4}$$

defined by the conditions

$$\deg_{\mathrm{s}}(f \otimes \alpha) = kf \otimes \alpha \quad \text{and} \quad \deg_{\mathrm{a}}(f \otimes \alpha) = \ell f \otimes \alpha \tag{2-5}$$

for $f \in \mathrm{S}^k\mathfrak{g}^*$ and $\alpha \in \Lambda^\ell \mathfrak{g}^*$. We extend these maps to formal power series by $\mathsf{R}[[t]]$-linearity. Then we can define the degree map $\deg_t$ by

$$\deg_t = t\frac{\partial}{\partial t}, \tag{2-6}$$

which is, however, not $\mathsf{R}[[t]]$-linear. Finally, the *total degree* is defined by

$$\mathrm{Deg} = \deg_{\mathrm{s}} + 2\deg_t. \tag{2-7}$$

It will be important that all these maps are derivations of the undeformed product μ of $\mathcal{W} \otimes \Lambda^\bullet$. We denote by

$$\mathcal{W}_k \otimes \Lambda^\bullet = \bigcup_{r \geq k} \{a \in \mathcal{W} \otimes \Lambda^\bullet \mid \mathrm{Deg}\, a = ra\} \tag{2-8}$$

the subspace of elements which have total degree bigger or equal to $+k$. This endows $\mathcal{W} \otimes \Lambda^\bullet$ with a complete filtration, a fact which we shall frequently use in the sequel. Moreover, the filtration is compatible with the undeformed product (2-2) in the sense that

$$ab \in \mathcal{W}_{k+\ell} \otimes \Lambda^\bullet \quad \text{for } a \in \mathcal{W}_k \otimes \Lambda^\bullet \text{ and } b \in \mathcal{W}_\ell \otimes \Lambda^\bullet. \tag{2-9}$$

Following the construction of Fedosov, we define the operators δ and δ^* by

$$\delta = e^i \wedge \mathrm{i}_{\mathrm{s}}(e_i) \quad \text{and} \quad \delta^* = e^i \vee \mathrm{i}_{\mathrm{a}}(e_i), \tag{2-10}$$

where i_{s} and i_{a} are the symmetric and antisymmetric insertion derivations. Both maps are graded derivations of μ with respect to the antisymmetric degree: δ lowers the symmetric degree by one and raises the antisymmetric degree by one; for δ^* it is the other way round. For homogeneous elements $a \in \mathrm{S}^k\mathfrak{g}^* \otimes \Lambda^\ell\mathfrak{g}^*$ we define

$$\delta^{-1}(a) = \begin{cases} 0 & \text{if } k + \ell = 0, \\ 1/(k+\ell)\delta^*(a) & \text{else,} \end{cases} \tag{2-11}$$

and extend this $\mathsf{R}[[t]]$-linearly. Notice that this map is not the inverse of δ; instead we have the following properties:

Lemma 2.3. *For* δ, δ^* *and* δ^{-1} *defined above*, $\delta^2 = (\delta^*)^2 = (\delta^{-1})^2 = 0$ *and*

$$\delta\delta^{-1} + \delta^{-1}\delta + \sigma = \mathsf{id}, \tag{2-12}$$

where σ *is the projection on the symmetric and antisymmetric degree zero.*

In fact, this can be seen as the polynomial version of the Poincaré lemma: δ corresponds to the exterior derivative and δ^{-1} is the standard homotopy.

The next step consists of deforming the product μ into a noncommutative one: we define the *star product* $\circ_\pi$ for $a, b \in \mathcal{W} \otimes \Lambda^\bullet$ by

$$a \circ_\pi b = \mu \circ \mathrm{e}^{(t/2)\mathcal{P}}(a \otimes b), \quad \text{where} \quad \mathcal{P} = \pi^{ij}\, \mathrm{i}_{\mathrm{s}}(e_i) \otimes \mathrm{i}_{\mathrm{s}}(e_j), \tag{2-13}$$

for $\pi^{ij} = r^{ij} + s^{ij}$, where r^{ij} are the coefficients of the r-matrix and $s^{ij} = s(e^i, e^j) \in \mathsf{R}$ are the coefficients of a *symmetric* bivector $s \in \mathrm{S}^2\mathfrak{g}$. When taking $s = 0$ we denote $\circ_\pi$ simply by $\circ_{\mathrm{Weyl}}$.

Proposition 2.4. *The star product* $\circ_\pi$ *is an associative* $\mathsf{R}[[t]]$*-bilinear product on* $\mathcal{W} \otimes \Lambda^\bullet$ *deforming* μ *in zeroth order of* t. *Moreover*, *the maps* δ, $\deg_{\mathrm{a}}$, *and* Deg *are graded derivations of* $\circ_\pi$ *of antisymmetric degree* $+1$ *for* δ *and* 0 *for* $\deg_{\mathrm{a}}$ *and* Deg, *respectively.*

Proof. The associativity follows from the fact that the insertion derivations are commuting; see [Gerstenhaber 1968, Theorem 8]. The statement about δ, $\deg_{\mathrm{a}}$ and Deg are immediate verifications. □

Next, we will need the graded commutator with respect to the antisymmetric degree, denoted by

$$\mathrm{ad}(a)(b) = [a, b] = a \circ_\pi b - (-1)^{k\ell} b \circ_\pi a \tag{2-14}$$

for any $a \in \mathcal{W} \otimes \Lambda^k$ and $b \in \mathcal{W} \otimes \Lambda^\ell$ and extended $\mathbb{K}[[t]]$-bilinearly as usual. Since $\circ_\pi$ deforms the graded commutative product μ, all graded commutators $[a, b]$ will vanish in the zeroth order of t. This allows one to define graded derivations $(1/t)\,\mathrm{ad}(a)$ of $\circ_\pi$.

Lemma 2.5. *An element* $a \in \mathcal{W} \otimes \Lambda^\bullet$ *is central*, *that is* $\mathrm{ad}(a) = 0$, *if and only if* $\deg_{\mathrm{s}}(a) = 0$.

By definition, a *covariant derivative* is an arbitrary bilinear map

$$\nabla : \mathfrak{g} \times \mathfrak{g} \ni (X, Y) \mapsto \nabla_X Y \in \mathfrak{g}. \tag{2-15}$$

The idea is that in the geometric interpretation the covariant derivative is uniquely determined by its values on the left invariant vector fields: we want an invariant covariant derivative and hence it should take values again in $\mathfrak{g}$. An arbitrary covariant derivative is called *torsion-free* if

$$\nabla_X Y - \nabla_Y X - [X, Y] = 0 \tag{2-16}$$

for all $X, Y \in \mathfrak{g}$. Having a covariant derivative, we can extend it to the tensor algebra over $\mathfrak{g}$ by requiring the maps

$$\nabla_X : \mathrm{T}^\bullet\mathfrak{g} \to \mathrm{T}^\bullet\mathfrak{g} \tag{2-17}$$

to be derivations for all $X \in \mathfrak{g}$. We also extend ∇_X to elements in the dual by

$$(\nabla_X \alpha)(Y) = -\alpha(\nabla_X Y) \tag{2-18}$$

for all $X, Y \in \mathfrak{g}$ and $\alpha \in \mathfrak{g}^*$. Finally, we can extend ∇_X to $\mathrm{T}^\bullet\mathfrak{g}^*$ as a derivation, too. Acting on symmetric or antisymmetric tensors, ∇_X will preserve the symmetry type and yields a derivation of the $\vee$- and $\wedge$-products, respectively. The fact that we extended ∇ as a derivation in a way which is compatible with natural pairings will lead to relations like

$$[\nabla_X, \mathrm{i}_\mathrm{s}(Y)] = \mathrm{i}_\mathrm{s}(\nabla_X Y) \tag{2-19}$$

for all $X, Y \in \mathfrak{g}$, as one can easily check on generators.

Sometimes it will be advantageous to use the basis of $\mathfrak{g}$ for computations. With respect to the basis we define the *Christoffel symbols*

$$\Gamma^k_{ij} = e^k(\nabla_{e_i} e_j) \tag{2-20}$$

of a covariant derivative, where $i, j, k = 1, \dots, n$. Clearly, ∇ is uniquely determined by its Christoffel symbols. Moreover, ∇ is torsion-free if and only if

$$\Gamma^k_{ij} - \Gamma^k_{ji} = C^k_{ij}, \tag{2-21}$$

with the usual structure constants $C^k_{ij} = e^k([e_i, e_j]) \in \mathsf{R}$ of the Lie algebra $\mathfrak{g}$.

As in symplectic geometry, the Hess trick [1981] shows the existence of a *symplectic* torsion-free covariant derivative:

Proposition 2.6 (Hess trick). *Let $(\mathfrak{g}, r)$ be a Lie algebra with nondegenerate r-matrix r and inverse ω. Then there exists a torsion-free covariant derivative ∇ such that for all $X \in \mathfrak{g}$ we have*

$$\nabla_X \omega = 0 \quad \textit{and} \quad \nabla_X r = 0. \tag{2-22}$$

Proof. The idea is to start with the half-commutator connection as in the geometric case and make it symplectic by means of the Hess trick. The covariant derivative

$$\tilde{\nabla} : \mathfrak{g} \times \mathfrak{g} \ni (X, Y) \mapsto \tfrac{1}{2}[X, Y] \in \mathfrak{g}$$

is clearly torsion-free. Since ω is nondegenerate, we can determine a map ∇_X uniquely by

$$\omega(\nabla_X Y, Z) = \omega(\tilde{\nabla}_X Y, Z) + \tfrac{1}{3}(\tilde{\nabla}_X \omega)(Y, Z) + \tfrac{1}{3}(\tilde{\nabla}_Y \omega)(X, Z). \tag{2-23}$$

It is then an immediate computation using the closedness $\delta_{\mathrm{CE}}\omega = 0$ of ω, that this map satisfies all requirements. $\square$

The curvature $\tilde{R}$ corresponding to ∇ is defined by

$$\tilde{R} : \mathfrak{g}\times\mathfrak{g}\times\mathfrak{g} \ni (X, Y, Z) \mapsto \tilde{R}(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \in \mathfrak{g}. \tag{2-24}$$

For a symplectic covariant derivative, we contract $\tilde{R}$ with the symplectic form ω and get

$$R : \mathfrak{g}\times\mathfrak{g}\times\mathfrak{g}\times\mathfrak{g} \ni (Z, U, X, Y) \mapsto \omega(Z, \tilde{R}(X, Y)U) \in \mathsf{R}, \tag{2-25}$$

which is symmetric in the first two components and antisymmetric in the last ones: this follows at once from ∇ being torsion-free and symplectic. In other words, $R \in \mathrm{S}^2(\mathfrak{g}^*) \otimes \Lambda^2\mathfrak{g}^*$ becomes an element of the formal Weyl algebra satisfying

$$\deg_{\mathrm{s}} R = 2R = \operatorname{Deg} R, \quad \deg_{\mathrm{a}} R = 2R, \quad \text{and} \quad \deg_t R = 0. \tag{2-26}$$

In the following, we will fix a symplectic torsion-free covariant derivative, the existence of which is granted by Proposition 2.6. Since ∇_X acts on all types of tensors already, we can use ∇ to define the following derivation D on the formal Weyl algebra

$$D : \mathcal{W}\otimes\Lambda^\bullet \ni (f\otimes\alpha) \mapsto \nabla_{e_i} f \otimes e^i \wedge \alpha + f \otimes e^i \wedge \nabla_{e_i}\alpha \in \mathcal{W}\otimes\Lambda^{\bullet+1}. \tag{2-27}$$

Notice that we do not use the explicit expression of ∇ given in (2-23). In fact, any other symplectic torsion-free covariant derivative will do the job as well.

For every torsion-free covariant derivative ∇ it is easy to check that

$$e^i \wedge \nabla_{e_i}\alpha = \delta_{\mathrm{CE}}\alpha \tag{2-28}$$

holds for all $\alpha \in \Lambda^\bullet\mathfrak{g}^*$: indeed, both sides define graded derivations of antisymmetric degree $+1$ and coincide on generators in $\mathfrak{g}^* \subseteq \Lambda^\bullet\mathfrak{g}^*$. Therefore, we can rewrite D as

$$D(f\otimes\alpha) = \nabla_{e_i} f \otimes e^i \wedge \alpha + f \otimes \delta_{\mathrm{CE}}\alpha. \tag{2-29}$$

From now on, unless clearly stated, we refer to $[\,\cdot\,,\,\cdot\,]$ as the supercommutator with respect to the antisymmetric degree.

Proposition 2.7. *Let ∇ be a symplectic torsion-free covariant derivative. If in addition s is covariantly constant, i.e., if $\nabla_X s = 0$ for all $X \in \mathfrak{g}$, the map $D : \mathcal{W}\otimes\Lambda^\bullet \to \mathcal{W}\otimes\Lambda^{\bullet+1}$ is a graded derivation of antisymmetric degree $+1$ of the star product $\circ_\pi$, i.e.,*

$$D(a \circ_\pi b) = D(a) \circ_\pi b + (-1)^k a \circ_\pi D(b) \tag{2-30}$$

for $a \in \mathcal{W}\otimes\Lambda^k$ and $b \in \mathcal{W}\otimes\Lambda^\bullet$. In addition, we have

$$\delta R = 0, \quad DR = 0, \quad [\delta, D] = \delta D + D\delta = 0, \quad D^2 = \frac{1}{2}[D, D] = \frac{1}{t}\operatorname{ad}(R). \tag{2-31}$$

Proof. For the operator $\mathcal{P}$ from (2-13) we have

$$
\begin{aligned}
&(\mathsf{id} \otimes \nabla_{e_k} + \nabla_{e_k} \otimes \mathsf{id})\mathcal{P}(a \otimes b) \\
&\qquad = \pi^{ij}\, \mathrm{i}_{\mathrm{s}}(e_i)a \otimes \nabla_{e_k} \mathrm{i}_{\mathrm{s}}(e_j)b + \pi^{ij} \nabla_{e_k} \mathrm{i}_{\mathrm{s}}(e_i)a \otimes \mathrm{i}_{\mathrm{s}}(e_j)b \\
&\qquad \overset{(a)}{=} (\pi^{\ell j}\Gamma^i_{k\ell} + \pi^{i\ell}\Gamma^j_{k\ell})\, \mathrm{i}_{\mathrm{s}}(e_i)a \otimes \mathrm{i}_{\mathrm{s}}(e_j)b + \mathcal{P}(\mathsf{id} \otimes \nabla_{e_k} + \nabla_{e_k} \otimes \mathsf{id})(a \otimes b) \\
&\qquad = \mathcal{P}(\mathsf{id} \otimes \nabla_{e_k} + \nabla_{e_k} \otimes \mathsf{id})(a \otimes b)
\end{aligned}
$$

for $a, b \in \mathcal{W} \otimes \Lambda^\bullet$. Here we used the relation $[\nabla_X, \mathrm{i}_{\mathrm{s}}(Y)] = \mathrm{i}_{\mathrm{s}}(\nabla_X Y)$ as well as the definition of the Christoffel symbols in (a). In the last step we used $\pi^{\ell j}\Gamma^i_{k\ell} + \pi^{i\ell}\Gamma^j_{k\ell} = 0$ which follows from $\nabla(r + s) = 0$. Therefore we have

$$
\nabla_{e_i} \circ \mu \circ \mathrm{e}^{\frac{1}{2}t\mathcal{P}} = \mu \circ (\mathsf{id} \otimes \nabla_{e_i} + \nabla_{e_i} \otimes \mathsf{id}) \circ \mathrm{e}^{\frac{1}{2}t\mathcal{P}} = \mu \circ \mathrm{e}^{\frac{1}{2}t\mathcal{P}} \circ (\mathsf{id} \otimes \nabla_{e_i} + \nabla_{e_i} \otimes \mathsf{id}).
$$

By $\wedge$-multiplying by the corresponding e^i it follows that D is a graded derivation of antisymmetric degree $+1$. Let $f \otimes \alpha \in \mathcal{W} \otimes \Lambda^\bullet$. Just using the definition of δ, (2-29) and the fact that ∇ is torsion-free we get

$$
\begin{aligned}
\delta D(f \otimes \alpha) &= \delta(\nabla_{e_k} f \otimes e^k \wedge \alpha + f \otimes \delta_{\mathrm{CE}}\alpha) \\
&= -D\delta(f \otimes \alpha) + \tfrac{1}{2}(\Gamma^\ell_{ik} - \Gamma^\ell_{ki} - C^\ell_{ik})\, \mathrm{i}_{\mathrm{s}}(e_\ell) f \otimes e^i \wedge e^k \wedge \alpha \\
&= -D\delta(f \otimes \alpha).
\end{aligned}
$$

Using a similar computation in coordinates, we get $D^2 = \frac{1}{2}[D, D] = (1/t)\,\mathrm{ad}(R)$. Finally, from the Jacobi identity of the graded commutator we get $(1/2t)\,\mathrm{ad}(\delta R) = [\delta, [D, D]\!] = 0$. Hence δR is central. Since δR has symmetric degree $+1$, this can only happen if $\delta R = 0$. With the same argument, $0 = [D, [D, D]\!]$ yields that DR is central, which again gives $DR = 0$ by counting degrees. □

Remark 2.8. In principle, we will mainly be interested in the case $s = 0$ in the following. However, if the Lie algebra allows for a covariantly constant s it might be interesting to incorporate this into the universal construction: already in the abelian case this leads to the freedom of choosing a different ordering than the Weyl ordering (total symmetrization). Here in particular the Wick ordering is of significance due to the better positivity properties; see [Bursztyn and Waldmann 2000] for a universal deformation formula in this context.

The core of Fedosov's construction is now to turn $-\delta + D$ into a differential: due to the curvature R the derivation $-\delta + D$ is not a differential directly. Nevertheless, from the above discussion we know that it is an inner derivation. Hence the idea is to compensate the defect of being a differential by inner derivations, leading to the following statement:

Proposition 2.9. *Let $\Omega \in t\Lambda^2\mathfrak{g}^*[\![t]\!]$ be a series of δ_{CE}-closed two-forms. Then there is a unique $\varrho \in \mathcal{W}_2 \otimes \Lambda^1$, such that*

$$\delta\varrho = R + D\varrho + \tfrac{1}{t}\varrho \circ_\pi \varrho + \Omega \tag{2-32}$$

and

$$\delta^{-1}\varrho = 0. \tag{2-33}$$

Moreover, the derivation $\mathscr{D}_{\mathrm{F}} = -\delta + D + (1/t)\,\mathrm{ad}(\varrho)$ satisfies $\mathscr{D}_{\mathrm{F}}^2 = 0$.

Proof. Let us first assume that (2-32) is satisfied and apply δ^{-1} to (2-33). This yields

$$\delta^{-1}\delta\varrho = \delta^{-1}\Big(R + Dx + \frac{1}{t}\varrho \circ_\pi \varrho + \Omega\Big).$$

From the Poincaré Lemma as in Lemma 2.3 we have

$$\varrho = \delta^{-1}\Big(R + D\varrho + \frac{1}{t}\varrho \circ_\pi \varrho + \Omega\Big). \tag{2-34}$$

Let us define the operator $B : \mathcal{W} \otimes \Lambda^1 \to \mathcal{W} \otimes \Lambda^1$ by

$$B(a) = \delta^{-1}\Big(R + Da + \frac{1}{t}a \circ_\pi a + \Omega\Big).$$

Thus the solutions of (2-33) coincide with the fixed points of the operator B. Now we want to show that B has indeed a unique fixed point. By a careful but straightforward counting of degrees we see that B maps $\mathcal{W}_2 \otimes \Lambda^1$ into $\mathcal{W}_2 \otimes \Lambda^1$. Second, we note that B is a contraction with respect to the total degree. Indeed, for $a, a' \in \mathcal{W}_2 \otimes \Lambda^1$ with $a - a' \in \mathcal{W}_k \otimes \Lambda^1$ we have

$$\begin{aligned} B(a) - B(a') &= \delta^{-1}D(a - a') + \frac{1}{t}(a \circ_\pi a - a' \circ_\pi a') \\ &= \delta^{-1}D(a - a') + \frac{1}{t}\delta^{-1}((a - a') \circ_\pi a' + a \circ_\pi (a - a')). \end{aligned}$$

The first term $\delta^{-1}D(a - a')$ is an element of $\mathcal{W}_{k+1} \otimes \Lambda^1$, because D does not change the total degree and δ^{-1} increases it by $+1$. Since Deg is a $\circ_\pi$-derivation and since a, a' have total degree at least 2 and their difference has total degree at least k, the second term has total degree at least $k+1$, as $1/t$ has total degree -2 but δ^{-1} raises the total degree by $+1$. This allows one to apply the Banach fixed-point theorem for the complete filtration by the total degree: we have a unique fixed-point $B(\varrho) = \varrho$ with $\varrho \in \mathcal{W}_2 \otimes \Lambda^1$, i.e., ϱ satisfies (2-34). Finally, we show that this ϱ fulfills (2-33). Define

$$A = \delta\varrho - R - D\varrho - \frac{1}{t}\varrho \circ_\pi \varrho - \Omega.$$

Applying δ to A and using Proposition 2.7, we obtain

$$\begin{aligned}
\delta A &= -\delta D\varrho - \frac{1}{t}(\delta\varrho \circ_\pi \varrho - \varrho \circ_\pi \delta\varrho) \\
&= D\delta\varrho + \frac{1}{t}\operatorname{ad}(\varrho)\delta\varrho \\
&= D\Big(A + R + D\varrho + \frac{1}{t}\varrho \circ_\pi \varrho + \Omega\Big) + \frac{1}{t}\operatorname{ad}(\varrho)\Big(A + R + D\varrho + \frac{1}{t}\varrho \circ_\pi \varrho + \Omega\Big) \\
&\overset{(a)}{=} DA + \frac{1}{t}\operatorname{ad}(\varrho)(A).
\end{aligned}$$

In (a) we used that $(-\delta + D + (1/t)\operatorname{ad}(\varrho))(R + D\varrho + (1/t)\varrho \circ_\pi \varrho + \Omega) = 0$, which can be seen as a version of the second Bianchi identity for $-\delta + D + (1/t)\operatorname{ad}(\varrho)$. This follows by an explicit computation for arbitrary ϱ. On the other hand

$$\delta^{-1}A = \delta^{-1}\Big(\delta\varrho - R - D\varrho - \frac{1}{t}\varrho \circ_\pi \varrho - \Omega\Big) = \delta^{-1}\delta\varrho - \varrho = \delta\delta^{-1}\varrho = 0$$

for ϱ being the fixed-point of the operator B. In other words,

$$A = \delta^{-1}\delta A = \delta^{-1}\Big(DA + \frac{1}{t}\operatorname{ad}(\varrho)(A)\Big)$$

is a fixed-point of the operator $K : \mathcal{W} \otimes \Lambda^\bullet \to \mathcal{W} \otimes \Lambda^\bullet$ defined by

$$Ka = \delta^{-1}\Big(Da + \frac{1}{t}\operatorname{ad}(\varrho)(a)\Big).$$

Using an analogous argument to the one above, this operator is a contraction with respect to the total degree, and has a unique fixed-point. Finally, since K is linear the fixed-point has to be zero, which means that $A = 0$. $\square$

Remark 2.10. It is important to note that the above construction of the element ϱ, which will be the crucial ingredient in the universal deformation formula below, is a fairly explicit recursion formula. Writing $\varrho = \sum_{r=3}^{\infty} \varrho^{(r)}$ with components $\varrho^{(r)}$ of homogeneous total degree $\operatorname{Deg}\varrho^{(r)} = r\varrho^{(r)}$ we see that $\varrho^{(3)} = \delta^{-1}(R + t\Omega_1)$ and

$$\varrho^{(r+3)} = \delta^{-1}\Bigg(D\varrho^{(r+2)} + \frac{1}{t}\sum_{\ell=1}^{r-1} \varrho^{(\ell+2)} \circ_\pi \varrho^{(r+2-\ell)} + \Omega^{(r+2)}\Bigg), \tag{2-35}$$

where $\Omega^{(2k)} = t^k\Omega_k$ for $k \in \mathbb{N}$ and $\Omega^{(2k+1)} = 0$. Moreover, if we find a *flat* ∇, i.e., if $R = 0$, then for trivial $\Omega = 0$ we have $\varrho = 0$ as solution.

3. Universal deformation formula

Let us consider a triangular Lie algebra $(\mathfrak{g}, r)$ acting on a generic associative algebra $(\mathscr{A}, \mu_{\mathscr{A}})$ via derivations. We denote by $\triangleright$ the corresponding Hopf algebra action

$\mathscr{U}(\mathfrak{g}) \to \mathsf{End}(\mathscr{A})$. In the following we refer to

$$\mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet = \prod_{k=0}^{\infty} (\mathscr{A} \otimes \mathrm{S}^k\mathfrak{g}^* \otimes \Lambda^\bullet\mathfrak{g}^*)[\![t]\!]$$

as the *enlarged Fedosov algebra*. The operators defined in the previous section are extended to $\mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet$ by acting trivially on the $\mathscr{A}$-factor and as before on the $\mathcal{W} \otimes \Lambda^\bullet$-factor.

The deformed product $\circ_\pi$ on $\mathcal{W} \otimes \Lambda^\bullet$ together with the product $\mu_{\mathscr{A}}$ of $\mathscr{A}$ yields a new (deformed) $\mathsf{R}[\![t]\!]$-bilinear product $m_\pi^{\mathscr{A}}$ for the extended Fedosov algebra. Explicitly, on factorizing tensors we have

$$m_\pi^{\mathscr{A}}(\xi_1 \otimes f_1 \otimes \alpha_1, \xi_2 \otimes f_2 \otimes \alpha_2) = (\xi_1 \cdot \xi_2) \otimes (f_1 \otimes \alpha_1) \circ_\pi (f_2 \otimes \alpha_2), \tag{3-1}$$

where $\xi_1, \xi_2 \in \mathscr{A}$, $f_1, f_2 \in \mathrm{S}^\bullet\mathfrak{g}^*$ and $\alpha_1, \alpha_2 \in \Lambda^\bullet\mathfrak{g}^*$. We simply write $\xi_1 \cdot \xi_2$ for the (undeformed) product $\mu_{\mathscr{A}}$ of $\mathscr{A}$. Clearly, this new product $m_\pi^{\mathscr{A}}$ is again associative.

As new ingredient we use the action $\triangleright$ to define the operator $L_{\mathscr{A}} : \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet \to \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet$ by

$$L_{\mathscr{A}}(\xi \otimes f \otimes \alpha) = e_i \triangleright \xi \otimes f \otimes e^i \wedge \alpha \tag{3-2}$$

on factorizing elements and extend it $\mathsf{R}[\![t]\!]$-linearly as usual. Since the action of Lie algebra elements is by derivations, we see that $L_{\mathscr{A}}$ is a derivation of $\mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet$ of antisymmetric degree $+1$. The sum

$$\mathscr{D}_{\mathscr{A}} = L_{\mathscr{A}} + \mathscr{D}_{\mathrm{F}} \tag{3-3}$$

is thus still a derivation of antisymmetric degree $+1$ which we call the *extended Fedosov derivation*. It turns out to be a differential, too:

Lemma 3.1. *The map $\mathscr{D}_{\mathscr{A}} = L_{\mathscr{A}} + \mathscr{D}_{\mathrm{F}}$ squares to zero.*

Proof. First, we observe that $\mathscr{D}_{\mathscr{A}}^2 = L_{\mathscr{A}}^2 + [\mathscr{D}_{\mathrm{F}}, L_{\mathscr{A}}]$, because $\mathscr{D}_{\mathrm{F}}^2 = 0$. Next, since $\triangleright$ is a Lie algebra action, we immediately obtain

$$L_{\mathscr{A}}^2(\xi \otimes f \otimes \alpha) = \tfrac{1}{2} C_{ij}^k e_k \triangleright \xi \otimes f \otimes e^i \wedge e^j \wedge \alpha$$

on factorizing elements. We clearly have $[\delta, L_{\mathscr{A}}] = 0 = [\mathrm{ad}(\varrho), L_{\mathscr{A}}]$ since the maps act on different tensor factors. It remains to compute the only nontrivial term in $[\mathscr{D}_{\mathrm{F}}, L_{\mathscr{A}}] = [D, L_{\mathscr{A}}]$. Using $\delta_{\mathrm{CE}} e^k = -\frac{1}{2} C_{ij}^k e^i \wedge e^j$, this results immediately in $[D, L_{\mathscr{A}}] = -L_{\mathscr{A}}^2$. □

The cohomology of this differential turns out to be almost trivial: we only have a nontrivial contribution in antisymmetric degree 0, the kernel of $\mathscr{D}_{\mathscr{A}}$. In higher antisymmetric degrees, the following homotopy formula shows that the cohomology is trivial:

Proposition 3.2. *The operator*

$$\mathscr{D}_{\mathscr{A}}^{-1} = \delta^{-1} \frac{1}{\mathsf{id} - \left[\delta^{-1},\, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\right]} \tag{3-4}$$

is a well-defined $\mathsf{R}[[t]]$*-linear endomorphism of* $\mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet$ *and we have*

$$a = \mathscr{D}_{\mathscr{A}}\mathscr{D}_{\mathscr{A}}^{-1} a + \mathscr{D}_{\mathscr{A}}^{-1}\mathscr{D}_{\mathscr{A}} a + \frac{1}{\mathsf{id} - \left[\delta^{-1},\, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\right]}\sigma(a). \tag{3-5}$$

for all $a \in \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet$.

Proof. Let us denote by A the operator $[\delta^{-1}, D + L_{\mathscr{A}} + (1/t)\,\mathrm{ad}(\varrho)]$. Since it increases the total degree by $+1$, the geometric series $(\mathsf{id} - A)^{-1}$ is well defined as a formal series in the total degree. We start with the Poincaré equation (2-12) and get

$$-\mathscr{D}_{\mathscr{A}}\delta^{-1} a - \delta^{-1}\mathscr{D}_{\mathscr{A}} a + \sigma(a) = (\mathsf{id} - A)a, \tag{3-6}$$

since $\mathcal{D}_{\mathcal{A}}$ deforms the differential $-\delta$ by higher order terms in the total degree. The usual homological perturbation argument then gives (3-4) by a standard computation; see, e.g., [Waldmann 2007, Proposition 6.4.17] for this computation. ☐

Corollary 3.3. *Let* $a \in \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^0$. *Then* $\mathscr{D}_{\mathscr{A}} a = 0$ *if and only if*

$$a = \frac{1}{\mathsf{id} - \left[\delta^{-1},\, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\right]}\sigma(a). \tag{3-7}$$

Since the element $a \in \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^0$ is completely determined in the symmetric and antisymmetric degree 0, we can use it to define the extended Fedosov Taylor series.

Definition 3.4 (Extended Fedosov Taylor series). Given the extended Fedosov derivation $\mathscr{D}_{\mathscr{A}} = -\delta + D + L_{\mathscr{A}} + (1/t)\,\mathrm{ad}(\varrho)$, the extended Fedosov Taylor series of $\xi \in \mathscr{A}[[t]]$ is defined by

$$\tau_{\mathscr{A}}(\xi) = \frac{1}{\mathsf{id} - \left[\delta^{-1},\, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\right]}\xi. \tag{3-8}$$

Lemma 3.5. *For* $\xi \in \mathscr{A}[[t]]$ *we have*

$$\sigma(\tau_{\mathscr{A}}(\xi)) = \xi. \tag{3-9}$$

Moreover, the map $\tau_{\mathscr{A}} : \mathscr{A}[[t]] \to \ker \mathscr{D}_{\mathscr{A}} \cap \ker \deg_{\mathrm{a}}$ *is a* $\mathsf{R}[[t]]$*-linear isomorphism starting with*

$$\tau_{\mathscr{A}}(\xi) = \sum_{k=0}^{\infty}\left[\delta^{-1},\, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\right]^k(\xi) = \xi \otimes 1 \otimes 1 + e_i \triangleright \xi \otimes e^i \otimes 1 + \cdots \tag{3-10}$$

in zeroth and first order of the total degree.

Proof. The isomorphism property follows directly from Corollary 3.3. The commutator $[\delta^{-1}, D + L_{\mathscr{A}} + (1/t)\,\mathrm{ad}(\varrho)]$ raises the total degree at least by one, thus the zeroth and first order terms in the total degree come from the terms with $k = 0$ and $k = 1$ in the geometric series in (3-10). Here it is easy to see that the only nontrivial contribution is

$$\Big[\delta^{-1}, D + L_{\mathscr{A}} + \frac{1}{t}\,\mathrm{ad}(\varrho)\Big]\xi = L_{\mathscr{A}}\xi,$$

proving the claim in (3-10). Note that already for $k = 2$ we also get contributions of S and $\mathrm{ad}(\varrho)$. □

Given the $\mathsf{R}[\![t]\!]$-linear isomorphism $\tau_{\mathscr{A}} : \mathscr{A}[\![t]\!] \to \ker \mathscr{D}_{\mathscr{A}} \cap \ker \deg_{\mathrm{a}}$ we can turn $\mathscr{A}[\![t]\!]$ into an algebra by pulling back the deformed product: note that the kernel of a derivation is always a subalgebra and hence the intersection $\ker \mathscr{D}_{\mathscr{A}} \cap \ker \deg_{\mathrm{a}}$ is also a subalgebra. This allows us to obtain a universal deformation formula for any $\mathscr{U}(\mathfrak{g})$-module algebra $\mathscr{A}$:

Theorem 3.6 (Universal deformation formula). *Let $\mathfrak{g}$ be a Lie algebra with nondegenerate r-matrix. Moreover, let $s \in \mathrm{S}^2\mathfrak{g}$ be such that there exists a symplectic torsion-free covariant derivative ∇ with s being covariantly constant. Consider then $\pi = r + s$. Finally, let $\Omega \in t\Lambda^2\mathfrak{g}^*[\![t]\!]$ be a formal series of δ_{CE}-closed two-forms. Then for every associative algebra $\mathscr{A}$ with action of $\mathfrak{g}$ by derivations one obtains an associative deformation $m_\star^{\mathscr{A}} : \mathscr{A}[\![t]\!] \times \mathscr{A}[\![t]\!] \to \mathscr{A}[\![t]\!]$ by*

$$m_\star^{\mathscr{A}}(\xi, \eta) = \sigma(m_\pi^{\mathscr{A}}(\tau_{\mathscr{A}}(\xi), \tau_{\mathscr{A}}(\eta))). \tag{3-11}$$

Writing simply $\star = \star_{\Omega,\nabla,s}$ for this new product, one has

$$\xi \star \eta = \xi \cdot \eta + \frac{t}{2}\pi^{ij}(e_i \triangleright \xi) \cdot (e_j \triangleright \eta) + \mathcal{O}(t^2) \quad \textit{for } \xi, \eta \in \mathscr{A}. \tag{3-12}$$

Proof. The product $m_\star^{\mathscr{A}}$ is associative, because $m_\pi^{\mathscr{A}}$ is associative and $\tau_{\mathscr{A}}$ is an isomorphism onto a subalgebra with inverse σ. The second part is a direct consequence of Lemma 3.5. □

Remark 3.7. The above theorem can be further generalized by observing that given a Poisson structure on $\mathscr{A}$ induced by a generic bivector on $\mathfrak{g}$, we can reduce to the quotient $\mathfrak{g}/\ker\triangleright$ and obtain an r-matrix on the quotient, inducing the same Poisson structure.

4. Universal construction for Drinfeld twists

Let us consider the particular case in which $\mathscr{A}$ is the tensor algebra $(\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g})), \otimes)$. In this case, we denote by L the operator $L_{\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))} : \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet \to$

$\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$, which is given by

$$L_{\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))}(\xi \otimes f \otimes \alpha) = L_{e_i}\xi \otimes f \otimes e^i \wedge \alpha. \tag{4-1}$$

Here L_{e_i} is the left multiplication in $\mathscr{U}(\mathfrak{g})$ of the element e_i extended as a derivation of the tensor product. Note that it is independent of the choice of the basis in $\mathfrak{g}$.

Applying the results discussed in the last section, we obtain a star product for the tensor algebra over $\mathscr{U}(\mathfrak{g})$ as a particular case of Theorem 3.6:

Corollary 4.1. *The map* $m_\star : \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]] \times \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]] \to \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]]$ *given by*

$$m_\star(\xi, \eta) = \xi \star \eta = \sigma(m_\pi(\tau(\xi), \tau(\eta))) \tag{4-2}$$

is an associative product and

$$\xi \star \eta = \xi \otimes \eta + \tfrac{1}{2} t \pi^{ij} L_{e_i}\xi \otimes L_{e_j}\eta + \mathcal{O}(t^2) \quad \textit{for } \xi, \eta \in \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g})). \tag{4-3}$$

In the following we prove that the star product $m_\star$ defined above allows one to construct a formal Drinfeld twist. Let us define, for any linear map

$$\Phi : \mathscr{U}(\mathfrak{g})^{\otimes k} \to \mathscr{U}(\mathfrak{g})^{\otimes \ell}, \tag{4-4}$$

the *lifted* map

$$\Phi^{\mathrm{Lift}} : \mathscr{U}(\mathfrak{g})^{\otimes k} \otimes \mathcal{W} \otimes \Lambda^\bullet \ni \xi \otimes f \otimes \alpha \mapsto \Phi(\xi) \otimes f \otimes \alpha \in \mathscr{U}(\mathfrak{g})^{\otimes \ell} \otimes \mathcal{W} \otimes \Lambda^\bullet, \tag{4-5}$$

obeying the following simple properties:

Lemma 4.2. *Let* $\Phi : \mathscr{U}(\mathfrak{g})^{\otimes k} \to \mathscr{U}(\mathfrak{g})^{\otimes \ell}$ *and* $\Psi : \mathscr{U}(\mathfrak{g})^{\otimes m} \to \mathscr{U}(\mathfrak{g})^{\otimes n}$ *be linear maps.*

(i) *The lifted map* Φ^{Lift} *commutes with* δ, δ^{-1}, D, *and* $\mathrm{ad}(x)$ *for all* $x \in \mathcal{W} \otimes \Lambda^\bullet$.

(ii) *We have*

$$\Phi \circ \sigma|_{\mathscr{U}(\mathfrak{g})^{\otimes k} \otimes \mathcal{W} \otimes \Lambda^\bullet} = \sigma|_{\mathscr{U}(\mathfrak{g})^{\otimes \ell} \otimes \mathcal{W} \otimes \Lambda^\bullet} \circ \Phi^{\mathrm{Lift}}. \tag{4-6}$$

(iii) *We have*

$$(\Phi \otimes \Psi)^{\mathrm{Lift}} m_\pi(a_1, a_2) = m_\pi(\Phi^{\mathrm{Lift}}(a_1), \Psi^{\mathrm{Lift}}(a_2)), \tag{4-7}$$

for any $a_1 \in \mathscr{U}(\mathfrak{g})^{\otimes k} \otimes \mathcal{W} \otimes \Lambda^\bullet$ *and* $a_2 \in \mathscr{U}(\mathfrak{g})^{\otimes m} \otimes \mathcal{W} \otimes \Lambda^\bullet$.

Let $\eta \in \mathscr{U}(\mathfrak{g})^{\otimes k}[[t]]$ be given. Then we can consider the right multiplication by η using the algebra structure of $\mathscr{U}(\mathfrak{g})^{\otimes k}[[t]]$ coming from the universal enveloping algebra as a map

$$\cdot\eta : \mathscr{U}(\mathfrak{g})^{\otimes k} \ni \xi \mapsto \xi \cdot \eta \in \mathscr{U}(\mathfrak{g})^{\otimes k}. \tag{4-8}$$

To this map we can apply the above lifting process and extend it this way to a $\mathsf{R}[\![t]\!]$-linear map such that on factorizing elements

$$\cdot\eta : \mathscr{U}(\mathfrak{g})^{\otimes k} \otimes \mathcal{W} \otimes \Lambda^\bullet \ni \xi \otimes f \otimes \alpha \mapsto (\xi \cdot \eta) \otimes f \otimes \alpha \in \mathscr{U}(\mathfrak{g})^{\otimes k}, \tag{4-9}$$

where we simply write $\cdot\eta$ instead of $(\cdot\eta)^{\mathrm{Lift}}$. Note that $a \cdot \eta$ is only defined if the tensor degrees k of $\eta \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))$ and a coincide since we use the algebra structure inherited from the universal enveloping algebra.

In the following we denote by $\mathscr{D}$ the derivation $\mathscr{D}_{\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))}$ as obtained in (3-3). We collect some properties how the lifted right multiplications match with the extended Fedosov derivation:

Lemma 4.3. (i) *For any* $a \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ *and* $\xi \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[\![t]\!]$, *we have* $\mathscr{D}(a \cdot \xi) = \mathscr{D}(a) \cdot \xi$

(ii) *The extended Fedosov-Taylor series* τ *preserves the tensor degree of elements in* $\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$.

(iii) *For any* $\xi, \eta \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[\![t]\!]$, *we have* $\tau(\xi \cdot \eta) = \tau(\xi) \cdot \eta$.

(iv) *For any* $a_1 \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ *and* $a_2 \in \mathrm{T}^\ell(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ *as well as* $\eta_1 \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[\![t]\!]$ *and* $\eta_2 \in \mathrm{T}^\ell(\mathscr{U}(\mathfrak{g}))[\![t]\!]$, *we have* $m_\pi(a_1 \cdot \eta_1, a_2 \cdot_l \eta_2) = m_\pi(a_1, a_2) \cdot (\eta_1 \otimes \eta_2)$.

Proof. Let $\xi \otimes a \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ and $\eta \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))$. Then we have

$$\begin{aligned} \mathscr{D}((\xi \otimes a) \cdot \eta) = \mathscr{D}((\xi \cdot \eta) \otimes a) & \\ &= L_{e_i}(\xi \cdot \eta) \otimes e^i \wedge a + (\xi \cdot \eta) \otimes \mathscr{D}_{\mathrm{F}}(a) \\ &= (L_{e_i}(\xi) \otimes e^i \wedge a) \cdot \eta + (\xi \otimes \mathscr{D}_{\mathrm{F}}(a)) \cdot \eta = \mathscr{D}(a) \cdot \eta. \end{aligned}$$

This proves the first claim. The second claim follows immediately from the fact that all operators defining τ do not change the tensor degree. In order to prove the claim (iii), let us consider $\xi, \eta \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[\![t]\!]$. Then we have

$$\mathscr{D}(\tau(\xi) \cdot \eta) = \mathscr{D}(\tau(\xi)) \cdot \eta = 0,$$

according to (i). Thus, $\tau(\xi) \cdot \eta \in \ker \mathscr{D} \cap \ker \deg_{\mathrm{a}}$ and therefore

$$\tau(\xi) \cdot \eta = \tau(\sigma(\tau(\xi) \cdot \eta)) = \tau(\sigma(\tau(\xi)) \cdot \eta) = \tau(\xi \cdot \eta).$$

Finally, to prove the last claim we choose $\xi_1 \otimes f_1 \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ and $\xi_2 \otimes f_2 \in \mathrm{T}^\ell(\mathscr{U}(\mathfrak{g})) \otimes \mathcal{W} \otimes \Lambda^\bullet$ as well as $\eta_1 \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[\![t]\!]$ and $\eta_2 \in \mathrm{T}^\ell(\mathscr{U}(\mathfrak{g}))[\![t]\!]$.

We obtain

$$\begin{aligned} m_\pi((\xi_1 \otimes f_1) \cdot \eta_1, (\xi_2 \otimes f_2) \cdot \eta_2) &= m_\pi((\xi_1 \cdot \eta_1) \otimes f_1, (\xi_2 \cdot \eta_2) \otimes f_2) \\ &= ((\xi_1 \cdot \eta_1) \otimes (\xi_2 \cdot \eta_2)) \otimes (f_1 \circ_\pi f_2) \\ &= ((\xi_1 \otimes \xi_2) \cdot (\eta_1 \otimes \eta_2)) \otimes (f_1 \circ_\pi f_2) \\ &= ((\xi_1 \otimes \xi_2) \otimes (f_1 \circ_\pi f_2)) \cdot (\eta_1 \otimes \eta_2). \end{aligned}$$

This concludes the proof. □

From the above lemma, we observe that the isomorphism τ can be computed for any element $\xi \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[[t]]$ via

$$\tau(\xi) = \tau(1^{\otimes k} \cdot \xi) = \tau(1^{\otimes k}) \cdot \xi, \tag{4-10}$$

where $1 \in \mathscr{U}(\mathfrak{g})$ is the unit element of the universal enveloping algebra. Moreover, from Lemma 4.2, we have

$$\xi \star \eta = \sigma(m_\pi(\tau(\xi) \otimes \tau(\eta))) = (1^{\otimes k} \star 1^{\otimes \ell}) \cdot (\xi \otimes \eta) \tag{4-11}$$

for $\xi \in \mathrm{T}^k(\mathscr{U}(\mathfrak{g}))[[t]]$ and $\eta \in \mathrm{T}^\ell(\mathscr{U}(\mathfrak{g}))[[t]]$. Thus $\star$ is entirely determined by the values on tensor powers of the unit element of the universal enveloping algebra. Note that the unit of $\star$ is the unit element in $\mathsf{R} \subseteq \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$ of the *tensor algebra* but not $1 \in \mathscr{U}(\mathfrak{g})$.

Lemma 4.4. *Let* $\Delta : \mathscr{U}(\mathfrak{g})[[t]] \to \mathscr{U}(\mathfrak{g})^{\otimes 2}[[t]]$ *be the coproduct of* $\mathscr{U}(\mathfrak{g})[[t]]$ *and* $\epsilon : \mathscr{U}(\mathfrak{g}) \to \mathsf{R}[[t]]$ *the counit.*

(i) *We have*

$$L|_{\mathscr{U}(\mathfrak{g})^{\otimes 2} \otimes \mathcal{W} \otimes \Lambda^\bullet} \circ \Delta^{\mathrm{Lift}} = \Delta^{\mathrm{Lift}} \circ L|_{\mathscr{U}(\mathfrak{g}) \otimes \mathcal{W} \otimes \Lambda^\bullet}. \tag{4-12}$$

(ii) *For the Fedosov-Taylor series one has*

$$\Delta^{\mathrm{Lift}} \circ \tau = \tau \circ \Delta. \tag{4-13}$$

(iii) *We have*

$$\epsilon^{\mathrm{Lift}} \circ L|_{\mathscr{U}(\mathfrak{g}) \otimes \mathcal{W} \otimes \Lambda^\bullet} = 0. \tag{4-14}$$

(iv) *For the Fedosov-Taylor series one has*

$$\epsilon^{\mathrm{Lift}} \circ \tau = \epsilon. \tag{4-15}$$

Proof. Let $\xi \otimes f \otimes \alpha \in \mathscr{U}(\mathfrak{g}) \otimes \mathcal{W} \otimes \Lambda^\bullet$. Then we get

$$\begin{aligned}
\Delta^{\mathrm{Lift}} L(\xi \otimes f \otimes \alpha) &= \Delta^{\mathrm{Lift}}(L_{e_i}(\xi) \otimes f \otimes e^i \wedge \alpha)\\
&= \Delta^{\mathrm{Lift}}(e_i\xi \otimes f \otimes e^i \wedge \alpha)\\
&= \Delta(e_i\xi) \otimes f \otimes e^i \wedge \alpha\\
&= \Delta(e_i) \cdot \Delta(\xi) \otimes f \otimes e^i \wedge \alpha\\
&= (e_i \otimes 1 + 1 \otimes e_i) \cdot \Delta(\xi) \otimes f \otimes e^i \wedge \alpha\\
&= L_{e_i}(\Delta(\xi)) \otimes f \otimes e^i \wedge \alpha\\
&= L\Delta^{\mathrm{Lift}}(\xi \otimes f \otimes \alpha),
\end{aligned}$$

since we extended the left multiplication by e_i as a *derivation* of the tensor product to higher tensor powers. Hence all the operators appearing in τ *commute* with Δ^{Lift} and therefore we get the second part. Similarly, we get

$$\begin{aligned}
\epsilon^{\mathrm{Lift}}(L(\xi \otimes f \otimes \alpha) &= \epsilon^{\mathrm{Lift}}(e_i\xi \otimes f \otimes e^i \wedge \alpha)\\
&= \epsilon(e_i\xi) \otimes f \otimes e^i \wedge \alpha = \epsilon(e_i)\epsilon(\xi) \otimes f \otimes e^i \wedge \alpha = 0,
\end{aligned}$$

where we used that ϵ vanishes on primitive elements of $\mathscr{U}(\mathfrak{g})$. Since ϵ^{Lift} commutes with all other operators δ^{-1}, D and $\mathrm{ad}(\varrho)$ according to Lemma 4.2, we first get

$$\epsilon^{\mathrm{Lift}} \circ \Big[\delta^{-1}, D + L + \frac{1}{t}\,\mathrm{ad}(\varrho)\Big] = \Big[\delta^{-1}, D + \frac{1}{t}\,\mathrm{ad}(\varrho)\Big] \circ \epsilon^{\mathrm{Lift}}.$$

Hence for $\xi \in \mathscr{U}(\mathfrak{g})[[t]]$ we have

$$\begin{aligned}
\epsilon^{\mathrm{Lift}}\tau(\xi) &= \epsilon^{\mathrm{Lift}}\bigg(\sum_{k=0}^{\infty}\Big[\delta^{-1}, D + L + \frac{1}{t}\,\mathrm{ad}(\varrho)\Big]^k \xi\bigg)\\
&= \sum_{k=0}^{\infty}\Big[\delta^{-1}, D + \frac{1}{t}\,\mathrm{ad}(\varrho)\Big]^k \epsilon^{\mathrm{Lift}}(\xi)\\
&= \epsilon(\xi),
\end{aligned}$$

since $\epsilon^{\mathrm{Lift}}(\xi) = \epsilon(\xi)$ is just a constant and hence unaffected by all the operators in the series. Thus only the zeroth term remains. □

This is now the last ingredient to show that the element $1 \star 1$ is the twist we are looking for:

Theorem 4.5. *The element* $1 \star 1 \in \mathscr{U}(\mathfrak{g})^{\otimes 2}[[t]]$ *is a twist such that*

$$1 \star 1 = 1 \otimes 1 + \frac{t}{2}\pi + \mathcal{O}(t^2). \tag{4-16}$$

Proof. First we see that

$$\begin{aligned}(\Delta \otimes \mathsf{id})(1 \star 1) &= (\Delta \otimes \mathsf{id})\sigma(m_\pi(\tau(1), \tau(1))) \\ &= \sigma((\Delta \otimes \mathsf{id})^{\mathrm{Lift}}(m_\pi(\tau(1), \tau(1)))) \\ &= \sigma(m_\pi(\Delta^{\mathrm{Lift}}\tau(1), \tau(1))) \\ &= \sigma(m_\pi(\tau(\Delta(1)), \tau(1))) \\ &= \sigma(m_\pi(\tau(1 \otimes 1), \tau(1))) \\ &= (1 \otimes 1) \star 1.\end{aligned}$$

Similarly, we get $(\mathsf{id} \otimes \Delta)(1 \star 1) = 1 \star (1 \otimes 1)$. Thus, using the associativity of $\star$ we obtain the first condition (1-2) for a twist as follows,

$$\begin{aligned}(\Delta \otimes \mathsf{id})(1 \star 1) \cdot ((1 \star 1) \otimes 1) &= ((1 \otimes 1) \star 1) \cdot ((1 \star 1) \otimes 1) \\ &= (1 \star 1) \star 1 \\ &= 1 \star (1 \star 1) \\ &= (\mathsf{id} \otimes \Delta)(1 \star 1) \cdot (1 \otimes (1 \star 1)).\end{aligned}$$

To check the normalization condition (1-3) we use Lemma 4.2 and Lemma 4.4 again to get

$$\begin{aligned}(\epsilon \otimes \mathsf{id})(1 \star 1) &= (\epsilon \otimes \mathsf{id})\sigma(m_\pi(\tau(1), \tau(1))) \\ &= \sigma((\epsilon \otimes \mathsf{id})^{\mathrm{Lift}}(m_\pi(\tau(1), \tau(1)))) \\ &= \sigma((m_\pi(\epsilon^{\mathrm{Lift}}\tau(1), \tau(1)))) \\ &= \sigma((m_\pi(\epsilon(1), \tau(1)))) \\ &= \epsilon(1)\sigma(\tau(1)) \\ &= 1,\end{aligned}$$

since $\epsilon(1)$ is the unit element of R and thus the unit element of $\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$, which serves as unit element for m_π as well. Similarly we obtain $(\mathsf{id} \otimes \epsilon)(1 \star 1) = 1$. Finally, the facts that the first term in t of $1 \star 1$ is given by π and that zero term in t is $1 \otimes 1$ follow from Corollary 4.1. □

Remark 4.6. From now on we refer to $1 \star 1$ as the *Fedosov twist*

$$\mathcal{F}_{\Omega,\nabla,s} = 1 \star 1, \tag{4-17}$$

corresponding to the choice of the δ_{CE}-closed form Ω, the choice of the torsion-free symplectic covariant derivative and the choice of the covariantly constant s. In the following we will be mainly interested in the dependence of $\mathcal{F}_{\Omega,\nabla,s}$ on the two-forms Ω and hence we shall write $\mathcal{F}_\Omega$ for simplicity. We also note that for $s = 0$ and $\Omega = 0$ we have a *preferred* choice for ∇, namely the one obtained from the Hess trick out of the half-commutator covariant derivative as described in Proposition 2.6. This gives a *canonical twist* $\mathcal{F}_0$ quantizing r.

The results discussed above allow us to give an alternative proof of the Drinfeld theorem [1983], stating the existence of twists for every r-matrix:

Corollary 4.7 (Drinfeld). *Let* $(\mathfrak{g}, r)$ *be a Lie algebra with* r*-matrix over a field* $\mathbb{K}$ *with characteristic* 0. *Then there exists a formal twist* $\mathcal{F} \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ *such that*

$$\mathcal{F} = 1 \otimes 1 + \frac{t}{2} r + \mathcal{O}(t^2).$$

To conclude this section we consider the question whether the two approaches of universal deformation formulas actually coincide: on the one hand we know that every twist gives a universal deformation formula by (1-1). On the other hand, we have constructed directly a universal deformation formula (3-11) in Theorem 3.6 based on the Fedosov construction. Since we also get a twist from the Fedosov construction, we are interested in the consistence of the two constructions. In order to answer this question, we need some preparation. Hence let $\mathscr{A}$ be an algebra with action of $\mathfrak{g}$ by derivations as before. Then we define the map

$$\bullet : \mathscr{U}(\mathfrak{g}) \otimes \mathcal{W} \otimes \Lambda^\bullet \times \mathscr{A} \ni (\xi \otimes \alpha, a) \mapsto (\xi \otimes \alpha) \bullet a = \xi \triangleright a \otimes \alpha \in \mathscr{A} \otimes \mathcal{W} \otimes \Lambda^\bullet \tag{4-18}$$

for any $a \in \mathscr{A}$ and $\alpha \in \mathcal{W} \otimes \Lambda^\bullet$. Then the following algebraic properties are obtained by a straightforward computation:

Lemma 4.8. *For any* $\xi \in \mathscr{U}(\mathfrak{g})$, $\alpha \in \mathcal{W} \otimes \Lambda^\bullet$ *and* $a \in \mathscr{A}$ *we have*

(i) $\sigma((\xi \otimes \alpha) \bullet a) = \sigma(\xi \otimes \alpha) \triangleright a$,

(ii) $L_{\mathscr{A}}(\xi \triangleright a \otimes \alpha) = L(\xi \otimes \alpha) \bullet a$,

(iii) $\tau_{\mathscr{A}}(a) = \tau(1) \bullet a$,

(iv) $m_\pi^{\mathscr{A}}(\xi_1 \otimes a_1 \otimes \alpha_1, \xi_2 \otimes a_2 \otimes \alpha_2) = (\mu_{\mathscr{A}} \otimes \mathsf{id} \otimes \mathsf{id})(m_\pi(\xi_1 \otimes \alpha_1, \xi_2 \otimes \alpha_2) \bullet (a_1 \otimes a_2))$.

For matching parameters Ω, ∇, and s of the Fedosov construction, the two approaches coincide:

Proposition 4.9. *For fixed choices of* Ω, ∇, *and* s *and for any* $a, b \in \mathscr{A}$ *we have*

$$a \star_{\Omega,\nabla,s} b = a \star_{\mathcal{F}_{\Omega,\nabla,s}} b. \tag{4-19}$$

Proof. This is now just a matter of computation. We have

$$\begin{aligned} a \star b &= \sigma(m_\pi^{\mathscr{A}}(\tau_{\mathscr{A}}(a) \otimes \tau_{\mathscr{A}}(b))) \\ &\overset{(a)}{=} \sigma(m_\pi((\tau(1) \otimes \tau(1)) \bullet (a \otimes b))) \\ &\overset{(b)}{=} \mu_{\mathscr{A}}(\sigma(m_\pi(\tau(1) \otimes \tau(1))) \triangleright (a \otimes b)) \\ &= \mu_{\mathscr{A}}((1 \star 1) \triangleright (a \otimes b)) \\ &= a \star_{\mathcal{F}} b, \end{aligned}$$

where in (a) we use the third claim of the above lemma and in (b) the first and the fourth. □

5. Classification of Drinfeld twists

In this section we discuss the classification of twists on universal enveloping algebras for a given Lie algebra $\mathfrak{g}$ with nondegenerate r-matrix. Recall that two twists $\mathcal{F}$ and $\mathcal{F}'$ are said to be *equivalent* and denoted by $\mathcal{F} \sim \mathcal{F}'$ if there exists an element $S \in \mathscr{U}(\mathfrak{g})[[t]]$, with $S = 1 + \mathcal{O}(t)$ and $\epsilon(S) = 1$ such that

$$\Delta(S)\mathcal{F}' = \mathcal{F}(S \otimes S). \tag{5-1}$$

In the following we prove that the set of equivalence classes of twists $\mathrm{Twist}(\mathscr{U}(\mathfrak{g}), r)$ with fixed r-matrix r is in bijection to the formal series in the second Chevalley–Eilenberg cohomology $\mathrm{H}^2_{\mathrm{CE}}(\mathfrak{g})[[t]]$.

We will fix the choice of ∇ and the symmetric part s in the Fedosov construction. Then the cohomological equivalence of the two-forms in the construction yields equivalent twists. In fact, an equivalence can even be computed recursively:

Lemma 5.1. *Let ϱ and ϱ' be the two elements in $\mathcal{W}_2 \otimes \Lambda^1$ uniquely determined from Proposition 2.9, corresponding to two closed two-forms $\Omega, \Omega' \in t\Lambda^2\mathfrak{g}^*[[t]]$, respectively, and let $\Omega - \Omega' = \delta_{\mathrm{CE}} C$ for a fixed $C \in t\mathfrak{g}^*[[t]]$. Then there is a unique solution $h \in \mathcal{W}_3 \otimes \Lambda^0$ of*

$$h = C \otimes 1 + \delta^{-1}\left(Dh - \frac{1}{t}\operatorname{ad}(\varrho)h - \frac{\frac{1}{t}\operatorname{ad}(h)}{\exp\left(\frac{1}{t}\operatorname{ad}(h)\right) - \mathsf{id}}(\varrho' - \varrho)\right), \quad \sigma(h) = 0. \tag{5-2}$$

For this h we have

$$\mathscr{D}'_{\mathrm{F}} = \mathcal{A}_h \mathscr{D}_{\mathrm{F}} \mathcal{A}_{-h},$$

with $\mathcal{A}_h = \exp((1/t)\operatorname{ad}(h))$ being an automorphism of $\circ_\pi$.

Proof. In the context of the Fedosov construction it is well known that cohomologous two-forms yield equivalent star products. The above approach with the explicit formula for h follows the arguments of [Reichert and Waldmann 2016, Lemma 3.5] which is based on [Neumaier 2001, §3.5.1.1]. □

Lemma 5.2. *Let $\Omega, \Omega' \in t\Lambda^2\mathfrak{g}^*[[t]]$ be δ_{CE}-cohomologous. Then the corresponding Fedosov twists are equivalent.*

Proof. By assumption, we can find an element $C \in t\mathfrak{g}^*[[t]]$, such that $\Omega - \Omega' = \delta_{\mathrm{CE}} C$. From Lemma 5.1 we get an element $h \in \mathcal{W}_3 \otimes \Lambda^0$ such that $\mathscr{D}'_F = \mathcal{A}_h \mathscr{D}_{\mathrm{F}} \mathcal{A}_{-h}$. An easy computation shows that $\mathcal{A}_h$ commutes with L, therefore

$$\mathscr{D}' = \mathcal{A}_h \mathscr{D} \mathcal{A}_{-h}.$$

Thus, $\mathcal{A}_h$ is an automorphism of m_π with $\mathcal{A}_h : \ker \mathscr{D} \to \ker \mathscr{D}'$ being a bijection between the two kernels. Let us consider the map

$$S_h : \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]] \ni \xi \mapsto (\sigma \circ \mathcal{A}_h \circ \tau)(\xi) \in \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]],$$

which defines an equivalence of star products, i.e.,

$$S_h(\xi \star \eta) = S_h(\xi) \star' S_h(\eta) \tag{5-3}$$

for any $\xi, \eta \in \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))[[t]]$. Let $\xi, \eta \in \mathscr{U}(\mathfrak{g})$. Then using Lemma 4.3,

$$\begin{aligned} S_h(\xi \otimes \eta) &= (\sigma \circ \mathcal{A}_h \circ \tau)(\xi \otimes \eta) \\ &= (\sigma \circ \mathcal{A}_h)(\tau(1 \otimes 1) \cdot (\xi \otimes \eta)) \\ &= \sigma((\mathcal{A}_h(\tau(1 \otimes 1))) \cdot (\xi \otimes \eta)) \\ &= \sigma(\mathcal{A}_h(\tau(1 \otimes 1))) \cdot (\xi \otimes \eta) \\ &= \sigma(\mathcal{A}_h(\Delta^{\mathrm{Lift}}\tau(1))) \cdot (\xi \otimes \eta) \\ &= \Delta(\sigma(\mathcal{A}_h(\tau(1)))) \cdot (\xi \otimes \eta) \\ &= \Delta(S_h(1)) \cdot (\xi \otimes \eta). \end{aligned}$$

From the linearity of S_h we immediately get $S_h(\xi \star \eta) = \Delta(S_h(1))(\xi \star \eta)$. Now, putting $\xi = \eta = 1$ in (5-3) and using (4-11) we obtain

$$\Delta(S_h(1)) \cdot (1 \star 1) = S_h(1 \star 1) = S_h(1) \star' S_h(1) = (1 \star' 1) \cdot (S_h(1) \otimes S_h(1)).$$

Thus, the twists $\mathcal{F}_\Omega = 1 \star 1$ and $\mathcal{F}_{\Omega'} = 1 \star' 1$ are equivalent since

$$\epsilon(S_h(1)) = 1.$$

□

Lemma 5.3. *Let $\Omega \in t\Lambda^2\mathfrak{g}^*$ with $\delta_{\mathrm{CE}}\Omega = 0$, x the element in $\mathcal{W}_2 \otimes \Lambda^1$ uniquely determined from Proposition 2.9 and $\mathcal{F}_\Omega$ the corresponding Fedosov twist.*

(i) *The lowest total degree of ϱ, where Ω_k appears, is $2k+1$, and*

$$\varrho^{(2k+1)} = t^k \delta^{-1} \Omega_k + \textit{terms not containing } \Omega_k. \tag{5-4}$$

(ii) *For $\xi \in \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$ the lowest total degree of $\tau(\xi)$ where Ω_k appears is $2k+1$, and*

$$\tau(\xi)^{(2k+1)} = \tfrac{1}{2}t^k(e_i \otimes \mathrm{i_a}((e^i)^\sharp)\Omega_k) + \textit{terms not containing } \Omega_k. \tag{5-5}$$

(iii) *The lowest t-degree of $\mathcal{F}_\Omega$ where Ω_k appears is $k+1$, and*

$$(F_\Omega)_{k+1} = -\tfrac{1}{2}(\Omega_k)^\sharp + \textit{terms not containing } \Omega_k.$$

(iv) *The map $\Omega \mapsto \mathcal{F}_\Omega$ is injective.*

Proof. The proof uses the recursion formula for ϱ as well as the explicit formulas for τ and $\star$ and consists of a careful counting of degrees. It follows along lines of [Waldmann 2007, Theorem 6.4.29]. □

Lemma 5.4. *Let $\mathcal{F}_\Omega$ and $\mathcal{F}_{\Omega'}$ be two equivalent Fedosov twists corresponding to the closed two-forms $\Omega, \Omega' \in t\Lambda^2\mathfrak{g}^*$. Then there exists an element $C \in t\mathfrak{g}^*[[t]]$, such that $\delta_{\mathrm{CE}}C = \Omega - \Omega'$.*

Proof. We can assume that Ω and Ω' coincide up to order $k-1$ for $k \in \mathbb{N}$, since they coincide at order 0. Due to Lemma 5.3,

$$(F_\Omega)_i = (F_{\Omega'})_i \quad \text{for any } i \in \{0, \dots, k\}$$

and

$$(F_\Omega)_{k+1} - (F_{\Omega'})_{k+1} = \tfrac{1}{2}(-\Omega_k^\sharp + \Omega'^\sharp_k).$$

From Lemma B.4, we know that we can find an element $\xi \in \mathfrak{g}^*$, such that

$$([(F_\Omega)_{k+1} - (F_{\Omega'})_{k+1}])^\flat = -\Omega_k^\sharp + \Omega'^\sharp_k = \delta_{\mathrm{CE}}\xi,$$

where $[(F_\Omega)_{k+1} - (F_{\Omega'})_{k+1}]$ denotes the skew-symmetrization of $(F_\Omega)_{k+1} - (F_{\Omega'})_{k+1}$. Let us define $\hat{\Omega} = \Omega - t^k\delta_{\mathrm{CE}}\xi$. From Lemma 5.3 we see that

$$(F_{\hat{\Omega}})_{k+1} - (F_{\Omega'})_{k+1} = 0.$$

Therefore the two twists $\mathcal{F}_{\hat{\Omega}}$ and $\mathcal{F}_{\Omega'}$ coincide up to order $k+1$. Finally, since $\mathcal{F}_{\hat{\Omega}}$ and $\mathcal{F}_\Omega$ are equivalent (from Lemma 5.2) and $\mathcal{F}_\Omega$ and $\mathcal{F}_{\Omega'}$ are equivalent by assumption, the two twists $\mathcal{F}_{\hat{\Omega}}$ and $\mathcal{F}_{\Omega'}$ are also equivalent. By induction, we find an element $C \in t\mathfrak{g}^*[[t]]$ such that

$$\mathcal{F}_{\Omega+\delta_{\mathrm{CE}}C} = \mathcal{F}_{\Omega'},$$

and therefore, from Lemma 5.3, $\Omega + \delta_{\mathrm{CE}}C = \Omega'$. □

Lemma 5.5. *Let $\mathcal{F} \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be a formal twist with r-matrix r. Then there exists a Fedosov twist $\mathcal{F}_\Omega$ such that $\mathcal{F} \sim \mathcal{F}_\Omega$.*

Proof. Let $\mathcal{F} \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be a given twist. We can assume that there is a Fedosov twist $\mathcal{F}_\Omega$, which is equivalent to $\mathcal{F}$ up to order k. Therefore we find a $\hat{\mathcal{F}}$ such that $\hat{\mathcal{F}}$ is equivalent to $\mathcal{F}$ and coincides with $\mathcal{F}_\Omega$ up to order k. Due to Lemma B.4, we can find an element $\xi \in \mathfrak{g}^*$ such that

$$[(F_\Omega)_{k+1} - \hat{F}_{k+1})] = (\delta_{\mathrm{CE}}\xi)^\sharp.$$

From Lemma 5.2, the twist $\mathcal{F}_{\Omega'}$ corresponding to $\Omega' = \Omega - t^k \delta_{\mathrm{CE}}\xi$ is equivalent to $\mathcal{F}_\Omega$. Moreover, $\mathcal{F}_{\Omega'}$ coincides with $\hat{\mathcal{F}}$ up to order k, since $\mathcal{F}_{\Omega'}$ coincides with $\mathcal{F}_\Omega$ and

$$(F_{\Omega'})_{k+1} = (F_\Omega)_{k+1} + \tfrac{1}{2}\delta_{\mathrm{CE}}\xi.$$

Therefore the skew-symmetric part of $(F_{\Omega'})_{k+1} - \hat{F}_{k+1}$ is vanishing and this difference is exact with respect to the differential defined in (A-1). Applying Lemma B.2, we can see that $\mathcal{F}_{\Omega'}$ is equivalent to $\hat{\mathcal{F}}$ up to order $k+1$. The claim follows by induction. □

Summing up all the above lemmas we obtain the following characterization of the equivalence classes of twists:

Theorem 5.6 (Classification of twists). *Let $\mathfrak{g}$ be a Lie algebra over* R *such that $\mathfrak{g}$ is free and finite-dimensional and let $r \in \Lambda^2\mathfrak{g}$ be a classical r-matrix such that $\sharp$ is bijective. Then the set of equivalence classes of twists* $\mathrm{Twist}(\mathscr{U}(\mathfrak{g}), r)$ *with r-matrix r is in bijection to* $\mathrm{H}^2_{\mathrm{CE}}(\mathfrak{g})[[t]]$ *via $\Omega \mapsto \mathcal{F}_\Omega$.*

It is important to remark that even for an abelian Lie algebra $\mathfrak{g}$ the second Chevalley–Eilenberg cohomology $\mathrm{H}^2_{\mathrm{CE}}(\mathfrak{g})[[t]]$ is different from zero. Thus, not all twists are equivalent. An example of a Lie algebra with trivial $\mathrm{H}^2_{\mathrm{CE}}(\mathfrak{g})[[t]]$ is the two-dimensional nonabelian Lie algebra:

Example 5.7 ($ax+b$). Let us consider the two-dimensional Lie algebra given by the R-span of the elements $X, Y \in \mathfrak{g}$ fulfilling

$$[X, Y] = Y, \tag{5-6}$$

with r-matrix $r = X \wedge Y$. We denote the dual basis of $\mathfrak{g}^*$ by $\{X^*, Y^*\}$. Since $\mathfrak{g}$ is two-dimensional, all elements of $\Lambda^2\mathfrak{g}^*$ are a multiple of $X^* \wedge Y^*$, which is closed for dimensional reasons. For Y^* we have

$$(\delta_{\mathrm{CE}}Y^*)(X, Y) = -Y^*([X, Y]) = -Y^*(Y) = -1. \tag{5-7}$$

Therefore $\delta_{\mathrm{CE}}Y^* = -X^* \wedge Y^*$ and $\mathrm{H}^2_{\mathrm{CE}}(\mathfrak{g}) = \{0\}$. From Theorem 5.6 we can therefore conclude that all twists with r-matrix r of $\mathfrak{g}$ are equivalent.

Remark 5.8 (Original construction of Drinfeld). Let us briefly recall the original construction of Drinfeld [1983, Theorem 6]: as a first step he uses the inverse

$B \in \Lambda^2\mathfrak{g}^*$ of r as a 2-cocycle to extend $\mathfrak{g}$ to $\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{R}$ by considering the new bracket

$$[(X, \lambda), (X', \lambda')]_{\tilde{\mathfrak{g}}} = ([X, X']_{\mathfrak{g}}, B(X, X')), \tag{5-8}$$

where $X, X' \in \mathfrak{g}$ and $\lambda, \lambda' \in \mathbb{R}$. On $\tilde{\mathfrak{g}}^*$ one has the canonical star product quantizing the linear Poisson structure $\star_{DG}$ according to Drinfeld and Gutt [Gutt 1983]. Inside $\tilde{\mathfrak{g}}^*$ one has an affine subspace defined by $H = \mathfrak{g}^* + \ell_0$ where ℓ_0 is the linear functional $\ell_0 : \tilde{\mathfrak{g}} \ni (X, \lambda) \mapsto \lambda$. Since the extension is central, $\star_{DG}$ turns out to be tangential to H, therefore it restricts to an associative star product on H. In a final step, Drinfeld then uses a local diffeomorphism $G \to H$ by mapping g to $\mathrm{Ad}^*_{g^{-1}} \ell_0$ to pull back the star product to G, which turns out to be left-invariant. By [Drinfeld 1983, Theorem 1] this gives a twist. Without major modification it should be possible to include also closed higher order terms $\Omega \in t\Lambda^2\mathfrak{g}^*[[t]]$ by considering $B + \Omega$ instead. We conjecture that

(i) this gives all possible classes of Drinfeld twists by modifying his construction including Ω,

(ii) the resulting classification matches the classification by our Fedosov construction.

Note that a direct comparison of the two approaches will be nontrivial due to the presence of the combinatorics in the BCH formula inside $\star_{DG}$ in the Drinfeld construction on the one hand and the recursion in our Fedosov approach on the other hand. We will come back to this in a future project.

6. Hermitian and completely positive deformations

In this section we bring aspects of positivity into the picture. In addition, let R be now an ordered ring and set $\mathsf{C} = \mathsf{R}(\mathrm{i})$ where $\mathrm{i}^2 = -1$. In C we have a complex conjugation as usual, denoted by $z \mapsto \bar{z}$. The Lie algebra $\mathfrak{g}$ will now be a Lie algebra over R, still being free as a R-module with finite dimension.

The formal power series $\mathsf{R}[[t]]$ are then again an ordered ring in the usual way and we have $\mathsf{C}[[t]] = (\mathsf{R}[[t]])(\mathrm{i})$. Moreover, we consider a *-algebra $\mathscr{A}$ over C which we would like to deform. Here we are interested in *Hermitian* deformations $\star$, where we require

$$(a \star b)^* = b^* \star a^* \quad \text{for all } a, b \in \mathscr{A}[[t]]. \tag{6-1}$$

Instead of the universal enveloping algebra directly, we consider now the complexified universal enveloping algebra $\mathscr{U}_{\mathsf{C}}(\mathfrak{g}) = \mathscr{U}(\mathfrak{g}) \otimes_{\mathsf{R}} \mathsf{C} = \mathscr{U}(\mathfrak{g}_{\mathsf{C}})$, where $\mathfrak{g}_{\mathsf{C}} = \mathfrak{g} \otimes_{\mathsf{R}} \mathsf{C}$ is the complexified Lie algebra. Then this is a *-Hopf algebra, where the *-involution is determined by the requirement

$$X^* = -X \tag{6-2}$$

for $X \in \mathfrak{g}$, i.e., the elements of $\mathfrak{g}$ are *anti-Hermitian*. The needed compatibility of the action of $\mathfrak{g}$ on $\mathscr{A}$ with the *-involution is then

$$(\xi \triangleright a)^* = S(\xi)^* \triangleright a^* \tag{6-3}$$

for all $\xi \in \mathscr{U}_{\mathbb{C}}(\mathfrak{g})$ and $a \in \mathscr{A}$. This is equivalent to $(X \triangleright a)^* = X \triangleright a^*$ for $X \in \mathfrak{g}$. We also set the elements of $\mathfrak{g}^* \subseteq \mathfrak{g}^*_{\mathbb{C}}$ to be *anti-Hermitian*.

In a first step we extend the complex conjugation to tensor powers of $\mathfrak{g}^*_{\mathbb{C}}$ and hence to the complexified Fedosov algebra

$$\mathcal{W}_{\mathbb{C}} \otimes \Lambda^\bullet_{\mathbb{C}} = \left(\prod_{k=0}^{\infty} \mathrm{S}^k \mathfrak{g}^*_{\mathbb{C}} \otimes \Lambda^\bullet \mathfrak{g}^*_{\mathbb{C}} \right) [\![t]\!] \tag{6-4}$$

and obtain a (graded) *-involution, i.e.,

$$((f \otimes \alpha) \cdot (g \otimes \beta))^* = (-1)^{ab} (g \otimes \beta)^* \cdot (f \otimes \alpha)^*, \tag{6-5}$$

where a and b are the antisymmetric degrees of α and β, respectively.

Let $\pi \in \mathfrak{g}_{\mathbb{C}} \otimes \mathfrak{g}_{\mathbb{C}}$ have antisymmetric and symmetric parts $\pi_- \in \Lambda^2 \mathfrak{g}_{\mathbb{C}}$ and $\pi_+ \in \Lambda^2 \mathfrak{g}_{\mathbb{C}}$, respectively. Then for the corresponding operator $\mathcal{P}_\pi$ as in (2-13),

$$\mathsf{T} \circ \overline{\mathcal{P}_\pi (a \otimes b)} = \mathcal{P}_{\tilde{\pi}} \circ \mathsf{T}(\bar{a} \otimes \bar{b}), \tag{6-6}$$

where $\tilde{\pi} = \bar{\pi}_+ - \bar{\pi}_-$. In particular, we have $\tilde{\pi} = \pi$ if and only if π_+ is Hermitian and π_- is anti-Hermitian. We set $t = \mathrm{i}t$ for the formal parameter as in the previous sections, i.e., we want to treat t as imaginary. Then we arrive at the following statement:

Lemma 6.1. *Let $\pi = \pi_+ + \pi_- \in \mathfrak{g}_{\mathbb{C}} \otimes \mathfrak{g}_{\mathbb{C}}$. Then the fiberwise product*

$$a \circ_\pi b = \mu \circ \mathrm{e}^{\frac{1}{2}\mathrm{i}t\mathcal{P}_\pi} (a \otimes b) \tag{6-7}$$

satisfies $(a \circ_\pi b)^* = (-1)^{ab} b^* \circ a^*$ if and only if π_+ is anti-Hermitian and π_- is Hermitian.

This lemma is now the motivation to take a *real* classical r-matrix $r \in \Lambda^2 \mathfrak{g} \subseteq \Lambda^2 \mathfrak{g}_{\mathbb{C}}$. Moreover, writing the symmetric part of π as $\pi_+ = \mathrm{i}s$, then $s = \bar{s} \in \mathrm{S}^2 \mathfrak{g}$ is Hermitian as well. In the following we shall assume that these reality conditions are satisfied.

It is now not very surprising that with such a Poisson tensor π on $\mathfrak{g}$ we can achieve a Hermitian deformation of a *-algebra $\mathscr{A}$ by the Fedosov construction. We summarize the relevant properties in the following proposition:

Proposition 6.2. *Let $\pi = r + \mathrm{i}s$ with a real strongly nondegenerate r-matrix $r \in \Lambda^2 \mathfrak{g}$ and a real symmetric $s \in \mathrm{S}^2 \mathfrak{g}$ such that there exists a symplectic torsion-free covariant derivative ∇ for $\mathfrak{g}$ with $\nabla s = 0$.*

(i) *The operators* δ, δ^{-1}, *and* σ *are real.*

(ii) *The operator* D *is real and* $D^2 = (1/\mathrm{i}t)\,\mathrm{ad}(R)$ *with a Hermitian curvature* $R = R^*$.

(iii) *Suppose that* $\Omega = \Omega^* \in \Lambda^2 \mathfrak{g}^*_{\mathsf{C}}[[t]]$ *is a formal series of Hermitian* δ_{CE}*-closed two-forms. Then the unique* $\varrho \in \mathcal{W}_2 \otimes \Lambda^1$ *with*

$$\delta\varrho = R + D\varrho + \frac{1}{\mathrm{i}t}\varrho \circ_\pi \varrho + \Omega \tag{6-8}$$

and $\delta^{-1}\varrho = 0$ *is Hermitian, too. In this case, the Fedosov derivative* $\mathscr{D}_{\mathrm{F}} = -\delta + D + 1/(\mathrm{i}t)\,\mathrm{ad}(\varrho)$ *is real.*

Suppose now in addition that $\mathscr{A}$ *is a* **-algebra over* C *with a* **-action of* $\mathfrak{g}$, *i.e.*, (6-3).

(iv) *The operator* $L_{\mathscr{A}}$ *as well as the extended Fedosov derivation* $\mathscr{D}_{\mathscr{A}}$ *are real.*

(v) *The Fedosov–Taylor series* $\tau_{\mathscr{A}}$ *is real.*

(vi) *The formal deformation* $\star$ *from Theorem 3.6 is a Hermitian deformation.*

When we apply this to the twist itself we first have to clarify which *-involution we take on the tensor algebra $\mathrm{T}^\bullet(\mathscr{U}_{\mathsf{C}}(\mathfrak{g}))$: by the universal property of the tensor algebra, there is a unique way to extend the *-involution of $\mathscr{U}_{\mathsf{C}}(\mathfrak{g})$ as a *-involution. With respect to this *-involution we have $r^* = -r$ since r is not only real as an element of $\mathfrak{g}_{\mathsf{C}} \otimes \mathfrak{g}_{\mathsf{C}}$ but also antisymmetric, causing an additional sign with respect to the *-involution of $\mathrm{T}^\bullet(\mathscr{U}_{\mathsf{C}}(\mathfrak{g}))$. Analogously, we have $s^* = s$ for the real and symmetric part of π.

Corollary 6.3. *The Fedosov twist* $\mathcal{F}$ *is Hermitian.*

Proof. Indeed, $1 \in \mathscr{U}_{\mathsf{C}}(\mathfrak{g})$ is Hermitian and hence $(1 \star 1)^* = 1^* \star 1^* = 1 \star 1$. □

Up to now we have not yet used the fact that R is ordered but only that we have a *-involution. The ordering of R allows one to transfer concepts of positivity from R to every *-algebra over C. Recall that a linear functional $\omega : \mathscr{A} \to \mathsf{C}$ is called *positive* if

$$\omega(a^* a) \geq 0 \tag{6-9}$$

for all $a \in \mathscr{A}$. This allows one to define an algebra element $a \in \mathscr{A}$ to be *positive* if $\omega(a) \geq 0$ for all positive ω. Note that the positive elements denoted by $\mathscr{A}^+$, form a convex cone in $\mathscr{A}$ and $a \in \mathscr{A}^+$ implies $b^* a b \in \mathscr{A}^+$ for all $b \in \mathscr{A}$. Moreover, elements of the form $a = b^* b$ are clearly positive: their convex combinations are denoted by $\mathscr{A}^{++}$ and are called *algebraically positive*. More details on these notions of positivity can be found in [Bursztyn and Waldmann 2001; 2005a; Waldmann 2005].

Since with R also $\mathsf{R}[[t]]$ is ordered, one can compare the positive elements of $\mathscr{A}$ and the ones of $(\mathscr{A}[[t]], \star)$, where $\star$ is a Hermitian deformation. The first trivial observation is that for a positive linear functional $\boldsymbol{\omega}=\omega_0+t\omega_1+\cdots$ of the deformed algebra, i.e., $\boldsymbol{\omega}(a^*\star a)\geq 0$ for all $a\in\mathscr{A}[\![t]\!]$ the classical limit ω_0 of $\boldsymbol{\omega}$ is a positive functional of the undeformed algebra. The converse need not be true: one has examples where a positive ω_0 is not directly positive for the deformed algebras, i.e., one needs higher order corrections, and one has examples where one simply can not find such higher order corrections at all; see [Bursztyn and Waldmann 2000; 2005b]. One calls the deformation $\star$ a *positive deformation* if every positive linear functional ω_0 of the undeformed algebra $\mathscr{A}$ can be deformed into a positive functional $\boldsymbol{\omega}=\omega_0+t\omega_1+\cdots$ of the deformed algebra $(\mathscr{A}[\![t]\!],\star)$. Moreover, since also $\mathrm{M}_n(\mathscr{A})$ is a *-algebra in a natural way we call $\star$ a *completely positive deformation* if for all n the canonical extension of $\star$ to $\mathrm{M}_n(\mathscr{A})[\![t]\!]$ is a positive deformation of $\mathrm{M}_n(\mathscr{A})$; see [Bursztyn and Waldmann 2005b]. Finally, if no higher order corrections are needed, then $\star$ is called a *strongly positive deformation*; see [Bursztyn and Waldmann 2000, Definition 4.1]

In a next step we want to use a Kähler structure for $\mathfrak{g}$. In general, this will not exist so we have to require it explicitly. In detail, we want to be able to find a basis $e_1,\ldots,e_n,f_1,\ldots,f_n\in\mathfrak{g}$ with the property that the r-matrix decomposes into

$$(e^k\otimes f^\ell)(r)=A^{k\ell}=-(f^\ell\otimes e^k)(r),\quad (e^k\otimes e^\ell)(r)=B^{k\ell}=-(f^k\otimes f^\ell)(r) \tag{6-10}$$

with a symmetric matrix $A=A^{\mathrm{T}}\in\mathrm{M}_n(\mathsf{R})$ and an antisymmetric matrix $B=-B^{\mathrm{T}}\in\mathrm{M}_n(\mathsf{R})$. We set

$$s=A^{k\ell}(e_k\otimes e_\ell+f_k\otimes f_\ell)+B^{k\ell}e_k\otimes f_\ell+B^{k\ell}f_\ell\otimes e_k. \tag{6-11}$$

The requirement of being *Kähler* is now that first we find a symplectic covariant derivative ∇ with $\nabla s=0$. Second, we require the symmetric two-tensor s to be positive in the sense that for all $x\in\mathfrak{g}^*$ we have $(x\otimes x)(s)\geq 0$. In this case we call s (and the compatible ∇) a Kähler structure for r. We have chosen this more coordinate-based formulation over the invariant one since in the case of an ordered ring R instead of the reals $\mathbb{R}$ it is more convenient to start directly with the nice basis we need later on.

As usual we consider now $\mathfrak{g}_{\mathsf{C}}$ with the vectors

$$Z_k=\tfrac{1}{2}(e_k-\mathrm{i}f_k)\quad\text{and}\quad \bar{Z}_\ell=\tfrac{1}{2}(e_\ell+\mathrm{i}f_\ell), \tag{6-12}$$

which together constitute a basis of the complexified Lie algebra. Finally, we have the complex matrix

$$g=A+\mathrm{i}B\in\mathrm{M}_n(\mathsf{C}), \tag{6-13}$$

which now satisfies the positivity requirement

$$\bar{z}_k g^{k\ell}z_\ell\geq 0\quad\text{for all } z_1,\ldots,z_n\in\mathsf{C}. \tag{6-14}$$

If our ring R has sufficiently many inverses and square roots, one can even find a basis $e_1, \dots, e_n, f_1, \dots, f_n$ such that g becomes the unit matrix. However, since we want to stay with an arbitrary ordered ring R we do not assume this.

We now use $\pi = r + \mathrm{i}s$ to obtain a fiberwise Hermitian product $\circ_{\mathrm{Wick}}$, called the fiberwise Wick product. Important now is the following explicit form of $\circ_{\mathrm{Wick}}$, which is a routine verification:

Lemma 6.4. *For the fiberwise Wick product $\circ_{\mathrm{Wick}}$ built out of $\pi = r + \mathrm{i}s$ with a Kähler structure s one has*

$$a \circ_{\mathrm{Wick}} b = \mu \circ \mathrm{e}^{2tg^{k\ell}\,\mathrm{i_s}(Z_k)\otimes \mathrm{i_s}(\bar{Z}_\ell)}(a \otimes b), \tag{6-15}$$

where g is the matrix from (6-13).

The first important observation is that the scalar matrix g can be viewed as an element of $\mathrm{M}_n(\mathscr{A})$ for any unital *-algebra. Then we have the following positivity property:

Lemma 6.5. *Let $\mathscr{A}$ be a unital *-algebra over C. Then for all $m \in \mathbb{N}$ and for all $a_{k_1 \cdots k_m} \in \mathscr{A}$ with $k_1, \dots, k_m = 1, \dots, n$*

$$\sum_{k_1,\ell_1,\dots,k_m,\ell_m=1}^{n} g^{k_1\ell_1} \cdots g^{k_m\ell_m} a^*_{k_1\cdots k_m} a_{\ell_1\cdots\ell_m} \in \mathscr{A}^+. \tag{6-16}$$

Proof. First we note that $g^{\otimes m} = g \otimes \cdots \otimes g \in \mathrm{M}_n(\mathsf{C}) \otimes \cdots \otimes \mathrm{M}_n(\mathsf{C}) = \mathrm{M}_{n^m}(\mathsf{C})$ still satisfies the positivity property

$$\sum_{k_1,\ell_1,\dots,k_m,\ell_m=1}^{n} g^{k_1\ell_1} \cdots g^{k_m\ell_m} \overline{z^{(1)}_{k_1}} \cdots \overline{z^{(m)}_{k_m}} z^{(1)}_{\ell_1} \cdots z^{(m)}_{\ell_m} \geq 0 \quad \text{for all } z^{(1)}, \dots, z^{(m)} \in \mathsf{C}^n$$

as the left-hand side clearly factorizes into m copies of the left hand side of (6-14). Hence $g^{\otimes m} \in \mathrm{M}_{n^m}(\mathsf{C})$ is a positive element. For a given positive linear functional $\omega : \mathscr{A} \to \mathsf{C}$ and $b_1, \dots, b_N \in \mathscr{A}$ we consider the matrix $(\omega(b_i^* b_j)) \in \mathrm{M}_N(\mathsf{C})$. We claim that this matrix is positive, too. Indeed, with the criterion from [Bursztyn and Waldmann 2001, App. A], for all $z_1, \dots, z_N \in \mathsf{C}$,

$$\sum_{i,j=1}^{N} \bar{z}_i \omega(b_i^* b_j) z_j = \omega\Bigg(\bigg(\sum_{i=1}^{N} z_i b_i\bigg)^* \bigg(\sum_{j=1}^{N} z_j b_j\bigg)\Bigg) \geq 0,$$

hence $(\omega(b_i^* b_j))$ is positive. Putting these statements together we see, for every positive linear functional $\omega : \mathscr{A} \to \mathsf{C}$, for the matrix $\Omega = (\omega(a^*_{k_1\cdots k_m} a_{\ell_1\cdots\ell_m})) \in \mathrm{M}_{n^m}(\mathsf{C})$,

$$\begin{aligned}\omega\Bigg(\sum_{k_1,\ell_1,\dots,k_m,\ell_m=1}^{n} g^{k_1\ell_1}\cdots g^{k_m\ell_m} a^*_{k_1\cdots k_m} a_{\ell_1\cdots\ell_m}\Bigg) &= \sum_{k_1,\ell_1,\dots,k_m,\ell_m=1}^{n} g^{k_1\ell_1}\cdots g^{k_m\ell_m}\,\omega(a^*_{k_1\cdots k_m} a_{\ell_1\cdots\ell_m}) \\ &= \mathsf{tr}(g^{\otimes m}\Omega) \geq 0,\end{aligned}$$

since the trace of the product of two positive matrices is positive by [Bursztyn and Waldmann 2001, Appendix A]. Note that for a *ring* R one has to use this slightly more complicated argumentation: for a field one could use the diagonalization of g instead. By definition of $\mathscr{A}^+$, this shows the positivity of (6-16). $\square$

Remark 6.6. Suppose that in addition $g = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is diagonal with positive $\lambda_1, \dots, \lambda_n > 0$. In this case one can directly see that the left-hand side of (6-16) is a convex combination of squares and hence in $\mathscr{A}^{++}$. This situation can often be achieved, e.g., for $\mathsf{R} = \mathbb{R}$.

We come now to the main theorem of this section: unlike the Weyl-type deformation, using the fiberwise Wick product yields a positive deformation in a universal way:

Theorem 6.7. *Let $\mathscr{A}$ be a unital *-algebra over $\mathsf{C} = \mathsf{R}(\mathrm{i})$ with a *-action of $\mathfrak{g}$ and let $\Omega = \Omega^* \in \Lambda^2\mathfrak{g}^*_{\mathsf{C}}$ be a formal series of Hermitian δ_{CE}-closed two-forms. Moreover, let s be a Kähler structure for the nondegenerate r-matrix $r \in \mathfrak{g}$ and consider the fiberwise Wick product $\circ_{\mathrm{Wick}}$ yielding the Hermitian deformation $\star_{\mathrm{Wick}}$ as in Proposition 6.2.*

(i) *For all $a \in \mathscr{A}$,*

$$a^* \star_{\mathrm{Wick}} a = \sum_{m=0}^{\infty} \frac{(2t)^m}{m!} \sum_{k_1,\dots,k_m,\ell_1,\dots,\ell_m=1}^{n} g^{k_1\ell_1}\cdots g^{k_m\ell_m} a^*_{k_1\cdots k_m} a_{\ell_1\cdots\ell_m}, \tag{6-17}$$

where $a_{k_1\cdots k_m} = \sigma(\mathrm{i_s}(\bar{Z}_{k_1})\cdots \mathrm{i_s}(\bar{Z}_{k_m})\tau_{\mathrm{Wick}}(a))$.

(ii) *The deformation $\star_{\mathrm{Wick}}$ is strongly positive.*

Proof. From Lemma 6.4 we immediately get (6-17). Now let $\omega : \mathscr{A} \to \mathsf{C}$ be positive. Then also the $\mathsf{C}[\![t]\!]$-linear extension $\omega : \mathscr{A}[\![t]\!] \to \mathsf{C}[\![t]\!]$ is positive with respect to the undeformed product: this is a simple consequence of the Cauchy–Schwarz inequality for ω. Then we apply Lemma 6.5 to conclude that $\omega(a^* \star a) \geq 0$. $\square$

Corollary 6.8. *The Wick-type twist $\mathcal{F}_{\mathrm{Wick}}$ in the Kähler situation is a convex series of positive elements.*

Remark 6.9 (Positive twist). Note that already for a Hermitian deformation, the twist $\mathcal{F} = 1 \star 1 = 1^* \star 1$ constructed as above is a *positive* element of the deformed algebra $\mathrm{T}^\bullet(\mathscr{U}_\mathbb{C}(\mathfrak{g}))[[t]]$. However, this seems to be not yet very significant: it is the statement of Corollary 6.8 and Theorem 6.7 which gives the additional and important feature of the corresponding universal deformation formula.

Appendix A: Hochschild-Kostant-Rosenberg theorem

Let us define the map

$$\partial : \mathscr{U}(\mathfrak{g}) \ni \xi \mapsto \xi \otimes 1 + 1 \otimes \xi - \Delta(\xi) \in \mathscr{U}(\mathfrak{g})^{\otimes 2}, \tag{A-1}$$

and extend it as a graded derivation of degree $+1$ of the tensor product to $\mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$. We recall that the map $\partial : \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g})) \to \mathrm{T}^\bullet(\mathscr{U}(\mathfrak{g}))$ is a differential. Its cohomology is described as follows:

Theorem A.1 (Hochschild–Kostant–Rosenberg). *Let $C \in \mathrm{T}^p(\mathscr{U}(\mathfrak{g}))$ such that $\partial C = 0$. Then there is a $X \in \Lambda^k \mathfrak{g}$ and a $S \in \mathrm{T}^{p-1}(\mathscr{U}(\mathfrak{g}))$ with*

$$C = X + \partial S \tag{A-2}$$

with $X = \mathrm{Alt}(C)$.

We do not prove the above theorem in full generality, since we need only the case $p = 2$. In this case the proof consists of the following two lemmas:

Lemma A.2. *Let $C \in \mathrm{T}^2(\mathscr{U}(\mathfrak{g}))$ with $\partial C = 0$. Then*

(i) $\partial \mathtt{T}(C) = 0$.

(ii) *The antisymmetric part satisfies $C - \mathtt{T}(C) \in \mathfrak{g} \wedge \mathfrak{g} \subseteq \mathrm{T}^2(\mathscr{U}(\mathfrak{g}))$.*

Proof. We have

$$\partial C = 0 \Longleftrightarrow C \otimes 1 + (\Delta \otimes \mathsf{id})(C) = 1 \otimes C + (\mathsf{id} \otimes \Delta)(C).$$

Thus,

$$\begin{aligned} \mathtt{T}(C) \otimes 1 &= (\mathtt{T} \otimes \mathsf{id})(C \otimes 1) \\ &= (\mathtt{T} \otimes \mathsf{id})(1 \otimes C + (\mathsf{id} \otimes \Delta)(C) - (\Delta \otimes \mathsf{id})(C)) \\ &= C_{13} + (\mathtt{T} \otimes \mathsf{id})(\mathsf{id} \otimes \Delta)(C) - (\Delta \otimes \mathsf{id})(C). \end{aligned}$$

Now we apply the cyclic permutation to this equation and get

$$1 \otimes \mathtt{T}(C) = \mathtt{T}(C) \otimes 1 + (\Delta \otimes \mathsf{id})(\mathtt{T}(C)) - (\mathsf{id} \otimes \Delta)(\mathtt{T}(C)),$$

which is equivalent to $\partial \mathtt{T}(C) = 0$. Since ∂ is linear, we get $\partial(T - \mathtt{T}(C)) = 0$ and denote $A = T - \mathtt{T}(C)$, which is now skew-symmetric. We define

$$Q = (\Delta \otimes \mathsf{id})A - A_{23} - A_{13}$$

and get that $Q = -\mathtt{Alt}(Q)$ with the fact that A is ∂-closed. Therefore we have $Q = \mathtt{Alt}^3 Q = (-1)^3 Q = -Q$ and we can conclude $Q = 0$. Thus, A has to be primitive in the first argument and with the skew-symmetry we get the same statement for the second argument. □

Lemma A.3. *Let $C \in \mathrm{T}^2(\mathscr{U}(\mathfrak{g}))$ with $\partial C = 0$. Then there exists a $S \in \mathscr{U}(\mathfrak{g})$ and a $X \in \mathfrak{g} \wedge \mathfrak{g}$, such that*

$$C = X + \partial S, \tag{A-3}$$

where $X = \frac{1}{2}(C - \mathtt{T}(C))$.

Proof. It is clear from Lemma A.2, that X is well defined and we have to prove that symmetric C are ∂-exact. So we assume that $C \in \mathrm{T}^2(\mathscr{U}(\mathfrak{g}))$ is ∂-closed and symmetric. Let k be the highest order appearing in C and assume the claim is true for all $r < k$ (in the sense of the filtration of $\mathscr{U}(\mathfrak{g}) = \bigcup_{n \in \mathbb{N}_0} \mathscr{U}(\mathfrak{g})_n$). Then we can write for a given basis $\{e_i\}_{i \in \{1,\ldots,n\}}$

$$C = \sum_{|\boldsymbol{i}|=k} e_{\boldsymbol{i}} \otimes D^{\boldsymbol{i}} + \text{l.o.t.}.$$

We mean lower order terms with respect to the filtration in the first tensor degree and the $\boldsymbol{i}$ are multi-indices such that $e_{\boldsymbol{i}} = e_{i_1} \cdots e_{i_k}$. We can assume that $D_{\boldsymbol{i}}$ is symmetric in the multi-index, because we can compensate for asymmetry by lower order terms. Since $\partial(\mathscr{U}(\mathfrak{g})_m) \subseteq \mathscr{U}(\mathfrak{g})_{m-1} \otimes \mathscr{U}(\mathfrak{g})_{m-1}$, we see that $\partial C = 0$ implies that $\partial D^{\boldsymbol{i}} = 0$, which is equivalent to $D^{\boldsymbol{i}} \in \mathfrak{g}$. Therefore, we can write

$$C = \sum_{|\boldsymbol{i}|=k} D^{\boldsymbol{i},j} e_{\boldsymbol{i}} \otimes e_j + H,$$

where $H \in \mathscr{U}(\mathfrak{g})_{k-1} \otimes \mathscr{U}(\mathfrak{g})$ is now of order strictly less then k in the first argument. Now we expand $H = \sum_{|\boldsymbol{i}_1|,|\boldsymbol{i}_2| \leq k-1} H_{\boldsymbol{i}_1,\boldsymbol{i}_2} e_{\boldsymbol{i}_1} \otimes e_{\boldsymbol{i}_2}$ and see, by using

$$\begin{aligned} 0 = \partial C &= \sum_{|\boldsymbol{i}|=k} D^{\boldsymbol{i},j} \partial(e_{\boldsymbol{i}}) \otimes e_j + \partial H \\ &= -D^{i_1,\ldots,i_k,j} \sum_r e_{i_1} \cdots \widehat{e_{i_r}} \cdots e_{i_k} \otimes e_{i_r} \otimes e_j + \partial H + \text{l.o.t.}, \end{aligned}$$

that H has to be of the form

$$H = \sum_{|\boldsymbol{i}_1|=k-1, |\boldsymbol{i}_2|=2} H_{\boldsymbol{i}_1,\boldsymbol{i}_2} e_{\boldsymbol{i}_1} \otimes e_{\boldsymbol{i}_2} + \text{l.o.t.},$$

and hence

$$\partial H = \sum_{|\boldsymbol{i}_1|=k-1, j_1, j_2} H_{\boldsymbol{i}_1, j_1, j_2} e_{\boldsymbol{i}_1} \otimes e_{j_1} \otimes e_{j_2} + \text{l.o.t.}.$$

This implies that $D^{i_1,\dots,i_k,j}$ is symmetric in all indices, since $\partial C = 0$ and $H_{\boldsymbol{i}_1,j_1,j_2} = H_{\boldsymbol{i}_1,j_2,j_1}$. Thus for

$$G = \frac{1}{k+1} D^{i_1,\dots,i_{k+1}} e_{i_1} \cdots e_{i_{k+1}},$$

we have

$$\partial G = -\sum_{|\boldsymbol{i}|=k} D^{\boldsymbol{i},j}(e_{\boldsymbol{i}} \otimes e_j + e_j \otimes e_{\boldsymbol{i}}) + \text{l.o.t.}.$$

Note that here the lower order terms are meant in both tensor arguments. Using the symmetry of C, we obtain

$$C = \sum_{|\boldsymbol{i}|=k} D^{\boldsymbol{i},j}(e_{\boldsymbol{i}} \otimes e_j + e_j \otimes e_{\boldsymbol{i}}) + \text{l.o.t.},$$

again the lower order terms are in both tensor factors. Thus,

$$C + \partial G \in \mathscr{U}(\mathfrak{g})_{k-1} \otimes \mathscr{U}(\mathfrak{g})_{k-1}.$$

This implies the lemma, because for $k = 0$ the statement is trivial. □

Corollary A.4. *Let $C \in \mathrm{T}^2(\mathscr{U}(\mathfrak{g}))$ with $\partial C = 0$ and $(\epsilon \otimes \mathsf{id})C = (\mathsf{id} \otimes \epsilon)C = 0$. Then we can find $S \in \mathscr{U}(\mathfrak{g})$ and $X \in \Lambda^2\mathfrak{g}$ such that $C = X + \partial S$ with $\epsilon(S) = 0$.*

Proof. The statement is clear from the construction of Lemma A.2. □

Appendix B: Technical Lemmas

In this section we prove several technical results necessary for the proofs is Section 5.

Lemma B.1. *Let $\mathcal{F}, \mathcal{F}' \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be two twists coinciding up to order k. Then*

$$\partial(F_{k+1} - F'_{k+1}) = 0. \tag{B-1}$$

Proof. We have

$$
\begin{aligned}
\partial(F_{k+1}) &= 1 \otimes F_{k+1} - F_{k+1} \otimes 1 + (\mathsf{id} \otimes \Delta)(F_{k+1}) - (\Delta \otimes \mathsf{id})(F_{k+1}) \\
&= \sum_{i=0}^{k+1} (1 \otimes F_i)(\mathsf{id} \otimes \Delta)(F_{k+1-i}) - \sum_{i=1}^{k} (1 \otimes F_i)(\mathsf{id} \otimes \Delta)(F_{k+1-i}) \\
&\quad + \sum_{i=1}^{k} (F_i \otimes 1)(\Delta \otimes \mathsf{id})(F_{k+1-i}) - \sum_{i=0}^{k+1} (F_i \otimes 1)(\Delta \otimes \mathsf{id})(F_{k+1-i}) \\
&= -\sum_{i=1}^{k} (1 \otimes F_i)(\mathsf{id} \otimes \Delta)(F_{k+1-i}) + \sum_{i=1}^{k} (F_i \otimes 1)(\Delta \otimes \mathsf{id})(F_{k+1-i}) \\
&= -\sum_{i=1}^{k} (1 \otimes F'_i)(\mathsf{id} \otimes \Delta)(F'_{k+1-i}) + \sum_{i=1}^{k} (F'_i \otimes 1)(\Delta \otimes \mathsf{id})(F'_{k+1-i}) \\
&= \partial(F'_{k+1}). \qquad \square
\end{aligned}
$$

Lemma B.2. *Let $\mathcal{F}, \mathcal{F}' \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be two twists coinciding up to order k such that*

$$
F_{k+1} - F'_{k+1} = \partial T_{k+1}. \tag{B-2}
$$

Then they are equivalent up to order $k+1$.

Proof. Consider $\exp(t^{k+1}T_{k+1}) = 1 + t^{k+1}T_{k+1} + \mathcal{O}(t^{k+2})$. Then we have

$$
(\Delta(\exp(t^{k+1}T_{k+1}))\mathcal{F})_i = (\mathcal{F}'(\exp(t^{k+1}T_{k+1}) \otimes \exp(t^{k+1}T_{k+1})))_i
$$

for any $i \leq k+1$. Note that, because

$$
(\epsilon \otimes \mathsf{id})(F_{k+1} - F'_{k+1}) = (\mathsf{id} \otimes \epsilon)(F_{k+1} - F'_{k+1}) = 0,
$$

we can choose T_{k+1} such that $\epsilon(T_{k+1}) = 0$ and therefore $\epsilon(\exp(t^{k+1}T_{k+1})) = 1$. $\square$

Lemma B.3. *Let $\mathcal{F}, \mathcal{F}' \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be two equivalent twists coinciding up to order k. Then there exists a $T = 1 + t^k T_k + \mathcal{O}(t^{k+1}) \in \mathscr{U}(\mathfrak{g})[[t]]$ such that*

$$
\Delta(T)\mathcal{F}' = \mathcal{F}(T \otimes T). \tag{B-3}
$$

Proof. Since the twists $\mathcal{F}$ and $\mathcal{F}'$ are equivalent, there is a $\tilde{T} = 1 + t^\ell \tilde{T}_\ell + \mathcal{O}(t^{\ell+1})$ such that

$$
\Delta(\tilde{T})F' = F(\tilde{T} \otimes \tilde{T}).
$$

Let us consider $\ell \leq k$. The above equation at order ℓ reads

$$
\Delta(\tilde{T}_\ell) + F'_\ell = F_\ell + \tilde{T}_\ell \otimes 1 + 1 \otimes \tilde{T}_\ell.
$$

Therefore, since $\mathcal{F}$ and $\mathcal{F}'$ coincide up to order k,

$$\Delta(\tilde{T}_\ell) = \tilde{T}_\ell \otimes 1 + 1 \otimes \tilde{T}_\ell,$$

and $\tilde{T}_\ell \in \mathfrak{g} \subseteq \mathscr{U}(\mathfrak{g})$. For $\ell < k$ we get at order $\ell + 1$

$$\Delta(\tilde{T}_{\ell+1}) + \Delta(\tilde{T}_\ell)F'_1 + F'_{\ell+1} = F_{\ell+1} + F_1(\tilde{T}_\ell \otimes 1 + 1 \otimes \tilde{T}_\ell) + \tilde{T}_{\ell+1} \otimes 1 + 1 \otimes \tilde{T}_{\ell+1}.$$

The skew-symmetrization of the above equation gives

$$(\tilde{T}_\ell \otimes 1 + 1 \otimes \tilde{T}_\ell)r = r(\tilde{T}_\ell \otimes 1 + 1 \otimes \tilde{T}_\ell).$$

An easy computation shows that this property is equivalent to $\delta_{\mathrm{CE}}\tilde{T}_\ell^\flat = 0$. Thus, we can define the map $S : \mathscr{U}(\mathfrak{g}) \to \mathscr{U}(\mathfrak{g})$ by defining it on primitive elements via

$$\mathfrak{g} \ni \xi \mapsto \tilde{T}_\ell^\flat(\xi) \cdot 1 \in \mathscr{U}(\mathfrak{g})$$

and extend it as a derivation of the product of $\mathscr{U}(\mathfrak{g})$. This map allows us to define an element

$$A = \frac{1}{t}(\epsilon \circ S \otimes \mathsf{id})[\mathcal{F}] = -\tilde{T}_\ell + \mathcal{O}(t),$$

which fulfills $\Delta(A)\mathcal{F} = \mathcal{F}(A \otimes 1 + 1 \otimes A)$ and $\epsilon(A) = 0$. Thus we get

$$\exp(t^\ell A)\mathcal{F} = \mathcal{F}(\exp(t^\ell A) \otimes \exp(t^\ell A)) \quad \text{as well as} \quad \epsilon(\exp(t^\ell A)) = 1.$$

We define $T = \exp(t^\ell A)\tilde{T}$ and obtain $\Delta(T)\mathcal{F}' = \mathcal{F}(T \otimes T)$ and

$$T = 1 + t^{\ell+1}T_{\ell+1} + \mathcal{O}(t^{\ell+2}).$$

Repeating this method $k - \ell$ times, we get an equivalence starting at order k. □

Lemma B.4. *Let $\mathcal{F}, \mathcal{F}' \in (\mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g}))[[t]]$ be two equivalent twists coinciding up to order k. Then there exists an element $\xi \in \mathfrak{g}^*$ such that*

$$([F_{k+1} - F'_{k+1}])^\flat = \delta_{\mathrm{CE}}\xi. \tag{B-4}$$

Proof. First, $[F_{k+1} - F'_{k+1}] \in \Lambda^2\mathfrak{g}$, because of Theorem A.1 and since, as in Lemma B.1, $\partial(F_{k+1} - F'_{k+1}) = 0$. From Lemma B.3 we know that we can find an element $T = 1 + t^k T_k + \mathcal{O}(t^{k+1})$ in $\mathscr{U}(\mathfrak{g})$ such that $\Delta(T)\mathcal{F}' = \mathcal{F}(T \otimes T)$. At order k this reads

$$\Delta(T_k) + F'_k = F_k + T_k \otimes 1 + 1 \otimes T_k,$$

which is equivalent to $T_k \in \mathfrak{g}$, because $F'_k = F_k$. At order $k + 1$, we can see that

$$\Delta(T_{k+1}) + \Delta(T_k)F'_1 + F'_{k+1} = F_{k+1} + F_1(T_k \otimes 1 + 1 \otimes T_k) + T_{k+1} \otimes 1 + 1 \otimes T_{k+1}.$$

For the skew-symmetric part we have

$$[F_{k+1} - F'_{k+1}] = (T_k \otimes 1 + 1 \otimes T_k)r - r(T_k \otimes 1 + 1 \otimes T_k) = [T_k \otimes 1 + 1 \otimes T_k, r],$$

which is equivalent to $([F_{k+1} - F'_{k+1}])^\flat = -\delta_{\mathrm{CE}} T_k^\flat$. □

Acknowledgements

We would like to thank Pierre Bieliavsky, Kenny De Commer, Alexander Karabegov, and Thorsten Reichert for the discussions and useful suggestions. Moreover, we would like to thank the referee for many useful comments and remarks.

References

[Aschieri and Schenkel 2014] P. Aschieri and A. Schenkel, "Noncommutative connections on bimodules and Drinfeld twist deformation", *Adv. Theor. Math. Phys.* **18**:3 (2014), 513–612. MR Zbl

[Bayen et al. 1978a] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, "Deformation theory and quantization, I: Deformations of symplectic structures", *Ann. Physics* **111**:1 (1978), 61–110. MR Zbl

[Bayen et al. 1978b] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, "Deformation theory and quantization, II: Physical applications", *Ann. Physics* **111**:1 (1978), 111–151. MR Zbl

[Bertelson et al. 1997] M. Bertelson, M. Cahen, and S. Gutt, "Equivalence of star products", *Classical Quantum Gravity* **14**:1A (1997), A93–A107. MR Zbl

[Bertelson et al. 1998] M. Bertelson, P. Bieliavsky, and S. Gutt, "Parametrizing equivalence classes of invariant star products", *Lett. Math. Phys.* **46**:4 (1998), 339–345. MR Zbl

[Bieliavsky and Gayral 2015] P. Bieliavsky and V. Gayral, *Deformation quantization for actions of Kählerian Lie groups*, Mem. Amer. Math. Soc. **1115**, American Mathematical Society, Providence, RI, 2015. MR Zbl

[Bordemann and Waldmann 1997] M. Bordemann and S. Waldmann, "A Fedosov star product of the Wick type for Kähler manifolds", *Lett. Math. Phys.* **41**:3 (1997), 243–253. MR Zbl

[Bordemann et al. 1998] M. Bordemann, N. Neumaier, and S. Waldmann, "Homogeneous Fedosov star products on cotangent bundles, I: Weyl and standard ordering with differential operator representation", *Comm. Math. Phys.* **198**:2 (1998), 363–396. MR Zbl

[Bordemann et al. 1999] M. Bordemann, N. Neumaier, and S. Waldmann, "Homogeneous Fedosov star products on cotangent bundles, II: GNS representations, the WKB expansion, traces, and applications", *J. Geom. Phys.* **29**:3 (1999), 199–234. MR Zbl

[Bordemann et al. 2003] M. Bordemann, N. Neumaier, M. J. Pflaum, and S. Waldmann, "On representations of star product algebras over cotangent spaces on Hermitian line bundles", *J. Funct. Anal.* **199**:1 (2003), 1–47. MR Zbl

[Bursztyn and Waldmann 2000] H. Bursztyn and S. Waldmann, "On positive deformations of $*$-algebras", pp. 69–80 in *Conférence Moshé Flato, II* (Dijon, 1999), edited by G. Dito and D. Sternheimer, Math. Phys. Stud. **22**, Kluwer, Dordrecht, 2000. MR Zbl

[Bursztyn and Waldmann 2001] H. Bursztyn and S. Waldmann, "Algebraic Rieffel induction, formal Morita equivalence, and applications to deformation quantization", *J. Geom. Phys.* **37**:4 (2001), 307–364. MR Zbl

[Bursztyn and Waldmann 2005a] H. Bursztyn and S. Waldmann, "Completely positive inner products and strong Morita equivalence", *Pacific J. Math.* **222**:2 (2005), 201–236. MR Zbl

[Bursztyn and Waldmann 2005b] H. Bursztyn and S. Waldmann, "Hermitian star products are completely positive deformations", *Lett. Math. Phys.* **72**:2 (2005), 143–152. MR Zbl

[Chari and Pressley 1994] V. Chari and A. Pressley, *A guide to quantum groups*, Cambridge Univ. Press, 1994. MR Zbl

[De Wilde and Lecomte 1983] M. De Wilde and P. B. A. Lecomte, "Existence of star-products and of formal deformations of the Poisson Lie algebra of arbitrary symplectic manifolds", *Lett. Math. Phys.* **7**:6 (1983), 487–496. MR Zbl

[Dolgushev et al. 2002] V. A. Dolgushev, A. P. Isaev, S. L. Lyakhovich, and A. A. Sharapov, "On the Fedosov deformation quantization beyond the regular Poisson manifolds", *Nuclear Phys. B* **645**:3 (2002), 457–476. MR Zbl

[Donin 2003] J. Donin, "Classification of polarized deformation quantizations", *J. Geom. Phys.* **48**:4 (2003), 546–579. MR Zbl

[Drinfeld 1983] V. G. Drinfeld, "Constant quasiclassical solutions of the Yang–Baxter quantum equation", *Dokl. Akad. Nauk SSSR* **273**:3 (1983), 531–535. In Russian; translated in *Soviet Math. Dokl.* **28**:3 (1983), 667–671. MR Zbl

[Drinfeld 1986] V. G. Drinfeld, "Quantum groups", *Zap. Nauchn. Sem. LOMI* **155** (1986), 18–49. In Russian; translated in *J. Soviet Math.* **41**:2 (1988), 898–915. MR Zbl

[Esposito 2015] C. Esposito, *Formality theory: from Poisson structures to deformation quantization*, SpringerBriefs in Mathematical Physics **2**, Springer, Cham, Switzerland, 2015. MR Zbl

[Etingof and Schiffmann 1998] P. Etingof and O. Schiffmann, *Lectures on quantum groups*, International Press, Boston, 1998. MR Zbl

[Fedosov 1986] B. V. Fedosov, "Quantization and the index", *Dokl. Akad. Nauk SSSR* **291**:1 (1986), 82–86. In Russian; translated in *Soviet Phys. Dokl.* **31**:11 (1986), 877–878. MR Zbl

[Fedosov 1994] B. V. Fedosov, "A simple geometrical construction of deformation quantization", *J. Differential Geom.* **40**:2 (1994), 213–238. MR Zbl

[Fedosov 1996] B. Fedosov, *Deformation quantization and index theory*, Mathematical Topics **9**, Akademie Verlag, Berlin, 1996. MR Zbl

[Gerstenhaber 1964] M. Gerstenhaber, "On the deformation of rings and algebras", *Ann. of Math.* (2) **79** (1964), 59–103. MR Zbl

[Gerstenhaber 1968] M. Gerstenhaber, "On the deformation of rings and algebras, III", *Ann. of Math.* (2) **88** (1968), 1–34. MR Zbl

[Giaquinto and Zhang 1998] A. Giaquinto and J. J. Zhang, "Bialgebra actions, twists, and universal deformation formulas", *J. Pure Appl. Algebra* **128**:2 (1998), 133–151. MR Zbl

[Gutt 1983] S. Gutt, "An explicit $*$-product on the cotangent bundle of a Lie group", *Lett. Math. Phys.* **7**:3 (1983), 249–258. MR Zbl

[Halbout 2006] G. Halbout, "Formality theorem for Lie bialgebras and quantization of twists and coboundary r-matrices", *Adv. Math.* **207**:2 (2006), 617–633. MR Zbl

[Hess 1981] H. Hess, *Symplectic connections in geometric quantization and factor orderings*, Ph.D. thesis, Freie Universität Berlin, 1981.

[Karabegov 1996] A. V. Karabegov, "Deformation quantizations with separation of variables on a Kähler manifold", *Comm. Math. Phys.* **180**:3 (1996), 745–755. MR Zbl

[Karabegov 2013] A. Karabegov, "On Gammelgaard's formula for a star product with separation of variables", *Comm. Math. Phys.* **322**:1 (2013), 229–253. MR Zbl

[Kassel 1995] C. Kassel, *Quantum groups*, Graduate Texts in Mathematics **155**, Springer, New York, 1995. MR Zbl

[Kontsevich 2003] M. Kontsevich, "Deformation quantization of Poisson manifolds", *Lett. Math. Phys.* **66**:3 (2003), 157–216. MR Zbl

[Nest and Tsygan 1995] R. Nest and B. Tsygan, "Algebraic index theorem", *Comm. Math. Phys.* **172**:2 (1995), 223–262. MR Zbl

[Neumaier 2001] N. Neumaier, *Klassifikationsergebnisse in der Deformationsquantisierung*, Ph.D. thesis, Albert-Ludwigs-Universität Freiburg, 2001, available at http://tinyurl.com/neumaier.

[Reichert and Waldmann 2016] T. Reichert and S. Waldmann, "Classification of equivariant star products on symplectic manifolds", *Lett. Math. Phys.* **106**:5 (2016), 675–692. MR Zbl

[Schnitzer 2016] J. Schnitzer, *A simple algebraic construction of Drinfeld twists*, master's thesis, University of Würzburg, 2016.

[Tamarkin 1998] D. E. Tamarkin, "Another proof of M. Kontsevich formality theorem", preprint, 1998. arXiv

[Waldmann 2005] S. Waldmann, "States and representations in deformation quantization", *Rev. Math. Phys.* **17**:1 (2005), 15–75. MR Zbl

[Waldmann 2007] S. Waldmann, *Poisson-Geometrie und Deformationsquantisierung: eine Einführung*, Springer, Berlin, 2007. Zbl

Received August 11, 2016. Revised April 25, 2017.

CHIARA ESPOSITO
INSTITUT FÜR MATHEMATIK, LEHRSTUHL FÜR MATHEMATIK X
UNIVERSITÄT WÜRZBURG
D-97074 WÜRZBURG
GERMANY

chiara.esposito@mathematik.uni-wuerzburg.de

JONAS SCHNITZER
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DEGLI STUDI DI SALERNO
I-84084 FISCIANO (SA)
ITALY

jschnitzer@unisa.it

STEFAN WALDMANN
INSTITUT FÜR MATHEMATIK, LEHRSTUHL FÜR MATHEMATIK X
UNIVERSITÄT WÜRZBURG
D-97074 WÜRZBURG
GERMANY

stefan.waldmann@mathematik.uni-wuerzburg.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.359

UNIFORM STABLE RADIUS, LÊ NUMBERS AND TOPOLOGICAL TRIVIALITY FOR LINE SINGULARITIES

CHRISTOPHE EYRAL

Let $\{f_t\}$ be a family of complex polynomial functions with *line* singularities. We show that if $\{f_t\}$ has a *uniform stable radius* (for the corresponding Milnor fibrations), then the Lê numbers of the functions f_t are independent of t for all small t. A similar assertion was proved by M. Oka and D. B. O'Shea in the case of isolated singularities — a case for which the only nonzero Lê number coincides with the Milnor number.

By combining our result with a theorem of J. Fernández de Bobadilla, we conclude that a family of line singularities in $\mathbb{C}^n$, $n \geq 5$, is topologically trivial if it has a uniform stable radius.

As an important example, we show that families of weighted homogeneous line singularities have a uniform stable radius if the nearby fibres $f_t^{-1}(\eta)$, $\eta \neq 0$, are "uniformly" nonsingular with respect to the deformation parameter t.

1. Introduction

Let $(t, z) := (t, z_1, \dots, z_n)$ be linear coordinates for $\mathbb{C} \times \mathbb{C}^n$ $(n \geq 2)$, and let

$$f : (\mathbb{C} \times \mathbb{C}^n, \mathbb{C} \times \{\mathbf{0}\}) \to (\mathbb{C}, 0), \quad (t, z) \mapsto f(t, z), \tag{1-1}$$

be a polynomial function. As usual, we write $f_t(z) := f(t, z)$, and for any $\eta \in \mathbb{C}$ we denote by $V(f_t - \eta)$ the hypersurface in $\mathbb{C}^n$ defined by the equation $f_t(z) = \eta$. (Note that (1-1) implies $f_t(\mathbf{0}) = f(t, \mathbf{0}) = 0$, so that the origin $\mathbf{0} \in \mathbb{C}^n$ belongs to the hypersurface $V(f_t) = f_t^{-1}(0)$ for all $t \in \mathbb{C}$.)

The purpose of this paper is to show that if the polynomial function f defines a family $\{f_t\}$ of hypersurfaces with *line* singularities and with a *uniform stable radius* (for the corresponding Milnor fibrations), then the Lê numbers

$$\lambda^0_{f_t,z}(\mathbf{0}), \dots, \lambda^{n-1}_{f_t,z}(\mathbf{0})$$

of the polynomial functions f_t at $\mathbf{0}$ with respect to the coordinates z — which do exist in this case — are independent of t for all small t (see Theorem 4.1). In the

MSC2010: primary 14B05, 14B07, 14J17, 14J70; secondary 32S05, 32S25.

Keywords: line singularities, uniform stable radius, Lê numbers, equisingularity.

case of hypersurfaces with *isolated* singularities — a case for which the constancy of the Lê numbers means the constancy of the Milnor number — a similar assertion was proved by M. Oka [1973] and D.B. O'Shea [1983a].

By combining Theorem 4.1 with a theorem of J. Fernández de Bobadilla [2013], to the effect that a family of hypersurfaces with line singularities in $\mathbb{C}^n$, $n \geq 5$, is topologically trivial if it has constant Lê numbers, it follows that a family of hypersurfaces with line singularities in $\mathbb{C}^n$, $n \geq 5$, is topologically trivial if it has a uniform stable radius (see Corollary 4.2).

Oka [1973] and O'Shea [1983a] also proved that, if $\{f_t\}$ is a family of *isolated* hypersurface singularities such that each f_t is *weighted homogeneous* with respect to a given system of weights, then $\{f_t\}$ has a uniform stable radius. In Theorem 5.1, we show this still holds true for weighted homogeneous hypersurfaces with *line* singularities provided that the nearby fibres $V(f_t - \eta)$, $\eta \neq 0$, are "uniformly" nonsingular with respect to the deformation parameter t — that is, nonsingular in a small ball the radius of which does not depends on t. (Note that this condition always holds true for isolated singularities.) In particular, by Theorem 4.1 and Corollary 4.2, such families have constant Lê numbers, and for $n \geq 5$, they are topologically trivial.

Finally, let us observe that by combining Corollary 4.2 with a theorem of Oka [1982] — which says that a family $\{f_t\}$ of nondegenerate functions with constant Newton boundary has a uniform stable radius — we get a new proof of a theorem of J. Damon [1983] which says that if $\{f_t\}$ is a family of nondegenerate line singularities in $\mathbb{C}^n$, $n \geq 5$, with constant Newton boundary, then $\{f_t\}$ is topologically trivial.

Notation 1.1. In this paper, we are only interested in the behaviour of functions (or hypersurfaces) near the origin $\mathbf{0} \in \mathbb{C}^n$. We denote by B_ε the closed ball centred at $\mathbf{0} \in \mathbb{C}^n$ with radius $\varepsilon > 0$, and we write $\mathring{B}_\varepsilon$ and S_ε for its interior and boundary, respectively. As usual, in $\mathbb{C}$, we write D_ε and $\mathring{D}_\varepsilon$ rather than B_ε and $\mathring{B}_\varepsilon$.

2. Uniform stable radius

By [Hamm and Lê 1973, lemme (2.1.4)], we know that for each t there exists a positive number $r_t > 0$ such that for any pair $(\varepsilon_t, \varepsilon'_t)$ with $0 < \varepsilon'_t \leq \varepsilon_t \leq r_t$, there exists $\delta(\varepsilon_t, \varepsilon'_t) > 0$ such that for any nonzero complex number η with $0 < |\eta| \leq \delta(\varepsilon_t, \varepsilon'_t)$, the hypersurface $V(f_t - \eta)$ is nonsingular in $\mathring{B}_{r_t}$ and transversely intersects with the sphere $S_{\varepsilon''}$ for any ε'' with $\varepsilon'_t \leq \varepsilon'' \leq \varepsilon_t$. Any such a number r_t is called a *stable radius* for the Milnor fibration of f_t at $\mathbf{0}$ [Oka 1982, §2].

Definition 2.1 [Oka 1982, §3]. We say that the family $\{f_t\}$ has a *uniform stable radius* (we also say that $\{f_t\}$ is *uniformly stable*) if there exist $\tau > 0$ and $r > 0$ such that for any pair $(\varepsilon, \varepsilon')$ with $0 < \varepsilon' \leq \varepsilon \leq r$, there exists $\delta(\varepsilon, \varepsilon') > 0$ such that for any nonzero complex number η with $0 < |\eta| \leq \delta(\varepsilon, \varepsilon')$, the hypersurface $V(f_t - \eta)$

is nonsingular in $\mathring{B}_r$ and transversely intersects with the sphere $S_{\varepsilon''}$ for any ε'' with $\varepsilon' \leq \varepsilon'' \leq \varepsilon$ and for any t with $0 \leq |t| \leq \tau$. Any such a number r is called a *uniform stable radius* for $\{f_t\}$.

In the special case where the polynomial function f defines a family $\{f_t\}$ of *isolated* hypersurface singularities (i.e., f_t has an isolated singularity at $\mathbf{0}$ for all small t), then, by [Milnor 1968], we also know that for each t there exists $R_t > 0$ such that the hypersurface $V(f_t)$ is nonsingular in $\mathring{B}_{R_t} \setminus \{\mathbf{0}\}$ and transversely intersects the sphere S_ρ for any ρ with $0 < \rho \leq R_t$.

Definition 2.2 [Oka 1973, §2]. Suppose that f defines a family $\{f_t\}$ of *isolated* hypersurface singularities. We say that $\{f_t\}$ satisfies *condition* (A) if there exist $\nu > 0$ and $R > 0$ such that $V(f_t)$ is nonsingular in $\mathring{B}_R \setminus \{\mathbf{0}\}$ and transversely intersects the sphere S_ρ for any ρ with $0 < \rho \leq R$ and for any t with $0 \leq |t| \leq \nu$.

It is easy to see that a family $\{f_t\}$ of isolated hypersurface singularities satisfies condition (A) if and only if it has no *vanishing fold* and no *nontrivial critical arc* in the sense of [O'Shea 1983a]. Also, it is worthwhile to observe that if $\{f_t\}$ satisfies condition (A), then it has a uniform stable radius [Oka 1973; O'Shea 1983a].

3. The Oka–O'Shea theorem for isolated singularities

Throughout this section we assume that the polynomial function f defines a family $\{f_t\}$ of *isolated* hypersurface singularities.

Theorem 3.1 [Oka 1973; O'Shea 1983a]. *Suppose that f defines a family $\{f_t\}$ of isolated hypersurface singularities. If furthermore $\{f_t\}$ satisfies condition (A) or has a uniform stable radius, then it is μ-constant — that is, the Milnor number $\mu_{f_t}(\mathbf{0})$ of f_t at $\mathbf{0}$ is independent of t for all small t.*

Actually Oka showed that if $\{f_t\}$ satisfies condition (A) or if it has a uniform stable radius, then the Milnor fibrations at $\mathbf{0}$ of f_0 and f_t are isomorphic.

Lê Dũng Tráng and C. P. Ramanujam [Lê and Ramanujam 1976] showed that for $n \neq 3$ any family of isolated hypersurface singularities with constant Milnor number is topologically $\mathscr{V}$-equisingular. With the same assumption, J. G. Timourian [1977] showed that the family is actually topologically trivial. We recall that a family $\{f_t\}$ is *topologically $\mathscr{V}$-equisingular* (respectively, *topologically trivial*) if there exist open neighbourhoods $D \subseteq \mathbb{C}$ and $U \subseteq \mathbb{C}^n$ of the origins in $\mathbb{C}$ and $\mathbb{C}^n$, together with a continuous map $\varphi\colon (D \times U, D \times \{\mathbf{0}\}) \to (\mathbb{C}^n, \mathbf{0})$ such that for all sufficiently small t, there is an open neighbourhood $U_t \subseteq U$ of $\mathbf{0} \in \mathbb{C}^n$ such that the map

$$\varphi_t\colon (U_t, \mathbf{0}) \to (\varphi(\{t\} \times U_t), \mathbf{0}), \quad z \mapsto \varphi_t(z) := \varphi(t, z),$$

is a homeomorphism satisfying the relation

$$\varphi_t(V(f_0) \cap U_t) = V(f_t) \cap \varphi_t(U_t)$$

(respectively, the relation $f_0 = f_t \circ \varphi_t$ on U_t).

Note that, in general, "μ-constant" does not imply condition (A) [Oka 1989; Briançon].

Finally, observe that the Briançon–Speder famous family shows that condition (A) does not imply the Whitney conditions along the t-axis [Briançon and Speder 1975].

4. Uniformly stable families of line singularities

Setup and statement of the main result. From now on we suppose that the polynomial function f defines a family $\{f_t\}$ of hypersurfaces with *line* singularities. As in [Massey 1988, §4], by such a family we mean a family $\{f_t\}$ such that for each t small enough, the singular locus Σf_t of f_t near the origin $\mathbf{0} \in \mathbb{C}^n$ is given by the z_1-axis, and the restriction of f_t to the hyperplane $V(z_1)$ defined by $z_1 = 0$ has an isolated singularity at the origin. Then, by [Massey 1995, Remark 1.29], the partition of $V(f_t)$ given by

$$\mathscr{S}_t := \big\{V(f_t) \setminus \Sigma f_t,\, \Sigma f_t \setminus \{\mathbf{0}\},\, \{\mathbf{0}\}\big\}$$

is a *good stratification* for f_t at $\mathbf{0}$, and the hyperplane $V(z_1)$ is a *prepolar slice* for f_t at $\mathbf{0}$ with respect to $\mathscr{S}_t$ for all t small enough. In particular, combined with [Massey 1995, Proposition 1.23], this implies that the *Lê numbers*

$$\lambda^0_{f_t,\mathbf{z}}(\mathbf{0}) \quad \text{and} \quad \lambda^1_{f_t,\mathbf{z}}(\mathbf{0})$$

of f_t at $\mathbf{0}$ with respect to the coordinates $\mathbf{z}$ do exist. (For the definitions of good stratifications, prepolarity and Lê numbers, we refer the reader to [Massey 1995].) Note that for line singularities, the only possible nonzero Lê numbers are precisely $\lambda^0_{f_t,\mathbf{z}}(\mathbf{0})$ and $\lambda^1_{f_t,\mathbf{z}}(\mathbf{0})$. All the other Lê numbers $\lambda^k_{f_t,\mathbf{z}}(\mathbf{0})$ for $2 \le k \le n-1$ are defined and equal to zero; see [Massey 1995].

Here is our main observation.

Theorem 4.1. *Suppose that f defines a family $\{f_t\}$ of hypersurfaces with line singularities. If furthermore $\{f_t\}$ has a uniform stable radius, then it is $\lambda_{\mathbf{z}}$-constant — that is, the Lê numbers $\lambda^0_{f_t,\mathbf{z}}(\mathbf{0})$ and $\lambda^1_{f_t,\mathbf{z}}(\mathbf{0})$ are independent of t for all small t.*

Theorem 4.1 extends to line singularities Oka and O'Shea's Theorem 3.1 concerning isolated singularities. Indeed, for isolated singularities, the only possible nonzero Lê number is $\lambda^0_{f_t,\mathbf{z}}(\mathbf{0})$ and the latter coincides with the Milnor number $\mu_{f_t}(\mathbf{0})$.

Note that if $\{f_t\}$ is a $\lambda_{\mathbf{z}}$-constant family of line singularities in $\mathbb{C}^n$ with $n \ge 5$, then, by a theorem of D. B. Massey [1988, Theorem (5.2)], the diffeomorphism type of the Milnor fibration of f_t at $\mathbf{0}$ is independent of t for all small t. Under the same assumption, Fernández de Bobadilla [2013, Theorem 42] showed that $\{f_t\}$ is actually topologically trivial. Combining this result with our Theorem 4.1 gives the following corollary.

Corollary 4.2. *Suppose that f defines a family $\{f_t\}$ of hypersurfaces with line singularities in $\mathbb{C}^n$ with $n \geq 5$. If furthermore $\{f_t\}$ has a uniform stable radius, then it is topologically trivial.*

Application to families of nondegenerate line singularities with constant Newton boundary. Oka [1982, Corollary 1] showed that if $\{f_t\}$ is a family of hypersurface singularities — not necessary line singularities — such that for all small t the polynomial function f_t is *nondegenerate* and the *Newton boundary* of f_t at $\mathbf{0}$ with respect to the coordinates z is independent of t, then $\{f_t\}$ has a uniform stable radius. (For the definitions of nondegeneracy and Newton boundary, see [Kouchnirenko 1976; Oka 1979].) Combined with Oka's result, Corollary 4.2 provides a new proof of the following result, which is a particular case of a more general theorem of Damon.

Theorem 4.3 [Damon 1983]. *Suppose that f defines a family $\{f_t\}$ of hypersurfaces with line singularities in $\mathbb{C}^n$ with $n \geq 5$. If furthermore for any sufficiently small t the polynomial function f_t is nondegenerate and the Newton boundary of f_t at $\mathbf{0}$ with respect to the coordinates z is independent of t, then the family $\{f_t\}$ is topologically trivial.*

Proof of Theorem 4.1. Consider the map $\Phi\colon \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^2$ defined by

$$(t, z) \mapsto \Phi(t, z) := (f(t, z), t),$$

and pick positive numbers τ and r which satisfy the condition of Definition 2.1. Then, in particular, the following property holds:

($\mathscr{P}$) For any ε with $0 < \varepsilon < r$, there exists $\delta(\varepsilon) > 0$ such that for any t with $0 \leq |t| \leq \tau$ and for any η with $0 < |\eta| \leq \delta(\varepsilon)$, the hypersurface $V(f_t - \eta)$ is nonsingular in $\mathring{B}_r$ and transversely intersects the sphere S_ε.

This property implies that the critical set $\Sigma\Phi$ of Φ does not intersect the set

$$U(\mathring{B}_r) := (\mathring{D}_\tau \times \mathring{B}_r) \cap \Phi^{-1}((\mathring{D}_{\delta(\varepsilon)} \setminus \{\mathbf{0}\}) \times \mathring{D}_\tau).$$

Indeed, suppose there is a point $(t_0, z_0) \in \Sigma\Phi \cap U(\mathring{B}_r)$. Then $z_0 \in \Sigma(f_{t_0} - f_{t_0}(z_0))$. But this is not possible, since by ($\mathscr{P}$) the hypersurface $V(f_{t_0} - f_{t_0}(z_0))$ is smooth. (We recall that a complex variety can never be a smooth manifold throughout a neighbourhood of a critical point; see [Milnor 1968, §2].)

It also follows from property ($\mathscr{P}$) that the map

$$\Phi|_{U(S_\varepsilon)}\colon U(S_\varepsilon) \to (\mathring{D}_{\delta(\varepsilon)} \setminus \{\mathbf{0}\}) \times \mathring{D}_\tau$$

(restriction of Φ to $U(S_\varepsilon) := (\mathring{D}_\tau \times S_\varepsilon) \cap \Phi^{-1}((\mathring{D}_{\delta(\varepsilon)} \setminus \{\mathbf{0}\}) \times \mathring{D}_\tau)$) is a submersion. Indeed, as $\Sigma\Phi \cap U(\mathring{B}_r) = \varnothing$ and $U(\mathring{B}_r)$ is an open subset of $\mathbb{C} \times \mathbb{C}^n$, the map

$$\Phi|_{U(\mathring{B}_r)}\colon U(\mathring{B}_r) \to (\mathring{D}_{\delta(\varepsilon)} \setminus \{\mathbf{0}\}) \times \mathring{D}_\tau$$

is a submersion. Thus, to show that $\Phi|_{U(S_\varepsilon)}$ is a submersion, it suffices to observe that the inclusion $U(S_\varepsilon) \hookrightarrow U(\mathring{B}_r)$ is transverse to the submanifold $\Phi|^{-1}_{U(\mathring{B}_r)}(f(t,z),t)$ for any point $(t,z) \in U(S_\varepsilon)$ — or equivalently that the submanifolds

$$\Phi|^{-1}_{U(\mathring{B}_r)}(f(t,z),t) \quad \text{and} \quad (\{t\} \times S_\varepsilon) \cap U(\mathring{B}_r)$$

are transverse to each other. This is exactly the content of $(\mathscr{P})$.

Now, as $\Phi|_{U(S_\varepsilon)}$ is also a proper map, a result of Massey and D. Siersma [1992, Proposition 1.10] shows that the Milnor number of a generic hyperplane slice of f_t at a point on Σf_t sufficiently close to the origin (which coincides with the Lê number $\lambda^1_{f_t,z}(\mathbf{0})$ for line singularities; see [Lê 1980; Massey 1988]) is independent of t for all small t.

Finally, since the family $\{f_t\}$ has a uniform stable radius — the full strength of this assumption is used here — it follows from [Oka 1982, Lemma 2] that the diffeomorphism type of the Milnor fibration of f_t at the origin is independent of t for all small t. In particular, the reduced Euler characteristic $\tilde{\chi}(F_{f_t,\mathbf{0}})$ of the Milnor fibre $F_{f_t,\mathbf{0}}$ of f_t at $\mathbf{0}$, which by [Massey 1995, Theorem 3.3] equals

$$(-1)^{n-1}\lambda^0_{f_t,z}(\mathbf{0}) + (-1)^{n-2}\lambda^1_{f_t,z}(\mathbf{0}),$$

is independent of t for all small t. The constancy of $\lambda^0_{f_t,z}(\mathbf{0})$ now follows from that of $\lambda^1_{f_t,z}(\mathbf{0})$.

5. Uniform stable radius and weighted homogeneous line singularities

By a result of Oka [1973] and O'Shea [1983a], we know that if $\{f_t\}$ is a family of *isolated* hypersurface singularities such that each f_t is *weighted homogeneous* with respect to a given system of weights, then $\{f_t\}$ satisfies condition (A), and hence, is uniformly stable. Our next observation says this still holds true for weighted homogeneous *line* singularities provided that the nearby fibres $V(f_t - \eta)$, $\eta \neq 0$, of the functions f_t are "uniformly" nonsingular with respect to the deformation parameter t — that is, nonsingular in a small ball the radius of which does *not* depends on t. (We recall that by [Hamm and Lê 1973] the nearby fibres are "individually" nonsingular — that is, nonsingular in a small ball the radius of which depends on t.)

Theorem 5.1. *Suppose that f defines a family $\{f_t\}$ of hypersurfaces with line singularities such that each f_t is weighted homogeneous with respect to a given system of weights $\boldsymbol{w} = (w_1, \dots, w_n)$ on the variables $(z_1, \dots, z_n)$, with $w_i \in \mathbb{N} \setminus \{0\}$. Also, assume that the nearby fibres $V(f_t - \eta)$, $\eta \neq 0$, of the functions f_t are uniformly nonsingular with respect to the deformation parameter t — that is, there exist positive numbers τ, r, δ such that for any $0 < |\eta| \leq \delta$ and $0 \leq |t| \leq \tau$, the hypersurface $V(f_t - \eta)$ is nonsingular in $\mathring{B}_r$. Under these assumptions, the family*

$\{f_t\}$ has a uniform stable radius. (*In particular*, *$\{f_t\}$ is λ_z-constant*, *and for $n \geq 5$, it is topologically trivial.*)

Proof. The argument is similar to those used in [Oka 1973; O'Shea 1983a]. Suppose that the family $\{f_t\}$ does not have a uniform stable radius. Then, as the nearby fibres of the functions f_t are uniformly nonsingular with respect to the deformation parameter t, for all $\tau > 0$ and all $r > 0$ small enough, there exist $0 < \varepsilon' \leq \varepsilon \leq r$ such that for all sufficiently small $\delta > 0$ there exist η_δ, ε_δ and t_δ, with $0 < |\eta_\delta| \leq \delta$, $\varepsilon' \leq \varepsilon_\delta \leq \varepsilon$ and $|t_\delta| \leq \tau$, such that $V(f_{t_\delta} - \eta_\delta)$ is nonsingular in $\mathring{B}_r$ and does not transversely intersect the sphere S_{ε_δ}. It follows that there is a point $z_\delta \in V(f_{t_\delta} - \eta_\delta) \cap S_{\varepsilon_\delta}$ which is a critical point of the restriction to $V(f_{t_\delta} - \eta_\delta) \cap B_r$ of the squared distance function:

$$z \in V(f_{t_\delta} - \eta_\delta) \cap B_r \mapsto \|z\|^2 = \sum_{1 \leq i \leq n} |z_i|^2.$$

In other words, the point (t_δ, z_δ) lies in the intersection of $D_\tau \times (B_\varepsilon \setminus \mathring{B}_{\varepsilon'})$ with the *real* algebraic set C consisting of the points (t, z) such that

$$\left(\frac{\partial f_t}{\partial z_1}(z), \dots, \frac{\partial f_t}{\partial z_n}(z)\right) = \lambda \bar{z} \tag{5-1}$$

for some $\lambda \in \mathbb{C} \setminus \{0\}$, where $\bar{z} := (\bar{z}_1, \dots, \bar{z}_n)$ and $\bar{z}_i$ denotes the complex conjugate of z_i (see e.g., [O'Shea 1983b, Lemma 1]). Let $C_{\tau,r} := C \cap (D_\tau \times (B_\varepsilon \setminus \mathring{B}_{\varepsilon'}))$. Take $\delta := \delta(m) := 1/m$ (where $m \in \mathbb{N} \setminus \{0\}$ is sufficiently large), and consider the corresponding sequence of points $(t_{\delta(m)}, z_{\delta(m)})$ in $C_{\tau,r}$. As $C_{\tau,r}$ is compact, taking a subsequence if necessary, we may assume that $(t_{\delta(m)}, z_{\delta(m)})$ converges to a point $(t_{\tau,r}, z_{\tau,r}) \in C_{\tau,r}$, and hence $\eta_{\delta(m)} := f(t_{\delta(m)}, z_{\delta(m)})$ tends to $f(t_{\tau,r}, z_{\tau,r})$ as $m \to \infty$. Since $0 < |\eta_{\delta(m)}| \leq \delta(m) = 1/m \to 0$ as $m \to \infty$, we have $f(t_{\tau,r}, z_{\tau,r}) = 0$. Thus $(t_{\tau,r}, z_{\tau,r}) \in V(f) \cap C_{\tau,r}$.

Now, since $f_{t_{\tau,r}}$ is weighted homogeneous with respect to the weights $\boldsymbol{w} = (w_1, \dots, w_n)$, the *Euler identity* implies the following contradiction:

$$d_{\boldsymbol{w}} \cdot \underbrace{f_{t_{\tau,r}}(z_{\tau,r})}_{=0} \overset{\text{Euler}}{=} \sum_{1 \leq i \leq n} w_i (z_{\tau,r})_i \frac{\partial f_{t_{\tau,r}}}{\partial z_i}(z_{\tau,r}) \overset{(5\text{-}1)}{=} \lambda \sum_{1 \leq i \leq n} w_i |(z_{\tau,r})_i|^2 \neq 0,$$

where $d_{\boldsymbol{w}}$ is the weighted degree of $f_{t_{\tau,r}}$ with respect to the weights $\boldsymbol{w}$ and $(z_{\tau,r})_i$ is the i-th component of $z_{\tau,r}$. □

Remark 5.2. Actually, the proof shows that if f defines a family $\{f_t\}$ of hypersurfaces — not necessarily with line singularities — such that each f_t is weighted homogeneous with respect to a given system of weights $\boldsymbol{w}$, and if furthermore, the nearby fibres $V(f_t - \eta)$, $\eta \neq 0$, of the functions f_t are uniformly nonsingular with respect to the deformation parameter t, then the family $\{f_t\}$ has a uniform stable radius.

References

[Briançon] J. Briançon, "Le théorème de Kouchnirenko", unpublished lecture notes.

[Briançon and Speder 1975] J. Briançon and J.-P. Speder, "La trivialité topologique n'implique pas les conditions de Whitney", *C. R. Acad. Sci. Paris Sér. A-B* **280**:6 (1975), 365–367. MR Zbl

[Damon 1983] J. Damon, "Newton filtrations, monomial algebras and nonisolated and equivariant singularities", pp. 267–276 in *Singularities, I* (Arcata, CA, 1981), edited by P. Orlik, Proc. Sympos. Pure Math. **40**, American Mathematical Society, Providence, RI, 1983. MR Zbl

[Fernández de Bobadilla 2013] J. Fernández de Bobadilla, "Topological equisingularity of hypersurfaces with 1-dimensional critical set", *Adv. Math.* **248** (2013), 1199–1253. MR Zbl

[Hamm and Lê 1973] H. A. Hamm and Lê D. T., "Un théorème de Zariski du type de Lefschetz", *Ann. Sci. École Norm. Sup.* (4) **6** (1973), 317–355. MR Zbl

[Kouchnirenko 1976] A. G. Kouchnirenko, "Polyèdres de Newton et nombres de Milnor", *Invent. Math.* **32**:1 (1976), 1–31. MR Zbl

[Lê 1980] Lê D. T., "Ensembles analytiques complexes avec lieu singulier de dimension un (d'après I. N. Iomdine)", pp. 87–95 in *Séminaire sur les singularités* (Paris, 1976/1977), edited by Lê D. T., Publ. Math. Univ. Paris VII **7**, Univ. Paris VII, 1980. MR Zbl

[Lê and Ramanujam 1976] Lê D. T. and C. P. Ramanujam, "The invariance of Milnor's number implies the invariance of the topological type", *Amer. J. Math.* **98**:1 (1976), 67–78. MR Zbl

[Massey 1988] D. B. Massey, "The Lê–Ramanujam problem for hypersurfaces with one-dimensional singular sets", *Math. Ann.* **282**:1 (1988), 33–49. MR Zbl

[Massey 1995] D. B. Massey, *Lê cycles and hypersurface singularities*, Lecture Notes in Mathematics **1615**, Springer, Berlin, 1995. MR Zbl

[Massey and Siersma 1992] D. B. Massey and D. Siersma, "Deformation of polar methods", *Ann. Inst. Fourier* (*Grenoble*) **42**:4 (1992), 737–778. MR Zbl

[Milnor 1968] J. Milnor, *Singular points of complex hypersurfaces*, Annals of Mathematics Studies **61**, Princeton Univ. Press, 1968. MR Zbl

[Oka 1973] M. Oka, "Deformation of Milnor fiberings", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **20** (1973), 397–400. Correction in **27**:2 (1980), 463–464. MR Zbl

[Oka 1979] M. Oka, "On the bifurcation of the multiplicity and topology of the Newton boundary", *J. Math. Soc. Japan* **31**:3 (1979), 435–450. MR Zbl

[Oka 1982] M. Oka, "On the topology of the Newton boundary, III", *J. Math. Soc. Japan* **34**:3 (1982), 541–549. MR Zbl

[Oka 1989] M. Oka, "On the weak simultaneous resolution of a negligible truncation of the Newton boundary", pp. 199–210 in *Singularities* (Iowa City, 1986), edited by R. Randell, Contemp. Math. **90**, American Mathematical Society, Providence, RI, 1989. MR Zbl

[O'Shea 1983a] D. B. O'Shea, "Finite jumps in Milnor number imply vanishing folds", *Proc. Amer. Math. Soc.* **87**:1 (1983), 15–18. MR Zbl

[O'Shea 1983b] D. B. O'Shea, "Vanishing folds in families of singularities", pp. 293–303 in *Singularities, II* (Arcata, CA, 1981), edited by P. Orlik, Proc. Sympos. Pure Math. **40**, American Mathematical Society, Providence, RI, 1983. MR Zbl

[Timourian 1977] J. G. Timourian, "The invariance of Milnor's number implies topological triviality", *Amer. J. Math.* **99**:2 (1977), 437–446. MR Zbl

Received April 27, 2017.

CHRISTOPHE EYRAL
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
00-656 WARSAW
POLAND
eyralchr@yahoo.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.369

ROST INVARIANT OF THE CENTER, REVISITED

SKIP GARIBALDI AND ALEXANDER S. MERKURJEV

The Rost invariant of the Galois cohomology of a simple simply connected algebraic group over a field F is defined regardless of the characteristic of F, but certain formulas for it have only been known under a hypothesis on the characteristic. We improve on those formulas by removing the hypothesis on the characteristic and removing an ad hoc pairing that appeared in those formulas. As a preliminary step of independent interest, we also extend the classification of invariants of quasitrivial tori to all fields.

1. Introduction

Cohomological invariants provide an important tool to distinguish elements of Galois cohomology groups such as $H^1(F, G)$ where G is a semisimple algebraic group. In the case where G is simple and simply connected there are no nonconstant invariants with values in $H^d(*, \mathbb{Q}/\mathbb{Z}(d-1))$ for $d < 3$. For $d = 3$, modulo constants the group of invariants $H^1(*, G) \to H^3(*, \mathbb{Q}/\mathbb{Z}(2))$ is finite cyclic with a canonical generator known as the Rost invariant and denoted by r_G; this was shown by Markus Rost in the 1990s and full details can be found in [Garibaldi et al. 2003]. Rost's theorem raised the questions: How do we calculate the Rost invariant of a class in $H^1(F, G)$? What is a formula for it?

At least for G of inner type A_n there is an obvious candidate for r_G, which is certainly equal to mr_G for some m relatively prime to $n+1$. The papers [Merkurjev et al. 2002; Garibaldi and Quéguiner-Mathieu 2007] studied the composition

$$H^1(F, C) \to H^1(F, G) \xrightarrow{r_G} H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \tag{1.1}$$

for C the center of G, and under some assumptions on char(F), computed the composition in terms of the value of m for type A. Eventually the value of m was determined in [Gille and Quéguiner-Mathieu 2011]. The main result of this paper is Theorem 1.2, which gives a formula for (1.1) that does not depend on the type of G nor on char(F). This improves on the results of [Merkurjev et al. 2002; Garibaldi and Quéguiner-Mathieu 2007] by removing the hypothesis on the characteristic

Merkurjev has been supported by the NSF grant DMS #1160206.
MSC2010: primary 20G15; secondary 11E72.
Keywords: Rost invariant, cohomological invariant, cup product, Tits class.

and avoiding the ad hoc type-by-type arguments used in those papers. We do rely on [Gille and Quéguiner-Mathieu 2011] for the computation of m for type A, but nothing more.

The strategy is to (1) extend the determination of invariants of quasitrivial tori from [Merkurjev et al. 2002] to all fields (see Theorem 3.7), (2) to follow the general outline of [Garibaldi and Quéguiner-Mathieu 2007] to reduce to the case of type A, and (3) to avoid the ad hoc formulas used in previous work by giving a formula independent of the Killing–Cartan type of G.

Specifically, there is a canonically defined element $t_G^\circ \in H^2(F, C^\circ)$, where C° denotes the dual multiplicative group scheme of C in a sense defined below, and a natural cup product $H^1(F, C) \otimes H^2(F, C^\circ) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2))$. We prove:

Theorem 1.2. *Let G be a semisimple and simply connected algebraic group over a field F, and $C \subset G$ be the center of G. Let t_G° be the image of the Tits class t_G under $\hat{\rho}^* : H^2(F, C) \to H^2(F, C^\circ)$. Then the diagram*

$$\begin{tikzcd} H^1(F, C) \arrow[r, "i^*"] \arrow[dr, "-t_G^\circ \cup"'] & H^1(F, G) \arrow[d, "r_G"] \\ & H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \end{tikzcd}$$

commutes, where the cup product map is the one defined in (2.9).

The map $\hat{\rho}^*$ is deduced from a natural map ρ defined in terms of the root system, see Section 5C.

Theorem 1.2 gives a general statement, which we state precisely in Theorem 6.4, for all invariants $H^1(*, G) \to H^3(*, \mathbb{Q}/\mathbb{Z}(2))$.

2. Cohomology of groups of multiplicative type

Let F be a field and M a group scheme of multiplicative type over F. Then M is uniquely determined by the Galois module M^* of characters over F_{sep}. In particular, we have

$$M(F_{\mathrm{sep}}) = \mathrm{Hom}(M^*, F_{\mathrm{sep}}^\times).$$

If M is a torus T, then T^* is a Galois lattice and we set $T_* = \mathrm{Hom}(T^*, \mathbb{Z})$. We have

$$T(F_{\mathrm{sep}}) = T_* \otimes F_{\mathrm{sep}}^\times. \tag{2.1}$$

If M is a finite group scheme C of multiplicative type, set $C_* := \mathrm{Hom}(C^*, \mathbb{Q}/\mathbb{Z})$, so we have a perfect pairing of Galois modules

$$C_* \otimes C^* \to \mathbb{Q}/\mathbb{Z}. \tag{2.2}$$

Write C° for the group of multiplicative type over F with the character module C_*. We call C° the group *dual to C*.

Example 2.3. We write μ_n for the sub-group-scheme of $\mathbb{G}_m$ of n-th roots of unity. The restriction of the natural generator of $\mathbb{G}_m^*$ (the identity $\mathbb{G}_m \to \mathbb{G}_m$) generates μ_n^* and thereby identifies μ_n^* with $\mathbb{Z}/n\mathbb{Z}$. Thus $\mu_n^* = \mathbb{Z}/n\mathbb{Z}$ via the pairing (2.2), hence $\mu_n^\circ = \mu_n$.

The change-of-sites map $\alpha : \mathrm{Spec}(F)_{\mathrm{fppf}} \to \mathrm{Spec}(F)_{\text{ét}}$ yields a functor

$$\alpha_* : \mathrm{Sh}_{\mathrm{fppf}}(F) \to \mathrm{Sh}_{\text{ét}}(F)$$

between the categories of sheaves over F and an exact functor

$$R\alpha_* : D^+ \mathrm{Sh}_{\mathrm{fppf}}(F) \to D^+ \mathrm{Sh}_{\text{ét}}(F)$$

between derived categories.

Every group M of multiplicative type can be viewed as a sheaf of abelian groups either in the étale or fppf topology. We have $\alpha_*(M) = M$ for every group M of multiplicative type. If M is smooth, we have $R^i\alpha_*(M) = 0$ for $i > 0$ by [Milne 1980, proof of Theorem 3.9]. It follows that $R\alpha_*(M) = M$, hence

$$H^i_{\text{ét}}(F, M) = H^i_{\text{ét}}(F, R\alpha_*(M)) = H^i_{\mathrm{fppf}}(F, M), \quad \text{for } M \text{ smooth.} \tag{2.4}$$

If

$$1 \to C \to T \to S \to 1$$

is an exact sequence of algebraic groups with C a finite group of multiplicative type and T and S tori, this sequence is exact in the fppf-topology but not in the étale topology (unless C is smooth). Applying $R\alpha_*$ to the exact triangle

$$C \to T \to S \to C[1]$$

in $D^+ \mathrm{Sh}_{\mathrm{fppf}}(F)$, we get an exact triangle,

$$R\alpha_*(C) \to T(F_{\mathrm{sep}}) \to S(F_{\mathrm{sep}}) \to R\alpha_*(C)[1],$$

in $D^+ \mathrm{Sh}_{\text{ét}}(F)$ since $R\alpha_*(T) = T(F_{\mathrm{sep}})$ and $R\alpha_*(S) = S(F_{\mathrm{sep}})$. In other words,

$$R\alpha_*(C) = \mathrm{cone}(T(F_{\mathrm{sep}}) \to S(F_{\mathrm{sep}}))[-1]. \tag{2.5}$$

Recall that $\mathbb{Z}(1)$ is the complex in $D^+ \mathrm{Sh}_{\text{ét}}(F)$ with only one nonzero term $F_{\mathrm{sep}}^\times$ placed in degree 1, i.e., $\mathbb{Z}(1) = F_{\mathrm{sep}}^\times[-1]$. Set

$$C_*(1) := C_* \overset{L}{\otimes} \mathbb{Z}(1), \qquad C^*(1) := C^* \overset{L}{\otimes} \mathbb{Z}(1),$$

where the derived tensor product is taken in the derived category $D^+ \mathrm{Sh}_{\text{ét}}(F)$. If T is an algebraic torus, we write

$$T_*(1) := T_* \overset{L}{\otimes} \mathbb{Z}(1) = T_* \otimes \mathbb{Z}(1) = T(F_{\mathrm{sep}})[-1].$$

Tensoring the exact sequence

$$0 \to T_* \to S_* \to C_* \to 0$$

with $\mathbb{Z}(1)$ and using (2.1), we get an exact triangle

$$C_*(1) \to T(F_{\text{sep}}) \to S(F_{\text{sep}}) \to C_*(1)[1].$$

It follows from (2.5) that

$$C_*(1) = R\alpha_*(C)$$

and therefore,

$$H^i_{\text{fppf}}(F, C) = H^i_{\text{ét}}(F, R\alpha_*(C)) = H^i_{\text{ét}}(F, C_*(1)).$$

Recall that we also have

$$H^i_{\text{fppf}}(F, T) = H^i_{\text{ét}}(F, T) = H^{i+1}_{\text{ét}}(F, T_*(1)).$$

Remark 2.6. There is a canonical isomorphism (see [Merkurjev 2016, §4c])

$$C_*(1) \simeq C(F_{\text{sep}}) \oplus (C_* \otimes F^{\times}_{\text{sep}})[-1].$$

The second term in the direct sum vanishes if $\operatorname{char}(F)$ does not divide the order of C_* or if F is perfect.

Notation 2.7. To simplify notation we will write $H^i(F, C)$ for $H^i_{\text{ét}}(F, C_*(1)) = H^i_{\text{fppf}}(F, C)$ and $H^i(F, C^\circ)$ for $H^i_{\text{ét}}(F, C^*(1)) = H^i_{\text{fppf}}(F, C^\circ)$.

Every C-torsor E over F has a class $c(E) \in H^1(F, C)$.

Example 2.8. Taking colimits of the connecting homomorphism arising from the sequences $1 \to \mathbb{G}_m \to \mathrm{GL}_d \to \mathrm{PGL}_d \to 1$ or $1 \to \mu_d \to \mathrm{SL}_d \to \mathrm{PGL}_d \to 1$ — which are exact in the fppf topology — gives isomorphisms $H^2(K, \mathbb{G}_m) \simeq \mathrm{Br}(K)$ and $H^2(K, \mu_n) \simeq {}_n\mathrm{Br}(K)$ as in [Gille and Szamuely 2006, 4.4.5][1], which we use.

In view of (2.4) and Notation 2.7, we work in the derived category of étale sheaves as in, for example, [Freitag and Kiehl 1988, Appendix A.II]. We use the motivic complex $\mathbb{Z}(2)$ of étale sheaves over F defined in [Lichtenbaum 1987; 1990]. Set

$$\mathbb{Q}/\mathbb{Z}(2) := \mathbb{Q}/\mathbb{Z} \overset{L}{\otimes} \mathbb{Z}(2).$$

The complex $\mathbb{Q}/\mathbb{Z}(2)$ is the direct sum of two complexes. The first complex is given by the locally constant étale sheaf (placed in degree 0) the colimit over n prime to $\operatorname{char}(F)$ of the Galois modules $\mu_n^{\otimes 2} := \mu_n \otimes \mu_n$. The second complex is nontrivial only in the case $p = \operatorname{char}(F) > 0$ and it is defined as

$$\operatorname*{colim}_n W_n \Omega^2_{\log}[-2]$$

[1]This reference assumes $\operatorname{char}(F)$ does not divide n, since it uses H^1 to denote Galois cohomology. With our notation, their arguments go through with no change.

with $W_n\Omega^2_{\log}$ the sheaf of *logarithmic de Rham–Witt differentials* (see [Kahn 1996]).

Note that
$$H^i(F, \mathbb{Q}/\mathbb{Z}(2)) \simeq H^{i+1}(F, \mathbb{Z}(2))$$
for $i \geq 3$.

Tensoring (2.2) with $\mathbb{Z}(2)$, we get the pairings

$$\begin{gathered} C_*(1) \overset{L}{\otimes} C^*(1) \to \mathbb{Q}/\mathbb{Z}(2) \quad \text{and} \\ H^i(F, C) \otimes H^j(F, C^\circ) \to H^{i+j}(F, \mathbb{Q}/\mathbb{Z}(2)). \end{gathered} \tag{2.9}$$

If S is a torus over F, we have $S_*(1) = S_* \otimes \mathbb{G}_m[-1] = S[-1]$ and the pairings

$$\begin{gathered} S_* \otimes S^* \to \mathbb{Z}, \quad S_*(1) \overset{L}{\otimes} S^*(1) \to \mathbb{Z}(2) \quad \text{and} \\ H^i(F, S) \otimes H^j(F, S^\circ) \to H^{i+j+2}(F, \mathbb{Z}(2)) = H^{i+j+1}(F, \mathbb{Q}/\mathbb{Z}(2)) \end{gathered} \tag{2.10}$$

if $i + j \geq 2$.

Let
$$1 \to C \to T \to S \to 1$$
be an exact sequence with T and S tori and C finite. Dualizing we get an exact sequence of dual groups

$$1 \to C^\circ \to S^\circ \to T^\circ \to 1. \tag{2.11}$$

We have the homomorphisms
$$\varphi : S(F) \to H^1(F, C), \qquad \psi : H^2(F, C^\circ) \to H^2(F, S^\circ).$$

Proposition 2.12. *For every $a \in S(F)$ and $b \in H^2(F, C^\circ)$, we have $\varphi(a) \cup b = a \cup \psi(b)$ in $H^3(F, \mathbb{Q}/\mathbb{Z}(2))$. Here the cup products are taken with respect to the pairings* (2.9) *and* (2.10) *respectively.*

Proof. The pairing $S_* \otimes S^* \to \mathbb{Z}$ extends uniquely to a pairing $S_* \otimes T^* \to \mathbb{Q}$. We have then a morphism of exact triangles

$$\begin{tikzcd} S_*(1)\overset{L}{\otimes}S^*(1) \arrow[r] \arrow[d] & S_*(1)\overset{L}{\otimes}T^*(1) \arrow[r] \arrow[d] & S_*(1)\overset{L}{\otimes}C^*(1) \arrow[r] \arrow[d, "s"] & S_*(1)\overset{L}{\otimes}S^*(1)[1] \arrow[d] \\ \mathbb{Z}(2) \arrow[r] & \mathbb{Q}(2) \arrow[r] & \mathbb{Q}/\mathbb{Z}(2) \arrow[r] & \mathbb{Z}(2)[1] \end{tikzcd}$$

and a commutative diagram

$$\begin{tikzcd} H^1(F, S_*(1)) \otimes H^2(F, C^*(1)) \arrow[r] \arrow[d] & H^1(F, S_*(1)) \otimes H^2(F, S^*(1)[1]) \arrow[d] \\ H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \arrow[r] & H^3(F, \mathbb{Z}(2)[1]) \end{tikzcd}$$

and therefore, a commutative diagram:

$$\begin{array}{ccc} H^0(F,S)\otimes H^2(F,C^\circ) & \longrightarrow & H^0(F,S)\otimes H^2(F,S^\circ) \\ \big\downarrow & & \big\downarrow \\ H^3(F,\mathbb{Q}/\mathbb{Z}(2)) & = & H^3(F,\mathbb{Q}/\mathbb{Z}(2)) \end{array}$$

On the other hand, the composition $S_*(1)\overset{L}{\otimes}C^*(1) \to C_*(1)\overset{L}{\otimes}C^*(1) \to \mathbb{Q}/\mathbb{Z}(2)$ coincides with s. Therefore, we have a commutative diagram

$$\begin{array}{ccc} H^1(F,S_*(1))\otimes H^2(F,C^*(1)) & \longrightarrow & H^1(F,C_*(1))\otimes H^2(F,C^*(1)) \\ \big\downarrow & & \big\downarrow \\ H^3(F,\mathbb{Q}/\mathbb{Z}(2)) & = & H^3(F,\mathbb{Q}/\mathbb{Z}(2)) \end{array}$$

and therefore, a diagram:

$$\begin{array}{ccc} H^0(F,S)\otimes H^2(F,C^\circ) & \longrightarrow & H^1(F,C)\otimes H^2(F,C^\circ) \\ \big\downarrow & & \big\downarrow \\ H^3(F,\mathbb{Q}/\mathbb{Z}(2)) & = & H^3(F,\mathbb{Q}/\mathbb{Z}(2)) \end{array}$$

The result follows. $\square$

Remark 2.13. We have used that the diagram

$$\begin{array}{ccc} H^i(A[a])\otimes H^j(B[b]) & \longrightarrow & H^{i+j}(A[a]\otimes B[b]) \\ \| & & \| \\ H^{i+a}(A)\otimes H^{j+b}(B) & \longrightarrow & H^{i+j+a+b}(A\otimes B) \end{array}$$

is $(-1)^{ib}$-commutative for all complexes A and B.

Let A be an étale algebra over F and C a finite group scheme of multiplicative type over A. Then $C' := R_{A/F}(C)$ is a finite group of multiplicative type over F. Moreover, $C'^\circ \simeq R_{A/F}(C^\circ)$ and there are canonical isomorphisms

$$\iota : H^i(A,C) \xrightarrow{\sim} H^i(F,C') \quad \text{and} \quad \iota^\circ : H^i(A,C^\circ) \xrightarrow{\sim} H^i(F,C'^\circ).$$

Lemma 2.14. *We have* $\iota(x)\cup\iota^\circ(y) = N_{A/F}(x\cup y)$ *in* $H^{i+j}(F,\mathbb{Q}/\mathbb{Z}(2))$ *for every* $x\in H^i(A,C)$ *and* $y\in H^j(A,C^\circ)$.

Proof. The group scheme C'_A is naturally isomorphic to the product $C_1\times C_2\times\cdots\times C_s$ of group schemes over A with $C_1 = C$. Let $\pi : C'_A \to C$ be the natural projection.

Similarly, $C'^{\circ} \simeq C_1^{\circ} \times C_2^{\circ} \times \cdots \times C_s^{\circ}$. Write $\varepsilon : C^{\circ} \to {C'_A}^{\circ}$ for the natural embedding. Then the inverse of ι coincides with the composition

$$H^i(F, C') \xrightarrow{\text{res}} H^i(A, C'_A) \xrightarrow{\pi^*} H^i(A, C)$$

and ι° coincides with the composition

$$H^i(A, C^{\circ}) \xrightarrow{\varepsilon^*} H^i(A, C'^{\circ}_A) \xrightarrow{N_{A/F}} H^i(F, C'^{\circ}).$$

Since $\pi^*(\iota(x)) = x$, we have $\text{res}(\iota(x)) = (x, x_2, \ldots, x_s)$ for some x_i. On the other hand, $\varepsilon^*(y) = (y, 0, \ldots, 0)$, hence

$$\text{res}(\iota(x)) \cup \varepsilon^*(y) = x \cup y. \tag{2.15}$$

Finally,

$$\begin{aligned}\iota(x) \cup \iota^{\circ}(y) &= \iota(x) \cup N_{A/F}(\varepsilon^*(y)) \\ &= N_{A/F}(\text{res}(\iota(x)) \cup \varepsilon^*(y)) \quad \text{by the projection formula} \\ &= N_{A/F}(x \cup y) \quad \text{by (2.15).}\end{aligned}$$ □

Lemma 2.16 (projection formula). *Let $f : C \to C'$ be a homomorphism of finite group schemes of multiplicative type. For $a \in H^m(F, C)$, the diagram*

$$\begin{array}{ccc} H^k(F, C'^{\circ}) & \xrightarrow{\cup f_*(a)} & H^{k+m}(F, \mathbb{Q}/\mathbb{Z}(2)) \\ {\scriptstyle f^*}\downarrow & & \| \\ H^k(F, C^{\circ}) & \xrightarrow{\cup a} & H^{k+m}(F, \mathbb{Q}/\mathbb{Z}(2)) \end{array}$$

commutes.

Proof. The pairings used in the diagram are induced by the pairings $C^* \otimes C_* \to \mathbb{Q}/\mathbb{Z}$ and $C'^* \otimes C'_* \to \mathbb{Q}/\mathbb{Z}$. The (obvious) projection formula for these pairings reads $\langle f^*(x), y \rangle = \langle x, f_*(y) \rangle$ for $x \in C'^*$ and $y \in C_*$. □

3. Invariants of quasitrivial tori

3A. ***Cohomological invariants.*** For a field F write $H^j(F)$ for the cohomology group $H^j(F, \mathbb{Q}/\mathbb{Z}(j-1))$, where $j \geq 1$ (see [Garibaldi et al. 2003]). In particular, $H^1(F)$ is the character group of continuous homomorphisms $\Gamma_F \to \mathbb{Q}/\mathbb{Z}$ and $H^2(F)$ is the Brauer group $\text{Br}(F)$.

The assignment $K \mapsto H^j(K)$ is functorial with respect to arbitrary field extensions. If K'/K is a finite separable field extension, we have a well-defined *norm map* $N_{K'/K} : H^j(K') \to H^j(K)$.

The graded group $H^*(F)$ is a (left) module over the Milnor ring $K_*(F)$.

Definition 3.1. Let $\mathcal{A}$ be a functor from the category of field extensions of F to pointed sets. A degree d *cohomological invariant of* $\mathcal{A}$ is a collection of maps of pointed sets

$$\iota_K : \mathcal{A}(K) \to H^d(K)$$

for all field extensions K/F, functorial in K. The degree d cohomological invariants of $\mathcal{G}$ form an abelian group denoted by $\operatorname{Inv}^d(\mathcal{A})$. If L/F is a field extension, we have a *restriction homomorphism*

$$\operatorname{Inv}^d(\mathcal{A}) \to \operatorname{Inv}^d(\mathcal{A}_L),$$

where $\mathcal{G}_L$ is the restriction of $\mathcal{G}$ to the category of field extensions of L.

If the functor $\mathcal{A}$ factors through the category of groups, we further consider the subgroup $\operatorname{Inv}^d_h(\mathcal{A})$ of $\operatorname{Inv}^d(\mathcal{A})$ consisting of those invariants ι such that ι_K is a group homomorphism for every K.

Example 3.2. If G is an algebraic group over F, we can view G as a functor taking a field extension K to the group $G(K)$ of K-points of G; in this case we consider $\operatorname{Inv}^d_h(G)$. We have also another functor $H^1(G) : K \to H^1(K, G)$ and we consider $\operatorname{Inv}^d(H^1(G))$. If G is commutative, then $H^1(K, G)$ is a group for every K, and we also consider $\operatorname{Inv}^d_h(H^1(G))$.

3B. *Residues.* Our goal is to prove Theorem 3.7 concerning the group $\operatorname{Inv}^d_h(T)$ for T a quasisplit torus. Such invariants of order not divisible by char(F) were determined in [Merkurjev et al. 2002]. We modify the method from [Merkurjev et al. 2002] so that it works in general. The difficulty is that the groups $H^j(K)$ do not form a cycle module, because the residue homomorphisms need not exist.

If K is a field with discrete valuation v and residue field $\kappa(v)$, write $H^j(F)_{nr,v}$ for the subgroup of all elements of $H^j(F)$ that are split by finite separable extensions K/F such that v admits an unramified extension to K. Note that every element in $H^j(F)_{nr,v}$ of order not divisible by char(F) belongs to $H^j(F)_{nr,v}$.

There are *residue homomorphisms* (see [Garibaldi et al. 2003] or [Kato 1982])

$$\partial_v : H^j(K)_{nr,v} \to H^{j-1}(\kappa(v)).$$

Example 3.3. Let $K = F(t)$ and let v be the discrete valuation associated with t. Then $\kappa(v) = F$ and $\partial_v(t \cdot h_K) = h$ for all $h \in H^{j-1}(F)$.

Lemma 3.4. *Let K'/K be a field extension and let v' be a discrete valuation on K' unramified over its restriction v on K. Then the diagram*

$$\begin{array}{ccc} H^j(K)_{nr,v} & \xrightarrow{\partial_v} & H^{j-1}(\kappa(v)) \\ \downarrow & & \downarrow \\ H^j(K')_{nr,v'} & \xrightarrow{\partial_{v'}} & H^{j-1}(\kappa(v')) \end{array}$$

commutes.

3C. ***Invariants of tori.*** Let A be an étale F-algebra and T^A the corresponding quasisplit torus, i.e.,

$$T^A(K) = (A \otimes_F K)^\times$$

for every field extension K/F. If B is another étale F-algebra, then

$$T^{A\times B} = T^A \times T^B$$

and

$$\mathrm{Inv}_h^d(T^{A\times B}) \simeq \mathrm{Inv}_h^d(T^A) \oplus \mathrm{Inv}_h^d(T^B).$$

Write A as a product of fields: $A = L_1 \times L_2 \times \cdots \times L_s$. Set

$$H^i(A) := H^i(L_1) \oplus H^i(L_2) \oplus \cdots \oplus H^i(L_s).$$

For $d \geq 2$ define a homomorphism

$$\alpha^A : H^{d-1}(A) \to \mathrm{Inv}_h^d(T^A)$$

as follows. If $h \in H^{d-1}(A)$, then the invariant $\alpha^A(h)$ is defined by

$$\alpha^A(h)(t) = N_{A\otimes K/K}(t \cdot h_{A\otimes K}) \in H^d(K)$$

for a field extension K/F and $t \in T^A(K) = (A \otimes_F K)^\times$.

Remark 3.5. In the notation of the previous section, $(T^A)^\circ \simeq T^A$, and we have

$$H^{d-1}(F, (T^A)^\circ) = H^{d-1}(F, T^A) = H^{d-1}(A, \mathbb{G}_m) = H^{d-1}(A).$$

The pairing (2.10) for the torus T^A, $i = 0$, and $j = 2$,

$$A^\times \otimes H^2(A) = T^A(F) \otimes H^2(F, (T^A)^\circ) \to H^3(F),$$

takes $t \otimes h$ to $N_{A/F}(t \cup h_A) = \alpha^A(h)(t)$. In other words, the map α^A coincides with the map

$$H^2(F, (T^A)^\circ) \to \mathrm{Inv}_h^3(T^A)$$

given by the cup product.

Note that every element $h \in H^{d-1}(A)$ is split by an étale extension of A, hence the invariant $\alpha^A(h)$ vanishes when restricted to F_{sep}.

Question 3.6. Do all invariants in $\mathrm{Inv}_h^d(T^A)$ vanish when restricted to F_{sep}?

The answer is "yes" when $\mathrm{char}(F) = 0$. For any prime $p \neq \mathrm{char}(F)$ and for F separably closed, the zero map is the only invariant $T^A(*) \to H^d(*, \mathbb{Q}_p/\mathbb{Z}_p(d-1))$ that is a homomorphism of groups [Merkurjev 1999, Proposition 2.5].

The main result of this section is:

Theorem 3.7. *The sequence*

$$0 \to H^{d-1}(A) \xrightarrow{\alpha^A} \operatorname{Inv}_h^d(T^A) \xrightarrow{\text{res}} \operatorname{Inv}_h^d(T^A_{\text{sep}})$$

is exact.

That is, defining $\widetilde{\operatorname{Inv}}_h^d(T^A) := \ker \text{res}$, we claim that $\alpha^A : H^{d-1}(A) \xrightarrow{\sim} \widetilde{\operatorname{Inv}}_h^d(T^A)$.

The torus T^A is embedded into the affine space $\mathbb{A}(A)$ as an open set. Let Z^A be the closed complement $\mathbb{A}(A) \setminus T^A$ and let S^A be the smooth locus of Z^A (see [Merkurjev et al. 2002]). Then S^A is a smooth scheme over A. In fact, S^A is a quasisplit torus over A of the A-algebra A' such that $A \times A' \simeq A \otimes_F A$. We have $A = L_1 \times L_2 \times \cdots \times L_s$, where the L_i are finite separable field extensions of F, and the connected components of S^A (as well as the irreducible components of Z^A) are in one-to-one correspondence with the factors L_i. Let v_i for $i = 1, 2, \ldots, s$ be the discrete valuation of the function field $F(T^A)$ corresponding to the i-th connected component S_i of S^A, or equivalently, to the i-th irreducible component Z_i of Z^A. The residue field of v_i is equal to the function field $F(Z_i) = F(S_i)$. We then have the residue homomorphisms

$$\partial_i : H^d(F(T^A))_{nr,v_i} \to H^{d-1}(F(Z_i)) = H^{d-1}(F(S_i)).$$

Write $\widetilde{H}^d(F(T^A))$ for the kernel of the natural homomorphism $H^d(F(T^A)) \to H^d(F_{\text{sep}}(T^A))$. Since every extension of the valuation v_i to $F_{\text{sep}}(T^A)$ is unramified, we have $\widetilde{H}^d(F(T^A)) \subset H^d(F(T^A))_{nr,v_i}$ for all i. Write $F(S^A)$ for the product of $F(S_i)$ over all i. The sum of the restrictions of the maps ∂_i on $\widetilde{H}^d(F(T^A))$ yields a homomorphism

$$\partial^A : \widetilde{H}^d(F(T^A)) \to H^{d-1}(F(S^A)).$$

Applying $u \in \widetilde{\operatorname{Inv}}_h^d(T^A)$ to the generic element g_{gen} of T^A over the function field $F(T^A)$, we get a cohomology class $u(g_{\text{gen}}) \in H^d(F(T^A))$. By assumption on u, we have $u(g_{\text{gen}}) \in \widetilde{H}^d(F(T^A))$. Applying ∂^A, we get a homomorphism

$$\beta^A : \widetilde{\operatorname{Inv}}_h^d(T^A) \to H^{d-1}(F(S^A)), \qquad u \mapsto \partial^A(u(g_{\text{gen}})).$$

If B is another étale F-algebra, we have (see [Merkurjev et al. 2002])

$$S^{A\times B} = S^A \times T^B + T^A \times S^B.$$

In particular, $F(S^A) \subset F(S^{A\times B}) \supset F(S^B)$. Lemma 3.4 then gives:

Lemma 3.8. *The diagram*

$$\begin{tikzcd}
\widetilde{\operatorname{Inv}}_h^d(T^A) \oplus \widetilde{\operatorname{Inv}}_h^d(T^B) \arrow[r, "\beta^A \oplus \beta^B"] \arrow[d, equal] & H^{d-1}(F(S^A)) \oplus H^{d-1}(F(S^B)) \arrow[d] \\
\widetilde{\operatorname{Inv}}_h^d(T^{A\times B}) \arrow[r, "\beta^{A\times B}"] & H^{d-1}(F(S^{A\times B}))
\end{tikzcd}$$

commutes.

Recall that S^A is a smooth scheme over A with an A-point. It follows that $A \subset F(S^A)$ and the natural homomorphism

$$H^j(A) \to H^j(F(S^A))$$

is injective by [Garibaldi et al. 2003, Proposition A.10]. We shall view $H^j(A)$ as a subgroup of $H^j(F(S^A))$.

Let $A = L_1 \times L_2 \times \cdots \times L_s$ be the decomposition of an étale F-algebra A into a product of fields. The *height* of A is the maximum of the degrees $[L_i : F]$. The height of A is 1 if and only if A is split. The following proposition will be proved by induction on the height of A.

Proposition 3.9. *The image of the homomorphism β^A is contained in $H^{d-1}(A)$.*

Proof. By Lemma 3.8 we may assume that $A = L$ is a field. If $L = F$, we have $S^A = \operatorname{Spec} F$, so $A = F(S^A)$ and the statement is clear.

Suppose $L \neq F$. The algebra L is a canonical direct factor of $L \otimes_F L$. It follows that the homomorphism β^L is a direct summand of $\beta^{L\otimes L}$. Since the height of the L-algebra $L \otimes_F L$ is less than the height of A, by the induction hypothesis, $\operatorname{Im}(\beta^{L\otimes L}) \subset H^{d-1}(L \otimes L)$. It follows that $\operatorname{Im}(\beta^L) \subset H^{d-1}(L)$. □

It follows from Proposition 3.9 that we can view β^A as a homomorphism

$$\beta^A : \widetilde{\operatorname{Inv}}_h^d(T^A) \to H^{d-1}(A).$$

We will show that α^A and β^A are isomorphisms inverse to each other. First consider the simplest case.

Lemma 3.10. *The maps α^A and β^A are isomorphisms inverse to each other in the case $A = F$.*

Proof. If $A = F$, then we have $T^A = \mathbb{G}_m$. The generic element g_{gen} is equal to $t \in F(t)^\times = F(\mathbb{G}_m)$. Let $h \in H^{d-1}(A) = H^{d-1}(F)$. Then the invariant $\alpha^F(h)$ takes t to $t \cdot h \in \widetilde{H}^d(F(t))$. By Example 3.3, $\beta^F(\alpha^F(h)) = \partial_v(t \cdot h) = h$, i.e., the composition $\beta^F \circ \alpha^F$ is the identity. It suffices to show that α^F is surjective.

Take $u \in \widetilde{\operatorname{Inv}}_h^d(\mathbb{G}_m)$. We consider t as an element of the complete field $L := F((t))$ and let $x = u_L(t) \in H^d(L)$. By assumption, x is split by the maximal unramified extension $L' := F_{\mathrm{sep}}((t))$ of L. By a theorem of Kato [1982],

$$x \in \operatorname{Ker}\big(H^d(L) \to H^d(L')\big) = H^d(F) \oplus t \cdot H^{d-1}(F),$$

i.e., $x = h'_L + t \cdot h_L$ for some $h' \in H^d(F)$ and $h \in H^{d-1}(F)$.

Let K/F be a field extension. We want to compute $u_K(a) \in H^d(K)$ for an element $a \in K^\times$. Consider the field homomorphism $\varphi : L \to M := K((t))$ taking a power series $f(t)$ to $f(at)$. By functoriality,

$$u_M(at) = u_M(\varphi(t)) = \varphi_*(u_L(t)) = \varphi_*(x) = \varphi_*(h'_L + t \cdot h_L) = h'_M + (at) \cdot h_M,$$

therefore,

$$u_M(a) = u_M(at) - u_M(t) = (h'_M + (at)\cdot h_M) - (h'_M + t\cdot h_M) = a\cdot h_M.$$

It follows that $u(a) = a\cdot h_K$ since the homomorphism $H^d(K) \to H^d(M)$ is injective by [Garibaldi et al. 2003, Proposition A.9]. We have proved that $u = \alpha^A(h)$, i.e., α^A is surjective. □

Lemma 3.11. *The homomorphism β^A is injective.*

Proof. The proof is similar to the proof of Proposition 3.9. We induct on the height of A. The right vertical homomorphism in Lemma 3.8 is isomorphic to the direct sum of the two homomorphisms $H^{d-1}(F(S^A)) \to H^{d-1}(F(S^A \times T^B))$ and $H^{d-1}(F(S^B)) \to H^{d-1}(F(T^A \times S^B))$. Both homomorphisms are injective by [Garibaldi et al. 2003, Proposition A.10]. It follows from Lemma 3.8 that we may assume that $A = L$ is a field.

The case $L = F$ follows from Lemma 3.10, so we may assume that $L \neq F$. The homomorphism β^L is a direct summand of $\beta^{L\otimes L}$. The latter is injective by the induction hypothesis, hence so is β^L. □

Lemma 3.12. *The composition $\beta^A \circ \alpha^A$ is the identity.*

Proof. We again induct by the height of A. By Lemma 3.8 that we may assume that $A = L$ is a field.

The case $L = F$ follows from Lemma 3.10, so we may assume that $L \neq F$. The homomorphisms α^L and β^L are direct summands of $\alpha^{L\otimes L}$ and $\beta^{L\otimes L}$, respectively. The composition $\beta^{L\otimes L} \circ \alpha^{L\otimes L}$ is the identity by the induction hypothesis, hence $\beta^A \circ \alpha^A$ is also the identity. □

It follows from Lemma 3.11 and Lemma 3.12 that α^A and β^A are isomorphisms inverse to each other. This completes the proof of Theorem 3.7.

4. Invariants of groups of multiplicative type

In this section, C denotes a group of multiplicative type over F such that there exists an exact sequence

$$1 \to C \to T \to S \to 1$$

such that S and T are quasitrivial tori. For example, this holds if C is the center of a simply connected semisimple group G over F, such as μ_n. In that case, C is isomorphic to the center of the quasisplit inner form G^q of G, and we take T to be any quasitrivial maximal torus in G^q. Then T^* is the weight lattice Λ_w and $S^* \simeq \Lambda_r$, where the Galois action permutes the fundamental weights and simple roots, respectively.

Proposition 4.1. *Every invariant in* $\widetilde{\mathrm{Inv}}^3_h(H^1(C))$ *is given by the cup product via the pairing* (2.9) *with a unique element in* $H^2(F, C^\circ)$.

Proof. Since $H^1(K, T) = 1$ for every K, the connecting homomorphism $S(K) \to H^1(K, C)$ is surjective for every K and therefore the natural homomorphism

$$\mathrm{Inv}^3_h(H^1(C)) \to \mathrm{Inv}^3_h(S)$$

is injective.

Consider the diagram

$$\begin{tikzcd}
H^2(F, C^\circ) \arrow[r, hook] \arrow[d] & H^2(F, S^\circ) \arrow[r] \arrow[d, "\wr"] & H^2(F, T^\circ) \arrow[d, "\wr"] \\
\widetilde{\mathrm{Inv}}^3_h(H^1(C)) \arrow[r, hook] & \widetilde{\mathrm{Inv}}^3_h(S) \arrow[r] & \widetilde{\mathrm{Inv}}^3_h(T)
\end{tikzcd}$$

where the vertical homomorphisms are given by cup products and the top row comes from the exact sequence (2.11); it is exact since $H^1(K, T^\circ) = 1$ for every field extension K/F. The bottom row comes from applying $\widetilde{\mathrm{Inv}}^3_h$ to the sequence $T(K) \to S(K) \to H^1(K, C)$; it is a complex. The vertical arrows are cup products, and the middle and right ones are isomorphisms by Theorem 3.7 and Remark 3.5. The diagram commutes by Proposition 2.12. By diagram chase, the left vertical map is an isomorphism. □

Note that the group $H^2(F, T)$ is a direct sum of the Brauer groups of finite extensions of F. Therefore, we have the following, a coarser version of [Garibaldi 2012, Proposition 7]:

Lemma 4.2. *The homomorphism* $H^2(F, C) \to \coprod \mathrm{Br}(K)$, *where the direct sum is taken over all field extensions* K/F *and all characters of* C *over* K, *is injective.*

Remark 4.3. The group G becomes quasisplit over the function field $F(X)$ of the variety X of Borel subgroups of G, so $F(X)$ kills t_G. But the kernel of $H^2(F, C) \to H^2(F(X), C)$ need not be generated by t_G, as can be seen by taking G of inner type D_n for n divisible by 4.

5. Root system preliminaries

5A. ***Notation.*** Let V be a real vector space and $R \subset V$ a root system (which we assume is reduced). Write $\Lambda_r \subset \Lambda_w$ for the *root* and *weight lattices*, respectively. For every root $\alpha \in R$, the reflection s_α with respect to α is given by the formula

$$s_\alpha(x) = x - \alpha^\vee(x) \cdot \alpha, \tag{5.1}$$

for every $x \in V$, where $\alpha^\vee \in V^* := \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$ is the *coroot* dual to α. Write $R^\vee \subset V^*$ for the dual root system and $\Lambda_r^\vee \subset \Lambda_w^\vee$ for the corresponding lattices.

We have
$$\Lambda_r^\vee = (\Lambda_w)^* := \mathrm{Hom}(\Lambda_w, \mathbb{Z}) \quad \text{and} \quad \Lambda_w^\vee = (\Lambda_r)^*.$$

The *Weyl group* W of R is a normal subgroup of the automorphism group $\mathrm{Aut}(R)$ of the root system R. The factor group $\mathrm{Aut}(R)/W$ is isomorphic to the automorphism group $\mathrm{Aut}(\mathrm{Dyn}(R))$ of the Dynkin diagram of R. There is a unique $\mathrm{Aut}(R)$-invariant scalar product $(\,,\,)$ on V normalized so that square-length $d_{\alpha^\vee} := (\alpha, \alpha)$ of short roots in every irreducible component of R is equal to 1. The formula (5.1) yields an equality
$$\alpha^\vee(x) = \frac{2(\alpha, x)}{(\alpha, \alpha)}$$
for all $x \in V$ and $\alpha \in R$.

We may repeat this construction with the dual root system $R^\vee$, defining $(\,,\,)^\vee$ on V^* so that the square-length $d_\alpha := (\alpha^\vee, \alpha^\vee)^\vee$ is 1 for short coroots $\alpha^\vee$ (equivalently, long roots α).

5B. ***The map φ.***

Proposition 5.2. *There is a unique $\mathbb{R}$-linear map $\varphi : V^* \to V$ such that $\varphi(\alpha^\vee) = \alpha$ for all short $\alpha^\vee$. Furthermore, φ is $\mathrm{Aut}(R)$-invariant, $\varphi(\alpha^\vee) = d_\alpha \alpha$ for all $\alpha^\vee \in R^\vee$, $\varphi(\Lambda_w^\vee) \subseteq \Lambda_w$, and $\varphi(\Lambda_r^\vee) \subseteq \Lambda_r$. Analogous statements hold for $\varphi^\vee : V \to V^*$. If R is irreducible, then $\varphi\varphi^\vee : V^* \to V^*$ and $\varphi^\vee\varphi : V \to V$ are multiplication by d_α for α a short root.*

Proof. Define $\varphi^\vee$ by $\langle \varphi^\vee(x), y \rangle = 2(x, y)$ for $x, y \in V$ and φ by $\langle x', \varphi(y') \rangle = 2(x', y')^\vee$ for $x', y' \in V^*$. We have
$$\langle \varphi^\vee(\alpha), x \rangle = 2(\alpha, x) = (\alpha, \alpha) \cdot \alpha^\vee(x) = d_{\alpha^\vee} \cdot \alpha^\vee(x),$$
hence $\varphi^\vee(\alpha) = d_{\alpha^\vee} \cdot \alpha^\vee$, and similarly for φ. For uniqueness of φ and $\varphi^\vee$, it suffices to note that the short roots generate V^*, which is obvious because they generate a subspace that is invariant under the Weyl group.

Let $x \in \Lambda_w$. By definition,
$$\mathbb{Z} \ni \alpha^\vee(x) = \frac{2(x, \alpha)}{(\alpha, \alpha)}$$
for all $\alpha \in R$. It follows that $\langle \varphi^\vee(x), \alpha \rangle = 2(x, \alpha) \in \mathbb{Z}$ since $(\alpha, \alpha) \in \mathbb{Z}$. Therefore,
$$\varphi(x) \in \Lambda_w^\vee.$$

For each root $\beta \in R$, $\varphi^\vee\varphi(\beta^\vee) = d_\beta d_{\beta^\vee} \beta^\vee$ and similarly for $\varphi\varphi^\vee$. As R is irreducible, either all roots have the same length (in which case $d_\beta d_{\beta^\vee} = 1$) or there are two lengths and β and $\beta^\vee$ have different lengths (in which case $d_\beta d_{\beta^\vee}$ is the square-length of a long root); in either case the product equals d_α as claimed. □

Remark 5.3. If the root system R is simply laced, then φ gives isomorphisms from V^*, $\Lambda_w^\vee$, and $\Lambda_r^\vee$ to V, Λ_w, and Λ_r, respectively, that agree with the canonical bijection $R^\vee \to R$ defined by $\alpha^\vee \leftrightarrow \alpha$.

Example 5.4. For $\alpha^\vee$ a simple coroot, we write $f_\alpha^\vee$ for the corresponding fundamental dominant weight of $R^\vee$. Consider an element $x' = \sum x_\beta \beta^\vee$ where β ranges over the simple roots. As $(f_\alpha^\vee, \beta^\vee)^\vee = \frac{1}{2}\langle f_\alpha^\vee, \beta\rangle(\beta^\vee, \beta^\vee)^\vee$, we have $(f_\alpha^\vee, x')^\vee = x_\alpha (f_\alpha^\vee, \alpha^\vee)^\vee = \frac{1}{2} d_\alpha x_\alpha$. That is, $\langle \varphi(f_\alpha^\vee), x'\rangle = d_\alpha x_\alpha = \langle d_\alpha f_\alpha, x'\rangle$ for all x', and we conclude that $\varphi(f_\alpha^\vee) = d_\alpha f_\alpha$.

Remark 5.5. Let $q \in S^2(\Lambda_w)^W$ be the only quadratic form on $\Lambda_r^\vee$ that is equal to 1 on every short coroot in every component of $R^\vee$. It is shown in [Merkurjev 2016, Lemma 2.1] that the polar form p of q in $\Lambda_w \otimes \Lambda_w$ in fact belongs to $\Lambda_r \otimes \Lambda_w$. Then the restriction of φ on $\Lambda_w^\vee$ coincides with the composition

$$\Lambda_w^\vee \xrightarrow{\mathrm{id}\otimes p} \Lambda_w^\vee \otimes (\Lambda_r \otimes \Lambda_w) = (\Lambda_w^\vee \otimes \Lambda_r) \otimes \Lambda_w \to \Lambda_w.$$

5C. ***The map ρ.*** Write $\Delta := \Lambda_w/\Lambda_r$ and $\Delta^\vee := \Lambda_w^\vee/\Lambda_r^\vee$. Note that Δ and $\Delta^\vee$ are dual to each other with respect to the pairing

$$\Delta \otimes \Delta^\vee \to \mathbb{Q}/\mathbb{Z}.$$

The group W acts trivially on Δ and $\Delta^\vee$, hence Δ and $\Delta^\vee$ are Aut(Dyn(R))-modules. The homomorphism φ yields an Aut(R)-equivariant homomorphism

$$\rho : \Delta^\vee \to \Delta.$$

The map ρ is an isomorphism if R is simply laced (because φ is an isomorphism) or if $\Lambda_w = \Lambda_r$. Similarly, $\rho = 0$ if and only if $\varphi(\Lambda_w^\vee) \subseteq \Lambda_r$, if and only if $p \in \Lambda_r \otimes \Lambda_r$.

Example 5.6. Suppose R has type C_n for some $n \geq 3$. Consulting the tables in [Bourbaki 2002], $f_n^\vee$, the fundamental weight of $R^\vee$ dual to the unique long simple root α_n, is the only fundamental weight of $R^\vee$ not in the root lattice. As α_n is long, $d_{\alpha_n} = 1$, so $\varphi(f_n^\vee) = f_n$, which belongs to Λ_r if and only if n is even. That is, $\rho = 0$ if and only if n is even; for n odd, ρ is an isomorphism.

Example 5.7. Suppose R has type B_n for some $n \geq 2$. For the unique short simple root α_n, $d_{\alpha_n} = 2$, and $\varphi(f_n^\vee) = 2f_n$ is in Λ_r. For $1 \leq i < n$, $\varphi(f_i^\vee) = f_i \in \Lambda_r$. We find that $\rho = 0$ regardless of n.

Thus we have determined $\ker \rho$ for every irreducible root system.

Example 5.8. Suppose R is irreducible and $\mathrm{char}(F) = d_\alpha$ for some short root α. Then for G, $G^\vee$ simple simply connected with root system R, $R^\vee$ respectively, there is a "very special" isogeny $\pi : G \to G^\vee$. The restriction of π to a maximal torus in G induces a $\mathbb{Z}$-linear map on the cocharacter lattices $\pi_* : \Lambda_r^\vee \to \Lambda_r$, which, by [Conrad et al. 2015, Proposition 7.1.5] or [Steinberg 1963, 10.1], equals φ.

In the case $R = B_n$, π is the composition of the natural map $G = \mathrm{Spin}_{2n+1} \to \mathrm{SO}_{2n+1}$ with the natural (characteristic 2 only) map $\mathrm{SO}_{2n+1} \to \mathrm{Sp}_{2n}$. As π vanishes on the center of G, it follows that $\rho = 0$ as in Example 5.7. Similarly, in case $R = C_n$, one can recover Example 5.6 by noting that the composition $\pi : G = \mathrm{Sp}_{2n} \to \mathrm{Spin}_{2n+1}$ with the spin representation $\mathrm{Spin}_{2n+1} \hookrightarrow \mathrm{GL}_{2^n}$ is the irreducible representation of G with highest weight f_n by [Steinberg 1963, §11].

Example 5.9. For $R = A_{n-1}$, define $\tau : \Delta \xrightarrow{\sim} \mathbb{Z}/n\mathbb{Z}$ via $\tau(f_1) = 1/n \in \mathbb{Q}/\mathbb{Z}$. As $\langle f_1, f_1^\vee \rangle = (n-1)/n \in \mathbb{Q}$, defining $\tau^\vee : \Delta^\vee \xrightarrow{\sim} \mathbb{Z}/n\mathbb{Z}$ via $\tau^\vee(f_1^\vee) = -1/n \in \mathbb{Q}/\mathbb{Z}$ gives a commutative diagram

$$\begin{tikzcd} \Delta \otimes \Delta^\vee \arrow[r, "\langle\, ,\, \rangle"] \arrow[d, "\tau\otimes\tau^\vee"', "\wr"] & \mathbb{Q}/\mathbb{Z} \\ \mathbb{Z}/n\mathbb{Z} \otimes \mathbb{Z}/n\mathbb{Z} \arrow[ur, "\text{natural}"'] & \end{tikzcd}$$

i.e., $\tau^\vee$ is the isomorphism induced by τ and the natural pairings. Furthermore, although ρ is induced by the canonical isomorphism $R^\vee \simeq R$, the previous discussion shows that the diagram

$$\begin{tikzcd} \Delta^\vee \arrow[r, "\rho"] \arrow[d, "\tau^\vee"', "\wr"] & \Delta \arrow[d, "\tau"', "\wr"] \\ \mathbb{Z}/n\mathbb{Z} \arrow[r, "-1"] & \mathbb{Z}/n\mathbb{Z} \end{tikzcd} \tag{5.10}$$

commutes, where the bottom arrow is multiplication by -1.

(The action of $\mathrm{Aut}(R)$ on Δ interchanges f_1 and f_{n-1}. Defining instead $\tau(-f_1) = \tau(f_{n-1}) = 1/n$ also gives the same commutative diagram (5.10). That is, the commutativity of (5.10) is invariant under $\mathrm{Aut}(R)$.)

6. Statement of the main result

6A. *The map ρ.* Let G be a simply connected semisimple group with root system R. Let C be the center of G. Then $C^* = \Lambda_w/\Lambda_r = \Delta$ and $C_* = \Lambda_w^\vee/\Lambda_r^\vee = \Delta^\vee$, and we get from Section 5C a homomorphism

$$\rho = \rho_G : C_* \to C^*$$

of Galois modules. Therefore, we have a group homomorphism

$$\hat\rho = \hat\rho_G : C \to C^\circ.$$

Note that $\hat\rho$ is an isomorphism if R is simply laced.

6B. *The Tits class.* Let G be a simply connected group over F with center C. Write t_G for the *Tits class* $t_G \in H^2(F, C)$. By definition, $t_G = -\partial(\xi_G)$, where

$$\partial : H^1(F, G/C) \to H^2(F, C)$$

is the connecting map for the exact sequence $1 \to C \to G \to G/C \to 1$ and $\xi_G \in H^1(F, G/C)$ is the unique class such that the twisted group ${}^{\xi}G$ is quasisplit.

6C. ***Rost invariant for an absolutely simple group.*** Let G be a simply connected group over F. Recall (see [Garibaldi et al. 2003]) that, for G absolutely simple, Rost defined an invariant $r_G \in \mathrm{Inv}^3(H^1(G))$ called the *Rost invariant*, i.e., a map

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

that is functorial in F.

Lemma 6.1. *If G is an absolutely simple and simply connected algebraic group, then $o(r_G) \cdot t_G = 0$.*

Proof. The order $o(r_G)$ of r_G is calculated in [Garibaldi et al. 2003], and in each case it is a multiple of the order of t_G. □

As mentioned in [Gille 2000, §2.3], there are several definitions of the Rost invariant that may differ by a sign. Gille and Quéguiner [2011] proved that for the definition of the Rost invariant r_G they chose, in the case $G = \mathbf{SL}_1(A)$ for a central simple algebra A of degree n over F, the value of r_G on the image of the class $aF^{\times n} \in F^\times/F^{\times n} = H^1(F, \mu_n)$ in $H^1(F, G)$ is equal to $(a) \cup [A]$ if n is not divisible by $\mathrm{char}(F)$ and to $-(a) \cup [A]$ if n is a power of $p = \mathrm{char}(F) > 0$. Therefore, *we normalize the Rost invariant by multiplying the p-primary component of the Rost invariant (of all groups) by -1 in the case $p = \mathrm{char}(F) > 0$.*

6D. ***The main theorem.*** For G semisimple and simply connected over F, there is an isomorphism

$$\psi : G \xrightarrow{\sim} \prod_{i=1}^{n} R_{F_i/F}(G_i), \tag{6.2}$$

where the F_i are finite separable extensions of F, and G_i is an absolutely simple and simply connected F_i-group. The product of the corestrictions of the r_{G_i} (in the sense of [Garibaldi et al. 2003, page 34]) is then an invariant of $H^1(G)$, which we also denote by r_G and call the Rost invariant of G. The map ψ identifies the center C of G with $\prod_i R_{F_i/F}(C_i)$ for C_i the center of G_i, and the Tits class $t_G \in H^2(F, C)$ with $\sum t_{G_i} \in \sum H^2(F_i, C_i)$.

The composition $r_G \circ i^*$ is a group homomorphism by [Merkurjev et al. 2002, Corollary 1.8] or [Garibaldi 2001, Lemma 7.1]. That is, the composition $r_G \circ i^*$ in Theorem 1.2 taken over all field extensions of F can be viewed not only as an invariant of $H^1(C)$, but as an element of $\mathrm{Inv}^3_h(H^1(C)$ as in Definition 3.1. Over a separable closure of F, the inclusion of C into G factors through a maximal split torus and hence this invariant is trivial by Hilbert's Theorem 90. By Proposition 4.1 the composition is given by the cup product with a unique element in $H^2(F, C^\circ)$. We will prove Theorem 1.2, which says that this element is equal to $-t_G^\circ$.

6E. ***Alternative formulation.*** Alternatively, we could formulate the main theorem as follows. The group of invariants $\operatorname{Inv}^3(H^1(G))$ is a sum of n cyclic groups with generators (the corestrictions of) the r_{G_i}, and in view of Lemma 6.1 we may define a homomorphism

$$\operatorname{Inv}^3(H^1(G)) \to H^2(F, C) \quad \text{via} \sum n_i r_{G_i} \mapsto \sum -n_i t_{G_i}. \tag{6.3}$$

Theorem 6.4. *For every invariant $s : H^1(*, G) \to H^3(*, \mathbb{Q}/\mathbb{Z}(2))$, the composition*

$$H^1(*, C) \to H^1(*, G) \to H^3(*, \mathbb{Q}/\mathbb{Z}(2))$$

equals the cup product with the image of s under the composition

$$\operatorname{Inv}^3(H^1(G)) \to H^2(F, C) \to H^2(F, C^\circ).$$

This will follow immediately from the main theorem, which we will prove over the course of the next few sections.

7. Rost invariant of transfers

The following statement is straightforward.

Lemma 7.1. *Let A be an étale F-algebra and G a simply connected semisimple group scheme over A, with C the center of G. Then $C' := R_{A/F}(C)$ is the center of $G' := R_{A/F}(G)$ and $C'^\circ \simeq R_{A/F}(C^\circ)$, and the diagram*

$$\begin{array}{ccc} H^i(A, C) & \xrightarrow{\hat{\rho}_G^*} & H^i(A, C^\circ) \\ \downarrow \wr & & \downarrow \wr \\ H^i(F, C') & \xrightarrow{\hat{\rho}_{G'}^*} & H^i(F, C'^\circ) \end{array}$$

commutes.

Lemma 7.2. *Set $C' := R_{A/F}(C)$ and $G' := R_{A/F}(G)$. Then the image of t_G under the isomorphism $H^2(A, C) \xrightarrow{\sim} H^2(F, C')$ is equal to $t_{G'}$.*

Proof. The corestriction of a quasisplit group is quasisplit. □

Lemma 7.3. *Let G be a simply connected semisimple algebraic group scheme over an étale F-algebra A. If Theorem 1.2 holds for G, then it also holds for $R_{A/F}(G)$.*

Proof. Let C be the center of G and $G' := R_{A/F}(G)$. The group $C' := R_{A/F}(C)$ is the center of G'. Let $x \in H^1(A, C)$ and let $x' \in H^1(F, C')$ be the image of x under the isomorphism $\nu : H^1(A, C) \xrightarrow{\sim} H^1(F, C')$. We have

$$\begin{aligned} r_{G'}(i'^*(x')) &= r_{G'}(\nu(i^*(x))) \\ &= N_{A/F}(r_G(i^*(x)) && \text{by [Garibaldi et al. 2003, Proposition 9.8]} \\ &= N_{A/F}(-t_G^\circ \cup x) && \text{by Theorem 1.2 for } x \\ &= -t_{G'}^\circ \cup x' && \text{by Lemmas 2.14, 7.1 and 7.2.} \end{aligned}$$

□

If Theorem 1.2 holds for semisimple groups G_1 and G_2, then it also holds for the group $G_1 \times G_2$. Combining this with Lemma 7.3 reduces the proof of Theorem 1.2 to the case where G is absolutely almost simple.

8. Rost invariant for groups of type A

In this section, we will prove Theorem 1.2 for G absolutely simple of type A_{n-1} for each $n \geq 2$.

8A. ***Inner type.*** Suppose G has inner type. Then there is an isomorphism $\psi : \mathbf{SL}_1(A) \xrightarrow{\sim} G$, where A is a central simple algebra of degree n over F. The map ψ restricts to an isomorphism $\mu_n \xrightarrow{\sim} C$, identifying C^* with $\mathbb{Z}/n\mathbb{Z}$, and induces $\psi^\circ : C^\circ \xrightarrow{\sim} \mu_n$. We find a commutative diagram

$$\begin{tikzcd} H^2(F, C^\circ) \otimes H^1(F, C) \arrow[r] \arrow[d, "\psi^\circ \otimes \psi^{-1}"'] & H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \arrow[d, equal] \\ H^2(F, \mu_n) \otimes H^1(F, \mu_n) \arrow[r] & H^3(F, \mathbb{Q}/\mathbb{Z}(2)) \end{tikzcd} \tag{8.1}$$

where the top and bottom arrows are the cup product from (2.9).

The connecting homomorphism arising from the Kummer sequence

$$1 \to \mu_n \to \mathbb{G}_m \to \mathbb{G}_m \to 1$$

gives an isomorphism $H^1(K, \mu_n) \simeq K^\times/K^{\times n}$ for every extension K/F. For each field extension K/F, the isomorphism ψ identifies the map $H^1(K, C) \to H^1(K, G)$ with the obvious map $K^\times/K^{\times n} = H^1(K, \mu_n) \to H^1(K, \mathbf{SL}_1(A)) = K^\times/\operatorname{Nrd}(A_K^\times)$. Further, $\psi^{-1}(t_G) \in H^2(K, \mu_n)$ is the Brauer class $[A]$ of A as in [Knus et al. 1998, pages 378 and 426]. By Example 5.9, the composition

$$H^1(F, \mu_n) \xrightarrow{\psi} H^1(F, C) \xrightarrow{\hat{\rho}^*} H^1(F, C^\circ) \xrightarrow{\psi^\circ} H^1(F, \mu_n)$$

is multiplication by -1 and in particular $[A] \mapsto t_G \mapsto t_G^\circ \mapsto -[A]$. That is, Theorem 1.2 claims that the diagram

$$\begin{tikzcd} H^1(K, \mu_n) \arrow[r, "\psi^{-1}"] \arrow[d, "{[A]\otimes}"'] & H^1(K, C) \arrow[r] & H^1(K, G) \arrow[d, "r_G"] \\ H^2(K, \mu_n) \otimes H^1(K, \mu_n) \arrow[rr] & & H^3(K, \mathbb{Q}/\mathbb{Z}(2)) \end{tikzcd} \tag{8.2}$$

commutes, where the bottom arrow is the same as in (8.1).

Let p be a prime integer and write m for the largest power of p dividing n. Both maps $H^1(K, \mu_n) \to H^3(K, \mathbb{Q}/\mathbb{Z}(2))$ in (8.2) are group homomorphisms, so it suffices to verify Theorem 1.2 on each p-primary component $r_G(x)_p$ of the Rost invariant with values in $\mathbb{Q}_p/\mathbb{Z}_p(2)$. In the case where p does not divide $\operatorname{char}(F)$, the commutativity of (8.2) is part of [Gille and Quéguiner-Mathieu 2011, Theorem 1.1]. (Note that the definition of cup product used in [Gille and Quéguiner-Mathieu 2011], the one from [Gille and Szamuely 2006, §3.4], is the same as (8.1), cf. [Freitag and Kiehl 1988, pages 302–303].)

Now let $p = \operatorname{char}(F) > 0$. Consider the sheaf $\nu_m(j)$ in the étale topology over F defined by $\nu_m(j)(L) = K_j(L)/p^m K_j(L)$. The natural morphisms $\mathbb{Z}(j) \to \nu_m(j)[-j]$ for $j \le 2$ are consistent with the products, hence we have a commutative diagram:

$$\begin{array}{ccc} (\mathbb{Z}/m\mathbb{Z})(1)\overset{L}{\otimes}(\mathbb{Z}/m\mathbb{Z})(1) & \longrightarrow & (\mathbb{Z}/m\mathbb{Z})(2) \\ \downarrow\wr & & \downarrow\wr \\ \nu_m(1)[-1] \otimes \nu_m(1)[-1] & \longrightarrow & \nu_m(2)[-2] \end{array}$$

Therefore, we have a commutative diagram

$$\begin{array}{ccc} H^2(F, \mu_m) \otimes H^1(F, \mu_m) & \longrightarrow & H^3(F, \mathbb{Z}/p^m\mathbb{Z}(2)) \\ \downarrow\wr & & \downarrow\wr \\ H^1(F, \nu_m(1)) \otimes H^0(F, \nu_m(1)) & \longrightarrow & H^1(F, \nu_m(2)) \end{array}$$

(see Remark 2.13 after Proposition 2.12). The bottom arrow is given by the cup product map

$$_m\mathrm{Br}(F) \otimes (F^\times/F^{\times m}) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2))$$

(see [Gille and Quéguiner-Mathieu 2011, 4D]). It is shown in [Gille and Quéguiner-Mathieu 2011, Theorem 1.1] that the p-component of the Rost invariant of G is given by the formula

$$r_G(x)_p = [A]_p \cup (x) \in H^3(K, \mathbb{Q}_p/\mathbb{Z}_p(2))$$

for every $x \in K^\times$. (The formula in [Gille and Quéguiner-Mathieu 2011] contains an additional minus sign, but it does not appear here due to the adjustment in the definition of r_G in Section 6C.) This completes the proof of Theorem 1.2 for groups of inner type A.

8B. ***Outer type.*** Now suppose that G has outer type A_{n-1}. There is an isomorphism $\psi : G \xrightarrow{\sim} \mathbf{SU}(B, \tau)$, where B is a central simple algebra of degree n over a separable quadratic field extension K/F with an involution τ of the second kind (τ restricts to a nontrivial automorphism of K/F). The map ψ identifies C with $\mu_{n,[K]}$, and $C = C^\circ$.

Suppose first that n is odd. Since $G_K \simeq \mathbf{SL}_1(B)$, the theorem holds over K. As K has degree 2 over F and C has odd exponent, the natural map $H^1(F, C) \to H^1(K, C)$ is injective, hence the theorem holds over F by the following general lemma.

Lemma 8.3. *Let $L_1, L_2, \dots, L_s$ be field extensions of F such that the natural homomorphism $H^2(F, C) \to \prod_i H^2(L_i, C)$ is injective. If Theorem 1.2 holds for G over all fields L_i, then it also holds over F.*

Proof. The left vertical map in Theorem 1.2 is multiplication by some element $h \in H^2(F, C^\circ)$. We need to show that $h = -t_G^\circ$. This equality holds over all fields L_i, hence it holds over F by the injectivity assumption. □

So we may assume that n is even. Then $H^1(F, C)$ is isomorphic to a factor group of the group of pairs $(a, z) \in F^\times \times K^\times$ such that $N_{K/F}(z) = a^n$ and $H^2(F, C)$ is isomorphic to a subgroup of $\mathrm{Br}(F) \oplus \mathrm{Br}(K)$ of all pairs (v, u) such that $v_K = mu$ and $N_{K/F}(u) = 0$, see [Merkurjev et al. 2002, pages 795–796].

Suppose that B is split; we follow the argument in [Knus et al. 1998, 31.44]. Then $\mathbf{SU}(B, \tau) = \mathbf{SU}(h)$, where h is a hermitian form of trivial discriminant on a vector K-space of dimension n for the quadratic extension K/F. Let $q(v) := h(v, v)$ be the associated quadratic form on V viewed as a $2n$-dimensional F-space. The quadratic form q is nondegenerate, and we can view $\mathbf{SU}(h)$ as a subgroup of $H := \mathbf{Spin}(V, q)$. The Dynkin index of G in H is 1, hence the composition $H^1(K, G) \to H^1(K, H) \xrightarrow{r_H} H^3(K)$ equals the Rost invariant of G. Then r_H is given by the Arason invariant, which has order 2. A computation shows that the image of the pair (a, z) representing an element $x \in H^1(F, C)$ under the composition

$$H^1(F, C) \to H^1(F, G) \xrightarrow{r_G} H^3(F)$$

coincides with $[D] \cup x$, where D is the class of the discriminant algebra of h. On the other hand, $[D] \cup x$ coincides with the cup product of x with the Tits class $t_G = -t_G^\circ$ represented by the pair $([D], 0)$ in $H^2(F, C^\circ)$, proving Theorem 1.2 in this case.

Now drop the assumption that B is split. As for the n odd case, the theorem holds over K. Note that there is an injective map $H^2(F, C) \to \mathrm{Br}(F) \oplus \mathrm{Br}(K)$. Let $X = R_{K/F}(SB(B))$. By [Merkurjev and Tignol 1995, 2.4.6], the map $\mathrm{Br}(F) \to \mathrm{Br}(F(X))$ is injective, hence the natural homomorphism

$$H^2(F, C) \to H^2(F, C_{F(X)}) \oplus H^2(F, C_K)$$

is injective. The theorem holds over K and by the preceding paragraph the theorem holds over $F(X)$. Therefore, by Lemma 8.3, the theorem holds over F.

9. Conclusion of the proof of Theorem 1.2

Choose a system of simple roots Π of G. Write Π_r for the subset of Π consisting of all simple roots whose fundamental weight belongs to Λ_r and let $\Pi' := \Pi \setminus \Pi_r$. Inspection of the Dynkin diagram shows that all connected components of Π' have type A.

Every element of Π_r is fixed by every automorphism of the Dynkin diagram, hence is fixed by the $*$-action of the absolute Galois group of F on Π. It follows that the variety X of parabolic subgroups of G_{sep} of type Π' is defined over F. By [Merkurjev and Tignol 1995], the kernel of the restriction map $\mathrm{Br}(K) \to \mathrm{Br}(K(X))$ for every field extension K/F is generated by the Tits algebras associated with the classes in C^* of the fundamental weights f_α corresponding to the simple roots $\alpha \in \Pi_r$. But $f_\alpha \in \Lambda_r$, so these Tits algebras are split [Tits 1971], hence the restriction map $\mathrm{Br}(K) \to \mathrm{Br}(K(X))$ is injective and, by Lemma 4.2, the natural homomorphism $H^2(F, C) \to H^2(F(X), C)$ is injective. In view of Lemma 8.3, it suffices to prove Theorem 1.2 over the field $F(X)$, i.e., we may assume that G has a parabolic subgroup of type Π'. The Levi subgroup G' of that parabolic is simply connected with Dynkin diagram Π', and its center C' contains C [Garibaldi and Quéguiner-Mathieu 2007, Proposition 5.5]. Write j for the embedding homomorphism $C \to C'$ and j° for the dual $C'^\circ \to C^\circ$.

Let $G' = \prod_i G'_i$ with G_i simply connected simple groups, $C = \prod C_i$, where C_i is the center of G_i, and $\Pi'_i \subset \Pi$ is the system of simple roots of G_i. Write j_i° for the composition $C_i'^\circ \to C'^\circ \to C^\circ$.

Lemma 9.1. *The map $j_i^* : H^2(F, C) \to H^2(F, C'_i)$ takes the Tits class t_G to $t_{G'_i}$.*

Proof #1. It suffices to check that $j^*(t_G) = t_{G'}$, for the projection

$$H^2(F, C') \to H^2(F, C'_i)$$

sends $t_{G'} \mapsto t_{G'_i}$.

There is a rank $|\Pi_r|$ split torus S in G whose centralizer is $S \cdot G'$. Arguing as in Tits' Witt-type theorem [Tits 1966, 2.7.1, 2.7.2(d)], one sees that the quasisplit inner form of G is obtained by twisting G by a 1-cocycle γ with values in $C_G(S)/C$, equivalently, in G'/C. Clearly, twisting G' by γ gives the quasisplit inner form of G'. The Tits class t_G is defined to be $-\partial_G(\gamma)$ where ∂_G is the connecting homomorphism $H^1(F, G/C) \to H^2(F, C)$ induced by the exact sequence $1 \to C \to G \to G/C \to 1$ and similarly for G' and C'. The diagram

$$\begin{tikzcd}
H^1(F, G'/C) \arrow[r] \arrow[d] & H^1(F, G/C) \arrow[r, "\partial_G"] & H^2(F, C) \arrow[d, "j^*"] \\
H^1(F, G'/C') \arrow[rr, "\partial_{G'}"] & & H^2(F, C')
\end{tikzcd}$$

commutes trivially, so $j^*(t_G) = j^*(-\partial_G(\gamma)) = -\partial_{G'}(\gamma) = t_{G'}$ as claimed. $\square$

Proof #2. For each $\chi \in T^*$, define $F(\chi)$ to be the subfield of F_{sep} of elements fixed by the stabilizer of χ under the Galois action. Note that because G is absolutely almost simple, the $*$-action fixes Π_r elementwise, and $F(\chi)$ equals the field extension

$F(\chi|_{T'})$ defined analogously for $\chi \in (T')^*$. The diagram

$$\begin{tikzcd} H^2(F,C) \arrow[r, "j^*"] \arrow[dr, "\chi|_C"'] & H^2(F,C') \arrow[d, "\chi|_{C'}"] \\ & H^2(F(\chi),\mathbb{G}_m) \end{tikzcd}$$

commutes. Now $\chi|_{C'}(t_{G'} - j^*(t_G)) = \chi|_{C'}(t_{G'}) - \chi|_C(t_G)$, which is zero for all $\chi \in T^*$ by [Tits 1971, §5.5]. As $\prod_{\chi\in(T')^*}\chi|_{C'}$ is injective by [Garibaldi 2012, Proposition 7], $t_{G'} = j^*(t_G)$ as claimed. □

The diagram Π'_i is simply laced. Write d_i for the square-length of $\alpha^\vee \in R^\vee$ for $\alpha \in \Pi'_i$.

Lemma 9.2. *The homomorphism $\hat\rho_G : C \to C^\circ$ coincides with the composition*

$$C \xrightarrow{j} C' \xrightarrow{\hat\rho_{G'}} C'^\circ = \prod_i C_i'^\circ \xrightarrow{\prod_i (j_i{}^\circ)^{d_i}} C^\circ,$$

where j_i is the composition $C \to C' \to C'_i$.

Proof. For every simple root $\alpha \in \Pi$ write f_α for the corresponding fundamental weight. Write Λ'_r and Λ'_w for the root and weight lattices, respectively, of the root system R' of G'. Let

$$\Phi := \{f_\alpha \mid \alpha \in \Pi_r\}.$$

Then Φ is a $\mathbb{Z}$-basis for the kernel of the natural surjection $\Lambda_w \to \Lambda'_w$. If $\alpha \in \Pi'$, we write α' for the image of α and f'_α for the image of f_α under this surjection. All α' (respectively, f'_α) form the system of simple roots (respectively, fundamental weights) of R'. If $\alpha \in \Pi'$, the image $\alpha'^\vee$ of $\alpha^\vee$ under the inclusion ${\Lambda'_r}^\vee \hookrightarrow {\Lambda_r}^\vee$ is a simple coroot of R'.

If V is the real vector space of R, then $R' \subset V' := V/\operatorname{span}(\Phi)$ and $R'^\vee \subset V'^* \subset V^*$. Let $x \in \Lambda_w^\vee$, i.e., $\langle x, \alpha\rangle \in \mathbb{Z}$ for all $\alpha \in \Pi$. Since $\Phi \subset \Lambda_r$, we have $a_\alpha := \langle x, f_\alpha\rangle \in \mathbb{Z}$ for all $\alpha \in \Pi_r$. Then the linear form $x' := x - \sum_{\alpha\in\Pi_r} a_\alpha\alpha^\vee$ vanishes on the subspace of V spanned by Φ, hence $x' \in {\Lambda'_w}^\vee$. We then have a well-defined homomorphism

$$s : \Lambda_w^\vee \to {\Lambda'_w}^\vee, \qquad x \mapsto x'. \tag{9.3}$$

If $\alpha \in \Pi'$, then $\langle x', \alpha\rangle = \langle x', \alpha'\rangle$. It follows that if $x' = \sum_{\alpha\in\Pi} b_\alpha f_\alpha^\vee$ in $\Lambda_w^\vee$ for $b_\alpha = \langle x', \alpha\rangle \in \mathbb{Z}$, then $x' = \sum_{\alpha\in\Pi'} b_\alpha f_\alpha'^\vee$ in ${\Lambda'_w}^\vee$.

Since $\Phi \subset \Lambda_r$, we have a surjective homomorphism

$$C'^* = \Lambda'_w/\Lambda'_r = \Lambda_w/\operatorname{span}(\Phi, \Pi') \to \Lambda_w/\Lambda_r = C^*$$

dual to the inclusion of C into C'. The dual homomorphism

$$C_* = \Lambda_w^\vee/\Lambda_r^\vee \to {\Lambda'_w}^\vee/{\Lambda'_r}^\vee = C'_*$$

is induced by s.

Consider the diagram

$$\begin{array}{ccc} \Lambda_w^\vee & \xrightarrow{\varphi} & \Lambda_w \\ {\scriptstyle s}\downarrow & & \uparrow{\scriptstyle t} \\ \Lambda_w'^\vee & \xrightarrow{\varphi'} & \Lambda_w' \end{array}$$

where the map t is defined by $t(f'_\alpha) = d_\alpha f_\alpha$ for all $\alpha \in \Pi'$ and the maps φ and φ' are defined in Proposition 5.2.

It suffices to prove that $\operatorname{Im}(t \circ \varphi' \circ s - \varphi) \subset \Lambda_r$.

Consider the other diagram

$$\begin{array}{ccc} \Lambda_w^\vee & \xrightarrow{\rho} & \Lambda_w \\ {\scriptstyle t^\vee}\uparrow & & \uparrow{\scriptstyle t} \\ \Lambda_w'^\vee & \xrightarrow{\rho'} & \Lambda_w' \end{array}$$

where $t^\vee(f_\alpha'^\vee) = f_\alpha^\vee$ for all $\alpha \in \Pi'$. This diagram is commutative. Indeed,

$$(\rho \circ t^\vee)(f_\alpha'^\vee) = \rho(f_\alpha^\vee) = d_\alpha f_\alpha = t(f'_\alpha) = (t \circ \rho')(f_\alpha'^\vee),$$

where the second equality is by Example 5.4. (Recall that the root system R' of G' is simply laced, hence $\rho'(f_\alpha'^\vee) = f'_\alpha$.)

We claim that

$$(t^\vee \circ s)(x) - x \in \operatorname{span}(\Phi^\vee) + \Lambda_r^\vee$$

for every $x \in \Lambda_w^\vee$, where $\Phi^\vee := \{f_\alpha^\vee \mid \alpha \in \Pi_r\}$. Indeed, in the notation of (9.3) we have

$$\begin{aligned} (t^\vee \circ s)(x) - x = t^\vee(x') - x &= t^\vee(x') - x' - \sum_{\alpha\in\Pi_r} a_\alpha \alpha^\vee \\ &= t^\vee\Big(\sum_{\alpha\in\Pi'} b_\alpha f_\alpha'^\vee\Big) - \sum_{\alpha\in\Pi} b_\alpha f_\alpha^\vee - \sum_{\alpha\in\Pi_r} a_\alpha \alpha^\vee \\ &= -\sum_{\alpha\in\Pi_r} b_\alpha f_\alpha^\vee - \sum_{\alpha\in\Pi_r} a_\alpha\alpha^\vee \in \operatorname{span}(\Phi^\vee) + \Lambda_r^\vee. \end{aligned}$$

It follows from the claim that

$$(t\circ\rho'\circ s)(x) - \rho(x) = (\rho\circ t^\vee\circ s)(x) - \rho(x) = \rho((t^\vee\circ s)(x) - x) \in \rho(\operatorname{span}(\Phi^\vee) + \Lambda_r^\vee).$$

As $\rho(f_\alpha^\vee) = d_\alpha f_\alpha \in \Lambda_r$ for all $f_\alpha \in \Phi$, this is contained in Λ_r, proving the claim. □

Lemmas 9.1 and 9.2 yield:

Corollary 9.4. *The element t_G° is equal to $\sum_i d_i \cdot j_i^{\circ *}(t_{G'_i}^\circ)$.*

Lemma 9.5. *The diagram*

$$\begin{tikzcd} H^1(F,G') \arrow[r, equal] \arrow[d] & \prod_i H^1(F,G'_i) \arrow[d, "\sum d_i \cdot r_{G'_i}"] \\ H^1(F,G) \arrow[r, "r_G"] & H^3(F,\mathbb{Q}/\mathbb{Z}(2)) \end{tikzcd}$$

commutes.

Proof. The composition

$$H^1(F,G'_i) \to H^1(F,G) \xrightarrow{r_G} H^3(F,\mathbb{Q}/\mathbb{Z}(2))$$

coincides with the k-th multiple of the Rost invariant $r_{G'_i}$, where k is the order of the cokernel of the map $Q(G) \to Q(G'_i)$ of infinite cyclic groups generated by positive definite quadratic forms q_G and $q_{G'_i}$ on the lattices of coroots normalized so that the forms have value 1 on the short coroots (see [Garibaldi et al. 2003]). Recall that all coroots of G'_i have the same length, hence $q_{G'_i}$ has value 1 on all coroots of G'_i. Therefore, k coincides with d_i, the square-length of all coroots of G'_i viewed as coroots of G. □

Write each $G'_i = R_{L_i/F}(H_i)$ for L_i a separable field extension of F and H_i a simply connected absolutely simple algebraic group of type A over L_i. Theorem 1.2 is proved for such groups in Section 8. By Lemma 7.3, Theorem 1.2 holds for the group G'_i and hence for G'.

Let $x \in H^1(F,C)$ and let $y \in H^1(F,G)$, $\prod x'_i \in H^1(F,C') = \prod H^1(F,C'_i)$ and $\prod y'_i \in \prod H^1(F,G'_i)$ denote its images under the natural maps. We find

$$\begin{aligned} r_G(y) &= \sum_i d_i \cdot r_{G'_i}(y_i) && \text{by Lemma 9.5} \\ &= \sum_i d_i \cdot (-t^\circ_{G'_i} \cup x'_i) && \text{by the main theorem for all } G'_i \\ &= \sum_i d_i \cdot j_i^{\circ *}(-t^\circ_{G'_i}) \cup x && \text{by Lemma 2.16} \\ &= -t^\circ_G \cup x && \text{by Corollary 9.4.} \end{aligned}$$

This completes the proof of Theorem 1.2.

10. Concrete formulas

The explicit formulas for the restriction of the Rost invariant to the center given in [Merkurjev et al. 2002; Garibaldi and Quéguiner-Mathieu 2007] (for F of good characteristic) relied on an ad hoc formula for a pairing $C \otimes C \to \mathbb{Q}/\mathbb{Z}(2)$ depending on the type of G. In this section, we deduce those formulas from Theorem 1.2; as a consequence we find that those formulas hold regardless of char(F).

10A. ***The pairing induced by ρ.*** The map ρ defines a bilinear pairing $\Delta^\vee \otimes \Delta^\vee \to \mathbb{Q}/\mathbb{Z}$ via

$$\Delta^\vee \otimes \Delta^\vee \xrightarrow{\mathrm{id}\otimes\rho} \Delta^\vee \otimes \Delta \to \mathbb{Q}/\mathbb{Z}. \tag{10.1}$$

We now determine this pairing for each simple root system R.

For R with different root lengths, ρ is zero and hence (10.1) is zero unless $R = C_n$ for odd $n \geq 3$. In that case (and also for $R = E_7$), $\Delta \simeq \mathbb{Z}/2 \simeq \Delta^\vee$ and ρ is the unique isomorphism, so (10.1) amounts to the product map $x \otimes y \mapsto xy$. Therefore we may assume that R has only one root length.

If $\Delta^\vee$ is cyclic, we pick a fundamental dominant weight $f_i^\vee$ that generates $\Delta^\vee$ and the pairing (10.1) is determined by the image of $f_i^\vee \otimes f_i^\vee$. The image of this under $\mathrm{id} \otimes \rho$ is $f_i^\vee \otimes f_i$ as in Example 5.4, so the image in $\mathbb{Q}/\mathbb{Z}$ is the same as that of the coefficient of the simple root α_i appearing in the expression for f_i in terms of simple roots, for which we refer to [Bourbaki 2002].

For $R = A_n$, we have $\Delta^\vee \simeq \mathbb{Z}/(n+1)$ generated by $f_1^\vee$ and $f_1^\vee \otimes f_1^\vee \mapsto n/(n+1)$, cf. Example 5.9.

For $R = D_n$ for odd $n > 4$, $\Delta^\vee \simeq \mathbb{Z}/4$ generated by $f_n^\vee$ and $f_n^\vee \otimes f_n^\vee \mapsto n/4$.

For $R = E_6$, we have $\Delta^\vee \simeq \mathbb{Z}/3$ generated by $f_1^\vee$ and $f_1^\vee \otimes f_1^\vee \mapsto 1/3$.

For $R = D_n$ for even $n \geq 4$, $\Delta^\vee$ is isomorphic to $\mathbb{Z}/2 \oplus \mathbb{Z}/2$ generated by $f_{n-1}^\vee$, $f_n^\vee$. The tables show that $f_{n-1}^\vee \otimes f_{n-1}^\vee$ and $f_n^\vee \otimes f_n^\vee$ map to $n/4$ whereas $f_n^\vee \otimes f_{n-1}^\vee$ and $f_{n-1}^\vee \otimes f_n^\vee$ map to $(n-2)/4$. That is, viewing (10.1) as a bilinear form on $\mathbb{F}_2 \oplus \mathbb{F}_2$, for $n \equiv 0 \bmod 4$ it is the wedge product (which is hyperbolic) and for $n \equiv 2 \bmod 4$ it is the unique (up to isomorphism) metabolic form that is not hyperbolic.

10B. ***The cup product on C.*** Let G be a simple simply connected algebraic group over F with center C. The pairing (10.1) reads as follows:

$$C_* \otimes C_* \xrightarrow{\mathrm{id}\otimes\rho} C_* \otimes C^* \to \mathbb{Q}/\mathbb{Z}.$$

Twisting (tensoring with $\mathbb{Z}(1)\overset{L}{\otimes}\mathbb{Z}(1)$) we get a composition

$$C_*(1) \overset{L}{\otimes} C_*(1) \to C_*(1) \overset{L}{\otimes} C^*(1) \to \mathbb{Q}/\mathbb{Z}(2),$$

where the second map was already defined in (2.9). Therefore, we have a pairing

$$H^1(F, C) \otimes H^2(F, C) \to H^1(F, C) \otimes H^2(F, C^\circ) \to H^3(F, \mathbb{Q}/\mathbb{Z}(2)), \tag{10.2}$$

which we denote by $\bullet$. In this language, Theorem 1.2 says that

$$r_G i^*(x) = -x \bullet t_G \quad \text{for } x \in H^1(F, C). \tag{10.3}$$

Combining this observation with the computation of (10.1) recovers the formulas given in [Merkurjev et al. 2002; Garibaldi and Quéguiner-Mathieu 2007], with no restriction on char(F).

Example 10.4. Suppose G has inner type D_n for some $n \geq 4$. Then G is isomorphic to $\mathbf{Spin}(A, \sigma, f)$ for some central simple algebra A with quadratic pair (σ, f) such that the (even) Clifford algebra of (A, σ, f) is a product $C_+ \times C_-$, see [Knus et al. 1998, 26.15]. Put μ_2 for the kernel of the map $\mathbf{Spin}(A, \sigma, f) \to \mathbf{SO}(A, \sigma, f)$ and write i_2 for the inclusion $\mu_2 \hookrightarrow G$. (The highest weights of the representations $\mathbf{Spin}(A, \sigma, f) \to \mathbf{GL}_1(C_\varepsilon)$ for $\varepsilon = \pm$ both restrict to the nonzero character on $i_2(\mu_2)$.)

We claim that, *for $z \in H^1(F, \mu_2)$, the equalities*

$$r_G i_2^*(z) = \begin{cases} z \cup [A] & \textit{if } n \textit{ even,} \\ z \cup [C_+] & \textit{if } n \textit{ odd,} \end{cases} \tag{10.5}$$

hold in $H^3(F, \mathbb{Z}/2\mathbb{Z}(2))$. This can be seen by combining (10.3) with the calculations in Section 10A. Alternatively, arguing as in the beginning of Section 9, it suffices to verify (10.5) in case the variety X has an F-point, where we may check the equality via Lemma 9.5 on the subgroup G'. Then Equation 12.2 of [Garibaldi and Quéguiner-Mathieu 2007] settles the n even case, and an analogous computation handles n odd. Note that for n odd, one could also write $z \cup [C_-]$ in (10.5), as $[C_-] = 3[C_+]$ and $2z = 0$.

Example 10.6. The exact sequence $1 \to C \xrightarrow{i} G \to G/C \to 1$ gives a connecting homomorphism $\partial : (G/C)(F) \to H^1(F, C)$. It follows from (10.3) that, for $y \in (G/C)(F)$, $\partial(y) \bullet t_G = r_G i^* \partial(y) = 0$, i.e.,

$$(\operatorname{im} \partial) \bullet t_G = 0 \quad \text{in } H^3(F, \mathbb{Q}/\mathbb{Z}(2)). \tag{10.7}$$

For G of inner type A_{n-1}, G is isomorphic to $\mathbf{SL}_1(A)$ for a central simple algebra A and we may identify $\operatorname{im} \partial$ with $\operatorname{Nrd}(A^\times) \subseteq H^1(F, \mu_n)$. In this case, (10.7) says: *If $x \in \operatorname{Nrd}(A^\times)$, then $(x) \cup [A] = 0$.*

For G of type C_n, G is isomorphic to $\mathbf{Sp}(A, \sigma)$ for a central simple algebra A with symplectic involution σ and we may identify $\operatorname{im} \partial$ with the group $G(A, \sigma)$ of multipliers of similitudes of (A, σ). If n is even, (10.7) is an empty claim because $\bullet$ is identically zero. If n is odd, (10.7) says that $G(A, \sigma) \cup [A] = 0$, i.e., since A is Brauer-equivalent to a quaternion algebra, $G(A, \sigma) \subseteq \operatorname{Nrd}(A^\times)$; this is proved in [Knus et al. 1998, 12.22].

References

[Bourbaki 2002] N. Bourbaki, *Lie groups and Lie algebras. Chapters 4–6*, Springer, Berlin, 2002. MR Zbl

[Conrad et al. 2015] B. Conrad, O. Gabber, and G. Prasad, *Pseudo-reductive groups*, 2nd ed., New Mathematical Monographs **26**, Cambridge Univ. Press, 2015. MR Zbl

[Freitag and Kiehl 1988] E. Freitag and R. Kiehl, *Étale cohomology and the Weil conjecture*, Ergebnisse der Mathematik (3) **13**, Springer, Berlin, 1988. MR Zbl

[Garibaldi 2001] S. Garibaldi, “The Rost invariant has trivial kernel for quasi-split groups of low rank”, *Comment. Math. Helv.* **76**:4 (2001), 684–711. MR Zbl

[Garibaldi 2012] S. Garibaldi, “Outer automorphisms of algebraic groups and determining groups by their maximal tori”, *Michigan Math. J.* **61**:2 (2012), 227–237. MR Zbl

[Garibaldi and Quéguiner-Mathieu 2007] S. Garibaldi and A. Quéguiner-Mathieu, “Restricting the Rost invariant to the center”, *Algebra i Analiz* **19**:2 (2007), 52–73. Reprinted in *St. Petersburg Math. J.* **19**:2 (2008), 197–213. MR Zbl

[Garibaldi et al. 2003] S. Garibaldi, A. Merkurjev, and J.-P. Serre, *Cohomological invariants in Galois cohomology*, University Lecture Series **28**, American Mathematical Society, Providence, RI, 2003. MR Zbl

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

[Gille and Quéguiner-Mathieu 2011] P. Gille and A. Quéguiner-Mathieu, “Formules pour l’invariant de Rost”, *Algebra Number Theory* **5**:1 (2011), 1–35. MR Zbl

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics **101**, Cambridge Univ. Press, 2006. MR Zbl

[Kahn 1996] B. Kahn, “Applications of weight-two motivic cohomology”, *Doc. Math.* **1** (1996), 395–416. MR Zbl

[Kato 1982] K. Kato, “Galois cohomology of complete discrete valuation fields”, pp. 215–238 in *Algebraic K-theory, II* (Oberwolfach, 1980), edited by R. K. Dennis, Lecture Notes in Math. **967**, Springer, Berlin, 1982. MR Zbl

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, Amer. Math. Soc. Colloquium Publications **44**, American Mathematical Society, Providence, RI, 1998. MR Zbl

[Lichtenbaum 1987] S. Lichtenbaum, “The construction of weight-two arithmetic cohomology”, *Invent. Math.* **88**:1 (1987), 183–215. MR Zbl

[Lichtenbaum 1990] S. Lichtenbaum, “New results on weight-two motivic cohomology”, pp. 35–55 in *The Grothendieck Festschrift, III*, edited by P. Cartier et al., Progr. Math. **88**, Birkhäuser, Boston, 1990. MR Zbl

[Merkurjev 1999] A. Merkurjev, “Invariants of algebraic groups”, *J. Reine Angew. Math.* **508** (1999), 127–156. MR Zbl

[Merkurjev 2016] A. Merkurjev, “Weight two motivic cohomology of classifying spaces for semi-simple groups”, *Amer. J. Math.* **138**:3 (2016), 763–792. MR Zbl

[Merkurjev and Tignol 1995] A. S. Merkurjev and J.-P. Tignol, “The multipliers of similitudes and the Brauer group of homogeneous varieties”, *J. Reine Angew. Math.* **461** (1995), 13–47. MR Zbl

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

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton Univ. Press, 1980. MR Zbl

[Steinberg 1963] R. Steinberg, “Representations of algebraic groups”, *Nagoya Math. J.* **22** (1963), 33–56. MR Zbl

[Tits 1966] J. Tits, “Classification of algebraic semisimple groups”, pp. 33–62 in *Algebraic groups and discontinuous subgroups* (Boulder, CO, 1965), edited by A. Borel and G. D. Mostow, Proc. Sympos. Pure Math. **9**, American Mathematical Society, Providence, RI, 1966. MR Zbl

[Tits 1971] J. Tits, “Représentations linéaires irréductibles d’un groupe réductif sur un corps quelconque”, *J. Reine Angew. Math.* **247** (1971), 196–220. MR Zbl

Received September 26, 2016. Revised February 8, 2017.

SKIP GARIBALDI
IDA CENTER FOR COMMUNICATIONS RESEARCH
SAN DIEGO, CA 92121
UNITED STATES
skip@member.ams.org

ALEXANDER S. MERKURJEV
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
LOS ANGELES, CA 90095-1555
UNITED STATES
merkurev@math.ucla.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.399

MODULI SPACES OF RANK 2 INSTANTON SHEAVES ON THE PROJECTIVE SPACE

MARCOS JARDIM, MARIO MAICAN AND ALEXANDER S. TIKHOMIROV

We study the irreducible components of the moduli space of instanton sheaves on $\mathbb{P}^3$, that is, μ-semistable rank 2 torsion-free sheaves E with $c_1(E) = c_3(E) = 0$ satisfying $h^1(E(-2)) = h^2(E(-2)) = 0$. In particular, we classify all instanton sheaves with $c_2(E) \leq 4$, describing all the irreducible components of their moduli space. A key ingredient for our argument is the study of the moduli space $\mathcal{T}(d)$ of stable sheaves on $\mathbb{P}^3$ with Hilbert polynomial $P(t) = d \cdot t$, which contains, as an open subset, the moduli space of rank 0 instanton sheaves of multiplicity d; we describe all the irreducible components of $\mathcal{T}(d)$ for $d \leq 4$.

1. Introduction

Instanton bundles on $\mathbb{CP}^3$ were introduced by Atiyah, Drinfeld, Hitchin and Manin in the late 1970s as the holomorphic counterparts, via twistor theory, to anti-self-dual connections with finite energy (instantons) on the four-dimensional round sphere S^4. To be more precise, an *instanton bundle of charge n* is a μ-stable rank 2 bundle E on $\mathbb{P}^3$ with $c_1(E) = 0$ and $c_2(E) = n$ satisfying the cohomological condition $h^1(E(-2)) = 0$; equivalently, an instanton bundle of charge n is a locally free sheaf which arises as cohomology of a linear monad of the form

$$(1) \qquad 0 \to n \cdot \mathcal{O}_{\mathbb{P}^3}(-1) \to (2+2n) \cdot \mathcal{O}_{\mathbb{P}^3} \to n \cdot \mathcal{O}_{\mathbb{P}^3}(1) \to 0.$$

The moduli space $\mathcal{I}(n)$ of such objects has been thoroughly studied in the past thirty-five years by various authors and it is now known to be an irreducible [Tikhomirov 2012; 2013], nonsingular [Jardim and Verbitsky 2014] affine [Costa and Ottaviani 2003] variety of dimension $8n - 3$.

The closure of $\mathcal{I}(n)$ within the moduli space $\mathcal{M}(n)$ of semistable rank 2 sheaves with Chern classes $c_1 = 0$, $c_2 = n$ and $c_3 = 0$ contains nonlocally free sheaves which also arise as cohomology of monads of the form (1). Such *instanton sheaves* can

MSC2010: 14D20, 14J60.
Keywords: Moduli spaces, semistable sheaves, instanton sheaves, sheaves on projective space.

alternatively be defined as rank 2 torsion-free sheaves satisfying the cohomological conditions

$$h^0(E(-1)) = h^1(E(-2)) = h^2(E(-2)) = h^3(E(-3)) = 0.$$

We prove that such sheaves are always stable (see Theorem 4 below), so they admit a moduli space $\mathcal{L}(n)$ regarded as an open subset of $\mathcal{M}(n)$ which, of course, contains $\mathcal{I}(n)$.

The spaces $\mathcal{L}(1)$ and $\mathcal{L}(2)$ were essentially known to be irreducible, see details in the first few paragraphs of Section 3 below. However, $\mathcal{L}(3)$ was observed to have at least two irreducible components [Jardim et al. 2015, Remark 8.6], while several new components of $\mathcal{L}(n)$ were constructed in [Jardim et al. 2017].

The main goal of this paper is to characterize the irreducible components of $\mathcal{L}(3)$ and $\mathcal{L}(4)$. We prove:

Main Theorem 1. (i) $\mathcal{L}(3)$ *is a connected quasiprojective variety consisting of exactly two irreducible components each of dimension* 21;

(ii) $\mathcal{L}(4)$ *is a connected quasiprojective variety consisting of exactly four irreducible components*, *three of dimension* 29 *and one of dimension* 32.

For every instanton sheaf E, the quotient $E^{\vee\vee}/E$ is a semistable sheaf with Hilbert polynomial $d\cdot(t+2)$, see Section 2 below. Therefore, an essential ingredient for the proof of Main Theorem 1 is the study of the moduli space $\mathcal{T}(d)$ of semistable sheaves on $\mathbb{P}^3$ with Hilbert polynomial $P(t) = d\cdot t$. Since these spaces are also interesting in their own right, we prove:

Main Theorem 2. (i) $\mathcal{T}(1)$ *is an irreducible projective variety of dimension* 5;

(ii) $\mathcal{T}(2)$ *is a connected projective variety consisting of exactly two irreducible components of dimension* 8;

(iii) $\mathcal{T}(3)$ *is a projective variety consisting of exactly four irreducible components*, *two of dimension* 12 *and two of dimension* 13;

(iv) $\mathcal{T}(4)$ *is a projective variety consisting of exactly eight irreducible components*, *four of dimension* 16, *two of dimension* 17, *one of dimension* 18 *and one of dimension* 20.

We also give a precise description of a generic point in each of the irreducible components mentioned in the statement of the theorem, see Section 4.

2. Stability of instanton sheaves

Recall from [Jardim 2006] that a torsion-free sheaf E on $\mathbb{P}^3$ is called an *instanton sheaf* if $c_1(E) = 0$ and the following cohomological conditions hold:

$$h^0(E(-1)) = h^1(E(-2)) = h^2(E(-2)) = h^3(E(-3)) = 0.$$

The integer $n := -\chi(E(-1))$ is called the charge of E; it is easy to check that $n = h^1(E(-1)) = c_2(E)$, and that $c_3(E) = 0$. The trivial sheaf $r \cdot \mathcal{O}_{\mathbb{P}^3}$ of rank r is considered as an instanton sheaf of charge zero. In this paper, we will only be interested in rank 2 instanton sheaves.

Recall that the singular locus $\mathrm{Sing}(G)$ of a coherent sheaf G on a nonsingular projective variety X is given by

$$\mathrm{Sing}(G) := \{x \in X \mid G_x \text{ is not free over } \mathcal{O}_{X,x}\},$$

where G_x denotes the stalk of G at a point x and $\mathcal{O}_{X,x}$ is its local ring. The following result, proved in [Gargate and Jardim 2016, Main Theorem], provides a key piece of information regarding the singular loci of rank 2 instanton sheaves.

Theorem 1. *If E is a nonlocally free instanton sheaf of rank 2 on $\mathbb{P}^3$, then*

(i) *its singular locus has pure dimension 1;*

(ii) *$E^{\vee\vee}$ is a (possibly trivial) locally free instanton sheaf.*

Remark 2. In fact, the quotient sheaf $Q_E := E^{\vee\vee}/E$ is a *rank 0 instanton sheaf*, in the sense of [Hauzer and Langer 2011, Section 6.1]; see also [Gargate and Jardim 2016, Section 3.2]. More precisely, a *rank* 0 *instanton sheaf* is a coherent sheaf Q on $\mathbb{P}^3$ such that $h^0(Q(-2)) = h^1(Q(-2)) = 0$; the integer $d := h^0(Q(-1))$ is called the *multiplicity* of Q.

The Hilbert polynomial of a rank 2 instanton sheaf E (in fact, of any coherent sheaf on $\mathbb{P}^3$ of rank 2 with $c_1 = 0,\ c_2 = n$ and $c_3 = 0$) is given by

$$P_E(t) = \tfrac{1}{3}(t+3)\cdot(t+2)\cdot(t+1) - n\cdot(t+2) = 2\cdot\chi(\mathcal{O}_{\mathbb{P}^3}(t)) - n\cdot(t+2). \tag{2}$$

Let $n' := c_2(E^{\vee\vee}) \geq 0$, then it follows from the standard sequence

$$0 \to E \to E^{\vee\vee} \to Q_E \to 0 \tag{3}$$

that

$$P_{Q_E}(t) = d\cdot(t+2) \quad \text{where } d := n - n'.$$

Note that $d = n - n'$ is precisely the multiplicity of Q_E as a rank 0 instanton sheaf.

Rank 0 instanton sheaves can be characterized in the following way.

Proposition 3. *Every rank* 0 *instanton sheaf Q admits a resolution of the form*

$$0 \to d\cdot\mathcal{O}_{\mathbb{P}^3}(-1) \to 2d\cdot\mathcal{O}_{\mathbb{P}^3} \to d\cdot\mathcal{O}_{\mathbb{P}^3}(1) \to Q \to 0. \tag{4}$$

Proof. Consider the Beilinson spectral sequence from [Choi et al. 2016, Section 6], applied to the sheaf $Q' := Q(-2)$. We have $H^0(Q') = 0$, and therefore also $H^0(Q' \otimes \Omega^1_{\mathbb{P}^3}(1)) = 0$ and $H^0(Q'(-1)) = 0$. We adopt the notations of [Choi et al. 2016, Section 6]. Since $\mathrm{Ker}(\varphi_5)/\mathrm{Im}(\varphi_4) = 0$, we deduce that $H^0(Q' \otimes \Omega^2_{\mathbb{P}^3}(2)) = 0$. Thus, the bottom row of the E^1-term of the spectral sequence vanishes. Since φ_7 is an isomorphism, we deduce that φ_1 is injective. Since φ_8 is injective, we deduce that $\operatorname{Ker}(\varphi_2) = \operatorname{Im}(\varphi_1)$. The top row of the E^1-term of the spectral sequence yields the resolution

$$0 \to H^1(Q'(-1)) \otimes \mathcal{O}_{\mathbb{P}^3}(-3) \xrightarrow{\varphi_1} H^1(Q' \otimes \Omega^2_{\mathbb{P}^3}(2)) \otimes \mathcal{O}_{\mathbb{P}^3}(-2) \xrightarrow{\varphi_2} H^1(Q' \otimes \Omega^1_{\mathbb{P}^3}(1)) \otimes \mathcal{O}_{\mathbb{P}^3}(-1) \to Q' \to 0.$$

We have

$$\chi(Q' \otimes \Omega^1_{\mathbb{P}^3}(1)) = -d, \qquad \chi(Q' \otimes \Omega^2_{\mathbb{P}^3}(2)) = -2d, \qquad \chi(Q'(-1)) = -d,$$

hence

$$h^1(Q' \otimes \Omega^1_{\mathbb{P}^3}(1)) = d, \qquad h^1(Q' \otimes \Omega^2_{\mathbb{P}^3}(2)) = 2d, \qquad h^1(Q'(-1)) = d.$$

The above exact sequence yields (4). □

Let now examine the stability properties of instanton sheaves.

Theorem 4. *Every nontrivial rank* 2 *instanton sheaf* E *is stable. In addition, a nontrivial instanton sheaf* E *is* μ*-stable if and only if its double dual* $E^{\vee\vee}$ *is nontrivial.*

Proof. Since rank 2 instanton sheaves have no global sections [Jardim 2006, Proposition 11], every nontrivial locally free rank 2 instanton sheaf is μ-stable; therefore, if $E^{\vee\vee}$ is nontrivial, then E is also μ-stable. Conversely, if E is μ-stable, then so is $E^{\vee\vee}$, hence it must be nontrivial.

Therefore, in order to prove the first claim of the Theorem, it is enough to consider *quasitrivial instanton sheaves*, i.e., rank 2 instanton sheaves E with $E^{\vee\vee} \simeq 2 \cdot \mathcal{O}_{\mathbb{P}^3}$; note that the multiplicity of Q_E is exactly $n = c_2(E)$.

Since E has no global sections, it can only be destabilized by the ideal sheaf I_C of a subscheme $C \subset \mathbb{P}^3$. Moreover, we can assume that the quotient sheaf E/I_C is torsion-free, thus it is also the ideal sheaf I_D of another subscheme $D \subset \mathbb{P}^3$. We obtain two exact sequences

$$\begin{tikzcd}
 & 0 \arrow[d] & & & \\
 & I_C \arrow[d] & & & \\
0 \arrow[r] & E \arrow[r] \arrow[d] & 2 \cdot \mathcal{O}_{\mathbb{P}^3} \arrow[r] & Q_E \arrow[r] & 0 \\
 & I_D \arrow[d] & & & \\
 & 0 & & &
\end{tikzcd}$$

Taking the double dual of the top vertical morphisms we obtain, also using the Snake Lemma, the following commutative diagram:

(5)
$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & I_C \arrow[r] \arrow[d] & \mathcal{O}_{\mathbb{P}^3} \arrow[r] \arrow[d] & \mathcal{O}_C \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & E \arrow[r] \arrow[d] & 2\cdot\mathcal{O}_{\mathbb{P}^3} \arrow[r] \arrow[d] & Q_E \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & I_D \arrow[r] \arrow[d] & \mathcal{O}_{\mathbb{P}^3} \arrow[r] \arrow[d] & \mathcal{O}_D \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

Since $h^0(Q_E(-2)) = 0$, then $h^0(\mathcal{O}_C(-2)) = 0$ also, hence C must have pure dimension 1. Moreover, note also that $h^1(Q_E(-2)) = 0$ implies $h^1(\mathcal{O}_D(-2)) = 0$.

We show that $\dim D = 0$. Indeed, assume that D has dimension 1. Let U be the maximal zero-dimensional subsheaf of $\mathcal{O}_D$, and set $\mathcal{O}_{D'} := \mathcal{O}_D/U$; clearly, D' has pure dimension 1. Next, let $D'' := D'_{\mathrm{red}}$ be the underlying reduced scheme. We end up with two exact sequences

$$0 \to U \to \mathcal{O}_D \to \mathcal{O}_{D'} \to 0 \quad \text{and} \quad 0 \to T \to \mathcal{O}_{D'} \to \mathcal{O}_{D''} \to 0,$$

so that the vanishing of $h^1(\mathcal{O}_D(-2))$ forces $h^1(\mathcal{O}_{D''}(-2)) = 0$.

Still, D'' may be reducible, so let $D'' := D''_1 \cup \cdots \cup D''_p$ be its decomposition into irreducible components. For each index $j = 1, \ldots, p$ we obtain a sequence,

$$0 \to S_j \to \mathcal{O}_{D''} \to \mathcal{O}_{D''_j} \to 0,$$

thus also $h^1(\mathcal{O}_{D''_j}(-2)) = 0$. Let d_j and p_j denote the degree and arithmetic genus of D''_j, respectively. It follows that

$$0 \le h^0(\mathcal{O}_{D''_j}(-2)) = \chi(\mathcal{O}_{D''_j}(-2)) = -2d_j + 1 - p_j,$$

thus $p_j \le -2d_j + 1 \le -1$, which is impossible for a reduced and irreducible curve.

Now let $\delta = h^0(\mathcal{O}_D)$ be the length of D. Since $\deg(C) = n$, we have

$$P_{I_C}(t) = \chi(\mathcal{O}_{\mathbb{P}^3}(t)) - \chi(\mathcal{O}_C(t)) = \chi(\mathcal{O}_{\mathbb{P}^3}(t)) - nt + (\delta - 2n).$$

Comparing with equation (2), we have

$$\frac{P_E(t)}{2} - P_{I_C}(t) = \frac{n}{2}t + n - \delta, \tag{6}$$

which is positive for t sufficiently large, and therefore E contains no destabilizing subsheaves. □

As a consequence of the proof above, we obtain the following interesting fact.

Corollary 5. *Every rank* 2 *quasitrivial instanton sheaf of charge n on $\mathbb{P}^3$ is an extension of the ideal of a zero-dimensional scheme D by the ideal of a pure one-dimensional scheme C of degree n containing D. In addition,*

$$\chi(\mathcal{O}_C) = 2n - h^0(\mathcal{O}_D).$$

On the other hand, it is easy to check that every rank 0 instanton sheaf is semistable.

Lemma 6. *Every rank* 0 *instanton sheaf is semistable.*

Proof. Let Z be a rank 0 instanton sheaf, and let T be a subsheaf of Z with Hilbert polynomial $P_T(t) = a \cdot t + \chi(T)$. Since $h^0(Z(-2)) = 0$, then $h^0(T(-2)) = 0$ and $-2a + \chi(T) = -h^1(T(-2)) \leq 0$. It follows that

$$\frac{\chi(T)}{a} \leq 2 = \frac{\chi(Z)}{d}. \qquad \square$$

Clearly, not every rank 0 instanton sheaf is stable: if Q_1 and Q_2 are rank 0 instanton sheaves, then so is any extension of Q_1 by Q_2, and this cannot possibly be stable.

Conversely, there are semistable sheaves with Hilbert polynomial $dt + 2d$ which are not rank 0 instanton sheaves: just consider $Q := \mathcal{O}_\Sigma(2)$ for an elliptic curve $\Sigma \hookrightarrow \mathbb{P}^3$, so that $h^0(Q(-2)) \neq 0$.

3. Moduli space of instanton sheaves

Let $\mathcal{L}(n)$ denote the open subscheme of the Maruyama moduli space $\mathcal{M}(n)$ of semistable rank 2 torsion-free sheaves with Chern classes $c_1 = 0,\ c_2 = n$ and $c_3 = 0$ consisting of instanton sheaves of charge n. Let $\mathcal{I}(n)$ denote the open subscheme of $\mathcal{M}(n)$ consisting of locally free instanton sheaves. Finally, let $\overline{\mathcal{L}(n)}$ and $\overline{\mathcal{I}(n)}$ denote the closures within $\mathcal{M}(n)$ of $\mathcal{L}(n)$ and $\mathcal{I}(n)$, respectively. We also consider the set $\mathcal{I}^0(n) := \overline{\mathcal{I}(n)} \cap \mathcal{L}(n)$, which consists of those instanton sheaves which either are locally free, or can be deformed into locally free ones.

It was shown in [Tikhomirov 2012; 2013] that $\mathcal{I}(n)$ is irreducible for every $n > 0$; its closure $\overline{\mathcal{I}(n)}$ is called the *instanton component* of $\mathcal{M}(n)$. However, the same is not true for $\mathcal{L}(n)$ as soon as $n \geq 3$. Indeed, it is well known that

$$\overline{\mathcal{I}(1)} = \mathcal{L}(1) = \mathcal{M}(1) \simeq \mathbb{P}^5,$$

see for instance [Jardim et al. 2017, Section 6].

The case $n = 2$ has also been understood.

Proposition 7. $\overline{\mathcal{L}(2)} = \overline{\mathcal{I}(2)}$.

In particular, $\mathcal{L}(2)$ possesses a single irreducible component of dimension 13.

Proof. Le Potier [1993b] showed that $\mathcal{M}(2)$ has exactly 3 irreducible components; according to the description of these components provided in [Jardim et al. 2017, Section 6], only the instanton component $\overline{\mathcal{I}(2)}$ contains instanton sheaves. $\square$

Let us now describe the irreducible components of $\mathcal{L}(n)$ for $n \geq 3$ introduced in [Jardim et al. 2017, Section 3].

Let Σ be an irreducible, nonsingular, complete intersection curve in $\mathbb{P}^3$, given as the intersection of a surface of degree d_1 with a surface of degree d_2, with $1 \leq d_1 \leq d_2$; denote by $\iota : \Sigma \hookrightarrow \mathbb{P}^3$ the inclusion morphism. Choose $L \in \operatorname{Pic}^{g-1}(\Sigma)$ such that $h^0(\Sigma, L) = h^1(\Sigma, L) = 0$. Given a (possibly trivial) locally free instanton sheaf F of charge $c \geq 0$ and an epimorphism $\varphi : F \twoheadrightarrow (\iota_* L)(2)$, the kernel $E := \operatorname{Ker} \varphi$ is an instanton sheaf of charge $c + d_1 d_2$. Thus we may consider the set

$$\mathcal{C}(d_1, d_2, c) := \big\{[E] \in \mathcal{M}(c + d_1 d_2) \mid E^{\vee\vee} \in \mathcal{I}(c),\, E^{\vee\vee}/E \simeq (\iota_* L)(2)\big\} \tag{7}$$

as a subvariety of $\mathcal{M}(c + d_1 d_2)$. The following result is proved in [Jardim et al. 2017], see Theorems 15, 17 and 23 of that paper.

Theorem 8. *For each $c \geq 0$ and $1 \leq d_1 \leq d_2$ such that $(d_1, d_2) \neq (1, 1), (1, 2)$, $\overline{\mathcal{C}(d_1, d_2, c)}$ is an irreducible component of $\mathcal{M}(c + d_1 d_2)$ of dimension*

$$\dim \overline{\mathcal{C}(d_1, d_2, c)} = 8c - 3 + \tfrac{1}{2} d_1 d_2 (d_1 + d_2 + 4) + h, \tag{8}$$

where

$$h = \begin{cases} 2\binom{d_1+3}{3} - 4 & \text{if } d_1 = d_2, \\ \binom{d_1+3}{3} + \binom{d_2+3}{3} - \binom{d_2-d_1+3}{3} - 2 & \text{if } d_1 < d_2. \end{cases}$$

In addition, $\overline{\mathcal{C}(d_1, d_2, c)} \cap \mathcal{I}^0(c + d_1 d_2) \neq \varnothing$.

We do not know whether the families $\mathcal{C}(d_1, d_2, c)$ exhaust all components of $\mathcal{L}(n)$, but we prove that this holds for $n = 3$ and 4 in Sections 5 and 6 below, respectively.

However, we remark that the previous result allows for a partial count of the number of components of $\mathcal{L}(n)$. Indeed, let $\tau(n)$ denote the number of irreducible components of the union

$$\overline{\mathcal{I}(n)} \bigcup \Bigg(\bigcup_{d_1 d_2 + c = n} \overline{\mathcal{C}(d_1, d_2, c)} \Bigg).$$

To estimate $\tau(n)$, we must count the different ways in which an integer $n \geq 3$ can be written as $n = d_1 d_2 + c$ with $c \geq 0$, and $1 \leq d_1 \leq d_2$ excluding the pairs $(d_1, d_2) = (1, 1), (1, 2)$. Consider the function

$$\delta(p) = \begin{cases} \frac{1}{2}(d(p) + 1) & \text{if } p \text{ is a perfect square,} \\ \frac{1}{2} d(p) & \text{otherwise,} \end{cases}$$

where $d(p)$ is the *divisor function*, i.e., the number of divisors of a positive integer p, including p itself. Note that $\delta(p)$ is the number of different ways in which we can write p as a product $d_1 d_2$ with $1 \le d_1 \le d_2$. Adding the instanton component, we have

$$\tau(n) = 1 + \sum_{p=3}^{n} \delta(p) = \frac{1}{2}\Big(\sum_{p=3}^{n} d(p) + \lfloor\sqrt{n}\rfloor + 1 \Big), \tag{9}$$

since $\lfloor\sqrt{n}\rfloor - 1$ accounts for the number of perfect squares between 3 and n.

Lemma 9. *Let $l(n)$ be the number of irreducible components of the moduli space of instanton sheaves of charge n. Then, for n sufficiently large, $l(n) > \frac{1}{2} n \cdot \log(n)$.*

Proof. Determining the asymptotic behavior of the sum of divisors function is a relevant problem in number theory called the *Dirichlet divisor problem*; indeed, it is known that

$$\sum_{p=1}^{n} d(p) = n \cdot \log(n) + (2\gamma - 1)n + O(n^{\theta}),$$

where γ denotes the Euler–Mascheroni constant, and $1/4 \le \theta \le 131/416$, see [Huxley 2003]. Comparing with equation (9), we obtain the desired estimate. □

Also relevant for us is a class of instanton sheaves studied in [Jardim et al. 2015]; more precisely, for $n > 0$ and each $m = 1, \dots, n$, consider the subset $\mathcal{D}(m,n)$ of $\mathcal{M}(n)$ consisting of the isomorphism classes $[E]$ of the sheaves E obtained in this way:

$$\mathcal{D}(m,n) := \big\{ [E] \in \mathcal{M}(n) \mid [E^{\vee\vee}] \in \mathcal{I}(n-m),\ \Gamma = \operatorname{supp}(E^{\vee\vee}/E) \in \mathcal{R}_0^*(m)_{E^{\vee\vee}}, \\ \text{and } E^{\vee\vee}/E \simeq \mathcal{O}_\Gamma((2m-1)\mathrm{pt}) \big\},$$

where the space $\mathcal{R}_0^*(m)_{E^{\vee\vee}}$ is described as follows: first, let $\mathcal{R}_0^*(m)$ denote the space of nonsingular rational curves $\Gamma \hookrightarrow \mathbb{P}^3$ of degree m whose normal bundle $N_{\Gamma/\mathbb{P}^3}$ is given by $2 \cdot \mathcal{O}_\Gamma((2m-1)\mathrm{pt})$; then, for any instanton sheaf F we set

$$\mathcal{R}_0^*(m)_F := \{ \Gamma \in \mathcal{R}_0^*(m) \mid F|_\Gamma \simeq 2 \cdot \mathcal{O}_\Gamma \}.$$

One can show that for every rank 2 instanton sheaf F, the space $\mathcal{R}_0^*(m)_F$ is a nonempty open subset of $\mathcal{R}_0^*(m)$, see [Jardim et al. 2015, Lemma 6.2].

Let $\overline{\mathcal{D}(m,n)}$ denote the closure of $\mathcal{D}(m,n)$ within $\mathcal{M}(n)$. Note that since $E^{\vee\vee}$ is a locally free instanton sheaf of charge $n-m$, and $\mathcal{O}_\Gamma(2m-1)$ is a rank 0 instanton sheaf of degree m, then E is an instanton sheaf of charge n, so that $\mathcal{D}(m,n) \subset \mathcal{L}(n)$. In fact, it is shown in [Jardim et al. 2017, Theorem 7.8] that $\overline{\mathcal{D}(m,n)} \subset \mathcal{I}^0(n)$. In addition, we prove:

Proposition 10. *Let $\Gamma_1, \dots, \Gamma_r$ be disjoint, smooth irreducible rational curves in $\mathbb{P}^3$ of degrees $m_1, \dots, m_r$, respectively; set $Q := \bigoplus_{j=1}^{r} \mathcal{O}_{\Gamma_j}(-\mathrm{pt})$. If F is a locally free instanton sheaf of charge c such that $F|_{\Gamma_j} \simeq 2 \cdot \mathcal{O}_{\Gamma_j}$ for each $j = 1, \dots, r$, and $\varphi : F \twoheadrightarrow Q(2)$ is an epimorphism, then $[\ker\varphi] \in \mathcal{I}^0(c + m_1 + \dots + m_r)$.*

The proof of the previous proposition requires the following technical lemma, proved in [Jardim et al. 2015, Lemma 7.1].

Lemma 11. *Let C be a smooth irreducible curve with a marked point* 0, *and set* $\boldsymbol{B} := C \times \mathbb{P}^3$. *Let $\boldsymbol{F}$ and $\boldsymbol{G}$ be $\mathcal{O}_{\boldsymbol{B}}$-sheaves, flat over C and such that $\boldsymbol{F}$ is locally free along* $\operatorname{supp}(\boldsymbol{G})$. *Denote*

$$\boldsymbol{G}_t := \boldsymbol{G}|_{\{t\}\times\mathbb{P}^3} \quad \text{and} \quad \boldsymbol{F}_t = \boldsymbol{F}|_{\{t\}\times\mathbb{P}^3} \quad \text{for } t \in C.$$

Assume that, for each $t \in C$,

$$H^i(\mathcal{H}om(\boldsymbol{F}_t, \boldsymbol{G}_t)) = 0, \quad i \geq 1. \tag{10}$$

If $s : \boldsymbol{F}_0 \to \boldsymbol{G}_0$ is an epimorphism, then, after possibly shrinking C, s extends to an epimorphism $\boldsymbol{s} : \boldsymbol{F} \twoheadrightarrow \boldsymbol{G}$.

Proof of Proposition 10. We argue by induction on r; the case $r = 1$ is just the aforementioned result, namely [Jardim et al. 2017, Theorem 7.8].

Let $Q' := \bigoplus_{j=1}^{r-1} \mathcal{O}_{\Gamma_j}(-\mathrm{pt})$, so that $Q = Q' \oplus \mathcal{O}_{\Gamma_r}(-\mathrm{pt})$. Let $E := \ker\varphi$, and let E' denote the kernel of the composition $F \xrightarrow{\varphi} Q(2) \twoheadrightarrow Q'(2)$. We obtain the following exact sequence:

$$0 \to E \to E' \xrightarrow{\varphi'} \mathcal{O}_{\Gamma_r}((2m_r - 1)\mathrm{pt}) \to 0.$$

By the induction hypothesis, $[E']$ is in $\mathcal{I}^0(c+m_1+\cdots+m_{r-1})$, thus one can find an affine open subset $0 \in U \subset \mathbb{A}^1$ and a coherent sheaf $\boldsymbol{E}$ on $\mathbb{P}^3 \times U$, flat over U, such that $\boldsymbol{E}_0 = E'$ and $\boldsymbol{E}_t$ is a locally free instanton sheaf of charge $c + m_1 + \cdots + m_{r-1}$ satisfying $\boldsymbol{E}_t|_{\Gamma_r} \simeq 2 \cdot \mathcal{O}_{\Gamma_r}$ for every $t \in U \setminus \{0\}$. Setting $\boldsymbol{G} := \pi^*(Q/Q'(2))$ where $\pi : \mathbb{P}^3 \times U \to \mathbb{P}^3$ is the projection onto the first factor, note that

$$H^i(\mathcal{H}om(\boldsymbol{E}_t, \boldsymbol{G}_t)) = H^i(2 \cdot \mathcal{O}_{\Gamma_r}((2m_r - 1)\mathrm{pt})) = 0, \quad \text{for } i \geq 1 \text{ and } t \in U.$$

This claim is clear for $t \neq 0$; when $t = 0$, simply observe that the sequence $0 \to E' \to F \to Q'(2) \to 0$ implies that $E'|_{\Gamma_r} \simeq F|_{\Gamma_r}$, since the support of Q' is disjoint from Γ_r.

By Lemma 11, there exists an epimorphism $\boldsymbol{s} : \boldsymbol{E} \twoheadrightarrow \boldsymbol{G}$ extending $\varphi' : E' \to \mathcal{O}_{\Gamma_r}((2m_r - 1)\mathrm{pt})$, so that $[\ker \boldsymbol{s}_t] \in \mathcal{D}(m_r, c + m_1 + \cdots + m_r)$, by construction. It then follows that $[E] \in \overline{\mathcal{D}(m_r, c + m_1 + \cdots + m_r)}$, hence, by [Jardim et al. 2017, Theorem 7.8], $[E] \in \mathcal{I}^0(c + m_1 + \cdots + m_r)$, as desired. □

Next, we consider the following situation: let Σ be an irreducible, nonsingular, complete intersection curve in $\mathbb{P}^3$, given as the intersection surfaces of degrees d_1 and d_2, with $1 \leq d_1 \leq d_2$ and $(d_1, d_2) \neq (1, 1), (1, 2)$, and let Γ be a smooth irreducible rational curve in $\mathbb{P}^3$ of degree m disjoint from Σ. Set $Q := L \oplus \mathcal{O}_\Gamma(-\mathrm{pt})$ for some $L \in \operatorname{Pic}^{g-1}(\Sigma)$ such that $h^0(\Sigma, L) = h^1(\Sigma, L) = 0$, where g is the genus of Σ.

Proposition 12. *If F is a locally free instanton sheaf of charge c such that $F|_\Gamma \simeq 2\cdot\mathcal{O}_\Gamma$, and $H^1(F^\vee|_\Sigma \otimes L(2)) = 0$. If $\varphi : F \twoheadrightarrow Q(2)$ is an epimorphism, then*

$$[\ker\varphi] \in \overline{\mathcal{C}(d_1, d_2, c+m)}.$$

Proof. The idea is the same as in the proof of Proposition 10. Let E' be the kernel of the composition $F \xrightarrow{\varphi} Q(2) \twoheadrightarrow \mathcal{O}_\Gamma((2m-1)\mathrm{pt})$, so that $E := \ker\varphi$ and E' are related via the following exact sequence:

$$0 \to E \to E' \xrightarrow{\varphi'} L(2) \to 0.$$

By [Jardim et al. 2017, Theorem 7.8], one can find an affine open subset $0 \in U \subset \mathbb{A}^1$ and a coherent sheaf $\boldsymbol{E}$ on $\mathbb{P}^3 \times U$, flat over U, such that $\boldsymbol{E}_0 = E'$ and $\boldsymbol{E}_t$ is a locally free instanton sheaf of charge $c+m$ for every $t \in U \setminus \{0\}$.

Setting $\boldsymbol{G} := \pi^* L(2)$, we must, in order to apply Lemma 11, check that

$$H^i(\mathcal{H}om(E_t, G_t)) = 0, \qquad \text{for } i \geq 1 \text{ and } t \in U.$$

Indeed, since $\dim G_t = 1$, it is enough to show that $H^1(\mathcal{H}om(E_t, G_t)) = 0$. Note

$$\mathcal{H}om(E_0, G_0) = \mathcal{H}om(E', L(2)) \simeq \mathcal{H}om(F, L(2)) \simeq F^\vee|_\Sigma \otimes L(2),$$

where the middle isomorphism follows from applying the functor $\mathcal{H}om(\,\cdot\,, L(2))$ to the sequence

$$0 \to E' \to F \to \mathcal{O}_\Gamma((2m-1)\mathrm{pt}) \to 0,$$

also exploring the fact that Σ and Γ are disjoint. It follows that $H^1(\mathcal{H}om(E_0, G_0)) = H^1(F^\vee|_\Sigma \otimes L(2)) = 0$ by hypothesis. By semicontinuity of $h^1(\mathcal{H}om(E_t, G_t))$, we can shrink U to another affine open subset $U' \subset \mathbb{A}^1$, if necessary, to guarantee that $H^1(\mathcal{H}om(E_t, G_t)) = 0$ for every $t \in U'$.

By Lemma 11, there exists an epimorphism $\boldsymbol{s} : \boldsymbol{F} \twoheadrightarrow \boldsymbol{G}$ extending $\varphi' : E' \to L(2)$, so that $[\ker \boldsymbol{s}_t] \in \mathcal{C}(d_1, d_2, c+m)$, by construction. Since $E \simeq \ker \boldsymbol{s}_0$, it follows that $[E] \in \overline{\mathcal{C}(d_1, d_2, c+m)}$. □

4. Moduli of sheaves of dimension one and Euler characteristic zero

Given two integers d and χ, $d \geq 1$, let $\mathcal{T}(d, \chi)$ be the moduli space of semistable coherent sheaves on $\mathbb{P}^3$ with Hilbert polynomial $P(t) = d \cdot t + \chi$. In this section, we focus on the space $\mathcal{T}(d) := \mathcal{T}(d, 0)$.

Apart from its intrinsic interest, the space $\mathcal{T}(d)$ is also relevant for the study of instanton sheaves, and the description of $\mathcal{T}(d)$ for $d \leq 4$ provided in this section will be a key ingredient for the proof of the Main Theorem 1.

In addition, let $\mathcal{Z}(d)$ denote the set of rank 0 instanton sheaves of degree d *modulo S-equivalence* (which makes sense, since, by Lemma 6, every rank 0 instanton sheaf is semistable). After a twist by $\mathcal{O}_{\mathbb{P}^3}(-2)$, $\mathcal{Z}(d)$ can be regarded as an open subscheme of the moduli space $\mathcal{T}(d)$ consisting of those sheaves Q satisfying $h^0(Q) = 0$.

The space $\mathcal{T}(d)$ has several distinguished subsets, which we now describe.

First, let $\mathcal{P}_d \subset \mathcal{T}(d)$ be the subset of planar sheaves; it is a fiber bundle over $(\mathbb{P}^3)^*$ with fiber being the moduli space of semistable coherent sheaves on $\mathbb{P}^2$ with Hilbert polynomial $P = d \cdot t$. In view of [Le Potier 1993a, Theorem 1.1], $\mathcal{P}_d$ is a projective irreducible variety of dimension $d^2 + 4$. In particular, $\mathcal{P}_d$ is closed.

Next, consider the subsets $\mathcal{R}_d^o, \mathcal{E}_d^o \subset \mathcal{T}(d)$ of sheaves supported on smooth rational curves of degree d, respectively, on smooth elliptic curves of degree d. Let $\mathcal{R}_d$ and $\mathcal{E}_d$ denote their closures.

Given a partition $(d_1, \dots, d_s)$ of d such that $d_1 \geq \cdots \geq d_s$, we denote by

$$\mathcal{T}_{d_1,\dots,d_s} \subset \mathcal{T}(d)$$

the locally closed subset of points of the form

$$[Q_1 \oplus \cdots \oplus Q_s], \tag{11}$$

where Q_i gives a stable point in $\mathcal{T}(d_i)$; in particular, $\mathcal{T}_d$ is the open subset of stable points in $\mathcal{T}(d)$. Let $\mathcal{T}^o_{d_1,\dots,d_s} \subset \mathcal{T}_{d_1,\dots,d_s}$ be the open dense subset given by the condition that $\mathrm{supp}(Q_i)$ be mutually disjoint. Clearly, each irreducible component of $\mathcal{T}^o_{d_1,\dots,d_s}$ is an open dense subset of an irreducible component of $\mathcal{T}(d)$. Hence the irreducible components of the closure of $\mathcal{T}_{d_1,\dots,d_s}$ within $\mathcal{T}(d)$, henceforth denoted by $\overline{\mathcal{T}}_{d_1,\dots,d_s}$, are also irreducible components of $\mathcal{T}(d)$. On the other hand, each point of $\mathcal{T}(d)$ is an S-equivalence class of a polystable (e.g., stable) sheaf of the form (11). Hence, Lemma 13 follows.

Lemma 13. (i) *All irreducible components of $\mathcal{T}(d)$ are exhausted by the irreducible components of the union*

$$\bigcup_{(d_1,\dots,d_s)} \overline{\mathcal{T}}_{d_1,\dots,d_s}, \tag{12}$$

this union being taken over all the partitions $(d_1, \dots, d_s)$ of d.

(ii) *For a given partition $(d_1, \dots, d_s)$ of d, each irreducible component of $\overline{\mathcal{T}}_{d_1,\dots,d_s}$ is birational to a symmetric product,*

$$(\mathcal{X}_1 \times \cdots \times \mathcal{X}_s)/\Sigma,$$

of irreducible components $\mathcal{X}_i$ of $\mathcal{T}_{d_i}$, where Σ is the subgroup of the full symmetric group Σ_s of degree s generated by the transpositions (i, j) for which $d_i = d_j$ and $\mathcal{X}_i = \mathcal{X}_j$.

Proof. We have only to prove statement (ii). Indeed, let $\Sigma' \subset \Sigma_s$ be the subgroup generated by the transpositions $(i, i+1)$ for which $d_i = d_{i+1}$. We have a bijective morphism

$$(\mathcal{T}_{d_1} \times \cdots \times \mathcal{T}_{d_s})/\Sigma' \to \mathcal{T}_{d_1,\dots,d_s}, \qquad ([Q_1], \dots, [Q_s]) \longmapsto [Q_1 \oplus \cdots \oplus Q_s],$$

which is an isomorphism over $\mathcal{T}^o_{d_1,\dots,d_s}$, because over this set we can construct local inverse maps. Whence, the statement (ii) follows. $\square$

Remark 14. Lemma 13 implies that the problem of finding the irreducible components of $\mathcal{T}(d)$ is reduced to the problem of finding the irreducible components of $\mathcal{T}_2, \dots, \mathcal{T}_d$.

Remark 15. It also follows from Lemma 13 that the number of irreducible components of $\mathcal{T}(d)$ is at least as large as the number of partitions of d, usually denoted $p(d)$. A well-known formula by Hardy and Ramanujan gives the following asymptotic expression

$$p(d) \sim \frac{1}{4\sqrt{3}\cdot d} \exp\Bigl(\pi\sqrt{\frac{2d}{3}}\Bigr).$$

Therefore, the number of irreducible components of $\mathcal{T}(d)$ grows at least exponentially on $\sqrt{d}$. However, as we shall see below in the cases $d = 3$ and $d = 4$, $p(d)$ is just a rough underestimate of the number of irreducible components of $\mathcal{T}(d)$.

Given a coherent sheaf Q on $\mathbb{P}^3$, we define

$$Q^{\mathrm{D}} := \mathcal{E}xt^c(Q, \omega_{\mathbb{P}^3}),$$

where $c = \operatorname{codim}(Q)$. We shall later use the following general result regarding stable sheaves in $\mathcal{T}(d)$.

Lemma 16. *Assume that $\mathcal{F}$ gives a stable point in $\mathcal{T}(d)$ and that $P \in \operatorname{supp}(\mathcal{F})$ is a closed point. Then there are exact sequences*

$$0 \to \mathcal{E} \to \mathcal{F} \to \mathbb{C}_P \to 0 \tag{13}$$

and

$$0 \to \mathcal{F} \to \mathcal{G} \to \mathbb{C}_P \to 0 \tag{14}$$

for some sheaves $\mathcal{E} \in \mathcal{T}(d, -1)$ and $\mathcal{G} \in \mathcal{T}(d, +1)$.

Proof. Choose a surjective morphism $\mathcal{F} \to \mathbb{C}_P$ and denote its kernel by $\mathcal{E}$. Since $\mathcal{F}$ is stable, $\mathcal{E}$ is semistable, so we have sequence (13). According to [Maican 2010, Theorem 13], the dual sheaf $\mathcal{F}^{\mathrm{D}}$ gives a stable point in $\mathcal{T}(d)$. Thus, we have an exact sequence

$$0 \to \mathcal{E}_1 \to \mathcal{F}^{\mathrm{D}} \to \mathbb{C}_P \to 0$$

with $\mathcal{E}_1 \in \mathcal{T}(d, -1)$. According to [Maican 2010, Remark 4], $\mathcal{F}$ is reflexive. According to [Maican 2010, Theorem 13], the sheaf $\mathcal{G} = \mathcal{E}_1^{\mathrm{D}}$ gives a point in $\mathcal{T}(d, 1)$. Since $\mathcal{F}^{\mathrm{D}}$ is pure, we can apply [Huybrechts and Lehn 1997, Proposition 1.1.10] to deduce that

$$\mathcal{E}xt^3(\mathcal{F}^{\mathrm{D}}, \omega_{\mathbb{P}^3}) = 0.$$

The long exact sequence of $\mathcal{E}xt$-sheaves associated to the above exact sequence yields (14). $\square$

The goal of this section is to describe the irreducible components of $\mathcal{T}(d)$ for $d \leq 4$. According to [Drézet and Maican 2011], for $\mathcal{F} \in \mathcal{T}(d)$ we have the following cohomological conditions

$$h^0(\mathcal{F}) = 0 \qquad \text{if } d = 1 \text{ or } 2,$$
$$h^0(\mathcal{F}) \leq 1 \qquad \text{if } d = 3 \text{ or } 4.$$

4.1. ***Moduli of sheaves of degree* 1 *and* 2.** The case $d = 1$ is straightforward: clearly, $\mathcal{T}(1) \simeq \mathcal{R}_1$, being isomorphic to the Grassmannian of lines in $\mathbb{P}^3$.

Proposition 17. *The moduli space $\mathcal{T}(1)$ is an irreducible projective variety of dimension* 4.

In addition, it is easy to see that $\mathcal{Z}(1) = \mathcal{T}(1)$.

Proposition 18. *The moduli space $\mathcal{T}(2)$ is connected, and has two irreducible components, each of dimension* 8: $\mathcal{P}_2$ (*which coincides with* $\mathcal{R}_2$) *and* $\overline{\mathcal{T}}_{1,1}$.

Proof. If $\mathcal{F} \in \mathcal{T}_2$, then we have the exact sequence (14) in which $\mathcal{G} \in \mathcal{T}(2, 1)$. Thus, $\mathcal{G}$ is the structure sheaf of a conic curve, hence $\mathcal{G}$ is planar, and hence $\mathcal{F}$ is planar. We conclude that $\mathcal{T}(2) = \mathcal{P}_2 \cup \overline{\mathcal{T}}_{1,1}$. The intersection $\mathcal{P}_2 \cap \overline{\mathcal{T}}_{1,1}$ consists of those points of the form $[\mathcal{O}_{\ell_1}(-1) \oplus \mathcal{O}_{\ell_2}(-1)]$ where ℓ_1 and ℓ_2 are two intersecting (and possibly coincident) lines. □

Note also that $\mathcal{Z}(2) = \mathcal{T}(2)$; the fact that $\mathcal{Z}(2)$ consists of two irreducible components of dimension 8 should be compared with [Hauzer and Langer 2011, Corollary 6.12], where the authors prove that the moduli space of *framed* rank 0 instanton sheaves of multiplicity 2 also consists of two irreducible components of dimension 8.

4.2. ***Moduli of sheaves of degree 3.***

Proposition 19. *The moduli space $\mathcal{T}(3)$ has four irreducible components $\mathcal{P}_3$, $\mathcal{R}_3$, $\overline{\mathcal{T}}_{2,1}$ and $\overline{\mathcal{T}}_{1,1,1}$, of dimension* 13, 13, 12, *and* 12, *respectively.*

Proof. By Proposition 18 we have $\overline{\mathcal{T}}_2 = \mathcal{P}_2$, so that in view of Lemma 13 we already obtain the irreducible components $\overline{\mathcal{T}}_{2,1}$ and $\overline{\mathcal{T}}_{1,1,1}$ of $\mathcal{T}(3)$. Therefore, by Remark 14, we only have to find the irreducible components of $\mathcal{T}_3$.

Thus, given $\mathcal{F} \in \mathcal{T}_3$, take a point $P \in \operatorname{supp}(\mathcal{F})$. We then have the exact sequence (14) for $\mathcal{G} \in \mathcal{T}(3, 1)$. According to [Freiermuth and Trautmann 2004, Theorem 1.1], $\mathcal{T}(3, 1)$ has two irreducible components: the subset $\mathcal{P}$ of planar sheaves and the subset $\mathcal{R}$ that is the closure of the set of structure sheaves of twisted cubics. Moreover, all sheaves in $\mathcal{R} \setminus \mathcal{P}$ are structure sheaves of curves $R \subset \mathbb{P}^3$ of degree 3 and arithmetic genus zero. If $\mathcal{G}$ is planar, then $\mathcal{F}$ is planar. If $\mathcal{G} = \mathcal{O}_R$, then $R = \operatorname{supp}(\mathcal{F})$, where the scheme-theoretic support $\operatorname{supp}(\mathcal{F})$ of the sheaf $\mathcal{F}$ is defined by the 0-th Fitting ideal $\operatorname{Fitt}^0(\mathcal{F}) : \mathcal{I}_{R/\mathbb{P}^3} = \operatorname{Fitt}^0(\mathcal{F})$. The morphism

$$\rho\colon \mathcal{T}_3 \setminus \mathcal{P}_3 \to \mathcal{R} \setminus \mathcal{P}, \qquad \rho([\mathcal{F}]) = [\mathcal{O}_{\operatorname{supp}(\mathcal{F})}],$$

is injective. Indeed, if $\rho([\mathcal{F}_1]) = \rho([\mathcal{F}_2])$, then $\operatorname{supp}(\mathcal{F}_1) = \operatorname{supp}(\mathcal{F}_2) = R$. Choose a point $P \in R$. We have exact sequences

$$0 \to \mathcal{F}_1 \to \mathcal{G}_1 \to \mathbb{C}_P \to 0, \qquad 0 \to \mathcal{F}_2 \to \mathcal{G}_2 \to \mathbb{C}_P \to 0,$$

with $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{T}(3, 1)$. Clearly, $\mathcal{G}_1$ and $\mathcal{G}_2$ are both isomorphic to $\mathcal{O}_R$, hence $\mathcal{F}_1$ and $\mathcal{F}_2$ are both isomorphic to the ideal sheaf $\mathcal{I}_{P,R}$ of P in R. The image of ρ is a constructible set of the irreducible variety $\mathcal{R} \setminus \mathcal{P}$ and contains an open subset of $\mathcal{R} \setminus \mathcal{P}$, namely the subset given by the condition that R be irreducible. Indeed, if R is irreducible, then it is easy to check that $\mathcal{I}_{P,R}$ is stable; we have $\rho([\mathcal{I}_{P,R}]) = [\mathcal{O}_R]$. We deduce that $\mathcal{T}_3 \setminus \mathcal{P}_3$ is irreducible. It follows that $\mathcal{R}_3^o$ is dense in $\mathcal{T}_3 \setminus \mathcal{P}_3$. Thus, $\mathcal{T}_3$ has two irreducible components, hence $\mathcal{T}(3)$ has the four irreducible components announced in the proposition. □

4.3. *Moduli of sheaves of degree 4.*

Proposition 20. *The moduli space* $\mathcal{T}(4)$ *has eight irreducible components*: $\mathcal{P}_4$, $\mathcal{E}_4$, $\mathcal{R}_4$, $\overline{\mathcal{T}}_{2,2}$, $\overline{\mathcal{T}}_{2,1,1}$, $\overline{\mathcal{T}}_{1,1,1,1}$ *and two irreducible components of* $\mathcal{T}_{3,1}$ *that are birational to* $\mathcal{P}_3 \times \mathcal{T}_1$ *and to* $\mathcal{R}_3 \times \mathcal{T}_1$, *respectively. Their dimensions are*, *respectively*, 20, 18, 16, 16, 16, 16, 17, 17.

Proof. By Propositions 18 and 19 and Lemma 13 we already have 5 irreducible components of $\mathcal{T}(4)$ which are $\overline{\mathcal{T}}_{2,2}$, $\overline{\mathcal{T}}_{2,1,1}$, $\overline{\mathcal{T}}_{1,1,1,1}$ and two irreducible components of $\mathcal{T}_{3,1}$ that are birational to $\mathcal{P}_3 \times \mathcal{T}_1$ and to $\mathcal{R}_3 \times \mathcal{T}_1$, respectively. Therefore by Remark 14 we have only to find the irreducible components of $\mathcal{T}_4$. Thus, given $\mathcal{F} \in \mathcal{T}_4$, take a point $P \in \operatorname{supp}(\mathcal{F})$. We then have the exact sequence (14) for $\mathcal{G} \in \mathcal{T}(4, 1)$. According to [Choi et al. 2016, Theorem 4.12], $\mathcal{T}(4, 1)$ has three irreducible components: the subset $\mathcal{P}$ of planar sheaves, the subset $\mathcal{R}$ that is the closure of the set of structure sheaves of rational quartic curves, and the set $\mathcal{E}$ that is the closure of the set of sheaves of the form $\mathcal{O}_E(P')$, where E is a smooth elliptic quartic curve and $P' \in E$. If $\mathcal{G} \in \mathcal{P}$, then $\mathcal{F} \in \mathcal{P}_4$. The sheaves in $\mathcal{R} \setminus (\mathcal{P} \cup \mathcal{E})$ are structure sheaves of quartic curves of arithmetic genus zero. The sheaves in $\mathcal{E} \setminus \mathcal{P}$ are supported on quartic curves of arithmetic genus 1. Let $\mathcal{T}_{4,\mathrm{rat}} \subset \mathcal{T}_4$ be the subset of sheaves whose support is a quartic curve of arithmetic genus zero. As in Proposition 19, we can construct an injective dominant morphism

$$\rho\colon \mathcal{T}_{4,\mathrm{rat}} \to \mathcal{R} \setminus (\mathcal{P} \cup \mathcal{E}), \qquad \rho([\mathcal{F}]) = [\mathcal{O}_{\operatorname{supp}(\mathcal{F})}].$$

It follows that $\mathcal{T}_{4,\mathrm{rat}}$ is irreducible, hence $\mathcal{T}_{4,\mathrm{rat}} \subset \mathcal{R}_4$. To finish the proof of the proposition we need to show that $\mathcal{T}_4 \setminus (\mathcal{P}_4 \cup \mathcal{T}_{4,\mathrm{rat}})$ is contained in $\mathcal{E}_4$.

According to [Maican 2017, Section 3], the sheaves $\mathcal{G}$ in $\mathcal{E} \setminus \mathcal{P}$ are of two kinds:

(i) $\mathcal{O}_E(P')$ for a curve E of arithmetic genus 1 given by an ideal of the form (q_1, q_2), where q_1, q_2 are quadratic forms, and $P' \in E$. Notice that

$$\mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathbb{C}_{P'}, \mathcal{O}_E) \xrightarrow{\sim} \mathbb{C},$$

so the notation $\mathcal{O}_E(P')$ is justified. Also note that $\mathcal{O}_E$ is stable.

(ii) Nonplanar extensions of the form

$$0 \to \mathcal{O}_L(-1) \to \mathcal{G} \to \mathcal{C} \to 0,$$

where L is a line and $\mathcal{C}$ gives a point in $\mathcal{T}_H(3, 1)$ for a plane H possibly containing L. (Here and below we use the notation $\mathcal{T}_S(d, \chi)$ for the moduli space of one-dimensional sheaves on a given surface S in $\mathbb{P}^3$ with Hilbert polynomial $P(t) = dt + \chi$. We also set $\mathcal{T}_S(d) := \mathcal{T}_S(d, 0)$.)

Claim 1: Case (ii) is unfeasible.

Assume, firstly, that $P \in H$. Tensoring (14) with $\mathcal{O}_H$, we get the exact sequence

$$\mathcal{F}|_H \to \mathcal{G}|_H \xrightarrow{\alpha} \mathbb{C}_P \to 0.$$

Thus, $\mathrm{Ker}(\alpha)$ is a quotient sheaf of $\mathcal{F}$ of slope zero. This contradicts the stability of $\mathcal{F}$. Assume, secondly, that $P \notin H$. According to [Maican 2017, Proposition 3.5], we have an exact sequence

$$0 \to \mathcal{E} \to \mathcal{G} \to \mathcal{O}_L \to 0$$

for some sheaf $\mathcal{E} \in \mathcal{T}_H(3)$. The composite map $\mathcal{E} \to \mathcal{G} \to \mathbb{C}_P$ is zero, hence $\mathcal{E}$ is a subsheaf of $\mathcal{F}$. This contradicts the stability of $\mathcal{F}$ and proves Claim 1.

It remains to deal with the sheaves from (i). We have one of the following possibilities:

(a) E is contained in a smooth quadric surface S.

(b) E is contained in an irreducible cone Σ but not in a smooth quadric surface.

(c) $\mathrm{span}\{q_1, q_2\}$ contains only reducible quadratic forms and q_1 and q_2 have no common factor.

Claim 2: In case (a), $\mathcal{F}$ belongs to $\mathcal{E}_4$.

Notice that $\mathcal{F} \in \mathcal{T}_S(4)$. According to [Ballico and Huh 2014, Proposition 7], $\mathcal{T}_S(4)$ has five disjoint irreducible components $\mathcal{T}_S(p, q, 4)$, where (p, q) is the type of the support of the one-dimensional sheaf with respect to $\mathrm{Pic}(S)$. Clearly, $\mathcal{F} \in \mathcal{T}_S(2, 2, 4)$. Thus, $\mathcal{F}$ is a limit of sheaves in $\mathcal{T}_S(2, 2, 4)$ supported on smooth curves of type $(2, 2)$, hence $\mathcal{F} \in \mathcal{E}_4$.

It remains to deal with cases (b) and (c). Next we reduce further to the case when $P = P'$. Notice that, if $P \neq P'$, then $\mathcal{F} \simeq \mathcal{O}_E(P') \otimes (\mathcal{O}_E(P))^{\mathrm{D}}$, hence the notation $\mathcal{F} = \mathcal{O}_E(P' - P)$ is justified.

Claim 3: Assume that $\mathcal{F} = \mathcal{O}_E(P' - P)$ for an elliptic quartic curve E and distinct closed points $P', P \in E$. Then $\mathcal{F}$ belongs to $\mathcal{E}_4$.

Let $Z_1, \dots, Z_m$ denote the irreducible components of E. Fix $i, j \in \{1, \dots, m\}$. Consider the locally closed subset $\mathcal{X} \subset \mathcal{E} \times \mathcal{E}$ of pairs $([\mathcal{O}_{E'}(P_1)], [\mathcal{O}_{E'}(P_2)])$, where E' is a quartic curve of arithmetic genus 1 whose ideal is generated by two quadratic polynomials, and P_1 and P_2 are distinct points on E' such that $P_1 \notin \cup_{k \neq i} Z_k$ and $P_2 \notin \cup_{k \neq j} Z_k$. Consider the morphisms

$$\xi : \mathcal{X} \to \mathcal{T}(4), \qquad ([\mathcal{O}_{E'}(P_1)], [\mathcal{O}_{E'}(P_2)]) \longmapsto [\mathcal{O}_{E'}(P_1 - P_2)],$$
$$\sigma : \mathcal{X} \to \mathrm{Hilb}_{\mathbb{P}^3}(4t), \qquad ([\mathcal{O}_{E'}(P_1)], [\mathcal{O}_{E'}(P_2)]) \longmapsto E',$$

where $\mathrm{Hilb}_{\mathbb{P}^3}(4t)$ is the Hilbert scheme of subschemes of $\mathbb{P}^3$ with Hilbert polynomial $P(t) = 4t$. According to [Chen and Nollet 2012, Examples 2.8 and 4.8], $\mathrm{Hilb}_{\mathbb{P}^3}(4t)$ consists of two irreducible components, denoted $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$. The generic member of $\boldsymbol{H}_1$ is a smooth elliptic quartic curve. The generic member of $\boldsymbol{H}_2$ is the disjoint union of a planar quartic curve and two isolated points. Note that $\boldsymbol{H}_2$ lies in the closed subset $\{E' \mid h^0(\mathcal{O}_{E'}) \geq 3\}$. Since E lies in the complement of this subset, we deduce that $E \in \boldsymbol{H}_1$. It follows that there exists an irreducible quasiprojective curve $\Gamma \subset \mathrm{Hilb}_{\mathbb{P}^3}(4t)$ containing E, such that $\Gamma \setminus \{E\}$ consists of smooth elliptic quartic curves (see the proof of [Maican 2017, Proposition 4.2]). The fibers of the map $\sigma^{-1}(\Gamma) \to \Gamma$ are irreducible of dimension 2. By [Shafarevich 1994, Theorem 8, p. 77], we deduce that $\sigma^{-1}(\Gamma)$ is irreducible. Thus, $\xi(\sigma^{-1}(\Gamma))$ is irreducible. This set contains $[\mathcal{O}_E(P' - P)]$ for $P' \in Z_i \setminus \cup_{k \neq i} Z_k$ and $P \in Z_j \setminus \cup_{k \neq j} Z_k$. The generic member of $\xi(\sigma^{-1}(\Gamma))$ is a sheaf supported on a smooth elliptic quartic curve. We conclude that $[\mathcal{O}_E(P' - P)] \in \mathcal{E}_4$. Since i and j are arbitrary, the result is true for all P' and P closed points on E.

Claim 4: In case (c), E is a quadruple line supported on a line L. More precisely, there are three distinct planes H, H', H'' containing L, such that

$$E = (H \cup H') \cap (2H'').$$

The claim will follow if we can show that there are linearly independent one-forms u, v such that $q_1, q_2 \in \mathbb{C}[u, v]$. Indeed, in this case (q_1, q_2) has the normal form $(uv, (u+v)^2)$. We argue by contradiction. Assume that $q_1 = XY$ and $q_2 = Zl$. Consider first the case when $l = aX + bY + cZ$. We will find $\lambda \in \mathbb{C}$ such that $f = XY + \lambda Zl$ is irreducible, which is equivalent to saying that

$$\frac{\partial f}{\partial X} = Y + a\lambda Z, \qquad \frac{\partial f}{\partial Y} = X + b\lambda Z, \qquad \frac{\partial f}{\partial Z} = \lambda(aX + bY + 2cZ)$$

have no common zero, or, equivalently,

$$\begin{vmatrix} 0 & 1 & a\lambda \\ 1 & 0 & b\lambda \\ a\lambda & b\lambda & 2c\lambda \end{vmatrix} \neq 0.$$

We have reduced to the inequality $2ab\lambda^2 - 2c\lambda \neq 0$. If $c \neq 0$ we can find a solution. If $c = 0$, then $ab \neq 0$, otherwise q_1 and q_2 would have a common factor, and we can choose any $\lambda \in \mathbb{C}^*$. Assume now that $l = aX + bY + cZ + dW$ with $d \neq 0$. Note that $f = XY + \lambda Zl$ is irreducible if its image in

$$\mathbb{C}[X, Y, Z, W]/\langle (c-1)Z + dW \rangle \simeq \mathbb{C}[X, Y, Z]$$

is irreducible. The above isomorphism sends f to $XY + \lambda Z(aX + bY + Z)$ which brings us to the case examined above.

Claim 5: In case (c), $\mathcal{F}$ belongs to $\mathcal{E}_4$.

We have $\mathcal{O}_E|_H \simeq \mathcal{O}_C$ and $\mathcal{O}_E|_{H'} \simeq \mathcal{O}_{C'}$ for conic curves C and C' supported on L. The kernel of the map $\mathcal{O}_E \to \mathcal{O}_C$ has Hilbert polynomial $2t - 1$ and is stable, because $\mathcal{O}_E$ is stable, hence it is isomorphic to $\mathcal{O}_{C'}(-1)$. We have a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_E \arrow[r] \arrow[d] & \mathcal{O}_E(P') \arrow[r] \arrow[d] & \mathbb{C}_{P'} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{O}_C \arrow[r] & \mathcal{O}_E(P')|_H \arrow[r] & \mathbb{C}_{P'} \arrow[r] & 0
\end{tikzcd}$$

in which the second row is obtained by restricting the first row to H. Applying the snake lemma, we obtain the first row of the following exact commutative diagram:

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_{C'}(-1) \arrow[r] & \mathcal{O}_E(P') \arrow[r] \arrow[d] & \mathcal{O}_E(P')|_H \arrow[r] \arrow[d, "\alpha"] & 0 \\
 & & \mathbb{C}_P \arrow[r, equal] & \mathbb{C}_P &
\end{tikzcd}$$

Applying the snake lemma to this diagram, we get the exact sequence

$$0 \to \mathcal{O}_{C'}(-1) \to \mathcal{F} \to \operatorname{Ker}(\alpha) \to 0.$$

Note that $\operatorname{Ker}(\alpha)$ has Hilbert polynomial $2t + 1$ and is semistable, being a quotient of the stable sheaf $\mathcal{F}$. It follows that $\operatorname{Ker}(\alpha) \simeq \mathcal{O}_C$. Thus, $\mathcal{F}$ gives a point in the set $\mathbb{P}(\operatorname{Ext}^1(\mathcal{O}_C, \mathcal{O}_{C'}(-1)))^{\mathrm{s}}$ of stable nonsplit extensions of $\mathcal{O}_C$ by $\mathcal{O}_{C'}(-1)$.

Consider the family of planes H''_t, $t \in \mathbb{P}^1 \setminus \{0, \infty\}$, containing L and different from H and H'. Denote $E_t = (H \cup H') \cap (2H''_t)$. We have a two-dimensional family of semistable sheaves

$$\left\{ \mathcal{O}_{E_t}(P' - P'') \mid t \in \mathbb{P}^1 \setminus \{0, \infty\},\ P'' \in L \setminus \{P'\} \right\} \subset \mathbb{P}(\operatorname{Ext}^1(\mathcal{O}_C, \mathcal{O}_{C'}(-1))).$$

This family is dense in the right-hand side because $\operatorname{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{C'}(-1)) \simeq \mathbb{C}^3$. To prove this we use the standard exact sequence obtained from Thomas' spectral sequence

$$
\begin{aligned}
0 \to \operatorname{Ext}^1_{\mathcal{O}_{H'}}(\mathcal{O}_C|_{H'}, \mathcal{O}_{C'}(-1)) &\to \operatorname{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{C'}(-1)) \\
&\to \operatorname{Hom}(\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{H'}), \mathcal{O}_{C'}(-1)) \to \operatorname{Ext}^2_{\mathcal{O}_{H'}}(\mathcal{O}_C|_{H'}, \mathcal{O}_{C'}(-1)),
\end{aligned}
$$

see also [Choi et al. 2016, Lemma 4.2]. Note that $\mathcal{O}_C|_{H'} \simeq \mathcal{O}_L$. Using Serre duality we obtain the isomorphisms

$$
\begin{aligned}
\operatorname{Ext}^2_{\mathcal{O}_{H'}}(\mathcal{O}_L, \mathcal{O}_{C'}(-1)) &\simeq \operatorname{Hom}_{\mathcal{O}_{H'}}(\mathcal{O}_{C'}(-1), \mathcal{O}_L(-3))^* = 0, \\
\operatorname{Ext}^1_{\mathcal{O}_{H'}}(\mathcal{O}_L, \mathcal{O}_{C'}(-1)) &\simeq \operatorname{Ext}^1_{\mathcal{O}_{H'}}(\mathcal{O}_{C'}(-1), \mathcal{O}_L(-3))^* \simeq \mathbb{C}^2.
\end{aligned}
$$

The last isomorphism follows from the long exact sequence of extension sheaves

$$
\begin{aligned}
0 = \operatorname{Hom}\big(\mathcal{O}_{H'}(-1), \mathcal{O}_L(-3)\big) &\to \operatorname{Hom}\big(\mathcal{O}_{H'}(-3), \mathcal{O}_L(-3)\big) \simeq H^0(\mathcal{O}_L) \simeq \mathbb{C} \\
\to \operatorname{Ext}^1_{\mathcal{O}_{H'}}\big(\mathcal{O}_{C'}(-1), \mathcal{O}_L(-3)\big) &\to \operatorname{Ext}^1_{\mathcal{O}_{H'}}\big(\mathcal{O}_{H'}(-1), \mathcal{O}_L(-3)\big) \simeq H^1\big(\mathcal{O}_L(-2)\big) \simeq \mathbb{C} \\
&\to \operatorname{Ext}^1_{\mathcal{O}_{H'}}\big(\mathcal{O}_{H'}(-3), \mathcal{O}_L(-3)\big) = 0
\end{aligned}
$$

derived from the short exact sequence

$$
0 \to \mathcal{O}_{H'}(-3) \to \mathcal{O}_{H'}(-1) \to \mathcal{O}_{C'}(-1) \to 0.
$$

Choose linear forms u and u' defining H and H'. Restricting the standard resolution

$$
0 \to \mathcal{O}(-3) \xrightarrow{\begin{bmatrix} -u \\ (u')^2 \end{bmatrix}} \mathcal{O}(-2) \oplus \mathcal{O}(-1) \xrightarrow{\begin{bmatrix} (u')^2 & u \end{bmatrix}} \mathcal{O} \to \mathcal{O}_C \to 0
$$

to H', we see that $\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{H'})$ is isomorphic to the cohomology of the complex

$$
\mathcal{O}_{H'}(-3) \xrightarrow{\begin{bmatrix} -u|_{H'} \\ 0 \end{bmatrix}} \mathcal{O}_{H'}(-2) \oplus \mathcal{O}_{H'}(-1) \xrightarrow{\begin{bmatrix} 0 & u|_{H'} \end{bmatrix}} \mathcal{O}_{H'}
$$

that is, to $\mathcal{O}_L(-2)$. Using the fact that $\mathcal{O}_{C'}(-1)$ and $\mathcal{O}_L(-2)$ are reflexive, we have the isomorphisms

$$
\operatorname{Hom}\big(\mathcal{O}_L(-2), \mathcal{O}_{C'}(-1)\big) \simeq \operatorname{Hom}\big(\mathcal{O}_{C'}(-1)^{\mathrm{D}}, \mathcal{O}_L(-2)^{\mathrm{D}}\big) \simeq \operatorname{Hom}(\mathcal{O}_{C'}, \mathcal{O}_L) \simeq \mathbb{C}.
$$

The above discussion shows that $[\mathcal{F}]$ is a limit of points in $\mathcal{T}(4)$ of the form $[\mathcal{O}_{E_t}(P' - P'')]$, with $P' \neq P''$. Claim 5 now follows from Claim 3.

It remains to consider sheaves $\mathcal{F}$ given by sequence (14) in which $\mathcal{G} = \mathcal{O}_E(P)$ and E is as at (b). We reduce further to the case when E has no regular points.

Claim 6: Assume E has a regular point. Then $\mathcal{F} \simeq \mathcal{O}_E$, hence $\mathcal{F}$ belongs to $\mathcal{E}_4$.

The proof of the claim is obvious because P in sequence (14) can be chosen arbitrarily on E. We choose $P \in \operatorname{reg}(E)$. The kernel of the map $\mathcal{O}_E(P) \to \mathbb{C}_P$ is $\mathcal{O}_E$. Note that E belongs to the irreducible component $\boldsymbol{H}_1$ of $\operatorname{Hilb}_{\mathbb{P}^3}(4t)$, hence it is the limit of smooth elliptic quartic curves.

Claim 7: Let $E \subset \mathbb{P}^3$ be a quartic curve of arithmetic genus 1 which is contained in an irreducible cone Σ, but not in a smooth quadric surface. Assume that E has no regular points. Then we have one of the following two possibilities:

(b1) $E = \Sigma \cap (H \cup H')$, where $H,\ H'$ are distinct planes each intersecting Σ along a double line.

(b2) $E = \Sigma \cap (2H)$, where H is a plane intersecting Σ along a double line.

To fix notations assume that Σ has vertex O and base a conic curve Γ contained in a plane Π. Assume first that $E = \Sigma \cap \Sigma'$ for Σ' another irreducible cone. If Σ and Σ' have distinct vertices, then E has regular points. Thus, Σ' has vertex O and base an irreducible conic curve Γ' contained in Π. Since E has no regular points, $\Gamma \cap \Gamma'$ is the union of two double points Q_1 and Q_2. Now E is the cone with vertex O and base $Q_1 \cup Q_2$, so E is as at (b1).

Assume next that $E = \Sigma \cap (H \cup H')$ for distinct planes H and H'. If H or H' does not contain O, then E has regular points. If H or H' is not tangent to Γ, then E has regular points. We deduce that E is as in (b1).

Assume, finally, that $E = \Sigma \cap (2H)$ for a double plane $2H$. If $O \notin H$, then it can be shown that E is contained in a smooth quadric surface. Indeed, assume that Σ has equation $X^2+Y^2+Z^2=0$ and H has equation $W=0$. Then E is contained in the smooth quadric surface with equation

$$X^2+Y^2+Z^2+W^2=0.$$

Thus, $O \in H$. If $\Gamma \cap H$ is the union of two distinct points, then $\Gamma \cap (2H)$ is the union of two double points Q_1 and Q_2 and E is as in (b1). If $\Gamma \cap H$ is a double point, then E is as in (b2).

Claim 8: In case (b1), $\mathcal{F}$ belongs to $\mathcal{E}_4$.

We have $\mathcal{O}_{E|H} \simeq \mathcal{O}_C$ and $\mathcal{O}_{E|H'} \simeq \mathcal{O}_{C'}$ for conic curves C and C' supported on lines L and L', respectively. Assume that $P \in L$ and choose a point $P' \in L$ not necessarily distinct from P. Let $\mathcal{F}' \in \mathcal{T}_4$ be given by the exact sequence

$$0 \to \mathcal{F}' \to \mathcal{O}_E(P') \to \mathbb{C}_P \to 0.$$

As in the first paragraph in the proof of Claim 5, we see that $\mathcal{F}'$ gives a point in the set $\mathbb{P}(\mathrm{Ext}^1(\mathcal{O}_C, \mathcal{O}_{C'}(-1)))^{\mathrm{s}}$. We have $\dim \mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{C'}(-1)) \le 2$. Indeed, start with the exact sequence

$$\begin{aligned}0 \to \mathrm{Ext}^1_{\mathcal{O}_{H'}}\big(\mathcal{O}_{C|H'}, \mathcal{O}_{C'}(-1)\big) &\to \mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}\big(\mathcal{O}_C, \mathcal{O}_{C'}(-1)\big)\\ &\to \mathrm{Hom}\big(\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{H'}), \mathcal{O}_{C'}(-1)\big).\end{aligned}$$

The group on the second line vanishes because $\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{H'})$ is supported on O while $\mathcal{O}_{C'}(-1)$ has no zero-dimensional torsion. It follows that

$$\mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C, \mathcal{O}_{C'}(-1)) \simeq \mathrm{Ext}^1_{\mathcal{O}_{H'}}(\mathcal{O}_{C|H'}, \mathcal{O}_{C'}(-1)).$$

The sheaf $\mathcal{O}_{C|H'}$ is the structure sheaf of a double point supported on O, hence we have the exact sequence

$$\mathbb{C}\simeq\mathrm{Ext}^1_{\mathcal{O}_{H'}}(\mathbb{C}_O,\mathcal{O}_{C'}(-1))\to\mathrm{Ext}^1_{\mathcal{O}_{H'}}(\mathcal{O}_{C|H'},\mathcal{O}_{C'}(-1))\to\mathrm{Ext}^1_{\mathcal{O}_{H'}}(\mathbb{C}_O,\mathcal{O}_{C'}(-1))\simeq\mathbb{C}$$

from which we get our estimate on the dimension of the middle vector space.

The one-dimensional family $\mathcal{O}_E(P'-P)$, $P'\in L\setminus\{P\}$, is therefore dense in $\mathbb{P}(\mathrm{Ext}^1(\mathcal{O}_C,\mathcal{O}_{C'}(-1)))^{\mathrm{s}}$, hence, in view of Claim 3, $\mathcal{F}$ is a limit of sheaves in $\mathcal{E}_4$. We conclude that $\mathcal{F}\in\mathcal{E}_4$.

Claim 9: In case (b2), $\mathcal{F}$ belongs to $\mathcal{E}_4$.

Let L be the reduced support of $\Sigma\cap H$. We have $\mathcal{O}_{E|H}\simeq\mathcal{O}_C$ for a conic curve supported on L. Choose a point $P'\in L$ not necessarily distinct from P and let $\mathcal{F}'\in\mathcal{T}_4$ be given by the exact sequence

$$0\to\mathcal{F}'\to\mathcal{O}_E(P')\to\mathbb{C}_P\to 0.$$

As in the first paragraph of the proof of Claim 5, we see that $\mathcal{F}'$ gives a point in the set $\mathbb{P}(\mathrm{Ext}^1(\mathcal{O}_C,\mathcal{O}_C(-1)))^{\mathrm{s}}$. We have $\dim\mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C,\mathcal{O}_C(-1))=5$. This follows from the exact sequence

$$\begin{aligned}0\to\mathrm{Ext}^1_{\mathcal{O}_H}(\mathcal{O}_C,\mathcal{O}_C(-1))&\to\mathrm{Ext}^1_{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C,\mathcal{O}_C(-1))\\ &\to\mathrm{Hom}(\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C,\mathcal{O}_H),\mathcal{O}_C(-1))\to\mathrm{Ext}^2_{\mathcal{O}_H}(\mathcal{O}_C,\mathcal{O}_C(-1)).\end{aligned}$$

From Serre duality we get

$$\mathrm{Ext}^2_{\mathcal{O}_H}(\mathcal{O}_C,\mathcal{O}_C(-1))\simeq\mathrm{Hom}_{\mathcal{O}_H}(\mathcal{O}_C(-1),\mathcal{O}_C(-3))^*\simeq H^0(\mathcal{O}_C(-2))^*=0.$$

We have $\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C,\mathcal{O}_H)\simeq\mathcal{O}_C(-1)$ hence $\mathrm{Hom}(\mathcal{T}or_1^{\mathcal{O}_{\mathbb{P}^3}}(\mathcal{O}_C,\mathcal{O}_H),\mathcal{O}_C(-1))\simeq\mathbb{C}$. Applying the functor $\mathrm{Hom}(-,\mathcal{O}_C(-1))$ to the short exact sequence

$$0\to\mathcal{O}_H(-2)\to\mathcal{O}_H\to\mathcal{O}_C\to 0,$$

we obtain the exact sequence,

$$\begin{aligned}0\to\mathrm{Hom}(\mathcal{O}_H(-2),\mathcal{O}_C(-1))\simeq H^0(\mathcal{O}_C(1))\simeq\mathbb{C}^3&\to\mathrm{Ext}^1_{\mathcal{O}_H}(\mathcal{O}_C,\mathcal{O}_C(-1))\\ &\to\mathrm{Ext}^1_{\mathcal{O}_H}(\mathcal{O}_H,\mathcal{O}_C(-1))\simeq H^1(\mathcal{O}_C(-1))\simeq\mathbb{C}\to 0,\end{aligned}$$

since $\mathrm{Hom}(\mathcal{O}_H,\mathcal{O}_C(-1))\simeq H^0(\mathcal{O}_C(-1))=0$, and $\mathrm{Ext}^1_{\mathcal{O}_H}(\mathcal{O}_H(-2),\mathcal{O}_C(-1))\simeq H^1(\mathcal{O}_C(1))=0$.

Denote $Q=L\cap\Pi$. We have a three-dimensional family Γ_t of irreducible conic curves in Π that contain Q and are tangent to H. Let Σ_t be the cone with vertex O and base Γ_t. Put $E_t=\Sigma_t\cap(2H)$. The four-dimensional family $\mathcal{O}_{E_t}(P'-P)$, $P'\in L\setminus\{P\}$ is dense in $\mathbb{P}(\mathrm{Ext}^1(\mathcal{O}_C,\mathcal{O}_C(-1)))^{\mathrm{s}}$, hence, in view of Claim 3, $\mathcal{F}$ is the limit of sheaves in $\mathcal{E}_4$. We conclude that $\mathcal{F}\in\mathcal{E}_4$. □

The proof of Main Theorem 2 is finally complete.

5. Components and connectedness of $\mathcal{L}(3)$

We are now ready to prove that the moduli space of rank 2 instanton sheaves of charge 3 on $\mathbb{P}^3$ is connected and has precisely two irreducible components. Indeed, two components of $\overline{\mathcal{L}(3)}$ have already been identified above:

(I) $\overline{\mathcal{I}(3)}$, whose generic point corresponds to a locally free instanton sheaf.

(II) $\overline{\mathcal{C}(1,3,0)}$, whose generic point corresponds to an instanton sheaf E fitting into an exact sequence of the form

$$0 \to E \to 2\cdot\mathcal{O}_{\mathbb{P}^3} \to \iota_* L(2) \to 0, \tag{15}$$

where $\iota : \Sigma \hookrightarrow \mathbb{P}^3$ is the inclusion of a nonsingular plane cubic Σ, and $L \in \mathrm{Pic}^0(\Sigma)$ is such that $h^0(\Sigma, L) = 0$.

Both components have dimension 21; this is a classical result for the component $\mathcal{I}^0(3)$, while the dimension of $\mathcal{C}(1,3,0)$ is given by Theorem 8. In addition, this same result also guarantees that the union $\overline{\mathcal{I}^0(3) \cup \mathcal{C}(1,3,0)}$ is connected.

Therefore, our task is to prove that $\overline{\mathcal{L}(3)}$ has no other irreducible components, i.e., that every instanton sheaf of charge 3 can be deformed either into a locally free instanton sheaf, or into an instanton sheaf given by a sequence of the form (15).

So let E be a nonlocally free instanton sheaf of charge 3, and let $Q_E := E^{\vee\vee}/E$ be the corresponding rank 0 instanton sheaf; let d_E denote the degree of Q_E. There are three possibilities to consider: $d_E = 1$, $d_E = 2$ and $d_E = 3$.

The first possibility is easy to deal with: if $d_E = 1$, then $Q_E = \mathcal{O}_\ell(1)$, where $\ell \hookrightarrow \mathbb{P}^3$ is a line in $\mathbb{P}^3$. It follows that E fits into an exact sequence of the form

$$0 \to E \to F \to \mathcal{O}_\ell(1) \to 0,$$

where F is a locally free instanton sheaf of charge 2. However, [Jardim et al. 2015, Proposition 7.2] ensures that E can be deformed in a ('t Hooft) locally free instanton sheaf of charge 3. In other words, if $d_E = 1$, then E lies within $\mathcal{I}^0(3)$.

Now, if $d_E = 2$, then, since Q_E is semistable and by Proposition 18 above, one can find an affine open subset $0 \in U \subset \mathbb{A}^1$ and a coherent sheaf $\boldsymbol{G}$ on $\mathbb{P}^3 \times U$ such that $G_0 = Q_E$ and, for $u \neq 0$, either

(i) $G_u = \mathcal{O}_\Gamma(3\mathrm{pt})$, where Γ is a nonsingular conic in $\mathbb{P}^3$; or

(ii) $G_u = \mathcal{O}_{\ell_1}(1) \oplus \mathcal{O}_{\ell_2}(1)$ where ℓ_1 and ℓ_2 are skew lines in $\mathbb{P}^3$.

Since $d_E = 2$, $E^{\vee\vee}$ is a locally free instanton sheaf of charge 1 (also known as a null-correlation bundle), we set $N := E^{\vee\vee}$. Take $\boldsymbol{F} := \pi^* N$, where $\pi : \mathbb{P}^3 \times U \to \mathbb{P}^3$ is the projection onto the first factor. Let $s : N \twoheadrightarrow Q_E$ be the epimorphism given by the standard sequence (3). For every $u \in U$, the sheaf $\mathcal{H}om(F_u, G_u) \simeq N \otimes G_u$ is supported in dimension 1, thus clearly $H^i(\mathcal{H}om(F_u, G_u)) = 0$ for $i = 2, 3$. For $u \neq 0$ we can, after possibly shrinking U, assume that either $N|_\Gamma \simeq 2\cdot\mathcal{O}_\Gamma$ or $N|_{\ell_1} \simeq 2\cdot\mathcal{O}_{\ell_1}$

and $N|_{\ell_2} \simeq 2\cdot\mathcal{O}_{\ell_2}$; in both situations, it is easy to check that $H^1(\mathcal{H}om(F_u, G_u)) = 0$. Finally, for $u = 0$, we twist the resolution of Q_E

$$0 \to 2\cdot\mathcal{O}_{\mathbb{P}^3}(-1) \xrightarrow{\alpha} 4\cdot\mathcal{O}_{\mathbb{P}^3} \xrightarrow{\beta} 2\cdot\mathcal{O}_{\mathbb{P}^3}(1) \to Q_E \to 0$$

by N and check that $H^1(\mathcal{H}om(F_0, G_0)) = H^1(N \otimes Q_E) \simeq H^2(N \otimes \operatorname{im}\beta) = 0$.

Therefore, it follows from Lemma 11 that there exists an epimorphism $\boldsymbol{s} : \boldsymbol{F} \twoheadrightarrow \boldsymbol{G}$ on $U \times \mathbb{P}^3$ extending $s : N \twoheadrightarrow Q_E$. Let $\boldsymbol{E} := \ker \boldsymbol{s}$; clearly, $E_0 := \boldsymbol{E}|_{\{0\}\times\mathbb{P}^3} = E$. For $0 \neq u \in U$, E_u fits into the exact sequence

$$0 \to E_u \to N \to G_u \to 0.$$

In the case (i) described above, E_u lies within $\mathcal{D}(2,3)$ for $u \neq 0$, hence $E = E_0$ lies within $\overline{\mathcal{D}(2,3)}$, which is contained in $\overline{\mathcal{I}(3)}$ by [Jardim et al. 2015, Theorem 7.8]. In other words, E can be deformed into a locally free instanton sheaf of charge 3, thus it lies within $\mathcal{I}^0(3)$.

In the case (ii), Proposition 10 also implies that $[E_0] \in \mathcal{I}^0(3)$.

An argument similar to the one used in the proof of [Jardim et al. 2015, Proposition 7.2] works to show that E can be deformed into a locally free ('t Hooft) instanton sheaf.

Summing up, we conclude that if $d_E = 2$, then E lies within $\mathcal{I}^0(3)$.

Finally, consider $d_E = 3$, so that $E^{\vee\vee} = 2\cdot\mathcal{O}_{\mathbb{P}^3}$. Since Q_E is semistable, it follows from Proposition 19 that one can find an affine open subset $0 \in U \subset \mathbb{A}^1$ and a coherent sheaf $\boldsymbol{G}$ on $\mathbb{P}^3 \times U$ such that $G_0 = Q_E$ and, for $u \neq 0$, either

(i) $G_u = \mathcal{O}_\Delta(5\text{pt})$, where Δ is a nonsingular twisted cubic in $\mathbb{P}^3$; or

(ii) $G_u = \mathcal{O}_\Gamma(3\text{pt}) \oplus \mathcal{O}_\ell(1)$, where Γ is a nonsingular conic and ℓ is a line disjoint from Γ; or

(iii) $G_u = \mathcal{O}_{\ell_1}(1) \oplus \mathcal{O}_{\ell_2}(1) \oplus \mathcal{O}_{\ell_3}(1)$ where ℓ_j are 3 skew lines in $\mathbb{P}^3$; or

(iv) $G_u = L(2)$, where $L \in \operatorname{Pic}^0(\Sigma)$, for some nonsingular plane cubic Σ in $\mathbb{P}^3$.

Now set $\boldsymbol{F} := 2\cdot\pi^*\mathcal{O}_{\mathbb{P}^3}$. Note that $H^i(\mathcal{H}om(F_u, G_u)) = H^i(2\cdot G_u)$, and this vanishes for $i = 1, 2, 3$ in all of the four cases outlined above for $u \neq 0$. For $u = 0$, $H^i(G_0) = H^i(Q_E)$ and this vanishes by dimension of Q_E when $i = 2, 3$, and by the vanishing of $h^1(Q_E(-2))$ when $i = 1$.

We complete the argument as before; again, it follows from Lemma 11 that there exists an epimorphism $\boldsymbol{s} : \boldsymbol{F} \twoheadrightarrow \boldsymbol{G}$ extending the epimorphism $s : 2\cdot\mathcal{O}_{\mathbb{P}^3} \twoheadrightarrow Q_E$ obtained from the standard sequence (3) for E. Let $\boldsymbol{E} := \ker \boldsymbol{s}$; then clearly, $E_0 := \boldsymbol{E}|_{\{0\}\times\mathbb{P}^3} = E$. For $u \neq 0$, E_u fits into the exact sequence

$$0 \to E_u \to 2\cdot\mathcal{O}_{\mathbb{P}^3} \to G_u \to 0.$$

In the cases (i) through (iii), we know from [Jardim et al. 2015, Theorem 7.8] and Proposition 10 above that $[E_0] \in \overline{\mathcal{D}(3,3)}$, thus also $[E] \in \mathcal{I}^0(3)$.

In the case (iv), E_u lies within $\mathcal{C}(1,3,0)$ for $u \neq 0$, by definition. It follows that $[E] \in \overline{\mathcal{C}(1,3,0)}$.

This completes the proof of the first part of Main Theorem 1.

6. Components and connectedness of $\mathcal{L}(4)$

In this section we prove the second part of Main Theorem 1, i.e., we enumerate the irreducible components of $\mathcal{L}(4)$, and show that $\mathcal{L}(4)$ is connected. Note that, from Theorem 8, we already know four irreducible components of $\overline{\mathcal{L}(4)}$:

(I) $\overline{\mathcal{I}(4)}$, whose generic point corresponds to a locally free instanton sheaf.

(II) $\overline{\mathcal{C}(1,3,1)}$, whose generic point corresponds to an instanton sheaf E fitting into an exact sequence of the form

$$0 \to E \to N \to \iota_* L(2) \to 0, \tag{16}$$

where N is a null-correlation bundle, $\iota : \Sigma \hookrightarrow \mathbb{P}^3$ is the inclusion of a nonsingular plane cubic Σ, and $L \in \mathrm{Pic}^0(\Sigma)$ is such that $h^0(\Sigma, L) = 0$.

(III) $\overline{\mathcal{C}(2,2,0)}$, whose generic point corresponds to an instanton sheaf E fitting into an exact sequence of the form

$$0 \to E \to 2 \cdot \mathcal{O}_{\mathbb{P}^3} \to \iota_* L(2) \to 0, \tag{17}$$

where $\iota : \Sigma \hookrightarrow \mathbb{P}^3$ is the inclusion of a nonsingular elliptic space quartic Σ, and $L \in \mathrm{Pic}^0(\Sigma)$ is such that $h^0(\Sigma, L) = 0$.

(IV) $\overline{\mathcal{C}(1,4,0)}$, whose generic point corresponds to an instanton sheaf E fitting into an exact sequence of the form

$$0 \to E \to 2 \cdot \mathcal{O}_{\mathbb{P}^3} \to \iota_* L(2) \to 0, \tag{18}$$

where $\iota : \Sigma \hookrightarrow \mathbb{P}^3$ is the inclusion of a nonsingular plane quartic Σ, and $L \in \mathrm{Pic}^2(\Sigma)$ is such that $h^0(\Sigma, L) = 0$.

The first three components have dimension 29, and the last one has dimension 32; this is a classical result for the component $\mathcal{I}^0(4)$, while the dimensions of $\mathcal{C}(1,3,1)$, $\mathcal{C}(2,2,0)$ and $\mathcal{C}(1,4,0)$ are given by Theorem 8 above. Furthermore, [Jardim et al. 2017, Theorem 23] implies that each of the last three components intersects $\mathcal{I}^0(4)$. Thus the union of these four components is connected.

To finish the proof of the second part of Main Theorem 1, it is again enough to show that there are no other irreducible components in $\mathcal{L}(4)$, except for those described above. The argument here is the same as before, exploring Theorem 1, Remark 2 and Proposition 20.

Take any $[E] \in \mathcal{L}(4)$ and consider the triple (3). Then, in view of Theorem 1 and Remark 2, Q_E is a rank 0 instanton sheaf of multiplicity $1 \le d_E \le 4$, and $E^{\vee\vee}$ is an instanton bundle of charge $4 - d_E$. Consider the possible cases for d_E.

The case $d_E = 1$. As in the similar case in Section 5, $Q_E = \mathcal{O}_\ell(1)$ where l is a line in $\mathbb{P}^3$. Respectively, $[E^{\vee\vee}] \in \mathcal{I}(3)$. Deforming ℓ in $\mathbb{P}^3$ we may assume that $E^{\vee\vee}|_\ell \simeq 2\cdot\mathcal{O}_\ell$, so that $[E] \in \mathcal{D}(1,4)$. Therefore, $[E] \in \mathcal{I}^0(4)$.

The case $d_E = 2$. As in the similar case in Section 5, Q_E can be deformed in a flat family either into a sheaf $\mathcal{O}_\Gamma(3\mathrm{pt})$, where Γ is a nonsingular conic in $\mathbb{P}^3$, or into a sheaf $\mathcal{O}_{\ell_1}(1) \oplus \mathcal{O}_{\ell_2}(1)$ where ℓ_1 and ℓ_2 are skew lines in $\mathbb{P}^3$. Respectively, $[E^{\vee\vee}] \in \mathcal{I}(2)$. Now the same argument as in Section 5 shows that $[E] \in \mathcal{I}^0(4)$.

The case $d_E = 3$. Then $E^{\vee\vee}$ is a null-correlation bundle and, as in the case $d_E = 3$ of Section 5, the sheaf Q_E deforms in a flat family to one of the sheaves:

(i) $L(2)$, where $L \in \mathrm{Pic}^0(\Sigma)$, for some nonsingular plane cubic Σ in $\mathbb{P}^3$.

(ii) $\mathcal{O}_\Delta(5\mathrm{pt})$, where Δ is a nonsingular twisted cubic in $\mathbb{P}^3$.

(iii) $\mathcal{O}_\Gamma(3\mathrm{pt}) \oplus \mathcal{O}_\ell(1)$, where Γ is a nonsingular conic and ℓ is a line disjoint from Γ.

(iv) $\mathcal{O}_{\ell_1}(1) \oplus \mathcal{O}_{\ell_2}(1) \oplus \mathcal{O}_{\ell_3}(1)$ where ℓ_j are 3 skew lines in $\mathbb{P}^3$.

By definition, $[E] \in \overline{\mathcal{C}(1,3,1)}$ in the case (i). The same argument as in Section 5, based on [Jardim et al. 2015, Theorem 7.8] and Proposition 10, shows that $[E] \in \overline{\mathcal{I}(4)}$ in the cases (ii) through (iv).

The case $d_E = 4$. Then $E^{\vee\vee} \simeq 2\cdot\mathcal{O}_{\mathbb{P}^3}$ and, according to Proposition 20, the sheaf Q_E deforms in a flat family to one of the sheaves:

(i) $L(2)$, where $L \in \mathrm{Pic}^2(\Sigma)$, for some nonsingular plane quartic Σ in $\mathbb{P}^3$, and L satisfies an open condition $h^1(L) = 0$.

(ii) $L(2)$, where $0 \neq L \in \mathrm{Pic}^0(\Delta)$, for some nonsingular space elliptic quartic Δ in $\mathbb{P}^3$.

(iii) $\mathcal{O}_\Delta(7\mathrm{pt})$ for some nonsingular rational space quartic Δ in $\mathbb{P}^3$.

(iv) $L(2) \oplus \mathcal{O}_\ell(1)$, where $L \in \mathrm{Pic}^0(\Sigma)$, for some nonsingular plane cubic Σ in $\mathbb{P}^3$ and a line ℓ disjoint from Σ.

(v) $\mathcal{O}_\Delta(5\mathrm{pt}) \oplus \mathcal{O}_\ell(1)$, where Δ is a nonsingular twisted cubic and ℓ is a line disjoint from Δ.

(vi) $\mathcal{O}_{\Gamma_1}(3\mathrm{pt}) \oplus \mathcal{O}_{\Gamma_2}(3\mathrm{pt})$, where Γ_1 and Γ_2 are nonsingular, disjoint conics.

(vii) $\mathcal{O}_\Gamma(3\mathrm{pt}) \oplus \mathcal{O}_{\ell_1}(1) \oplus \mathcal{O}_{\ell_2}(1)$, where Γ is a nonsingular conic, and ℓ_1 and ℓ_2 are two skew lines disjoint from Γ.

(viii) $\mathcal{O}_\ell(1) \oplus \mathcal{O}_{\ell_2}(1) \oplus \mathcal{O}_{\ell_3}(1) \oplus \mathcal{O}_{\ell_4}(1)$, where ℓ_1, ℓ_2, ℓ_3, ℓ_4 are four disjoint lines in $\mathbb{P}^3$.

In the case (i), since, in the notation of Lemma 11, $F_0 = 2\cdot\mathcal{O}_{\mathbb{P}^3}$ and $G_0 = L$, $H^i(\mathcal{H}om(F_0, G_0)) = 0$, where $i \geq 1$, and therefore the condition (10) is satisfied by the semicontinuity, so that the deformation argument as above shows that $E \in \overline{\mathcal{C}(1,4,0)}$.

In case (ii), by the same reason, $[E] \in \overline{\mathcal{C}(2,2,0)}$.

In case (iii), a similar argument shows that $[E] \in \overline{\mathcal{D}(4,4)}$, and thus $[E] \in \overline{\mathcal{I}(4)}$.

In case (iv), Proposition 12 guarantees that $[E] \in \overline{\mathcal{C}(1,3,1)}$.

In the cases remaining, (v) through (viii), as in cases (ii) and (iii) for $d_E = 3$ above, we again obtain $[E] \in \overline{\mathcal{I}(4)}$.

Main Theorem 1 is finally proved.

Acknowledgements

Jardim is partially supported by the CNPq grant number 303332/2014-0, and the FAPESP grants number 2014/14743-8 and 2016/03759-6; this work was completed during a visit to the University of Edinburgh, and he is grateful for its hospitality. Jardim also thanks Daniele Faenzi and Simone Marchesi for their help in the proof of Theorem 4. Tikhomirov was supported by a subsidy to the HSE from the Government of the Russian Federation for the implementation of the Global Competitiveness Program. Tikhomirov also acknowledges the support from the Max Planck Institute for Mathematics in Bonn, where this work was finished during the winter of 2017.

References

[Ballico and Huh 2014] E. Ballico and S. Huh, "Stable sheaves on a smooth quadratic surface with linear Hilbert bipolynomials", *Sci. World J.* **2014** (2014), art. id. 346126.

[Chen and Nollet 2012] D. Chen and S. Nollet, "Detaching embedded points", *Algebra Number Theory* **6**:4 (2012), 731–756. MR Zbl

[Choi et al. 2016] J. Choi, K. Chung, and M. Maican, "Moduli of sheaves supported on quartic space curves", *Michigan Math. J.* **65**:3 (2016), 637–671. MR Zbl

[Costa and Ottaviani 2003] L. Costa and G. Ottaviani, "Nondegenerate multidimensional matrices and instanton bundles", *Trans. Amer. Math. Soc.* **355**:1 (2003), 49–55. MR Zbl

[Drézet and Maican 2011] J.-M. Drézet and M. Maican, "On the geometry of the moduli spaces of semi-stable sheaves supported on plane quartics", *Geom. Dedicata* **152** (2011), 17–49. MR Zbl

[Freiermuth and Trautmann 2004] H. G. Freiermuth and G. Trautmann, "On the moduli scheme of stable sheaves supported on cubic space curves", *Amer. J. Math.* **126**:2 (2004), 363–393. MR

[Gargate and Jardim 2016] M. Gargate and M. Jardim, "Singular loci of instanton sheaves on projective space", *Internat. J. Math.* **27**:7 (2016), 1640006, 18. MR Zbl

[Hauzer and Langer 2011] M. Hauzer and A. Langer, "Moduli spaces of framed perverse instantons on $\mathbb{P}^3$", *Glasg. Math. J.* **53**:1 (2011), 51–96. MR Zbl

[Huxley 2003] M. N. Huxley, "Exponential sums and lattice points, III", *Proc. London Math. Soc.* (3) **87**:3 (2003), 591–609. MR Zbl

[Huybrechts and Lehn 1997] D. Huybrechts and M. Lehn, *The geometry of moduli spaces of sheaves*, Aspects of Mathematics **E31**, Friedr. Vieweg Sohn, Braunschweig, 1997. MR Zbl

[Jardim 2006] M. Jardim, "Instanton sheaves on complex projective spaces", *Collect. Math.* **57**:1 (2006), 69–91. MR Zbl

[Jardim and Verbitsky 2014] M. Jardim and M. Verbitsky, "Trihyperkähler reduction and instanton bundles on $\mathbb{CP}^3$", *Compos. Math.* **150**:11 (2014), 1836–1868. MR Zbl

[Jardim et al. 2015] M. Jardim, D. Markushevich, and A. S. Tikhomirov, “New divisors in the boundary of the instanton moduli space”, preprint, 2015. arXiv

[Jardim et al. 2017] M. Jardim, D. Markushevich, and A. S. Tikhomirov, “Two infinite series of moduli spaces of rank 2 sheaves on $\mathbb{P}^3$”, *Ann. Mat. Pura Appl.* (2017).

[Le Potier 1993a] J. Le Potier, “Faisceaux semi-stables de dimension 1 sur le plan projectif”, *Rev. Roumaine Math. Pures Appl.* **38**:7-8 (1993), 635–678. MR Zbl

[Le Potier 1993b] J. Le Potier, *Systèmes cohérents et structures de niveau*, Astérisque **214**, Société Mathématique de France, Paris, 1993. MR Zbl

[Maican 2010] M. Maican, “A duality result for moduli spaces of semistable sheaves supported on projective curves”, *Rend. Semin. Mat. Univ. Padova* **123** (2010), 55–68. MR Zbl

[Maican 2017] M. Maican, “Moduli of space sheaves with Hilbert polynomial $4m + 1$”, *Can. Math. Bull.* (online publication April 2017).

[Shafarevich 1994] I. R. Shafarevich, *Basic algebraic geometry, I*, 2nd ed., Springer, Berlin, 1994. MR Zbl

[Tikhomirov 2012] A. S. Tikhomirov, “Moduli of mathematical instanton vector bundles with odd c_2 on projective space”, *Izv. Ross. Akad. Nauk Ser. Mat.* **76**:5 (2012), 143–224. In Russian; translated in *Izv. Math.* **76**:5 (2012), 991–1073. MR Zbl

[Tikhomirov 2013] A. S. Tikhomirov, “Moduli of mathematical instanton vector bundles with even c_2 on projective space”, *Izv. Ross. Akad. Nauk Ser. Mat.* **77**:6 (2013), 139–168. In Russian; translated in **77**:6 (2013), 1195–1223. MR Zbl

Received February 23, 2017. Revised May 10, 2017.

MARCOS JARDIM
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
UNIVERSITY OF CAMPINAS
RUA SÉGIO BUARQUE DE HOLANDA, 651
CIDADE UNIVERSITÁRIA
13083-859 CAMPINAS-
BRAZIL
jardim@ime.unicamp.br

MARIO MAICAN
INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY
CALEA GRIVITEI 21
010702 BUCHAREST
ROMANIA
maican@imar.ro

ALEXANDER S. TIKHOMIROV
DEPARTMENT OF MATHEMATICS
NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS
6 USACHEVA STREET
MOSCOW
119048
RUSSIA
astikhomirov@mail.ru

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.425

A SYMMETRIC 2-TENSOR CANONICALLY ASSOCIATED TO Q-CURVATURE AND ITS APPLICATIONS

YUEH-JU LIN AND WEI YUAN

We define a symmetric 2-tensor, called the J-tensor, canonically associated to the Q-curvature on any Riemannian manifold with dimension at least three. The relation between the J-tensor and the Q-curvature is like that between the Ricci tensor and the scalar curvature. Thus the J-tensor can be interpreted as a higher-order analogue of the Ricci tensor. This tensor can be used to understand the Chang–Gursky–Yang theorem on 4-dimensional Q-singular metrics. We show that an *almost-Schur lemma* holds for the Q-curvature, yielding an estimate of the Q-curvature on closed manifolds.

1. Introduction

Let M be a smooth manifold and $\mathcal{M}$ be the space of all metrics on M. Consider scalar curvature as a nonlinear map

$$R : \mathcal{M} \to C^\infty(M), \quad g \mapsto R_g.$$

It is well known that the linearization of scalar curvature at a given metric g is

$$\gamma_g h := DR_g \cdot h = -\Delta_g \operatorname{tr}_g h + \delta_g^2 h - \operatorname{Ric}_g \cdot h, \tag{1-1}$$

where $h \in S_2(M)$ is a symmetric 2-tensor and $\delta_g = -\operatorname{div}_g$; see [Besse 1987; Chow et al. 2006; Fischer and Marsden 1975]. Thus, its L^2-formal adjoint is given by

$$\gamma_g^* f = \nabla_g^2 f - g\Delta_g f - f \operatorname{Ric}_g \tag{1-2}$$

for any smooth function $f \in C^\infty(M)$.

An interesting observation is that, if we take f to be constantly 1, we get

$$\operatorname{Ric}_g = -\gamma_g^* 1.$$

This work was partially supported by NSF (Grant No. DMS-1440140), NSFC (Grant No. 11521101, No. 11601531), *The Fundamental Research Funds for the Central Universities* (Grant No.2016-34000-31610258).
MSC2010: 53C20, 53C25.
Keywords: J-tensor, Q-curvature, Q-singular metric.

That means we can recover Ricci tensor from γ_g^*. Furthermore, the scalar curvature is given by

$$R_g = -\operatorname{tr}_g \gamma_g^* 1.$$

Now let (M^n, g) be an n-dimensional Riemannian manifold ($n \geq 3$). We can define the Q-curvature to be

$$Q_g = A_n \Delta_g R_g + B_n |\mathrm{Ric}_g|_g^2 + C_n R_g^2, \tag{1-3}$$

where

$$A_n = -\frac{1}{2(n-1)}, \quad B_n = -\frac{2}{(n-2)^2}, \quad C_n = \frac{n^2(n-4)+16(n-1)}{8(n-1)^2(n-2)^2}.$$

In fact, Q-curvature was introduced originally to generalize the classic *Gauss-Bonnet theorem* on surfaces to closed 4-manifolds (M^4, g):

$$\int_{M^4} \left(Q_g + \tfrac{1}{4}|W_g|_g^2\right) dv_g = 8\pi^2 \chi(M), \tag{1-4}$$

where W_g is the Weyl tensor.

Paneitz and Branson extended Q-curvature to any dimension $n \geq 3$ (see [Branson 1985; Paneitz 2008]) such that it satisfies certain conformal invariant properties. For more details, please refer to the appendix of [Lin and Yuan 2016].

Like the scalar curvature, we can also view Q-curvature as a nonlinear map

$$Q : \mathcal{M} \to C^\infty(M), \quad g \mapsto Q_g.$$

Let $\Gamma_g : S_2(M) \to C^\infty(M)$ be the linearization of Q-curvature at the metric g and $\Gamma_g^* : C^\infty(M) \to S_2(M)$ be its L^2-formal adjoint.

Now we can define the central notion in this article:

Definition 1.1. Let (M^n, g) be a Riemannian manifold ($n \geq 3$). We define the symmetric 2-tensor

$$J_g := -\tfrac{1}{2}\Gamma_g^* 1.$$

We say (M, g) is *J-Einstein* if $J_g = \Lambda g$ for some smooth function $\Lambda \in C^\infty(M)$. In particular, it is *J-flat* if $\Lambda = 0$.

In [Lin and Yuan 2016], we calculated the explicit expression of Γ_g^* and showed

$$\operatorname{tr}_g \Gamma_g^* f = \frac{1}{2}\left(P_g - \frac{n+4}{2} Q_g\right) f, \tag{1-5}$$

for any $f \in C^\infty(M)$. Here P_g is the *Paneitz operator* defined by

$$P_g = \Delta_g^2 - \operatorname{div}_g[(a_n R_g g + b_n \operatorname{Ric}_g)d] + \frac{n-4}{2} Q_g, \tag{1-6}$$

where

$$a_n = \frac{(n-2)^2+4}{2(n-1)(n-2)} \quad \text{and} \quad b_n = -\frac{4}{n-2}.$$

In particular, $\operatorname{tr}_g \Gamma_g^* 1 = -2Q_g$. Thus

$$\operatorname{tr}_g J_g = Q_g. \tag{1-7}$$

On the other hand, for any smooth vector field $X \in \mathscr{X}(M)$ on M,

$$\begin{aligned}\int_M \langle X, \delta_g \Gamma_g^* f\rangle \, dv_g &= \frac{1}{2}\int_M \langle L_X g, \Gamma_g^* f\rangle \, dv_g \\ &= \frac{1}{2}\int_M f\Gamma_g(L_X g)\, dv_g = \frac{1}{2}\int_M \langle f\, dQ_g, X\rangle \, dv_g.\end{aligned}$$

Thus

$$\delta_g \Gamma_g^* f = \tfrac{1}{2} f\, dQ_g$$

on M. Hence,

$$\operatorname{div}_g J_g = \tfrac{1}{2}\delta_g \Gamma_g^* 1 = \tfrac{1}{4}\, dQ_g. \tag{1-8}$$

Recall that for Ricci tensor, we have

$$\operatorname{tr}_g \operatorname{Ric}_g = R_g \quad \text{and} \quad \operatorname{div}_g \operatorname{Ric}_g = \tfrac{1}{2} dR_g.$$

Therefore, if we consider Q-curvature as a higher-order analogue of scalar curvature, we can interpret J_g as a higher-order analogue of Ricci curvature on Riemannian manifolds.

A notion closely related to the J-tensor is the *Q-singular metric*, which refers to a metric satisfying $\ker \Gamma_g^* \neq \{0\}$. Clearly, J-flat metrics are Q-singular, since it is equivalent to $1 \in \ker \Gamma_g^*$.

One of the motivations for us to study J-flat manifolds is to understand the following theorem by Chang, Gursky and Yang:

Theorem 1.2 [Chang et al. 2002]. *Let (M^4, g) be a Q-singular 4-manifold. Then $1 \in \ker \Gamma_g^*$ if and only if (M^4, g) is Bach flat with vanishing Q-curvature.*

To achieve our goal, we need to give an explicit expression of the J-tensor:

Theorem 1.3. *For $n \geq 3$,*

$$J_g = \frac{1}{n} Q_g g - \frac{1}{n-2} B_g - \frac{n-4}{4(n-1)(n-2)} T_g, \tag{1-9}$$

where B_g is the Bach tensor and

$$\begin{aligned}T_g := (n-2)\Big(\nabla^2 \operatorname{tr}_g S_g - \frac{1}{n} g \Delta_g \operatorname{tr}_g S_g\Big) \\ + 4(n-1)\Big(S_g \times S_g - \frac{1}{n}|S_g|^2 g\Big) - n^2(\operatorname{tr}_g S_g)\mathring{S}_g.\end{aligned}$$

Here $(S \times S)_{jk} = S^i_j S_{ik}$, S_g is the Schouten tensor and $\mathring{S}_g$ is its traceless part.

Remark 1.4. Note that both the Bach tensor and the tensor T are traceless, thus the traceless part of J is given by

$$\mathring{J}_g = J_g - \frac{1}{n} Q_g g = -\frac{1}{n-2}\Big(B_g + \frac{n-4}{4(n-1)} T_g\Big). \tag{1-10}$$

Thus, an equivalent definition for a metric g being J-Einstein is

$$B_g = -\frac{n-4}{4(n-1)} T_g. \tag{1-11}$$

In particular, when $n = 4$, J-Einstein metrics are exactly Bach flat ones. Hence we can also interpret that J-Einstein metric is a generalization of Bach flat metric on 4-dimensional manifolds.

Remark 1.5. Gursky [1997] introduced a similar tensor for 4-manifolds from the viewpoint of functional determinants. In the same article, he also remarked this tensor can be introduced from the perspective of first variations of total Q-curvature when dimension is at least 5 (see [Case 2012] for a detailed calculation).

With the similar perspective, Gover and Ørsted introduced an abstract tensor called *higher Einstein tensor*, which coincides with our J-tensor in one of its special case. We refer interested readers to their article [Gover and Ørsted 2013].

Note that for any Einstein metric g, its Q-curvature is given by

$$Q_g = B_n |\mathrm{Ric}_g|^2 + C_n R_g^2 = \Big(\frac{1}{n} B_n + C_n\Big) R_g^2 = \frac{(n+2)(n-2)}{8n(n-1)^2} R_g^2,$$

which is a nonnegative constant and vanishes if and only if g is Ricci flat.

It is easy to check that $T_g = 0$ for any Einstein metric g. Combining this with the well-known fact that any Einstein metric is Bach flat, we can easily deduce that any nonflat Einstein metrics are also positive J-Einstein and Ricci flat metrics are J-flat as well.

With the aid of this notion, we can recover and generalize Theorem 1.2 to any dimension $n \geq 3$:

Corollary 1.6. *Let (M^n, g) be a Q-singular n-dimensional Riemannian manifold. Then $1 \in \ker \Gamma_g^*$ if and only if (M^n, g) is J-flat or equivalently (M^n, g) satisfies*

$$B_g = -\frac{n-4}{4(n-1)} T_g$$

with vanishing Q-curvature.

Remark 1.7. In [Chang et al. 2002], Bach flatness in Theorem 1.2 is derived using the variational property of the Bach tensor for 4-manifolds.

As another application of J-tensor, we can derive the *Schur lemma for Q-curvature* as follows:

Theorem 1.8 (Schur lemma). *Let (M^n, g) be an n-dimensional J-Einstein manifold with $n \neq 4$ or equivalently,*

$$B_g = -\frac{n-4}{4(n-1)} T_g.$$

Then Q_g is a constant on M.

Moreover, the following *almost-Schur lemma* holds exactly like the case for Ricci tensor and scalar curvature, cf., [Cheng 2013; De Lellis and Topping 2012; Ge and Wang 2012].

Theorem 1.9 (almost-Schur lemma). *For $n \neq 4$, let (M^n, g) be an n-dimensional closed Riemannian manifold with positive Ricci curvature. Then*

$$\int_M (Q_g - \bar{Q}_g)^2 \, dv_g \leq \frac{16n(n-1)}{(n-4)^2} \int_M |\mathring{J}_g|^2 \, dv_g, \tag{1-12}$$

where $\bar{Q}_g$ is the average of Q_g. Moreover, the equality holds if and only if (M, g) is J-Einstein.

In order to derive an equivalent form of above inequality, we need to define the *J-Schouten tensor* as follows:

$$S_J = \frac{1}{n-4}\Big(J_g - \frac{3}{4(n-1)} Q_g g\Big). \tag{1-13}$$

Immediately, we have

$$\mathrm{tr}_g \, S_J = \frac{1}{4(n-1)} Q_g \tag{1-14}$$

and

$$\mathrm{div}_g \, S_J = \frac{1}{4(n-1)} d Q_g = d \, \mathrm{tr}_g \, S_J. \tag{1-15}$$

Remark 1.10. Recall the definition of classic Schouten tensor

$$S_g = \frac{1}{n-2}\Big(\mathrm{Ric}_g - \frac{1}{2(n-1)} R_g g\Big). \tag{1-16}$$

We have

$$\mathrm{tr}_g \, S_g = \frac{1}{2(n-1)} R_g \tag{1-17}$$

and

$$\mathrm{div}_g \, S_g = \frac{1}{2(n-1)} d R_g = d \, \mathrm{tr}_g \, S_g. \tag{1-18}$$

We can see the tensor S_J shares similar properties with the classic Schouten tensor.

Following the observation in [Ge and Wang 2012], we get immediately the following result by rewriting Theorem 1.9 with J-Schouten tensor:

Corollary 1.11. *For $n \neq 4$, let (M^n, g) be an n-dimensional closed Riemannian manifold with positive Ricci curvature. Then*

$$(\mathrm{Vol}_g M)^{-(n-8)/n} \int_M \sigma_2^J(g)\, dv_g \leq \frac{n-1}{2n} Y_Q^2(g), \tag{1-19}$$

where

$$Y_Q(g) := \frac{\int_M \sigma_1^J(g)\, dv_g}{(\mathrm{Vol}_g M)^{(n-4)/n}}$$

is the ***Q-Yamabe quotient*** *and $\sigma_i^J(g) = \sigma_i(S_J(g))$, $i = 1, 2$ are the i-th symmetric polynomial of $S_J(g)$. Moreover, the equality holds if and only if (M, g) is J-Einstein.*

Remark 1.12. Our *almost-Schur lemma* can be generalized to a broader setting by combining it with the work [Gover and Ørsted 2013]. More detailed discussion together with some related topics will be presented in a subsequent article.

This article is organized as follows: In Section 2, we derive an explicit formula for the J-tensor and with it we prove Theorem 1.3 and Corollary 1.6. We then prove Theorem 1.8 (Schur lemma) and Theorem 1.9 (almost-Schur lemma) in Section 3.

2. J-flatness and Q-singular metrics

We begin with some discussion of conformal tensors. Let

$$S_{jk} = \frac{1}{n-2}\Big(R_{jk} - \frac{1}{2(n-1)} R g_{jk}\Big) \tag{2-1}$$

be the Schouten tensor.

For $n \geq 4$, the Bach tensor is defined to be

$$B_{jk} = \frac{1}{n-3} \nabla^i \nabla^l W_{ijkl} + W_{ijkl} S^{il}. \tag{2-2}$$

In order to extend the definition to $n = 3$, we introduce the Cotton tensor

$$C_{ijk} = \nabla_i S_{jk} - \nabla_j S_{ik}. \tag{2-3}$$

It is related to Weyl tensor by the equation

$$\nabla^l W_{ijkl} = (n-3) C_{ijk}. \tag{2-4}$$

Therefore, for any $n \geq 3$, we can define the Bach tensor as

$$B_{jk} = \nabla^i C_{ijk} + W_{ijkl} S^{il}. \tag{2-5}$$

The following identity is well known for experts; we include calculations here for the convenience of readers.

Proposition 2.1. *The Bach tensor can be written as*

$$B_g = \Delta_g S - \nabla^2 \operatorname{tr} S + 2\mathring{Rm}\cdot S - (n-4)S\times S - |S|^2 g - 2(\operatorname{tr} S)S, \tag{2-6}$$

where $(\mathring{Rm}\cdot S)_{jk} = R_{ijkl}S^{il}$ *and* $(S\times S)_{jk} = S^i_j S_{ik}$. *Equivalently,*

$$B_g = \Delta_L S - \nabla^2 \operatorname{tr} S + n\Big(S\times S - \frac{1}{n}|S|^2 g\Big), \tag{2-7}$$

where Δ_L *is the Lichnerowicz Laplacian.*

Proof. By the *second contracted Bianchi identity*,

$$\begin{aligned}\nabla^i S_{ik} = \frac{1}{n-2}\Big(\nabla^i R_{ik} - \frac{1}{2(n-1)}\nabla_k R\Big) &= \frac{1}{n-2}\Big(\frac{1}{2}\nabla_k R - \frac{1}{2(n-1)}\nabla_k R\Big)\\ &= \frac{1}{2(n-1)}\nabla_k R\\ &= \nabla_k \operatorname{tr} S\end{aligned}$$

and

$$\operatorname{tr} S = \frac{1}{n-2}\Big(R - \frac{n}{2(n-1)}R\Big) = \frac{1}{2(n-1)}R,$$

we have

$$\mathrm{Ric} = (n-2)S + (\operatorname{tr} S)g.$$

Using these facts,

$$\begin{aligned}\nabla^i C_{ijk} &= \nabla^i(\nabla_i S_{jk} - \nabla_j S_{ik})\\ &= \Delta_g S_{jk} - (\nabla_j\nabla_i S^i_k + R^i_{ijp}S^p_k - R^p_{ijk}S^i_p)\\ &= \Delta_g S_{jk} - \nabla_j\nabla_k \operatorname{tr} S - (\mathrm{Ric}\times S)_{jk} + (\mathring{Rm}\cdot S)_{jk}\\ &= \Delta_g S_{jk} - \nabla_j\nabla_k \operatorname{tr} S - (((n-2)S + (\operatorname{tr} S)g)\times S)_{jk} + (\mathring{Rm}\cdot S)_{jk}\\ &= \Delta_g S_{jk} - \nabla_j\nabla_k \operatorname{tr} S - (n-2)(S\times S)_{jk} - (\operatorname{tr} S)S_{jk} + (\mathring{Rm}\cdot S)_{jk}\end{aligned}$$

and

$$\begin{aligned}W_{ijkl}S^{il} &= (Rm - S \owedge g)_{ijkl}S^{il}\\ &= R_{ijkl}S^{il} - (S_{il}g_{jk} + S_{jk}g_{il} - S_{ik}g_{jl} - S_{jl}g_{ik})S^{il}\\ &= (\mathring{Rm}\cdot S)_{jk} - |S|^2 g_{jk} + 2(S\times S)_{jk} - (\operatorname{tr} S)S_{jk},\end{aligned}$$

where $\owedge$ is the *Kulkarni–Nomizu product*:

$$(\alpha \owedge \beta)_{ijkl} := \alpha_{il}\beta_{jk} + \alpha_{jk}\beta_{il} - \alpha_{ik}\beta_{jl} - \alpha_{jl}\beta_{ik}$$

for any symmetric 2-tensor $\alpha,\beta \in S_2(M)$.

Combining them, we get

$$B_{jk} = \Delta_g S_{jk} - \nabla_j\nabla_k \operatorname{tr} S + 2(\mathring{Rm}\cdot S)_{jk} - (n-4)(S\times S)_{jk} - |S|^2 g_{jk} - 2(\operatorname{tr} S)S_{jk}.$$

From this,

$$\begin{aligned} B_{jk} &= \Delta_L S_{jk} + 2(\mathrm{Ric}\times S)_{jk} - \nabla_j\nabla_k \operatorname{tr} S - (n-4)(S\times S)_{jk} - |S|^2 g_{jk} - 2(\operatorname{tr} S) S_{jk} \\ &= \Delta_L S_{jk} + 2((\mathrm{Ric} - (\operatorname{tr} S) g)\times S)_{jk} - \nabla_j\nabla_k \operatorname{tr} S - (n-4)(S\times S)_{jk} - |S|^2 g_{jk} \\ &= \Delta_L S - \nabla^2 \operatorname{tr} S + n(S\times S) - |S|^2 g \\ &= \Delta_L S - \nabla^2 \operatorname{tr} S + n\Big(S\times S - \frac{1}{n}|S|^2 g\Big). \qquad \square \end{aligned}$$

The Q-curvature can also be rewritten using Schouten tensor:

Lemma 2.2. $$Q_g = -\Delta_g \operatorname{tr} S - 2|S|^2 + \frac{n}{2}(\operatorname{tr} S)^2.$$

Proof. Using the equalities $\mathrm{Ric} = (n-2)S + (\operatorname{tr} S)g$ and $R = 2(n-1)\operatorname{tr} S$,

$$\begin{aligned} Q_g &= A_n \Delta_g R + B_n |\mathrm{Ric}|^2 + C_n R^2 \\ &= 2(n-1) A_n \Delta_g \operatorname{tr} S + B_n |(n-2)S + (\operatorname{tr} S) g|^2 + 4(n-1)^2 C_n (\operatorname{tr} S)^2 \\ &= -\Delta_g \operatorname{tr} S - 2|S|^2 + ((3n-4)B_n + 4(n-1)^2 C_n)(\operatorname{tr} S)^2 \\ &= -\Delta_g \operatorname{tr} S - 2|S|^2 + \frac{n}{2}(\operatorname{tr} S)^2. \qquad \square \end{aligned}$$

We recall the expression of Γ_g^* in [Lin and Yuan 2016] as follows:

Lemma 2.3.

$$\begin{aligned} \Gamma_g^* f := A_n\big(&-g\Delta^2 f + \nabla^2\Delta f - \mathrm{Ric}\,\Delta f + \tfrac{1}{2} g\delta(f\,dR) + \nabla(f\,dR) - f\nabla^2 R\big) \\ &- B_n(\Delta(f\,\mathrm{Ric}) + 2f\mathring{Rm}\cdot\mathrm{Ric} + g\delta^2(f\,\mathrm{Ric}) + 2\nabla\delta(f\,\mathrm{Ric})) \\ &- 2C_n(g\Delta(fR) - \nabla^2(fR) + fR\,\mathrm{Ric}). \end{aligned} \tag{2-8}$$

Now we can calculate an explicit expression of J_g:

Theorem 2.4. *For $n \geq 3$,*

$$J_g = \frac{1}{n} Q_g g - \frac{1}{n-2} B_g - \frac{n-4}{4(n-1)(n-2)} T_g, \tag{2-9}$$

where

$$\begin{aligned} T_g := (n-2)\Big(\nabla^2 \operatorname{tr}_g S_g - \frac{1}{n} g\Delta_g \operatorname{tr}_g S_g\Big) \\ + 4(n-1)\Big(S_g\times S_g - \frac{1}{n}|S_g|^2 g\Big) - n^2(\operatorname{tr}_g S_g)\mathring{S}_g. \end{aligned}$$

Here $\mathring{S}_g = S_g - (1/n)\operatorname{tr}_g S_g g$ is the traceless part of Schouten tensor.

Proof. By Lemma 2.3,

$$\begin{aligned} \Gamma_g^* 1 = -\big(\tfrac{1}{2}A_n + \tfrac{1}{2}B_n + 2C_n\big) g\Delta R + (B_n + 2C_n)\nabla^2 R \\ - B_n(\Delta\,\mathrm{Ric} + 2\mathring{Rm}\cdot\mathrm{Ric}) - 2C_n R\,\mathrm{Ric}. \end{aligned}$$

Applying equalities $\mathrm{Ric} = (n-2)S + (\mathrm{tr}\, S)g$ and $R = 2(n-1)\,\mathrm{tr}\, S$,

$$\begin{aligned}
\Gamma_g^* 1 &= -((n-1)A_n + nB_n + 4(n-1)C_n) g \Delta\, \mathrm{tr}\, S + 2(n-1)(B_n + 2C_n)\nabla^2\, \mathrm{tr}\, S \\
&\qquad - (n-2)B_n(\Delta S + 2\mathring{Rm}\cdot S) - 2(n-2)(B_n + 2(n-1)C_n)(\mathrm{tr}\, S)S \\
&\qquad - 2(B_n + 2(n-1)C_n)(\mathrm{tr}\, S)^2 g \\
&= \frac{3}{2(n-1)} g\Delta\, \mathrm{tr}\, S + \frac{2}{n-2}(\Delta S + 2\mathring{Rm}\cdot S) + \frac{n^2-10n+12}{2(n-1)(n-2)}\nabla^2\, \mathrm{tr}\, S \\
&\qquad - \frac{n^2-2n+4}{2(n-1)}(\mathrm{tr}\, S)S - \frac{n^2-2n+4}{2(n-1)(n-2)}(\mathrm{tr}\, S)^2 g.
\end{aligned}$$

Since $\mathrm{tr}\, \Gamma_g^* 1 = -2Q_g$, by Lemma 2.2,

$$\begin{aligned}
&\Gamma_g^* 1 + \frac{2}{n}Q_g g \\
&\quad= \Big(\frac{3}{2(n-1)} - \frac{2}{n}\Big) g\Delta\, \mathrm{tr}\, S + \frac{2}{n-2}(\Delta S + 2\mathring{Rm}\cdot S) + \frac{n^2-10n+12}{2(n-1)(n-2)}\nabla^2\, \mathrm{tr}\, S \\
&\qquad\quad - \frac{4}{n}|S|^2 g - \frac{n^2-2n+4}{2(n-1)}(\mathrm{tr}\, S)S + \Big(1 - \frac{n^2-2n+4}{2(n-1)(n-2)}\Big)(\mathrm{tr}\, S)^2 g \\
&\quad= -\frac{n-4}{2n(n-1)} g\Delta\, \mathrm{tr}\, S + \frac{2}{n-2}(\Delta S + 2\mathring{Rm}\cdot S) + \frac{n^2-10n+12}{2(n-1)(n-2)}\nabla^2\, \mathrm{tr}\, S \\
&\qquad\quad - \frac{4}{n}|S|^2 g - \frac{n^2-2n+4}{2(n-1)}(\mathrm{tr}\, S)S + \frac{n(n-4)}{2(n-1)(n-2)}(\mathrm{tr}\, S)^2 g.
\end{aligned}$$

Applying Proposition 2.1,

$$\begin{aligned}
\Gamma_g^* 1 + \frac{2}{n}Q_g g &= \frac{2}{n-2}B_g - \frac{n-4}{2n(n-1)} g\Delta\, \mathrm{tr}\, S + \Big(\frac{2}{n-2} + \frac{n^2-10n+12}{2(n-1)(n-2)}\Big)\nabla^2\, \mathrm{tr}\, S \\
&\qquad + \frac{2(n-4)}{n-2} S\times S + \Big(\frac{2}{n-2} - \frac{4}{n}\Big)|S|^2 g \\
&\qquad + \Big(\frac{4}{n-2} - \frac{n^2-2n+4}{2(n-1)}\Big)(\mathrm{tr}\, S)S + \frac{n(n-4)}{2(n-1)(n-2)}(\mathrm{tr}\, S)^2 g.
\end{aligned}$$

That is,

$$\begin{aligned}
\Gamma_g^* 1 + \frac{2}{n}Q_g g &= \frac{2}{n-2}B_g - \frac{n-4}{2n(n-1)} g\Delta \mathrm{tr} S + \frac{n-4}{2(n-1)}\nabla^2 \mathrm{tr} S + \frac{2(n-4)}{n-2} S\times S \\
&\qquad - \frac{2(n-4)}{n(n-2)}|S|^2 g - \frac{n^2(n-4)}{2(n-1)(n-2)}(\mathrm{tr} S)S + \frac{n(n-4)}{2(n-1)(n-2)}(\mathrm{tr} S)^2 g \\
&= \frac{2}{n-2}B_g + \frac{n-4}{2(n-1)}\Big(\nabla^2 \mathrm{tr} S - \frac{1}{n} g\Delta \mathrm{tr} S\Big) + \frac{2(n-4)}{n-2}\Big(S\times S - \frac{1}{n}|S|^2 g\Big) \\
&\qquad - \frac{n^2(n-4)}{2(n-1)(n-2)}(\mathrm{tr} S)\Big(S - \frac{1}{n}(\mathrm{tr} S) g\Big) \\
&= \frac{2}{n-2}B_g + \frac{n-4}{2(n-1)(n-2)}T_g,
\end{aligned}$$

where

$$T_g := (n-2)\Big(\nabla^2 \operatorname{tr}_g S_g - \frac{1}{n} g \Delta_g \operatorname{tr}_g S_g\Big) \\ + 4(n-1)\Big(S_g \times S_g - \frac{1}{n}|S_g|^2 g\Big) - n^2 (\operatorname{tr}_g S_g) \mathring{S}_g.$$

Therefore,

$$J_g = -\frac{1}{2}\Gamma_g^* 1 = \frac{1}{n} Q_g g - \frac{1}{n-2} B_g - \frac{n-4}{4(n-1)(n-2)} T_g. \qquad \square$$

Immediately, we have the following generalization of Theorem 1.2:

Corollary 2.5. *Let (M^n, g) be a Q-singular n-dimensional Riemannian manifold. Then $1 \in \ker \Gamma_g^*$ if and only if (M^n, g) is J-flat or equivalently (M^n, g) satisfies*

$$B_g = -\frac{n-4}{4(n-1)} T_g \tag{2-10}$$

with vanishing Q-curvature.

Remark 2.6. A similar result holds for Ricci curvature: a vacuum static space admits a constant static potential if and only if it is Ricci flat, cf., [Fischer and Marsden 1975].

3. An almost-Schur lemma for Q-curvature

Since the tensor J_g can be interpreted as a higher-order analogue of Ricci tensor, we can also derive the Schur lemma for J_g as follows:

Theorem 3.1 (Schur lemma). *Let (M^n, g) be an n-dimensional J-Einstein manifold with $n \neq 4$ or equivalently,*

$$B_g = -\frac{n-4}{4(n-1)} T_g.$$

Then Q_g is a constant on M.

Proof. By the assumption, $J_g = \Lambda g$ for some smooth function Λ on M. Then

$$\Lambda = \frac{1}{n} \operatorname{tr}_g J_g = \frac{1}{n} Q_g \quad \text{and} \quad d\Lambda = \operatorname{div}_g J_g = \frac{1}{4} dQ_g.$$

Therefore,

$$\frac{n-4}{4n} dQ_g = 0$$

on M, which implies that Q_g is a constant on M provided $n \neq 4$. $\square$

Remark 3.2. When $n = 4$, J-Einstein metrics are exactly Bach flat ones. Due to the conformal invariance of Bach flatness in dimension 4, we can easily see that the constancy of Q-curvature can not always be achieved. Thus the above Schur

Lemma does not hold for 4-dimensional manifolds, which is exactly like the classic Schur lemma for surfaces.

In fact, a more general result can be derived:

Theorem 3.3 (almost-Schur lemma). *For $n \neq 4$, let (M^n, g) be an n-dimensional closed Riemannian manifold with positive Ricci curvature. Then*

$$\int_M (Q_g - \bar{Q}_g)^2 \, dv_g \leq \frac{16n(n-1)}{(n-4)^2} \int_M |\mathring{J}_g|^2 \, dv_g, \tag{3-1}$$

where $\bar{Q}_g$ is the average of Q_g. Moreover, the equality holds if and only if (M^n, g) is J-Einstein.

The proof is along the same lines as in [De Lellis and Topping 2012]. For completeness, we include it here. For more details, please refer to that work.

Proof. Let u be the unique solution to

$$\begin{cases} \Delta_g u & = Q_g - \bar{Q}_g, \\ \int_M u \, dv_g & = 0. \end{cases}$$

Then

$$\begin{aligned} \int_M (Q_g - \bar{Q}_g)^2 \, dv_g = \int_M (Q_g - \bar{Q}_g) \Delta_g u \, dv_g &= -\int_M \langle \nabla Q_g, \nabla u \rangle \, dv_g \\ &= -\frac{4n}{n-4} \int_M \langle \operatorname{div}_g \mathring{J}_g, \nabla u \rangle, \end{aligned}$$

where for the last step we use the fact

$$\operatorname{div}_g \mathring{J}_g = \operatorname{div}_g \Big(J_g - \frac{1}{n} Q_g g \Big) = \frac{1}{4} \, d Q_g - \frac{1}{n} \, d Q_g = \frac{n-4}{4n} \, d Q_g.$$

Integrating by parts,

$$\begin{aligned} -\frac{4n}{n-4} \int_M \langle \operatorname{div}_g \mathring{J}_g, \nabla u \rangle \, dv_g &= \frac{4n}{n-4} \int_M \langle \mathring{J}_g, \nabla^2 u \rangle \, dv_g \\ &= \frac{4n}{n-4} \int_M \Big\langle \mathring{J}_g, \nabla^2 u - \frac{1}{n} g \Delta_g u \Big\rangle \, dv_g \\ &\leq \frac{4n}{n-4} \Big(\int_M |\mathring{J}_g|^2 \, dv_g \Big)^{1/2} \Big(\int_M \Big| \nabla^2 u - \frac{1}{n} g \Delta_g u \Big|^2 \, dv_g \Big)^{1/2} \\ &= \frac{4n}{n-4} \Big(\int_M |\mathring{J}_g|^2 dv_g \Big)^{1/2} \Big(\int_M |\nabla^2 u|^2 - \frac{1}{n} (\Delta_g u)^2 \, dv_g \Big)^{1/2}. \end{aligned}$$

From the *Bochner formula* and the assumption $\operatorname{Ric}_g > 0$,

$$\int_M |\nabla^2 u|^2 \, dv_g = \int_M (\Delta_g u)^2 \, dv_g - \int_M \operatorname{Ric}_g(\nabla u, \nabla u) \, dv_g \leq \int_M (\Delta_g u)^2 \, dv_g.$$

Thus,

$$\begin{aligned}\int_M (Q_g - \bar{Q}_g)^2\, dv_g &\le \frac{4n}{n-4}\Big(\int_M |\mathring{J}_g|^2\, dv_g\Big)^{1/2}\Big(\frac{n-1}{n}(\Delta_g u)^2\, dv_g\Big)^{1/2}\\ &= \frac{4n}{n-4}\Big(\int_M |\mathring{J}_g|^2\, dv_g\Big)^{1/2}\Big(\frac{n-1}{n}(Q_g - \bar{Q}_g)^2\, dv_g\Big)^{1/2}.\end{aligned}$$

That is,

$$\int_M (Q_g - \bar{Q}_g)^2\, dv_g \le \frac{16n(n-1)}{(n-4)^2}\int_M |\mathring{J}_g|^2\, dv_g.$$

Now we consider the equality case.

If g is J-Einstein, then Q_g is a constant by the *Schur lemma* (Theorem 1.8). Thus both sides of inequality (3-1) vanish and equality is achieved.

On the contrary, assume in (3-1) equality is achieved:

$$\int_M (Q_g - \bar{Q}_g)^2\, dv_g = \frac{16n(n-1)}{(n-4)^2}\int_M |\mathring{J}_g|^2\, dv_g.$$

Then in particular we have $\mathrm{Ric}(\nabla u, \nabla u) = 0$, which implies that $\nabla u = 0$ and hence u is a constant on M, since we assume $\mathrm{Ric}_g > 0$.

Thus $Q \equiv \bar{Q}$ on M and

$$\int_M |\mathring{J}_g|^2\, dv_g = \frac{(n-4)^2}{16n(n-1)}\int_M (Q_g - \bar{Q}_g)^2\, dv_g = 0.$$

Therefore, $\mathring{J}_g \equiv 0$ on M, i.e., (M, g) is J-Einstein. □

Remark 3.4. By assuming $\mathrm{Ric} \ge -(n-1)Kg$ for some constant $K \ge 0$ and following the proof in [Cheng 2013], the inequality (3-1) can be improved to

$$\int_M (Q_g - \bar{Q}_g)^2\, dv_g \le \frac{16n(n-1)}{(n-4)^2}\Big(1 + \frac{nK}{\lambda_1}\Big)\int_M |\mathring{J}_g|^2\, dv_g, \tag{3-2}$$

where $\lambda_1 > 0$ is the first nonzero eigenvalue of $(-\Delta_g)$.

Now we can derive an equivalent form of inequality (3-1):

Corollary 3.5. *For $n \ne 4$, let (M^n, g) be an n-dimensional closed Riemannian manifold with positive Ricci curvature. Then*

$$(\mathrm{Vol}_g\, M)^{-(n-8)/n}\int_M \sigma_2^J(g)\, dv_g \le \frac{n-1}{2n} Y_Q^2(g). \tag{3-3}$$

Moreover, the equality holds if and only if (M^n, g) is J-Einstein.

Proof. Note that

$$\sigma_1^J(g) = \mathrm{tr}_g\, S_J = \frac{1}{4(n-1)} Q_g$$

and

$$\sigma_2^J(g) = \frac{1}{2}((\sigma_1^J)^2 - |S_J|^2) = \frac{n-1}{2n}(\sigma_1^J)^2 - \frac{1}{2(n-4)^2}|\mathring{J}_g|^2,$$

where we use the fact

$$|S_J|^2 = \left|\mathring{S}_J + \frac{1}{n}(\operatorname{tr}_g S_J)g\right|^2 = \left|\frac{1}{n-4}\mathring{J}_g + \frac{1}{n}(\sigma_1^J)g\right|^2 = \frac{1}{(n-4)^2}|\mathring{J}_g|^2 + \frac{1}{n}(\sigma_1^J)^2.$$

By substituting these terms in the inequality (3-1), we get

$$\left(\int_M \sigma_1^J(g)\,dv_g\right)^2 \geq \frac{2n}{n-1}\operatorname{Vol}_g(M)\int_M \sigma_2^J(g)\,dv_g.$$

Therefore,

$$\begin{aligned}\int_M \sigma_2^J(g)\,dv_g &\leq \frac{n-1}{2n}(\operatorname{Vol}_g M)^{-1}\left(\int_M \sigma_1^J(g)\,dv_g\right)^2\\ &= \frac{n-1}{2n}(\operatorname{Vol}_g M)^{(n-8)/n}\left(\frac{\int_M \sigma_1^J(g)\,dv_g}{(\operatorname{Vol}_g M)^{(n-4)/n}}\right)^2\\ &= \frac{n-1}{2n}(\operatorname{Vol}_g M)^{(n-8)/n}Y_Q^2(g).\end{aligned}$$ □

Remark 3.6. Note that the *Q-Yamabe quotient*

$$Y_Q(g) := \frac{\int_M \sigma_1^J(g)\,dv_g}{(\operatorname{Vol}_g M)^{(n-4)/n}}$$

is scaling invariant and in particular, when $n = 8$,

$$\int_M \sigma_2^J(g)\,dv_g \leq \frac{7}{16}Y_Q^2(g),$$

provided that $\operatorname{Ric}_g > 0$, where the equality holds if and only if (M, g) is J-Einstein.

Acknowledgement

We would like to thank Professor Sun-Yung Alice Chang, Professor Matthew Gursky and Professor Jeffrey Case for their interest in this work and inspiring discussions. We would especially like to thank Professor Case for bringing the work [Gover and Ørsted 2013; Gursky 1997] to our attention and his wonderful comments.

Part of the work was done when Lin was in residence at the Mathematical Sciences Research Institute in Berkeley, supported by the NSF grant DMS-1440140 during Spring 2016. Also, part of the work was done when Yuan visited Institut Henri Poincaré in Paris during Fall 2015. We would like to express our deepest appreciations to both MSRI and IHP for their sponsorship and hospitality.

References

[Besse 1987] A. L. Besse, *Einstein manifolds*, Ergebnisse der Mathematik (3) **10**, Springer, Berlin, 1987. MR Zbl

[Branson 1985] T. P. Branson, "Differential operators canonically associated to a conformal structure", *Math. Scand.* **57**:2 (1985), 293–345. MR Zbl

[Case 2012] J. S. Case, "Some computations with the Q-curvature", preprint, 2012, Available at http://www.personal.psu.edu/jqc5026/notes/symmetric.pdf.

[Chang et al. 2002] S.-Y. A. Chang, M. Gursky, and P. C. Yang, "Remarks on a fourth order invariant in conformal geometry", pp. 373–372 in *Proceedings of International Conference on Aspects of Mathematics* (Hong Kong, 1996), edited by N. Mok, Hong Kong University, 2002.

[Cheng 2013] X. Cheng, "A generalization of almost-Schur lemma for closed Riemannian manifolds", *Ann. Global Anal. Geom.* **43**:2 (2013), 153–160. MR Zbl

[Chow et al. 2006] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Graduate Studies in Mathematics **77**, American Mathematical Society, Providence, RI, 2006. MR Zbl

[De Lellis and Topping 2012] C. De Lellis and P. M. Topping, "Almost-Schur lemma", *Calc. Var. Partial Differential Equations* **43**:3-4 (2012), 347–354. MR Zbl

[Fischer and Marsden 1975] A. E. Fischer and J. E. Marsden, "Deformations of the scalar curvature", *Duke Math. J.* **42**:3 (1975), 519–547. MR Zbl

[Ge and Wang 2012] Y. Ge and G. Wang, "An almost Schur theorem on 4-dimensional manifolds", *Proc. Amer. Math. Soc.* **140**:3 (2012), 1041–1044. MR Zbl

[Gover and Ørsted 2013] A. R. Gover and B. Ørsted, "Universal principles for Kazdan–Warner and Pohozaev–Schoen type identities", *Commun. Contemp. Math.* **15**:4 (2013), art. id. 1350002. MR Zbl

[Gursky 1997] M. J. Gursky, "Uniqueness of the functional determinant", *Comm. Math. Phys.* **189**:3 (1997), 655–665. MR Zbl

[Lin and Yuan 2016] Y.-J. Lin and W. Yuan, "Deformations of Q-curvature, I", *Calc. Var. Partial Differential Equations* **55**:4 (2016), art. id. 101. MR Zbl

[Paneitz 2008] S. M. Paneitz, "A quartic conformally covariant differential operator for arbitrary pseudo-Riemannian manifolds (summary)", *SIGMA Symm. Integ. Geom. Methods Appl.* **4** (2008), art. id. 036. MR Zbl

Received April 13, 2016.

YUEH-JU LIN
MSRI AND DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MI 48109
UNITED STATES
yuehjul@umich.edu

WEI YUAN
DEPARTMENT OF MATHEMATICS
SUN YAT-SEN UNIVERSITY
510275 GUANGZHOU
CHINA
gnr-x@163.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.439

GAUGE INVARIANTS FROM THE POWERS OF ANTIPODES

CRIS NEGRON AND SIU-HUNG NG

We prove that the trace of the n-th power of the antipode of a Hopf algebra with the Chevalley property is a gauge invariant, for each integer n. As a consequence, the order of the antipode, and its square, are invariant under Drinfeld twists. The invariance of the order of the antipode is closely related to a question of Shimizu on the pivotal covers of finite tensor categories, which we affirmatively answer for representation categories of Hopf algebras with the Chevalley property.

1. Introduction

This paper is dedicated to a study of the traces of the powers of the antipode of a Hopf algebra, and an approach to the Frobenius–Schur indicators of nonsemisimple Hopf algebras.

The antipode of a Hopf algebra has emerged as an object of importance in the study of Hopf algebras. It has been proved by Radford [1976] that the order of the antipode S of any finite-dimensional Hopf algebra H is finite. Moreover, the trace of S^2 is nonzero if, and only if, H is semisimple and cosemisimple [Larson and Radford 1988a]. If the base field $\Bbbk$ is of characteristic zero, $\mathrm{Tr}(S^2) = \dim H$ or 0, which characterizes respectively whether H is semisimple or nonsemisimple [Larson and Radford 1988b]. This means semisimplicity of H is characterized by the value of $\mathrm{Tr}(S^2)$. In particular, $\mathrm{Tr}(S^2)$ is an invariant of the finite tensor category H-mod. The invariance of $\mathrm{Tr}(S^2)$ and $\mathrm{Tr}(S)$ can also be obtained in any characteristic via Frobenius–Schur indicators.

A generalized notion of the n-th Frobenius–Schur (FS-)indicator $\nu_n^{\mathrm{KMN}}(H)$ has been introduced in [Kashina et al. 2012] for studying finite-dimensional Hopf algebras H, which are not necessarily semisimple or *pivotal*. However, $\nu_n^{\mathrm{KMN}}(H)$ coincides with the n-th FS-indicator of the regular representation of H when H is semisimple, defined in [Linchenko and Montgomery 2000]. These indicators are

Negron was supported by NSF Postdoctoral Fellowship DMS-1503147. Ng was partially supported by NSF grant DMS-1501179.
MSC2010: 16G99, 16T05, 18D20.
Keywords: tensor categories, Hopf algebras, gauge invariants, Frobenius–Schur indicators.

invariants of the finite tensor categories H-mod. In particular, $\nu_2^{\mathrm{KMN}}(H) = \mathrm{Tr}(S)$ and $\nu_0^{\mathrm{KMN}}(H) = \mathrm{Tr}(S^2)$ (see [Shimizu 2015a]) are invariants of H-mod.

The invariance of $\mathrm{Tr}(S)$ and $\mathrm{Tr}(S^2)$ alludes to the following question to be investigated in this paper:

Question 1.1. For any finite-dimensional Hopf algebra H with the antipode S, is the sequence $\{\mathrm{Tr}(S^n)\}_{n\in\mathbb{N}}$ an invariant of the finite tensor category H-mod?

For the purposes of this paper, we will always assume $\Bbbk$ to be an algebraically closed field of characteristic zero, and all Hopf algebras are finite-dimensional over $\Bbbk$.

Recall that a finite-dimensional Hopf algebra H has the Chevalley property if its Jacobson radical is a Hopf ideal. Equivalently, H has the Chevalley property if the full subcategory of sums of irreducible modules in H-mod forms a tensor subcategory. We provide a positive answer to Question 1.1 for Hopf algebras with the Chevalley property.

Theorem I (Theorem 4.3). Let H and K be finite-dimensional Hopf algebras over $\Bbbk$ with antipodes S_H and S_K respectively. Suppose H has the Chevalley property and that H-mod and K-mod are equivalent as tensor categories. Then we have

$$\mathrm{Tr}(S_H^n) = \mathrm{Tr}(S_K^n)$$

for all integers n.

In a categorial language, the theorem tells us that for any finite tensor category $\mathscr{C}$ with the Chevalley property which admits a fiber functor to the category of vector spaces, the "traces of the powers of the antipode" are well-defined invariants which are independent of the choice of fiber functor. One naturally asks whether these scalars can be expressed purely in terms of categorial data of $\mathscr{C}$.

Etingof asked the question whether, for any finite-dimensional H, $\mathrm{Tr}(S^{2m}) = 0$ provided $\mathrm{ord}(S^2) \nmid m$ [Radford and Schneider 2002, p. 186]. This question is affirmatively answered for pointed and dual pointed Hopf algebras in [Radford and Schneider 2002]. However, the odd powers of the antipode may have nonzero traces in general. We note that the above result covers both the even and odd powers of the antipode.

Theorem I also implies that the orders of the first two powers of the antipode of a Hopf algebra with the Chevalley property are also invariants.

Corollary I (Corollary 4.4). Let H and K be finite-dimensional Hopf algebras over $\Bbbk$ with antipodes S_H and S_K respectively. Suppose H has the Chevalley property and that H-mod and K-mod are equivalent as tensor categories. Then $\mathrm{ord}(S_H) = \mathrm{ord}(S_K)$ and hence $\mathrm{ord}(S_H^2) = \mathrm{ord}(S_K^2)$.

The order of S^2 is related to a known invariant called the *quasiexponent* $\mathrm{qexp}(H)$ [Etingof and Gelaki 2002]. Namely, for any finite-dimensional Hopf algebra,

$\mathrm{ord}(S^2)$ divides $\mathrm{qexp}(H)$. However, we still do not know whether or not the order of S^2 is an invariant in general.

The questions under consideration here are closely related to some recent investigations of Frobenius–Schur indicators for nonsemisimple Hopf algebras. The 2nd Frobenius–Schur indicator $\nu_2(V)$ of an irreducible complex representation of a finite group was introduced in [Frobenius and Schur 1906]; the notion was then extended to semisimple Hopf algebras, quasi-Hopf algebras, certain C^*-fusion categories and conformal field theory (see [Linchenko and Montgomery 2000; Mason and Ng 2005; Fuchs et al. 1999; Bantay 1997]). Higher Frobenius–Schur indicators $\nu_n(V)$ for semisimple Hopf algebra have been extensively studied in [Kashina et al. 2006]. In the most general context, FS-indicators can be defined for each object V in a *pivotal* tensor category $\mathscr{C}$, and they are invariants of these tensor categories [Ng and Schauenburg 2007b].

The n-th Frobenius–Schur indicators $\nu_n(H)$ of the regular representation of a semisimple Hopf algebra H, defined in [Linchenko and Montgomery 2000], in particular is an invariant of the fusion category H-mod (see [Ng and Schauenburg 2007b; 2008, Theorem 2.2]). For this special representation it is obtained in [Kashina et al. 2006] that

$$\nu_n(H) = \mathrm{Tr}(S \circ P_{n-1}), \tag{1-1}$$

where P_k denotes the k-th convolution power of the identity map id_H in $\mathrm{End}_{\Bbbk}(H)$. On elements, the map $S \circ P_{n-1}$ is given by $h \mapsto S(h_1 \dots h_{n-1})$.

The importance of the FS-indicators is illustrated in their applications to semisimple Hopf algebras and *spherical* fusion categories (see for examples [Bruillard et al. 2016; Dong et al. 2015; Kashina et al. 2006; Ng and Schauenburg 2007a; 2010; Ostrik 2015; Tucker 2015]). The arithmetic properties of the values of the FS-indicators have played an integral role in all these applications, and remains the main interest of FS-indicators (see for example [Guralnick and Montgomery 2009; Iovanov et al. 2014; Montgomery et al. 2016; Schauenburg 2016; Shimizu 2015a]).

It would be tempting to extend the notion of FS-indicators for the study of finite tensor categories or nonsemisimple Hopf algebras. One would expect that such a *generalized* indicator for a general Hopf algebra H should coincide with the existing one when H is semisimple.

The introduction of (what we refer to as) the KMN-indicators $\nu_n^{\mathrm{KMN}}(H)$ in [Kashina et al. 2012] is an attempt at this endeavor. Note that the right-hand side of (1-1), $\mathrm{Tr}(S \circ P_{n-1})$, is well defined for any finite-dimensional Hopf algebra over any base field, and we denote it as $\nu_n^{\mathrm{KMN}}(H)$. It has been shown in [Kashina et al. 2012] that the scalar $\nu_n^{\mathrm{KMN}}(H)$ is an invariant of the finite tensor category H-mod for each positive integer n. However, this definition of indicators for the regular representation in H-mod cannot be extended to other objects in H-mod.

Shimizu [2015b] lays out an alternative categorial approach to generalized indicators for a nonsemisimple Hopf algebra H. He first constructs a *universal pivotalization* $(H\text{-mod})^{\text{piv}}$ of H-mod, i.e., a pivotal tensor category with a fixed monoidal functor $\Pi : (H\text{-mod})^{\text{piv}} \to H\text{-mod}$ which is universal among all such categories. The pivotal category $(H\text{-mod})^{\text{piv}}$ has a *regular object* $\boldsymbol{R}_H$, and the scalar $\nu_n^{\text{KMN}}(H)$ can be recovered from a new version of the n-th indicator $\nu_n^{\text{Sh}}(\boldsymbol{R}_H^*)$. The universal pivotalization is natural in the sense that for any monoidal functor $\mathcal{F} : H\text{-mod} \to K\text{-mod}$, where K is a Hopf algebra, there exists a unique pivotal functor

$$\mathcal{F}^{\text{piv}} : (H\text{-mod})^{\text{piv}} \to (K\text{-mod})^{\text{piv}}$$

compatible with both Π and $\mathcal{F}$.

However, the invariance of $\nu_n^{\text{KMN}}(H)$ does not follow immediately from this categorical framework. Instead, it would be a consequence of a proposed isomorphism $\mathcal{F}^{\text{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$ associated to any monoidal equivalence $\mathcal{F} : H\text{-mod} \to K\text{-mod}$. While the latter condition remains open in general, we show below that the regular objects are preserved under monoidal equivalence for Hopf algebras with the Chevalley property.

Theorem II (Theorem 7.4). Let H and K be Hopf algebras with the Chevalley property and $\mathcal{F} : H\text{-mod} \to K\text{-mod}$ an equivalence of tensor categories. Then the induced pivotal equivalence $\mathcal{F}^{\text{piv}} : (H\text{-mod})^{\text{piv}} \to (K\text{-mod})^{\text{piv}}$ on the universal pivotalizations satisfies $\mathcal{F}^{\text{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$.

This gives a positive solution to Question 5.12 of [Shimizu 2015b]. From Theorem II we recover the gauge invariance result of [Kashina et al. 2012], in the specific case of Hopf algebras with the Chevalley property.

Corollary II [Kashina et al. 2012, Theorem 2.2]. Suppose H and K are Hopf algebras with the Chevalley property and have equivalent tensor categories of representations. Then $\nu_n^{\text{KMN}}(H) = \nu_n^{\text{KMN}}(K)$.

The paper is organized as follows: Section 2 recalls some basic notions and results on Hopf algebras and pivotal tensor categories. In Section 3, we prove that a specific element γ_F associated to a Drinfeld twist F of a semisimple Hopf algebra H is fixed by the antipode of H, using the pseudounitary structure of H-mod. We proceed to prove Theorem I and Corollary I in Section 4. In Section 5, we recall the construction of the universal pivotalization $(H\text{-mod})^{\text{piv}}$, the corresponding definition of n-th indicators for an object in $(H\text{-mod})^{\text{piv}}$ and their relations to $\nu_n^{\text{KMN}}(H)$. In Section 6, we introduce finite pivotalizations of H-mod and, in particular, the exponential pivotalization which contains all the possible pivotal categories defined on H-mod. In Section 7, we answer a question of Shimizu on the preservation of regular objects for Hopf algebras with the Chevalley property.

2. Preliminaries

Throughout this paper, we assume some basic definitions on Hopf algebras and monoidal categories. We denote the antipode of a Hopf algebra H by S_H or, when no confusion will arise, simply by S. A tensor category in this paper is a $\Bbbk$-linear abelian monoidal category with simple unit object $\mathbf{1}$. A monoidal functor between two tensor categories is a pair $(\mathcal{F}, \xi)$ in which $\mathcal{F}$ is a $\Bbbk$-linear functor satisfying $\mathcal{F}(\mathbf{1}) = \mathbf{1}$, and

$$\xi_{V,W} : \mathcal{F}(V) \otimes \mathcal{F}(W) \to \mathcal{F}(V \otimes W)$$

is the coherence isomorphism. If the context is clear, we may simply write $\mathcal{F}$ for the pair $(\mathcal{F}, \xi)$. The readers are referred to [Kassel 1995; Montgomery 1993] for the details.

Gauge equivalence, twists, and the antipode. Let H be a finite-dimensional Hopf algebra over $\Bbbk$ with antipode S, comultiplication Δ and counit ϵ. The category H-mod of finite-dimensional representations of H is a finite tensor category in the sense of [Etingof and Ostrik 2004]. For $V \in H$-mod, the dual vector space V' of V admits the natural right H-action $\leftharpoonup$ given by

$$(v^* \leftharpoonup h)(v) = v^*(hv)$$

for $h \in H$, $v^* \in V'$ and $v \in V$. The left dual V^* of V is the vector space V' endowed with the left H-action defined by

$$hv^* = v^* \leftharpoonup S(h)$$

for $h \in H$ and $v^* \in V'$, with the usual evaluation $\mathrm{ev} : V^* \otimes V \to \Bbbk$ and the dual basis map as the coevaluation $\mathrm{coev} : \Bbbk \to V \otimes V^*$. The right dual of V is defined similarly, with S replaced by S^{-1}.

Suppose K is another finite-dimensional Hopf algebra over $\Bbbk$ such that K-mod and H-mod are equivalent tensor categories. It follows from [Ng and Schauenburg 2008, Theorem 2.2] that there is a gauge transformation $F = \sum_i f_i \otimes g_i \in H \otimes H$ (see [Kassel 1995]), which is an invertible element satisfying

$$(\epsilon \otimes \mathrm{id})(F) = 1 = (\mathrm{id} \otimes \epsilon)(F),$$

such that the map $\Delta^F : H \to H \otimes H, h \mapsto F\Delta(h)F^{-1}$ together with the counit ϵ and the algebra structure of H form a bialgebra H^F and that $K \overset{\sigma}{\cong} H^F$ as bialgebras. In particular, H^F is a Hopf algebra with the antipode give by

$$S_F(h) = \beta_F S(h) \beta_F^{-1}, \tag{2-1}$$

where $\beta_F = \sum_i f_i S(g_i)$. Following the terminology of [Kassel 1995] (see [Kashina et al. 2012]), we say that K and H are *gauge equivalent* if the categories of their finite-dimensional representations are equivalent tensor categories. A quantity $f(H)$

obtained from a finite-dimensional Hopf algebra H is called a *gauge invariant* if $f(H) = f(K)$ for any Hopf algebra K gauge equivalent to H. For instance, $\mathrm{Tr}(S)$ and $\mathrm{Tr}(S^2)$ are gauge invariants of H.

If $F^{-1} = \sum_i d_i \otimes e_i$, then $\beta_F^{-1} = \sum_i S(d_i)e_i$. For the purpose of this paper, we set $\gamma_F = \beta_F S(\beta_F^{-1})$ and so, by (2-1), we have

$$S_F^2(h) = \gamma_F S^2(h)\gamma_F^{-1} \tag{2-2}$$

for $h \in H$.

Since the associativities of K and H are given by $1 \otimes 1 \otimes 1$, the gauge transformation F satisfies the condition

$$(1 \otimes F)(\mathrm{id} \otimes \Delta)(F) = (F \otimes 1)(\Delta \otimes \mathrm{id})(F). \tag{2-3}$$

This is a necessary and sufficient condition for Δ^F to be coassociative. A gauge transformation $F \in H \otimes H$ satisfying (2-3) is often called a *Drinfeld twist* or simply a *twist*.

Suppose $F \in H \otimes H$ is a twist and $K \overset{\sigma}{\cong} H^F$ as Hopf algebras. Following [Kassel 1995], one can define an equivalence $(\mathcal{F}_\sigma, \xi^F) : H\text{-mod} \to K\text{-mod}$ of tensor categories. For $V \in H\text{-mod}$, $\mathcal{F}_\sigma(V)$ is the left K-module with the action given by $k \cdot v := \sigma(k)v$ for $k \in K$ and $v \in V$. The assignment $V \mapsto \mathcal{F}_\sigma(V)$ defines a $\Bbbk$-linear equivalence from H-mod to K-mod with identity action on the morphisms. Together with the natural isomorphism

$$\xi^F : \mathcal{F}_\sigma(V) \otimes \mathcal{F}_\sigma(W) \to \mathcal{F}_\sigma(V \otimes W)$$

defined by the action of F^{-1} on $V \otimes W$, the pair $(\mathcal{F}_\sigma, \xi^F) : H\text{-mod} \to K\text{-mod}$ is an equivalence of tensor categories. If $K = H^F$ for some twist $F \in H \otimes H$, then $(\mathrm{Id}, \xi^F) : H\text{-mod} \to H^F\text{-mod}$ is an equivalence of tensor categories since $\mathcal{F}_{\mathrm{id}}$ is the identity functor Id.

Pivotal categories. For any finite tensor category $\mathscr{C}$ with the unit object $\mathbf{1}$, the left duality can define a functor $(-)^* : \mathscr{C} \to \mathscr{C}^{\mathrm{op}}$ and the double dual functor $(-)^{**} : \mathscr{C} \to \mathscr{C}$ is an equivalence of tensor categories. A pivotal structure of $\mathscr{C}$ is an isomorphism $j : \mathrm{Id} \to (-)^{**}$ of monoidal functors. Associated with a pivotal structure j are the notions of *trace* and *dimension*: For any $V \in \mathscr{C}$ and $f : V \to V$, one can define $\underline{\mathrm{ptr}}(f)$ as the scalar of the composition

$$\underline{\mathrm{ptr}}(f) := (\mathbf{1} \xrightarrow{\mathrm{coev}} V \otimes V^* \xrightarrow{f \otimes V^*} V \otimes V^* \xrightarrow{j \otimes V^*} V^{**} \otimes V^* \xrightarrow{\mathrm{ev}} \mathbf{1})$$

and $d(V) = \underline{\mathrm{ptr}}(\mathrm{id}_V)$. A finite tensor category with a specified pivotal structure is called a *pivotal category*.

Suppose $\mathscr{C}$ and $\mathscr{D}$ are pivotal categories with the pivotal structures j and j' respectively, and $(\mathcal{F}, \xi) : \mathscr{C} \to \mathscr{D}$ is a monoidal functor. Then there exists a unique natural isomorphism $\tilde{\xi} : \mathcal{F}(V^*) \to \mathcal{F}(V)^*$ which is determined by either of the following commutative diagrams (see [Ng and Schauenburg 2007b, p. 67]):

$$\begin{tikzcd} \mathcal{F}(V^*)\otimes\mathcal{F}(V) \arrow[r, "\tilde{\xi}\otimes\mathcal{F}(V)"] \arrow[d, "\xi"] & \mathcal{F}(V)^*\otimes\mathcal{F}(V) \arrow[d, "\mathrm{ev}"] \\ \mathcal{F}(V^*\otimes V) \arrow[r, "\mathcal{F}(\mathrm{ev})"] & \mathbf{1} \end{tikzcd} \quad \text{or} \quad \begin{tikzcd} \mathcal{F}(V)\otimes\mathcal{F}(V^*) \arrow[r, "\mathcal{F}(V)\otimes\tilde{\xi}"] & \mathcal{F}(V)\otimes\mathcal{F}(V)^* \\ \mathcal{F}(V\otimes V^*) \arrow[u, "\xi^{-1}"] & \mathbf{1} \arrow[l, "\mathcal{F}(\mathrm{coev})"] \arrow[u, "\mathrm{coev}"'] \end{tikzcd} \tag{2-4}$$

The monoidal functor $(\mathcal{F}, \xi)$ is said to be *pivotal* if it preserves the pivotal structures, which means the commutative diagram

$$\begin{tikzcd} \mathcal{F}(V) \arrow[r, "\mathcal{F}(j_V)"] \arrow[d, "j'_{\mathcal{F}(V)}"'] & \mathcal{F}(V^{**}) \arrow[d, "\tilde{\xi}"] \\ \mathcal{F}(V)^{**} \arrow[r, "\tilde{\xi}^*"] & \mathcal{F}(V^*)^* \end{tikzcd} \tag{2-5}$$

is satisfied for $V \in \mathscr{C}$. It follows from [Ng and Schauenburg 2007b, Lemma 6.1] that pivotal monoidal equivalence preserves dimensions. More precisely, if $\mathcal{F} : \mathscr{C} \to \mathscr{D}$ is an equivalence of pivotal categories, then $d(V) = d(\mathcal{F}(V))$ for $V \in \mathscr{C}$.

3. Semisimple Hopf algebras and pseudounitary fusion categories

In general, a finite tensor category may not have a pivotal structure. However, all the known semisimple finite tensor categories, also called *fusion categories*, over $\Bbbk$, admit a pivotal structure. It remains an open question whether every fusion category admits a pivotal structure (see [Etingof et al. 2005]). We present an equivalent definition of *pseudounitary* fusion categories obtained in [Etingof et al. 2005] or more generally in [Drinfeld et al. 2010] as in the following proposition.

Proposition 3.1 [Etingof et al. 2005]. *Let $\Bbbk_c$ denote the subfield of $\Bbbk$ generated by $\mathbb{Q}$ and all the roots of unity in $\Bbbk$. A fusion category $\mathscr{C}$ over $\Bbbk$ is called (ϕ-)pseudounitary if there exist a pivotal structure $j^{\mathscr{C}}$ and a field monomorphism $\phi : \Bbbk_c \to \mathbb{C}$ such that $\phi(d(V))$ is real and nonnegative for all simple $V \in \mathscr{C}$, where $d(V)$ is the dimension of V associated with $j^{\mathscr{C}}$. In this case, this pivotal structure $j^{\mathscr{C}}$ is unique and $\phi(d(V))$ is identical to the Frobenius–Perron dimension of V.*

The reference of ϕ becomes irrelevant when the dimensions associated with the pivotal structure $j^{\mathscr{C}}$ of $\mathscr{C}$ are nonnegative integers. In this case, $\mathscr{C}$ is simply said to be pseudounitary, and $j^{\mathscr{C}}$ is called the *canonical* pivotal structure of $\mathscr{C}$. In particular, the fusion category H-mod of a finite-dimensional semisimple quasi-Hopf algebra H is pseudounitary and the pivotal dimension of an H-module V associated with the canonical pivotal structure of H-mod is simply the ordinary dimension of V (see [Etingof et al. 2005]).

The canonical pivotal structure j^{Vec} on the *trivial* fusion category Vec of finite-dimensional $\Bbbk$-linear space is just the usual vector space isomorphism $V \to V^{**}$, which sends an element $v \in V$ to the evaluation function $\hat{v} : V^* \to \Bbbk,\ f \mapsto f(v)$.

Let H be a finite-dimensional semisimple Hopf algebra over $\Bbbk$. Then the antipode S of H satisfies $S^2 = \mathrm{id}$ (see [Larson and Radford 1988b]). Thus, for $V \in H$-mod, the natural isomorphism $j^{\mathrm{Vec}} : V \to V^{**}$ of vector spaces is an H-module map. In fact, j^{Vec} provides a pivotal structure of H-mod and the associated pivotal dimension $d(V)$ of V, given by the composition map

$$\Bbbk \xrightarrow{\mathrm{coev}} V \otimes V^* \xrightarrow{j \otimes V^*} V^{**} \otimes V^* \xrightarrow{\mathrm{ev}} \Bbbk,$$

is equal to its ordinary dimension dim V, which is a nonnegative integer. Therefore, j^{Vec} is the canonical pivotal structure of H-mod.

By [Ng and Schauenburg 2007b, Corollary 6.2], the canonical pivotal structure of a pseudounitary fusion category is preserved by any monoidal equivalence of fusion categories. For the purpose of this article, we restate this statement in the context of semisimple Hopf algebras.

Corollary 3.2 [Ng and Schauenburg 2007b, Corollary 6.2]. *Let H and K be finite-dimensional semisimple Hopf algebras over $\Bbbk$. If*

$$(\mathcal{F}, \xi) : H\text{-}mod \to K\text{-}mod$$

defines a monoidal equivalence, then $(\mathcal{F}, \xi)$ preserves their canonical pivotal structures, i.e., they satisfy the commutative diagram (2-5). *In particular, if $K \overset{\sigma}{\cong} H^F$ as Hopf algebras for some twist $F \in H \otimes H$, then the monoidal equivalence $(\mathcal{F}_\sigma, \xi^F) : H\text{-}mod \to K\text{-}mod$ preserves their canonical pivotal structures.*

Now, we can prove the following on a twist of a semisimple Hopf algebra:

Theorem 3.3. *Let H be a semisimple Hopf algebra over $\Bbbk$ with antipode S, $F = \sum_i f_i \otimes g_i \in H \otimes H$ a twist and $\beta_F = \sum_i f_i S(g_i)$. Then*

$$S(\beta_F) = \beta_F.$$

Proof. Let $F^{-1} = \sum_i d_i \otimes e_i$. Then $\beta^{-1} = \sum_i S(d_i) e_i$ (see Section 2), where β_F is simply abbreviated as β. For $V \in H$-mod, we denote by V^* and $V^\vee$ respectively the left duals of V in H-mod and H^F-mod. It follows from (2-4) that the duality transformation $\tilde{\xi}^F : V^* \to V^\vee$, for $V \in H$-mod, of the monoidal equivalence $(\mathrm{Id}, \xi^F) : H\text{-mod} \to H^F$-mod, is given by

$$\tilde{\xi}^F(v^*) = v^* \leftharpoonup \beta^{-1} \tag{3-1}$$

for all $v^* \in V^*$. Since both H and H^F are semisimple, their canonical pivotal structures are the same as the usual natural isomorphism j^{Vec} of finite-dimensional vector spaces over $\Bbbk$. Since (Id, ξ^F) preserves the canonical pivotal structures, by (2-5), we have

$$\begin{aligned}\widetilde{\xi}^F(j^{\mathrm{Vec}}(v))(v^*) &= (\widetilde{\xi}^F)^*(j^{\mathrm{Vec}}(v))(v^*)\\ &= j^{\mathrm{Vec}}(v)(\widetilde{\xi}^F(v^*)) = (v^* \leftharpoonup \beta^{-1})(v) = v^*(\beta^{-1}v),\end{aligned}$$

for all $v \in V$ and $v^* \in V^*$. Rewriting the first term of this equation, we find

$$v^*(S(\beta^{-1})v) = v^*(\beta^{-1}v).$$

This implies $\beta^{-1} = S(\beta^{-1})$ by taking $V = H$ and $v = 1$. $\square$

4. Hopf algebras with the Chevalley property

A finite-dimensional Hopf algebra H over $\Bbbk$ is said to have the *Chevalley property* if the Jacobson radical $J(H)$ of H is a Hopf ideal. In this case, $\overline{H} = H/J(H)$ is a semisimple Hopf algebra and the natural surjection $\pi : H \to \overline{H}$ is a Hopf algebra map. Let $F \in H \otimes H$ be a twist of H. Then

$$\overline{F} := (\pi \otimes \pi)(F) \in \overline{H} \otimes \overline{H}$$

is a twist and so

$$\pi(\beta_F) = \beta_{\overline{F}} = \overline{S}(\beta_{\overline{F}}) = \pi(S(\beta_F))$$

by Theorem 3.3, where $\overline{S}$ denotes the antipode of $\overline{H}$. Therefore, $S(\beta_F) \in \beta_F + J(H)$, and this proves the next result:

Lemma 4.1. *Let H be a finite-dimensional Hopf algebra over $\Bbbk$ with the Chevalley property. For any twist $F \in H \otimes H$,*

$$S(\beta_F) \in \beta_F + J(H).$$

We will need the following lemma.

Lemma 4.2. *Let A be a finite-dimensional algebra over $\Bbbk$ and T an algebra endomorphism or antiendomorphism of A.*

(i) *For any $x \in J(A)$ and $a \in A$,*

$$l(x)r(a)T \quad \text{and} \quad l(a)r(x)T$$

are nilpotent operators, where $l(x)$ and $r(x)$ respectively denote the left and the right multiplication by x.

(ii) *For any $a, a', b, b' \in A$ such that $a' \in a + J(A)$ and $b' \in b + J(A)$, we have* $\mathrm{Tr}(l(a)r(b)T) = \mathrm{Tr}(l(a')r(b')T)$.

Proof. (i) Let n be a positive integer such that $J(A)^n = 0$. We first consider the case when T is an algebra endomorphism of A. Then

$$\begin{aligned}(l(a)r(x)T)^n &= l(a)l(T(a))\cdots l(T^{n-1}(a))r(x)\cdots r(T^{n-1}(x))T^n\\ &= l(aT(a)\cdots T^{n-1}(a))r(T^{n-1}(x)\cdots T(x)x)T^n.\end{aligned}$$

Since $J(A)^n = 0$ and $x, T(x), \dots, T^{n-1}(x) \in J(A)$,

$$T^{n-1}(x) \cdots T(x)x = 0.$$

Therefore, $(l(a)r(x)T)^n = 0$. We can show that $(l(x)r(a)T)^n = 0$ by the same argument. In particular, they are nilpotent operators.

If T is an algebra antiendomorphism of A, then

$$(l(a)r(x)T)^2 = l(aT(x))r(T(a)x)T^2.$$

Since T^2 is an algebra endomorphism of A and $aT(x) \in J(A)$, we have that $(l(a)r(x)T)^{2n}$ is equal to 0. Similarly, $(l(x)r(a)T)^{2n} = 0$.

(ii) Let $a' = a + x$ and $b' = b + y$ for some $x, y \in J(A)$.

$$l(a')r(b')T = l(a)r(b)T + l(x)r(b')T + l(a)r(y)T.$$

By (i), $l(x)r(b')T$ and $l(a)r(y)T$ are nilpotent operators, and the result follows. □

We can now prove that the traces of the powers of the antipode of a Hopf algebra with the Chevalley property are gauge invariants.

Theorem 4.3. *Let H be a Hopf algebra over $\Bbbk$ with the antipode S. Suppose H has the Chevalley property. Then for any twist $F \in H \otimes H$, we have*

$$\mathrm{Tr}(S_F^n) = \mathrm{Tr}(S^n)$$

for all integers n, where S_F is the antipode of H^F. Moreover, if K is another Hopf algebra over $\Bbbk$ with antipode S' which is gauge equivalent to H, then

$$\mathrm{Tr}(S^n) = \mathrm{Tr}(S'^n)$$

for all integers n.

Proof. By (2-1), the antipode S_F of H^F is given by

$$S_F(h) = \beta_F S(h)\beta_F^{-1}$$

for $h \in H$. Recall from (2-2) that

$$S_F^2(h) = \gamma_F S^2(h)\gamma_F^{-1}$$

where $\gamma_F = \beta_F S(\beta_F^{-1})$. Then, for any nonnegative integer n, we can write $S_F^n = l(u_n)r(u_n^{-1})S^n$ where $u_0 = 1$ and

$$u_n = \begin{cases} \gamma_F S^2(\gamma_F) \cdots S^{n-2}(\gamma_F) & \text{if } n \text{ is positive and even,} \\ \beta_F S(u_{n-1}^{-1}) & \text{if } n \text{ is odd.} \end{cases}$$

Thus, if n is an even positive integer, $u_n \in 1 + J(H)$ by Lemma 4.1. It follows from Lemma 4.2 that

$$\mathrm{Tr}(S_F^n) = \mathrm{Tr}(l(u_n)r(u_n^{-1})S^n) = \mathrm{Tr}(l(1)r(1)S^n) = \mathrm{Tr}(S^n).$$

From now, we assume n is odd. Then $u_n \in \beta_F + J(H)$ and so we have

$$\begin{aligned}\operatorname{Tr}(S_F^n) &= \operatorname{Tr}(l(u_n)r(u_n^{-1})S^n) = \operatorname{Tr}(l(\beta_F)r(\beta_F^{-1})S^n) \\ &= \operatorname{Tr}(l(\beta_F)r(S^n(\beta_F^{-1}))S^n).\end{aligned} \tag{4-1}$$

The last equality of the above equation follows from Lemmas 4.1 and 4.2(ii).

Let Λ be a left integral of H and λ a right integral of H^* such that $\lambda(\Lambda) = 1$. By [Radford 1994, Theorem 2],

$$\operatorname{Tr}(T) = \lambda(S(\Lambda_2)T(\Lambda_1))$$

for any $\Bbbk$-linear endomorphism T on H, where $\Delta(\Lambda) = \Lambda_1 \otimes \Lambda_2$ is the Sweedler notation with the summation suppressed. Thus, by (4-1), we have

$$\begin{aligned}\operatorname{Tr}(S_F^n) &= \lambda(S(\Lambda_2)\beta_F S^n(\Lambda_1)S^n(\beta_F^{-1})) \\ &= \lambda(S(\Lambda_2)\beta_F S^n(\beta_F^{-1}\Lambda_1)).\end{aligned} \tag{4-2}$$

Recall from [Radford 1994, p. 591] that

$$\Lambda_1 \otimes a\Lambda_2 = S(a)\Lambda_1 \otimes \Lambda_2$$

for all $a \in H$. Using this equality and (4-2), we find

$$\begin{aligned}\operatorname{Tr}(S_F^n) &= \lambda(S(\Lambda_2)\beta_F S^n(\beta_F^{-1}\Lambda_1)) = \lambda(S(S^{-1}(\beta_F^{-1})\Lambda_2)\beta_F S^n(\Lambda_1)) \\ &= \lambda(S(\Lambda_2)\beta_F^{-1}\beta_F S^n(\Lambda_1)) = \lambda(S(\Lambda_2)S^n(\Lambda_1)) = \operatorname{Tr}(S^n).\end{aligned}$$

The second part of the theorem then follows immediately from Corollary 3.2. □

Corollary 4.4. *If H is a finite-dimensional Hopf algebra over $\Bbbk$ with the Chevalley property, then* $\operatorname{ord}(S)$ *is a gauge invariant. In particular,* $\operatorname{ord}(S^2)$ *is a gauge invariant.*

Proof. Since $\Bbbk$ is of characteristic zero, $\operatorname{Tr}(S^n) = \dim H$ if, and only if, $S^n = \mathrm{id}$. In particular, $\operatorname{ord}(S)$ is the smallest positive integer n such that $\operatorname{Tr}(S^n) = \dim H$. If K is a Hopf algebra (over $\Bbbk$) with the antipode S' and is gauge equivalent to H, then $\dim K = \dim H$ by Corollary 3.2. Hence, by Theorem 4.3, $\operatorname{ord}(S) = \operatorname{ord}(S')$. Note that S has odd order if, and only if, S is the identity. Therefore, the last statement follows. □

5. Pivotalization and indicators

KMN-indicators. For the regular representation H of a semisimple Hopf algebra H over $\Bbbk$ with the antipode S, the formula of the n-th Frobenius–Schur indicator $\nu_n(H)$ was obtained in [Kashina et al. 2006] and is given by (1-1). Since a monoidal equivalence between the module categories of two finite-dimensional Hopf algebras preserves their regular representation [Ng and Schauenburg 2008, Theorem 2.2] and Frobenius–Schur indicators are invariant under monoidal equivalences (see [Ng and

Schauenburg 2007b, Corollary 4.4] or [Ng and Schauenburg 2008, Proposition 3.2]), $\nu_n(H)$ is an invariant of $\mathrm{Rep}(H)$ if H is semisimple.

The formula (1-1) is well defined even for a nonsemisimple Hopf algebra H without any pivotal structure in H-mod. In fact, the gauge invariance of these scalars has been recently proved in [Kashina et al. 2012] which is stated as the following theorem.

Theorem 5.1 [Kashina et al. 2012, Theorem 2.2]. *For any finite-dimensional Hopf algebra H over any field $\Bbbk$, we define $\nu_n^{\mathrm{KMN}}(H)$ as in* (1-1). *If H and K are gauge equivalent finite-dimensional Hopf algebras over $\Bbbk$, then we have*

$$\nu_n^{\mathrm{KMN}}(H) = \nu_n^{\mathrm{KMN}}(K).$$

In general, these indicators $\nu_n^{\mathrm{KMN}}(H)$ can *only* be defined for the regular representation of H. The proof of Theorem 5.1 relies heavily on Corollary 3.2 and theory of Hopf algebras. We would like to have a categorial framework for the definition of $\nu_n^{\mathrm{KMN}}(H)$ in order to extend the definitions of the indicators to other objects in H-mod and give a categorial proof of gauge invariance of these indicators.

The universal pivotalization. In [Shimizu 2015b] the notion of universal pivotalization $\mathscr{C}^{\mathrm{piv}}$ of a finite tensor category $\mathscr{C}$ is proposed in order to produce indicators for pairs consisting of an object V in $\mathscr{C}$ along with a chosen isomorphism to its double dual. Under this categorical framework, $\nu_n^{\mathrm{KMN}}(H)$ is the n-th indicator of a special (or regular) object in $(H\text{-mod})^{\mathrm{piv}}$. We recall some constructions and results from [Shimizu 2015b] here.

For a finite tensor category $\mathscr{C}$ one can construct the universal pivotalization $\Pi_{\mathscr{C}} : \mathscr{C}^{\mathrm{piv}} \to \mathscr{C}$ of $\mathscr{C}$, which is referred to as the *pivotal cover* of $\mathscr{C}$ in [Shimizu 2015b].[1] The category $\mathscr{C}^{\mathrm{piv}}$ is the abelian, rigid, monoidal category of pairs (V, ϕ_V) of an object V and an isomorphism $\phi_V : V \to V^{**}$ in $\mathscr{C}$. Morphisms $(V, \phi_V) \to (W, \phi_W)$ in $\mathscr{C}^{\mathrm{piv}}$ are maps $f : V \to W$ in $\mathscr{C}$ which satisfy $\phi_W f = f^{**}\phi_V$. Note that the forgetful functor $\Pi_{\mathscr{C}} : \mathscr{C}^{\mathrm{piv}} \to \mathscr{C}$ is faithful.

The category $\mathscr{C}^{\mathrm{piv}}$ will be monoidal under the obvious tensor product

$$(V, \phi_V) \otimes (W, \phi_W) := (V \otimes W, \phi_V \otimes \phi_W)$$

(where we suppress the natural isomorphism $(V \otimes W)^{**} \cong V^{**} \otimes W^{**}$), and (left) rigid under the dual $(V, \phi_V)^* = (V^*, (\phi_V^{-1})^*)$. There is a natural pivotal structure $j : \mathrm{Id}_{\mathscr{C}^{\mathrm{piv}}} \to (-)^{**}$ on $\mathscr{C}^{\mathrm{piv}}$ which, on each object (V, ϕ_V), is simply given by $j_{(V,\phi_V)} := \phi_V$.

The construction $\mathscr{C}^{\mathrm{piv}}$ is universal in the sense that any monoidal functor $\mathcal{F} : \mathscr{D} \to \mathscr{C}$ from a pivotal tensor category $\mathscr{D}$ factors uniquely through $\mathscr{C}^{\mathrm{piv}}$. By

[1] We accept the term pivotal cover, but adopt the term pivotalization as it is consistent with the constructions of [Etingof et al. 2015] and admits adjectives more readily.

faithfulness of the forgetful functor $\Pi_{\mathscr{C}} : \mathscr{C}^{\text{piv}} \to \mathscr{C}$, the factorization $\widetilde{\mathcal{F}} : \mathscr{D} \to \mathscr{C}^{\text{piv}}$, which is a monoidal functor preserving the pivotal structures, is determined uniquely by where it sends objects. This factorization is described as follows.

Theorem 5.2 [Shimizu 2015b, Theorem 4.3]. *Let j denote the pivotal structure on $\mathscr{D}$ and $(\mathcal{F}, \xi) : \mathscr{D} \to \mathscr{C}$ a monoidal functor. Then the factorization $\widetilde{\mathcal{F}} : \mathscr{D} \to \mathscr{C}^{\text{piv}}$ sends each object V in $\mathscr{D}$ to the pair $(\mathcal{F}(V), (\widetilde{\xi}^*)^{-1}\widetilde{\xi}\mathcal{F}(j_V))$, where $\widetilde{\xi}$ is the duality transformation as in Section 2.*

From the universal property for $\mathscr{C}^{\text{piv}}$ one can conclude that the construction $(-)^{\text{piv}}$ is functorial, which means a monoidal functor $\mathcal{F} : \mathscr{D} \to \mathscr{C}$ induces a unique pivotal functor $\mathcal{F}^{\text{piv}} : \mathscr{D}^{\text{piv}} \to \mathscr{C}^{\text{piv}}$ which satisfies the commutative diagram

$$\begin{tikzcd} \mathscr{D}^{\text{piv}} \arrow[r, "\mathcal{F}^{\text{piv}}"] \arrow[d, "\Pi_{\mathscr{D}}"'] & \mathscr{C}^{\text{piv}} \arrow[d, "\Pi_{\mathscr{C}}"] \\ \mathscr{D} \arrow[r, "\mathcal{F}"] & \mathscr{C} \end{tikzcd}$$

of monoidal functors.

Indicators via $\mathscr{C}^{\text{piv}}$. Following [Ng and Schauenburg 2007b], for any $V, W \in \mathscr{C}$, we denote by $A_{V,W}$ and $D_{V,W}$ the natural isomorphisms $\text{Hom}_{\mathscr{C}}(\mathbf{1}, V \otimes W) \to \text{Hom}_{\mathscr{C}}(V^*, W)$ and $\text{Hom}_{\mathscr{C}}(V, W) \to \text{Hom}_{\mathscr{C}}(W^*, V^*)$ respectively. Thus,

$$T_{V,W} := A^{-1}_{W,V^{**}} \circ D_{V^*,W} \circ A_{V,W}$$

is a natural isomorphism from $\text{Hom}_{\mathscr{C}}(\mathbf{1}, V \otimes W) \to \text{Hom}_{\mathscr{C}}(\mathbf{1}, W \otimes V^{**})$. We also define $V^{\otimes 0} = \mathbf{1}$ and $V^{\otimes n} = V \otimes V^{\otimes(n-1)}$ for any positive integer n inductively.

Similar to the definition provided in [Ng and Schauenburg 2007b, p. 71], for any $\boldsymbol{V} = (V, \phi_V) \in \mathscr{C}^{\text{piv}}$ and positive integer n, one can define the map

$$E^{(n)}_{\boldsymbol{V}} : \text{Hom}_{\mathscr{C}}(\mathbf{1}, V^{\otimes n}) \to \text{Hom}_{\mathscr{C}}(\mathbf{1}, V^{\otimes n})$$

by

$$E^{(n)}_{\boldsymbol{V}}(f) := \Phi^{(n)} \circ (\text{id} \otimes \phi_V^{-1}) \circ T_{V,W}(f),$$

where $W = V^{\otimes(n-1)}$ and $\Phi^{(n)} : W \otimes V \to V \otimes W$ is the unique map obtained by the associativity isomorphisms. Shimizu's version of the n-th FS-indicator of $\boldsymbol{V}$ is defined as

$$\nu^{\text{Sh}}_n(\boldsymbol{V}) = \text{Tr}(E^{(n)}_{\boldsymbol{V}}).$$

This indicator is preserved by monoidal equivalence in the following sense:

Theorem 5.3 [Shimizu 2015b, Theorem 5.3]. *If $\mathcal{F} : \mathscr{C} \to \mathscr{D}$ is an equivalence of monoidal categories, for any $\boldsymbol{V} \in \mathscr{C}^{\text{piv}}$ and positive integer n, we have*

$$\nu^{\text{Sh}}_n(\boldsymbol{V}) = \nu^{\text{Sh}}_n(\mathcal{F}^{\text{piv}}(\boldsymbol{V})).$$

Remark 5.4. The definition of the n-th FS-indicator $\nu_n^{\mathrm{Sh}}(\boldsymbol{V})$ of $\boldsymbol{V}$ is different from the definition $\nu_n(\boldsymbol{V})$ introduced in [Ng and Schauenburg 2007b], in which $E_{\boldsymbol{V}}^{(n)}$ is defined on the space $\mathrm{Hom}_{\mathscr{C}^{\mathrm{piv}}}(\mathbf{1}, \boldsymbol{V}^{\otimes n})$ instead. It is natural to ask the question whether or how these two notions of indicators are related.

In the case of a finite-dimensional Hopf algebra $\mathscr{C} = H$-mod, we take $\boldsymbol{R}_H = (H, \phi_H)$ to be the object in $\mathscr{C}^{\mathrm{piv}}$, in which H is the left regular H-module and $\phi_H : H \to H^{**}$ is the composition $j^{\mathrm{Vec}} \circ S^2 : H \to \mathcal{F}_{S^2}(H) \cong H^{**}$. We call $\boldsymbol{R}_H$ the *regular object* in $\mathscr{C}^{\mathrm{piv}}$, and we have the following theorem:

Theorem 5.5 [Shimizu 2015b, Theorem 5.7]. *Suppose $\mathscr{C} = H$-mod. Then for each integer n we have $\nu_n^{\mathrm{Sh}}(\boldsymbol{R}_H^*) = \nu_n^{\mathrm{KMN}}(H)$.*

The theorem provides a convincing argument to pursue this categorical framework of FS-indicator for nonsemisimple Hopf algebras. However, this framework does not yield another proof for the gauge invariance of $\nu_n^{\mathrm{KMN}}(H)$ (see Theorem 5.1). The gauge invariance of $\nu_n^{\mathrm{KMN}}(H)$ will follow if this question, raised in [Shimizu 2015b], can be positively answered:

Question 5.6 [Shimizu 2015b]. Let H and K be two gauge equivalent Hopf algebras, and let $\mathcal{F} : H$-mod $\to K$-mod be a monoidal equivalence. Do we have $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$ in $(K\text{-mod})^{\mathrm{piv}}$?

If the question is affirmatively answered for gauge equivalent Hopf algebras H and K, then we have $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$ in $(K\text{-mod})^{\mathrm{piv}}$ for any monoidal equivalence $\mathcal{F} : H$-mod $\to K$-mod. Thus,

$$\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H^*) \cong (\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H))^* \cong \boldsymbol{R}_K^*.$$

It follows from [Shimizu 2015b, Theorem 5.3] that

$$\nu_n^{\mathrm{KMN}}(H) = \nu_n^{\mathrm{Sh}}(\boldsymbol{R}_H^*) = \nu_n^{\mathrm{Sh}}(\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H^*)) = \nu_n^{\mathrm{Sh}}(\boldsymbol{R}_K^*) = \nu_n^{\mathrm{KMN}}(K).$$

An affirmative answer to the question for semisimple H has been provided in [Shimizu 2015b, Proposition 5.10], and we will give in Theorem 7.4 a positive answer for H having the Chevalley property. As discussed above, an affirmative answer to the above question yields a categorial proof of Theorem 5.1.

6. Finite pivotalizations for Hopf algebras

Let $\mathscr{C} = H$-mod. In this section we remark that the universal pivotalization $\mathscr{C}^{\mathrm{piv}}$, which is not a finite tensor category in general, has a finite alternative for module categories of Hopf algebras.

For any $\Bbbk$-linear map $\tau : V \to V^{**}$ we let $\underline{\tau} \in \mathrm{Aut}_{\Bbbk}(V)$ denote the automorphism $\underline{\tau} := (j^{\mathrm{Vec}})^{-1} \circ \tau$.

Definition 6.1. For a Hopf algebra H we let H^{piv} denote the smash product $H \rtimes \mathbb{Z}$, where the generator x of $\mathbb{Z}$ acts on H by S^2. Similarly, for any positive integer N with $\operatorname{ord}(S^2)|N$, we take $H^{N\mathrm{piv}} = H \rtimes (\mathbb{Z}/N\mathbb{Z})$, where again the generator x of $\mathbb{Z}/N\mathbb{Z}$ acts as S^2.

The smash products H^{piv} and $H^{N\mathrm{piv}}$ admit a unique Hopf structure so that the inclusions $H \to H^{\mathrm{piv}}$ and $H \to H^{N\mathrm{piv}}$ are Hopf algebra maps and x is grouplike.

It has been pointed out in [Shimizu 2015b, Remark 4.5] that H^{piv}-mod is isomorphic to $(H\text{-mod})^{\mathrm{piv}}$ as pivotal tensor categories. To realize the identification $\Theta : H^{\mathrm{piv}}\text{-mod} \xrightarrow{\cong} \mathscr{C}^{\mathrm{piv}}$ one takes an H^{piv}-module V to the H-module V along with the isomorphism $\phi_V := j^{\mathrm{Vec}} \circ l(x) : V \to \mathcal{F}_{S^2}(V) \cong V^{**}$. On elements, $\phi_V(v) = j^{\mathrm{Vec}}(x \cdot v)$. So we see that the inverse functor $\Theta^{-1} : \mathscr{C}^{\mathrm{piv}} \to H^{\mathrm{piv}}$-mod takes the pair (V, ϕ_V) to the H-module V along with the action of the grouplike $x \in H^{\mathrm{piv}}$ by $x \cdot v = \underline{\phi_V}(v)$.

From the above description of $\mathscr{C}^{\mathrm{piv}}$ for Hopf algebras we see that $\mathscr{C}^{\mathrm{piv}}$ will not usually be a finite tensor category.

Note that, for any integer N as above, we have the Hopf projection $H^{\mathrm{piv}} \to H^{N\mathrm{piv}}$ which is the identity on H and sends x (in H^{piv}) to x (in $H^{N\mathrm{piv}}$). Dually, we get a fully faithful embedding of tensor categories $H^{N\mathrm{piv}}\text{-mod} \to H^{\mathrm{piv}}\text{-mod}$.

Definition 6.2. For any positive integer N which is divisible by the order of S^2, we let $\mathscr{C}^{N\mathrm{piv}}$ denote the full subcategory of $\mathscr{C}^{\mathrm{piv}}$ which is the image of

$$H^{N\mathrm{piv}}\text{-mod} \subset H^{\mathrm{piv}}\text{-mod}$$

along the isomorphism $\Theta : H^{\mathrm{piv}}\text{-mod} \to \mathscr{C}^{\mathrm{piv}}$.

From this point on if we write $H^{N\mathrm{piv}}$ or $\mathscr{C}^{N\mathrm{piv}}$ we are assuming that N is a positive integer with $\operatorname{ord}(S^2)|N$. We see, from the descriptions of the isomorphisms Θ and Θ^{-1} given above, that $\mathscr{C}^{N\mathrm{piv}}$ is the full subcategory consisting of all pairs (V, ϕ_V) so that the associated automorphism $\underline{\phi_V} \in \operatorname{Aut}_{\Bbbk}(V)$ has order dividing N.

Lemma 6.3. *The category $\mathscr{C}^{N\mathrm{piv}}$ is a pivotal finite tensor subcategory in the pivotal (nonfinite) tensor category $\mathscr{C}^{\mathrm{piv}}$ which contains $\boldsymbol{R}_H$.*

Proof. Since the map $\Theta : H^{\mathrm{piv}}\text{-mod} \to \mathscr{C}^{\mathrm{piv}}$ is a tensor equivalence, it follows that $\mathscr{C}^{N\mathrm{piv}}$, which is defined as the image of $H^{N\mathrm{piv}}$-mod in $\mathscr{C}^{\mathrm{piv}}$, is a full tensor subcategory in $\mathscr{C}^{\mathrm{piv}}$. The category $\mathscr{C}^{N\mathrm{piv}}$ is pivotal with its pivotal structure inherited from $\mathscr{C}^{\mathrm{piv}}$. The fact that $\boldsymbol{R}_H = (H, j^{\mathrm{Vec}} \circ S^2)$ is in $\mathscr{C}^{N\mathrm{piv}}$ just follows from the fact the order of $S^2 = \underline{\phi_{\boldsymbol{R}_H}}$ is assumed to divide N. □

Remark 6.4. There is another interesting object $\boldsymbol{A}_H$ introduced in [Shimizu 2015b, Section 6.1 and Theorem 7.1]. This object is the adjoint representation H_{ad} of H along with the isomorphism $\phi_{\boldsymbol{A}_H} = j^{\mathrm{Vec}} \circ S^2$. We will have that $\boldsymbol{A}_H$ is also in $\mathscr{C}^{N\mathrm{piv}}$ for any N.

Some choices for N which are of particular interest are $N = \mathrm{ord}(S^2)$ or $N = \mathrm{qexp}(H)$, where $\mathrm{qexp}(H)$ is the quasiexponent of H. Recall that the quasiexponent $\mathrm{qexp}(H)$ of H is defined as the unipotency index of the Drinfeld element u in the Drinfeld double $D(H)$ of H (see [Etingof and Gelaki 2002]). This number is always finite and divisible by the order of S^2 [Etingof and Gelaki 2002, Proposition 2.5]. More importantly, $\mathrm{qexp}(H)$ is a gauge invariant of H.

When we would like to pivotalize with respect to the quasiexponent we take $H^{E\mathrm{piv}} = H^{\mathrm{qexp}(H)\mathrm{piv}}$ and $\mathscr{C}^{E\mathrm{piv}} = \mathscr{C}^{\mathrm{qexp}(H)\mathrm{piv}}$. We call $\mathscr{C}^{E\mathrm{piv}}$ the *exponential pivotalization* of $\mathscr{C} = H$-mod.

If $\mathscr{C}$ admits any pivotal structures, one can show that the exponential pivotalization contains a copy of $(\mathscr{C}, j)$ for any choice of pivotal structure j on $\mathscr{C}$ as a full pivotal subcategory. More specifically, for any choice of pivotal structure j on $\mathscr{C}$ the induced map $(\mathscr{C}, j) \to \mathscr{C}^{\mathrm{piv}}$ will necessarily have image in $\mathscr{C}^{E\mathrm{piv}}$. In this way, the indicators for $\mathscr{C}$ calculated with respect to any choice of pivotal structure can be recovered from the (Shimizu-)indicators on $\mathscr{C}^{E\mathrm{piv}}$.

For some Hopf algebras H, the integer $\mathrm{qexp}(H)$ is minimal so that $\mathscr{C}^{N\mathrm{piv}}$ has this property. For example, when we take the generalized Taft algebra

$$H_{n,d}(\zeta) = k\langle g, x\rangle/(g^{nd} - 1, x^d, gx - \zeta xg),$$

where ζ is a *primitive* d-th root of unity (see [Taft 1971; Etingof and Walton 2016, Definition 3.1]). We have $\mathrm{ord}(S^2) = d$ and $nd = \mathrm{qexp}(H_{n,d}(\zeta))$ by [Etingof and Gelaki 2002, Theorem 4.6]. The grouplike element g provides a pivotal structure j on $H_{n,d}(\zeta)$-mod, and the resulting map into $(H_{n,d}(\zeta)\text{-mod})^{\mathrm{piv}}$ has image in $(H_{n,d}(\zeta)\text{-mod})^{N\mathrm{piv}}$ if, and only if, $\mathrm{qexp}(H_{n,d}(\zeta))|N$. This relationship can be seen as a consequence of the general fact that $\mathrm{qexp}(H) = \exp(G(H))$ for any pointed Hopf algebra H [Etingof and Gelaki 2002, Theorem 4.6].

Our functoriality result for the finite pivotalizations is the following.

Proposition 6.5. *For any monoidal equivalence $\mathcal{F} : H\text{-}mod \to K\text{-}mod$, where H and K are Hopf algebras, the functor $\mathcal{F}^{\mathrm{piv}}$ restricts to an equivalence*

$$\mathcal{F}^{E\mathrm{piv}} : (H\text{-}mod)^{E\mathrm{piv}} \to (K\text{-}mod)^{E\mathrm{piv}}.$$

Furthermore, when H has the Chevalley property $\mathcal{F}^{\mathrm{piv}}$ restricts to an equivalence $\mathcal{F}^{N\mathrm{piv}} : (H\text{-}mod)^{N\mathrm{piv}} \to (K\text{-}mod)^{N\mathrm{piv}}$ for each N (in particular $N = \mathrm{ord}(S_H^2) = \mathrm{ord}(S_K^2)$).

The proof of the proposition is given in the appendix.

7. Preservation of the regular object

In this section we show that for a monoidal equivalence $\mathcal{F} : H$-mod $\to K$-mod of Hopf algebras H and K with the Chevalley property we will have $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$. From this we recover Theorem 5.1 for Hopf algebras with the Chevalley property.

Let H be a finite-dimensional Hopf algebra with antipode S, and $F \in H \otimes H$ a twist of H. We let $\mathscr{C} = H$-mod, $\mathscr{C}_F = H^F$-mod, and let $F = (\mathcal{F}_{\mathrm{id}}, \xi^F)$ denote the associated equivalence from $\mathscr{C}$ to $\mathscr{C}_F$, by abuse of notation.

For this section we will be making copious use of the isomorphism $j^{\mathrm{Vec}} : V \to V^{**}$, and adopt the shorthand $\hat{v} = j^{\mathrm{Vec}}(v) \in V^{**}$ for $v \in V$. Recall that $\hat{v}$ is just the evaluation map $V^* \to \Bbbk$, $\eta \mapsto \eta(v)$.

Preservation of regular objects. Recall that the antipode S_F of H^F is given by $S_F(h) = \beta_F S(h) \beta_F^{-1}$ and that $\gamma_F = \beta_F S(\beta_F)^{-1}$. For any positive integer k, define

$$\gamma_F^{(k)} = \gamma_F S^2(\gamma_F) \cdots S^{2k-2}(\gamma_F).$$

Then we have $S_F^{2k}(h) = \gamma_F^{(k)} S^{2k}(h) (\gamma_F^{(k)})^{-1}$ for all positive integers k and $h \in H$. The following lemma is well known and it follows immediately from [Aljadeff et al. 2002, Equation (6)].

Lemma 7.1. *The element $\gamma_F^{(\mathrm{ord}(S^2))}$ is a grouplike element in H^F.*

Proof. Take $N = \mathrm{ord}(S^2)$. We have from [Aljadeff et al. 2002, Equation (6)] that

$$\Delta(\gamma_F) = F^{-1}(\gamma_F \otimes \gamma_F)(S^2 \otimes S^2)(F)$$

(see also [Majid 1995]). Hence

$$\Delta(\gamma_F^{(n)}) = F^{-1}(\gamma_F^{(n)} \otimes \gamma_F^{(n)})(S^{2n} \otimes S^{2n})(F)$$

for each n and therefore

$$\Delta_F(\gamma_F^{(N)}) = F \Delta(\gamma_F^{(N)}) F^{-1} = \gamma_F^{(N)} \otimes \gamma_F^{(N)}. \qquad \square$$

We have the following concrete description of the (universal) pivotalization of an equivalence $F : \mathscr{C} \to \mathscr{C}_F$ induced by a twist F on H.

Lemma 7.2. *The functor $F^{\mathrm{piv}} : \mathscr{C}^{\mathrm{piv}} \to \mathscr{C}_F^{\mathrm{piv}}$ sends an object (V, ϕ_V) in $\mathscr{C}^{\mathrm{piv}}$ to the pair consisting of the object V along with the isomorphism*

$$V \to V^{**}, \qquad v \mapsto j^{\mathrm{Vec}}(\gamma_F \underline{\phi_V}(v)).$$

In particular, $F^{\mathrm{piv}}(\boldsymbol{R}_H) = (H^F, j^{\mathrm{Vec}} \circ l(\gamma_F) \circ S^2)$.

Proof. Take $\beta = \beta_F$, $\gamma = \gamma_F$ and $\xi = \xi^F$. Recall that $F(V^*) = F(V)^* = V^*$ as vector spaces for each V in $\mathscr{C}$. It follows from (3-1) that, for any object V in $\mathscr{C}$,

$$\widetilde{\xi} : F(V^*) \to F(V)^*$$

is given by

$$\widetilde{\xi}(f) = f \leftharpoonup \beta^{-1} \ \text{ for } f \in V^*.$$

This implies

$$\widetilde{\xi}(\hat{v})(f) = (\hat{v} \leftharpoonup \beta^{-1})(f) = \hat{v}(\beta^{-1} \cdot f) = f(S(\beta^{-1})v) = j^{\mathrm{Vec}}(S(\beta^{-1})v)(f)$$

for $\hat{v} \in F(V^{**})$ and $f \in F(V^*)$. Thus,

$$(\widetilde{\xi}^*)^{-1}\widetilde{\xi}(\hat{v})(f) = (\widetilde{\xi}^*)^{-1} j^{\mathrm{Vec}}(S(\beta^{-1})v)(f) = j^{\mathrm{Vec}}(S(\beta^{-1})v)(\widetilde{\xi}^{-1}(f))$$
$$= j^{\mathrm{Vec}}(S(\beta^{-1})v)(f \leftharpoonup \beta) = f(\beta S(\beta^{-1})v) = f(\gamma v) = j^{\mathrm{Vec}}(\gamma v)(f)$$

for $\hat{v} \in F(V^{**})$ and $f \in F(V)^*$. By Theorem 5.2, $F^{\mathrm{piv}}(V, \phi_V) = (V, (\widetilde{\xi}^*)^{-1}\widetilde{\xi}\phi_V)$ and

$$(\widetilde{\xi}^*)^{-1}\widetilde{\xi}\phi_V(v) = (\widetilde{\xi}^*)^{-1}\widetilde{\xi} j^{\mathrm{Vec}}\underline{\phi_V}(v) = j^{\mathrm{Vec}}(\gamma\underline{\phi_V}(v))$$

for $v \in V$. The last statement follows immediately from the definition of $\boldsymbol{R}_H = (H, j^{\mathrm{Vec}} \circ S^2)$. This completes the proof. □

In the following proposition we let $\bar{S}^2$ denote the automorphism of $H/J(H)$ induced by S^2.

Proposition 7.3. *Let $F \in H \otimes H$ be a twist. The following statements are equivalent.*

(i) $F^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_{H^F}$ *in* $\mathscr{C}_F^{\mathrm{piv}}$.

(ii) *There is a unit t in H which satisfies the equation*

$$S^2(t)\gamma_F^{-1} - t = 0. \tag{7-1}$$

(iii) *There is a unit $\bar{t}$ in $H/J(H)$ which satisfies the equation*

$$\bar{S}^2(\bar{t})\bar{\gamma}_F^{-1} - \bar{t} = 0. \tag{7-2}$$

Proof. We take $N = \mathrm{ord}(S^2)$. By Lemma 7.2, $F^{\mathrm{piv}}(\boldsymbol{R}_H) = (H^F, j^{\mathrm{Vec}} \circ l(\gamma_F) \circ S^2)$. An isomorphism $F^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_{H^F}$ is determined by a H^F-module automorphism of H^F, which is necessarily given by right multiplication by a unit $t \in H^F$, producing a diagram

$$\begin{tikzcd}
H^F \arrow[r, "l(\gamma_F)S^2"] \arrow[d, "r(t)"'] & H^F \arrow[r, "j^{\mathrm{Vec}}"] \arrow[d, "r(t)"] & (H^F)^{**} \arrow[d, "r(t)^{**}"] \\
H^F \arrow[r, "S_F^2"'] & H^F \arrow[r, "j^{\mathrm{Vec}}"] & (H^F)^{**}
\end{tikzcd}$$

Equivalently, we are looking for a unit t such that

$$\gamma_F S^2(h)t = S_F^2(ht) = \gamma_F S^2(h) S^2(t)\gamma_F^{-1}$$

for all $h \in H$. This equation is equivalent to

$$S^2(t)\gamma_F^{-1} = t. \tag{7-3}$$

Let σ denote the $\Bbbk$-linear automorphism $r(\gamma_F^{-1}) \circ S^2 = r(\gamma_F)^{-1} \circ S^2$ of H^F, and let Σ be the subgroup generated by σ in $\mathrm{Aut}_{\Bbbk}(H^F)$. Then we have

$$\sigma^N = r(\gamma_F^{(N)})^{-1} \circ S^{2N} = r(\gamma_F^{(N)})^{-1}.$$

Since $\gamma_F^{(N)}$ is grouplike in H^F, it has a finite order. Therefore σ^N has finite order, as does σ, and Σ is a finite cyclic group.

Since $J(H)$ is a σ-invariant, the exact sequence

$$0 \to J(H) \to H \to H/J(H) \to 0$$

is in $\mathrm{Rep}(\Sigma)$. Applying the exact functor $(-)^\Sigma$, we get another exact sequence

$$0 \to J(H)^\Sigma \to H^\Sigma \to (H/J(H))^\Sigma \to 0. \tag{7-4}$$

Recall that an element in H is a unit if, and only if, its image in $H/J(H)$ is a unit. So from the exact sequence (7-4), we conclude that there is a unit in $(H/J(H))^\Sigma$ if and only if there is a unit in H^Σ. Rather, there exists a unit $\bar{t}$ solving the equation $\sigma \cdot X - X = 0$ in $H/J(H)$ if, and only if, there exists a unit t solving the equation in H. Since $\sigma \cdot \bar{t} = \bar{S}^2(\bar{t})\bar{\gamma}_F^{-1}$ and $\sigma \cdot t = S^2(t)\gamma_F^{-1}$, the equation $\bar{S}^2(X)\bar{\gamma}_F^{-1} - X = 0$ has a unit solution in $\overline{H}$ if, and only if, the equation $S^2(X)\gamma_F^{-1} - X = 0$ has a unit solution in H. □

As an immediate consequence of this proposition, we can prove preservation of regular objects for Hopf algebras with the Chevalley property.

Theorem 7.4. *Suppose H and K are gauge equivalent finite-dimensional Hopf algebras with the Chevalley property, and $\mathcal{F} : H$-mod $\to K$-mod is a monoidal equivalence. Then we have $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$ in $(K$-mod$)^{\mathrm{piv}}$.*

Proof. In view of [Ng and Schauenburg 2008, Theorem 2.2], it suffices to assume $K = H^F$ for some twist $F \in H \otimes H$, and that $\mathcal{F}$ is the associated equivalence

$$F : H\text{-mod} \to H^F\text{-mod}.$$

Let S be the antipode of H. It follows from Lemma 4.1 that $\bar{\gamma}_F = \bar{1}$ and $\bar{S}^2 = \mathrm{id}$. Therefore, every unit $t \in H/J(H)$ satisfies $\bar{S}^2(t)\bar{\gamma}_F^{-1} - t = 0$. The proof is then completed by Proposition 7.3. □

As a corollary we recover Theorem 5.1 for Hopf algebras with the Chevalley property.

Corollary 7.5 [Kashina et al. 2012, Theorem 2.2]. *If $\mathcal{F} : H$-mod $\to K$-mod is a gauge equivalence and H has the Chevalley property then we have*

$$\nu_n^{\mathrm{KMN}}(H) = \nu_n^{\mathrm{KMN}}(K)$$

for all $n \geq 0$.

Proof. We have $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_K$ by Theorem 7.4. Since a gauge equivalence preserves duals this implies $\mathcal{F}^{\mathrm{piv}}(\boldsymbol{R}_H^*) \cong \boldsymbol{R}_K^*$ as well. Hence, using [Shimizu 2015b, Theorems 5.3 and 5.7], we have

$$\nu_n^{\mathrm{KMN}}(H) = \nu_n^{\mathrm{Sh}}(\boldsymbol{R}_H^*) = \nu_n^{\mathrm{Sh}}(\boldsymbol{R}_K^*) = \nu_n^{\mathrm{KMN}}(K). \qquad \square$$

Appendix: Functoriality of finite pivotalizations

We adopt the notation introduced at the beginning of Section 6. Recall that the subcategory $\mathscr{C}^{N\mathrm{piv}} \subset \mathscr{C}^{\mathrm{piv}}$ is the full subcategory consisting of all pairs (V, ϕ_V) such that the associated automorphism $\underline{\phi_V} \in \mathrm{Aut}_{\Bbbk}(V)$ satisfies $\mathrm{ord}(\underline{\phi_V})|N$.

Lemma A.1. *Let $F \in H \otimes H$ be a twist and consider the functor $F : \mathscr{C} \to \mathscr{C}_F$. Then, for any N divisible by* $\mathrm{ord}(S^2)$, *the following statements are equivalent*:

(i) *F^{piv} restricts to an equivalence $F^{N\mathrm{piv}} : \mathscr{C}^{N\mathrm{piv}} \to \mathscr{C}_F^{N\mathrm{piv}}$.*

(ii) $\gamma_F^{(N)} = 1$.

Furthermore, the existence of an isomorphism $F^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_{H^F}$ implies (i) *and* (ii) *for all such N.*

Proof. Consider any (V, ϕ_V) in $\mathscr{C}^{N\mathrm{piv}}$. We have $F^{\mathrm{piv}}(V, \phi_V) = (V, j^{\mathrm{Vec}} \circ l(\gamma_F) \circ \underline{\phi_V})$, by Lemma 7.2. So $\underline{\phi_{F^{\mathrm{piv}}(V,\phi_V)}} = l(\gamma_F) \circ \underline{\phi_V}$. Since $\underline{\phi_V}$, considered as an H-module map, is a map from V to $\mathcal{F}_{S^2}(V)$, we find by induction that

$$(l(\gamma_F) \circ \underline{\phi_V})^n = l\big(\gamma_F^{(n)}\big) \circ \underline{\phi_V}^n$$

for each n. In particular,

$$(l(\gamma_F) \circ \underline{\phi_V})^N = l\big(\gamma_F^{(N)}\big) \tag{A-1}$$

since $\underline{\phi_V}^N = 1$.

From Equation (A-1) we see that $F^{\mathrm{piv}}(V, \phi_V)$ lies in $\mathscr{C}_F^{N\mathrm{piv}}$ if, and only if, $l(\gamma_F^{(N)}) = \mathrm{id}_V$, whence we have the implication (ii) $\Rightarrow$ (i). Applying (A-1) to the case $(V, \phi_V) = \boldsymbol{R}_H$ gives the converse implication (i) $\Rightarrow$ (ii) as well as the implication $F^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_{H^F} \Rightarrow$ (ii), since $\boldsymbol{R}_{H^F}$ is in each $\mathscr{C}_F^{N\mathrm{piv}}$. $\square$

We can now give the following proof:

Proof of Proposition 6.5. In view of [Ng and Schauenburg 2008, Theorem 2.2], it suffices to assume $K = H^F$ for some twist $F \in H \otimes H$ and consider the monoidal equivalence $F : H\text{-mod} \to H^F\text{-mod}$.

For Hopf algebras with the Chevalley property: Recall $\mathrm{ord}(S^2) = \mathrm{ord}(S_F^2)$ by Corollary 4.4. So we can pivotalize both H and H^F with respect to any N divisible by $\mathrm{ord}(S^2)$. We have already seen that $F^{\mathrm{piv}}(\boldsymbol{R}_H) \cong \boldsymbol{R}_{H^F}$. It follows, by Lemma A.1, that F^{piv} restricts to an equivalence $F^{N\mathrm{piv}} : \mathscr{C}^{N\mathrm{piv}} \to \mathscr{C}_F^{N\mathrm{piv}}$.

For the general case: From [Etingof and Gelaki 2002, Proposition 3.2] and the proof of [Etingof and Gelaki 2002, Proposition 3.3], $\gamma_F^{(\mathrm{qexp}(H))} = 1$. By Lemma A.1 it follows that F^{piv} restricts to an equivalence $F^{E\mathrm{piv}} : \mathscr{C}^{E\mathrm{piv}} \to \mathscr{C}_F^{E\mathrm{piv}}$. $\square$

Acknowledgements. The authors would like to thank the referee for the careful reading and suggestions on the previous version of this paper.

References

[Aljadeff et al. 2002] E. Aljadeff, P. Etingof, S. Gelaki, and D. Nikshych, "On twisting of finite-dimensional Hopf algebras", *J. Algebra* **256**:2 (2002), 484–501. MR Zbl

[Bantay 1997] P. Bantay, "The Frobenius–Schur indicator in conformal field theory", *Phys. Lett. B* **394**:1-2 (1997), 87–88. MR Zbl

[Bruillard et al. 2016] P. Bruillard, S.-H. Ng, E. C. Rowell, and Z. Wang, "Rank-finiteness for modular categories", *J. Amer. Math. Soc.* **29**:3 (2016), 857–881. MR Zbl

[Dong et al. 2015] C. Dong, X. Lin, and S.-H. Ng, "Congruence property in conformal field theory", *Algebra Number Theory* **9**:9 (2015), 2121–2166. MR Zbl

[Drinfeld et al. 2010] V. Drinfeld, S. Gelaki, D. Nikshych, and V. Ostrik, "On braided fusion categories, I", *Selecta Math.* (*N.S.*) **16**:1 (2010), 1–119. MR Zbl

[Etingof and Gelaki 2002] P. Etingof and S. Gelaki, "On the quasi-exponent of finite-dimensional Hopf algebras", *Math. Res. Lett.* **9**:2-3 (2002), 277–287. MR Zbl

[Etingof and Ostrik 2004] P. Etingof and V. Ostrik, "Finite tensor categories", *Mosc. Math. J.* **4**:3 (2004), 627–654. MR Zbl

[Etingof and Walton 2016] P. Etingof and C. Walton, "Pointed Hopf actions on fields, II", *J. Algebra* **460** (2016), 253–283. MR Zbl

[Etingof et al. 2005] P. Etingof, D. Nikshych, and V. Ostrik, "On fusion categories", *Ann. of Math.* (2) **162**:2 (2005), 581–642. MR Zbl

[Etingof et al. 2015] P. Etingof, S. Gelaki, D. Nikshych, and V. Ostrik, *Tensor categories*, Mathematical Surveys and Monographs **205**, American Mathematical Society, Providence, RI, 2015. MR Zbl

[Frobenius and Schur 1906] G. Frobenius and I. Schur, "Über die reellen Darstellungen der endlichen Gruppen", *Berl. Ber.* **1906** (1906), 186–208. JFM

[Fuchs et al. 1999] J. Fuchs, A. C. Ganchev, K. Szlachányi, and P. Vecsernyés, "S_4 symmetry of $6j$ symbols and Frobenius–Schur indicators in rigid monoidal C^* categories", *J. Math. Phys.* **40**:1 (1999), 408–426. MR Zbl

[Guralnick and Montgomery 2009] R. Guralnick and S. Montgomery, "Frobenius–Schur indicators for subgroups and the Drinfeld double of Weyl groups", *Trans. Amer. Math. Soc.* **361**:7 (2009), 3611–3632. MR Zbl

[Iovanov et al. 2014] M. Iovanov, G. Mason, and S. Montgomery, "FSZ-groups and Frobenius–Schur indicators of quantum doubles", *Math. Res. Lett.* **21**:4 (2014), 757–779. MR Zbl

[Kashina et al. 2006] Y. Kashina, Y. Sommerhäuser, and Y. Zhu, *On higher Frobenius–Schur indicators*, Mem. Amer. Math. Soc. **855**, 2006. MR Zbl

[Kashina et al. 2012] Y. Kashina, S. Montgomery, and S.-H. Ng, "On the trace of the antipode and higher indicators", *Israel J. Math.* **188** (2012), 57–89. MR Zbl

[Kassel 1995] C. Kassel, *Quantum groups*, Graduate Texts in Mathematics **155**, Springer, New York, 1995. MR Zbl

[Larson and Radford 1988a] R. G. Larson and D. E. Radford, "Semisimple cosemisimple Hopf algebras", *Amer. J. Math.* **110**:1 (1988), 187–195. MR Zbl

[Larson and Radford 1988b] R. G. Larson and D. E. Radford, "Finite-dimensional cosemisimple Hopf algebras in characteristic 0 are semisimple", *J. Algebra* **117**:2 (1988), 267–289. MR Zbl

[Linchenko and Montgomery 2000] V. Linchenko and S. Montgomery, "A Frobenius–Schur theorem for Hopf algebras", *Algebr. Represent. Theory* **3**:4 (2000), 347–355. MR Zbl

[Majid 1995] S. Majid, *Foundations of quantum group theory*, Cambridge Univ. Press, 1995. MR Zbl

[Mason and Ng 2005] G. Mason and S.-H. Ng, "Central invariants and Frobenius–Schur indicators for semisimple quasi-Hopf algebras", *Adv. Math.* **190**:1 (2005), 161–195. MR Zbl

[Montgomery 1993] S. Montgomery, *Hopf algebras and their actions on rings*, CBMS Regional Conference Series in Mathematics **82**, American Mathematical Society, Providence, RI, 1993. MR Zbl

[Montgomery et al. 2016] S. Montgomery, M. D. Vega, and S. Witherspoon, "Hopf automorphisms and twisted extensions", *J. Algebra Appl.* **15**:6 (2016), art. id. 1650103. MR Zbl

[Ng and Schauenburg 2007a] S.-H. Ng and P. Schauenburg, "Frobenius–Schur indicators and exponents of spherical categories", *Adv. Math.* **211**:1 (2007), 34–71. MR Zbl

[Ng and Schauenburg 2007b] S.-H. Ng and P. Schauenburg, "Higher Frobenius–Schur indicators for pivotal categories", pp. 63–90 in *Hopf algebras and generalizations*, edited by L. H. Kauffman et al., Contemp. Math. **441**, American Mathematical Society, Providence, RI, 2007. MR Zbl

[Ng and Schauenburg 2008] S.-H. Ng and P. Schauenburg, "Central invariants and higher indicators for semisimple quasi-Hopf algebras", *Trans. Amer. Math. Soc.* **360**:4 (2008), 1839–1860. MR Zbl

[Ng and Schauenburg 2010] S.-H. Ng and P. Schauenburg, "Congruence subgroups and generalized Frobenius–Schur indicators", *Comm. Math. Phys.* **300**:1 (2010), 1–46. MR Zbl

[Ostrik 2015] V. Ostrik, "Pivotal fusion categories of rank 3", *Mosc. Math. J.* **15**:2 (2015), 373–396. MR Zbl

[Radford 1976] D. E. Radford, "The order of the antipode of a finite dimensional Hopf algebra is finite", *Amer. J. Math.* **98**:2 (1976), 333–355. MR Zbl

[Radford 1994] D. E. Radford, "The trace function and Hopf algebras", *J. Algebra* **163**:3 (1994), 583–622. MR Zbl

[Radford and Schneider 2002] D. E. Radford and H.-J. Schneider, "On the even powers of the antipode of a finite-dimensional Hopf algebra", *J. Algebra* **251**:1 (2002), 185–212. MR Zbl

[Schauenburg 2016] P. Schauenburg, "Frobenius–Schur indicators for some fusion categories associated to symmetric and alternating groups", *Algebr. Represent. Theory* **19**:3 (2016), 645–656. MR Zbl

[Shimizu 2015a] K. Shimizu, "On indicators of Hopf algebras", *Israel J. Math.* **207**:1 (2015), 155–201. MR Zbl

[Shimizu 2015b] K. Shimizu, "The pivotal cover and Frobenius–Schur indicators", *J. Algebra* **428** (2015), 357–402. MR Zbl

[Taft 1971] E. J. Taft, "The order of the antipode of finite-dimensional Hopf algebra", *Proc. Nat. Acad. Sci. U.S.A.* **68** (1971), 2631–2633. MR Zbl

[Tucker 2015] H. Tucker, "Frobenius–Schur indicators for near-group and Haagerup–Izumi fusion categories", preprint, 2015. arXiv

Received October 17, 2016. Revised April 11, 2017.

CRIS NEGRON
DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MA 02139
UNITED STATES
negronc@mit.edu

SIU-HUNG NG
DEPARTMENT OF MATHEMATICS
LOUISIANA STATE UNIVERSITY
BATON ROUGE, LA 70803
UNITED STATES
rng@math.lsu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.461

BRANCHING LAWS FOR THE METAPLECTIC COVER OF GL_2

SHIV PRAKASH PATEL

Let F be a nonarchimedean local field of characteristic zero and E/F be a quadratic extension. The aim of this article is to study the multiplicity of an irreducible admissible representation of $\mathrm{GL}_2(F)$ occurring in an irreducible admissible genuine representation of the nontrivial two-fold covering $\widetilde{\mathrm{GL}}_2(E)$ of $\mathrm{GL}_2(E)$.

1. Introduction

Let F be a nonarchimedean local field of characteristic zero and let E be a quadratic extension of F. The branching laws for restriction of representations of $\mathrm{SO}_{n+1}(F)$ to $\mathrm{SO}_n(F)$ were formulated as conjectures by B. Gross and D. Prasad [1992], and these are widely known as Gross–Prasad conjectures although they have been completely proved by Mœglin and Waldspurger [2012]. The first case of these conjectures is for the restriction of representations of $\mathrm{GL}_2(F)$ to its maximal tori, which was considered by J. B. Tunnell [1983] and H. Saito [1993]. A metaplectic analog of this result was recently considered by the author in a joint work with Prasad, where the restriction of representations of metaplectic $\mathrm{GL}_2(F)$ to inverse images of the maximal tori was studied [Patel and Prasad 2017]. The results of Tunnell and Saito have, in particular, a multiplicity one result which is then refined in terms of certain ϵ-factors. The metaplectic case of this restriction loses the multiplicity one property, but still one has finite multiplicities which are bounded by some explicit constants. The next case of Gross–Prasad conjectures can be considered to be the restriction of representations of $\mathrm{GL}_2(E)$ to $\mathrm{GL}_2(F)$ which was studied by Prasad [1992]. These cases played an important role in the formulation of Gross–Prasad conjectures. Our aim in this paper is to study an analogous restriction of representations of metaplectic $\mathrm{GL}_2(E)$ to $\mathrm{GL}_2(F)$.

The problem of decomposing a representation of $\mathrm{GL}_2(E)$ restricted to $\mathrm{GL}_2(F)$ was considered and solved by Prasad [1992], proving a multiplicity one theorem, and giving an explicit classification of representations π_1 of $\mathrm{GL}_2(E)$ and π_2 of $\mathrm{GL}_2(F)$ such that there exists a nonzero $\mathrm{GL}_2(F)$ invariant linear form:

$$l : \pi_1 \otimes \pi_2 \to \mathbb{C}.$$

MSC2010: primary 22E35; secondary 22E50.
Keywords: metaplectic group, branching laws.

This problem is closely related to a similar branching law from $\mathrm{GL}_2(E)$ to $D_F^\times$, where D_F is the unique quaternion division algebra which is central over F, and $D_F^\times \hookrightarrow \mathrm{GL}_2(E)$. We recall that the embedding $D_F^\times \hookrightarrow \mathrm{GL}_2(E)$ is given by fixing an isomorphism $D_F \otimes E \cong M_2(E)$, by the Skolem–Noether theorem, which is unique up to conjugation by elements of $\mathrm{GL}_2(E)$. Henceforth, we fix one such embedding of $D_F^\times$ inside $\mathrm{GL}_2(E)$. The restriction problems for the pair $(\mathrm{GL}_2(E), \mathrm{GL}_2(F))$ and $(\mathrm{GL}_2(E), D_F^\times)$ are related by a certain dichotomy. More precisely, the following result was proved in [Prasad 1992]:

Theorem 1.1 (Prasad). *Let π_1 and π_2 be irreducible admissible infinite-dimensional representations of $\mathrm{GL}_2(E)$ and $\mathrm{GL}_2(F)$, respectively, such that the central character of π_1 restricted to the center of $\mathrm{GL}_2(F)$ is the same as the central character of π_2. Then*:

(1) *For a principal series representation π_2 of $\mathrm{GL}_2(F)$, we have*

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) = 1.$$

(2) *For a discrete series representation π_2 of $\mathrm{GL}_2(F)$, letting π_2' be the finite-dimensional representation of $D_F^\times$ associated to π_2 by the Jacquet–Langlands correspondence, we have*

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) + \dim \mathrm{Hom}_{D_F^\times}(\pi_1, \pi_2') = 1.$$

In this paper, we study the analogous problem in the metaplectic setting. More precisely, instead of considering $\mathrm{GL}_2(E)$ we will consider the group $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$ which is a topological central extension of $\mathrm{GL}_2(E)$ by $\mathbb{C}^\times$, which is obtained from the two-fold topological central extension $\widetilde{\mathrm{GL}}_2(E)$ described below. We recall that there is unique (up to isomorphism) two-fold cover of $\mathrm{SL}_2(E)$ called the metaplectic cover and denoted by $\widetilde{\mathrm{SL}}_2(E)$ in this paper, but there are many inequivalent two-fold coverings of $\mathrm{GL}_2(E)$ which extend this two-fold covering of $\mathrm{SL}_2(E)$. We fix a covering of $\mathrm{GL}_2(E)$ as follows. Observe that $\mathrm{GL}_2(E)$ is a semidirect product of $\mathrm{SL}_2(E)$ and $E^\times$, where $E^\times$ sits inside $\mathrm{GL}_2(E)$ by $e \mapsto \left(\begin{smallmatrix} e & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The action of $E^\times$ on $\mathrm{SL}_2(E)$ lifts to an action on $\widetilde{\mathrm{SL}}_2(E)$. Denote $\widetilde{\mathrm{GL}}_2(E)$ to be $\widetilde{\mathrm{SL}}_2(E) \rtimes E^\times$ which we call “the” metaplectic cover of $\mathrm{GL}_2(E)$. This cover can be described by an explicit 2-cocycle on $\mathrm{GL}_2(E)$ with values in $\{\pm 1\}$, see [Kubota 1969]. The group $\widetilde{\mathrm{GL}}_2(E)$ is a topological central extension of $\mathrm{GL}_2(E)$ by $\mu_2 := \{\pm 1\}$, i.e., we have an exact sequence of topological groups:

$$1 \to \mu_2 \to \widetilde{\mathrm{GL}}_2(E) \to \mathrm{GL}_2(E) \to 1.$$

The group $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} := \widetilde{\mathrm{GL}}_2(E) \times_{\mu_2} \mathbb{C}^\times$ is called the $\mathbb{C}^\times$-cover of $\mathrm{GL}_2(E)$ obtained from the two-fold cover $\widetilde{\mathrm{GL}}_2(E)$, and is a topological central extension of $\mathrm{GL}_2(E)$

by $\mathbb{C}^\times$, i.e., we have an exact sequence of topological groups:

$$1 \to \mathbb{C}^\times \to \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \to \mathrm{GL}_2(E) \to 1.$$

Now we recall the following result regarding splitting of this cover when restricted to certain subgroups. This makes it possible to consider an analog of the Prasad's restriction problem in the metaplectic case.

Theorem 1.2 [Patel 2016]. *Let E be a quadratic extension of a nonarchimedean local field and $\widetilde{\mathrm{GL}}_2(E)$ be the two-fold metaplectic covering of $\mathrm{GL}_2(E)$. Then*:

(1) *The two-fold metaplectic covering $\widetilde{\mathrm{GL}}_2(E)$ splits over the subgroup $\mathrm{GL}_2(F)$.*

(2) *The $\mathbb{C}^\times$-covering obtained from $\widetilde{\mathrm{GL}}_2(E)$ splits over the subgroup $D_F^\times$.*

Note that the splittings over $\mathrm{GL}_2(F)$ and $D_F^\times$ in Theorem 1.2 are not unique. As there is more than one splitting in each case, to study the problem of decomposing a representation of $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$ restricted to $\mathrm{GL}_2(F)$ and $D_F^\times$, we must fix one splitting of each of the subgroups $\mathrm{GL}_2(F)$ and $D_F^\times$, which are related to each other. We make the following working hypothesis, which has been formulated by Prasad.

Working Hypothesis 1.3. Let L be a quadratic extension of F. Write R for the restriction of scalars torus $R_{L/F}\mathbb{G}_m$. Thus $R(F) = L^\times$. Fix embeddings of R into GL_2 and $D_F^\times$ (viewed as algebraic groups over F). The sets of splittings

$$\begin{tikzcd} & \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \arrow[d] \\ \mathrm{GL}_2(F) \arrow[ur, hook, dashed, "s"] \arrow[r] & \mathrm{GL}_2(E) \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} & \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \arrow[d] \\ D_F^\times \arrow[ur, hook, dashed, "s'"] \arrow[r] & \mathrm{GL}_2(E) \end{tikzcd}$$

are principal homogeneous spaces over the group $\mathrm{Hom}(F^\times, \mathbb{C}^\times)$. More explicitly, two splittings s_1, s_2 of $\mathrm{GL}_2(F)$ will be related by

$$s_2(g) = \chi(\det g) \cdot s_1(g)$$

for some character $\chi \in \mathrm{Hom}(F^\times, \mathbb{C}^\times)$ (for $D_F^\times$ the det should be replaced by Nm the reduced norm map). A pair (s, s') of splittings, where

$$s : \mathrm{GL}_2(F) \to \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \quad \text{and} \quad s' : D_F^\times \to \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times},$$

is called a pair of "compatible splittings" if for any quadratic extension L/F with the fixed embedding of R into GL_2 and $D_F^\times$ the restriction of s and s' to $L^\times$ as in the following diagrams

$$\begin{tikzcd} & & \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \arrow[d] \\ L^\times \arrow[r] \arrow[urr, hook, dashed] & \mathrm{GL}_2(F) \arrow[r] \arrow[ur, hook, dashed, "s"'] & \mathrm{GL}_2(E) \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} & & \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times} \arrow[d] \\ L^\times \arrow[r] \arrow[urr, hook, dashed] & D_F^\times \arrow[r] \arrow[ur, hook, dashed, "s'"'] & \mathrm{GL}_2(E) \end{tikzcd}$$

are conjugate in $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$, i.e., there is an element $g \in \widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$ such that $s(L^\times) = g \cdot s'(L^\times) \cdot g^{-1}$. Then we assume that

$$\text{there exists a pair } (s, s') \text{ of compatible splittings.}$$

If (s, s') is a pair of compatible splittings and χ is a character of $F^\times$ then the pair of splittings $(\chi(\det(\bullet)s, \chi(\mathrm{Nm}(\bullet)s')$ is also compatible. Thus, given a single pair (s, s') of compatible splittings, we have a $\mathrm{Hom}(F^\times, \mathbb{C}^\times)$-equivariant bijection between the sets of splittings, in such a way that all pairs matched by the bijection are compatible.

Definition 1.4. A representation of $\widetilde{\mathrm{GL}}_2(E)$ (respectively, $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$) is called genuine if μ_2 acts nontrivially (respectively, $\mathbb{C}^\times$ acts by identity).

In particular, a genuine representation does not factor through $\mathrm{GL}_2(E)$. In what follows, we always consider genuine representations of the metaplectic group $\widetilde{\mathrm{GL}}_2(E)$. Let $B(E)$, $A(E)$ and $N(E)$ be the Borel subgroup, maximal torus and maximal unipotent subgroup of $\mathrm{GL}_2(E)$ consisting of all upper triangular matrices, diagonal matrices and upper triangular unipotent matrices respectively. Let $B(F)$, $A(F)$ and $N(F)$ denote the corresponding subgroups of $\mathrm{GL}_2(F)$. Let Z be the center of $\mathrm{GL}_2(E)$ and $\widetilde{Z}$ the inverse image of Z in $\widetilde{\mathrm{GL}}_2(E)$. Note that $\widetilde{Z}$ is an abelian subgroup of $\widetilde{\mathrm{GL}}_2(E)$ but is not the center of $\widetilde{\mathrm{GL}}_2(E)$; the center of $\widetilde{\mathrm{GL}}_2(E)$ is $\widetilde{Z}^2$, the inverse image of $Z^2 := \{z^2 \mid z \in Z\}$.

Let ψ be a nontrivial additive character of E. Note that the metaplectic covering splits when restricted to the subgroup $N(E)$ and hence ψ gives a character of $N(E)$. Let π be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and $\pi_{N(E),\psi}$, the ψ-twisted Jacquet module which is a $\widetilde{Z}$-module. Let ω_π be the central character of π. A character of $\widetilde{Z}$ appearing in $\pi_{N(E),\psi}$ agrees with ω_π when restricted to $\widetilde{Z}^2$. Let $\Omega(\omega_\pi)$ be the set of genuine characters of $\widetilde{Z}$ whose restriction to $\widetilde{Z}^2$ agrees with ω_π. We also realize $\Omega(\omega_\pi)$ as a $\widetilde{Z}$-module, i.e., as direct sum of characters in $\Omega(\omega_\pi)$ with multiplicity one. From [Gelbart et al. 1979, Theorem 4.1], one knows that the multiplicity of a character $\mu \in \Omega(\omega_\pi)$ in the $\widetilde{Z}$-module $\pi_{N(E),\psi}$ is at most one. Hence $\pi_{N(E),\psi}$ is a $\widetilde{Z}$-submodule of $\Omega(\omega_\pi)$. Now we state the main result of this paper.

We abuse notation and write $\widetilde{\mathrm{GL}}_2(E)$ for $\widetilde{\mathrm{GL}}_2(E)_{\mathbb{C}^\times}$.

Theorem 1.5. *Let π_1 be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and let π_2 be an infinite-dimensional irreducible admissible representation of $\mathrm{GL}_2(F)$. Assume that the central characters ω_{π_1} of π_1 and ω_{π_2} of π_2 agree on $E^{\times 2} \cap F^\times$. Fix a nontrivial additive character ψ of E such that $\psi|_F = 1$. Let $Q = (\pi_1)_{N(E)}$ be the Jacquet module of π_1. Assume that Working Hypothesis 1.3 holds.*

(A) *Let $\pi_2 = \mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}(\chi)$ be a principal series representation of $\mathrm{GL}_2(F)$. Assume $\mathrm{Hom}_{A(F)}(Q, \chi \cdot \delta^{1/2}) = 0$. Then*

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) = \dim \mathrm{Hom}_{Z(F)}((\pi_1)_{N(E),\psi}, \omega_{\pi_2}).$$

(B) *Let $\pi_1 = \mathrm{Ind}_{\widetilde{B}(E)}^{\widetilde{\mathrm{GL}}_2(E)}(\tilde{\tau})$ be a principal series representation of $\widetilde{\mathrm{GL}}_2(E)$ and π_2 a discrete series representation of $\mathrm{GL}_2(F)$. Let π_2' be the finite-dimensional representation of $D_F^\times$ associated to π_2 by the Jacquet–Langlands correspondence. Assume that*

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}\big(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}(\tilde{\tau}.\delta^{1/2}), \pi_2\big) = 0.$$

Then

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) + \dim \mathrm{Hom}_{D_F^\times}(\pi_1, \pi_2') = \big[E^\times : F^\times E^{\times 2}\big].$$

(C) *Let π_1 be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and π_2 a supercuspidal representation of $\mathrm{GL}_2(F)$. Let π_1' be a genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ which has the same central character as that of π_1 and as a $\widetilde{Z}$-module $(\pi_1)_{N(E),\psi} \oplus (\pi_1')_{N(E),\psi} = \Omega(\omega_{\pi_1})$. Let π_2' be the finite-dimensional representation of $D_F^\times$ associated to π_2 by the Jacquet–Langlands correspondence. Then*

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1 \oplus \pi_1', \pi_2) + \dim \mathrm{Hom}_{D_F^\times}(\pi_1 \oplus \pi_1', \pi_2') = [E^\times : F^\times E^{\times 2}].$$

The strategy to prove this theorem is similar to that in [Prasad 1992]. We recall it briefly. Part (A) of this theorem is proved by looking at the Kirillov model of an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and its Jacquet module with respect to $N(F)$. Part (B) makes use of Mackey theory. For the third part (C), we use a trick of Prasad [1992], where we "transfer" the results of a principal series representation (from part (B)) to those which do not belong to principal series. Prasad transfers the results from principal series representations to discrete series representations. This is done by using character theory and an analog of a result of Casselman and Prasad [Prasad 1992, Theorem 2.7] for $\widetilde{\mathrm{GL}}_2(E)$ which we study in Section 4.

2. Part A of Theorem 1.5

Let $\pi_2 = \mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}(\chi)$ be a principal series representation of $\mathrm{GL}_2(F)$ where χ is a character of $A(F)$. By Frobenius reciprocity [Bernstein and Zelevinskii 1976, Theorem 2.28], we get

$$\begin{aligned}\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) &= \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}(\chi)) \\ &= \mathrm{Hom}_{A(F)}((\pi_1)_{N(F)}, \chi . \delta^{1/2}),\end{aligned}$$

where $(\pi_1)_{N(F)}$ is the Jacquet module of π_1 with respect to $N(F)$. We can describe $(\pi_1)_{N(F)}$ by realizing π_1 in the Kirillov model. Now depending on whether π_1 is a supercuspidal representation or not, we consider them separately.

2A. *Kirillov model and Jacquet module.* Now we describe the Kirillov model of an irreducible admissible genuine representation π of $\widetilde{\mathrm{GL}}_2(E)$. Let $l : \pi \to \pi_{N(E),\psi}$ be the canonical map. Let $\mathcal{C}^\infty(E^\times, \pi_{N(E),\psi})$ denote the space of smooth functions on $E^\times$ with values in $\pi_{N(E),\psi}$. Define the Kirillov mapping

$$\mathtt{K} : \pi \to \mathcal{C}^\infty(E^\times, \pi_{N(E),\psi})$$

given by $v \mapsto \xi_v$ where $\xi_v(x) = l\big(\pi\big(\big(\begin{smallmatrix} x & 0 \\ 0 & 1\end{smallmatrix}\big), 1\big)v\big)$. More conceptually, $\pi_{N(E),\psi}$ is a representation of $\widetilde{Z} \cdot N(E)$, and by Frobenius reciprocity, there exists a natural map

$$\pi|_{\widetilde{B}(E)} \to \mathrm{Ind}_{\widetilde{Z}\cdot N(E)}^{\widetilde{B}(E)} \pi_{N(E),\psi}.$$

Since $\widetilde{B}(E)/\widetilde{Z} \cdot N(E)$ can be identified with $E^\times$ sitting as $\big\{\big(\begin{smallmatrix} e & 0 \\ 0 & 1\end{smallmatrix}\big) : e \in E^\times\big\}$ in $\widetilde{B}(E)$, we get a map of $\widetilde{B}(E)$-modules:

$$\pi|_{\widetilde{B}(E)} \to C^\infty(E^\times, \pi_{N(E),\psi}).$$

We summarize some of the properties of the Kirillov mapping in the following proposition.

Proposition 2.1. (1) *If $v' = \pi\big(\big(\begin{smallmatrix} a & b \\ 0 & d\end{smallmatrix}\big), 1\big)v$ for $v \in \pi$ then*

$$\xi_{v'}(x) = (x, d)\psi(bd^{-1}x)\pi\left(\begin{pmatrix} d & 0 \\ 0 & d\end{pmatrix}, 1\right)\xi_v(ad^{-1}x).$$

(2) *For $v \in \pi$ the function ξ_v is a locally constant function on $E^\times$ which vanishes outside a compact subset of E.*

(3) *The map $\mathtt{K}$ is an injective linear map.*

(4) *The image $\mathtt{K}(\pi)$ of the map $\mathtt{K}$ contains the space $\mathcal{S}(E^\times, \pi_{N(E),\psi})$ of smooth functions on $E^\times$ with compact support with values in $\pi_{N(E),\psi}$.*

(5) *The Jacquet module $\pi_{N(E)}$ of π is isomorphic to $\mathtt{K}(\pi)/\mathcal{S}(E^\times, \pi_{N(E),\psi})$.*

(6) *The representation π is supercuspidal if and only if $\mathtt{K}(\pi) = \mathcal{S}(E^\times, \pi_{N(E),\psi})$.*

Proof. Part (1) follows from the definition. The proofs of parts (2) and (3) are verbatim those of Lemma 2 and Lemma 3 in [Godement 1970]. The proofs of parts (4), (5) and (6) follow from the proofs of the corresponding statements of [Prasad and Raghuram 2000, Theorem 3.1]. □

Since the map $\mathtt{K}$ is injective, we can transfer the action of $\widetilde{\mathrm{GL}}_2(E)$ on the space of π to $\mathtt{K}(\pi)$ using the map $\mathtt{K}$. The realization of the representation π on the space $\mathtt{K}(\pi)$ is called the Kirillov model, on which the action of $\widetilde{B}(E)$ is explicitly given by part (1) in Proposition 2.1. It is clear that $\mathcal{S}(E^\times, \pi_{N(E),\psi})$ is $\widetilde{B}(E)$-stable, which gives rise to the following short exact sequence of $\widetilde{B}(E)$-modules

$$0 \to \mathcal{S}(E^\times, \pi_{N(E),\psi}) \to \mathtt{K}(\pi) \to \pi_{N(E)} \to 0. \tag{1}$$

2B. *The Jacquet module with respect to $N(F)$.* In this section, we try to understand the restriction of an irreducible admissible genuine representation π of $\widetilde{\mathrm{GL}}_2(E)$ to $B(F)$. For this, we describe the Jacquet module $\pi_{N(F)}$ of π. We utilize the short exact sequence in equation (1) of $\widetilde{B}(E)$-modules arising from the Kirillov model of π, which is also a short exact sequence of $B(F)$-modules. By the exactness of the Jacquet functor with respect to $N(F)$, we get the following short exact sequence from equation (1),

$$0 \to \mathcal{S}(E^\times, \pi_{N(E),\psi})_{N(F)} \to \mathtt{K}(\pi)_{N(F)} \to \pi_{N(E)} \to 0.$$

Let us first describe $\mathcal{S}(E^\times, \pi_{N(E),\psi})_{N(F)}$, the Jacquet module of $\mathcal{S}(E^\times, \pi_{N(E),\psi})$ with respect to $N(F)$.

Proposition 2.2. *There exists an isomorphism*

$$\mathcal{S}(E^\times, \pi_{N(E),\psi})_{N(F)} \cong \mathcal{S}(F^\times, \pi_{N(E),\psi})$$

of $F^\times$-modules where $F^\times$ acts by its natural action on $\mathcal{S}(F^\times, \pi_{N(E),\psi})$.

Proposition 2.2 follows from the proposition below. The author thanks Professor Prasad for suggesting the proof.

Proposition 2.3. *Let ψ be a nontrivial additive character of E such that $\psi|_F = 1$. Let $\mathcal{S}(E^\times)$ be a representation of E where the action of E on $\mathcal{S}(E^\times)$ is given by*

$$(n \cdot f)(x) = \psi(nx) f(x)$$

for $n \in E$, $f \in \mathcal{S}(E^\times)$ and $x \in E^\times$. Then the restriction map

$$\mathcal{S}(E^\times) \longrightarrow \mathcal{S}(F^\times) \tag{2}$$

realizes $\mathcal{S}(E^\times)_F$ the maximal F-coinvariant quotient of $\mathcal{S}(E^\times)$ as $\mathcal{S}(F^\times)$.

Proof. Note that $\mathcal{S}(E^\times) \hookrightarrow \mathcal{S}(E)$. For a fixed Haar measure dw on E, we define the Fourier transform $\mathcal{F}_\psi : \mathcal{S}(E) \to \mathcal{S}(E)$ with respect to the character ψ by

$$\mathcal{F}_\psi(f)(z) := \int_E f(w)\psi(zw)\, dw.$$

As is well known, $\mathcal{F}_\psi : \mathcal{S}(E) \to \mathcal{S}(E)$ is an isomorphism of vector spaces, and the image of $\mathcal{S}(E^\times)$ can be identified with those functions in $\mathcal{S}(E)$ whose integral on

E is zero. The Fourier transform takes the action of E on $\mathcal{S}(E^\times)$ to the restriction of the action of E on $\mathcal{S}(E)$ given by $(n \cdot f)(x) = f(x+n)$ for $n \in E$, $f \in \mathcal{S}(E)$ and $x \in E$. Write $E = F(\sqrt{d})$ for a suitable $d \in F^\times$. Define $\phi : E \to F$ given by

$$\phi(e) = (e - \bar{e})/(2\sqrt{d}),$$

where $\bar{e}$ is the nontrivial Galois conjugate of $e \in E$, i.e., $\bar{e} = x - \sqrt{d}y$ for $e = x + \sqrt{d}y$ with $x, y \in F$. Clearly $\phi(z_1) = \phi(z_2)$ for $z_1, z_2 \in E$ if and only if $z_1 - z_2 \in F$. We define the integration along the fibers of the map $\phi : E \to F$, to be denoted by $I : \mathcal{S}(E) \to \mathcal{S}(F)$, as follows:

$$I(f)(y) := \int_F f(x + \sqrt{d}y)\, dx \quad \text{for all } y \in F.$$

Clearly $I(f)$ belongs to $\mathcal{S}(F)$. Note that the maximal quotient of $\mathcal{S}(E)$ on which F acts trivially (F acting by translation on $\mathcal{S}(E)$) can be identified with $\mathcal{S}(F)$ by integration along the fibers of the map ϕ. Since $\psi|_F = 1$, the restriction of the character $\psi_{\sqrt{d}}$ (given by $x \mapsto \psi(\sqrt{d}x)$ for $x \in E$) from E to F is a nontrivial character of F. The proposition will follow if we prove the commutativity of the diagram

$$\begin{array}{ccc} \mathcal{S}(E) & \xrightarrow{\mathcal{F}_\psi} & \mathcal{S}(E) \\ {\scriptstyle \mathrm{Res}}\big\downarrow & & \big\downarrow{\scriptstyle I} \\ \mathcal{S}(F) & \xrightarrow{\mathcal{F}_{\psi_{\sqrt{d}}}} & \mathcal{S}(F) \end{array}$$

where $\mathcal{F}_\psi$ is the Fourier transform on $\mathcal{S}(E)$ with respect to the character ψ, $\mathcal{F}_{\psi_{\sqrt{d}}}$ is the Fourier transform on $\mathcal{S}(F)$ with respect to $\psi_{\sqrt{d}} = (\psi_{\sqrt{d}})|_F$), Res denotes the restriction of functions from E to F, and I denotes the integration along the fibers mentioned above. Recall that $\mathcal{F}_{\psi_{\sqrt{d}}} : \mathcal{S}(F) \to \mathcal{S}(F)$ is defined by

$$\mathcal{F}_{\psi_{\sqrt{d}}}(\phi)(x) := \int_F \phi(y)\psi_{\sqrt{d}}(xy)\, dy = \int_F \phi(y)\psi(\sqrt{d}xy)\, dy \text{ for all } x \in F.$$

We prove that the above diagram is commutative. Let $f \in \mathcal{S}(E)$. We want to show that $I \circ \mathcal{F}_\psi(f)(y) = \mathcal{F}_{\psi_{\sqrt{d}}} \circ \mathrm{Res}(f)(y)$ for all $y \in F$. We write an element of E as $x + \sqrt{d}y$ with $x, y \in F$. We choose a measure dx on F which is self dual with respect to $\psi_{\sqrt{d}}$ in the sense that $\mathcal{F}_{\psi_{\sqrt{d}}}(\mathcal{F}_{\psi_{\sqrt{d}}}(\phi))(x) = \phi(-x)$ for all $\phi \in \mathcal{S}(F)$ and $x \in F$. We identify E with $F \times F$ as a vector space. Consider the product measure $dx\, dy$ on $E = F \times F$. Using Fubini's theorem we have

$$\int_F \int_F \phi(z_2)\psi_{\sqrt{d}}(xz_2)\, dz_2\, dx = \mathcal{F}_{\psi_{\sqrt{d}}}(\mathcal{F}_{\psi_{\sqrt{d}}}(\phi))(0) = \phi(0)$$

for $\phi \in \mathcal{S}(F)$. Therefore,

$$
\begin{aligned}
I \circ \mathcal{F}_{\psi}(f)(y) &= \int_F \mathcal{F}_{\psi}(f)\big(x+\sqrt{d}y\big)\,dx \\
&= \int_F \int_{E=F\times F} f\big(z_1+\sqrt{d}z_2\big)\psi\big(\big(x+\sqrt{d}y\big)\big(z_1+\sqrt{d}z_2\big)\big)\,dz_1\,dz_2\,dx \\
&= \int_F \int_F \int_F f\big(z_1+\sqrt{d}z_2\big)\psi_{\sqrt{d}}(yz_1+xz_2)\,dz_1\,dz_2\,dx \\
&= \int_F \bigg(\int_F \int_F f\big(z_1+\sqrt{d}z_2\big)\psi_{\sqrt{d}}(xz_2)\,dz_2\,dx\bigg)\psi_{\sqrt{d}}(yz_1)\,dz_1 \\
&= \int_F f(z_1)\psi_{\sqrt{d}}(yz_1)\,dz_1 = \mathcal{F}_{\psi_{\sqrt{d}}} \circ \mathrm{Res}(f)(y).
\end{aligned}
$$

This proves the commutativity of the above diagram. $\square$

2C. ***Completion of the proof of Part (A).*** First we consider the case when π_1 is a supercuspidal representation of $\widetilde{\mathrm{GL}}_2(E)$. Then one knows that the functions in the Kirillov model for π_1 have compact support in $E^\times$ and one has

$$\pi_1 \cong \mathcal{S}(E^\times, (\pi_1)_{N(E),\psi})$$

as $\widetilde{B}(E)$modules by Proposition 2.1. Now using Proposition 2.2, we get

$$
\begin{aligned}
\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1,\pi_2) &= \mathrm{Hom}_{A(F)}\big((\pi_1)_{N(F)},\chi.\delta^{1/2}\big) \\
&= \mathrm{Hom}_{A(F)}\big(\mathcal{S}(E^\times,(\pi_1)_{N(E),\psi})_{N(F)},\chi.\delta^{1/2}\big) \\
&= \mathrm{Hom}_{A(F)}\big(\mathcal{S}(F^\times,(\pi_1)_{N(E),\psi}),\chi.\delta^{1/2}\big).
\end{aligned}
$$

Since $\mathcal{S}(F^\times,(\pi_1)_{N,\psi}) \cong \mathrm{ind}_{Z(F)}^{A(F)}(\pi_1)_{N(E),\psi}$ as $A(F)$-modules, by Frobenius reciprocity [Bernstein and Zelevinskii 1976, Proposition 2.29], we get

$$
\begin{aligned}
\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1,\pi_2) &= \mathrm{Hom}_{A(F)}\big(\mathrm{ind}_{Z(F)}^{A(F)}(\pi_1)_{N(E),\psi},\chi.\delta^{1/2}\big) \\
&= \mathrm{Hom}_{Z(F)}\big((\pi_1)_{N(E),\psi},(\chi.\delta^{1/2})|_{Z(F)}\big) \\
&= \mathrm{Hom}_{Z(F)}((\pi_1)_{N(E),\psi},\omega_{\pi_2}).
\end{aligned}
$$

This proves part (A) of Theorem 1.5 for π_1 a supercuspidal representation.

Now we consider the case when π_1 is not a supercuspidal representation of $\widetilde{\mathrm{GL}}_2(E)$. Then from equation (1) we get the following short exact sequence of $A(F)$-modules:

$$0 \to \mathcal{S}(F^\times,(\pi_1)_{N(E),\psi}) \to (\pi_1)_{N(F)} \to Q \longrightarrow 0.$$

Now applying the functor $\mathrm{Hom}_{A(F)}\big(-,\chi.\delta^{1/2}\big)$, we get the long exact sequence

$$
\begin{aligned}
0 \to \mathrm{Hom}_{A(F)}\big(Q,\chi.\delta^{1/2}\big) &\to \mathrm{Hom}_{A(F)}\big((\pi_1)_{N(F)},\chi.\delta^{1/2}\big) \\
&\to \mathrm{Hom}_{A(F)}\big(\mathcal{S}(F^\times,(\pi_1)_{N(E),\psi}),\chi.\delta^{1/2}\big) \to \mathrm{Ext}^1_{A(F)}\big(Q,\chi.\delta^{1/2}\big) \to \cdots
\end{aligned}
$$

Lemma 2.4. $\mathrm{Hom}_{A(F)}\big(Q, \chi.\delta^{1/2}\big) = 0$ *if and only if* $\mathrm{Ext}^1_{A(F)}\big(Q, \chi.\delta^{1/2}\big) = 0$.

Proof. The space Q is finite-dimensional and completely reducible. So it is enough to prove the lemma for the one-dimensional representations, i.e., for characters of $A(F)$. Moreover one can regard these representations as representations of $F^\times$ (after tensoring by a suitable character of $A(F)$ so that it descends to a representation of $A(F)/Z(F) \cong F^\times$). Then our lemma follows from the following lemma due to Prasad. □

Lemma 2.5. *If χ_1 and χ_2 are two characters of $F^\times$, then*

$$\dim \mathrm{Hom}_{F^\times}(\chi_1, \chi_2) = \dim \mathrm{Ext}^1_{F^\times}(\chi_1, \chi_2).$$

Proof. Let $\mathcal{O}$ be the ring of integers of F and ϖ a uniformizer of F. Since $F^\times \cong \mathcal{O}^\times \times \varpi^{\mathbb{Z}}$ and $\mathcal{O}^\times$ is compact, $\mathrm{Ext}^i_{F^\times}(\chi_1, \chi_2) = H^i(\mathbb{Z}, \mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2))$. If $\mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2) = 0$, then the lemma is obvious. Hence suppose $\mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2) \neq 0$. Then $\mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2)$ is a certain one dimensional vector space with an action of $\varpi^{\mathbb{Z}}$. If the action of $\varpi^{\mathbb{Z}}$ on $\mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2)$ is nontrivial then $H^i(\mathbb{Z}, \mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2)) = 0$ for all $i \geq 0$. Whereas if the action of $\varpi^{\mathbb{Z}}$ on $\mathrm{Hom}_{\mathcal{O}^\times}(\chi_1, \chi_2)$ is trivial, then $H^0(\mathbb{Z}, \mathbb{C}) \cong H^1(\mathbb{Z}, \mathbb{C}) \cong \mathbb{C}$. □

We have made an assumption that $\mathrm{Hom}_{A(F)}(Q, \chi.\delta^{1/2}) = 0$ and hence by the lemma above, $\mathrm{Ext}^1_{A(F)}(Q, \chi.\delta^{1/2}) = 0$. So in this case

$$\begin{aligned} \mathrm{Hom}_{A(F)}\big((\pi_1)_{N(F)}, \chi.\delta^{1/2}\big) &\cong \mathrm{Hom}_{A(F)}\big(\mathcal{S}(F^\times, (\pi_1)_{N(E),\psi}), \chi.\delta^{1/2}\big) \\ &= \mathrm{Hom}_{Z(F)}((\pi_1)_{N(E),\psi}, \omega_{\pi_2}). \end{aligned}$$

Hence

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) = \dim \mathrm{Hom}_{Z(F)}((\pi_1)_{N(E),\psi}, \omega_{\pi_2}).$$

Remark 2.6. Recall that $Q := (\pi_1)_{N(E)}$ is a finite-dimensional representation of $\tilde{A}(E)$ and we have assumed that $\mathrm{Hom}_{A(F)}(Q, \chi.\delta^{1/2}) = 0$. The number of characters χ of $A(F)$ for which $\mathrm{Hom}_{A(F)}(Q, \chi.\delta^{1/2}) \neq 0$ is at most the dimension of Q. The maximum possible dimension of Q is $2[E^\times : E^{\times 2}]$ (the maximum occurs only if π_1 is a principal series representation). Therefore for a given π_1 we leave out finitely many ($\leq 2[E^\times : E^{\times 2}]$) representations π_2 in our analysis.

3. Part B of Theorem 1.5

In this section, we consider the case when π_1 is a principal series representation of $\widetilde{\mathrm{GL}}_2(E)$ and π_2 a discrete series representation of $\mathrm{GL}_2(F)$.

Let $\pi_1 = \mathrm{Ind}_{\tilde{B}(E)}^{\widetilde{\mathrm{GL}}_2(E)}(\tilde{\tau})$, where $(\tilde{\tau}, V)$ is a genuine irreducible representation of $\tilde{A} = \tilde{A}(E)$. The group $\tilde{A}$ sits in the central extension

$$1 \to A^2 \times \{\pm 1\} \to \tilde{A} \xrightarrow{p} A/A^2 \to 1,$$

where A/A^2 equals $E^\times/E^{\times 2} \times E^\times/E^{\times 2}$, and the commutator of two elements $\tilde{a}_1$ and $\tilde{a}_2$ of $\tilde{A}$ whose images in A/A^2 are $a_1 = (e_1, f_1)$ and $a_2 = (e_2, f_2)$, is

$$[\tilde{a}_1, \tilde{a}_2] = (e_1, f_2)(e_2, f_1) \in \{\pm 1\} \subset A^2 \times \{\pm 1\},$$

which is the product of Hilbert symbols (e_i, f_j) of E. Since the Hilbert symbol is a nondegenerate bilinear form on $E^\times/E^{\times 2}$, it follows that

$$[\tilde{a}_1, \tilde{a}_2] : A/A^2 \times A/A^2 \to \{\pm 1\}$$

is also a nondegenerate (skew-symmetric) bilinear form. Thus $\tilde{A}$ is closely related to the "usual Heisenberg" groups, and its representation theory is closely related to the representation theory of the "usual Heisenberg" groups. In particular, given a character $\chi : A^2 \times \{\pm 1\} \to \mathbb{C}^\times$ which is nontrivial on $\{\pm 1\}$, there exists a unique irreducible representation of $\tilde{A}$ which contains χ. Further, for any subgroup $A_0 \subset A/A^2$ for which the commutator map $[\tilde{a}_1, \tilde{a}_2]$, $a_i \in A_0$, is identically trivial, and for which A_0 is maximal for this property, $\tilde{A}_0 = p^{-1}(A_0)$ is a maximal abelian subgroup of $\tilde{A}$, and the restriction of an irreducible genuine representation $\tilde{\tau}$ of $\tilde{A}$ to $\tilde{A}_0$ contains all characters of $\tilde{A}_0$ with multiplicity one whose restriction to the center $A^2 \times \{\pm 1\}$ is the central character of $\tilde{\tau}$. Further, $\tilde{\tau} = \mathrm{Ind}_{\tilde{A}_0}^{\tilde{A}} \chi$ where χ is any character of $\tilde{A}_0$ appearing in $\tilde{\tau}$. All the assertions here are consequences of the fact that the inner conjugation action of $\tilde{A}$ on $\tilde{A}_0$ is transitive on the set of characters of $\tilde{A}_0$ with a given restriction on $A^2 \times \{\pm 1\}$; this itself is a consequence of the nondegeneracy of the Hilbert symbol.

It follows that the set of equivalence classes of irreducible genuine representations $\tilde{\tau}$ of $\tilde{A}$ is parametrized by the set of characters of A^2, i.e., a pair of characters of $E^{\times 2}$.

Lemma 3.1. *The subgroup $\widetilde{Z} \cdot A^2$ of $\tilde{A}$ is a maximal abelian subgroup. Let $\tilde{\tau}$ be an irreducible genuine representation of $\tilde{A}$. Then $\tilde{\tau}|_{\widetilde{Z}}$ contains all the genuine characters of $\widetilde{Z}$ which agree with the central character of τ when restricted to $\widetilde{Z}^2$.*

Proof. By explicit description of the commutation relation recalled above it is easy to see that $\widetilde{Z} \cdot A^2$ is a maximal abelian subgroup of $\tilde{A}$. The rest of the statements follow from preceding discussion. □

Proposition 3.2 [Gelbart and Piatetski-Shapiro 1980, Theorem 2.4]. *For some irreducible genuine representation $\tilde{\tau}$ of $\tilde{A}$, let $\pi_1 = \mathrm{Ind}_{\widetilde{B}(E)}^{\widetilde{\mathrm{GL}}_2(E)}(\tilde{\tau})$. Then*

$$(\pi_1)_{N,\psi} \cong \Omega(\pi_1) \cong \tilde{\tau}|_{\widetilde{Z}}.$$

Now as in [Prasad 1992], we use Mackey theory to understand its restriction to $\mathrm{GL}_2(F)$. We have $\widetilde{\mathrm{GL}}_2(E)/\widetilde{B}(E) \cong \mathbb{P}^1_E$ and this has two orbits under the left action of $\mathrm{GL}_2(F)$. One of the orbits is closed, and naturally identified with $\mathbb{P}^1_F \cong \mathrm{GL}_2(F)/B(F)$. The other orbit is open, and can be identified with $\mathbb{P}^1_E - \mathbb{P}^1_F \cong \mathrm{GL}_2(F)/E^\times$. By Mackey theory, we get this exact sequence of $\mathrm{GL}_2(F)$-modules:

$$0 \to \mathrm{ind}_{E^\times}^{\mathrm{GL}_2(F)}(\tilde{\tau}'|_{E^\times}) \to \pi_1 \to \mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}(\tilde{\tau}|_{B(F)}\delta^{1/2}) \to 0, \tag{3}$$

where $\tilde{\tau}'|_{E^\times}$ is the representation of $E^\times$ obtained from the embedding $E^\times \hookrightarrow \tilde{A}$ which comes from conjugating the embedding $E^\times \hookrightarrow \mathrm{GL}_2(F) \hookrightarrow \widetilde{\mathrm{GL}}_2(E)$. We now identify $E^\times$ with its image inside $\tilde{A}$ which is given by $x \mapsto \bigl(\bigl(\begin{smallmatrix} x & 0 \\ 0 & \bar{x}\end{smallmatrix}\bigr), \epsilon(x)\bigr)$ where $\bar{x}$ is the nontrivial $\mathrm{Gal}(E/F)$-conjugate of x and $\epsilon(x) \in \{\pm 1\}$. Now let π_2 be any irreducible admissible representation of $\mathrm{GL}_2(F)$. By applying the functor $\mathrm{Hom}_{\mathrm{GL}_2(F)}(-, \pi_2)$ to the short exact sequence (3), we get the long exact sequence

$$\begin{aligned} &0 \to \mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}|_{B(F)}\delta^{1/2}\bigr), \pi_2\bigr) \\ &\to \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) \to \mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{ind}_{E^\times}^{\mathrm{GL}_2(F)}(\tilde{\tau}'|_{E^\times}), \pi_2\bigr) \\ &\to \mathrm{Ext}^1_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}|_{B(F)}\delta^{1/2}\bigr), \pi_2\bigr) \to \cdots \end{aligned} \tag{4}$$

From [Prasad 1990, Corollary 5.9], we know that

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\chi.\delta^{1/2}\bigr), \pi_2\bigr) = 0 \Leftrightarrow \mathrm{Ext}^1_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\chi.\delta^{1/2}\bigr), \pi_2\bigr) = 0.$$

Since $\tilde{\tau}|_{B(F)}$ factors through $T(F)$, which is direct sum of $[E^\times : E^{\times 2}]$ characters of $T(F)$, we can use the above result of Prasad with χ replaced by $\tilde{\tau}|_{B(F)}$. Then from the exactness of (4), it follows that

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) = 0$$

if and only if

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}|_{B(F)}\delta^{1/2}\bigr), \pi_2\bigr) = 0$$

and

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{ind}_{E^\times}^{\mathrm{GL}_2(F)}(\tilde{\tau}'|_{E^\times}), \pi_2\bigr) = 0.$$

Note that the representation $\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}|_{B(F)}\delta^{1/2}\bigr)$ consists of exactly $[E^\times : E^{\times 2}]$ principal series representations of $\mathrm{GL}_2(F)$. Since we have made the assumption that $\mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}.\delta^{1/2}\bigr), \pi_2\bigr) = 0$, it follows that

$$\mathrm{Ext}^1_{\mathrm{GL}_2(F)}\bigl(\mathrm{Ind}_{B(F)}^{\mathrm{GL}_2(F)}\bigl(\tilde{\tau}.\delta^{1/2}\bigr), \pi_2\bigr) = 0.$$

This gives

$$\begin{aligned} \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) &\cong \mathrm{Hom}_{\mathrm{GL}_2(F)}\bigl(\mathrm{ind}_{E^\times}^{\mathrm{GL}_2(F)}(\tilde{\tau}'|_{E^\times}), \pi_2\bigr) \\ &\cong \mathrm{Hom}_{E^\times}(\tilde{\tau}'|_{E^\times}, \pi_2|_{E^\times}). \end{aligned}$$

The following lemma describes $\tilde{\tau}'|_{E^\times}$.

Lemma 3.3. *If we identify $E^\times$ with its image $\bigl\{\bigl(\bigl(\begin{smallmatrix} x & 0 \\ 0 & \bar{x}\end{smallmatrix}\bigr), \epsilon(x)\bigr) \mid x \in E^\times\bigr\}$ inside $\tilde{A}$ as above then the subgroup $E^\times \cdot \tilde{A}^2$ inside $\tilde{A}$ is a maximal abelian subgroup. Moreover, $\tilde{\tau}'|_{E^\times}$ contains all the characters of $E^\times$ which are same as $\omega_{\tilde{\tau}}|_{E^{\times 2}}$ when restricted to $E^{\times 2}$, where $\omega_{\tilde{\tau}}$ is the central character of $\tilde{\tau}$.*

Proof. From the explicit cocycle description and the nondegeneracy of the quadratic Hilbert symbol, it is easy to verify that $E^\times \cdot \tilde{A}^2$ is a maximal abelian subgroup of $\tilde{A}$. The rest follows from the discussion preceding Lemma 3.1. $\square$

As π_2 is a discrete series representation, it is not always true (unlike what happens in case of a principal series representation) that any character of $E^\times$, whose restriction to $F^\times$ is the same as the central character of π_2, appears in π_2. Let π_2' be the finite dimensional representation of $D_F^\times$ associated to π_2 by the Jacquet–Langlands correspondence. Considering the left action of $D_F^\times$ on

$$\mathbb{P}_E^1 \cong \widetilde{\mathrm{GL}}_2(E)/\widetilde{B}(E)$$

induced by $D_F^\times \hookrightarrow \widetilde{\mathrm{GL}}_2(E)$ it is easy to verify that $\mathbb{P}_E^1 \cong D_F^\times/E^\times$. Then by Mackey theory, when restricted to $D_F^\times$, the principal series representation π_1 becomes isomorphic to $\mathrm{ind}_{E^\times}^{D_F^\times}(\tilde{\tau}'|_{E^\times})$. Therefore,

$$\begin{aligned}\mathrm{Hom}_{D_{F^\times}}(\pi_1, \pi_2') &\cong \mathrm{Hom}_{D_{F^\times}}\left(\mathrm{ind}_{E^\times}^{D_F^\times}(\tilde{\tau}'|_{E^\times}), \pi_2'\right)\\ &\cong \mathrm{Hom}_{E^\times}(\tilde{\tau}'|_{E^\times}, \pi_2'|_{E^\times}).\end{aligned}$$

In order to prove

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1, \pi_2) + \dim \mathrm{Hom}_{D_F^\times}(\pi_1, \pi_2') = [E^\times : F^\times E^{\times 2}] \tag{5}$$

we shall prove

$$\dim \mathrm{Hom}_{E^\times}(\tilde{\tau}'|_{E^\times}, \pi_2|_{E^\times}) + \dim \mathrm{Hom}_{E^\times}(\tilde{\tau}'|_{E^\times}, \pi_2'|_{E^\times}) = [E^\times : F^\times E^{\times 2}]. \tag{6}$$

By Remark 2.9 in [Prasad 1992], a character of $E^\times$ whose restriction to $F^\times$ is the same as the central character of π_2 appears either in π_2 with multiplicity one or in π_2' with multiplicity one, and exactly one of the two possibilities hold. Note that we are assuming that the two embeddings of $E^\times$, one via $\mathrm{GL}_2(F)$ and the other via $D_F^\times$ are conjugate in $\widetilde{\mathrm{GL}}_2(E)$. Then the left-hand side of equation (6) is the same as the number of characters of $E^\times$ appearing in $(\tilde{\tau}, V)$ which upon restriction to $F^\times$ coincide with the central character of π_2, which equals $\dim \mathrm{Hom}_{F^\times}(\tilde{\tau}|_{F^\times}, \omega_{\pi_2})$. We are reduced to the following lemma.

Lemma 3.4. *Let $(\tilde{\tau}, V)$ be an irreducible genuine representation of $\tilde{A}$ and let χ be a character of $Z(F) = F^\times$ such that $\chi|_{E^{\times 2} \cap F^\times} = \tilde{\tau}|_{E^{\times 2} \cap F^\times}$. Then*

$$\dim \mathrm{Hom}_{F^\times}(\tilde{\tau}, \chi) = [E^\times : F^\times E^{\times 2}].$$

Proof. Note that $E^{\times 2} \cap F^\times = Z^{\times 2} \cap F^\times$. From Proposition 3.2, $\tilde{\tau}|_{\tilde{Z}} \cong \Omega(\omega_{\pi_1})$. If a character $\mu \in \Omega(\omega_{\pi_1})$ is specified on $F^\times$ then it is specified on $F^\times E^{\times 2}$. Therefore the number of characters in $\Omega(\omega_{\pi_1})$ which agree with χ when restricted to $F^\times$ is equal to $[E^\times : F^\times E^{\times 2}]$. $\square$

4. A theorem of Casselman and Prasad

As mentioned in the introduction, we use results of part (B) involving principal series representation and "transfer" these to the other cases, as stated in part (C) which involves restriction of the two representations. To make such a transfer possible Prasad used a result which says that if two irreducible representations of $\mathrm{GL}_2(E)$ have the same central characters then the difference of their characters is a smooth function on $\mathrm{GL}_2(E)$. We will need a similar theorem for $\widetilde{\mathrm{GL}}_2(E)$, which we prove in this section. In order to do this, we recall a variant of a theorem of Rodier which is true for covering groups in general; this variant is proved in [Patel 2015]. Let us first recall some facts about germ expansions, restricted only to $\widetilde{\mathrm{SL}}_2(E)$.

For any nonzero nilpotent orbit in $\mathfrak{sl}_2(E)$ there is a lower triangular nilpotent matrix $Y_a = \left(\begin{smallmatrix} 0 & 0 \\ a & 0 \end{smallmatrix}\right)$ such that Y_a belongs to the nilpotent orbit. For a given nonzero nilpotent orbit, the element a is uniquely determined modulo $E^{\times 2}$. We write $\mathcal{N}_a$ for the nilpotent orbit which contains Y_a. Thus the set of all nonzero nilpotent orbits is $\{\mathcal{N}_a \mid a \in E^\times/E^{\times 2}\}$.

Let τ be an irreducible admissible genuine representation of $\widetilde{\mathrm{SL}}_2(E)$. Recall that for an irreducible admissible genuine representation τ of $\widetilde{\mathrm{SL}}_2(E)$, the character distribution Θ_τ is a smooth function on the set of regular semisimple elements. The Harish-Chandra–Howe character expansion of Θ_τ in a neighborhood of the identity is given as follows:

$$\Theta_\tau \circ \exp = c_0(\tau) + \sum_{a \in E^\times/E^{\times 2}} c_a(\tau) \cdot \hat{\mu}_{\mathcal{N}_a}$$

where $c_0(\tau), c_a(\tau)$ are constants and $\hat{\mu}_{\mathcal{N}_a}$ is the Fourier transform of a suitably chosen $\mathrm{SL}_2(E)$-invariant (under the adjoint action) measure on $\mathcal{N}_a$.

Fix a nontrivial additive character ψ of E. Define a character χ of N by $\chi\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right) = \psi(x)$. For $a \in E^\times$ we write ψ_a for the character of E given by $\psi_a(x) = \psi(ax)$. We write (N, ψ) for the nondegenerate Whittaker datum (N, χ). It can be seen that the set of conjugacy classes of nondegenerate Whittaker data has a set of representatives $\{(N, \psi_a) \mid a \in E^\times/E^{\times 2}\}$.

From the proof of the main theorem in [Patel 2015], the bijection between $\{\mathcal{N}_a \mid a \in E^\times/E^{\times 2}\}$ and $\{(N, \psi_a) \mid a \in E^\times/E^{\times 2}\}$ given by $\mathcal{N}_a \leftrightarrow (N, \psi_a)$ satisfies the following property: $c_a \neq 0$ if and only if the representation τ of $\widetilde{\mathrm{SL}}_2(E)$ admits a nonzero (N, ψ_a)-Whittaker functional.

It follows from [Gelbart et al. 1979, Theorem 4.1] that for any nontrivial additive character ψ' of N, the dimension of the space of (N, ψ')-Whittaker functionals for τ is at most one. Therefore, from the theorem of Rodier, as extended in [Patel 2015], each $c_a(\tau)$ is either 1 or 0 depending on whether τ admits a nonzero Whittaker functional corresponding to the nondegenerate Whittaker datum (N, ψ_a) or not.

Remark 4.1. Let $\widetilde{G}$ be a topological central extension of a connected reductive group G by μ_r, a cyclic group of order r. For $g \in \widetilde{G}$ there exists a semisimple element $g_s \in \widetilde{G}$ such that g belongs to any conjugation invariant neighborhood of $g_s \in \widetilde{G}$.

Let τ_1 and τ_2 be two irreducible admissible genuine representations of $\widetilde{\mathrm{SL}}_2(E)$. Note that $\widetilde{\{\pm 1\}}$ is the center of $\widetilde{\mathrm{SL}}_2(E)$ and these are the only nonregular semisimple elements of $\widetilde{\mathrm{SL}}_2(E)$. It is known that the character distributions Θ_{τ_1} and Θ_{τ_2} are given by smooth functions at regular semisimple elements. Therefore $\Theta_{\tau_1} - \Theta_{\tau_2}$ is also a smooth function at regular semisimple elements. For $i = 1, 2$, and any element $z \in \widetilde{\{\pm 1\}}$, the character expansion of τ_i in a neighborhood of z is given by the $\omega_{\tau_i}(z)$ multiplied by the character expansion of τ_i in a neighborhood of the identity. Therefore, if we know that $\Theta_{\tau_1} - \Theta_{\tau_2}$ is also smooth in a neighborhood of the identity and both the representations τ_1 and τ_2 have the same central characters then $\Theta_{\tau_1} - \Theta_{\tau_2}$ is a smooth function on the whole of $\widetilde{\mathrm{SL}}_2(E)$.

For any nontrivial additive character ψ' of E, let us assume that τ_1 admits a nonzero Whittaker functional for (N, ψ') if and only if τ_2 does so too. Under this assumption $c_a(\tau_1) = c_a(\tau_2)$ for all $a \in E^\times/E^{\times 2}$. Then we have the following result.

Theorem 4.2. *Let τ_1, τ_2 be two irreducible admissible genuine representations of $\widetilde{\mathrm{SL}}_2(E)$ with the same central characters. For a nontrivial additive character ψ' of E, assume that τ_1 admits a nonzero Whittaker functional with respect to (N, ψ') if and only if τ_2 admits a nonzero Whittaker functional with respect to (N, ψ'). Then $\Theta_{\tau_1} - \Theta_{\tau_2}$ is constant in a neighborhood of identity and hence extends to a smooth function on all of $\widetilde{\mathrm{SL}}_2(E)$.*

Using Theorem 4.2, we prove an extension of a theorem of Casselman and Prasad [Prasad 1992, Theorem 2.7]. From [Patel and Prasad 2016], let us recall the following lemma.

Lemma 4.3. *Let π be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$. Write $\widetilde{\mathrm{GL}}_2(E)_+ = \widetilde{Z} \cdot \widetilde{\mathrm{SL}}_2(E)$. Then there exists an irreducible admissible genuine representation τ of $\widetilde{\mathrm{SL}}_2(E)$ and a genuine character μ of $\widetilde{Z}$ with $\mu|_{\widetilde{\{\pm 1\}}} = \omega_\tau$ and*

$$\pi \cong \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)} \mu\tau.$$

Moreover, we have

$$\pi|_{\widetilde{\mathrm{GL}}_2(E)_+} \cong \bigoplus_{a \in \widetilde{\mathrm{GL}}_2(E)/\widetilde{\mathrm{GL}}_2(E)_+} \mu^a \tau^a.$$

Now we prove the theorem of Casselman and Prasad for the $\widetilde{\mathrm{GL}}_2(E)$.

Theorem 4.4. *Let ψ be a nontrivial character of E. Let π_1 and π_2 be two irreducible admissible genuine representations of $\widetilde{\mathrm{GL}}_2(E)$ with the same central* *characters such that* $(\pi_1)_{N,\psi} \cong (\pi_2)_{N,\psi}$ *as* $\widetilde{Z}$*-modules. Then* $\Theta_{\pi_1} - \Theta_{\pi_2}$*, initially defined on regular semisimple elements of* $\widetilde{\mathrm{GL}}_2(E)$*, extends to a smooth function on all of* $\widetilde{\mathrm{GL}}_2(E)$.

Proof. We know that Θ_{π_1} and Θ_{π_2} are smooth on the set of regular semisimple elements, so is $\Theta_{\pi_1} - \Theta_{\pi_2}$. To prove the smoothness of $\Theta_{\pi_1} - \Theta_{\pi_2}$ on all of $\widetilde{\mathrm{GL}}_2(E)$, we need to prove the smoothness at every point in $\widetilde{Z}$. As $\widetilde{Z}$ is not the center of $\widetilde{\mathrm{GL}}_2(E)$, the smoothness at the identity is not enough to imply the smoothness at every point in $\widetilde{Z}$. Note that $\widetilde{Z}$ is the center of $\widetilde{\mathrm{GL}}_2(E)_+ := \widetilde{Z} \cdot \widetilde{\mathrm{SL}}_2(E)$ and $\widetilde{\mathrm{GL}}_2(E)_+$ is an open and normal subgroup of $\widetilde{\mathrm{GL}}_2(E)$ of index $[E^\times : E^{\times 2}]$.

Using Lemma 4.3, choose irreducible admissible genuine representations τ_1 and τ_2 of $\widetilde{\mathrm{SL}}_2(E)$ and genuine characters μ_1, μ_2 of $\widetilde{Z}$ such that

$$\pi_1 = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu_1\tau_1) \quad \text{and} \quad \pi_2 = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu_2\tau_2). \tag{7}$$

From Lemma 4.3, we have

$$\pi_1|_{\widetilde{\mathrm{GL}}_2(E)_+} = \bigoplus_{a\in E^\times/E^{\times 2}} (\mu_1\tau_1)^a \quad \text{and} \quad \pi_2|_{\widetilde{\mathrm{GL}}_2(E)_+} = \bigoplus_{a\in E^\times/E^{\times 2}} (\mu_2\tau_2)^a. \tag{8}$$

We also know that all the characters μ_1^a for $a \in E^\times/E^{\times 2}$ are distinct. From the identity (8) we find that

$$(\pi_1)_{N(E),\psi} = \bigoplus_{a\in E^\times/E^{\times 2}} \mu_1^a(\tau_1^a)_{N(E),\psi} \text{ and } (\pi_2)_{N(E),\psi} = \bigoplus_{a\in E^\times/E^{\times 2}} \mu_2^a(\tau_2^a)_{N(E),\psi}. \tag{9}$$

Since $(\pi_1)_{N,\psi} \cong (\pi_2)_{N,\psi}$ as $\widetilde{Z}$-modules, in particular, the parts corresponding to μ^a-eigenspaces are isomorphic for all $a \in E^\times/E^{\times 2}$. Therefore $\mu_1 = \mu_2^b$ for some $b \in E^\times/E^{\times 2}$. Since

$$\pi_2 = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu_2\tau_2) = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu_2^b\tau_2^b),$$

by changing τ_2 by τ_2^b, we can assume $\pi_1 = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu\tau_1)$, and $\pi_2 = \mathrm{ind}_{\widetilde{\mathrm{GL}}_2(E)_+}^{\widetilde{\mathrm{GL}}_2(E)}(\mu\tau_2)$. Now $(\pi_1)_{N(E),\psi} \cong (\pi_2)_{N(E),\psi}$ as $\widetilde{Z}$-modules translates into $(\tau_1^a)_{N(E),\psi} \cong (\tau_2^a)_{N(E),\psi}$ for all $a \in E^\times/E^{\times 2}$. Therefore, by Theorem 4.2, $\Theta_{\tau_1^a} - \Theta_{\tau_2^a}$ is constant in a neighborhood of the identity for all $a \in E^\times/E^{\times 2}$.

Let $\Theta_{\rho,g}$ denote the character expansion of an irreducible admissible representation ρ in a neighborhood of the point g. Then

$$\Theta_{\pi_1,\tilde{z}} = \sum_{a\in E^\times/E^{\times 2}} \Theta_{(\mu\tau_1)^a,\,\tilde{z}} = \sum_{a\in E^\times/E^{\times 2}} \mu^a(\tilde{z})\Theta_{\tau_1^a,1}$$

and

$$\Theta_{\pi_2,\tilde{z}} = \sum_{a\in E^\times/E^{\times 2}} \Theta_{(\mu\tau_2)^a,\,\tilde{z}} = \sum_{a\in E^\times/E^{\times 2}} \mu^a(\tilde{z})\Theta_{\tau_2^a,1}.$$

This proves that $\Theta_{\pi_1} - \Theta_{\pi_2}$ is a constant function on regular semisimple points in some neighborhood of $\tilde{z}$ for all $\tilde{z} \in \widetilde{Z} \subset \widetilde{\mathrm{GL}}_2(E)$, and therefore it extends to a smooth function in that neighborhood of $\tilde{z}$. Thus $\Theta_{\pi_1} - \Theta_{\pi_2}$, which is initially defined on regular semisimple elements of $\widetilde{\mathrm{GL}}_2(E)$, extends to a smooth function on all of $\widetilde{\mathrm{GL}}_2(E)$. $\square$

Corollary 4.5. *Let π_1, π_2 be two irreducible admissible genuine representations of $\widetilde{\mathrm{GL}}_2(E)$ with the same central character such that $(\pi_1)_{N,\psi} \cong (\pi_2)_{N,\psi}$ as $\widetilde{Z}$-modules. Let H be a subgroup of $\widetilde{\mathrm{GL}}_2(E)$ that is compact modulo center. Then there exist finite-dimensional representations σ_1, σ_2 of H such that*

$$\pi_1|_H \oplus \sigma_1 \cong \pi_2|_H \oplus \sigma_2.$$

In other words, this corollary says that the virtual representation $(\pi_1 - \pi_2)|_H$ is finite-dimensional and hence the multiplicity of an irreducible representation of H in $(\pi_1 - \pi_2)|_H$ will be finite.

5. Part C of Theorem 1.5

Let π_1 be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$. We take another admissible genuine representation π_1' having the same central character as that of π_1 and satisfying $(\pi_1)_{N(E),\psi} \oplus (\pi_1')_{N(E),\psi} \cong \Omega(\omega_{\pi_1})$ as $\widetilde{Z}$-modules. From Proposition 3.2, if π_1 is a principal series representation then we can take $\pi_1' = 0$. It can be seen that if π_1 is not a principal series representation then $(\pi_1)_{N(E),\psi}$ is a proper $\widetilde{Z}$-submodule of $\Omega(\omega_{\pi_1})$ forcing $\pi_1' \neq 0$. In particular, if π_1 is one of the Jordan–Hölder factors of a reducible principal series representation then one can take π_1' to be the other Jordan–Hölder factor of the principal series representation. It should be noted that for a supercuspidal representation π_1 we do not have any obvious choice for π_1'.

Let π_2 be a supercuspidal representation of $\mathrm{GL}_2(F)$. To prove Theorem 1.5 in this case, we use character theory and deduce the result by using the result of restriction of a principal series representation of $\widetilde{\mathrm{GL}}_2(E)$ which has already been proved in Section 3. We can assume, if necessary after twisting by a character of $F^\times$, that π_2 is a minimal representation. Recall that an irreducible representation π_2 of $\mathrm{GL}_2(F)$ is called minimal if the conductor of π_2 is less than or equal to the conductor of $\pi_2 \otimes \chi$ for any character χ of $F^\times$. By a theorem of Kutzko [1978], a minimal supercuspidal representation π_2 of $\mathrm{GL}_2(F)$ is of the form $\mathrm{ind}_{\mathcal{K}}^{\mathrm{GL}_2(F)}(W_2)$, where W_2 is a representation of a maximal compact modulo center subgroup $\mathcal{K}$ of $\mathrm{GL}_2(F)$. By Frobenius reciprocity,

$$\begin{aligned}\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1 \oplus \pi_1', \pi_2) &= \mathrm{Hom}_{\mathrm{GL}_2(F)}\big(\pi_1 \oplus \pi_1', \mathrm{ind}_{\mathcal{K}}^{\mathrm{GL}_2(F)}(W_2)\big)\\ &= \mathrm{Hom}_{\mathcal{K}}\big((\pi_1 \oplus \pi_1')|_{\mathcal{K}}, W_2\big).\end{aligned}$$

To prove Theorem 1.5, it suffices to prove that

$$\dim \mathrm{Hom}_{\mathcal{K}}((\pi_1 \oplus \pi_1')|_{\mathcal{K}}, W_2) + \dim \mathrm{Hom}_{D_F^\times}(\pi_1 \oplus \pi_1', \pi_2') = [E^\times : F^\times E^{\times 2}].$$

For any (virtual) representation π of $\widetilde{\mathrm{GL}}_2(E)$, let $m(\pi, W_2) = \dim \mathrm{Hom}_{\mathcal{K}}[\pi|_{\mathcal{K}}, W_2]$ and $m(\pi, \pi_2') = \dim \mathrm{Hom}_{D_F^\times}[\pi, \pi_2']$. With these notations we will prove

$$m(\pi_1 \oplus \pi_1', W_2) + m(\pi_1 \oplus \pi_1', \pi_2') = [E^\times : F^\times E^{\times 2}]. \tag{10}$$

Let Ps be an irreducible principal series representation of $\widetilde{\mathrm{GL}}_2(E)$ whose central character ω_{Ps} is the same as the central character ω_{π_1} of π_1 (it is clear that one exists). By Proposition 3.2, we know that $(Ps)_{N(E),\psi} \cong \Omega(\omega_{Ps})$ as a $\widetilde{Z}$-module. On the other hand, the representation π_1' has been chosen in such a way that $(\pi_1)_{N(E),\psi} \oplus (\pi_1')_{N(E),\psi} = \Omega(\omega_{\pi_1})$ as a $\widetilde{Z}$-module. Then, as a $\widetilde{Z}$-module we have

$$(\pi_1 \oplus \pi_1')_{N(E),\psi} = (\pi_1)_{N(E),\psi} \oplus (\pi_1')_{N(E),\psi} = \Omega(\omega_{\pi_1}) = \Omega(\omega_{Ps}) = (Ps)_{N(E),\psi}.$$

We have already proved in Section 3 that

$$m(Ps, W_2) + m(Ps, \pi_2') = [E^\times : F^\times E^{\times 2}].$$

In order to prove equation (10), we prove

$$m(\pi_1 \oplus \pi_1' - Ps, W_2) + m(\pi_1 \oplus \pi_1' - Ps, \pi_2') = 0. \tag{11}$$

The relation in equation (11) follows from the following theorem:

Theorem 5.1. *Let Π_1, Π_2 be two genuine representations of $\widetilde{\mathrm{GL}}_2(E)$ of finite length, having the same central characters, and such that $(\Pi_1)_{N(E),\psi} \cong (\Pi_2)_{N(E),\psi}$ as $\widetilde{Z}$-modules. Let π_2 be an irreducible supercuspidal representation of $\mathrm{GL}_2(F)$ such that the central characters ω_{Π_1} of Π_1 and ω_{π_2} of π_2 agree on $F^\times \cap E^{\times 2}$. Let π_2' be the finite-dimensional representation of $D_F^\times$ associated to π_2 by the Jacquet–Langlands correspondence. Then*

$$m(\Pi_1 - \Pi_2, \pi_2) + m(\Pi_1 - \Pi_2, \pi_2') = 0.$$

We will use character theory to prove this relation, following [Prasad 1992] very closely. First of all, by Theorem 4.4, $\Theta_{\Pi_1 - \Pi_2}$ is given by a smooth function on $\widetilde{\mathrm{GL}}_2(E)$. Now we recall the Weyl integration formula for $\mathrm{GL}_2(F)$.

5A. *Weyl integration formula.*

Lemma 5.2 [Jacquet and Langlands 1970, Formula 7.2.2]. *For a smooth and compactly supported function f on $\mathrm{GL}_2(F)$ we have*

$$\int_{\mathrm{GL}_2(F)} f(y)dy = \sum_{E_i} \int_{E_i} \triangle(x) \left(\frac{1}{2} \int_{E_i \backslash \mathrm{GL}_2(F)} f(\bar{g}^{-1} x \bar{g})\, d\bar{g} \right) dx, \tag{12}$$

where the E_i are representatives for the distinct conjugacy classes of maximal tori in $\mathrm{GL}_2(F)$ and

$$\triangle(x) = \left\| \frac{(x_1 - x_2)^2}{x_1 x_2} \right\|_F$$

where x_1 and x_2 are the eigenvalues of x.

We will use this formula to integrate the function $f(x) = \Theta_{\Pi_1-\Pi_2} \cdot \Theta_{W_2}(x)$ on $\mathcal{K}$ which is extended to $\mathrm{GL}_2(F)$ by setting it to be zero outside $\mathcal{K}$. In addition, we also need the following result of Harish-Chandra, cf. [Prasad 1992, Proposition 4.3.2].

Lemma 5.3 (Harish-Chandra). *Let $F(g) = (gv, v)$ be a matrix coefficient of a supercuspidal representation π of a reductive p-adic group G with center Z. Then the orbital integrals of F at regular nonelliptic elements vanish. Moreover, the orbital integral of Fat a regular elliptic element x contained in a torus T is given by the formula*

$$\int_{T\backslash G} F(\bar{g}^{-1}x\bar{g})d\bar{g} = \frac{(v, v)\cdot\Theta_\pi(x)}{d(\pi)\cdot \mathrm{vol}(T/Z)}, \tag{13}$$

where $d(\pi)$ denotes the formal degree of the representation π (which depends on a choice of Haar measure on $T\backslash G$).

Since π_2 is obtained by induction from W_2, a matrix coefficient of W_2 (extended to $\mathrm{GL}_2(F)$ by setting it to be zero outside $\mathcal{K}$) is also a matrix coefficient of π_2. It follows that

(1) for the choice of Haar measure on $\mathrm{GL}_2(F)/F^\times$ giving $\mathcal{K}/F^\times$ measure 1,

$$\dim W_2 = d(\pi_2),$$

(2) for a separable quadratic field extension E_i of F and a regular elliptic element x of $\mathrm{GL}_2(E)$ which generates E_i, and for the above Haar measure $d\bar{g}$,

$$\int_{E_i^\times\backslash \mathrm{GL}_2(F)} \Theta_{W_2}(\bar{g}^{-1}x\bar{g})d\bar{g} = \frac{\Theta_{\pi_2}(x)}{\mathrm{vol}(E_i^\times/F^\times)}. \tag{14}$$

Equation (14) can be explained as follows. Let the dimension of W_2 be n and let $\{e_1, \dots, e_n\}$ be an orthonormal basis of W_2. For $g \in \mathcal{K}$ the map $g \mapsto F_i(g) := \langle ge_i, e_i\rangle$ defines a matrix coefficient of W_2 for all $i = 1, \dots, n$. Then $\Theta_{W_2}(g) = \sum_{i=1}^n F_i(g)$. Now consider all these F_i as matrix coefficients of π_2. Apply Lemma 5.3 for $F = F_i$ and sum up over all $i = 1, \dots, n$ then we get equation (14), since $d(\pi_2) = \dim W_2 = n$.

5B. ***Completion of the proof of Theorem 1.5.*** We recall the following important observations from Section 5A and Theorem 4.4:

(1) The virtual representation $(\Pi_1 - \Pi_2)|_{\mathcal{K}}$ is finite-dimensional.

(2) Θ_{W_2} is a matrix coefficient of π_2 (extended to $\mathrm{GL}_2(F)$ by zero outside $\mathcal{K}$).

(3) There is Haar measure on $\mathrm{GL}_2(F)/F^\times$ giving $\mathrm{vol}(\mathcal{K}/F^\times) = 1$ such that the (14) is satisfied.

(4) The orbital integral in equation (13) vanishes if T is the maximal split torus.

Let the E_i's be the quadratic extensions of F. Then these observations together with Lemma 5.3 imply the following

$$
\begin{aligned}
&m(\Pi_1-\Pi_2,W_2)\\
&\qquad=\frac{1}{\operatorname{vol}(\mathcal{K}/F^\times)}\int_{\mathcal{K}/F^\times}\Theta_{\Pi_1-\Pi_2}\cdot\Theta_{W_2}(x)\,dx\\
&\qquad=\frac{1}{\operatorname{vol}(\mathcal{K}/F^\times)}\int_{\mathrm{GL}_2(F)/F^\times}\Theta_{\Pi_1-\Pi_2}\cdot\Theta_{W_2}(x)\,dx\\
&\qquad=\frac{1}{\operatorname{vol}(\mathcal{K}/F^\times)}\sum_{E_i}\int_{E_i^\times/F^\times}\triangle(x)\Bigl[\frac{1}{2}\int_{E_i^\times\backslash\mathrm{GL}_2(F)}\Theta_{\Pi_1-\Pi_2}\cdot\Theta_{W_2}(\bar g^{-1}x\bar g)\,d\bar g\Bigr]\,dx\\
&\qquad=\sum_{E_i}\frac{1}{2\operatorname{vol}(E_i^\times/F^\times)}\int_{E_i^\times/F^\times}(\triangle\cdot\Theta_{\Pi_1-\Pi_2}\cdot\Theta_{\pi_2})(x)\,dx.
\end{aligned}
$$

Similarly, we have the equality

$$
m(\Pi_1-\Pi_2,\pi_2')=\sum_{E_i}\frac{1}{2\operatorname{vol}(E_i^\times/F^\times)}\int_{E_i^\times/F^\times}(\triangle\cdot\Theta_{\Pi_1-\Pi_2}\cdot\Theta_{\pi_2'})(x)dx.
$$

Note that the E_i's correspond to quadratic extensions of F and the embeddings of $\mathrm{GL}_2(F)$ and $D_F^\times$ have been fixed so that Working Hypothesis 1.3 (as stated in the introduction) is satisfied, i.e., the embeddings of the E_i in $\mathrm{GL}_2(F)$ and in $D_F^\times$ are conjugate in $\widetilde{\mathrm{GL}}_2(E)$. Then the value of $\Theta_{\Pi_1-\Pi_2}(x)$ for $x\in E_i$, does not depend on the inclusion of E_i inside $\widetilde{\mathrm{GL}}_2(E)$, i.e., on whether the inclusion is via $\mathrm{GL}_2(F)$ or via $D_F^\times$. Now using the relation $\Theta_{\pi_2}(x)=-\Theta_{\pi_2'}(x)$ on regular elliptic elements x [Jacquet and Langlands 1970, Proposition 15.5], we conclude the following, which proves equation (11):

$$
m(\Pi_1-\Pi_2,W_2)+m(\Pi_1-\Pi_2,\pi_2')=0.
$$

6. A remark on higher multiplicity

We have shown that the restriction of an irreducible admissible representation of $\widetilde{\mathrm{GL}}_2(E)$, for example a principal series representation, to the subgroup $\mathrm{GL}_2(F)$ has multiplicity more than one. Given the important role multiplicity one theorems play, it would be desirable to modify the situation so that multiplicity one might be true. One natural way to do this is to decrease the larger group, and increase the smaller group. In this section we discuss some natural subgroups of the group $\widetilde{\mathrm{GL}}_2(E)$ which can be used, but unfortunately, it still does not help one to achieve multiplicity one situation. We discuss this modification in this section in some detail.

Let us take the subgroup of $\widetilde{\mathrm{GL}}_2(E)$ which is generated by $\mathrm{GL}_2(F)$ and $\widetilde{Z}$. We will prove that this subgroup also fails to achieve multiplicity one for the restriction problem from $\widetilde{\mathrm{GL}}_2(E)$ to $\mathrm{GL}_2(F)\cdot\widetilde{Z}$. Let

$$H=\mathrm{GL}_2(F)\subset H_+=Z\cdot\mathrm{GL}_2(F)\subset\mathrm{GL}_2(E).$$

We will show that the restriction of an irreducible admissible representation of $\widetilde{\mathrm{GL}}_2(E)$ to the subgroup $\widetilde{H}_+$ has higher multiplicity. Note that the subgroups $\widetilde{Z}$ and $\mathrm{GL}_2(F)$ do not commute but $\widetilde{Z}^2$ commutes with $\mathrm{GL}_2(F)$. In fact, the commutator relation is given by

$$[\tilde{e},\tilde{g}]=(e,\det g)_E\in\{\pm 1\}\subset\widetilde{\mathrm{GL}}_2(E), \tag{15}$$

where $\tilde{e}\in\widetilde{Z}$ and $\tilde{g}\in\widetilde{\mathrm{GL}}_2(F)$ lie over elements $e\in Z$ and $g\in\mathrm{GL}_2(F)$ respectively, and $(-,-)_E$ denotes the Hilbert symbol for the field E. The lemma below proves that the center of $\widetilde{H}_+$ is $\widetilde{Z^2F^\times}$.

Lemma 6.1. *For an element $e\in E^\times$, the map $F^\times\to\{\pm1\}$ defined by $f\mapsto(e,f)_E$ is trivial if and only if $e\in F^\times E^{\times 2}$.*

Proof. Let $(\cdot,\cdot)_E$ and $(\cdot,\cdot)_F$ denote the Hilbert symbol of the field E and F respectively. For $e\in E^\times$ and $f\in F^\times$, the following is well known [Bender 1973]:

$$(e,f)_E=(N_{E/F}(e),f)_F,$$

where $N_{E/F}$ is the norm map of the extension E/F. Therefore, if $(e,f)_E=1$ is true for all $f\in F^\times$, then by the nondegeneracy of the Hilbert symbol $(\cdot,\cdot)_F$ one will have $N_{E/F}(e)\in F^{\times 2}$. The inverse image of $F^{\times 2}$ under the norm map $N_{E/F}$ is now seen to be $E^{\times 2}F^\times$ since this subgroup surjects onto $F^{\times 2}$ under the norm mapping, and contains the kernel $\{z/\bar{z}=z^2/z\bar{z}:z\in E^\times\}$ of $N_{E/F}$. □

Let σ be an irreducible admissible representation of $\mathrm{GL}_2(F)$. For any character χ of $F^\times$ let us abuse the notation and simply write $\sigma\otimes\chi$ for $\sigma\otimes(\chi\circ\det)$. By the commutator relation (15), for $a\in Z$ and $g\in\mathrm{GL}_2(F)$ we have

$$a(g,\epsilon)a^{-1}=(g,\chi_a(\det g)\epsilon),$$

where χ_a is given by $x\mapsto(x,a)_E$ for all $x\in E^\times$. Therefore, the conjugation action by $a\in Z$ takes σ to the quadratic twist $\sigma\otimes\chi_a$. We have the following lemma which easily follows from Clifford theory.

Lemma 6.2. *Let $\widetilde{H}_0=\widetilde{Z}^2\cdot\mathrm{GL}_2(F)$. Let σ be an irreducible admissible representation of $\mathrm{GL}_2(F)$. Assume that $\sigma\otimes\chi_a\not\cong\sigma$ for any nontrivial element $a\in E^\times/F^\times E^{\times 2}$. Fix a genuine character η of $\widetilde{Z}^2$ such that $\eta|_{F^\times\cap\widetilde{Z}^2}=\omega_\sigma|_{F^\times\cap\widetilde{Z}^2}$. Then $\rho=\mathrm{Ind}_{\widetilde{H}_0}^{\widetilde{H}_+}(\eta\sigma)$ is an irreducible representation of $\widetilde{H}_+$. The representation ρ is the only irreducible representation of $\widetilde{H}_+$ whose central character restricted* *to $\widetilde{Z}^2$ is η and also contains σ. Moreover, $\rho|_{\widetilde{H}_0} \cong \bigoplus_{a \in E^\times/F^\times E^{\times 2}} \eta(\sigma \otimes \chi_a)$. In particular, from Lemma 6.1, the restriction of ρ to $\widetilde{H}_0$ is multiplicity free.*

Note that if σ is a principal series representation of $\mathrm{GL}_2(F)$ which is not of the form $Ps(\chi_1, \chi_2)$ with χ_1/χ_2 a quadratic character, then such principal series representation of $\mathrm{GL}_2(F)$ have no nontrivial self twist, i.e., for any character χ of $F^\times$ the relation $Ps(\chi_1, \chi_2) \otimes (\chi \circ \det) \cong Ps(\chi_1, \chi_2)$ implies that χ is trivial. Let π be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ such that $\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi, \sigma) \geq 2$. Let η be the central character of π. Note that the central character of any irreducible representation of $\widetilde{H}_+$, which is contained in π, agrees with η when restricted to $\widetilde{Z}^2$. As in the previous lemma, we let $\rho = \mathrm{Ind}_{\widetilde{H}_0}^{\widetilde{H}_+}(\eta\sigma)$. The representation ρ is the only representation of $\widetilde{H}_+$ which appears in π and contains σ. So the multiplicity of such a principal series representation σ of $\mathrm{GL}_2(F)$ in the restriction of an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ is the same as the multiplicity of the corresponding irreducible representation of $\widetilde{H}_+$, i.e., $\dim \mathrm{Hom}_{\widetilde{H}_+}(\pi, \rho) = \dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi, \sigma) \geq 2$. Thus we conclude that the restriction of representations of $\widetilde{\mathrm{GL}}_2(E)$ to $\widetilde{H}_+$ has higher multiplicity.

On the other hand, let us take the group $G = \{g \in \mathrm{GL}_2(E) : \det g \in F^\times E^{\times 2}\}$. Note that this subgroup G contains $\mathrm{GL}_2(E)_+ = Z \cdot \mathrm{SL}_2(E)$. We will prove that the pair $(\widetilde{G}, \mathrm{GL}_2(F))$ also fails to achieve multiplicity one for the restriction problem from $\widetilde{G}$ to $\mathrm{GL}_2(F)$. From the commutation relation (15), it follows that the center of the group $\widetilde{G}$ is $\widetilde{F^\times Z^2}$. Recall that the restriction from $\widetilde{\mathrm{GL}}_2(E)$ to $\widetilde{\mathrm{GL}}_2(E)_+$ is multiplicity free and $\widetilde{G} \supset \widetilde{\mathrm{GL}}_2(E)_+$, thus the restriction from $\widetilde{\mathrm{GL}}_2(E)$ to $\widetilde{G}$ is also multiplicity free. Let π be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and ρ be an irreducible admissible genuine representation of $\widetilde{G}$ such that $\rho \hookrightarrow \pi|_{\widetilde{G}}$. Then we have

$$\pi|_{\widetilde{G}} = \bigoplus_{a \in E^\times/F^\times E^{\times 2}} \rho^a.$$

For $a_1 \neq a_2$ in $E^\times/F^\times E^{\times 2}$, $\rho^{a_1} \ncong \rho^{a_2}$. In fact, the central characters of ρ^{a_1} and ρ^{a_2} are different when restricted to $F^\times$.

Let π be an irreducible admissible genuine representation of $\widetilde{\mathrm{GL}}_2(E)$ and σ an irreducible admissible representation of $\mathrm{GL}_2(F)$ such that

$$\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi, \sigma) \geq 2.$$

If $\mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho^{a_1}, \sigma) \neq 0$ then $\mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho^{a_2}, \sigma) = 0$ for $a_2 \neq a_1$ in $E^\times/F^\times E^{\times 2}$, since the central character of ρ^{a_2} restricted to $F^\times$ will be different from the central character of σ. Thus there exists only one $a \in E^\times/F^\times E^{\times 2}$ such that $\mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho^a, \sigma) \neq 0$. We can assume that $\mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho, \sigma) \neq 0$. We have

$$\mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho, \sigma) = \mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi, \sigma)$$

and hence $\dim \mathrm{Hom}_{\mathrm{GL}_2(F)}(\rho, \sigma) \geq 2$.

Acknowledgements. The author would like to express his indebtedness to Professor Dipendra Prasad. Without his continuous encouragement and help this paper would not have appeared in the present form. The author is thankful to Professor Sandeep Varma for several helpful discussions, and the anonymous referee for several helpful comments.

References

[Bender 1973] E. A. Bender, "A lifting formula for the Hilbert symbol", *Proc. Amer. Math. Soc.* **40** (1973), 63–65. MR Zbl

[Bernstein and Zelevinskii 1976] I. N. Bernstein and A. V. Zelevinskii, "Representations of the group $GL(n, F)$, where F is a local non-Archimedean field", *Uspehi Mat. Nauk* **31**:3 (1976), 5–70. In Russian; translated in *Russian Math. Surv.* **31**:3 (1976), 1–68. MR Zbl

[Gelbart and Piatetski-Shapiro 1980] S. Gelbart and I. I. Piatetski-Shapiro, "Distinguished representations and modular forms of half-integral weight", *Invent. Math.* **59**:2 (1980), 145–188. MR Zbl

[Gelbart et al. 1979] S. Gelbart, R. Howe, and I. Piatetski-Shapiro, "Uniqueness and existence of Whittaker models for the metaplectic group", *Israel J. Math.* **34**:1-2 (1979), 21–37. MR Zbl

[Godement 1970] R. Godement, *Notes on Jacquet–Langlands' theory*, Inst. Advanced Study, Princeton, 1970.

[Gross and Prasad 1992] B. H. Gross and D. Prasad, "On the decomposition of a representation of SO_n when restricted to SO_{n-1}", *Canad. J. Math.* **44**:5 (1992), 974–1002. MR Zbl

[Jacquet and Langlands 1970] H. Jacquet and R. P. Langlands, *Automorphic forms on* GL(2), Lecture Notes in Math. **114**, Springer, Berlin, 1970. MR Zbl

[Kubota 1969] T. Kubota, *On automorphic functions and the reciprocity law in a number field*, Lectures in Math., Kyoto Univ. **2**, Kinokuniya, Tokyo, 1969. MR Zbl

[Kutzko 1978] P. C. Kutzko, "On the supercuspidal representations of GL_2", *Amer. J. Math.* **100**:1 (1978), 43–60. MR Zbl

[Mœglin and Waldspurger 2012] C. Mœglin and J.-L. Waldspurger, "La conjecture locale de Gross–Prasad pour les groupes spéciaux orthogonaux: le cas général", pp. 167–216 in *Sur les conjectures de Gross et Prasad, II*, Astérisque **347**, Société Mathématique de France, Paris, 2012. MR Zbl

[Patel 2015] S. P. Patel, "A theorem of Mœglin and Waldspurger for covering groups", *Pacific J. Math.* **273**:1 (2015), 225–239. MR Zbl

[Patel 2016] S. P. Patel, "A question on splitting of metaplectic covers", *Comm. Algebra* **44**:12 (2016), 5095–5104. MR Zbl

[Patel and Prasad 2016] S. P. Patel and D. Prasad, "Multiplicity formula for restriction of representations of $\widetilde{GL_2}(F)$ to $\widetilde{SL_2}(F)$", *Proc. Amer. Math. Soc.* **144**:2 (2016), 903–908. MR Zbl

[Patel and Prasad 2017] S. P. Patel and D. Prasad, "Restriction of representations of metaplectic GL_2 to tori", preprint, 2017. To appear in *Israel J. Math.* arXiv

[Prasad 1990] D. Prasad, "Trilinear forms for representations of GL(2) and local ϵ-factors", *Compositio Math.* **75**:1 (1990), 1–46. MR Zbl

[Prasad 1992] D. Prasad, "Invariant forms for representations of GL_2 over a local field", *Amer. J. Math.* **114**:6 (1992), 1317–1363. MR Zbl

[Prasad and Raghuram 2000] D. Prasad and A. Raghuram, "Kirillov theory for $GL_2(\mathcal{D})$ where $\mathcal{D}$ is a division algebra over a non-Archimedean local field", *Duke Math. J.* **104**:1 (2000), 19–44. MR Zbl

[Saito 1993] H. Saito, "On Tunnell's formula for characters of GL(2)", *Compositio Math.* **85**:1 (1993), 99–108. MR Zbl

[Tunnell 1983] J. B. Tunnell, "Local ϵ-factors and characters of GL(2)", *Amer. J. Math.* **105**:6 (1983), 1277–1307. MR Zbl

Received September 21, 2016. Revised May 9, 2017.

SHIV PRAKASH PATEL
DEPARTMENT OF MATHEMATICS
BEN-GURION UNIVERSITY OF THE NEGEV
P.O.B. 653
8410501
ISRAEL
shivprakashpatel@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 291, No. 2, 2017
dx.doi.org/10.2140/pjm.2017.291.485

HESSIAN EQUATIONS ON CLOSED HERMITIAN MANIFOLDS

DEKAI ZHANG

We solve the complex Hessian equation on closed Hermitian manifolds, which generalizes the Kähler case proven by Hou, Ma and Wu and Dinew and Kołodziej. Solving the equation can be reduced to the derivation of a priori second-order estimates. We introduce a new method to prove the C^0 estimate. The C^2 estimate can be derived if we use the auxiliary function which is mainly due to Hou, Ma and Wu and Tosatti and Weinkove.

1. Introduction

Let (M, ω) be a closed Hermitian manifold of complex dimension $n \geq 2$. In this paper, we study the Hessian equation

$$(1\text{-}1)\qquad \begin{cases} \binom{n}{k}\omega_u^k \wedge \omega^{n-k} = e^f \omega^n, \\ \sup_M u = 0, \\ \omega_u = \omega + \sqrt{-1}\partial\bar{\partial}u \in \Gamma_k(M), \end{cases}$$

where $\binom{n}{k} = n!/(k!(n-k)!)$, $\Gamma_k(M)$ is a convex cone (see (2-2) in Section 2) and $1 \leq k \leq n$.

The complex Hessian equation is an important class of fully nonlinear elliptic equations. It arises naturally from many significant geometric problems. When $k = 1$, it is the classical Laplacian equation. For $k = n$, equation (1-1) is the complex Monge–Ampère equation

$$(1\text{-}2)\qquad \omega_u^n = e^f \omega^n, \quad \sup_M u = 0.$$

Yau [1978] solved equation (1-2) on compact Kähler manifolds, and his solution is now known as Calabi–Yau theorem. For general Hermitian manifolds, (1-2) has been solved by Cherrier [1987] for dimension 2. Guan and Li [2010] and Zhang [2010] obtained C^1 and C^2 estimates for dimension $n \geq 2$. Finally, Tosatti and Weinkove [2010] derived the C^0 estimate and thus solved (1-2) for arbitrary dimension.

While $1 < k < n$, equation (1-1) has more complicated structure and also is closely related to many important geometric problems. For example, for $k = 2$, it

MSC2010: 35J60, 53C55.
Keywords: Hessian equations, closed Hermitian manifolds, a priori estimates.

relates to Fu and Yau's [2008] generalization of the Strominger system which comes from superstring theory. Several significant results about the Fu–Yau equation have been obtained by Phong, Picard and Zhang [Phong et al. 2016a; 2016b; 2017]. When $k = n-1$, it has similar features to the Monge–Ampère type equation in the study of Gauduchon conjecture by Tosatti and Weinkove [2015; 2017] and Tosatti, Weinkove and Székelyhidi [Székelyhidi et al. 2015].

We now come back to the complex Hessian equation. To solve it, it is crucial to derive the *a priori* estimates up to second-order. If (M, ω) is a Kähler manifold, Hou, Ma and Wu [Hou et al. 2010] proved

$$\max |\partial\bar{\partial}u|_g \leq C(1+\max |\nabla u|_g^2), \tag{1-3}$$

where C does not depend on the gradient bound of the solution.

They also pointed out that (1-3) may be adapted to the blow up analysis to get the gradient estimate. Later on, combining (1-3) with a blow up argument, Dinew and Kołodziej [2017] obtained the gradient estimate. Then equation (1-1) can be solved on Kähler manifolds.

In this paper, we solve the complex Hessian equation on closed Hermitian manifolds. More precisely,

Theorem 1.1. *Let (M, g) be a closed Hermitian manifold of complex dimension $n \geq 2$ and f be a smooth real function on M. Then there exist a unique real number b and a unique smooth real function u on M solving*

$$\binom{n}{k}\omega_u^k \wedge \omega^{n-k} = e^{f+b}\omega^n, \quad \omega_u \in \Gamma_k(M), \quad \sup_M u = 0. \tag{1-4}$$

We use the continuity method to solve problem (1-4). The openness follows from implicit function theory. The closeness argument can be reduced to *a priori* estimates up to the second order by the standard Evans–Krylov theory. Actually, we can derive the zero-order estimate and the second-order estimate of solutions of equation (1-1) and thus use the blow up method to obtain the gradient estimate.

For the complex Monge–Ampère equation on closed Hermitian manifolds, Tosatti and Weinkove [2010] derived C^0 estimate by proving a Cherrier-type inequality which was originally proved in [Cherrier 1987]. For the Hessian equation (1-1), we can prove a similar Cherrier-type inequality by a new method which combines an inductive argument with key inequalities for k-th elementary symmetric functions in [Chou and Wang 2001]. For the C^2 estimate, the main difficulty is that there are new terms of the form $T * D^3u$, where T is the torsion of ω. To control these terms, we use the auxiliary function due to Tosatti and Weinkove [2013]. The auxiliary function originally comes from Hou, Ma and Wu [Hou et al. 2010]. For the Hessian equation, the main difference is that for equation (1-1) we need to apply some lemmas for the k-th elementary symmetric functions which were proved by Hou, Ma and Wu [Hou et al. 2010].

The rest of the paper is organized as follows. In Section 2, we give some preliminaries. In Section 3, a Cherrier-type inequality is derived, and then we obtain the C^0 estimate. In Section 4, we prove the C^2 estimate by a similar auxiliary function used in [Tosatti and Weinkove 2013].

Székelyhidi [2015] has also obtained similar results, but our methods are different.

2. Preliminaries

Let (M, g) be a closed Hermitian manifold and let ∇ denote the Chern connection of g. In this section we give some preliminaries about the k-th elementary symmetric function and the commutation formula of covariant derivatives.

Elementary symmetric functions. The k-th elementary symmetric function is defined by

$$\sigma_k(\lambda) = \sum_{1 \le i_1 < \cdots < i_k \le n} \lambda_{i_1} \cdots \lambda_{i_k},$$

where $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$. Let $\lambda(a_{i\bar{j}})$ denote the eigenvalues of the Hermitian matrix $\{a_{i\bar{j}}\}$; we define

$$\sigma_k(a_{i\bar{j}}) = \sigma_k(\lambda\{a_{i\bar{j}}\}).$$

The definition of σ_k can be naturally extended to a Hermitian manifold. Indeed, let $A^{1,1}(M, \mathbb{R})$ be the space of smooth real $(1, 1)$-forms on M; for $\chi \in A^{1,1}(M, \mathbb{R})$ we define

$$\sigma_k(\chi) = \binom{n}{k} \frac{\chi^k \wedge \omega^{n-k}}{\omega^n}.$$

Definition 2.1.

$$\Gamma_k := \{\lambda \in \mathbb{R}^n : \sigma_j(\lambda) > 0,\ j = 1, \ldots, k\}. \tag{2-1}$$

Similarly, we define Γ_k on M as follows

$$\Gamma_k(M) := \{\chi \in A^{1,1}(M, \mathbb{R}) : \sigma_j(\chi) > 0,\ j = 1, \ldots, k\}. \tag{2-2}$$

Furthermore, $\sigma_r(\lambda | i_1 \cdots i_l)$, with $i_1, \ldots, i_l$ being distinct, denotes the r-th symmetric function with $\lambda_{i_1} = \cdots = \lambda_{i_l} = 0$. For more details about elementary symmetric functions, one can see the lecture notes [Wang 2009].

To prove the C^0 estimate, we need the following lemma for elementary symmetric functions:

Lemma 2.2. *Suppose that $\lambda \in \Gamma_k$, $3 \le k \le n$ and $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. Then there exists a positive constant C depending only on k and n, such that for $1 \le i \le k-2$,*

$$|\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}| \le (n-3)^{k-2}\sigma_i(\lambda|j), \quad 1 \le j_1 < j_2 < \cdots < j_i \le n, \quad j_l \ne j, 1 \le l \le i, 1 \le j \le n. \tag{2-3}$$

Proof. Since

$$\sum_{p=k}^{n} \lambda_p = \sigma_1(\lambda|12\cdots k-1) > 0, \quad \text{and} \quad \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n,$$

then

$$|\lambda_p| \le (n-k)\lambda_k, k+1 \le p \le n. \tag{2-4}$$

We first prove the lemma for $k=3$. In this case, one needs to prove

$$|\lambda_l| \le C\sigma_1(\lambda|j) \quad \text{for } 1 \le j, l \le n \text{ and } l \ne j.$$

Since $\sigma_1(\lambda|j) = \lambda_l + \sigma_1(\lambda|jl)$, $\lambda_l \le \sigma_1(\lambda|j)$. Now, if $\lambda_l < 0$, then $l \ge 4$. By (2-4),

$$|\lambda_l| \le (n-3)\lambda_3 \le \sigma_1(\lambda|j), 4 \le l \le n.$$

Then the lemma follows for $k=3$.

Next we prove the lemma for the general k, $3 \le k \le n$.

If $j > i$, since $i \le k-2$, $\lambda|j \in \Gamma_{i+1}$, applying [Lin and Trudinger 1994, p. 322, (19)] yields $\sigma_i(\lambda|j) \ge \lambda_1\cdots\lambda_i$. Since $1 \le l \le i \le k-2$, by (2-4) we have

$$|\lambda_{j_l}| \le \max\{\lambda_l, (n-k)\lambda_k\} \le (n-k)\lambda_l.$$

Then

$$|\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}| \le (n-k)^i\lambda_1\cdots\lambda_i \le (n-k)^{k-2}\sigma_i(\lambda|j). \tag{2-5}$$

If $j \le i$, applying [Lin and Trudinger 1994, p. 322, (19)] yields

$$\sigma_i(\lambda|j) \ge \lambda_1\cdots\lambda_{j-1}\lambda_{j+1}\cdots\lambda_{i+1}.$$

Note $j_l \ne j$, so

$$|\lambda_{j_l}| \le \begin{cases} (n-3)\lambda_l, & j_l < j, \\ (n-3)\lambda_{l+1}, & j_l > j. \end{cases}$$

Therefore, we have

$$|\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}| \le (n-k)^i\lambda_1\cdots\lambda_{j-1}\lambda_{j+1}\cdots\lambda_{i+1} \le (n-k)^{k-2}\sigma_i(\lambda|j). \tag{2-6}$$

Combining (2-5) and (2-6), we obtain

$$|\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}| \le (n-3)^{k-2}\sigma_i(\lambda|j), \quad 1 \le j_1 < j_2 < \cdots < j_i \le n, \quad j_l \ne j, \quad 1 \le l \le i, \quad 1 \le j \le n. \qquad \square$$

By Lemma 2.2, we immediately obtain the following lemma which is a key ingredient in proving Lemma 3.2:

Lemma 2.3. *There exists a positive constant C depending only on (M, ω) and n such that*

$$\left|\frac{\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i-1}\wedge T_i}{\omega^n}\right|\le C\frac{\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i}\wedge\omega^{n-i-1}}{\omega^n}, \tag{2-7}$$

where T_i is defined as the combinations of ω, $\partial\omega$, $\partial\bar\partial\omega$; more precisely,

$$T_i=\sum_{0\le 3p+2q\le n-i}\omega^{n-i-3p-2q}\wedge(\sqrt{-1})^p(\partial\omega)^p\wedge(\bar\partial\omega)^p\wedge(\sqrt{-1})^q(\partial\bar\partial\omega)^q$$

for $1\le i\le k-1$.

Proof. For $x\in M$, we choose the coordinates such that

$$\omega(x)=\sqrt{-1}\sum_{j=1}^n dz^j\wedge d\bar z^j,\quad \omega_u(x)=\sqrt{-1}\sum_{j=1}^n\lambda_j dz^j\wedge d\bar z^j,$$

and $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_n$. Write T_i as follows:

$$T_i=(\sqrt{-1})^{n-i}(T_i)_{l_1\cdots l_{n-i},\bar m_1,\cdots\bar m_{n-i}}dz^{l_1}\wedge\cdots\wedge dz^{l_{n-i}}\wedge d\bar z^{m_1}\wedge\cdots\wedge d\bar z^{m_{n-i}}.$$

Then

$$\begin{aligned}\left|\frac{\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i}\wedge T_i}{\omega^n}\right|&\le C\sum_{j,l=1}^n\sum_{1\le j_1<\cdots<j_i\le n,\neq j,l}|u_j||u_{\bar l}||\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}|\\&\le C\sum_{j=1}^n\sum_{\substack{1\le j_1<\cdots<j_i\le n\\ j_l\neq j,1\le l\le i}}|u_j|^2|\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i}|\\&\le C\sum_{j=1}^n\sigma_i(\lambda|j)|u_j|^2\\&=C\frac{\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i}\wedge\omega^{n-i-1}}{\omega^n},\end{aligned}\tag{2-8}$$

where we have used Lemma 2.2 in the last inequality and C depends on the bound of the torsion and the curvature of (M, ω). □

Commutation formula of covariant derivatives. We have, in local complex coordinates $z_1,\dots,z_n$,

$$g_{i\bar j}=g\Big(\frac{\partial}{\partial z^i},\frac{\partial}{\partial\bar z^j}\Big),\quad \{g^{i\bar j}\}=\{g_{i\bar j}\}^{-1}\tag{2-9}$$

For the Chern connection ∇, we denote the covariant derivatives

$$u_i=\nabla_{\partial/\partial z^i}u,\quad u_{i\bar j}=\nabla_{\partial/\partial\bar z^j}\nabla_{\partial/\partial z^i}u,\quad u_{i\bar jk}=\nabla_{\partial/\partial z^k}\nabla_{\partial/\partial\bar z^j}\nabla_{\partial/\partial z^i}u.\tag{2-10}$$

We use commutation formulas for covariant derivatives on Hermitian manifolds which can be found in [Tosatti and Weinkove 2013]:

$$u_{i\bar{j}l} = u_{l\bar{j}i} - T_{li}^{p} u_{p\bar{j}}, \quad u_{pi\bar{j}} = u_{p\bar{j}i} + u_q R_{i\bar{j}p}{}^{q}, \quad u_{i\bar{p}\bar{j}} = u_{i\bar{j}\bar{p}} - \overline{T_{jp}^{q}} u_{i\bar{q}}. \tag{2-11}$$

$$u_{i\bar{j}l\bar{m}} = u_{l\bar{m}i\bar{j}} + u_{p\bar{j}} R_{l\bar{m}i}{}^{p} - u_{p\bar{m}} R_{i\bar{j}l}{}^{p} - T_{li}^{p} u_{p\bar{m}\bar{j}} - \overline{T_{mj}^{p}} u_{l\bar{p}i} - T_{li}^{p} \overline{T_{mj}^{q}} u_{p\bar{q}}. \tag{2-12}$$

3. Zero-order estimate

In this section we derive the zero-order estimate by proving a Cherrier-type inequality and the lemmas in [Tosatti and Weinkove 2010]. Since the constant b in Theorem 1.1 satisfies

$$|b| \le \sup |f| + C,$$

where C is a positive constant depending only on (M, ω). Thus, we will assume $b = 0$ for convenience.

Theorem 3.1. *Let u be a solution of Theorem 1.1. Then there exists a constant C depending only on (M, ω), n, k and $\sup_M |f|$ such that*

$$\sup_M |u| \le C.$$

Due to Tosatti and Weinkove's results, finding the zero-order estimate can be reduced to deriving a Cherrier-type inequality which was first proved by Cherrier [1987]. For the Hessian equation, we use a new method which combines an inductive argument with the key Lemma 2.3. Even for the Monge–Ampère equation, our proof is different from that in [Tosatti and Weinkove 2010].

Lemma 3.2. *There exist constants p_0 and C depending only on (M, ω), n, k and $\sup_M |f|$ such that for any $p \ge p_0$*

$$\int_M |\partial e^{-(p/2)u}|_g^2 \omega^n \le Cp \int_M e^{-pu} \omega^n.$$

Remark 3.3. Recently, applying our key Lemma 2.2, Sun [2017] also proved the lemma above.

Proof. By the equation, we have

$$\omega_u^k \wedge \omega^{n-k} - \omega^n = \left(\frac{e^f}{\binom{n}{k}} - 1 \right) \omega^n \le C_0 \omega^n,$$

where C_0 is a constant depending only on $\sup f$, n and k. On the other hand,

$$\omega_u^k \wedge \omega^{n-k} - \omega^n = (\omega_u^k - \omega^k) \wedge \omega^{n-k} = \sqrt{-1}\partial\bar{\partial}u \wedge \alpha, \tag{3-1}$$

where $\alpha = \sum_{i=1}^{k} \omega_u^{i-1} \wedge \omega^{n-i}$.

Now multiply both sides in (3-1) by e^{-pu} and integrate by parts:

$$
\begin{aligned}
\text{(3-2)} \qquad C_0\int_M e^{-pu}\omega^n &\geq \int_M e^{-pu}\sqrt{-1}\partial\bar{\partial}u\wedge\alpha \\
&= -\int_M \partial e^{-pu}\sqrt{-1}\bar{\partial}u\wedge\alpha+\int_M e^{-pu}\sqrt{-1}\bar{\partial}u\wedge\partial\alpha \\
&= p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar{\partial}u\wedge\alpha-\frac{1}{p}\int_M\sqrt{-1}\bar{\partial}e^{-pu}\wedge\partial\alpha \\
&= p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar{\partial}u\wedge\alpha+\frac{1}{p}\int_M e^{-pu}\sqrt{-1}\bar{\partial}\partial\alpha \\
&:= A+B,
\end{aligned}
$$

where we denote

$$
A=p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar{\partial}u\wedge\left(\sum_{i=1}^{k}\omega_u^{i-1}\wedge\omega^{n-i}\right),\qquad B=\frac{1}{p}\int_M e^{-pu}\sqrt{-1}\bar{\partial}\partial\alpha.
$$

Our goal is to use term A to control term B. Direct calculation gives

$$
\partial\alpha=n\sum_{i=1}^{k-1}\omega_u^{i-1}\wedge\omega^{n-i-1}\wedge\partial\omega+(n-k)\omega_u^{k-1}\wedge\omega^{n-k-1}\wedge\partial\omega,
$$

and

$$
\begin{aligned}
\bar{\partial}\partial\alpha=&(n-k)(n-k-1)\omega_u^{k-1}\wedge\omega^{n-k-2}\wedge\bar{\partial}\omega\wedge\partial\omega+(n-k)\omega_u^{k-1}\wedge\omega^{n-k-1}\wedge\bar{\partial}\partial\omega \\
&+(n-k)(n+k-1)\omega_u^{k-2}\wedge\omega^{n-k-1}\wedge\bar{\partial}\omega\wedge\partial\omega \\
&+n(n-1)\sum_{i=0}^{k-3}\omega_u^{i}\wedge\omega^{n-i-3}\wedge\bar{\partial}\omega\wedge\partial\omega+n\sum_{i=0}^{k-2}\omega_u^{i}\wedge\omega^{n-i-2}\wedge\bar{\partial}\partial\omega.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
B=\frac{(n-k)(n-k-1)}{p}&\int_M e^{-pu}\omega_u^{k-1}\wedge\omega^{n-k-2}\wedge\sqrt{-1}\bar{\partial}\omega\wedge\partial\omega \\
&+\frac{(n-k)}{p}\int_M e^{-pu}\omega_u^{k-1}\wedge\omega^{n-k-1}\wedge\sqrt{-1}\bar{\partial}\partial\omega \\
&+\frac{(n+k-1)(n-k)}{p}\int_M e^{-pu}\omega_u^{k-2}\wedge\omega^{n-k-1}\wedge\sqrt{-1}\bar{\partial}\omega\wedge\partial\omega \\
&+\frac{n(n-1)}{p}\sum_{i=0}^{k-3}\int_M e^{-pu}\omega_u^{i}\wedge\omega^{n-i-3}\wedge\sqrt{-1}\bar{\partial}\omega\wedge\partial\omega \\
&+\frac{n}{p}\sum_{i=0}^{k-2}\int_M e^{-pu}\omega_u^{i}\wedge\omega^{n-i-2}\wedge\sqrt{-1}\bar{\partial}\partial\omega.
\end{aligned}
$$

When $k=2$, term B just becomes

$$
\begin{aligned}
B &= \frac{(n-2)(n-3)}{p}\int_M e^{-pu}\omega_u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\quad + \frac{(n-2)}{p}\int_M e^{-pu}\omega_u\wedge\omega^{n-3}\wedge\sqrt{-1}\bar\partial\partial\omega \\
&\quad + \frac{(n+1)(n-2)}{p}\int_M e^{-pu}\omega^{n-3}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&= \frac{(n-2)(n-3)}{p}\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\quad + \frac{(n-2)}{p}\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-3}\wedge\sqrt{-1}\bar\partial\partial\omega \\
&\quad + \frac{2(n-1)(n-2)}{p}\int_M e^{-pu}\omega^{n-3}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\quad + \frac{(n-2)}{p}\int_M e^{-pu}\omega^{n-2}\wedge\sqrt{-1}\bar\partial\partial\omega \\
&\geq \frac{(n-2)(n-3)}{p}\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\quad + \frac{(n-2)}{p}\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-3}\wedge\sqrt{-1}\bar\partial\partial\omega - \frac{C_1}{p}\int_M e^{-pu}\omega^n.
\end{aligned}
\tag{3-3}
$$

We next use integration by parts to deal with the first term and second term on the right-hand side of the above equality. Indeed,

$$
\begin{aligned}
&\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\quad = p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\qquad + \int_M e^{-pu}\sqrt{-1}\bar\partial u\wedge\sqrt{-1}\partial(\omega^{n-4}\wedge\bar\partial\omega\wedge\partial\omega) \\
&\quad = p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega \\
&\qquad + \frac{1}{p}\int_M e^{-pu}\sqrt{-1}\bar\partial\partial(\omega^{n-4}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial\omega) \\
&\quad \geq -pC_1\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-1} - \frac{C_1}{p}\int_M e^{-pu}\omega^n \\
&\quad \geq -C_1A - \frac{C_1}{p}\int_M e^{-pu}\omega^n.
\end{aligned}
\tag{3-4}
$$

A similar calculation gives

$$\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega^{n-3}\wedge\sqrt{-1}\bar\partial\partial\omega \ge -C_1A-\frac{C_1}{p}\int_M e^{-pu}\omega^n. \tag{3-5}$$

Inserting (3-4) and (3-5) into (3-3), we have

$$B\ge -\frac{C_1}{p}A-\frac{C_1}{p}\int_M e^{-pu}\omega^n.$$

By (3-2) and choosing $p_0=2C_1+1$, we obtain for $p\ge p_0$

$$\frac{A}{2}\le\Big(1-\frac{C_1}{p}\Big)A\le\Big(\frac{C_1}{p}+C_0\Big)\int_M e^{-pu}\omega^n\le (C_0+1)\int_M e^{-pu}\omega^n.$$

By (3-7) below, we thus prove the lemma.

For the general k, $3\le k\le n$, we claim that there exist constants C_{1i} depending only on n, k and (M,ω) such that the following holds for $0\le i\le k-1$:

$$\begin{aligned}\int_M e^{-pu}\omega_u^i\wedge T_i\\ &\ge -pC_{1i}\sum_{j=0}^{k-2}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^j\wedge\omega^{n-j-1}-C_{1i}\int_M e^{-pu}\omega^n,\end{aligned}\tag{3-6}$$

where T_i is defined as the combinations of ω, $\partial\omega$, $\partial\bar\partial\omega$; more precisely

$$T_i=\sum_{0\le 3p+2q\le n-i}\omega^{n-i-3p-2q}\wedge(\sqrt{-1})^p(\partial\omega)^p\wedge(\bar\partial\omega)^p\wedge(\sqrt{-1})^q(\partial\bar\partial\omega)^q.$$

We use the claim (3-6) to prove the lemma:

$$\begin{aligned}B&\ge -C_1\sum_{i=2}^{k}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{k-i}\wedge\omega^{n+i-k-1}-\frac{C_1}{p}\int_M e^{-pu}\omega^n\\ &\ge -\frac{C_1}{p}A-\frac{C_1}{p}\int_M e^{-pu}\omega^n.\end{aligned}$$

Thus we have

$$\Big(1-\frac{C_1}{p}\Big)A\le\Big(\frac{C_1}{p}+C_0\Big)\int_M e^{-pu}\omega^n.$$

Now we choose $p_0=2C_1+1$, then for any $p\ge p_0$,

$$p^2\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-1}\le 2p(C_0+1)\int_M e^{-pu}\omega^n.$$

Therefore we have

$$
\begin{aligned}
\text{(3-7)}\qquad \int_M |\partial e^{-(p/2)u}|_g^2\omega^n &= \frac{np^2}{4}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-1}\\
&\le \frac{np(C_0+1)}{2}\int_M e^{-pu}\omega^n\\
&= pC\int_M e^{-pu}\omega^n.
\end{aligned}
$$

Now, we prove the claim (3-6) by an inductive argument. When $i=1$,

$$
\begin{aligned}
&\int_M e^{-pu}\omega_u\wedge T_1\\
&\quad= \int_M e^{-pu}\omega\wedge T_1+\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge T_1\\
&\quad= \int_M e^{-pu}\omega\wedge T_1-\int_M \partial e^{-pu}\wedge\sqrt{-1}\bar\partial u\wedge T_1+\int_M e^{-pu}\sqrt{-1}\bar\partial u\wedge\partial T_1\\
&\quad= \int_M e^{-pu}\omega\wedge T_1+p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge T_1-\frac{1}{p}\int_M \sqrt{-1}\bar\partial e^{-pu}\wedge\partial T_1\\
&\quad= p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge T_1+\int_M e^{-pu}\omega\wedge T_1-\frac{1}{p}\int_M e^{-pu}\wedge\sqrt{-1}\partial\bar\partial T_1\\
&\quad\ge -C_1p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega^{n-1}-C_1\int_M e^{-pu}\omega^n.
\end{aligned}
$$

Suppose that the claim is true for $l\le i-1$; we will prove that the claim is also true for $l=i$. Indeed,

$$
\begin{aligned}
\int_M e^{-pu}\omega_u^i\wedge T_i &= \int_M e^{-pu}\omega_u^{i-1}\wedge\omega\wedge T_i+\int_M e^{-pu}\sqrt{-1}\partial\bar\partial u\wedge\omega_u^{i-1}\wedge T_i\\
&= \int_M e^{-pu}\omega_u^{i-1}\wedge\omega\wedge T_i+p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i-1}\wedge T_i\\
&\qquad\qquad +\int_M e^{-pu}\bar\partial u\wedge\sqrt{-1}\partial(\omega_u^{i-1}\wedge T_i)\\
&:= A_{i,1}+A_{i,2}+A_{i,3}.
\end{aligned}
$$

By the induction,

$$
\begin{aligned}
A_{i,1} &= \int_M e^{-pu}\omega_u^{i-1}\wedge\omega\wedge T_i\\
&\ge -pC_{1i}(n,k,\omega)\sum_{j=0}^{k-2}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^j\wedge\omega^{n-j-1}\\
&\qquad\qquad -C_{1i}(n,k,\omega)\int_M e^{-pu}\omega^n.
\end{aligned}
$$

By the inequality (2-7) in Lemma 2.3, we have

$$
\begin{aligned}
A_{i,2} &= p\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i-1}\wedge T_i \\
&\geq -pC_{2i}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{i-1}\wedge\omega^{n-i}.
\end{aligned}\tag{3-8}
$$

Now we deal with the term $A_{i,3}$:

$$
\begin{aligned}
A_{i,3} &= \int_M e^{-pu}\bar\partial u\wedge\sqrt{-1}\partial(\omega_u^{i-1}\wedge T_i) \\
&= \frac{1}{p}\int_M e^{-pu}\sqrt{-1}\bar\partial\partial(\omega_u^{i-1}\wedge T_i) \\
&= \frac{(i-1)(i-2)}{p}\int_M e^{-pu}\sqrt{-1}\omega_u^{i-3}\wedge\bar\partial\omega\wedge\partial\omega\wedge T_i \\
&\quad + \frac{i-1}{p}\int_M e^{-pu}\omega_u^{i-2}\wedge\sqrt{-1}\bar\partial(\partial\omega\wedge T_i) + \frac{i-1}{p}\int_M e^{-pu}\omega_u^{i-2}\wedge\sqrt{-1}\bar\partial\omega\wedge\partial T_i \\
&\quad - \frac{1}{p}\int_M e^{-pu}\omega_u^{i-1}\wedge\sqrt{-1}\partial\bar\partial T_i \\
&= \frac{(i-1)(i-2)}{p}\int_M e^{-pu}\sqrt{-1}\omega_u^{i-3}\wedge\bar\partial\omega\wedge\partial\omega\wedge T_i \\
&\quad + \frac{i-1}{p}\int_M e^{-pu}\omega_u^{i-2}\wedge[\sqrt{-1}\bar\partial(\partial\omega\wedge T_i) + \sqrt{-1}\bar\partial\omega\wedge\partial T_i] \\
&\quad - \frac{1}{p}\int_M e^{-pu}\omega_u^{i-1}\wedge\sqrt{-1}\partial\bar\partial T_i \\
&\geq -pC_{3i}\sum_{j=0}^{k-2}\int_M e^{-pu}\sqrt{-1}\partial u\wedge\bar\partial u\wedge\omega_u^{j}\wedge\omega^{n-j-1} - C_{3i}(n,k,\omega)\int_M e^{-pu}\omega^n.
\end{aligned}
$$

For the last inequality, we have used the induction. □

4. Second-order estimate

In this section we use the auxiliary function in [Tosatti and Weinkove 2013] which is modified by the auxiliary function in [Hou et al. 2010] to derive the second-order estimate of the form (1-3). The difficulty arises from the third-order derivatives. Locally the equation is

$$\sigma_k(\omega_u) = e^f. \tag{4-1}$$

Theorem 4.1. *There exists a uniform constant C depending only on (M,ω), n, k and f such that*

$$\max|\partial\bar\partial u|_g \leq C(1+\max|\nabla u|_g^2). \tag{4-2}$$

Proof. Denote $w_{i\bar{j}} = g_{i\bar{j}} + u_{i\bar{j}}$ and let $\xi \in T^{1,0}M$, $|\xi|_g^2 = 1$.

We use the auxiliary function similar to that in [Tosatti and Weinkove 2013]:

$$H(x,\xi) = \log(w_{k\bar{l}}\xi^k\bar{\xi}^l) + c_0 \log(g^{k\bar{l}} w_{p\bar{l}} w_{k\bar{q}} \xi^p \bar{\xi}^q) + \varphi(|\nabla u|_g^2) + \psi(u),$$

where φ, ψ are given by

$$\varphi(s) = -\frac{1}{2}\log\Big(1 - \frac{s}{2K}\Big), \qquad 0 \le s \le K-1,$$
$$\psi(t) = -A\log\Big(1 + \frac{t}{2L}\Big), \quad -L+1 \le t \le 0,$$

for

$$K := \sup_M |\nabla u|_g^2 + 1, \quad L = \sup_M |u| + 1, \quad A := 2L(C_0 + 1),$$

where A_0 is a constant to be determined later and c_0 is a small positive constant depending only on n and will be determined later. By [Hou et al. 2010], we have

$$\frac{1}{2K} \ge \varphi' \ge \frac{1}{4K} > 0, \qquad \varphi'' = 2(\varphi')^2 > 0. \tag{4-3}$$

$$\frac{A}{L} \ge -\psi' \ge \frac{A}{2L} = C_0 + 1, \quad \psi'' \ge \frac{2\varepsilon_0}{1-\varepsilon_0}(\psi')^2, \quad \text{for } \varepsilon_0 \le \frac{1}{2A+1}. \tag{4-4}$$

These inequalities will be used below.

Suppose $H(x,\xi)$ attains its maximum at the point x_0 in the direction ξ_0. Then we choose local coordinates $\{\partial/\partial z^1, \dots, \partial/\partial z^n\}$ near x_0 such that

$$g_{i\bar{j}}(x_0) = \delta_{ij}, \quad u_{i\bar{j}} = u_{i\bar{i}}(x_0)\delta_{ij}, \quad \lambda_i = w_{i\bar{i}}(x_0) = 1 + u_{i\bar{i}}(x_0) \quad \text{with } \lambda_1 \ge \cdots \ge \lambda_n.$$

We want to prove that

$$H(x_0,\xi) \le H\Big(x_0, \frac{\partial}{\partial z^1}\Big) \quad \text{for all } \xi \in T^{1,0}M,\ |\xi|_g^2 = 1, \sum_{i,j} w_{i\bar{j}}(x_0)\xi^i\xi^{\bar{j}} > 0$$

by choosing c_0 small enough. In fact, at x_0 we have

$$\log(w_{k\bar{l}}\xi^k\bar{\xi}^l) + c_0\log(g^{k\bar{l}} w_{p\bar{l}} w_{k\bar{q}}\xi^p\bar{\xi}^q) = \log\Big(\sum_{k=1}^n w_{k\bar{k}}|\xi^k|^2\Big) + c_0\log\Big(\sum_{k=1}^n |w_{k\bar{k}}|^2|\xi^k|^2\Big).$$

If $w_{n\bar{n}} \ge -w_{1\bar{1}}$, which is always satisfied when $n \le 3$, then $w_{i\bar{i}}^2 \le w_{1\bar{1}}$. Thus we have $H(x_0,\xi) \le H(x_0, \partial/\partial z^1)$.

Now we suppose that $w_{n\bar{n}} < -w_{1\bar{1}}$. Thus we have $n \geq 4$. Let i_0 be the smallest integer satisfying $w_{i\bar{i}} < -w_{1\bar{1}}$. Then $i_0 \geq k+1$. By $|w_{i\bar{i}}| < (n-2)w_{1\bar{1}}$, so

$$\begin{aligned}
&\log \sum_{i=1}^{n} w_{i\bar{i}}|\xi^i|^2 + c_0 \log \sum_{i=1}^{n} |w_{i\bar{i}}|^2|\xi^i|^2 \\
&\leq \log w_{1\bar{1}}\Bigg(\sum_{i=1}^{i_0-1}|\xi^i|^2 - \sum_{i=i_0}^{n}|\xi^i|^2\Bigg) + c_0 \log\Bigg(w_{1\bar{1}}^2\sum_{i=1}^{i_0-1}|\xi^i|^2 + (n-2)^2 w_{1\bar{1}}^2 \sum_{i=1}^{i_0-1}|\xi^i|^2\Bigg) \\
&= \log w_{1\bar{1}}(1-2t) + c_0 \log w_{1\bar{1}}^2(1-t+(n-2)^2 t) := h(t),
\end{aligned}$$

where $t = \sum_{i=i_0}^{n} |\xi^i|^2 \in \big(0, \tfrac{1}{2}\big)$.

By choosing $c_0 = 2/((n-2)^2-1)$, we have $h'(t) \leq 0$. Then

$$h(t) \leq h(0) = \log(w_{1\bar{1}}) + c_0 \log w_{1\bar{1}}^2.$$

Consequently, we obtain

$$H(x_0, \xi) \leq H\Big(x_0, \frac{\partial}{\partial z^1}\Big) \quad \text{for all } \xi \in T^{1,0}M,\ |\xi|_g^2 = 1,\ \sum_{i,j} \eta_{i\bar{j}}(x_0)\xi^i \xi^{\bar{j}} > 0,$$

by choosing $c_0 = 2/((n-2)^2-1)$ when $n \geq 4$ and $c_0 = 1$ when $n \leq 3$.

We extend ξ_0 near x_0 by $\xi_0 = (g_{1\bar{1}})^{-1/2}(\partial/\partial z^1)$. Consider the function

$$Q(x) = H(x, \xi_0) = \log(g_{1\bar{1}}^{-1} w_{1\bar{1}}) + c_0 \log(g_{1\bar{1}}^{-1} g^{k\bar{l}} w_{1\bar{l}} w_{k\bar{1}}) + \varphi(|\nabla u|_g^2) + \psi(u).$$

We will calculate $F^{i\bar{j}} Q_{i\bar{j}}$ at x_0 to get the estimate; all the calculations are taken at x_0. For simplicity, we denote $\xi = \xi_0$ in the following. By $\langle \xi, \bar{\xi}\rangle_g = |\xi|_g^2 = 1$, differentiating both sides, we obtain at x_0

$$\begin{aligned}
0 = \frac{\partial}{\partial z^i}\langle \xi, \bar{\xi}\rangle_g &= \langle \nabla_{\partial/\partial z^i}\xi, \bar{\xi}\rangle_g + \langle \xi, \nabla_{\partial/\partial z^i}\bar{\xi}\rangle_g \\
&= \Big\langle \xi^k{}_{,i}\frac{\partial}{\partial z^k}, \overline{\xi^l}\frac{\partial}{\bar{\partial} z^l}\Big\rangle_g + \Big\langle \xi^k \frac{\partial}{\partial z^k}, \overline{\xi^l}{}_{,i}\frac{\partial}{\bar{\partial} z^l}\Big\rangle_g \\
&= g_{k\bar{l}}\xi^k{}_{,i}\overline{\xi^l} + g_{k\bar{l}}\xi^k \overline{\xi^l}{}_{,i} \\
&= \xi^1{}_{,i} + \overline{\xi^1}{}_{,i}.
\end{aligned} \tag{4-5}$$

We also have the basic formula for $\xi \in T^{1,0}M$:

$$\begin{aligned}
\overline{\xi^k}{}_{,i} &= \frac{\partial \overline{\xi^k}}{\partial z^i} = \overline{\frac{\partial \xi^k}{\partial \overline{z^i}}} = \overline{\xi^k{}_{,\bar{i}}}, \quad \xi^k{}_{,\bar{i}} = \frac{\partial \xi^k}{\partial \overline{z^i}} = \overline{\frac{\partial \overline{\xi^k}}{\partial z^i}} = \overline{\overline{\xi^k}{}_{,i}} \\
\overline{\xi^k}{}_{,i} &= \frac{\partial \overline{\xi^k}}{\partial z^i} = \overline{\frac{\partial \xi^k}{\partial \overline{z^i}}} = \overline{\xi^k{}_{,\bar{i}}}, \quad \xi^k{}_{,\bar{i}} = \frac{\partial \xi^k}{\partial \overline{z^i}} = \overline{\frac{\partial \overline{\xi^k}}{\partial z^i}} = \overline{\overline{\xi^k}{}_{,i}}
\end{aligned} \tag{4-6}$$

Direct calculations give

$$Q_i = \frac{(w_{k\bar l}\xi^k\overline{\xi^l})_i}{w_{k\bar l}\xi^k\overline{\xi^l}} + c_0\frac{(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})_i}{g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l}} + \varphi_i + \psi_i,$$

$$\begin{aligned}Q_{i\bar i} = {}& \frac{(w_{k\bar l}\xi^k\overline{\xi^l})_{i\bar i}}{w_{k\bar l}\xi^k\overline{\xi^l}} - \frac{(w_{k\bar l}\xi^k\overline{\xi^l})_i(w_{k\bar l}\xi^k\overline{\xi^l})_{\bar i}}{(w_{k\bar l}\xi^k\overline{\xi^l})^2} + c_0\frac{(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})_{i\bar i}}{g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l}}\\ &- c_0\frac{(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})_i(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})_{\bar i}}{(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})^2} + \varphi_{i\bar i} + \psi_{i\bar i}.\end{aligned}$$

Next, we want to simplify Q_i and $Q_{i\bar i}$. By (4-5), we have

$$(w_{k\bar l}\xi^k\overline{\xi^l})_i = w_{k\bar l,i}\xi^k\overline{\xi^l} + w_{k\bar l}\xi^k{}_{,i}\overline{\xi^l} + w_{k\bar l}\xi^k\overline{\xi^l}{}_{,i} = w_{1\bar 1,i} + w_{1\bar 1}(\xi^1{}_{,i} + \overline{\xi^1}{}_{,i}) = w_{1\bar 1 i},$$

Thus we have

$$\begin{aligned}&(g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l})_i\\ &\quad= g^{p\bar q}w_{k\bar q i}w_{p\bar l}\xi^k\overline{\xi^l} + g^{p\bar q}w_{k\bar q}w_{p\bar l i}\xi^k\overline{\xi^l} + g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k{}_i\overline{\xi^l} + g^{p\bar q}w_{k\bar q}w_{p\bar l}\xi^k\overline{\xi^l}{}_{,i}\\ &\quad= w_{1\bar 1}(w_{1\bar 1 i} + w_{1\bar 1 i}) + w_{1\bar 1}{}^2(\xi^1{}_{,i} + \overline{\xi^1}{}_{,i})\\ &\quad= 2w_{1\bar 1}w_{1\bar 1 i}.\end{aligned}$$

Therefore, we obtain the simplified formula for Q_i at x_0:

$$Q_i = \frac{w_{1\bar 1 i}}{w_{1\bar 1}} + c_0\frac{2w_{1\bar 1 i}}{w_{1\bar 1}} + \varphi_i + \psi_i = (1+2c_0)\frac{w_{1\bar 1 i}}{w_{1\bar 1}} + \varphi_i + \psi_i = 0 \tag{4-7}$$

Similar calculations give

$$\begin{aligned}(w_{k\bar l}\xi^k\overline{\xi^l})_{i\bar i} &= [w_{k\bar l i}\xi^k\overline{\xi^l} + w_{k\bar l}(\xi^k{}_i\overline{\xi^l} + \xi^k\overline{\xi^l}{}_i)]_{\bar i}\\ &= w_{k\bar l i\bar i}\xi^k\overline{\xi^l} + w_{k\bar l i}(\xi^k{}_{\bar i}\overline{\xi^l} + \xi^k\overline{\xi^l}{}_{\bar i}) + w_{k\bar l\bar i}(\xi^k{}_i\overline{\xi^l} + \xi^k\overline{\xi^l}{}_i)\\ &\qquad + w_{k\bar l}(\xi^k{}_{i\bar i}\overline{\xi^l} + \xi^k{}_{\bar i}\overline{\xi^l}{}_i + \xi^k{}_i\overline{\xi^l}{}_{\bar i} + \xi^k\overline{\xi^l}{}_{i\bar i})\\ &= w_{1\bar 1 i\bar i} + w_{k\bar 1 i}\xi^k{}_{\bar i} + w_{1\bar l i}\overline{\xi^l}{}_{\bar i} + w_{k\bar 1\bar i}\xi^k{}_i + w_{1\bar l\bar i}\overline{\xi^l}{}_i\\ &\qquad + w_{1\bar 1}(\xi^1{}_{i\bar i} + \overline{\xi^1}{}_{i\bar i}) + w_{k\bar k}(\xi^k{}_{\bar i}\overline{\xi^k}{}_i + \xi^k{}_i\overline{\xi^k}{}_{\bar i})\\ &= w_{1\bar 1 i\bar i} + 2\sum_{k\neq 1}\mathrm{Re}(w_{k\bar 1 i}\xi^k{}_{\bar i} + w_{1\bar k i}\overline{\xi^k}{}_i) + w_{1\bar 1}(\xi^1{}_{i\bar i} + \overline{\xi^1}{}_{i\bar i})\\ &\qquad + w_{k\bar k}(|\xi^k{}_{\bar i}|^2 + |\xi^k{}_i|^2).\end{aligned}$$

The last equality holds because we use (4-2) and (4-5) and the fact

$$w_{k\bar 1 i}\xi^k{}_{\bar i} + w_{1\bar l\bar i}\overline{\xi^l}{}_i = 2\,\mathrm{Re}(w_{k\bar 1 i}\xi^k{}_{\bar i}),\quad w_{1\bar l i}\overline{\xi^l}{}_{\bar i} + w_{k\bar 1\bar i}\xi^k{}_i = 2\,\mathrm{Re}(w_{1\bar k i}\overline{\xi^k}{}_i).$$

We can also calculate

$$
\begin{aligned}
&(g^{p\bar q} w_{k\bar q} w_{p\bar l} \xi^k \overline{\xi^l})_{i\bar i}\\
&\quad = g^{p\bar q}(w_{k\bar q i} w_{p\bar l} \xi^k \overline{\xi^l} + w_{k\bar q} w_{p\bar l i} \xi^k \overline{\xi^l} + w_{k\bar q} w_{p\bar l} \xi^k{}_i \overline{\xi^l} + w_{k\bar q} w_{p\bar l} \xi^k \overline{\xi^l}{}_i)_{\bar i}\\
&\quad = g^{p\bar q}(w_{k\bar q i\bar i} w_{p\bar l} \xi^k \overline{\xi^l} + w_{k\bar q i} w_{p\bar l\bar i} \xi^k \overline{\xi^l} + w_{k\bar q i} w_{p\bar l} \xi^k{}_{\bar i} \overline{\xi^l} + w_{k\bar q i} w_{p\bar l} \xi^k \overline{\xi^l}{}_{\bar i})\\
&\qquad + g^{p\bar q}(w_{k\bar q\bar i} w_{p\bar l i} \xi^k \overline{\xi^l} + w_{k\bar q} w_{p\bar l i\bar i} \xi^k \overline{\xi^l} + w_{k\bar q} w_{p\bar l i} \xi^k{}_{\bar i} \overline{\xi^l} + w_{k\bar q} w_{p\bar l i} \xi^k \overline{\xi^l}{}_{\bar i})\\
&\qquad\quad + g^{p\bar q}(w_{k\bar q\bar i} w_{p\bar l} \xi^k{}_i \overline{\xi^l} + w_{k\bar q} w_{p\bar l\bar i} \xi^k{}_i \overline{\xi^l} + w_{k\bar q} w_{p\bar l} \xi^k{}_{i\bar i} \overline{\xi^l} + w_{k\bar q} w_{p\bar l} \xi^k{}_i \overline{\xi^l}{}_{\bar i})\\
&\qquad\qquad + g^{p\bar q}(w_{k\bar q\bar i} w_{p\bar l} \xi^k \overline{\xi^l}{}_i + w_{k\bar q} w_{p\bar l\bar i} \xi^k \overline{\xi^l}{}_i + w_{k\bar q} w_{p\bar l} \xi^k{}_{\bar i} \overline{\xi^l}{}_i + w_{k\bar q} w_{p\bar l} \xi^k \overline{\xi^l}{}_{i\bar i})\\
&\quad = w_{1\bar 1 i\bar i} w_{1\bar 1} + w_{1\bar p i} w_{p\bar 1\bar i} + w_{k\bar 1 i} w_{1\bar 1} \xi^k{}_{\bar i} + w_{1\bar p i} w_{p\bar p} \overline{\xi^p}{}_{\bar i}\\
&\qquad + w_{1\bar p\bar i} w_{p\bar 1 i} + w_{1\bar 1} w_{1\bar 1 i\bar i} + w_{p\bar p} w_{p\bar 1 i} \xi^p{}_{\bar i} + w_{1\bar 1} w_{1\bar l i} \overline{\xi^l}{}_{\bar i}\\
&\qquad\quad + w_{k\bar 1\bar i} w_{1\bar 1} \xi^k{}_i + w_{p\bar p} w_{p\bar 1\bar i} \xi^p{}_i + w_{1\bar 1}{}^2 \xi^1{}_{i\bar i} + w_{p\bar p}{}^2 \xi^p{}_i \overline{\xi^p}{}_{\bar i}\\
&\qquad\qquad + w_{1\bar p\bar i} w_{p\bar p} \overline{\xi^p}{}_i + w_{1\bar 1} w_{1\bar l\bar i} \overline{\xi^l}{}_i + w_{p\bar p}{}^2 \xi^p{}_{\bar i} \overline{\xi^p}{}_i + w_{1\bar 1}{}^2 \overline{\xi^1}{}_{i\bar i}\\
&\quad = 2w_{1\bar 1} w_{1\bar 1 i\bar i} + |w_{1\bar p i}|^2 + |w_{1\bar p\bar i}|^2 + 2w_{1\bar 1}\operatorname{Re}(w_{p\bar 1 i} \xi^p{}_{\bar i} + w_{p\bar 1\bar i} \xi^p{}_i)\\
&\qquad + 2w_{p\bar p}\operatorname{Re}(w_{1\bar p i} \overline{\xi^p}{}_{\bar i} + w_{p\bar 1 i} \xi^p{}_{\bar i}) + w_{p\bar p}{}^2(|\xi^p{}_i|^2 + |\xi^p{}_{\bar i}|^2) + w_{1\bar 1}{}^2(\xi^1{}_{i\bar i} + \overline{\xi^1}{}_{i\bar i})
\end{aligned}
$$

Therefore we simplify $Q_{i\bar i}$ at x_0 as follows

$$
\begin{aligned}
Q_{i\bar i} = (1+2c_0)\frac{w_{1\bar 1 i\bar i}}{w_{1\bar 1}} + \frac{c_0}{w_{1\bar 1}^2}\sum_{p\neq 1}(|w_{1\bar p i}|^2 + |w_{1\bar p\bar i}|^2)&\\
-(1+2c_0)\frac{|w_{1\bar 1 i}|^2}{w_{1\bar 1}{}^2} + (**)_{i\bar i} + \varphi_{i\bar i} + \psi_{i\bar i}&,
\end{aligned}
$$

where $(**)_{i\bar i}$ is given by

$$
\begin{aligned}
(**)_{i\bar i} = {}& \frac{2}{w_{1\bar 1}}\sum_{k\neq 1}\operatorname{Re}(w_{k\bar 1 i}\xi^k{}_{\bar i} + w_{1\bar k i}\overline{\xi^k}{}_i) + \xi^1{}_{i\bar i} + \overline{\xi^1}{}_{i\bar i} + \frac{w_{k\bar k}}{w_{1\bar 1}}(|\xi^k{}_{\bar i}|^2 + |\xi^k{}_i|^2)\\
&+ \frac{2c_0}{w_{1\bar 1}}\sum_{p\neq 1}\operatorname{Re}(w_{p\bar 1 i}\xi^p{}_{\bar i} + w_{p\bar 1\bar i}\xi^p{}_i) + \sum_{p\neq 1}\frac{2c_0 w_{p\bar p}}{w_{1\bar 1}{}^2}\operatorname{Re}(w_{1\bar p i}\overline{\xi^p}{}_{\bar i} + w_{p\bar 1 i}\xi^p{}_{\bar i})\\
&+ \frac{2c_0 w_{p\bar p}{}^2}{w_{1\bar 1}{}^2}(|\xi^p{}_i|^2 + |\xi^p{}_{\bar i}|^2) + c_0(\xi^1{}_{i\bar i} + \overline{\xi^1}{}_{i\bar i}).
\end{aligned}
$$

For this term $(**)_{i\bar i}$, we have the estimate

$$
(**)_{i\bar i} \geq -\frac{c_0}{2w_{1\bar 1}{}^2}\sum_{p\neq 1}(|w_{1\bar p i}|^2 + |w_{1\bar p\bar i}|^2) - C,
$$

where C is a positive constant depending only on (M,ω).

Let $F(\omega_u)=(\sigma_k(\omega_u))^{1/k}$. We denote by

$$F^{i\bar{j}}=\frac{\partial F}{\partial w_{i\bar{j}}},\qquad F^{i\bar{j},p\bar{q}}=\frac{\partial^2 F}{\partial w_{i\bar{j}}\partial w_{p\bar{q}}},$$

where $(w_u)_{i\bar{j}}=g_{i\bar{j}}+u_{i\bar{j}}$. Then, the positive definite matrix $(F^{i\bar{j}}(\omega_u))$ is diagonalized at the point x_0. More precisely,

$$F^{i\bar{j}}(\omega_u)=\delta_{ij}F^{i\bar{i}}(\omega_u)=\frac{1}{k}[\sigma_k(\lambda)]^{1/k-1}\sigma_{k-1}(\lambda|i)\delta_{ij}. \tag{4-8}$$

Furthermore, at x_0,

$$F^{i\bar{j},p\bar{q}}(\omega_u)=\begin{cases}F^{i\bar{i},p\bar{p}}, & \text{if } i=j,\ p=q;\\ F^{i\bar{p},p\bar{i}}, & \text{if } i=q,\ p=j,\ i\neq p;\\ 0, & \text{otherwise,}\end{cases} \tag{4-9}$$

in which

$$\begin{aligned}F^{i\bar{i},p\bar{p}}&=\frac{1}{k}[\sigma_k(\lambda)]^{1/k-1}(1-\delta_{ip})\sigma_{k-2}(\lambda|ip)\\&\qquad+\frac{1}{k}\Big(\frac{1}{k}-1\Big)[\sigma_k(\lambda)]^{1/k-2}\sigma_{k-1}(\lambda|i)\sigma_{k-1}(\lambda|p),\\ F^{i\bar{p},p\bar{i}}&=-\frac{1}{k}[\sigma_k(\lambda)]^{1/k-1}\sigma_{k-2}(\lambda|ip).\end{aligned}$$

We have, in addition, at x_0

$$\sum_{i=1}^n F^{i\bar{i}}w_{i\bar{i}}=\sum_{i=1}^n F^{i\bar{i}}\lambda_i=\sigma_k^{1/k}=e^{f/k}. \tag{4-10}$$

By the maximum principal, we have

$$\begin{aligned}0\geq F^{i\bar{j}}Q_{i\bar{j}}&=F^{i\bar{i}}Q_{i\bar{i}} \\ &\geq(1+2c_0)\sum_{i=1}^n\frac{F^{i\bar{i}}u_{1\bar{1}i\bar{i}}}{w_{1\bar{1}}}+\frac{c_0}{2}\sum_{i=1}^n\sum_{p\neq 1}\frac{F^{i\bar{i}}|u_{1\bar{p}i}|^2}{w_{1\bar{1}}^2}\\ &\quad-(1+2c_0)\sum_{i=1}^n\frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w_{1\bar{1}}^2}+\psi'\sum_{i=1}^nF^{i\bar{i}}u_{i\bar{i}}+\psi''\sum_{i=1}^nF^{i\bar{i}}|u_i|^2\\ &\quad+\varphi''\sum_{i=1}^nF^{i\bar{i}}|\nabla u|_i^2|\nabla u|_{\bar{i}}^2+\varphi'\sum_{i,p=1}^nF^{i\bar{i}}(|u_{p\bar{i}}|^2+|u_{pi}|^2)\\ &\quad+\varphi'\sum_{i,p=1}^nF^{i\bar{i}}(u_{p\bar{i}i}u_{\bar{p}}+u_{\bar{p}i\bar{i}}u_p)-C_1\sum_{i=1}^nF^{i\bar{i}}\\ &:=I_1+I_2+I_3+I_4+I_5+I_6+I_7+I_8-C_1\sum_{i=1}^nF^{i\bar{i}}\end{aligned} \tag{4-11}$$

The equation can be written as $F(\omega_u) = e^{f/k} := h$. Differentiating this, we get

$$\sum_{i,j=1}^{n} F^{i\bar{j}} u_{i\bar{j}l} = \nabla_l F = h_l, \quad \sum_{i,j=1}^{n} F^{i\bar{j}} u_{i\bar{j}l\bar{m}} + \sum_{i,j,p,q=1}^{n} F^{i\bar{j},p\bar{q}} u_{i\bar{j}l} u_{p\bar{q}\bar{m}} = h_{l\bar{m}}.$$

and

$$\sum_{i=1}^{n} F^{i\bar{i}} u_{i\bar{i}1\bar{1}} = h_{1\bar{1}} - \sum_{i,j,p,q=1}^{n} F^{i\bar{j},p\bar{q}} u_{i\bar{j}1} u_{p\bar{q}\bar{1}}.$$

By commuting the covariant derivatives formula (2-12), we have

$$\begin{aligned}\sum_{i=1}^{n} F^{i\bar{i}} u_{1\bar{1}i\bar{i}} &= \sum_{i=1}^{n} F^{i\bar{i}} u_{i\bar{i}1\bar{1}} + \sum_{i=1}^{n} F^{i\bar{i}} \Big(u_{1\bar{1}} - \sum_{i=1}^{n} u_{i\bar{i}}\Big) R_{i\bar{i}1\bar{1}} \\ &\quad + \sum_{i=1}^{n} F^{i\bar{i}} \bigg(\sum_{p=1}^{n} T_{1i}^{p} u_{p\bar{1}\bar{i}} + \sum_{q=1}^{n} \overline{T_{1i}^{q}} u_{1\bar{q}i} - \sum_{p=1}^{n} |T_{1i}^{p}|^2 u_{p\bar{p}} \bigg).\end{aligned} \tag{4-12}$$

Inserting (4-12) into the term I_1, we have

$$\begin{aligned} I_1 &= (1+2c_0) \sum_{i=1}^{n} \frac{F^{i\bar{i}} u_{1\bar{1}i\bar{i}}}{w_{1\bar{1}}} \\ &= (1+2c_0) \sum_{i=1}^{n} \frac{F^{i\bar{i}} u_{i\bar{i}1\bar{1}}}{w_{1\bar{1}}} + (1+2c_0) \sum_{i=1}^{n} \frac{F^{i\bar{i}} (u_{1\bar{1}} - u_{i\bar{i}}) R_{i\bar{i}1\bar{1}}}{w_{1\bar{1}}} \\ &\qquad + 2(1+2c_0) \sum_{i,p=1}^{n} F^{i\bar{i}} \operatorname{Re}\Big(\frac{T_{1i}^{p} u_{p\bar{1}\bar{i}}}{w_{1\bar{1}}} \Big) - (1+2c_0) \sum_{i,p=1}^{n} F^{i\bar{i}} \frac{|T_{1i}^{p}|^2 u_{p\bar{p}}}{w_{1\bar{1}}} \\ &= (1+2c_0) \frac{h_{1\bar{1}}}{w_{1\bar{1}}} - (1+2c_0) \sum_{i,j,p,q=1}^{n} \frac{F^{i\bar{j},p\bar{q}} u_{i\bar{j}1} u_{p\bar{q}\bar{1}}}{w_{1\bar{1}}} \\ &\quad + (1+2c_0) \sum_{i=1}^{n} \frac{F^{i\bar{i}} (u_{1\bar{1}} - u_{i\bar{i}}) R_{i\bar{i}1\bar{1}}}{w_{1\bar{1}}} + 2(1+2c_0) \sum_{i}^{n} F^{i\bar{i}} \operatorname{Re}\Big(\frac{T_{1i}^{1} u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}} \Big) \\ &\quad + 2(1+2c_0) \sum_{i=1}^{n} F^{i\bar{i}} \operatorname{Re}\Big(\sum_{p\neq 1} \frac{T_{1i}^{p} u_{p\bar{1}\bar{i}}}{w_{1\bar{1}}} \Big) - (1+2c_0) \sum_{i,p=1}^{n} F^{i\bar{i}} \frac{|T_{1i}^{p}|^2 u_{p\bar{p}}}{w_{1\bar{1}}} \\ &:= I_{11} + I_{12} + I_{13} + I_{14} + I_{15} + I_{16}. \end{aligned} \tag{4-13}$$

We estimate each term in this sum. First we have

$$I_{11} + I_{13} + I_{16} \geq -C_1 - 3(nC_2 + C_3) \sum_{i=1}^{n} F^{i\bar{i}},$$

where we have supposed that $\sup_M |T|_g^2 \leq C_2$, $\sup_M |R| \leq C_3$.

We claim $I_{15}+I_2 \geq -18n^2C_2\sum_{i=1}^n F^{i\bar{i}}$. Indeed, since $\frac{1}{n^2} \leq c_0 \leq 1$, we have

$$\begin{aligned}
I_{15}+I_2 &= \frac{c_0}{2}\sum_{i=1}^n\sum_{p\neq 1}\frac{F^{i\bar{i}}|u_{1\bar{p}i}|^2}{w_{1\bar{1}}^2}+2(1+2c_0)\sum_{i=1}^n F^{i\bar{i}}\,\mathrm{Re}\Big(\sum_{p\neq 1}\frac{T_{1i}^p u_{p\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big)\\
&= \frac{c_0}{2}\sum_{i=1}^n F^{i\bar{i}}\sum_{p\neq 1}\Big|\frac{u_{1\bar{p}i}}{w_{1\bar{1}}}+\frac{2(1+2c_0)}{c_0}T_{1i}^p\Big|^2-\frac{2(1+2c_0)^2}{c_0}\sum_{i=1}^n\sum_{p\neq 1}F^{i\bar{i}}|T_{1i}^p|^2\\
&\geq -\frac{2(1+2c_0)^2}{c_0}\sum_{i=1}^n\sum_{p\neq 1}F^{i\bar{i}}|T_{1i}^p|^2\\
&\geq -18n^2C_2\sum_{i=1}^n F^{i\bar{i}}.
\end{aligned}$$

Then we obtain

$$I_1+I_2 \geq -(1+2c_0)\sum_{i,j,p,q=1}^n\frac{F^{i\bar{j},p\bar{q}}u_{i\bar{j}1}u_{p\bar{q}\bar{1}}}{w_{1\bar{1}}}+2(1+2c_0)\sum_{i=1}^n F^{i\bar{i}}\mathrm{Re}\Big(\frac{T_{1i}^1 u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big) - (21n^2C_2+3C_3)\sum_{i=1}^n F^{i\bar{i}}-C_1. \tag{4-14}$$

For terms I_7+I_8, we claim

$$I_7+I_8 \geq \frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar{i}}|u_{i\bar{i}}|_1^2-(C_2+C_3)\sum_{i=1}^n F^{i\bar{i}}-C_1. \tag{4-15}$$

Indeed, by the commutation formula for covariant derivatives (2-11) in Section 2,

$$u_{pi\bar{i}} = u_{i\bar{i}p}+T_{pi}^i u_{i\bar{i}}+u_q R_{i\bar{i}p\bar{q}}, \quad u_{\bar{p}i\bar{i}} = u_{i\bar{p}\bar{i}} = u_{i\bar{i}\bar{p}}-\overline{T_{ip}^i}u_{i\bar{i}}.$$

Then

$$\begin{aligned}
\sum_{i=1}^n F^{i\bar{i}}u_{pi\bar{i}} &= \sum_{i=1}^n F^{i\bar{i}}u_{i\bar{i}p}+\sum_{i=1}^n F^{i\bar{i}}(T_{pi}^i u_{i\bar{i}}+u_q R_{i\bar{i}p\bar{q}})\\
&= h_p+\sum_{i=1}^n F^{i\bar{i}}(T_{pi}^i u_{i\bar{i}}+u_q R_{i\bar{i}p\bar{q}})\\
\sum_{i=1}^n F^{i\bar{i}}u_{\bar{p}i\bar{i}} &= \sum_{i=1}^n F^{i\bar{i}}u_{i\bar{i}\bar{p}}+\sum_{i=1}^n F^{i\bar{i}}\overline{T_{ip}^i}u_{i\bar{i}} = h_{\bar{p}}+\sum_{i=1}^n F^{i\bar{i}}\overline{T_{ip}^i}u_{i\bar{i}}
\end{aligned}$$

Inserting the above formula into I_8, we obtain

$$
\begin{aligned}
(4\text{-}16)\quad I_8 &= \varphi' \sum_{i,p=1}^{n} F^{i\bar{i}}(u_{p\bar{i}i}u_{\bar{p}} + u_{\bar{p}i\bar{i}}u_p) \\
&= \varphi' \sum_{p=1}^{n} u_{\bar{p}} \Big[h_p + \sum_{i=1}^{n} F^{i\bar{i}}(T^i_{pi}u_{i\bar{i}} + u_q R_{i\bar{i}p\bar{q}}) \Big] + \varphi' \sum_{p=1}^{n} u_p \Big[h_{\bar{p}} - \sum_{i=1}^{n} F^{i\bar{i}} \overline{T^i_{ip}} u_{i\bar{i}} \Big] \\
&= 2\varphi' \sum_{i,p=1}^{n} F^{i\bar{i}} u_{i\bar{i}} \operatorname{Re}(u_{\bar{p}} T^i_{pi}) + \varphi' \sum_{p=1}^{n} \Big[2\operatorname{Re}(u_{\bar{p}} h_p) + \sum_{i,q=1}^{n} u_{\bar{p}} u_q F^{i\bar{i}} R_{i\bar{i}p\bar{q}} \Big] \\
&= I_{81} + I_{82}.
\end{aligned}
$$

For the term I_{82}, we have

$$
I_{82} \geq -C_3 \sum_{i=1}^{n} F^{i\bar{i}} - C_1.
$$

For the term I_{81}, we obtain

$$
\begin{aligned}
I_{81} + I_7 &= 2\varphi' \sum_{i,p=1}^{n} F^{i\bar{i}} u_{i\bar{i}} \operatorname{Re}(u_{\bar{p}} T^i_{pi}) + \varphi' \sum_{i,p=1}^{n} F^{i\bar{i}}(|u_{p\bar{i}}|^2 + |u_{pi}|^2) \\
&\geq \varphi' \sum_{i=1}^{n} F^{i\bar{i}} \Big[|u_{i\bar{i}}|^2 + 2u_{i\bar{i}} \operatorname{Re}\Big(\sum_{p=1}^{n} u_{\bar{p}} T^i_{pi} \Big) \Big] \\
&= \varphi' \sum_{i=1}^{n} F^{i\bar{i}} \Big| \frac{u_{i\bar{i}}}{2} + 2\sum_{p=1}^{n} u_p \overline{T^i_{pi}} \Big|^2 + \frac{3}{4}\varphi' \sum_{i=1}^{n} F^{i\bar{i}} |u_{i\bar{i}}|^2 - 4\varphi' \sum_{i=1}^{n} F^{i\bar{i}} \Big| \sum_{p=1}^{n} u_p \overline{T^i_{pi}} \Big|^2 \\
&\geq \frac{1}{2}\varphi' \sum_{i=1}^{n} F^{i\bar{i}} |u_{i\bar{i}}|^2 - C_2 \sum_{i=1}^{n} F^{i\bar{i}}.
\end{aligned}
$$

Thus we have proved the above claim (4-15). Moreover, applying (4-10) yields

$$
\begin{aligned}
\psi' \sum_{i=1}^{n} F^{i\bar{i}} u_{i\bar{i}} &= \psi' \sum_{i=1}^{n} F^{i\bar{i}} (\lambda_i - 1) \\
&= \psi' h - \psi' \sum_{i=1}^{n} F^{i\bar{i}} \geq -2(C_0 + 1) \sup_M e^{f/k} - \psi' \sum_{i=1}^{n} F^{i\bar{i}}.
\end{aligned}
$$

Similarly,

$$\begin{aligned}\frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}|u_{i\bar i}|^2 &= \frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}(\lambda_i-1)^2\\ &= \frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\lambda_i^2-\varphi'\sum_{i=1}^n F^{i\bar i}\lambda_i+\frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\\ &= \frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\lambda_i^2-\varphi' h+\frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\\ &\geq \frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\lambda_i^2-\frac{1}{2}\sup_M e^{f/k}+\frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}.\end{aligned}$$

Inserting these terms into (4-11), we obtain

$$\begin{aligned}0 &\geq F^{i\bar i}Q_{i\bar i}\\ &\geq -(1+2c_0)\sum_{i,j,p,q=1}^n \frac{F^{i\bar j,p\bar q}u_{i\bar j 1}u_{p\bar q\bar 1}}{w_{1\bar 1}}+2(1+2c_0)\sum_{i=1}^n F^{i\bar i}\,\mathrm{Re}\Big(\frac{T^1_{1i}u_{1\bar 1\bar i}}{w_{1\bar 1}}\Big)\\ &\qquad -(1+2c_0)\sum_{i=1}^n\frac{F^{i\bar i}|u_{1\bar 1 i}|^2}{w_{1\bar 1}^2}\\ &\qquad +\varphi''\sum_{i=1}^n F^{i\bar i}|\nabla u|_i^2|\nabla u|_{\bar i}^2+\psi''\sum_{i=1}^n F^{i\bar i}|u_i|^2+\frac{1}{2}\varphi'\sum_{i=1}^n F^{i\bar i}\lambda_i^2\\ &\qquad +\Big(-\psi'+\frac{1}{2}\varphi'-22n^2C_2-4C_3\Big)\sum_{i=1}^n F^{i\bar i}-C_1\\ &= A_1+A_2+A_3+A_4+A_5+A_6\\ &\qquad +\Big(-\psi'+\frac{1}{2}\varphi'-22n^2C_2-4C_3\Big)\sum_{i=1}^n F^{i\bar i}-C_1,\end{aligned}\tag{4-17}$$

where C_1 is a positive constant depending only on C_0, $\sup e^{f/k}$, $\sup|\nabla(e^{f/k})|^2$ and $\sup|\partial\bar\partial(e^{f/k})|$.

Let $\varepsilon=\frac{1}{4}\delta\leq\frac{1}{16}$ and $\delta=1/(2A+1)$, where $A=2L(C_0+1)$ and $C_0=31n^2C_2+4C_3$. We divide into two cases to derive the estimate, which is similar to [Hou et al. 2010].

Case 1: $\lambda_n<-\varepsilon\lambda_1$.

By condition (4-7), for $1\leq i\leq n$, we have

$$\begin{aligned}-(1+2c_0)^2\Big|\frac{u_{1\bar 1 i}}{w_{1\bar 1}}\Big|^2 &= -|\varphi'|\nabla u|_i^2+\psi' u_i|^2\geq -2(\varphi')^2|\nabla u|_i^2|\nabla u|_{\bar i}^2-2(\psi')^2|u_i|^2\\ &= -\varphi''|\nabla u|_i^2|\nabla u|_{\bar i}^2-2(\psi')^2|u_i|^2.\end{aligned}$$

This gives

$$
\begin{aligned}
A_2 &= 2(1+2c_0)\sum_{i\neq 1} F^{i\bar{i}} \operatorname{Re}\Big(\frac{T^1_{1i}u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big) \\
&\geq -c_0 \sum_{i\neq 1} F^{i\bar{i}}\Big|\frac{u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big|^2 - \frac{(1+2c_0)^2}{c_0}\sum_{i\neq 1} F^{i\bar{i}}|T^1_{1i}|^2 \\
&\geq -c_0 \sum_{i\neq 1} F^{i\bar{i}}\Big|\frac{u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big|^2 - 9n^2 C_2 \sum_{i\neq 1} F^{i\bar{i}}|T^1_{1i}|^2
\end{aligned}
$$

Thus

$$
\begin{aligned}
A_2 + A_3 &\geq -(1+3c_0)\sum_{i=1}^{n} \frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w^2_{1\bar{1}}} - 9n^2 C_2 \sum_{i\neq 1} F^{i\bar{i}}|T^1_{1i}|^2 \\
&\geq -(1+2c_0)^2 \sum_{i=1}^{n} \frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w^2_{1\bar{1}}} - 9n^2 C_2 \sum_{i=1}^{n} F^{i\bar{i}} \\
&= -A_4 - 2(\psi')^2 \sum_{i=1}^{n} F^{i\bar{i}}|u_i|^2 - 9n^2 C_2 \sum_{i=1}^{n} F^{i\bar{i}}.
\end{aligned}
$$

We therefore obtain

$$
A_2 + A_3 + A_4 \geq -2(\psi')^2 \sum_{i=1}^{n} F^{i\bar{i}}|u_i|^2 - 9n^2 C_2 \sum_{i=1}^{n} F^{i\bar{i}}. \tag{4-18}
$$

Using the inequality

$$
\sum_{i=1}^{n} F^{i\bar{i}}\lambda_i^2 \geq F^{n\bar{n}}\lambda_n^2 > \varepsilon^2 F^{n\bar{n}}\lambda_1^2 \geq \frac{\varepsilon^2}{n}\sum_{i=1}^{n} F^{i\bar{i}}\lambda_1^2,
$$

we have

$$
A_6 = \frac{1}{2}\varphi' \sum_{i=1}^{n} F^{i\bar{i}}\lambda_i^2 \geq \frac{\varepsilon^2}{2n}\varphi' \sum_{i=1}^{n} F^{i\bar{i}}\lambda_1^2. \tag{4-19}
$$

Combining (4-17) and (4-18) (4-19), we obtain

$$
\begin{aligned}
0 \geq \sum_{i=1}^{n} F^{i\bar{i}} Q_{i\bar{i}} &\geq \frac{\varepsilon^2}{2n}\varphi' \sum_{i=1}^{n} F^{i\bar{i}}\lambda_1^2 - 2(\psi')^2 \sum_{i=1}^{n} F^{i\bar{i}}|u_i|^2 \\
&\qquad + \Big(-\psi' + \frac{1}{2}\varphi' - 31n^2 C_2 - 4C_3\Big)\sum_{i=1}^{n} F^{i\bar{i}} - C_1 \\
&\geq \Big(\frac{\varepsilon^2}{8nK}\lambda_1^2 - 8K(C_0+1)^2\Big)\sum_{i=1}^{n} F^{i\bar{i}} - C_1 \\
&\geq \frac{\varepsilon^2}{8nK}\lambda_1^2 - 8K(C_0+1)^2 - C_1,
\end{aligned}
$$

where we use the fact that $\sum_{i=1}^{n} F^{i\bar{i}} \geq 1$, which follows from the definition of F^{ii} and the Newton–Maclaurin inequality.

Hence, we obtain

$$\lambda_1 \leq 8\sqrt{2}(2A+1)\sqrt{nK(8K(C_0+1)^2+C_1)} \leq CK.$$

Case 2: $\lambda_n > -\varepsilon\lambda_1$.

Let $I = \{i \in \{1,\ldots,n\} | \sigma_{k-1}(\lambda|i) \geq \varepsilon^{-1}\sigma_{k-1}(\lambda|1)\}$. Obviously, $1 \notin I$ and $i \in I$ if and only if $F^{i\bar{i}} > \varepsilon^{-1}F^{1\bar{1}}$. We first treat those indices which are not in I. By (4-7), we have

$$\begin{aligned}
-(1+2c_0)&\sum_{i\notin I}\frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w_{1\bar{1}}^2} + 2(1+2c_0)\sum_{i\notin I}F^{i\bar{i}}\operatorname{Re}\frac{T_{1i}^1 u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}} \\
&\geq -(1+2c_0)^2\sum_{i\notin I}\frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w_{1\bar{1}}^2} - \frac{(1+2c_0)^2}{c_0}\sum_{i\notin I}F^{i\bar{i}}|T_{1i}^1|^2 \\
&= -\varphi''\sum_{i\notin I}F^{i\bar{i}}|\nabla u|^2{}_i|\nabla u|^2{}_{\bar{i}} - 2(\psi')^2\sum_{i\notin I}F^{i\bar{i}}|u_i|^2 - 9n^2C_2\sum_{i\notin I}F^{i\bar{i}}|T_{1i}^1|^2 \\
&\geq -\varphi''\sum_{i\notin I}F^{i\bar{i}}|\nabla u|^2{}_i|\nabla u|^2{}_{\bar{i}} - 2\varepsilon^{-1}K(\psi')^2F^{1\bar{1}} - 9n^2C_2\sum_{i=1}^{n}F^{i\bar{i}}.
\end{aligned}$$

Substituting the above inequality into (4-17) yields

$$\begin{aligned}
0 &\geq F^{i\bar{i}}Q_{i\bar{i}} \qquad\qquad (4\text{-}20)\\
&\geq -(1+2c_0)\sum_{i,j,p,q=1}^{n}\frac{F^{i\bar{j},p\bar{q}}u_{i\bar{j}1}u_{p\bar{q}\bar{1}}}{w_{1\bar{1}}} + 2(1+2c_0)\sum_{i\in I}F^{i\bar{i}}\operatorname{Re}\Big(\frac{T_{1i}^1 u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}}\Big) \\
&\qquad - (1+2c_0)\sum_{i\in I}\frac{F^{i\bar{i}}|u_{1\bar{1}i}|^2}{w_{1\bar{1}}^2} + \varphi''\sum_{i\in I}F^{i\bar{i}}|\nabla u|_i^2|\nabla u|_{\bar{i}}^2 + \psi''\sum_{i=1}^{n}F^{i\bar{i}}|u_i|^2 \\
&\qquad + \frac{1}{2}\varphi'\sum_{i=1}^{n}F^{i\bar{i}}\lambda_i^2 - 2\varepsilon^{-1}K(\psi')^2F^{1\bar{1}} \\
&\qquad + \Big(-\psi' + \frac{1}{2}\varphi' - 31n^2C_2 - 4C_3\Big)\sum_{i=1}^{n}F^{i\bar{i}} - C_1 \\
&= B_1+B_2+B_3+B_4+B_5+B_6+B_7+B_8.
\end{aligned}$$

Firstly, we have

$$B_6+B_7 = \frac{1}{2}\varphi'\sum_{i=1}^{n}F^{i\bar{i}}\lambda_i^2 - 2\varepsilon^{-1}K(\psi')^2F^{1\bar{1}} \geq \frac{1}{4}\varphi'\sum_{i=1}^{n}F^{i\bar{i}}\lambda_i^2,$$

where we assume that $\frac{1}{4}\varphi' F^{1\bar{1}}\lambda_1^2 \geq 2\varepsilon^{-1}K(\psi')^2 F^{1\bar{1}}$ (for otherwise $\frac{1}{4}\varphi' F^{1\bar{1}}\lambda_1^2 \leq 2\varepsilon^{-1}K(\psi')^2 F^{1\bar{1}}$, i.e., $\lambda_1 \leq CK$).

We next use B_1 to cancel the other terms containing the third derivatives of u.

As the proof in [Hou et al. 2010, p. 559], we have

$$\lambda_1\sigma_{k-2}(\lambda|1i) \geq (1-2\varepsilon)\sigma_{k-1}(\lambda|i) \quad \text{for } i \in I.$$

Then

$$-\lambda_1 F^{i\bar{1},1\bar{i}} = \frac{F^{1-k}}{k}\lambda_1\sigma_{k-2}(\lambda|1i) \geq \frac{F^{1-k}}{k}(1-2\varepsilon)\sigma_{k-1}(\lambda|i) = (1-2\varepsilon)F^{i\bar{i}}.$$

Since $u_{i\bar{1}1} = u_{1\bar{1}i} - T^1_{1i}(\lambda_1 - 1)$, we get

$$\begin{aligned} B_1 &= -\frac{1+2c_0}{\lambda_1}\sum_{i,j,p,q=1}^{n} F^{i\bar{j},p\bar{q}} u_{i\bar{j}1}u_{p\bar{q}\bar{1}} \\ &\geq -\frac{1+2c_0}{\lambda_1^2}\sum_{i\in I}\lambda_1 F^{i\bar{1},1\bar{i}} u_{i\bar{1}1}u_{1\bar{i}\bar{1}} \\ &\geq \frac{1+2c_0}{\lambda_1^2}(1-2\varepsilon)\sum_{i\in I} F^{i\bar{i}}|u_{1\bar{1}i} - T^1_{1i}(\lambda_1-1)|^2, \end{aligned}$$

and

$$B_2 = \frac{2(1+2c_0)}{\lambda_1}\sum_{i\in I} F^{i\bar{i}}\,\mathrm{Re}(T^1_{1i}u_{1\bar{1}\bar{i}}).$$

From (4-7), we have

$$\begin{aligned} B_4 &= \varphi''\sum_{i\in I} F^{i\bar{i}}|\nabla u|^2_i|\nabla u|^2_{\bar{i}} \\ &= 2\sum_{i\in I} F^{i\bar{i}}\left|(1+2c_0)\frac{u_{1\bar{1}\bar{i}}}{w_{1\bar{1}}} + \psi' u_i\right|^2 \\ &\geq 2(1+2c_0)^2\delta\sum_{i\in I} F^{i\bar{i}}\frac{|u_{1\bar{1}\bar{i}}|^2}{w^2_{1\bar{1}}} - \frac{2\delta}{1-\delta}(\psi')^2\sum_{i\in I} F^{i\bar{i}}|u_i|^2 \\ &\geq 2(1+2c_0)^2\delta\sum_{i\in I} F^{i\bar{i}}\frac{|u_{1\bar{1}\bar{i}}|^2}{w^2_{1\bar{1}}} - B_5, \end{aligned}$$

where we use $(2\delta/(1-\delta))(\psi')^2 = \psi''$ by choosing $\delta = 1/(2A+1)$.

So we get

$$B_3 + B_4 + B_5 \geq -(1+2c_0)\frac{[1-2(1+2c_0)\delta]}{\lambda_1^2}\sum_{i\in I} F^{i\bar{i}}|u_{1\bar{1}\bar{i}}|^2.$$

Then we conclude

$$
\begin{aligned}
&B_1+B_2+B_3+B_4+B_5\\
&\qquad \geq \frac{1+2c_0}{\lambda_1^2}(1-2\varepsilon)\sum_{i\in I}F^{i\bar i}|u_{1\bar 1 i}-T^1_{1i}(\lambda_1-1)|^2\\
&\qquad\quad -(1+2c_0)\frac{[1-2(1+2c_0)\delta]}{\lambda_1^2}\sum_{i\in I}F^{i\bar i}|u_{1\bar 1\bar i}|^2\\
&\qquad\quad +\frac{2(1+2c_0)}{\lambda_1^2}\sum_{i\in I}F^{i\bar i}\,\mathrm{Re}(\lambda_1 T^1_{1i}u_{1\bar 1\bar i})\\
&\qquad =\frac{1+2c_0}{\lambda_1^2}\sum_{i\in I}F^{i\bar i}\{(1-2\varepsilon)|u_{1\bar 1 i}-T^1_{1i}(\lambda_1-1)|^2\\
&\qquad\quad -(1-2(1+2c_0)\delta)|u_{1\bar 1 i}|^2+2\,\mathrm{Re}(\lambda_1 T^1_{1i}u_{1\bar 1\bar i})\}\\
&\qquad =\frac{1+2c_0}{\lambda_1^2}\sum_{i\in I}F^{i\bar i}\{(2(1+2c_0)\delta-2\varepsilon)|u_{1\bar 1 i}|^2\\
&\qquad\quad +2[2\varepsilon(\lambda_1-1)+1]\,\mathrm{Re}(T^1_{1i}u_{1\bar 1\bar i})+(1-2\varepsilon)(\lambda_1-1)^2|T^1_{1i}|^2\}\\
&\qquad \geq 0,
\end{aligned}
$$

where the last inequality holds if we choose $\varepsilon=\frac{1}{4}\delta\leq\frac{1}{16}$. In fact,

$$
\begin{aligned}
\Delta=B^2-4AC&=4[2\varepsilon(\lambda_1-1)+1]^2-4(1-2\varepsilon)(\lambda_1-1)^2(2(1+2c_0)\delta-2\varepsilon)\\
&\leq 36\varepsilon^2(\lambda_1-1)^2-4(1-2\varepsilon)(\lambda_1-1)^2(2(1+2c_0)\delta-2\varepsilon)\\
&\leq 4(\lambda_1-1)^2(9\varepsilon^2-2(1-2\varepsilon)((1+2c_0)\delta)+2\varepsilon(1-2\varepsilon))\\
&\leq 4(\lambda_1-1)^2(5\varepsilon^2+2\varepsilon-\delta)\\
&\leq 4(\lambda_1-1)^2(4\varepsilon-\delta)\\
&=0.
\end{aligned}
$$

Then we finally obtain

$$
\begin{aligned}
0&\geq\frac{1}{4}\varphi'\sum_{i=1}^n F^{i\bar i}|u_{i\bar i}|^2+\Big(-\psi'+\frac{1}{2}\varphi'-C_2-C_3\Big)\sum_{i=1}^n F^{i\bar i}-C_1\\
&=\Big(-\psi'+\frac{1}{2}\varphi'-C_2-C_3\Big)\sum_{i=1}^n F^{i\bar i}+\frac{1}{4}\varphi'\sum_{i=1}^n F^{i\bar i}|u_{i\bar i}|^2-C_1\\
&\geq\sum_{i=1}^n F^{i\bar i}+\frac{1}{16K}\sum_{i=1}^n F^{i\bar i}{\lambda_i}^2-C_1,
\end{aligned}
$$

where we use $-\psi'\geq C_0+1$ and $C_0=31n^2C_2+4C_3$.

In particular, $\sum_{i=1}^{n} F^{i\bar{i}} \le C$. By Lemma 2.2 in [Hou et al. 2010], we have $F^{1\bar{1}} \ge c(n,k)/C_1^{k-1}$, where $c(n,k)$ is a positive constant depending only on n and k. Then we get the desired estimate

$$\lambda_1 \le \frac{4C_1^{k/2}}{c(n,k)^{1/2}}\sqrt{K},$$

where C_1 is given in (4-17). □

Acknowledgment

We would like to thank Professor Xinan Ma for suggesting the problem and his great encouragement. We thank Professor Shengli Kong and Professor Chuanqiang Chen for careful reading and suggestions. We are also grateful to the referees for their careful reading and many helpful comments.

References

[Cherrier 1987] P. Cherrier, "Équations de Monge–Ampère sur les variétés hermitiennes compactes", *Bull. Sci. Math.* (2) **111**:4 (1987), 343–385. MR Zbl

[Chou and Wang 2001] K.-S. Chou and X.-J. Wang, "A variational theory of the Hessian equation", *Comm. Pure Appl. Math.* **54**:9 (2001), 1029–1064. MR Zbl

[Dinew and Kołodziej 2017] S. Dinew and S. Kołodziej, "Liouville and Calabi–Yau type theorems for complex Hessian equations", *Amer. J. Math.* **139**:2 (2017), 403–415. MR Zbl

[Fu and Yau 2008] J.-X. Fu and S.-T. Yau, "The theory of superstring with flux on non-Kähler manifolds and the complex Monge–Ampère equation", *J. Differential Geom.* **78**:3 (2008), 369–428. MR Zbl

[Guan and Li 2010] B. Guan and Q. Li, "Complex Monge–Ampère equations and totally real submanifolds", *Adv. Math.* **225**:3 (2010), 1185–1223. MR Zbl

[Hou et al. 2010] Z. Hou, X.-N. Ma, and D. Wu, "A second order estimate for complex Hessian equations on a compact Kähler manifold", *Math. Res. Lett.* **17**:3 (2010), 547–561. MR Zbl

[Lin and Trudinger 1994] M. Lin and N. S. Trudinger, "On some inequalities for elementary symmetric functions", *Bull. Austral. Math. Soc.* **50**:2 (1994), 317–326. MR Zbl

[Phong et al. 2016a] D. H. Phong, S. Picard, and X. Zhang, "The Fu–Yau equation with negative slope parameter", *Invent. Math.* (online publication December 2016).

[Phong et al. 2016b] D. H. Phong, S. Picard, and X. Zhang, "On estimates for the Fu–Yau generalization of a Strominger system", *J. Reine Angew. Math.* (online publication October 2016).

[Phong et al. 2017] D. H. Phong, S. Picard, and X. Zhang, "Geometric flows and Strominger systems", *Math. Z.* (online publication March 2017).

[Sun 2017] W. Sun, "On uniform estimate of complex elliptic equations on closed Hermitian manifolds", *Commun. Pure Appl. Anal.* **16**:5 (2017), 1553–1570. MR Zbl

[Székelyhidi 2015] G. Székelyhidi, "Fully non-linear elliptic equations on compact Hermitian manifolds", preprint, 2015. To appear in *J. Differential Geom.* arXiv

[Székelyhidi et al. 2015] G. Székelyhidi, V. Tosatti, and B. Weinkove, "Gauduchon metrics with prescribed volume form", preprint, 2015. arXiv

[Tosatti and Weinkove 2010] V. Tosatti and B. Weinkove, "The complex Monge–Ampère equation on compact Hermitian manifolds", *J. Amer. Math. Soc.* **23**:4 (2010), 1187–1195. MR Zbl

[Tosatti and Weinkove 2013] V. Tosatti and B. Weinkove, "Hermitian metrics, $(n-1,\ n-1)$ forms and Monge–Ampère equations", preprint, 2013. arXiv

[Tosatti and Weinkove 2015] V. Tosatti and B. Weinkove, "On the evolution of a Hermitian metric by its Chern–Ricci form", *J. Differential Geom.* **99**:1 (2015), 125–163. MR Zbl

[Tosatti and Weinkove 2017] V. Tosatti and B. Weinkove, "The Monge–Ampère equation for $(n-1)$-plurisubharmonic functions on a compact Kähler manifold", *J. Amer. Math. Soc.* **30**:2 (2017), 311–346. MR Zbl

[Wang 2009] X.-J. Wang, "The k-Hessian equation", pp. 177–252 in *Geometric analysis and PDEs*, edited by A. Ambrosetti et al., Lecture Notes in Math. **1977**, Springer, Dordrecht, 2009. MR Zbl

[Yau 1978] S. T. Yau, "On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I", *Comm. Pure Appl. Math.* **31**:3 (1978), 339–411. MR Zbl

[Zhang 2010] X. Zhang, "A priori estimates for complex Monge–Ampère equation on Hermitian manifolds", *Int. Math. Res. Not.* **2010**:19 (2010), 3814–3836. MR Zbl

Received October 30, 2015. Revised January 18, 2017.

DEKAI ZHANG
SCHOOL OF MATHEMATICAL SCIENCE
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, ANHUI
CHINA

dekzhang@mail.ustc.edu.cn

CONTENTS

Volume 291, no. 1 and no. 2

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

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.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 291 No. 2 December 2017

0030-8730(201712)291:2;1-W