Pacific Journal of Mathematics

Volume 260 No. 1 November 2012

PACIFIC JOURNAL OF MATHEMATICS

http://pacificmath.org

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

See inside back cover or pacificmath.org for submission instructions.

The subscription price for 2012 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. 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
http://msp.org/
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.1

THE DECOMPOSITION OF GLOBAL CONFORMAL INVARIANTS: SOME TECHNICAL PROOFS II

SPYROS ALEXAKIS

This paper complements our research monograph *The decomposition of global conformal invariants* (Princeton University Press, 2012) in proving a conjecture of Deser and Schwimmer regarding the algebraic structure of "global conformal invariants"; these are defined to be conformally invariant integrals of geometric scalars. The conjecture asserts that the integrand of any such integral can be expressed as a linear combination of a local conformal invariant, a divergence and of the Chern–Gauss–Bonnet integrand.

The present paper provides a proof of certain purely algebraic statements announced in our previous work and whose rather technical proof was deferred to this paper; the lemmas proven here serve to reduce "main algebraic propositions" to certain technical inductive statements.

1. Introduction

This paper complements our earlier work [A 2010, 2011, 2012] in proving a conjecture of Deser and Schwimmer [1993] regarding the algebraic structure of "global conformal invariants". It provides a (rather technical) proof of certain lemmas announced in [A 2010, 2012].

MSC2010: primary 53A55, 53B20; secondary 53A30.
Keywords: conformal invariant, global conformal invariant, Deser–Schwimmer conjecture, Riemannian invariant.

We recall that a global conformal invariant is an integral of a natural scalar-valued function of Riemannian metrics, $\int_{M^n} P(g)\, dV_g$, which remains invariant under conformal rescalings of the underlying metric.[1] More precisely, $P(g)$ is assumed to be a linear combination, $P(g) = \sum_{l\in L} a_l C^l(g)$, where each $C^l(g)$ is a complete contraction in the form

$$contr^l\big(\nabla^{(m_1)} R \otimes \cdots \otimes \nabla^{(m_s)} R\big); \tag{1-1}$$

here each factor $\nabla^{(m)} R$ stands for the m^{th} iterated covariant derivative of the curvature tensor R, ∇ is the Levi-Civita connection of the metric g and R is the curvature associated to this connection. The contractions are taken with respect to the quadratic form g^{ij}. In the present paper, along with [A 2011, 2012], we prove:

Theorem. *Assume that $P(g) = \sum_{l\in L} a_l C^l(g)$, where each $C^l(g)$ is a complete contraction in the form (1-1), with weight $-n$. Assume that for every closed Riemannian manifold (M^n, g) and every $\phi \in C^\infty(M^n)$,*

$$\int_{M^n} P(e^{2\phi} g)\, dV_{e^{2\phi} g} = \int_{M^n} P(g)\, dV_g.$$

We claim that $P(g)$ can then be expressed in the form

$$P(g) = W(g) + \mathrm{div}_i T^i(g) + \mathrm{Pfaff}(R_{ijkl}).$$

Here $W(g)$ stands for a local conformal invariant of weight $-n$ (meaning that $W(e^{2\phi} g) = e^{-n\phi} W(g)$ for every $\phi \in C^\infty(M^n)$), $\mathrm{div}_i\, T^i(g)$ is the divergence of a Riemannian vector field of weight $-n+1$, and $\mathrm{Pfaff}(R_{ijkl})$ is the Pfaffian of the curvature tensor.

Before we discuss the position of the present paper in the entire work [A 2010, 2011, 2012] we digress to describe the relation between the present series of papers with classical and recent work on scalar local invariants in various geometries.

Broad discussion. The theory of *local* invariants of Riemannian structures (and indeed, of more general geometries, such as conformal, projective, or CR) has a long history. As discussed in [A 2012], the original foundations of this field were laid in the work of Hermann Weyl and Élie Cartan; see [Weyl 1939; Cartan 1896]. The task of writing out local invariants of a given geometry is intimately connected with understanding polynomials in a space of tensors with given symmetries; these polynomials are required to remain invariant under the action of a Lie group on the components of the tensors. In particular, the problem of writing down all local Riemannian invariants reduces to understanding the invariants of the orthogonal group.

[1] See the introduction of [A 2012] for a detailed discussion of the Deser–Schwimmer conjecture, and for background on scalar Riemannian invariants.

In more recent times, a major program was laid out by C. Fefferman [1976] aimed at finding all scalar local invariants in CR geometry. This was motivated by the problem of understanding the local invariants that appear in the asymptotic expansion of the Bergman and Szegő kernels of strictly pseudoconvex CR manifolds, in a similar way to Riemannian invariants that appear in the asymptotic expansion of the heat kernel; the study of the local invariants in the singularities of these kernels led to important breakthroughs in [Bailey et al. 1994b] and more recently by Hirachi [2000]. This program was later extended to conformal geometry in [Fefferman and Graham 1985]. Both these geometries belong to a broader class of structures, the *parabolic geometries*; these admit a principal bundle whose structure group is a parabolic subgroup P of a semisimple Lie group G, and a Cartan connection on that principle bundle (see the introduction in [Čap and Gover 2002]). An important question in the study of these structures is the problem of constructing all their local invariants, which can be thought of as the *natural, intrinsic* scalars of these structures.

In the context of conformal geometry, the first (modern) landmark in understanding *local conformal invariants* was the work of Fefferman and Graham [1985], where they introduced the *ambient metric*. This allows one to construct local conformal invariants of any order in odd dimensions, and up to order $\frac{n}{2}$ in even dimensions. The question is then whether *all* invariants arise via this construction.

The subsequent work of Bailey–Eastwood–Graham [1994b] proved that this is indeed true in odd dimensions; in even dimensions, they proved that the result holds when the weight (in absolute value) is bounded by the dimension. The ambient metric construction in even dimensions was recently extended by Graham and Hirachi [2008]; this enables them to identify in a satisfactory way *all* local conformal invariants, even when the weight (in absolute value) exceeds the dimension.

An alternative construction of local conformal invariants can be obtained via the *tractor calculus* introduced by Bailey et al. [1994a]. This construction bears a strong resemblance to the Cartan conformal connection, and to the work of T.Y. Thomas [1934]. The tractor calculus has proven to be very universal; tractor bundles have been constructed [Čap and Gover 2002] for an entire class of parabolic geometries. The relation between the conformal tractor calculus and the Fefferman–Graham ambient metric has been elucidated in [Čap and Gover 2003].

The present paper, along with [A 2010, 2011, 2012], while pertaining to the question above (given that it ultimately deals with the algebraic form of local *Riemannian* and *conformal* invariants), nonetheless addresses a different *type* of problem: We here consider Riemannian invariants $P(g)$ for which the *integral* $\int_{M^n} P(g)\, dV_g$ remains invariant under conformal changes of the underlying metric; we then seek to understand the possible algebraic form of the *integrand* $P(g)$, ultimately proving that it can be decomposed in the way that Deser and Schwimmer asserted. It is thus not surprising that the prior work on the construction and understanding of local *conformal* invariants, in [A 2011] and in the second chapter of [A 2012], plays a central role in this endeavor.

On the other hand, a central element of our proof are the "Main algebraic propositions" 2.28, 3.27, 3.28 in [A 2012]; these deal *exclusively* with algebraic properties of the *classical* scalar Riemannian invariants. (These "main algebraic propositions" are discussed in brief below. A *generalization* of these propositions is the Proposition 1.1 below). The "Fundamental proposition 1.1" makes no reference to integration; it is purely a statement concerning *local Riemannian invariants*. Thus, while the author was led to led to the main algebraic propositions in [A 2012] out of the strategy that he felt was necessary to solve the Deser–Schwimmer conjecture, they can be thought of as results of an independent interest. The *proof* of these propositions, presented in the second part of [A 2012] (and certain claims announced there proven in the present paper), is in fact not particularly intuitive. It is the author's sincere hope that deeper insight (and hopefully a more intuitive proof) will be obtained in the future as to *why* these algebraic propositions hold.

Let us now discuss the position of the present paper in the entire work [A 2010, 2011, 2012] in more detail: In the first part of [A 2012] (complemented by [A 2011]) we proved that the Deser–Schwimmer conjecture holds, *provided* one can show certain "main algebraic propositions," announced in Chapters 2 and 3 in [A 2012]. In [A 2010] (which is reproduced in Chapter 4 of [A 2012] — for convenience we refer to propositions and lemmas in [A 2010]; the same propositions can be found in [A 2012] with different numbering) we claimed a more general proposition which *implies* the "main algebraic propositions;" this new "Fundamental proposition" 2.1 in [A 2010][2] is to be proven by an induction of four parameters. In [A 2010] we also reduced the inductive step of Proposition 2.1 to three lemmas (in particular we distinguished Cases I, II, III of Proposition 2.1 by examining the tensor fields appearing in its hypothesis, see (1-7) below; Lemmas 3.1, 3.2, 3.5 in [A 2010][3] correspond to these three cases). We proved that these three lemmas *imply* the inductive step of the Fundamental proposition in Cases I, II, III respectively, apart from certain *special cases* which were deferred to the present paper. In these special cases we will derive Proposition 2.1 in [A 2010] directly,[4] in Section 3. Now, in *proving* that the inductive step of Proposition 1.1 follows from Lemmas 3.1, 3.2, 3.5 in [A 2010] we asserted certain *technical lemmas*, whose proof was deferred to the present paper. These were Lemmas 4.6, 4.8, and 4.7, 4.9

[2]This is reproduced as Proposition 4.13 in [A 2012].

[3]These correspond respectively to Lemmas 4.16, 4.19 and 4.24 in [A 2012].

[4]By this we mean *without* recourse to the Lemmas 3.1, 3.2, 3.5 in [A 2010].

in [A 2010];[5] also, the proof of Lemma A.1 in [A 2010] was deferred to the present paper. We prove all these lemmas from [A 2010] in Section 2.

For reference purposes, and for the reader's convenience, we recall the precise formulation of the "Fundamental proposition" 2.1 in [A 2010], referring the reader to [A 2010] for a definition of many of the terms appearing in the formulation. First however, we will recall (schematically) the first "Main algebraic proposition" 2.28 in [A 2012]; this is a *special case* of Proposition 2.1 in [A 2010], and provides a simpler version of it.

A simpler version of Proposition 2.1 in [A 2010]. Given a Riemannian metric g over an n-dimensional manifold M^n and auxiliary C^∞ scalar-valued functions $\Omega_1, \dots, \Omega_p$ defined over M^n, the objects of study are linear combinations of tensor fields $\sum_{l\in L} a_l C^{l,i_1\dots i_\alpha}_g$, where each $C^{l,i_1\dots i_\alpha}_g$ is a *partial contraction* with α free indices, in the form

$$pcontr\big(\nabla^{(m)} R\otimes\cdots\otimes\nabla^{(m_s)} R\otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_m)}\Omega_p\big); \tag{1-2}$$

here $\nabla^{(m)} R$ stands for the m^{th} covariant derivative of the curvature tensor R,[6] and $\nabla^{(b)}\Omega_h$ stands for the b^{th} covariant derivative of the function Ω_h. A *partial contraction* means that we have list of pairs of indices $({}_a, {}_b), \dots, ({}_c, {}_d)$ in (1-2), that are contracted against each other via the metric g^{ij}. The remaining indices (which are not contracted against another index in (1-2)) are the *free indices* ${}_{i_1}, \dots, {}_{i_\alpha}$.

The "Main algebraic proposition" 2.28 in [A 2012] (roughly) asserts the following: Let $\sum_{l\in L_\mu} a_l C^{l,i_1\dots i_\mu}_g$ stand for a linear combination of partial contractions in the form (1-2), where each $C^{l,i_1\dots i_\mu}_g$ has a given number σ_1 of factors and a given number p of factor $\nabla^{(b)}\Omega_h$. Assume also that $\sigma_1+p\ge 3$, each $b_i\ge 2$,[7] and that for each contracting pair of indices $({}_a, {}_b)$ in any given $C^{l,i_1\dots i_\mu}_g$, the indices ${}_a, {}_b$ do not belong to the same factor. Assume also the rank $\mu>0$ is fixed and each partial contraction $C^{l,i_1\dots i_\mu}_g$, $l\in L_\mu$ has a given *weight* $-n+\mu$.[8] Let also $\sum_{l\in L_{>\mu}} a_l C^{l,i_1\dots i_{y_l}}_g$ stand for a (formal) linear combination of partial contractions of weight $-n+y_l$, with all the properties of the terms indexed in L_μ, *except* that now all the partial contractions have a different rank y_l, and each $y_l>\mu$.

Assume also that the local equation

$$\sum_{l\in L_\mu} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\mu} C^{l,i_1\dots i_\mu}_g + \sum_{l\in L_{>\mu}} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_{y_l}} C^{l,i_1\dots i_{y_l}}_g = 0 \tag{1-3}$$

[5] These correspond to Lemmas 4.35, 4.41, 4.37 and 4.42 in [A 2012].

[6] In particular it is a tensor of rank $m+4$; if we write out its free indices it would be in the form $\nabla^{(m)}_{r_1\dots r_m} R_{ijkl}$.

[7] $b_i\ge 2$ means that the function Ω_i is differentiated at least twice.

[8] See [A 2012] for a precise definition of weight.

holds *modulo* complete contractions with $\sigma+1$ factors. Here, given a partial contraction $C^{l,i_1\dots i_\alpha}_g$ in the form (1-2) $X\,\mathrm{div}_{i_s}[C^{l,i_1\dots i_\alpha}_g]$ stands for sum of $\sigma-1$ terms in $\mathrm{div}_{i_s}[C^{l,i_1\dots i_\alpha}_g]$ where the derivative ∇^{i_s} is *not* allowed to hit the factor to which the free index ${}_{i_s}$ belongs.[9]

The "Main algebraic proposition" 2.28 in [A 2012] says that there exists a linear combination of partial contractions in the form (1-2), $\sum_{h\in H} a_h C^{h,i_1\dots i_{\mu+1}}_g$, with all the properties of the terms indexed in $L_{>\mu}$, and all with rank $(\mu+1)$, so that

$$\sum_{l\in L_1} a_l C^{l,(i_1\dots i_\mu)}_g + \sum_{h\in H} a_h X\,\mathrm{div}_{i_{\mu+1}}\, C^{l,(i_1\dots i_\mu)i_{\mu+1}}_g = 0; \tag{1-4}$$

the above holds modulo terms of length $\sigma+1$. The symbol $(\dots)$ means that we are *symmetrizing* over the indices between parentheses.

In [A 2010] we set up a multiple induction by which we will prove the "main algebraic propositions" in Chapters 2, 3 in [A 2012]. The inductive step is proven in the "Fundamental proposition" 2.1 in [A 2010], which we reproduce here in Proposition 1.1. This deals with tensor fields in the forms

$$\begin{aligned} &pcontr\big(\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_s)}R_{ijkl}\otimes\nabla^{(b_1)}\Omega_1 \\ &\qquad \otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla\phi_1\otimes\cdots\otimes\nabla\phi_u\big), \end{aligned} \tag{1-5}$$

$$\begin{aligned} &pcontr\big(\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_{\sigma_1})}R_{ijkl} \\ &\qquad \otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\cdots\otimes S_*\nabla^{(\nu_t)}R_{ijkl} \\ &\qquad \otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla\phi_{z_1}\otimes\cdots\otimes\nabla\phi_{z_w} \\ &\qquad \otimes\nabla\phi'_{z_{w+1}}\otimes\cdots\otimes\nabla\phi'_{z_{w+d}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{w+d+1}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{w+d+y}}\big). \end{aligned} \tag{1-6}$$

(See the introduction in [A 2010] for a detailed description of the above form.) In keeping with the conventions introduced in [A 2010], we remark that a complete or partial contraction in the above form will be called "acceptable" if each $b_i\ge 2$, for $1\le i\le p$.[10]

Proposition 1.1. *Consider two linear combinations of acceptable tensor fields in the form* (1-6),

$$\sum_{l\in L_\mu} a_l C^{l,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),$$

[9]Recall that given a partial contraction $C^{l,i_1\dots i_\alpha}_g$ in the form (1-2) with σ factors, $\mathrm{div}_{i_s} C^{l,i_1\dots i_\alpha}_g$ is a sum of σ partial contractions of rank $\alpha-1$. The first summand arises by adding a derivative ∇^{i_s} onto the first factor T_1 and then contracting the upper index ^{i_s} against the free index ${}_{i_s}$; the second summand arises by adding a derivative ∇^{i_s} onto the second factor T_2 and then contracting the upper index ^{i_s} against the free index ${}_{i_s}$ and so on.

[10]In other words, we are requiring each function Ω_i is differentiated at least twice.

$$\sum_{l\in L_{>\mu}} a_l C_g^{l,i_1\dots i_{\beta_l}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),$$

where each tensor field above has real length $\sigma\geq 3$ *and a given simple character* $\vec{\kappa}_{\text{simp}}$. *We assume that for each* $l\in L_{>\mu}$, $\beta_l\geq\mu+1$. *We also assume that none of the tensor fields of maximal refined double character in* L_μ *are "forbidden"* (*see Definition* 2.12 *in [A 2010]*).

We denote by

$$\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

a generic linear combination of complete contractions (*not necessarily acceptable*) *in the form* (1-5) *that are simply subsequent to* $\vec{\kappa}_{\text{simp}}$.[11] *We assume that*

$$\begin{aligned}\text{(1-7)}\quad \sum_{l\in L_\mu} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha}&\\ \times C_g^{l,i_1\dots i_\mu}&(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\\ +\sum_{l\in L_{>\mu}} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_{\beta_l}}&\\ \times C_g^{l,i_1\dots i_{\beta_l}}&(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\\ +\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&=0.\end{aligned}$$

We draw our conclusion with a little more notation: *We break the index set* L_μ *into subsets* L^z, $z\in Z$ (Z *is finite*), *with the rule that each* L^z *indexes tensor fields with the same refined double character, and conversely two tensor fields with the same refined double character must be indexed in the same* L^z. *For each index set* L^z, *we denote the refined double character in question by* $\vec{L}^z$. *Consider the subsets* L^z *that index the tensor fields of maximal refined double character.*[12] *We assume that the index set of those* z *is* $Z_{Max}\subset Z$.

We claim that for each $z\in Z_{Max}$ *there is some linear combination of acceptable* $(\mu{+}1)$-*tensor fields*,

$$\sum_{r\in R^z} a_r C_g^{r,i_1\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),$$

where each

$$C_g^{r,i_1\dots i_{\mu+1}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

[11] Of course if $Def(\vec{\kappa}_{\text{simp}})=\varnothing$ then by definition $\sum_{j\in J}\dots=0$.

[12] Note that in any set S of μ-refined double characters with the same simple character there is going to be a subset S' consisting of the maximal refined double characters.

has a μ-double character $\vec{L}^z_1$ and also the same set of factors $S_*\nabla^{(\nu)}R_{ijkl}$ as in $\vec{L}^z$ contain special free indices, so that

$$\begin{aligned}(1\text{-}8)\quad \sum_{l\in L^z} a_l C^{l,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\upsilon\dots\nabla_{i_\mu}\upsilon \\ -\sum_{r\in R^z} a_r X\,\mathrm{div}_{i_{\mu+1}} C^{r,i_1\dots i_{\mu+1}}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\upsilon\dots\nabla_{i_\mu}\upsilon \\ =\sum_{t\in T_1} a_t C^{t,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,,\phi_1,\dots,\phi_u)\nabla_{i_1}\upsilon\dots\nabla_{i_\mu}\upsilon,\end{aligned}$$

modulo complete contractions of length $\geq \sigma+u+\mu+1$. Here each

$$C^{t,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

is acceptable and is either simply or doubly subsequent to $\vec{L}^z$.[13]

(See the first section in [A 2010] for a description of the notions of *real length, acceptable tensor fields, simple character, refined double character, maximal refined double character, simply subsequent, strongly doubly subsequent.*) We prove Proposition 1.1 by a multiple induction on the parameters $-n$ (the *weight* of the complete contractions appearing in (1-7)), σ (the total number of factors in the form $\nabla^{(m)}R_{ijkl}$, $S_*\nabla^{(\nu)}R_{ijkl}$, $\nabla^{(A)}\Omega_h$ among the partial contractions in (1-7)),[14] Φ (the number of factors $\nabla\phi_1,\dots,\nabla\phi_u$ appearing in (1-7)), and $\sigma_1+\sigma_2$ (the total number of factors $\nabla^{(m)}R_{ijkl}$, $S_*\nabla^{(\nu)}R_{ijkl}$). When $\Phi=0$, Proposition 1.1 *coincides* with the "Main algebraic proposition" 2.28 in [A 2012] outlined above.[15]

2. Proof of the technical lemmas from [A 2010]

2A. ***Restatement of the technical Lemmas 4.6–4.9 from [A 2010].*** We start by recalling a definition from [A 2010] that will be used frequently in the present paper:

Definition. Consider any partial contraction in the form (1-6). We consider any set of indices, $\{_{x_1},\dots,_{x_s}\}$ belonging to a factor T, which is either in the form $\nabla^{(B)}\Omega_h$ or $\nabla^{(m)}R_{ijkl}$. We assume that these indices are not free and are not contracting against a factor $\nabla\phi_h$.

If the indices belong to a factor T in the form $\nabla^{(B)}\Omega_h$ then $\{_{x_1},\dots,_{x_s}\}$ are removable provided $B\geq s+2$.

[13] Recall that "simply subsequent" means that the simple character of $C^{t,i_1\dots i_\mu}_g$ is subsequent to $Simp(\vec{L}^z)$.

[14] The partial contractions in (1-7) are assumed to all have the same simple character–this implies that they all have the same number of factors $\nabla^{(m)}R_{ijkl}$, $S_*\nabla^{(\nu)}R_{ijkl}$, $\nabla^{(A)}\Omega_h$ respectively.

[15] Similarly, the "Main algebraic propositions" 3.27 and 3.28 in Chapter 3 of [A 2012] *coincide* with Proposition 1.1 above when $\Phi=1$.

Now, we consider indices that belong to a factor T in the form $\nabla^{(m)}R_{ijkl}$ (and are not free and do not contract against a factor $\nabla\phi_h$). Any such index ${}_x$ which is a derivative index will be removable. Furthermore, if T has at least two free derivative indices, then if neither of the indices ${}_i, {}_j$ are free we will say one of ${}_i, {}_j$ is removable; accordingly, if neither of ${}_k, {}_l$ is free then we will say that one of ${}_k, {}_l$ is removable. Moreover, if T has one free derivative index then: if none of the indices ${}_i, {}_j$ are free we will say that one of the indices ${}_i, {}_j$ is removable; on the other hand if one of the indices ${}_i, {}_j$ is also free and none of the indices ${}_k, {}_l$ are free then we will say that one of the indices ${}_k, {}_l$ is removable.

Now, we consider a set of indices $\{{}_{x_1}, \dots, {}_{x_s}\}$ that belong to a factor T in the form $S_*\nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}$; if $\{x_1, \dots, x_s\} \subset \{r_1 \dots r_\nu, j\}$ and none of them are free and none of them contract against a factor $\nabla\phi_x$, then we will say this set of indices is removable. Furthermore, we will say that the indices ${}_k, {}_l$ in such a factor are removable if neither ${}_k$ nor ${}_l$ is free and $\nu > 0$ and at least one of the indices ${}_{r_1}, \dots, {}_{r_\nu}, {}_j$ is free.

For the two Lemmas 2.1 and 2.2 we will consider tensor fields in the form

$$\begin{aligned}
\text{(2-1)}\quad pcontr\big(&\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_{\sigma_1})}R_{ijkl}\\
&\otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\cdots\otimes S_*\nabla^{(\nu_t)}R_{ijkl}\otimes\nabla Y\\
&\otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla\phi_{z_1}\cdots\otimes\nabla\phi_{z_w}\\
&\otimes\nabla\phi'_{z_{w+1}}\otimes\cdots\otimes\nabla\phi'_{z_{w+d}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{w+d+1}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{w+d+y}}\big).
\end{aligned}$$

(Notice this is the same as the form (1-6), but for the fact that we have inserted a factor ∇Y in the second line.) We recall that for a partial contraction $C^{i_1\dots i_h}_g$ in the above form, $X_*\,\mathrm{div}_{i_r}\,C^{i_1\dots i_h}_g$ stands for the sublinear combination in the divergence $\mathrm{div}_{i_r}\,C^{i_1\dots i_h}_g$ where the derivative ∇^{i_r} is not allowed to hit the factor to which ${}_{i_r}$ belongs, nor any factor $\nabla\phi_x$, nor the factor ∇Y.

The claims whose proof was deferred to the present paper are then as follows:

Lemma 2.1. *Assume that*

$$\begin{aligned}
\text{(2-2)}\quad \sum_{h\in H_2} a_h X_*\,\mathrm{div}_{i_1}\dots X_*\,\mathrm{div}_{i_{a_h}}\,C^{h,i_1\dots i_{a_h}}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'})&\\
=\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_{u'}),&
\end{aligned}$$

where all tensor fields have rank $a_h \geq \alpha$. All tensor fields have a given u-simple character $\vec{\kappa}'_{\mathrm{simp}}$, for which $\sigma \geq 4$. Moreover, we assume that if we formally treat the factor ∇Y as a factor $\nabla\phi_{u'+1}$ in the above equation, then the inductive assumption of Proposition 1.1 can be applied. (See Subsection 3.1 in [A 2010] for a strict discussion of the multiparameter induction by which we prove Proposition 1.1.)

The conclusion (under various assumptions which we will explain below) is: Denote by $H_{2,\alpha} \subset H_2$ the index set of tensor fields with rank α in (2-2).

We claim that there is a linear combination of acceptable[16] *tensor fields*

$$\sum_{d\in D} a_d C_g^{d,i_1\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u),$$

each with a simple character $\vec{\kappa}'_{\mathrm{simp}}$, *so that*

$$\sum_{h\in H_{2,\alpha}} a_h C_g^{h,i_1\dots i_\alpha}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'})\nabla_{i_1}\upsilon\dots\nabla_{i_\alpha}\upsilon - X_*\,\mathrm{div}_{i_{\alpha+1}}\sum_{d\in D} a_d C_g^{d,i_1\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'})\nabla_{i_1}\upsilon\dots\nabla_{i_\alpha}\upsilon = \sum_{t\in T} a_t C_g^t(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'},\upsilon^\alpha). \tag{2-3}$$

The linear combination on the right-hand side stands for a generic linear combination of complete contractions in the form (2-1) with a factor ∇Y *and with a simple character that is subsequent to* $\vec{\kappa}'_{\mathrm{simp}}$.

The assumption under which (2-3) holds is that there should be no tensor fields of rank α *in (2-2) that are "bad". Here "bad" means the following:*

If $\sigma_2 = 0$ *in* $\vec{\kappa}'_{\mathrm{simp}}$ *then a tensor field in the form (2-1) is "bad" provided:*

(1) *The factor* ∇Y *contains a free index.*

(2) *If we formally erase the factor* ∇Y *(which contains a free index), then the resulting tensor field should have no removable indices,*[17] *and no free indices.*[18] *Moreover, any factors* S_*R_{ijkl} *should be simple.*

If $\sigma_2 > 0$ *in* $\vec{\kappa}'_{\mathrm{simp}}$ *then a tensor field in the form (2-1) is "bad" provided:*

(1) *The factor* ∇Y *contains a free index.*

(2) *If we formally erase the factor* ∇Y *(which contains a free index), then the resulting tensor field should have no removable indices, any factors* S_*R_{ijkl} *should be simple, any factor* $\nabla^{(2)}_{ab}\Omega_h$ *should have at most one of the indices* $_a, _b$ *free or contracting against a factor* $\nabla\phi_s$.

(3) *Any factor* $\nabla^{(m)}R_{ijkl}$ *can contain at most one (necessarily special, by virtue of (2)) free index.*

Furthermore, we claim that the proof of this lemma will only rely on the inductive assumption of Proposition 1.1. Moreover, we claim that if none of the tensor fields indexed in H_2 *(in (2-2)) have a free index in* ∇Y, *then we may assume that none of the tensor fields indexed in* D *in (2-3) have that property either.*

---

[16] "Acceptable" in the sense that each factor Ω_i is differentiated at least twice).

[17] Thus, the tensor field should consist of factors S_*R_{ijkl}, $\nabla^{(2)}\Omega_h$, and factors $\nabla^{(m)}_{r_1\dots r_m}R_{ijkl}$ with all the indices $_{r_1},\dots,_{r_m}$ contracting against factors $\nabla\phi_h$.

[18] That is $\alpha = 1$ in (2-2).

Lemma 2.2. *We assume that (2-2) holds for $\sigma = 3$. We also assume that for each of the tensor fields in $H_2^{\alpha,*}$*[19] *there is at least one removable index. We then have two claims*:

First, *the conclusion of Lemma 2.1 holds in this setting. Second*, *we can write*

$$\sum_{h\in H_2} a_h X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha} C_g^{h,i_1\dots i_\alpha}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'}) \\ = \sum_{q\in Q} a_q X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_{a'}} C_g^{q,i_1\dots i_{a'}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'}) \\ + \sum_{t\in T} a_t C_g^t(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'}), \tag{2-4}$$

where the linear combination $\sum_{q\in Q} a_q C_g^{q,i_1\dots i_{a'}}$ stands for a generic linear combination of tensor fields in the form

$$pcontr\big(\nabla^{(m_1)}R_{ijkl}\otimes\dots\otimes\nabla^{(m_{\sigma_1})}R_{ijkl} \\ \otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\dots\otimes S_*\nabla^{(\nu_t)}R_{ijkl}\otimes\nabla^{(B)}Y \\ \otimes\nabla^{(b_1)}\Omega_1\otimes\dots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla\phi_{z_1}\dots\otimes\nabla\phi_{z_w} \\ \otimes\nabla\phi'_{z_{w+1}}\otimes\dots\otimes\nabla\phi'_{z_{w+d}}\otimes\dots\otimes\nabla\tilde{\phi}_{z_{w+d+1}}\otimes\dots\otimes\nabla\tilde{\phi}_{z_{w+d+y}}\big), \tag{2-5}$$

with $B\ge 2$, with a simple character $\vec{\kappa}'_{\mathrm{simp}}$ and with each $a'\ge\alpha$. The acceptable complete contractions $C_g^t(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_{u'})$ are simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$. $X\,\mathrm{div}_i$ here means that ∇_i is not allowed to hit the factors $\nabla\phi_h$ (but it is allowed to hit $\nabla^{(B)}Y$).

For our next two lemmas, we will be considering tensor fields in the general form

$$contr\big(\nabla^{(m_1)}R_{ijkl}\otimes\dots\otimes\nabla^{(m_s)}R_{ijkl} \\ \otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\dots\otimes S_*\nabla^{(\nu_b)}R_{ijkl} \\ \otimes\nabla^{(B,+)}_{r_1\dots r_B}(\nabla_a\omega_1\nabla_b\omega_2-\nabla_b\omega_1\nabla_a\omega_1) \\ \otimes\nabla^{(d_1)}\Omega_p\otimes\dots\otimes\nabla^{(d_p)}\Omega_p\otimes\nabla\phi_1\otimes\dots\otimes\nabla\phi_u\big); \tag{2-6}$$

here $\nabla^{(B,+)}_{r_1\dots r_B}(\dots)$ stands for the sublinear combination in $\nabla^{(B)}_{r_1\dots r_B}(\dots)$ where each derivative ∇_{r_i} is not allowed to hit the factor $\nabla\omega_2$.

We also recall from [A 2010] that $X_+\,\mathrm{div}_i$ stands for the sublinear combination in $X\,\mathrm{div}_i$ where ∇_i is in addition not allowed to hit the factor $\nabla\omega_2$ (it is allowed to hit the factor $\nabla^{(B)}\omega_1$).

[19] Recall from [A 2010] that $H_2^{\alpha,*}$ is the index set of tensor fields of rank α in (2-2) with a free index in the factor ∇Y.

Lemma 2.3. *Consider a linear combination of partial contractions,*

$$\sum_{x\in X} a_x C^{x,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_{u'}),$$

where each of the tensor fields $C^{x,i_1\dots i_a}_g$ is in the form (2-6) with $B=0$ (and is antisymmetric in the factors $\nabla_a\omega_1$, $\nabla_b\omega_2$ by definition), with rank $a\geq\alpha$ and real length $\sigma\geq 4$.[20] *We assume that all these tensor fields have a given simple character which we denote by $\vec{\kappa}'_{\mathrm{simp}}$ (we use u' instead of u to stress that this lemma holds in generality). We assume that*

$$\sum_{x\in X} a_x X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_a} C^{x,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u) + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)=0, \tag{2-7}$$

where $X_*\operatorname{div}_i$ stands for the sublinear combination in $X\operatorname{div}_i$ where ∇_i is in addition not allowed to hit the factors $\nabla\omega_1$, $\nabla\omega_2$. The contractions C^j here are simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$. We assume that if we formally treat the factors $\nabla\omega_1$, $\nabla\omega_2$ as factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ (disregarding whether they are contracting against special indices) in the above, then the inductive assumption of Proposition 1.1 applies.

The conclusion we will draw (under various hypotheses that we will explain below) is that we can write

$$\begin{aligned}\sum_{x\in X} a_x X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_a} C^{x,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)\\ =\sum_{x\in X'} a_x X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_a} C^{x,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)\\ +\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)=0,\end{aligned} \tag{2-8}$$

where the tensor fields indexed in X' on the right-hand side are in the form (2-6) with $B>0$. All the other sublinear combinations are as above.

Assumptions needed for (2-8): We claim that (2-8) holds under certain assumptions on the α-tensor fields in (2-7) that have rank α and have a free index in one of the factors $\nabla\omega_1$, $\nabla\omega_2$ (say in $\nabla\omega_1$ without loss of generality)—we denote the index set of those tensor fields by $X^{\alpha,*}\subset X$.

The assumption we need in order for the claim to hold is that no tensor field indexed in $X^{\alpha,*}$ should be "bad". A tensor field is "bad" if it has the property that when we erase the expression $\nabla_{[a}\omega_1\nabla_{b]}\omega_2$ (and make the index that contracted

[20] Recall that in the definition of "real length" in this setting, we count each factor $\nabla^{(m)}R$, $S_*\nabla^{(\nu)}R$, $\nabla^{(B)}\Omega_x$ once, the two factors $\nabla^{(a)}\omega_1$, $\nabla\omega_2$ as one, and the factors $\nabla\phi$, $\nabla\phi'$, $\nabla\tilde{\phi}$ as nothing.

against $_b$ into a free index) then the resulting tensor field will have no removable indices, and all factors $S_* R_{ijkl}$ will be simple.

Lemma 2.4. *We assume that (2-7) holds, where now the tensor fields have length $\sigma = 3$. We also assume that for each of the tensor fields indexed in X, there is a removable index in each of the real factors. We then claim that the conclusion of Lemma 2.3 is still true in this setting.*

For the most part, the remainder of this paper is devoted to proving the above lemmas. However, we first state and prove some further technical claims, one of which appeared as Lemma A.1 in [A 2010].[21]

2B. ***Two more technical lemmas.*** We claim that an analogue of Lemma 4.10 in [A 2010][22] can be derived for tensor fields with a given simple character $\vec{\kappa}_{\text{simp}}$, and where rather than having *one* additional factor $\nabla\phi_{u+1}$ (which is not encoded in the simple character $\vec{\kappa}_{\text{simp}}$), we have *two additional factors* $\nabla_a\phi_{u+1}$, $\nabla_b\phi_{u+2}$.

Lemma 2.5. *Consider a linear combination*

$$\sum_{l\in L} a_l C_g^{l,i_1\dots i_\beta}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

of acceptable tensor fields in the form (1-6) with a given u-simple character $\vec{\kappa}_{\text{simp}}$. Assume that the minimum rank among those tensor fields above is $\alpha \geq 2$. Assume that

$$\begin{aligned}\sum_{l\in L} a_l X_*\operatorname{div}_{i_3}\dots X_*\operatorname{div}_{i_\beta} C_g^{l,i_1\dots i_\beta}&(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\\ &\times\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\\ +\sum_{j\in J} a_j C_g^{j,i_1 i_2}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}=0;\end{aligned}\tag{2-9}$$

here $X_*\operatorname{div}_i$ means that ∇^i is in addition not allowed to hit either of the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$. We also assume that if we formally treat the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ as factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ then (2-9) falls under the inductive assumption of Proposition 1.1 (with respect to the parameters (n,σ,Φ,u)). Denote by $L^\alpha \subset L$ the index set of terms with rank α. We additionally assume that none of the tensor fields $C_g^{l,i_1\dots i_\beta}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$ are "forbidden," in the sense defined above Proposition 2.1 *in* [A 2010].

[21] Its proof was also deferred to the present paper.

[22] This corresponds to Lemma 4.44 in [A 2012].

We then claim that there exists a linear combination of $(\alpha+1)$*-tensor fields with a u-simple character* $\vec{\kappa}_{\text{simp}}$ (*indexed in* Y *below*) *so that*

$$\sum_{l\in L^\alpha} a_l C_g^{l,i_1\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon$$
$$= X_*\operatorname{div}_{i_{\alpha+1}}\sum_{y\in Y} a_y C_g^{l,i_1\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\times\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon$$
$$+\sum_{j\in J} a_j C_g^{j,i_1\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\times\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon. \tag{2-10}$$

We also claim that we can write

$$\sum_{l\in L} a_l X\operatorname{div}_{i_3}\dots X\operatorname{div}_{i_\beta} C_g^{l,i_1\dots i_\beta}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}$$
$$=\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$
$$+\sum_{q\in Q_1} a_q X\operatorname{div}_{i_3}\dots X\operatorname{div}_{i_\alpha}\times C_g^{q,i_3\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_{u+2})\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}$$
$$+\sum_{q\in Q_2} a_q X\operatorname{div}_{i_3}\dots X\operatorname{div}_{i_\alpha}\times C_g^{q,i_3\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_{u+2}), \tag{2-11}$$

where the tensor fields indexed in Q_1 *are acceptable with a u-simple character* $\vec{\kappa}_{\text{simp}}$ *and with a factor* $\nabla^{(2)}\phi_{u+1}$ *and a factor* $\nabla\phi_{u+2}$. *The tensor fields indexed in* Q_2 *are acceptable with a u-simple character* $\vec{\kappa}_{\text{simp}}$ *and with a factor* $\nabla^{(2)}\phi_{u+2}$ *and a factor* $\nabla\phi_{u+1}$.

Proof of Lemma 2.5. We may divide the index set L^α into subsets L^α_I, L^α_{II} according to whether the two factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against the same factor or not — we will then prove our claim for those two index sets separately. Our claim for the index set L^α_{II} follows by a straightforward adaptation of the proof of Lemma 4.10 in [A 2010]. (Notice that the forbidden cases of the present lemma are exactly in correspondence with the forbidden cases of that lemma.) Therefore, we now prove our claim for the index set L^α_I:

We denote by $L_I\subset L$, $J_I\subset J$ the index sets of terms for which the two factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against the same factor. It then follows that (2-9) holds with the index sets L, J replaced by L_I, J_I — denote the resulting new equation by New(2-9). Now, for each tensor field $C_g^{l,i_1\dots i_\beta}$ and each complete contraction C^j_g, we let $\text{Sym}[C_g^{l,i_1\dots i_\beta}]$, $\text{Sym}[C_g^{l,i_1\dots i_\beta}]$, $\text{AntSym}[C^j_g]$, $\text{AntSym}[C^j_g]$ stand for the tensor field/complete contraction that arises from $C^{l,i_1\dots i_\beta}_g$, C^j_g by symmetrizing (respectively antisymmetrizing) the indices ${}_a, {}_b$ in the two factors $\nabla_a\phi_{u+1}$, $\nabla_b\phi_{u+2}$. We accordingly derive two new equations from New(2-9), which we denote by New(2-9)$_{\mathrm{Sym}}$ and New(2-9)$_{\mathrm{AntSym}}$.

We will then prove the claim separately for the tensor fields in the sublinear combination $\sum_{l\in L^\alpha_I} a_l \,\mathrm{Sym}[C]^{l,i_1\dots i_\alpha}_g$ and the tensor fields in the sublinear combination $\sum_{l\in L^\alpha_I} a_l \,\mathrm{AntSym}[C]^{l,i_1\dots i_\alpha}_g$.

The claim (2-10) for the sublinear combination $\sum_{l\in L^\alpha_I} a_l \,\mathrm{AntSym}[C]^{l,i_1\dots i_\alpha}_g$ follows directly from the arguments in the proof of Lemma 2.3 (see this proof below). Therefore it suffices to show our claim for $\sum_{l\in L^\alpha_I} a_l \,\mathrm{Sym}[C]^{l,i_1\dots i_\alpha}_g$.

We prove this claim as follows: We divide the index set L^α_I according to the *form* of the factor against which the two factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ contract: List out the nongeneric factors $\{T_1,\dots,T_a\}$ in $\vec{\kappa}_{\mathrm{simp}}$.[23] Then, for each $k\le a$ we let $L^\alpha_{I,k}$ stand for the index set of terms for which the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against the factor T_k. We also let $L^\alpha_{I,a+1}$ stand for the index set of terms for which the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against a generic factor $\nabla^{(m)}R_{ijkl}$. We will prove our claim for each of the sublinear combinations $\sum_{l\in L^\alpha_{I,a+1}} a_l \,\mathrm{Sym}[C]^{l,i_1\dots i_\alpha}_g$ separately.

We first observe that for each $k\le a+1$, we may obtain a new true equation from (2-9) by replacing L by $L_{I,a+1}$ — denote the resulting equation by (2-9)$_{I,\mathrm{Sym},k}$. Therefore, for each $k\le a+1$ for which T_k is in the form $\nabla^{(p)}\Omega_h$, our claim follows straightforwardly by applying Corollary 1 from [A 2010].[24]

Now, we consider the case where the factor T_k is in the form $S_*\nabla^{(\nu)}R_{abcd}$: In that case we denote by $L_{I,k,\sharp}$ the index set of terms for which one of the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ is contracting against a special index in T_k. In particular, we will let $L^\alpha_{I,k,\sharp}\subset L_{I,k,\sharp}$ stand for the index set of terms with rank α. We will then show that two equations hold:

First, we claim that there exists a linear combination of tensor fields as claimed in (2-10) so that

$$
\begin{aligned}
\text{(2-12)}\qquad \sum_{l\in L^\alpha_{I,k,\sharp}} a_l \,\mathrm{Sym}[C]^{l,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&\\
\times \nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon&\\
-\sum_{y\in Y} a_y X\,\mathrm{div}_{i_{\alpha+1}} C^{y,i_1\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&\\
\times \nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon&
\end{aligned}
$$

[23] Recall from [A 2010] that the nongeneric factors in $\vec{\kappa}_{\mathrm{simp}}$ are all the factors in the form $\nabla^{(A)}\Omega_h$, $S_*\nabla^{(\nu)}R_{ijkl}$, and also all the factors $\nabla^{(m)}R_{ijkl}$ that contract against at least one factor $\nabla\phi_s$.

[24] This corresponds to Corollary 4.14 in [A 2012]. There is no danger of falling under a "forbidden case," since we started with tensor fields which were not forbidden.

$$= \sum_{l\in L^{\alpha}_{OK}} a_l X\,\mathrm{div}_{i_{\alpha+1}}\, C^{l,i_1\dots i_\alpha i_{\alpha+1}}_g \nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon$$
$$+\sum_{j\in J} a_j C^{j,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u) \times \nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon,$$

where the tensor fields in L^{α}_{OK} have all the properties of the terms in $L_{I,k}$, rank α and furthermore none of the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against a special index.

Then (under the assumption that $L^{\alpha}_{I,k,\sharp}=\varnothing$) we claim that we can write

$$\begin{aligned}\text{(2-13)}\quad &\sum_{l\in L_{I,k,\sharp}} a_l X\,\mathrm{div}_{i_3}\dots X\,\mathrm{div}_{i_\beta}\\ &\qquad\times \mathrm{Sym}[C]^{l,i_1\dots i_\beta}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\\ &= \sum_{l\in L_{I,k,OK}} a_l X\,\mathrm{div}_{i_3}\dots X\,\mathrm{div}_{i_\beta}\\ &\qquad\times \mathrm{Sym}[C]^{l,i_1\dots i_\beta}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2}\\ &\quad+\sum_{j\in J} a_j\,\mathrm{Sym}[C]^{j,i_1 i_2}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2},\end{aligned}$$

where the tensor fields in $L_{I,k,OK}$ have all the properties of the terms in $L_{I,k}$, but they additionally have rank $\geq \alpha+1$ and furthermore none of the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ are contracting against a special index.

If we can show the above two equations, then we are reduced to showing our claim under the additional assumption that no tensor field indexed in L in Sym(2-9) has any factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ that contract against a special index in T_k. Under that assumption, we may additionally assume that none of the complete contractions indexed in J in (2-9) have that property.[25] Therefore, we may then *erase* the factor $\nabla\phi_{u+1}$ from all the complete contractions and tensor fields in (2-9)$_k$ by virtue of the operation Erase, introduced in the Appendix in [A 2012] — our claim then follows by applying Corollary 1 from [A 2010] to the resulting equation and then reintroducing the erased factor $\nabla\phi_{u+1}$. □

Outline of the proofs of (2-12) *and* (2-13). First we prove (2-12): Suppose without loss of generality that T_k contracts against $\nabla\tilde{\phi}_1$ and $\nabla\phi'_2,\dots,\nabla\phi'_h$; then replace the two factors $\nabla_a\phi_1$, $\nabla_b\phi_{u+1}$ by g_{ab} and then apply Ricto Ω_{p+1},[26] (obtaining a new true equation) an then apply the eraser to the resulting true equation. We then apply Corollary 1 from [A 2010] to the resulting equation,[27] and finally we replace

[25] This can be derived by repeating the *proof* of (2-12), (2-13).

[26] See the relevant lemma in the Appendix of [A 2012].

[27] Since the factor $\nabla\phi_{u+2}$ survives this operation, and since we started out with terms that were not "forbidden," there is no danger of falling under a "forbidden case" of Corollary 1 from [A 2010].

the factor $\nabla^{(b)}_{r_1\dots r_b}\Omega_{p+1}$ by an expression

$$S_*\nabla^{(b+h-1)}_{y_2\dots y_h r_1\dots r_{b-1}} R_{ijkr_b}\nabla^i\tilde{\phi}_1\nabla^j\phi_{u+2}\nabla^k\phi_{u+1}\nabla^{y_2}\phi'_2\dots\nabla^{y_h}\phi'_h.$$

As in the proof of Lemma 4.10 in [A 2010], we derive our claim. Then (2-13) is proven by iteratively applying this step and making each $\nabla\upsilon$ into an X div at every stage.

We analogously show our claim when the factor T_k is in the form $\nabla^{(m)}R_{ijkl}$: In that case we denote by $L_{I,k,\sharp}$ the index set of terms for which *both* the factors $\nabla\phi_{u+1}, \nabla\phi_{u+2}$ are contracting against a special index in T_k. We will then show two claims:

First, that there exists a linear combination of partial contractions (indexed in Y below) as claimed in (2-10) so that

$$\begin{aligned}\sum_{l\in L^\alpha_{I,k,\sharp}} a_l \operatorname{Sym}[C]^{l,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon \\ = \sum_{y\in Y} a_y X\operatorname{div}_{i_{\alpha+1}} C^{y,i_1\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u) \\ \times\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon \\ +\sum_{l\in L^\alpha_{OK}} a_l X\operatorname{div}_{i_{\alpha+1}} C^{l,i_1\dots i_\alpha i_{\alpha+1}}_g\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon \\ +\sum_{j\in J} a_j C^{j,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u) \\ \times\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\omega\nabla_{i_3}\upsilon\dots\nabla_{i_\alpha}\upsilon,\end{aligned}\tag{2-14}$$

where the tensor fields in L^α_{OK} have all the properties of the terms in $L_{I,k}$, but they additionally have rank α and furthermore one of the factors $\nabla\phi_{u+1}, \nabla\phi_{u+2}$ does not contract against a special index. Then (under the assumption that $L^\alpha_{I,k,\sharp}=\varnothing$) we denote by $L_{I,k,\sharp}$ the sublinear combination of terms in $L_{I,k}$ where *both* factors $\nabla\phi_{u+1}$ or $\nabla\phi_{u+1}$ contract against a special index in T_k. We claim that we can write

$$\begin{aligned}\sum_{l\in L_{I,k,\sharp}} a_l X\operatorname{div}_{i_3}\dots X\operatorname{div}_{i_\beta} \\ \times\operatorname{Sym}[C]^{l,i_1\dots i_\beta}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2} \\ =\sum_{l\in L_{I,k,OK}} a_l X\operatorname{div}_{i_3}\dots X\operatorname{div}_{i_\beta} \\ \times\operatorname{Sym}[C]^{l,i_1\dots i_\beta}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2} \\ +\sum_{j\in J} a_j\operatorname{Sym}[C]^{j,i_1 i_2}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\nabla_{i_2}\phi_{u+2},\end{aligned}\tag{2-15}$$

where the tensor fields in $L_{I,k,OK}$ have all the properties of the terms in $L_{I,k}$, but they additionally have rank $\geq\alpha+1$ and furthermore one of the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ does not contract against a special index.

If we can show the above two equations, then we are reduced to showing our claim under the additional assumption that no tensor field indexed in L in Sym(2-9) has the two factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$, contracting against a special index in T_k. Under that assumption, we may additionally assume that none of the complete contractions indexed in J in (2-9) have that property. Therefore, we may then *erase* the factor $\nabla\phi_{u+1}$ from all the complete contractions and tensor fields in $(2\text{-}9)_k$ — our claim then follows by applying Lemma 4.10 in [A 2010] to the resulting equation[28] and then reintroducing the erased factor $\nabla\phi_{u+1}$. □

Outline of the proofs of (2-14) *and* (2-15). First we prove (2-14). Suppose without loss of generality that T_k contracts against $\nabla\phi_1, \dots, \nabla\phi_h$ (possibly with $h = 0$); then replace the two factors $\nabla_a\phi_1$, $\nabla_b\phi_{u+1}$ by g_{ab} and then apply Ricto Ω_{p+1} (obtaining a new true equation), and then apply the eraser to the factors $\nabla\phi_1, \dots, \nabla\phi_h$ in the resulting true equation. Then (apart from the cases, discussed below, where the above operation may lead to a "forbidden case" of Corollary 1 in [A 2010]), we apply that corollary to the resulting equation, and finally we replace the factor $\nabla^{(b)}_{r_1\dots r_b}\Omega_{p+1}$ by an expression

$$\nabla^{(b+h)}_{s_1\dots s_h r_1\dots r_{b-2}} R_{i r_{b-1} k r_b} \nabla^i\phi_{u+1}\nabla^k\phi_{u+2}\nabla^{s_1}\phi_1\dots\nabla^{s_h}\phi_h.$$

As in the proof of Lemma 4.10 in [A 2010], we derive our claim. Then (2-14) is proven by iteratively applying this step and making each $\nabla\upsilon$ into an X div at every stage (again, provided we never encounter "forbidden cases"). If we do encounter forbidden cases, then our claims follow by just making the factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ into X divs and then applying Corollary 1 in [A 2012] to the resulting equation (the resulting equation is not forbidden, since it will contain a factor $\nabla^{(m)}R_{ijkl}$ with two free indices), and in the end renaming two factors $\nabla\upsilon$ as $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$. □

A further generalization: Proof of Lemma A.1 from [A 2010]. We remark that on a few occasions later in this series of papers we will be using a generalized version of the Lemma 2.5. The generalized version asserts that the claim of Lemma 2.5 remains true, for the general case where rather than one or two "additional" factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$ we have $\beta \geq 3$ "additional" factors $\nabla\phi_{u+1}, \dots, \nabla\phi_{u+\beta}$. Moreover, in that case there are no "forbidden cases".

Lemma 2.6. *Let*

$$\sum_{l\in L_1} a_l C_g^{l,i_1\dots i_\mu, i_{\mu+1}\dots i_{\mu+\beta}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),$$

$$\sum_{l\in L_2} a_l C_g^{l,i_1\dots i_{b_l}, i_{b_l+1}\dots i_{b_l+\beta}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),$$

[28] Notice that there is no danger of falling under a "forbidden case" of that lemma, since there will be a nonsimple factor $S_*\nabla^{(\nu)}R_{ijkl}$ by virtue of the factor $\nabla\phi_{u+2}$.

stand for two linear combinations of acceptable tensor fields in the form (1-6), *each with u-simple character* $\vec{\kappa}_{\mathrm{simp}}$. *We assume that the terms indexed in* L_1 *have rank* $\mu+\beta$, *while the ones indexed in* L_2 *have rank greater than* $\mu+\beta$.

Assume that

$$\begin{aligned}\sum_{l\in L_1} a_l X\,\mathrm{div}_{i_{\beta+1}}\dots X\,\mathrm{div}_{i_{\mu+\beta}}&\\ \times C_g^{l,i_1\dots i_{\mu+\beta}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}&\\ +\sum_{l\in L_2} a_l X\,\mathrm{div}_{i_{\beta+1}}\dots X\,\mathrm{div}_{i_{b_l}}&\\ \times C_g^{l,i_1\dots i_{b_l+\beta}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}&\\ +\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_{u+\beta})=0&,\end{aligned}\tag{2-16}$$

modulo terms of length $\geq \sigma+u+\beta+1$. *Furthermore, we assume that the above equation falls under the inductive assumption of Proposition* 2.1 *in [A 2010]* (*with regard to the parameter weights*, σ, Φ, p). *We are not excluding any "forbidden cases".*

We claim that there exists a linear combination of $(\mu+\beta+1)$*-tensor fields in the form* (1-6) *with u-simple character* $\vec{\kappa}_{\mathrm{simp}}$ *and length* $\sigma+u$ (*indexed in* H *below*) *such that*

$$\begin{aligned}\sum_{l\in L_1} a_l C_g^{l,i_1\dots i_{\mu+\beta}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&\\ \times\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}\nabla_{i_{\beta+1}}\upsilon\dots\nabla_{i_{\beta+\mu}}\upsilon&\\ +\sum_{h\in H} a_h X\,\mathrm{div}_{i_{\mu+\beta+1}} C_g^{l,i_1\dots i_{\mu+\beta+1}}(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)&\\ \times\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{\beta+1}\nabla_{i_1}\upsilon\dots\nabla_{i_{\beta+\mu}}\upsilon&\\ +\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_{u+\beta},\upsilon^\mu)=0&,\end{aligned}\tag{2-17}$$

modulo terms of length $\geq \sigma+u+\beta+\mu+1$. *The terms indexed in* J *here are u-simply subsequent to* $\vec{\kappa}_{\mathrm{simp}}$.

Proof of Lemma 2.6. The proof of the above is a straightforward adaptation of the proof of Lemma 2.5, *except* for the cases where the tensor fields $C_g^{l,i_1\dots i_\mu,i_{\mu+1}\dots i_{\mu+\beta}}$ are "bad," where "bad" in this case means that all factors are in the form R_{ijkl}, S_*R_{ijkl}, $\nabla^{(2)}\Omega_h$,[29] and in addition each factor $\nabla^{(2)}\Omega_h$ contracts against at most one factor $\nabla\phi_h$, $1\leq h\leq u+\beta$. So we now focus on that case.

[29] Notice that if this property holds for one of the terms $C_g^{l,i_1\dots i_\mu,i_{\mu+1}\dots i_{\mu+\beta}}$, then it will hold for all of them by weight considerations.

The "bad" case. Let us observe that by weight considerations, *all* tensor fields in (2-9) must now have rank μ.

We recall that this special proof applies only in the case where there are special free indices in factors $S_* R_{ijkl}$ among the tensor fields of minimum rank in (2-9). (If there were no such terms, then the regular proof of Lemma 2.5 would apply.) We distinguish three cases: Either $p > 0$, or $p = 0$ and $\sigma_1 > 0$, or $p = \sigma_1 = 0$ and $\sigma_2 > 0$. We will prove the above by an induction on the parameters (*weight*), σ: Suppose that the weight of the terms in (2-16) is $-K$ and the real length is $\sigma \geq 3$. We assume that the lemma holds when the Equation (2-16) consists of terms with weight $-K'$, $K' < K$, or of terms with weight $-K$ and real length σ', $3 \leq \sigma' < \sigma$.

The case $p > 0$. We first consider the μ-tensor fields in (2-9) with the extra factor $\nabla\phi_{u+1}$ contracting against a factor $\nabla^{(2)}\Omega_h$. Denote the index set of those terms by $\bar{L}_\mu$. We will first prove that

$$\sum_{l \in \bar{L}_\mu} a_l C_g^{l,i_1 \dots i_{\mu+\beta}}(\Omega_1, \dots, \Omega_p, \phi_1, \dots, \phi_u) \times \nabla_{i_1}\phi_{u+1} \dots \nabla_{i_\beta}\phi_{u+\beta} \nabla_{i_{\beta+1}} \upsilon \dots \nabla_{i_{\beta+\mu}} \upsilon = 0. \tag{2-18}$$

It suffices to prove the above for the sublinear combination of μ-tensor fields where $\nabla\phi_{u+1}$ contracts against $\nabla^{(2)}\Omega_1$. (2-18) will then follow by relabeling the functions $\Omega_1, \dots, \Omega_p$ and repeating this step p times.

We start by a preparatory claim: Let us denote by $\bar{L}_{\mu,\sharp} \subset \bar{L}_\mu$ the index set of μ-tensor fields for which the factor $\nabla^{(2)}\Omega_1$ contains a free index, say the index $_{i_1}$ without loss of generality. We will first prove that

$$\sum_{l \in \bar{L}_{\mu,\sharp}} a_l C_g^{l,i_1 \dots i_\mu} \nabla_{i_1}\phi_{u+1} \dots \nabla_{i_\beta}\phi_{u+\beta} \nabla_{i_{\beta+1}} \upsilon \dots \nabla_{i_{\beta+\mu}} \upsilon = 0. \tag{2-19}$$

Proof of (2-19). We will use the technique (introduced in Subsection 3.1 of [A 2011]) of "inverse integration by parts" followed by the silly divergence formula.

Let us denote by $\hat{C}_g^l$ the complete contraction that arises from each $C_g^{l,i_1 \dots i_{\mu+\beta}}$ by formally erasing the expression $\nabla^{(2)}_{si_1}\Omega_1 \nabla^s \phi_{u+1}$ and then making all free indices $_{i_{\beta+1}}, \dots, _{i_{\beta+\mu}}$ into internal contractions.[30] Then, the "inverse integration by parts" implies a new integral equation

$$\int_{M^n} \sum_{l \in L_\mu} a_l \hat{C}_g^l + \sum_{j \in J} a_j C_g^j + \sum_{z \in Z} a_z C_g^z \, dV_g = 0. \tag{2-20}$$

Here the complete contractions indexed in J have length $\sigma + u$ and u factors $\nabla\phi_u$, but they are simply subsequent to the simple character $\vec{\kappa}_{\text{simp}}$. The terms indexed in

[30] We recall that to "make a free index $_{i_y}$ into an internal contraction" means that we add a derivative ∇_{i_y} onto the factor T_{i_y} to which the free index $_{i_y}$ belongs. The new derivative index ∇^{i_y} is then contracted against the index $_{i_y}$ in T_{i_y}.

Z either have length $\geq \sigma+u+1$ or have length $\sigma+u$, but also have at least one factor $\nabla^{(B)}\phi_h$ with $B\geq 2$.

Now, in the above, we consider the complete contractions indexed in $\bar{L}_{\mu,\sharp}\subset \bar{L}_\mu$ and we "pull out" the expression $\Delta\nabla_t\Omega_1\nabla^t\phi_{u+1}$ to write

$$\sum_{l\in L_{\mu,\sharp}} a_l\hat{C}^l_g=\sum_{l\in L_{\mu,\sharp}} a_l\bar{C}^l_g\cdot(\Delta\nabla_t\Omega_1\nabla^t\phi_1).$$

Now, we consider the silly divergence formula applied to (2-20) obtained by integrating by parts with respect to the function Ω_1. If we denote the integrand in (2-20) by F_g, we denote the resulting (local) equation by $silly[F_g]=0$. We consider the sublinear combination $silly^*[F_g]$ which consists of terms with length $\sigma+u$, μ internal contractions and $u-1+\beta$ factors $\nabla\phi_h$, $h\geq 2$, and a factor $\Delta\phi_{u+1}$. Clearly, this sublinear combination must vanish separately modulo longer terms,

$$silly^*[F_g]=0.$$

The above equation can be expressed as

$$\text{Spread}^{\nabla^s,\nabla_s}\Bigg[\sum_{l\in\bar{L}_{\mu,\sharp}} a_l\bar{C}^l_g\Bigg]\cdot\Omega_1\cdot\Delta\phi_{u+1}=0. \tag{2-21}$$

(Here $\text{Spread}^{\nabla^s,\nabla_s}$ is a formal operation that acts on complete contractions in the form (1-5) by hitting a factor T in the form $\nabla^{(m)}R_{ijkl}$ or $\nabla^{(p)}\Omega_h$ with a derivative ∇^s and then hitting another factor $T'\neq T$ in the form $\nabla^{(m)}R_{ijkl}$ or $\nabla^{(p)}\Omega_h$ with a derivative ∇_s that contracts against ∇^s and then adding over all the terms we can thus obtain.) Now, using the fact that (2-21) holds formally, we derive[31]

$$\sum_{l\in L_{\mu,\sharp}} a_l\bar{C}^l_g=0. \tag{2-22}$$

Thus, applying the operation Sub_υ $\mu-1$ times to the above and then multiplying by $\nabla_{i_1i_2}\Omega_1\nabla^{i_1}\upsilon\nabla^{i_2}\phi_{u+1}$ we derive (2-19). So for the rest of this proof we may assume that $\bar{L}_{\mu,\sharp}=\varnothing$. $\square$

Now we prove our claim under the additional assumption that for the tensor fields indexed in $\bar{L}_\mu$, the factor $\nabla^{(2)}\Omega_1$ contains no free index.

We again refer to (2-20) and perform integrations by parts with respect to the factor $\nabla^{(B)}\Omega_1$. We denote the resulting local equation by $silly[L_g]=0$. We pick out the sublinear combination $silly^*[L_g]$ of terms with $\sigma+u$ factors, $u+\beta$ factors $\nabla\phi_h$, μ internal contractions, with $u+\beta-1$ factors $\nabla\phi_h$, $h\geq 2$, and a factor $\Delta\phi_1$. This sublinear combination must vanish separately, $silly^*[L_g]=0$; the resulting new true equation can be described easily: Let us denote by $\hat{C}^{l,j_1}_g$ the 1-vector

[31] This can be proven by using the operation Erase[...], see the Appendix in [A 2010].

field that arises from $C^{l,i_1\dots i_\mu}_g$, $l\in L_{\mu,*}$ by formally erasing the factor $\nabla^{(2)}_{js}\Omega_1\nabla^s\phi_1$, making the index ^{j} that contracted against ${}_j$ into a free index ${}_{j_1}$, and making all the free indices ${}_{i_1},\dots,{}_{i_\mu}$ into internal contractions. (Denote by $\vec{\kappa}'_{\text{simp}}$ the simple character of these vector fields.) Then the equation $\mathit{silly}^*[L_g]=0$ can be expressed in the form

$$\sum_{l\in L_{\mu,*}} a_l\{X\,\mathrm{div}_{j_1}\hat{C}^{l,j_1}_g\}\Delta\phi_1+\sum_{j\in J}a_jC^j_g\Delta\phi_1=0; \tag{2-23}$$

here the complete contractions C^j_g are simply subsequent to $\vec{\kappa}'_{\text{simp}}$. The above holds modulo terms of length $\geq\sigma+u+1$. Now, we apply the operation Sub_ω μ times (see the Appendix in [A 2012]). In the case $\sigma>3$, we apply the inductive assumption of our Lemma 2.6 to the resulting equation (notice that the above falls under the inductive assumption of this lemma since we have lowered the weight in absolute value); we ensure that Lemma 2.6 can be applied by just labeling one of the factors $\nabla\omega$ into $\nabla\phi_{u+1}$. We derive (due to weight considerations) that there *can not be* tensor fields of higher rank, thus

$$\sum_{l\in L_\mu} a_l\,\mathrm{Sub}^{\mu-1}_\omega\big[\hat{C}^{l,j_1}_g\big]\nabla_{i_1}\upsilon\Delta\phi_1=0. \tag{2-24}$$

Now, formally replacing the factor $\nabla_{i_1}\upsilon$ by $\nabla^{(2)}_{j_1t}\Omega_1\nabla^t\phi_1$, and then setting $\omega=\upsilon$, we derive the claim of our lemma. In the case $\sigma=3$ (2-24) follows by inspection, since the only two possible cases are $\sigma_2=2$ and $\sigma_1=2$; in the first case there are only two possible partial contractions in $\bar{L}_\mu$ while in the second there are four. Equation (2-23) (by inspection) implies that the coefficients of all these tensor fields must vanish, which is equivalent to (2-24).

Now, we will prove our claim under the additional assumption $\bar{L}_\mu=\varnothing$ (still for $p>0$). We again refer to (2-20) and again consider the same equation $\mathit{silly}[L_g]=0$ as above. We now pick out the sublinear combination of terms with $\sigma+u$ factors, $u+\beta$ factors $\nabla\phi_h$, and μ internal contractions. We derive that

$$\sum_{l\in L_\mu} a_l X\,\mathrm{div}_{j_1} X\,\mathrm{div}_{j_2}\hat{C}^{l,j_1j_2}_g+\sum_{j\in J}a_jC^j_g=0; \tag{2-25}$$

here the terms $\hat{C}^{l,j_1j_2}_g$ arise from the μ-tensor fields $C^{li_1\dots i_\mu}_g$ by replacing all μ free indices by internal contractions, erasing the factor $\nabla^{(2)}_{jk}\Omega_1$ and making the indices ${}^j,{}^k$ into free indices ${}^{j_1},{}^{j_2}$. Now, applying Sub_ω μ times, and then applying the inductive assumption of Lemma 4.10 in [A 2010] (this applies by length considerations as above for $\sigma>3$; while if $\sigma=3$ the claim (2-26) will again follow by inspection) we derive that

$$\sum_{l\in L_\mu} a_l \hat{C}^{l,j_1 j_2}_g \nabla_{j_1}\upsilon\nabla_{j_2}\upsilon = 0. \tag{2-26}$$

Replacing the expression $\nabla_{j_1}\upsilon\nabla_{j_2}\upsilon$ by a factor $\nabla^{(2)}_{j_1 j_2}\Omega_2$ and then setting $\omega=\upsilon$, we derive our claim in the case $p>0$.

The case $p=0, \sigma_1>0$. We will reduce to the previous case: We let L^1_μ be the index set of μ-tensor fields where the factor $T_1=S_*R_{ijkl}\nabla^i\tilde{\phi}_1$ contains a special free index (say the index $_k$ is the free index $_{i_{\beta+1}}$ without loss of generality). We will prove our claim for the index set L^1_μ; if we can prove this, then clearly our lemma will follow by induction.

To prove this claim, we consider the first conformal variation of our hypothesis, $Image^1_Y[L_g]=0$, and we pick out the sublinear combination of terms with length $\sigma+u+\beta$, where the factor $\nabla^{(\nu)}S_*R_{ijkl}\nabla^i\tilde{\phi}_1$ has been replaced by a factor $\nabla^{(\nu+2)}Y$, and the factor $\nabla\phi_1$ now contracts against a factor $T_2=R_{ijkl}$. This sublinear combination vanishes separately, thus we derive a new local equation. To describe the resulting equation, we denote by

$$\hat{C}^{l,i_1\dots\hat{i}_{\beta+1}\dots i_{\mu+\beta}}_g(Y,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}$$

the $(\mu-1)$-tensor field that arises from

$$C^{l,i_1\dots\hat{i}_{\beta+1}\dots i_{\mu+\beta}}_g(Y,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}$$

by formally replacing the factor $T_1=S_*R_{ijkl}\nabla^i\tilde{\phi}_1$ by $\nabla^{(2)}_{jl}Y$ and also adding a derivative index ∇_{i_*} onto the factor $T_2=R_{ijkl}$ and then contracting that index $_{i_*}$ against an (added anew) factor $\nabla\phi_1$. Denote the $(u-1)$-simple character of the above (the one defined by $\nabla\phi_2,\dots,\nabla\phi_u$) by $\vec{\kappa}'_{\text{simp}}$. We then have an equation

(2-27)
$$\sum_{l\in L^1_\mu} a_l X\,\text{div}_{i_{\beta+2}}\dots X\,\text{div}_{i_{\beta+\mu}}\,\hat{C}^{l,i_1\dots\hat{i}_{\beta+1}\dots i_{\mu+\beta}}_g(Y,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}$$
$$+\sum_{h\in H} a_h X\,\text{div}_{i_{\beta+1}}\dots X\,\text{div}_{i_{\mu+\beta}}\,C^{l,i_2\dots i_{\mu+\beta}}_g(Y,\phi_1,\dots,\phi_u)\nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}$$
$$=\sum_{j\in J} a_j C^{j,i_{\mu+1}\dots i_\beta}_g(Y,\phi_1,\dots,\phi_u)\nabla_{i_{\mu+1}}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}.$$

The terms indexed in H are acceptable, have a $(u-1)$-simple character $\vec{\kappa}'_{\text{simp}}$ and the factor $\nabla\phi_1$ contracts against an internal index (without loss of generality, say the index $_i$ in the factor $T_2=R_{ijkl}$); writing that factor as $S_*R_{ijkl}\nabla^i\tilde{\phi}_1$, we denote the resulting u-simple factor by $\tilde{\kappa}_{\text{simp}}$. The terms indexed in J are simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$. Now, applying the inductive assumption of Lemma 2.6,[32] we derive that

$$\sum_{h\in H} a_h C_g^{l,i_2\dots i_{\mu+\beta}}(Y,\phi_1,\dots,\phi_u) \times \nabla_{i_1}\phi_{u+1}\dots\nabla_{i_{\beta+1}}\phi_{u+\beta}\nabla_{i_{\beta+1}}\upsilon\dots\nabla_{i_{\beta+\mu}}\upsilon = 0. \tag{2-28}$$

Thus, we may assume without loss of generality that $H=\varnothing$ in (2-27). Now, we again apply Lemma 2.6 to (2-27) (under that additional assumption), and we derive that

$$\sum_{l\in L^1_\mu} a_l \hat{C}_g^{l,i_1\dots\hat{i}_{\beta+1}\dots i_{\mu+\beta}}(Y,\phi_1,\dots,\phi_u) \times \nabla_{i_1}\phi_{u+1}\dots\nabla_{i_\beta}\phi_{u+\beta}\nabla_{i_{\beta+1}}\upsilon\dots\nabla_{i_{\beta+\mu}}\upsilon = 0. \tag{2-29}$$

Now, erasing the factor $\nabla\phi_1$ from the above, and then formally replacing the factor $\nabla^{(2)}_{ab}Y$ by $S_*R_{i(ab)l}\nabla^i\tilde{\phi}_1\nabla^l\upsilon$, we derive our claim.

The case $p=0, \sigma_1=0$. In this case $\sigma=\sigma_2$. In other words, all factors in $\vec{\kappa}_{\mathrm{simp}}$ are simple factors in the form $S_*R_{ijkl}\nabla^i\tilde{\phi}_h$. We recall that in this case all μ-tensor fields in (2-9) must have at most one free index in any factor S_*R_{ijkl}. In that case, we will prove our claim in a more convoluted manner, again reducing ourselves to the inductive assumption of Proposition 2.1 in [A 2010].

A key observation is that by the definition of the special cases, $\mu+\beta\le\sigma_2$. In the case of strict inequality, we see (by a counting argument) that at least one of the special indices in one of the factors S_*R_{ijkl} must contract against a special index in another factor S_*R_{abcd}. In the case $\mu+\beta=\sigma_2$ this remains true, *except* for the terms for which the β factors $\nabla\phi_{u+h}$ contract against special indices, say the indices ${}_k$, in β factors $T_y=S_*R_{ikl}\nabla^i\tilde{\phi}_y$, and moreover these factors must not contain a free index, and all other factors S_*R_{ikl} contain exactly one free index, which must be special. In this subcase, we will prove our claim for all μ-tensor fields *excluding* this particular "bad" sublinear combination; we will prove our claim for this sublinear combination in the end.

We will now proceed to *normalize* the different $(\mu+\beta)$-tensor fields in (2-9). A *normalized* tensor field will be in the form (1-6), with possibly certain pairs of indices in certain of the factors S_*R_{ijkl} being symmetrized over.

Let us first introduce a few definitions: Given each $C_g^{l,i_1\dots i_\mu}$, we list out the factors $T_1,\dots T_{\sigma_2}$ in the form S_*R_{ikl}. Here T_a is the factor for which the index ${}_i$ is contracting against the factor $\nabla\tilde{\phi}_a$. We say that factors S_*R_{ikl} are of type I if they contain no free index. We say they are of type II if they contain a special free index. We say they are of type III if they contain a nonspecial free index.

Given any tensor field $C_g^{l,i_1\dots i_\mu}$ in the form (1-6), pick out the pairs of factors T_α, T_β in the form S_*R_{ijkl} for which a special index in T_α contracts against a special

[32] The terms indexed in L^1_μ are now simply subsequent to $\tilde{\kappa}_{\mathrm{simp}}$.

index in T_β. (Call such particular contractions "special-to-special" particular contractions.) Now, in any $C^{l,i_1\dots i_\mu}_g$ we define an ordering among all its factors S_*R_{ijkl}: The factor $T_a=S_*R_{ikl}\nabla^i\tilde{\phi}_a$ is more important than $T_b=S_*R_{i'j'k'l'}\nabla^{i'}\tilde{\phi}_b$ if $a<b$.

Now, consider a tensor field $C^{l,i_1\dots i_\mu}_g$ and list out all the pairs of factors T_a, T_b with a special-to-special particular contraction. We say that (T_a, T_b) is the *most important* pair of factors with a special-to-special particular contraction[33] if any *other* such pair (T_c, T_d)[34] has either T_c less important than T_a or $T_a=T_c$ and T_d less important than T_b.

Now, consider a tensor field $C^{l,i_1\dots i_\mu}_g$ and consider the most important pair of factors (T_a, T_b) with a special-to-special particular contraction. Assume without loss of generality that the index ${}_l$ in $T_a=S_*R_{ijkl}\nabla^i\tilde{\phi}_a$ contracts against the index ${}_{l'}$ in $T_b=S_*R_{i'j'k'l'}\nabla^{i'}\tilde{\phi}_b$. We say that $C^{l,i_1\dots i_\mu}_g$ is *normalized* if both factors T_a, T_b are normalized. The factor $T_a=S_*R_{ikl}\nabla^i\tilde{\phi}_a$ is normalized if: Either the index ${}_j$ contracts against a factor T_c that is more important than T_b, *or* if the indices ${}_j, {}_k$ are symmetrized. If T_a is of type II, then we require that the index ${}_j$ in $T_b=S_*R_{ij(free)l}$ *must* contract against a special index of some other factor T_c, and moreover T_c must be more important than T_b. If T_a is of type III, then it is automatically normalized. The same definition applies to T_b, where any reference to T_b must be replaced by a reference to T_a.

Let us now prove that we may assume without loss of generality that all μ-tensor fields in (2-9) are normalized: Consider a $C^{l,i_1\dots i_\mu}_g$ in (2-9) for which the most important pair of factors with a special-to-special particular contraction is the pair (T_a, T_b). We will prove that we can write

$$C^{l,i_1\dots i_{\mu+\beta}}_g=\tilde{C}^{l,i_1\dots i_{\mu+\beta}}_g+\sum_{t\in T}a_tC^{t,i_1\dots i_{\mu+\beta}}_g; \tag{2-30}$$

here the term $\tilde{C}^{l,i_1\dots i_{\mu+\beta}}_g$ is normalized, the most important pair of factors with a special-to-special particular contraction is the pair (T_a, T_b), and moreover its refined double character is the same as for $C^{l,i_1\dots i_{\mu+\beta}}_g$. Each term $C^{t,i_1\dots i_{\mu+\beta}}_g$ has either the same, or a doubly subsequent refined double character to $C^{l,i_1\dots i_{\mu+\beta}}_g$; moreover in the first case its most important pair of factors with a special-to-special particular contraction will be less important than the pair (T_a, T_b). In the second case the most important pair will either be (T_a, T_b) or a less important pair.

Clearly, if we can prove the above, then by iterative repetition we may assume without loss of generality that all $(\mu+\beta)$-tensor fields in (2-9) are normalized.

Proof of (2-30). Pick out the most important pair of factors with a special-to-special particular contraction is the pair (T_a, T_b) in $C^{l,i_1\dots i_{\mu+\beta}}_g$. Let us first normalize

[33] Assume without loss of generality that T_a is more important than T_b.

[34] Again assume without loss of generality that T_c is more important than T_d.

T_a. If T_a is of type III, there is nothing to do. If it is of type II and already normalized, there is again nothing to do. If it is of type II and not normalized, then we *interchange* the indices ${}_j, {}_k$. The resulting factor *is* normalized. The correction term we obtain by virtue of the first Bianchi identity is also normalized (it is of type III). Moreover, the resulting tensor field is doubly subsequent to $C^{l,i_1\dots i_{\mu+\beta}}_g$. Finally, if the factor T_a is of type I, we inquire on the factor T_c against which ${}_j$ in $T_a = S_* R_{ijkl}$ contracts: If it is more important than T_b, then we leave T_a as it is; it is already normalized. If not, we symmetrize ${}_j, {}_k$. The resulting tensor field is normalized. The correction term we obtain by virtue of the first Bianchi identity will then have the same refined double character as $C^{l,i_1\dots i_{\mu+\beta}}_g$, and moreover its most important pair of factors with a special-to-special particular contraction is less important than that pair (T_a, T_b). □

We may now prove the claim of Lemma 2.6 in this special case, under the additional assumption that all tensor fields in (2-9) are normalized. We list out the most important pair of special-to-special particular contractions in each $C^{l,i_1\dots i_{\mu+\beta}}_g$, and denote it by $(a, b)_l$. We let (α, β) stand for the lexicographically minimal pair among the list $(a, b)_l$, $l \in L_\mu$. We denote by $L^{(\alpha,\beta)}_\mu \subset L_\mu$ the index set of terms with a special-to-special particular contraction among the terms T_α, T_β. We will prove that

$$\sum_{l\in L^{(\alpha,\beta)}_\mu} a_l C^{l,i_1\dots i_{\mu+\beta}}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon = 0. \tag{2-31}$$

Clearly, the above will imply our claim, by iterative repetition.[35]

Proof of (2-31). Consider $Image^2_{Y_1,Y_2}[L_g] = 0$ and pick out the sublinear combination where the factors T_α, T_β are replaced by $\nabla^{(A)}Y_1 \otimes g$, $\nabla^{(B)}Y_2 \otimes g$, and the two factors $\nabla\tilde{\phi}_\alpha$, $\nabla\tilde{\phi}_\beta$ contract against each other. The resulting sublinear combination must vanish separately. We erase the expression $\nabla_t\tilde{\phi}_\alpha\nabla^t\tilde{\phi}_\beta$,[36] and derive a new true equation in the form

$$\sum_{l\in L^{(\alpha,\beta)}_\mu} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\mu}\tilde{C}^{l,i_1\dots i_{\mu+\beta}}_g(\Omega_1, Y_1, Y_2) + \sum_{j\in J} a_j C^j_g(\Omega_1, Y_1, Y_2) = 0; \tag{2-32}$$

here the tensor fields $\tilde{C}^{l,i_1\dots i_{\mu+\beta}}_g(\Omega_1, Y_1, Y_2)$ arise from the tensor fields $C^{l,i_1\dots i_{\mu+\beta}}_g$ by replacing the expression $\nabla^i\tilde{\phi}_\alpha S_* R_{ijkl} \otimes S_* R_{i'jk}{}^l \nabla^{i'}\tilde{\phi}_\beta$ with $\nabla_{jk}Y_1 \otimes \nabla_{j'k'}Y_2$ (notice we have lowered the weight in absolute value).

[35] In the subcase $\mu + \beta = \sigma_2$ it will only imply it for the "excluded" sublinear combination defined above.

[36] Denote the resulting $(u-2)$-simple character by $\vec{\kappa}'''_{\mathrm{simp}}$.

Now, applying the inductive assumption of Lemma 2.6 to the above,[37] we derive

$$\sum_{l\in L_\mu^{(\alpha,\beta)}} a_l \tilde{C}_g^{l,i_1\dots i_{\mu+\beta}}(\Omega_1, Y_1, Y_2)\nabla_{i_1}\upsilon\dots\nabla_{i_\mu}\upsilon=0. \tag{2-33}$$

The proof of (2-31) is only one step away. Let us start with an important observation: For each given complete contraction above, examine the factor $\nabla^{(2)}_{zx}Y_1$; it either contracts against no factor $\nabla\upsilon$ or against one factor $\nabla\upsilon$.[38] In the first case, the factor $\nabla^{(2)}_{zx}Y_1$ must have arisen from a factor S_*R_{ijkl} of type I. In fact, the indices ${}_z, {}_x$ correspond to the indices ${}_j, {}_k$ in the original factor, and we can even determine their position: Since the pair (α,β) is the most important pair in (2-9), at most one of the indices ${}_z, {}_x$ can contract against a special index in a more important factor than T_β. If one of them does (say ${}_z$), then that index must have been the index ${}_j$ in $T_\alpha = S_*R_{ikl}$. If none of them does, then the two indices ${}_z, {}_x$ must be symmetrized over, since the two indices ${}_j, {}_k$ in T_α to which they correspond were symmetrized over. Now, these two separate sublinear combinations in (2-33) must vanish separately (this can be proven using the eraser from the Appendix in [A 2012]), and furthermore in the first case, we may assume that the index ${}_z$ (which contracts against a special index in a more important factor than T_β) occupies the leftmost position in $\nabla^{(2)}_{zx}Y_1$ and is *not permuted* in the formal permutations of indices that make (2-33) hold *formally*).

On the other hand, consider the terms in (2-33) with the factor $\nabla^{(2)}Y_1$ contracting against a factor $\nabla\upsilon$. By examining the index ${}_y$ in the factor $\nabla^{(2)}_{yt}Y_1\nabla^t\upsilon$, we can determine the *type* of factor in $C_g^{l,i_1\dots i_{\mu+\beta}}$ from which the factor $\nabla^{(2)}Y_1$ arose: If the index ${}_y$ contracts against a special index in a factor S_*R_{ijkl} which is *more important* than T_β, then $\nabla^{(2)}Y_1$ can only have arisen from a factor of type II in $C_g^{l,i_1\dots i_{\mu+\beta}}$. In fact, the index ${}_y$ in $\nabla^{(2)}Y_1$ must correspond to the index ${}_j$ in $S_*R_{ij(\mathit{free})l}$ in T_α. If the index ${}_y$ in $\nabla^{(2)}_{yt}Y_1\nabla^t\upsilon$ does not contract against a special index in a factor T_c which is more important than T_β, then the factor $\nabla^{(2)}Y_1$ can only have arisen from a factor of type III in $C_g^{l,i_1\dots i_{\mu+\beta}}$. In fact, the index ${}_y$ in $\nabla^{(2)}Y_1$ must correspond to the index ${}_k$ in $S_*R_{i(\mathit{free})kl}$ in T_α.

The same analysis can be repeated for the factor $\nabla^{(2)}Y_2$, with any reference to the factor T_β now replaced by the factor T_α.

In view of the above analysis, we can break the left-hand side of (2-33) into four sublinear combinations that vanish separately (depending on whether $\nabla^{(2)}Y_1$, $\nabla^{(2)}Y_2$ contract against a factor $\nabla\upsilon$ or not). Then in each of the four sublinear combinations, we can arrange that in the formal permutations that make the left-hand side of (2-33) formally zero, the two indices in the factors $\nabla^{(2)}Y_1$, $\nabla^{(2)}Y_2$ are

[37] We have lowered the weight in absolute value.

[38] The two corresponding sublinear combinations vanish separately, of course.

not permuted (by virtue of the remarks above). In view of this and the analysis in the previous paragraph, we can then *replace* the two factors $\nabla^{(2)}_{zx}Y_1$, $\nabla^{(2)}_{qw}Y_2$ by an expression $\nabla^i\tilde{\phi}_\alpha S_*R_{izxl}\otimes S_*R_{i'qw}{}^l\nabla^{i'}\tilde{\phi}_\beta$, in such a way that the resulting linear combination *vanishes formally without permuting the two indices* ${}_q, {}_w, {}_{q'}, {}_{w'}$. This proves our claim, except for the subcase $\mu+\beta=\sigma_2$ where we only derive our claim for all terms except for the "bad sublinear combination". We now prove our claim for that case.

The "bad sublinear combination". We break up the left-hand side of (2-16) according to which factor T_s the factor $\nabla\phi_{u+1}$ contracts — denote the index set of those terms by L^K_μ. Denote the resulting sublinear combinations by L^K_g, $K=1,\dots,\sigma_2$. Given any K, we consider the equation $Image^1_Y[L_g]=0$, and we pick out the sublinear combination where the term $\nabla^{(B)}S_*R_{ijkl}\nabla^i\tilde{\phi}_K$ is replaced by $\nabla^{(B+2)}Y$, and the factor $\nabla\phi_K$ now contracts against the factor $\nabla\phi_{u+1}$. This sublinear combination must vanish separately. We then again perform the "inverse integration by parts" to this true equation (deriving an integral equation), and then we consider the silly divergence formula for this integral equation, obtained by integrating by parts with respect to $\nabla^{(B)}Y$. We pick out the sublinear combination with $\sigma+u+\beta$ factors, μ internal contractions and $u+\beta$ factors $\nabla\phi_h$, and an expression $\nabla_s\phi_{u+1}\nabla^s\tilde{\phi}_K$ This gives us a new true local equation,

$$\sum_{l\in L^K_\mu} a_l X_*\operatorname{div}_{j_1} X_*\operatorname{div}_{j_2}\tilde{C}^{l,j_1j_2}_g+\sum_{j\in J}a_jC^j_g=0. \tag{2-34}$$

Here the tensor fields $\tilde{C}^{l,j_1j_2}_g$ arise from $C^{l,i_1\dots i_\mu}_g$ by formally replacing all μ free indices with internal contractions, and also replacing $\nabla_x\phi_{u+1}\otimes S_*R_{i(jk)}{}^x\nabla^i\tilde{\phi}_K$ with $\nabla_x\phi_{u+1}\nabla^s\tilde{\phi}_K\otimes Y$, and then making the indices ${}^j, {}^k$ that contracted against ${}_j, {}_k$ into free indices ${}^{j_1}, {}^{j_2}$. $X_*\operatorname{div}_j$ stands for the sublinear combination in $X\operatorname{div}_j$ where ∇^j is not allowed to hit the factor Y. Now, applying the inductive assumption of Lemma 2.6 to the above,[39] we derive that

$$\sum_{l\in L^K_\mu} a_l\tilde{C}^{l,{}_1j_2}_g\nabla_{j_1}\omega\nabla_{j_2}\omega=0.$$

We replace $\nabla^x\phi_K\nabla_x\phi_{u+1}\nabla_{j_1}\omega\nabla_{j_2}\omega\nabla_lY$ with $\nabla^l\phi_{u+1}S_*R_{i(j_1j_2)l}\nabla^i\tilde{\phi}_K$ and then replace all internal contractions by factors $\nabla\upsilon$ (applying the operation Sub_υ from the Appendix in [A 2012]). The resulting (true) equation is precisely our remaining claim for the "bad" sublinear combination. □

[39] We have lowered the weight in absolute value.

2C. *Proof of Lemmas 4.6, 4.8 in [A 2010]: The main part.* We first write down the form of the complete and partial contractions that we are dealing with in Lemmas 2.1 and 2.3. In the setting of Lemma 2.1 we recall that the tensor fields $C^{h,i_1\dots i_\alpha}$ indexed in H_2 (in the hypothesis of Lemma 2.1) are all partial contractions in the form

$$pcontr\big(\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_{\sigma_1})}R_{ijkl}\otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\cdots\otimes S_*\nabla^{(\nu_t)}R_{ijkl} \otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla Y\otimes\nabla\phi_{z_1}\cdots\otimes\nabla\phi_{z_f} \otimes\nabla\phi'_{z_{f+1}}\otimes\cdots\otimes\nabla\phi'_{z_{f+d}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{f+d+1}}\otimes\cdots\otimes\nabla\tilde{\phi}_{z_{f+d+y}}\big), \tag{2-35}$$

where we let $f+d+y=u'$. The main assumption here is that all tensor fields have the same u'-simple character (the one defined by $\nabla\phi_1,\dots,\nabla\phi_{u'}$), which we denote by $\vec{\kappa}^+_{\text{simp}}$. The other main assumption is that if we formally treat the factor ∇Y as a function $\nabla\phi_{u+1}$, then the hypothesis of Lemma 2.1 falls under the inductive assumptions of Proposition 1.1 (i.e., the weight, real length, Φ and p are as in our inductive assumption of Proposition 1.1).

In the setting of Lemma 2.3 we recall that we are dealing with complete and partial contractions in the form

$$contr\big(\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_{\sigma_1})}R_{ijkl}\otimes S_*\nabla^{(\nu_1)}R_{ijkl}\otimes\cdots\otimes S_*\nabla^{(\nu_t)}R_{ijkl} \otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes[\nabla\omega_1\otimes\nabla\omega_2]\otimes\nabla\phi_{z_1}\cdots\otimes\nabla\phi_{z_f} \otimes\nabla\phi'_{z_{f+1}}\otimes\cdots\otimes\nabla\phi'_{z_{f+d}}\otimes\cdots\otimes\tilde{\phi}_{z_{f+d+1}}\otimes\cdots\otimes\tilde{\phi}_{z_{f+d+y}}\big), \tag{2-36}$$

where we let $f+d+y=u'$. The main assumption here is that all partial contractions have the same u'-simple character (the one defined by $\nabla\phi_1,\dots,\nabla\phi_{u'}$), which we denote by $\vec{\kappa}^+_{\text{simp}}$. The other main assumption is that if we formally treat the factors $\nabla\omega_1$, $\nabla\omega_2$ as factors $\nabla\phi_{u+1}$, $\nabla\phi_{u+2}$, then the hypothesis of Lemma 2.3 falls under the inductive assumptions of Proposition 1.1 (i.e., the weight, real length, Φ and p are as in our inductive assumption of Proposition 1.1).

Note. From now on, we will be writing $u'=u$, for simplicity. We will also be writing $\vec{\kappa}^+_{\text{simp}}=\vec{\kappa}_{\text{simp}}$, for simplicity. We will also be labeling the indices ${}_{i_1},\dots,{}_{i_\alpha}$ as ${}_{i_{\pi+1}},\dots,{}_{i_{\alpha+1}}$.

New induction. We will now prove the two Lemmas 2.1 and 2.3 by a new induction on the weight of the complete contractions in the hypotheses of those lemmas. We will assume that these two lemmas are true when the weight of the complete contractions in their hypotheses is $-W$, for any $W<K\le n$. We will then show our lemmas for weight $-K$.

Reduce Lemma 2.1 to two lemmas. In order to show Lemma 2.1, we further break up H_2 into subsets: We say that $h\in H_2^a$ if $C^{h,i_{\pi+1}\dots i_{\alpha+1}}$ has a free index (say the free index ${}_{i_{\alpha+1}}$ without loss of generality) belonging to the factor ∇Y. On the other hand, we say that $h \in H_2^b$ if the index in the factor ∇Y is not free. Lemma 2.1 will then follow from Lemmas 2.7 and 2.8 below.

Lemma 2.7. *There exists a linear combination of acceptable $(\alpha-\pi+1)$-tensor fields, $\sum_{v\in V} a_v C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$, where the index $_{i_{\alpha+1}}$ belongs to the factor ∇Y, with a simple character $\vec{\kappa}_{\text{simp}}$, so that*

$$\sum_{h\in H_2^a} a_h C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon$$
$$= \sum_{v\in V} a_v X_* \operatorname{div}_{i_{\alpha+2}} C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u) \times \nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon$$
$$+\sum_{j\in J} a_j C_g^{j,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon. \tag{2-37}$$

Each C^j is simply subsequent to $\vec{\kappa}_{\text{simp}}$.

We observe that if we can show our first claim, then we can assume, with no loss of generality, that $H_2^a=\varnothing$, since it immediately follows from the above that

$$\sum_{h\in H_2^a} a_h X_* \operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$$
$$= \sum_{v\in V} a_v X_* \operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}} X_*\operatorname{div}_{i_{\alpha+2}} \times C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u), \tag{2-38}$$

where each complete contraction C^j is subsequent to $\vec{\kappa}_{\text{simp}}$. (Note that one of the free indices in the tensor fields $C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}$ will belong to the factor ∇Y.)

The second claim, in the setting of Lemma 2.1 is:

Lemma 2.8. *We assume $H_2^a=\varnothing$. We then claim that modulo complete contractions of length $\ge \sigma+u+1$,*

$$\sum_{h\in H_2} a_h C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon$$
$$= \sum_{t\in T} a_t X_* \operatorname{div}_{i_{\alpha+2}} C_g^{t,i_{\pi+1}\dots i_{\alpha+2}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u) \times \nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u), \tag{2-39}$$

where each C^j is acceptable and subsequent to $\vec{\kappa}_{\text{simp}}$.

We observe that if we can show the above two lemmas then Lemma 2.1 will follow. (Notice that replacing by the right-hand side of (2-38) into the hypothesis of Lemma 2.1, we do not introduce 1-forbidden terms.)

We make two analogous claims for Lemma 2.3:

Reduce Lemma 2.3 to two lemmas. We say that $h \in H_2^a$ if $C^{h,i_{\pi+1}\dots i_{\alpha+1}}$ has a free index belonging to one of the factors $\nabla\omega_1, \nabla\omega_2$. On the other hand, we say that $h \in H_2^b$ if none of the factors $\nabla\omega_1, \nabla\omega_2$ in $C^{h,i_{\pi+1}\dots i_{\alpha+1}}$ contains a free index. (Observe that we may assume with no loss of generality that there are no tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}$ with free indices in both factors $\nabla\omega_1, \nabla\omega_2$ — this is by virtue of the antisymmetry of the factors $\nabla\omega_1, \nabla\omega_2$.) We make two claims. First:

Lemma 2.9. *There is a linear combination of acceptable* $(\alpha-\pi+1)$*-tensor fields,* $\sum_{v\in V} a_v C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$, *in the form* (2-35) *with a simple character* $\vec{\kappa}_{\text{simp}}$, *so that*

$$
\begin{aligned}
\text{(2-40)}\quad \sum_{h\in H_2^a} a_h X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}}& \\
\times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}&(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u) \\
= \sum_{v\in V} a_v X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}} X_+ \operatorname{div}_{i_{\alpha+2}}& \\
\times C_g^{v,i_{\pi+1}\dots i_{\alpha+2}}&(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u) \\
+ \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}}& \\
\times C_g^{q,i_{\pi+1}\dots i_{\alpha+1}}&(\Omega_1,\dots,\Omega_p,\nabla_+[\omega_1,\omega_2],\phi_1,\dots,\phi_u) \\
+ \sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u).&
\end{aligned}
$$

(*Recall that by definition the complete contractions indexed in Q have a factor* $\nabla^{(2)}\omega_1$.)

We observe that if we can show our first claim, then we can, with no loss of generality, assume that $H_2^a = \varnothing$.

Second claim:

Lemma 2.10. *We assume* $H_2^a = \varnothing$, *and that for some* $k \ge 1$, *we can write*

$$
\begin{aligned}
\text{(2-41)}\quad \sum_{h\in H_2^b} a_x X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}}& \\
\times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}&(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u) \\
= \sum_{t\in T_k} a_t X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+k}}& \\
\times C_g^{t,i_{\pi+1}\dots i_{\alpha+k}}&(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)
\end{aligned}
$$

$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{\pi+1} \dots X_+ \operatorname{div}_{i_{\alpha+1}} \times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\nabla_+[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$$
$$+\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u),$$

where the last two linear combinations on the left-hand side of the equality are generic linear combinations in the form described in the claim of Lemma 2.3.[40] *On the other hand,*

$$\sum_{t\in T_k} a_t C_g^{t,i_{\pi+1}\dots i_{\alpha+k}}(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$$

is a linear combination of acceptable $(\alpha-\pi+k)$*-tensor fields in the form (2-36) with a simple character* $\vec{\kappa}_{\text{simp}}$*, and with two antisymmetric factors* $\nabla\omega_1$, $\nabla\omega_2$ *that do not contain a free index. We then claim that modulo complete contractions of length* $\geq \sigma+u+1$ *we can write*

$$\sum_{t\in T_k} a_t X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_{a+k}} \times C_g^{t,i_1\dots i_{a+k}}(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$$
$$=\sum_{t\in T_{k+1}} a_t X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{a+k+1}} \times C_g^{t,i_{\pi+1}\dots i_{a+k+1}}(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$$
$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_{a+1}} \times C_g^{q,i_1\dots i_{a+1}}(\Omega_1,\dots,\Omega_p,\nabla_+[\omega_1,\omega_2],\phi_1,\dots,\phi_u)$$
$$+\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u), \tag{2-42}$$

with the same notational conventions as above.

We observe that if we can show the above two claims, then Lemma 2.3 will follow by iterative repetition of the second claim.

We will now show the four lemmas above.

Proof of Lemmas 2.8 and 2.10. Lemma 2.8 is a direct consequence of Lemma 4.10 in [A 2010].[41] Lemma 2.10 can be proven in two steps: First, by Lemma 2.5 we derive that there exists a linear combination of acceptable $(a+k+1)$-tensor fields

[40] In Lemma 2.3, Q is called V.

[41] Observe that our hypotheses on the tensor fields in the equation in Lemma 2.1 not being "bad" ensure that we do not fall under the "forbidden" cases of Lemma 4.10 in [A 2010].

(indexed in X below) with a u-simple character $\vec{\kappa}_{\text{simp}}$ so that

$$\begin{aligned}(2\text{-}43)\quad &\sum_{t\in T_k} a_t C^{t,i_1\dots i_{a+k}}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)\nabla_{i_1}\upsilon\dots\nabla_{i_{a+k}}\upsilon\\ &\quad-\sum_{t\in T_{k+1}} a_t X_*\operatorname{div}_{i_{a+k+1}} C^{t,i_{\pi+1}\dots i_{a+k+1}}_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u)\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\times\nabla_{i_1}\upsilon\dots\nabla_{i_{a+k}}\upsilon\\ &\quad=\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u,\upsilon^{a+k}),\end{aligned}$$

where the complete contractions indexed in J have length $\sigma+a+k+1$ and are simply subsequent to $\vec{\kappa}_{\text{simp}}$. Then, making each factor $\nabla\upsilon$ in the above into an X_+ div, we derive Lemma 2.10. □

Proof of Lemma 2.7. We have denoted by $\vec{\kappa}_{\text{simp}}$ the simple character of our tensor fields. We distinguish two cases: In Case A there is a factor $\nabla^{(m)}R_{ijkl}$ in $\vec{\kappa}_{\text{simp}}$, and in Case B there is no such factor.

We denote $\alpha+1=\gamma$, for brevity.

Now we break the set H^b_2 into subsets: In Case A we say that $h\in H^{b,+}_2$ if ∇Y contracts against an internal index of a factor $\nabla^{(m)}R_{ijkl}$. In Case B we say that $h\in H^{b,+}_2$ if ∇Y contracts against one of the indices ${}_k, {}_l$ in a factor $S_*\nabla^{(\nu)}R_{ijkl}$.

We define $H^{b,-}_2=H^b_2\setminus H^{b,+}_2$.

In each of the above cases and subcases we treat the term ∇Y as a term $\nabla\phi_{u+1}$ in our lemma hypothesis. Then, by applying the first claim in Lemma 4.10 in [A 2010][42] to our lemma hypothesis and then making each $\nabla\upsilon$ into an X_* div, we derive that we can write

$$\begin{aligned}(2\text{-}44)\quad &X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_\gamma}\sum_{h\in H^{b,+}_2} a_h C^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ &\quad=X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_\gamma}\sum_{h\in H^{b,*,-}_2} a_h C^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ &\qquad\qquad\qquad\qquad\qquad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u),\end{aligned}$$

where $\sum_{h\in H^{b,*,-}_2} a_h C^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ stands for a *generic* linear combination as defined above (i.e., it is in the general form $\sum_{h\in H^b_2}\dots$ but the factor ∇Y is *not* contracting against a special index in any factor $\nabla^{(m)}R_{ijkl}$

[42]By weight considerations, since we started out with no “bad terms” in Lemma 2.1, we will not encounter no “forbidden tensor fields” for Lemma 4.10 in [A 2010].

or $S_*\nabla^{(\nu)}R_{ijkl}$).[43] On the other hand, each $C^j_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ is a complete contraction with a simple character that is subsequent to $\vec{\kappa}_{\text{simp}}$.

Thus, by virtue of (2-44), we reduce ourselves to the case where $H_2^{b,+}=\varnothing$. We will then show Lemma 2.7 separately in Cases A and B, under the assumption that $H_2^{b,+}=\varnothing$.

Proof of Lemma 2.7 in Case A. We will define the C-crucial factor, for the purposes of this proof only: We denote by *Set* the set of numbers u for which $\nabla\phi_u$ contracts against one of the factors $\nabla^{(m)}R_{ijkl}$. If $Set\neq\varnothing$, we define u_+ to be the minimum element of *Set*, and we pick out the factor $\nabla^{(m)}R_{ijkl}$ in each C^h against which $\nabla\phi_{u_+}$ contracts. We call that factor $\nabla^{(m)}R_{ijkl}$ *C-crucial*. If $Set=\varnothing$, we will say *the C-crucial factors* and will mean *any* of the factors $\nabla^{(m)}R_{ijkl}$.

Now we pick out the subset $H_2^{b,*}\subset H_2^b$, that is defined by the rule $h\in H_2^{b,*}$ if ∇Y contracts against the (one of the) C-crucial factor.

Now, for each $h\in H_2^a$ we denote by

$$\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$$

the sublinear combination in $X_*\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ that arises when ∇_{i_γ} hits the (one of the) C-crucial factor.[44] It then follows that

(2-45)
$$\begin{aligned}\sum_{h\in H_2^a} a_h X_*\,\text{div}_{i_{\pi+1}}\dots X_*\,\text{div}_{i_\alpha}\,\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ +\sum_{h\in H_2^{b,*}} a_h X_*\,\text{div}_{i_{\pi+1}}\dots X_*\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ =\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u),\end{aligned}$$

where each C^j_g has the factor ∇Y contracting against the C-crucial factor $\nabla^{(m)}R_{ijkl}$ and is simply subsequent to $\vec{\kappa}_{\text{simp}}$.

Denote the $(u+1)$-simple character (the one defined by $\nabla\phi_1,\dots,\nabla\phi_{u+1}=\nabla Y$) of the tensor fields $\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ by $\vec{\kappa}'_{\text{simp}}$. (Observe that they all have the same $(u+1)$-simple character.)

[43] Recall that a special index in a factor $\nabla^{(m)}R_{ijkl}$ is an internal index, while a special index in a factor $S_*\nabla^{(\nu)}R_{ijkl}$ is an index ${}_{k},{}_{l}$.

[44] Recall that ${}_{i_\gamma}$ is the free index that belongs to ∇Y.

We observe that by applying Corollary 1 in [A 2010] to (2-45) (all tensor fields are acceptable and have the same simple character $\vec{\kappa}'_{\mathrm{simp}}$),[45] we obtain

$$\begin{aligned}(2\text{-}46)\quad &\sum_{h\in H_2^a} a_h \,\mathrm{Hit}\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon\\ &+\sum_{u\in U} a_u X\,\mathrm{div}_{i_{\alpha+1}} C_g^{u,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ &\qquad\qquad\times\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon\\ &=\sum_{j\in J} a_j C_g^{j,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon=0,\end{aligned}$$

where the tensor fields indexed in U are acceptable (we are treating ∇Y as a factor $\nabla\phi_{u+1}$), have a simple character $\vec{\kappa}'_{\mathrm{simp}}$ and each C^j is simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$.

But then, our first claim follows almost immediately. We recall the operation $\mathrm{Erase}_{\nabla Y}[\dots]$ from the Appendix in [A 2012] which acts on the complete contractions in the above by erasing the factor ∇Y and the (derivative) index that it contracts against. Then, since (2-46) holds formally, we have that the tensor field required for Lemma 2.7 is

$$\sum_{u\in U} a_u\,\mathrm{Erase}_{\nabla Y}[C_g^{u,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)]\cdot\nabla_{i_\gamma}Y.$$

Proof of Lemma 2.7 in Case B. We again distinguish two subcases: In Subcase (i) there is some nonsimple factor $S_*\nabla^{(\nu)}R_{ijkl}$ in $\vec{\kappa}_{\mathrm{simp}}$ or a nonsimple factor $\nabla^{(B)}\Omega_x$ contracting against two factors $\nabla\phi'_h$ in $\vec{\kappa}_{\mathrm{simp}}$. In Subcase (ii) there are no such factors.

In Subcase (i), we arbitrarily pick out one factor $S_*\nabla^{(\nu)}R_{ijkl}$ or $\nabla^{(B)}\Omega_x$ with the properties described above and call it the D-crucial factor. In this first subcase we will show our claim for the whole sublinear combination $\sum_{h\in H_2^a}\dots$ in one piece.

In Subcase (ii), we will introduce some notation: We will examine each factor $T=S_*\nabla^{(\nu)}R_{ijkl}$, $T=\nabla^{(B)}\Omega_x$ in each tensor field $C_g^{h,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}$ and define its "measure" as follows: If $T=S_*\nabla^{(\nu)}R_{ijkl}$ then its "measure" will stand for its total number of free indices *plus* $\frac{1}{2}$. If $T=\nabla^{(B)}\Omega_x$ then its "measure" will stand for its total number of free indices *plus* the number of factors $\nabla\phi_h$ against which it contracts.

We divide the index set H_2^a into subsets according to the measure of any given factor. We denote by M the maximum measure among all factors among the tensor fields $C_g^{h,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}$, $h\in H_2^a$. We denote by $H_a^{2,*}\subset H_2^a$ the index set of the tensor fields that contain a factor of maximum measure. We will show the claim of

[45] Notice that by weight considerations, since we started out with no "bad" terms in the hypothesis of Lemma 2.1, there is no danger of falling under a "forbidden case" of that corollary.

Lemma 2.7 for the sublinear combination $\sum_{h\in H_a^{2,*}}\dots$. Clearly, if we can do this, then Lemma 2.7 will follow by induction.

We will prove Lemma 2.7 in the second subcase (which is the hardest). The proof in the first subcase follows by the same argument, only by disregarding any reference to M free indices belonging to a given factor and so on.

Proof of Lemma 2.7 in Case B for the sublinear combination $\sum_{h\in H_a^{2,*}}\dots$. We will further divide $H_a^{2,*}$ into subsets, $H_a^{2,*,k}$, $k=1,\dots,\sigma$, according to the factor of maximum measure: First, we order the factors $S_*\nabla^{(\nu)}R_{ijkl},\dots\nabla^{(p)}\Omega_h$ in $\vec{\kappa}_{\text{simp}}$, and label them $T_1,\dots,T_\sigma$ (observe each factor is well-defined in $\vec{\kappa}_{\text{simp}}$, because we are in Case B). We then say that $h\in H_2^{a,*,1}$ if in $C_g^{u,i_{\pi+1}\dots i_\alpha}$ the factor T_1 has measure M. We say that $h\in H_2^{a,*,2}$ if in $C_g^{u,i_{\pi+1}\dots i_\alpha}$ the factor T_2 has measure M and T_1 has measure less than M, and so on. We will then prove our claim for each of the index sets $h\in H_2^{a,*,k}$:[46] We arbitrarily pick a $k\le K$ and show our claim for $\sum_{h\in H_a^{2,*,k}}\dots$.

For the purposes of this proof, we call the factor T_k the D-crucial factor.

Now we pick out the subset $H_2^{b,k}\subset H_2^b$, that is defined by the rule $h\in H_2^{b,k}$ if and only if ∇Y is contracting against the D-crucial factor T_k.

Now, for each $h\in H_2^a$ we denote by

$$\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$$

the sublinear combination in $X\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ that arises when ∇_{i_γ} hits the D-crucial factor.[47] It then follows that

$$\sum_{h\in H_2^a} a_h X\,\text{div}_{i_{\pi+1}}\dots X\,\text{div}_{i_\alpha}\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u) \\ +\sum_{h\in H_2^{b,k}} a_h X\,\text{div}_{i_{\pi+1}}\dots X\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u) \\ =\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u), \tag{2-47}$$

where each C_g^j has the factor ∇Y contracting against the D-crucial factor and is simply subsequent to $\vec{\kappa}_{\text{simp}}$.

Denote the $(u+1)$-simple character (the one defined by $\nabla\phi_1,\dots,\nabla\phi_{u+1}=\nabla Y$) of the tensor fields $\text{Hit}\,\text{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)$ by $\vec{\kappa}'_{\text{simp}}$. (Observe that they all have the same $(u+1)$-simple character.)

We apply Corollary 1 in [A 2010] to (2-47) (all tensor fields are acceptable and have the same simple character $\vec{\kappa}'_{\text{simp}}$) and then pick out the sublinear combination

[46] Again we observe that if we can prove this then Lemma 2.7 in Case B will follow by induction.

[47] Recall that $i_\gamma=i_{\alpha+1}$ belongs to ∇Y by hypothesis.

where there are M factors $\nabla\upsilon$ or $\nabla\phi_h$ or $\nabla\phi'_h$ contracting against T_k, obtaining

$$\sum_{h\in H_2^{a,*,k}} a_h \,\mathrm{Hit}\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon \tag{2-48}$$
$$+\sum_{u\in U} a_u X\,\mathrm{div}_{i_{\alpha+1}} C_g^{h,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u) \times\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon$$
$$=\sum_{j\in J} a_j C_g^{j,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon=0,$$

where the tensor fields indexed in U are acceptable and have a simple character $\vec{\kappa}'_{\mathrm{simp}}$ and each C^j is simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$.

Now, observe that if $M\ge\frac{3}{2}$, we can apply the eraser to ∇Y (see the Appendix in [A 2012]) and the index it contracts against in the D-crucial factor and derive our conclusion as in Case A.

On the other hand, in the remaining cases[48] the above argument cannot be directly applied. In those cases, we derive our claim as follows:

In the case $M=1$ the D-crucial factor is of the form $\nabla^{(p)}\Omega_h$, then we cannot directly derive our claim by the above argument, because if for some tensor fields in U above we have ∇Y contracting according to the pattern $\nabla_i Y\nabla^{ij}\Omega_h\nabla_j\psi$ (where $\psi=\upsilon$ or $\psi=\phi_h$), then we will not obtain acceptable tensor fields after we apply the eraser. Therefore, if $M=1$ and the D-crucial factor is of the form $\nabla^{(p)}\Omega_h$, we apply Lemma 4.6 in [A 2010] to (2-48) (treating the factors $\nabla\upsilon$ as factors $\nabla\phi$)[49] to obtain a new equation in the form (2-48), where for any tensor field indexed in U the factor ∇Y contracts against a factor $\nabla^{(l)}\Omega_h$, $l\ge 3$.[50] Then, applying the eraser as explained, we derive our Lemma 2.7 in this case.

When $M=\frac{1}{2}$ or $M=0$, then we first apply the inductive assumptions of Corollaries 3 and 2 in [A 2010] (respectively) to (2-48),[51] in order to assume with no loss of generality that for each tensor field indexed in U there, the factor ∇Y either contracts against a factor $\nabla^{(B)}\Omega_h$, $B\ge 3$ or a factor $S_*\nabla^{(\nu)}R_{ijkl}$, $\nu\ge 1$. Then the eraser can be applied and it produces acceptable tensor fields. Hence, applying $\mathrm{Erase}_{\nabla Y}$ to (2-48) we derive our claim. ☐

[48]Observe that the remaining cases are when $M=0$, $M=\frac{1}{2}$, $M=1$.

[49]Furthermore, we can observe that we do not fall under a "forbidden case" of Lemma 4.1 in [A 2010], by weight considerations, and since the tensor fields in our lemma assumption are not "bad".

[50]Note that the weight becomes less negative, hence Lemma 4.10 in [A 2010] applies.

[51]By our assumptions there will be a removable index in these cases. Hence our extra requirements of those lemmas are fulfilled.

Proof of Lemma 2.9. We rewrite the hypothesis of Lemma 2.3 (which is also the hypothesis of Lemma 2.9) as

$$\begin{aligned}(2\text{-}49)\quad \sum_{h\in H_2} a_h X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}}\big\{&C^{h,i_1\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &-\mathrm{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\\ =\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u).\end{aligned}$$

Here the operation Switch interchanges the indices $_a$ and $_b$ in the two factors $\nabla_a\omega_1$, $\nabla_b\omega_2$.

Notational conventions: We have again denoted by $H_2^a\subset H_2$ the index set of those vector fields for which one of the free indices (say $_{i_{\alpha+1}}$) belongs to a factor $\nabla\omega_1$ or $\nabla\omega_2$. With no loss of generality we assume that for each $h\in H_2^a$, the index $_{i_{\alpha+1}}$ belongs to the factor $\nabla\omega_1$. We can clearly do this, due to the antisymmetry of the factors $\nabla\omega_1$, $\nabla\omega_2$.

We have defined $H_2^b=H_2\setminus H_2^a$. For each $h\in H_2^b$ we denote by T_{ω_1}, T_{ω_2} the factors against which $\nabla\omega_1$, $\nabla\omega_2$ contract. Also, for each $h\in H_2^a$ we will denote by T_{ω_2} the factor against which $\nabla\omega_2$ contracts.[52]

For each $h\in H_2$, we will call the factors T_{ω_1}, T_{ω_2} against which $\nabla\omega_1$ or $\nabla\omega_2$ are contracting "problematic" in the following cases: If T_{ω_1} or T_{ω_2} is of the form $\nabla^{(m)}R_{ijkl}$ and $\nabla\omega_1$ or $\nabla\omega_2$ contracts against an internal index; or if T_{ω_1} or T_{ω_2} is of the form $S_*\nabla^{(\nu)}R_{ijkl}$ and the factor $\nabla\omega_1$ or $\nabla\omega_2$ contracts against one of the indices $_k$ or $_l$.

We then define a few subsets of H_2^a, H_2^b:

Definition. We define $H^b_{2,**}$ to be the index set of the tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g$ for which $\nabla\omega_1$, $\nabla\omega_2$ contract against different factors and both T_{ω_1} and T_{ω_2} are problematic.

We define $H^a_{2,*}\subset H^a_2$ to be the index set of the tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g$s for which T_{ω_2} is problematic.

We define $H^b_{2,*}$ to stand for the index set of the tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g$s for which either $T_{\omega_1}=T_{\omega_2}$ or $T_{\omega_1}\neq T_{\omega_2}$ and one of the factors T_{ω_1}, T_{ω_2} is problematic.

Abusing notation, we will use the symbols $\sum_{h\in H^b_{2,*}}$ and so on to denote *generic* linear combinations as above, when these symbols appear in the right-hand sides of the equations below.

[52] Note that the definition of T_{ω_1}, T_{ω_2} depends on h; however, to simplify notation we suppress the index h that should appear in T_{ω_1}, T_{ω_2}.

We then state three preparatory claims. First, we claim that we can write

$$\begin{aligned}(2\text{-}50)\quad &\sum_{h\in H^b_{2,**}} a_h X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}}\\ &\quad\times \big\{C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &\qquad - \text{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\\ &= \sum_{h\in H^b_{2,*}} a_h X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}}\\ &\quad\times \big\{C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &\qquad - \text{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\\ &\quad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u),\end{aligned}$$

where the linear combination $\sum_{h\in H^b_{2,*}}\dots$ on the right-hand side stands for a generic linear combination in the form described above. Observe that if we can show (2-50) then we may assume with no loss of generality that $H^b_{2,**}=\varnothing$ in our lemma hypothesis.

Then, assuming that $H^b_{2,**}=\varnothing$ in our lemma hypothesis we will show that there exists a linear combination of $(\alpha-\pi+1)$-tensor fields (indexed in X below) which are in the form (2-5) with a simple character $\vec{\kappa}_{\text{simp}}$ so that

$$\begin{aligned}(2\text{-}51)\quad &\sum_{h\in H^a_{2,*}} a_h\big\{C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &- \text{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon\\ &- X_*\operatorname{div}_{i_{\alpha+2}}\sum_{x\in X} a_x\big\{C^{x,i_1\dots i_{\alpha+1}i_{\alpha+2}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &- \text{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon\\ &+\sum_{h\in H^b_{2,*}} a_h\big\{C^{h,i_1\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &- \text{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_{\alpha+1}}\upsilon\\ &\qquad=\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u,\upsilon^{\alpha-\pi}).\end{aligned}$$

We observe that if we can show the above, we may then assume that $H^a_{2,*}=\varnothing$ (and $H^b_{2,**}=\varnothing$) in the hypothesis of Lemma 2.9.

Finally, under the assumption that $H^b_{2,**} = H^a_{2,*} = \varnothing$ in our lemma hypothesis, we will show that we can write

$$
\begin{aligned}
(2\text{-}52)\quad & \sum_{h\in H^b_{2,*}} a_h X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}} \\
&\qquad \times \big\{ C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u) \\
&\qquad\qquad - \operatorname{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\} \\
&= \sum_{h\in H^b_{2,OK}} a_h X_+ \operatorname{div}_{i_{\pi+1}} \dots X_+ \operatorname{div}_{i_{\alpha+1}} \\
&\qquad\qquad \times \big\{ C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u) \\
&\qquad\qquad\qquad - \operatorname{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\} \\
&\qquad + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,[\omega_1,\omega_2],\phi_1,\dots,\phi_u),
\end{aligned}
$$

where the sublinear combination $\sum_{h\in H^b_{2,OK}}\dots$ on the right-hand side stands for a generic linear combination of acceptable tensor fields in the form (2-5) with simple character $\vec{\kappa}_{\text{simp}}$, with no free indices in the factors $\nabla\omega_1$, $\nabla\omega_2$ and where the factors T_{ω_1}, T_{ω_2} are not problematic. Therefore, if we can show the above equations, we are reduced to showing Lemma 2.9 under the assumptions that $H^2_{a,*} = H^2_{b,**} = H^2_{b,*} = \varnothing$.

Sketch of the proof of (2-50)*,* (2-51)*,* (2-52)*.* Equation (2-50) follows by reiterating the proof of the first claim of Lemma 4.10 in [A 2010].[53] (2-51) follows by reiterating the proof of the first claim of Lemma 4.10 in [A 2010], but rather than applying Corollary 1 [A 2010] in that proof, we now apply Lemma 2.7 (which we have shown).[54] Finally, the claim of (2-52) *for the sublinear combination in $H^b_{2,*}$ where $T_{\omega_1} \neq T_{\omega_2}$* follows by applying Lemma 2.5.[55] We can then show that the remaining sublinear combination in $\sum_{h\in H^b_{2,*}}\dots$ must vanish separately (modulo a linear combination $\sum_{j\in J}\dots$) by picking out the sublinear combination in the hypothesis of Lemma 2.10 where both factors $\nabla\omega_1$, $\nabla\omega_2$ are contracting against the same factor. □

Now, under these additional assumptions that $H^2_{a,*} = H^2_{b,**} = H^2_{b,*} = \varnothing$, we will show our claim by distinguishing two cases: In Case A there is a factor $\nabla^{(m)}R_{ijkl}$

[53] By the additional restrictions imposed on the assumption of Lemma 2.3 there is no danger of falling under a "forbidden case" of Corollary 1 in [A 2010].

[54] Observe that the assumption that Lemma 2.3 does not include "forbidden cases" ensures that we will not need to apply Lemma 2.7 in a "forbidden case".

[55] In this case there will be a factor $\nabla\omega_1$ or $\nabla\omega_2$ contracting against a nonspecial index; therefore there is no danger of falling under a "forbidden" case of Lemma 2.7.

in $\vec{\kappa}_{\mathrm{simp}}$; in Case B there is no such factor. An important note: We may now *use* Lemma 2.7, which we have proven earlier in this section.

Proof of Lemma 2.9 in Case A. We define the (set of) C-crucial factors (which will necessarily be of the form $\nabla^{(m)}R_{ijkl}$) as in the setting of Lemma 2.7. First we prove a mini-claim which only applies to the case where the C-crucial factor is unique.

Mini-claim, when the C-crucial factor is unique. We then consider the tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g$, $h\in H^a_2$ for which $\nabla\omega_2$ contracts against the C-crucial factor. Notice that by our hypothesis that $H^2_{a,*}=\varnothing$, it follows that $\nabla\omega_2$ contracts against a derivative index in the C-crucial factor. Denote by $H^{a,+}_2\subset H^a_2$ the index set of these tensor fields.

We observe that for each $h\in H^{a,+}_2$ we can now construct a tensor field by erasing the index in the factor $\nabla^{(m)}R_{ijkl}$ that contracts against the factor $\nabla\omega_2$ and making the index in $\nabla\omega_2$ into a free index ${}_{i_\beta}$. We denote this tensor field by $C^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)$. By the analogous operation we obtain a tensor field $\mathrm{Switch}[C^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)]$.

It follows that in the case where the C-crucial factor is unique, for each $h\in H^{a,+}_2$,

$$\begin{aligned}
(2\text{-}53)\quad & X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}}\bigl\{C^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\
&\qquad -\mathrm{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\bigr\}\\
&= X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}}X_*\operatorname{div}_{i_\beta}\\
&\qquad\times\bigl\{C^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\
&\qquad\quad -\mathrm{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\bigr\}\\
&\quad+\sum_{r\in R}a_rX_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}}\\
&\qquad\times\bigl\{C^{r,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\
&\qquad\quad -\mathrm{Switch}[C]^{r,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\bigr\}\\
&\quad+\sum_{j\in J}a_jC^j_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u),
\end{aligned}$$

where each tensor field $C^{r,i_{\pi+1}\dots i_{\alpha+1}}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)$ has the factor $\nabla\omega_2$ contracting against some factor other than the C-crucial factor.

But we observe that

$$\begin{aligned}
(2\text{-}54)\quad & X_*\operatorname{div}_{i_{\pi+1}}\dots X_*\operatorname{div}_{i_{\alpha+1}}X_*\operatorname{div}_{i_\beta}\\
&\quad\times\bigl\{C^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\
&\quad -\mathrm{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}i_\beta}_g(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\bigr\}=0.
\end{aligned}$$

Therefore, in the case $Set \neq \varnothing$ or $Set = \varnothing$ and $\sigma_1 = 1$, we have now reduced Lemma 2.9 to the case where $H_2^{a,+} = \varnothing$.

Now (under the assumption that $H_2^{a,+} = \varnothing$ when the C-crucial factor is unique) we consider the sublinear combination *Special* in the hypothesis of Lemma 2.9 that consists of complete contractions with $\nabla\omega_1$ contracting against the C-crucial factor while the factor $\nabla\omega_2$ is contracting against some other factor. (If $Set = \varnothing$ and $\sigma_1 > 1$ *Special* stands for the sublinear combination where $\nabla\omega_1$ is contracting against a generic C-crucial factor and $\nabla\omega_2$ is contracting against some other factor.) In particular, for each $h \in H_2^a$, since $H_2^{a,+} = \varnothing$ we see that the sublinear combination in

$$\begin{aligned}\tag{2-55} \sum_{h\in H_2^a} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_{\alpha+1}} &\\ \times \big\{C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u) &\\ - \operatorname{Switch}[C]_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\big\}&\end{aligned}$$

that belongs to *Special* is precisely

$$\sum_{h\in H_2^a} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_\alpha} \times \operatorname{Hit}\operatorname{div}_{i_{\alpha+1}} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u);$$

(in the case $Set = \varnothing$ and $\sigma_1 > 1$, $\operatorname{Hit}\operatorname{div}_{i_{\alpha+1}}$ just means that ∇_{i_γ} can hit any factor $\nabla^{(m)}R_{ijkl}$ that is not contracting against $\nabla\omega_2$; recall that in the other cases it means that it must hit the unique C-crucial factor).

We also consider the tensor fields $C^{h,i_{\pi+1}\dots i_{\alpha+1}}$, $\operatorname{Switch}[C]^{h,i_{\pi+1}\dots i_{\alpha+1}}$, $h \in H_2^b$, for which $\nabla\omega_1$ contracts against the C-crucial factor and $\nabla\omega_2$ does not (or, if there are multiple C-crucial factors, where $\nabla\omega_1$, $\nabla\omega_2$ contract against different C-crucial factors). For this proof, we index all those tensor fields in $H_2^{b,\Psi}$ and we will denote them by $C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}$.

Thus we derive

$$\begin{aligned}\tag{2-56} \sum_{h\in H_2^a} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_\alpha} \operatorname{Hit}\operatorname{div}_{i_{\alpha+1}} &\\ \times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u) &\\ + \sum_{h\in H_2^{b,\Psi}} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_{\alpha+1}} &\\ \times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u) &\\ = \sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u).&\end{aligned}$$

We group up the vector fields on the left-hand side according to their *weak* $(u+2)$*-characters*[56] (defined by $\nabla\phi_1,\dots,\nabla\phi_u,\nabla\omega_1,\nabla\omega_2$). (Recall that we started off

[56] See [A 2010] for a definition of this notion.

with complete contractions with the same u-simple characters — so the only new information that we are taking into account is what *type* of factor $\nabla\omega_2$ contracts against.) We consider the set of weak simple characters that we have obtained. We denote this set by $\{\vec{\kappa}_1, \dots \vec{\kappa}_B\}$, and we respectively have the index sets $H_2^{a,\vec{\kappa}_f}$ and $H_2^{b,\vec{\kappa}_f}$.

We will show our Lemma 2.9 by replacing the index set H_2^a by any $H_2^{a,\vec{\kappa}_f}$, $f \le B$.

It follows that for each $f \le B$,

$$\begin{aligned}(2\text{-}57)\quad &\sum_{h\in H_2^{a,\vec{\kappa}_f}} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_\alpha} \operatorname{Hit}\operatorname{div}_{i_{\alpha+1}} \\ &\qquad\qquad \times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1, \dots, \Omega_p, \omega_1, \omega_2, \phi_1, \dots, \phi_u) \\ &\quad + \sum_{h\in H_2^{b,\vec{\kappa}_f}} a_h X_* \operatorname{div}_{i_{\pi+1}} \dots X_* \operatorname{div}_{i_{\alpha+1}} \\ &\qquad\qquad \times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1, \dots, \Omega_p, \omega_1, \omega_2, \phi_1, \dots, \phi_u) \\ &\qquad\qquad\qquad = \sum_{j\in J} a_j C_g^j(\Omega_1, \dots, \Omega_p, \omega_1, \omega_2, \phi_1, \dots, \phi_u),\end{aligned}$$

where the complete contractions C_g^j have a u-simple character that is subsequent to $\vec{\kappa}_{\text{simp}}$. We will show our claim for each of the index sets $H_2^{b,\vec{\kappa}_f}$ separately.

Now, we treat the factors $\nabla\omega_1, \nabla\omega_2$ in the above as factors $\nabla\phi_{u+1}, \nabla\phi_{u+2}$. We see that since $H^b_{2,**} = H^2_{b,*} = H^2_{a,*} = \varnothing$, all the tensor fields in the above have the same $(u+2)$-simple character.

Our claim (Lemma 2.9) for the index set $H_2^{a,\vec{\kappa}_f}$ then follows: First, apply the operator $\operatorname{Erase}_{\nabla\omega_1}[\dots]$ to (2-57).[57] We are then left with tensor fields (denote them by

$$C_g^{h,i_{\pi+1}\dots i_\alpha}(\Omega_1, \dots, \Omega_p, \omega_2, \phi_1, \dots, \phi_u), \quad h \in H_2^{a,\vec{\kappa}_f},$$
$$C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1, \dots, \Omega_p, \omega_2, \phi_1, \dots, \phi_u), \quad h \in H_2^{b,\vec{\kappa}_f},$$

respectively) with the same $(u+1)$-simple character; say $\vec{\kappa}_{\text{simp},f}$. We can then apply Corollary 1 from [A 2010] (since we have weight $-n+2k$, $k>0$ by virtue of the eraser — notice that by weight considerations, since we started out with no "bad" tensor fields, there is no danger of falling under a "forbidden case"), to derive that there is a linear combination of acceptable α-tensor fields indexed in V below, with $(u+1)$-simple character $\vec{\kappa}_{\text{simp},f}$, so that

$$(2\text{-}58)\quad \sum_{h\in H_2^{a,\vec{\kappa}_f}} a_h C_g^{h,i_{\pi+1}\dots i_\alpha}(\Omega_1, \dots, \Omega_p, \omega_2, \phi_1, \dots, \phi_u)\nabla_{i_{\pi+1}}\upsilon \dots \nabla_{i_\alpha}\upsilon$$

[57] See the relevant lemma in the Appendix of [A 2012].

$$-\sum_{v\in V} a_v X_* \operatorname{div}_{i_{\alpha+1}} C_g^{v,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_2,\phi_1,\dots,\phi_u) \\ \times \nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon \\ =\sum_{j\in J} a_j C_g^{j,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\omega_2,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon,$$

where each complete contraction indexed in J is $(u+1)$-subsequent to $\vec{\kappa}_{\text{simp},f}$. In this setting $X_*\operatorname{div}_i$ just means that in addition to the restrictions imposed on $X\operatorname{div}_i$ we are not allowed to hit the factor $\nabla\omega_2$.

Then, if we multiply the above equation by an expression $\nabla_i\omega_1\nabla^i\upsilon$ and then antisymmetrize the indices $_a$, $_b$ in the factors $\nabla_a\omega_1$, $\nabla_b\omega_2$ and finally make all $\nabla\upsilon$s into X_+ divs, we derive our claim. □

Proof of Lemma 2.9 in Case B (when $\sigma_1=0$). Our proof follows the same pattern as the proof of Lemma 2.7 in Case B.

We again define the "measure" of each factor in each tensor field $C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}$ as in the proof of Case B in Lemma 2.7. Again, let M stand for the maximum measure among all factors in all tensor fields $C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}$, $h\in H_2^a$. We denote by $H_2^{a,M}\subset H_2^a$ the index set of the tensor fields for which some factor has measure M.

We will further divide $H_a^{2,M}$ into subsets $H_a^{2,M,k}$, $k=1,\dots,\sigma$, according to the factor which has measure M: First, we order the factors $S_*\nabla^{(\nu)}R_{ijkl},\dots\nabla^{(p)}\Omega_h$ in $\vec{\kappa}_{\text{simp}}$, and label them $T_1,\dots,T_\sigma$ (observe each factor is well-defined in $\vec{\kappa}_{\text{simp}}$, because we are in Case B). We then say that $h\in H_2^{a,M,1}$ if in $C_g^{h,i_{\pi+1}\dots i_\alpha}$, T_1 has measure M. We say that $h\in H_2^{a,M,2}$ if in $C_g^{h,i_{\pi+1}\dots i_\alpha}$, T_2 has measure M and T_1 has measure less than M, and so on. We will then prove our claim for each of the index sets $h\in H_2^{a,M,k}$.[58] We arbitrarily pick a $k\le\sigma$ and show our claim for $\sum_{h\in H_a^{2,M,k}}\dots$.

For the purposes of this proof, we call the factor T_k the D-crucial factor (in this setting the D-crucial factor is unique).

Now, we pick out the subset $H_2^{b,k,+}\subset H_2^b$ that is defined by the rule $h\in H_2^{b,k}$ if and only if $\nabla\omega_1$ contracts against the D-crucial factor T_k. We also pick out the subset $H_2^{b,k,-}\subset H_2^b$ that is defined by the rule $h\in H_2^{b,k}$ if and only if $\nabla\omega_2$ contracts against the D-crucial factor T_k. Finally, we define $H_2^{a,\tilde{}}\subset H_2^a$, $H_2^{a,-}\subset H_2^a$ to stand for the index set of tensor fields for which $\nabla\omega_2$ contracts against the D-crucial factor.

Now, for each $h\in H_2^a$ we denote by

$$\operatorname{Hit}\operatorname{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)$$

[58] Again we observe that if we can prove this then Lemma 2.9 in Case B will follow by induction.

the sublinear combination in $X\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)$ that arises when ∇_{i_γ} hits the D-crucial factor. It then follows that

$$\begin{aligned}(2\text{-}59)\quad &\sum_{h\in H_2^a} a_h X_*\,\mathrm{div}_{i_{\pi+1}}\dots X_*\,\mathrm{div}_{i_\alpha}\\ &\qquad\times \mathrm{Hit}\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &-\sum_{h\in H_2^{a,\tilde{}}} a_h X_*\,\mathrm{div}_{i_{\pi+1}}\dots X_*\,\mathrm{div}_{i_{\alpha+1}}\\ &\qquad\times \mathrm{Switch}[C]_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &+\sum_{h\in H_2^{b,k,+}} a_h X\,\mathrm{div}_{i_{\pi+1}}\dots X\,\mathrm{div}_{i_\gamma}\\ &\qquad\times C_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &-\sum_{h\in H_2^{b,k,-}} a_h X\,\mathrm{div}_{i_{\pi+1}}\dots X\,\mathrm{div}_{i_\gamma}\\ &\qquad\times \mathrm{Switch}[C]_g^{h,i_{\pi+1}\dots i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &=\sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u),\end{aligned}$$

where each C_g^j has the factor $\nabla\omega_1$ contracting against the D-crucial factor and is simply subsequent to $\vec{\kappa}_{\mathrm{simp}}$.

We now denote the $(u+1)$-simple character (the one defined by $\nabla\phi_1,\dots,\nabla\omega_1$) of the tensor fields $\mathrm{Hit}\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_\alpha,i_\gamma}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)$ by $\vec{\kappa}'_{\mathrm{simp}}$. (Observe that they all have the same $(u+1)$-simple character.)

We observe that just applying Lemma 2.1 to (2-59) (all tensor fields are acceptable and have the same simple character $\vec{\kappa}'_{\mathrm{simp}}$ — we treat $\nabla\omega_1$ as a factor $\nabla\phi_{u+1}$ and the factor $\nabla\omega_2$ as a factor ∇Y) and we then pick out the sublinear combination where there are M factors $\nabla\upsilon$ contracting against T_k, we obtain

$$\begin{aligned}(2\text{-}60)\quad &\sum_{h\in H_2^{a,*,k}} a_h\,\mathrm{Hit}\,\mathrm{div}_{i_\gamma} C_g^{h,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon\\ &+\sum_{x\in X} a_x X\,\mathrm{div}_{i_{\alpha+1}} C_g^{x,i_{\pi+1}\dots i_\alpha,i_{\alpha+1}}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\\ &\qquad\qquad\times\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon\\ &+\sum_{j\in J} a_j C_g^{j,i_{\pi+1}\dots i_\alpha}(\Omega_1,\dots,\Omega_p,\omega_1,\omega_2,\phi_1,\dots,\phi_u)\nabla_{i_{\pi+1}}\upsilon\dots\nabla_{i_\alpha}\upsilon=0,\end{aligned}$$

where the tensor fields indexed in X are acceptable and have a $(u+1)$-simple character $\vec{\kappa}'_{\mathrm{simp}}$ and each C^j is simply subsequent to $\vec{\kappa}'_{\mathrm{simp}}$.

Now, observe that *if* $M\geq\frac{3}{2}$ then we can apply the Eraser (from the Appendix in [A 2012]) to $\nabla\omega_1$ and the index it contracts against in the D-crucial factor and derive our conclusion as in Case A.

The remaining cases are when $M=1$, $M=\frac{1}{2}$ and $M=0$. The first one is easier, so we proceed to show our claim in that case. The two subcases $M=\frac{1}{2}$, $M=0$ will be discussed in the next subsection.

In the case $M=1$, i.e., the D-crucial factor is of the form $\nabla^{(p)}\Omega_h$, then we cannot derive our claim, because of the possibility that some tensor fields indexed in X above have $\nabla\omega_1$ contracting according to the pattern $\nabla_i\omega_1\nabla^{ij}\Omega_h\nabla_j\psi$, where $\psi=\upsilon$ or $\psi=\phi_h$. Therefore, in this setting, we first apply the eraser twice to remove the expression $\nabla^{(2)}_{ij}\Omega_h\nabla^i\psi\nabla^j\omega_1$ and then apply Corollary 2 from [A 2010][59] to (2-60) (observe that (2-60) now falls under the inductive assumption of Lemma 4.6 in [A 2010] since we have lowered the weight[60] to obtain a new equation in the form (2-60), where each tensor field in X has the factor $\nabla\omega_1$ contracting against a factor $\nabla^{(l)}\Omega_h$, $l\geq 3$. Then, applying the eraser as explained, we derive our Lemma 2.9 in this case.

The cases $M=\frac{1}{2}$, $M=0$. Notice that in these cases we must have $\alpha=\pi$, by virtue of the definition of maximal "measure" above. We will then prove our claim by proving a more general claim by induction, in the next subsection. □

2D. ***The remaining cases of Lemma 2.9.*** We prove our claim in these cases via an induction. In order to give a detailed proof, we will restate our lemma hypothesis in this case (with a slight change of notation).

The hypothesis of the remaining cases of Lemma 2.9. Recall that we assume that

$$\begin{aligned}\sum_{x\in X_a} a_x X_*\operatorname{div}_{i_1} C^{x,i_1}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,[\omega_1,\omega_2])\\ +\sum_{x\in X_b} a_x X_*\operatorname{div}_{i_1} C^{x,i_1}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,[\omega_1,\omega_2])\\ +\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)=0\end{aligned}\tag{2-61}$$

holds modulo complete contractions of length $\geq\sigma+u+3$ ($\sigma\geq 3$ — here σ stands for $u+p$ — see the next equation). We denote the weight of the complete contractions in the above by $-K$. The tensor fields in the above equation are each in the form

$$\begin{aligned}pcontr\big(S_*\nabla^{(\nu_1)}R_{x_1jkl}\otimes\cdots\otimes S_*\nabla^{(\nu_u)}R_{x_zj'k'l'}\\ \otimes\nabla^{(a_1)}\Omega_1\otimes\cdots\otimes\nabla^{(a_p)}\Omega_p\otimes[\nabla\omega_1\otimes\nabla\omega_2]\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\cdots\otimes\nabla^{x_u}\tilde{\phi}_z\big).\end{aligned}\tag{2-62}$$

[59] Recall that we showed in [A 2010] that this is a corollary of Lemma 4.6 in [A 2010], which we have now shown.

[60] There is no danger of falling under a "forbidden case" of Lemma 2.1 by weight considerations since we are assuming that none of the tensor fields of minimum rank in the assumption of Lemma 2.3 are "bad".

We recall that the u-simple character of the above has been denoted by $\vec{\kappa}_{\mathrm{simp}}$. Recall that we are now assuming that all the factors $\nabla^{(a_i)}\Omega_x$ in $\vec{\kappa}_{\mathrm{simp}}$ are acceptable.[61] The complete contractions indexed in J in (2-61) are simply subsequent to $\vec{\kappa}_{\mathrm{simp}}$. We also recall that $X_*\,\mathrm{div}_i$ stands for the sublinear combination in $X\,\mathrm{div}_i$ where ∇_i is not allowed to hit either of the factors $\nabla\omega_1$, $\nabla\omega_2$.

We recall that the tensor fields indexed in X_a have the free index $_{i_1}$ belonging to the factor $\nabla\omega_1$. The tensor fields indexed in X_b have the free index $_{i_1}$ *not* belonging to any of the factors $\nabla\omega_1$, $\nabla\omega_2$.

We recall the key assumption that for each of the tensor fields indexed in X_a, there is at least one removable index in each tensor field

$$C^{x,i_1}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,[\omega_1,\omega_2]),$$

$x\in X_a$.[62]

In order to complete our proof of Lemma 2.9, we will show that we can write

$$\begin{aligned}\text{(2-63)}\quad &\sum_{x\in X_a} a_x C^{x,i_1}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,[\omega_1,\omega_2])\nabla_{i_1}\upsilon\\ &=\sum_{x\in X'} a_x X_*\,\mathrm{div}_{i_2}\dots X_*\,\mathrm{div}_{i_a}\\ &\qquad\times C^{x,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,[\omega_1,\omega_2])\nabla_{i_1}\upsilon\\ &\quad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u),\end{aligned}$$

where the tensor fields indexed in X' are acceptable in the form (2-62), each with rank $a\geq 2$. Note that this will imply the remaining cases of Lemma 2.9, completing the proof of Lemma 2.3.

We recall that we are proving this claim when the assumption (2-61) formally falls under our inductive assumption of Proposition 1.1 (if we formally treat $\nabla\omega_1$, $\nabla\omega_2$ as factors $\nabla\phi_{z+1}$, $\nabla\phi_{z+2}$).

We will prove (2-63) by inductively proving a more general statement.

Assumptions. We consider vector fields (that is, partial contractions with one free index)

$$C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),$$
$$C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau),$$

[61] Meaning that each $a_i\geq 2$.

[62] Recall the definition of a "removable" index from page 8.

in the following forms, respectively,

$$\begin{aligned}pcontr\big(&S_*\nabla^{(\nu_1)}R_{x_1jkl}\otimes\cdots\otimes S_*\nabla^{(\nu_v)}R_{x_vj'k'l'}\\&\otimes\nabla^{(a_1)}\Omega_1\otimes\dots\nabla^{(a_b)}\Omega_b\otimes\nabla Y\\&\otimes\nabla\psi_1\otimes\cdots\otimes\nabla\psi_\tau\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\cdots\otimes\nabla^{x_v}\tilde{\phi}_v\big),\end{aligned}\tag{2-64}$$

$$\begin{aligned}pcontr\big(&S_*\nabla^{(\nu_1)}R_{x_1jkl}\otimes\cdots\otimes S_*\nabla^{(\nu_v)}R_{x_vj'k'l'}\\&\otimes\nabla^{(a_1)}\Omega_1\otimes\dots\nabla^{(a_b)}\Omega_b\otimes[\nabla\chi_1\otimes\nabla\chi_2]\\&\otimes\nabla\psi_1\otimes\cdots\otimes\nabla\psi_\tau\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\cdots\otimes\nabla^{x_v}\tilde{\phi}_v\big),\end{aligned}\tag{2-65}$$

for which the weight is $-W+1$, $W\le K$. We also assume $v+b\ge 2$. *Note:* the bracket $[\dots]$ stands for the antisymmetrization of the indices ${}_a$, ${}_b$ in the expression $\nabla_a\omega_1\nabla_b\omega_2$.

We assume (respectively) that

$$\begin{aligned}&\sum_{\zeta\in Z_a}a_\zeta X_*\operatorname{div}_{i_1}C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\&+\sum_{\zeta\in\bar{Z}_a}a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_\gamma}\\&\qquad\times C^{\zeta,i_1\dots i_\gamma}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\&+\sum_{\zeta\in Z_b}a_\zeta X_*\operatorname{div}_{i_1}C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\&+\sum_{j\in J}a_jC^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)=0,\end{aligned}\tag{2-66}$$

and

$$\begin{aligned}&\sum_{\zeta\in Z_a}a_\zeta X_*\operatorname{div}_{i_1}C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\&+\sum_{\zeta\in\bar{Z}_a}a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_\gamma}\\&\qquad\times C^{\zeta,i_1\dots i_\gamma}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\&+\sum_{\zeta\in Z_b}a_\zeta X_*\operatorname{div}_{i_1}C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\&+\sum_{j\in J}a_jC^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)=0,\end{aligned}\tag{2-67}$$

hold modulo complete contractions of length $\ge v+b+\tau+3$.

The tensor fields indexed in Z_a are assumed to have the free index in one of the factors $\nabla Y,\nabla\psi_1,\dots,\nabla\psi_\tau$, or one of the factors $\nabla\chi_1,\nabla\chi_2,\nabla\psi_1,\dots,\nabla\psi_\tau$, respectively. The tensor fields indexed in $\bar{Z}_a$ have rank $\gamma\ge 2$ and all their free indices belong to the factors $\nabla Y, \nabla\psi_1, \dots, \nabla\psi_\tau$, or the factors $\nabla\chi_1, \nabla\chi_2, \nabla\psi_1, \dots, \nabla\psi_\tau$, respectively. The tensor fields indexed in Z_b have the property that ${}_{i_1}$ does not belong to any of the factors

$$\nabla Y, \nabla\psi_1, \dots, \nabla\psi_\tau \quad \text{or} \quad \nabla\chi_1, \nabla\chi_2, \nabla\psi_1, \dots, \nabla\psi_\tau,$$

respectively. We also assume that for the tensor fields indexed in $Z_a \cup Z_b \cup \bar{Z}_a$, none of the factors $\nabla\psi_1, \dots, \nabla\psi_\tau$ are contracting against a special index in any factor $S_*\nabla^{(\nu)}R_{ijkl}$ and none of them are contracting against the rightmost index in any $\nabla^{(a_h)}\Omega_h$ (we will refer to this property as the $\mathfrak{p}$-property). We assume that $v+b \geq 2$, and furthermore if $v+b=2$ then for each $\zeta \in Z_a \cup Z_b$, the factors ∇Y (or $\nabla\chi_1, \nabla\chi_2$) are also not contracting against a special index in any $S_*\nabla^{(\nu)}R_{ijkl}$ and are not contracting against the rightmost index in any $\nabla^{(a_h)}\Omega_h$. Finally (and importantly) we assume that for the tensor fields indexed in Z_a, there is at least one removable index in each C^{ζ,i_1}. (In this setting, for a tensor field indexed in Z_a, a "removable" index is either a nonspecial index in a factor $S_*\nabla^{(\nu)}R_{ijkl}$, with $\nu > 0$ or an index in a factor $\nabla^{(B)}\Omega_h$, $B \geq 3$.)

Convention. In this subsection only, for tensor fields in the forms (2-66), (2-67) we say then an index is *special* if it is one of the indices ${}_k, {}_l$ in a factor $S_*\nabla^{(\nu)}R_{ijkl}$ (this is the usual convention), *or* if it is an index in a factor $\nabla^{(B)}_{r_1\dots r_B}\Omega_h$ *for which all the other indices are contracting against factors* $\nabla\psi_1, \dots, \nabla\psi_\tau$.

All tensor fields in (2-66), (2-67) have a given v-simple character $\bar{\kappa}_{\text{simp}}$. We assume the complete contractions indexed in J have a weak v-character $Weak(\bar{\kappa}_{\text{simp}})$ and are simply subsequent to $\bar{\kappa}_{\text{simp}}$. Here $X_*\operatorname{div}_i$ stands for the sublinear combination in $X\operatorname{div}_i$ where ∇_i is not allowed to hit any of the factors

$$\nabla Y, \nabla\psi_1, \dots, \nabla\psi_\tau \quad \text{or} \quad \nabla\chi_1, \nabla\chi_2, \nabla\psi_1, \dots, \nabla\psi_\tau,$$

respectively.

The claims of the general statement. We claim that under the assumption (2-67), there exists a linear combination of acceptable 2-tensor fields in the form (2-64), (2-65) respectively (indexed in W below), for which the $\mathfrak{p}$-property is satisfied, so that (respectively)

$$\begin{aligned}(2\text{-}68)\quad &\sum_{\zeta\in Z_a} a_\zeta C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon\\ &-\sum_{w\in W} a_w X_*\operatorname{div}_{i_2} C^{w,i_1i_2}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon\\ &+\sum_{j\in J} a_j C^{j,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon=0,\end{aligned}$$

and

$$\sum_{\zeta\in Z_a} a_\zeta C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon + \sum_{w\in W} a_w X_*\operatorname{div}_{i_2} \times C^{w,i_1i_2}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon + \sum_{j\in J} a_j C^{j,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon = 0. \tag{2-69}$$

We observe that when $\tau = 0$ and $v + b \geq 3$, (2-69) coincides with (2-63).[63] Therefore, if we can prove this general statement, we will have shown Lemma 2.9 in full generality, thus also completing the proof of Lemma 2.3.

We also have a further claim, when we assume (2-66), (2-67) with $v + b = 2$. In that case, we also claim that we can write

$$\sum_{\zeta\in Z_a\cup Z_b\cup\bar{Z}_a} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) = \sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^{j}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau), \tag{2-70}$$

and

$$\sum_{\zeta\in Z_a\cup\bar{Z}_a\cup Z_b} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) = \sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^{j}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau), \tag{2-71}$$

where the tensor fields indexed in Q are in the same form as (2-64) or (2-65) respectively, but have a factor (expression) $\nabla^{(2)}Y$ or $\nabla^{(2)}_{a[i}\omega_1\nabla_{j]}\omega_2$, respectively, and satisfy all the other properties of the tensor fields in Z_a.

Consequence of (2-68), (2-69) *when* $v+b\geq 3$. We here codify an implication one can derive from (2-68), (2-69). This implication will be useful further down in this subsection. We see that by making the factors $\nabla\upsilon$ into $X_*\operatorname{div}$ in (2-66),[64] (2-67)

[63] Also, the assumption of existence of a non removable index coincides with the corresponding assumption of Lemma 2.3.

[64] See the Appendix in [A 2012].

and replacing into (2-68), (2-69), we obtain

$$\sum_{\zeta\in Z'_a} a_\zeta X_* \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) + \sum_{\zeta\in Z_b} a_\zeta X_* \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) = 0, \tag{2-72}$$

and

$$\sum_{\zeta\in Z'_a} a_\zeta X_* \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{\zeta\in Z_b} a_\zeta X_* \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) = 0, \tag{2-73}$$

where here the tensor fields indexed in Z'_a are like the tensor fields indexed in Z_a in (2-66), (2-67) but have the additional feature that no free index belongs to the factor $\nabla\psi_1$ (and all the other assumptions of equations (2-66), (2-67) continue to hold).

We then claim that we can derive new equations

$$\sum_{\zeta\in Z'_a} a_\zeta X_+ \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) + \sum_{\zeta\in Z_b} a_\zeta X_+ \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) = \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) \tag{2-74}$$

and

$$\sum_{\zeta\in Z'_a} a_\zeta X_+ \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{\zeta\in Z_b} a_\zeta X_+ \operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) = \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau), \tag{2-75}$$

where here $X_+\operatorname{div}_i$ stands for the sublinear combination in $X\operatorname{div}_i$ where ∇_i *is* allowed to hit the factor ∇Y or $\nabla\chi_1$ (respectively), but not the factors $\nabla\psi_1,\dots,\nabla\phi_\tau$, $(\nabla\chi_2)$. Furthermore, the linear combinations indexed in Q stand for generic linear combinations of vector fields in the form (2-64) or (2-65), only with the expressions ∇Y or $\nabla_{[a}\omega_1\nabla_{b]}\omega_2$ replaced by expressions $\nabla^{(2)}Y$, $\nabla^{(2)}_{c[a}\omega_1\nabla_{b]}\omega_2$.

Proof that (2-74), (2-75) *follow from* (2-68), (2-69). We prove the above by an induction. We will first subdivide Z'_a, Z_b into subsets as follows: $\zeta\in Z'_{a,\mathfrak{p}}$ or $\zeta\in Z_{b,\mathfrak{p}}$ if the factor ∇Y (or one of the factors $\nabla\chi_1,\nabla\chi_2$) contracts against a special index in the same factor against which $\nabla\psi_1$ contracts.

Now, if $Z'_{a,\mathfrak{p}}\cup Z_{b,\mathfrak{p}}\neq\varnothing$ our inductive statement will be that we can write

$$\begin{aligned}
(2\text{-}76)\quad &\sum_{\zeta\in Z'_{a,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_\gamma}\\
&\qquad\times C^{\zeta,i_1\dots i_\gamma}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^k} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k}\\
&\qquad\times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{\zeta\in Z'_{a,No\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_\gamma}\\
&\qquad\times C^{\zeta,i_1\dots i_\gamma}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),
\end{aligned}$$

and

$$\begin{aligned}
(2\text{-}77)\quad &\sum_{\zeta\in Z'_{a,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^k} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k}\\
&\qquad\qquad\times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{\zeta\in Z'_{a,No\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_\gamma}\\
&\qquad\qquad\times C^{\zeta,i_1\dots i_\gamma}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)
\end{aligned}$$

$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau),$$

where the tensor fields indexed in T^k have all the properties of the tensor fields indexed in $Z'_{a,\mathfrak{p}}$ (in particular the index in $\nabla\psi_1$ is not free) and in addition have rank k. The tensor fields indexed in $Z'_{a,No\mathfrak{p}}$ in the right-hand side have all the regular features of the terms indexed in Z'_a (in particular rank $\gamma\geq 1$ and the factor $\nabla\psi_1$ does not contain a free index) and in addition none of the factors ∇Y (or $\nabla\chi_1,\nabla\chi_2$) contract against a special index.

Our inductive claim is that we can write

$$\sum_{\zeta\in Z'_a} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$=\sum_{\zeta\in Z_b} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{t\in T^{k+1}} a_t X_+ \operatorname{div}_{i_1}\dots X_+ \operatorname{div}_{i_{k+1}}$$
$$\times C_g^{t,i_1\dots i_{k+1}}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{\zeta\in Z'_{a,No\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1}\dots X_+ \operatorname{div}_{i_\gamma}$$
$$\times C_g^{\zeta,i_1\dots i_\gamma}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau), \tag{2-78}$$

and

$$\sum_{\zeta\in Z'_a} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$=\sum_{\zeta\in Z_b} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{t\in T^{k+1}} a_t X_+ \operatorname{div}_{i_1}\dots X_+ \operatorname{div}_{i_k}$$
$$\times C_g^{t,i_1\dots i_{k+1}}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{\zeta\in Z'_{a,No\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1}\dots X_+ \operatorname{div}_{i_\gamma}$$
$$\times C_g^{\zeta,i_1\dots i_\gamma}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \tag{2-79}$$

$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)=0.$$

We will derive (2-78), (2-79) momentarily. For now, we observe that by iterative repetition of the above inductive step we are reduced to showing (2-74), (2-75) under the additional assumption that $Z'_{a,\mathfrak{p}}=\varnothing$.

Under that assumption, we denote by $Z_{b,\mathfrak{p}}\subset Z_b$ the index set of vector fields for which the factor ∇Y (or one of the factors $\nabla\chi_1,\nabla\chi_2$) contracts against a special index. We will then prove another inductive statement: We assume that we can write

$$\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$=\sum_{t\in V^k} a_t X_+ \operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k} \times C_g^{t,i_1\dots i_k}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{\zeta\in Z_{b,No\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau), \tag{2-80}$$

and

$$\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$=\sum_{t\in V^k} a_t X_+ \operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k} \times C_g^{t,i_1\dots i_k}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{\zeta\in Z_{b,No\mathfrak{p}}} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau), \tag{2-81}$$

where the tensor fields indexed in V^k have all the features of the tensor fields indexed in $Z_{b,\mathfrak{p}}$ but in addition have all the k free indices *not* belonging to factors

$\nabla\psi_1,\dots,\nabla\psi_\tau$. The tensor fields indexed in $Z_{b,No\mathfrak{p}}$ have all the regular features of the tensor fields in Z_b and in addition have the factor ∇Y (or the factors $\nabla\chi_1,\nabla\chi_2$) *not* contracting against special indices. The terms indexed in Q are as required in the right-hand side of (2-74), (2-75) (which are the equations that we are proving).

We will then show that we can write

$$\begin{aligned}\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\ =\sum_{t\in V^{k+1}} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_{k+1}}\\ \times C^{t,i_1\dots i_{k+1}}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\ +\sum_{\zeta\in Z_{b,No\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\ +\sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\ +\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),\end{aligned}\tag{2-82}$$

and

$$\begin{aligned}\sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\ =\sum_{t\in V^{k+1}} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_{k+1}}\\ \times C^{t,i_1\dots i_{k+1}}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\ +\sum_{\zeta\in Z_{b,No\mathfrak{p}}} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\ +\sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\ +\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau).\end{aligned}\tag{2-83}$$

(Here the tensor fields indexed in V^{k+1} have all the features described above and moreover have rank $k+1$.)

Thus, by iterative repetition of this step we are reduced to showing our claim under the additional assumption that $Z'_{a,\mathfrak{p}}=Z_{b,\mathfrak{p}}=\varnothing$.

We prove (2-82), (2-83) below. Now, we present the rest of our claims under the assumption that $Z'_{a,\mathfrak{p}}=Z_{b,\mathfrak{p}}=\varnothing$. For the rest of this proof we may assume that all tensor fields have the factor ∇Y (or the factors $\nabla\chi_1,\nabla\chi_2$) not contracting against special indices.

We then perform a new induction: We assume that we can write

$$\begin{aligned}
(2\text{-}84)\quad &\sum_{\zeta\in Z'_a} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_b} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^k} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k}\\
&\qquad\times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),
\end{aligned}$$

and

$$\begin{aligned}
(2\text{-}85)\quad &\sum_{\zeta\in Z'_a} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_b} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^k} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_k}\\
&\qquad\times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau),
\end{aligned}$$

where the tensor fields indexed in T^k have all the properties of the tensor fields indexed in Z'_a (in particular the index in $\nabla\psi_1$ is not free) and in addition have rank k. We then show that we can write

$$\begin{aligned}
(2\text{-}86)\quad &\sum_{\zeta\in Z'_a} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_b} a_\zeta X_+\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^{k+1}} a_t X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_{k+1}}\\
&\qquad\times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)
\end{aligned}$$

$$+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)$$
$$+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),$$

and

$$\begin{aligned}
&\sum_{\zeta\in Z'_a} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&=\sum_{\zeta\in Z_b} a_\zeta X_+ \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{t\in T^{k+1}} a_t X_+ \operatorname{div}_{i_1}\dots X_+ \operatorname{div}_{i_{k+1}}\\
&\qquad\times C_g^{t,i_1\dots i_{k+1}}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\\
&\quad+\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau).
\end{aligned}\tag{2-87}$$

We will derive (2-86), (2-87) momentarily. For now, we observe that by iterative repetition of the above we are reduced to showing (2-74), (2-75) under the further assumption that $Z'_a=\varnothing$. In that setting, we can just repeatedly apply the eraser (see the Appendix in [A 2012] for a definition of this notion) to as many factors $\nabla\psi_\tau$ as needed in order to reduce ourselves to a new true equation where each of the real factors contracts against at most one of the factors $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla Y$ (or $\nabla\chi_1,\nabla\chi_2$).[65] Then, by invoking Corollary 1 from [A 2010][66] and then reintroducing the factors we erased, we derive our claim.

Proof of (2-86) *and* (2-87). Picking out the sublinear combination in (2-84), (2-85) with one derivative on ∇Y or $\nabla\chi_1$ and substituting into (2-72), (2-73) we derive

$$\begin{aligned}
&\sum_{t\in T^k} a_t X_* \operatorname{div}_{i_1}\dots X_* \operatorname{div}_{i_k} C_g^{t,i_1\dots i_k}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\qquad+\sum_{\zeta\in Z_b} a_\zeta X_* \operatorname{div}_{i_1} C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\\
&\qquad\qquad=\sum_{j\in J} a_j C_g^{j}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau),
\end{aligned}\tag{2-88}$$

---

[65] All remaining factors $\nabla\psi_1,\dots,\nabla\psi_\tau$ and also the factor(s) ∇Y (or $\nabla\chi_1,\nabla\chi_2$) are treated as factors $\nabla\phi_h$.

[66] Notice that there will necessarily be at least one nonsimple factor $S_*\nabla^{(\nu)}R_{ijkl}$ or $\nabla^{(B)}\Omega_h$, by virtue of the factors ∇Y (or $\nabla\omega_1,\nabla\omega_2$), therefore that corollary can be applied.

and

$$\sum_{t\in T^k} a_t X_* \operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_k} \times C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) + \sum_{\zeta\in Z_b} a_\zeta X_*\operatorname{div}_{i_1} C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) = \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau); \tag{2-89}$$

the sublinear combination $\sum_{\zeta\in Z_b}\dots$ above is *generic*.

Split the index set T^k according to which of the factors $\nabla\psi_2,\dots,\nabla\psi_\tau,\nabla Y$ (or $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla\chi_1$) contain the k free indices. Thus we write $T^k=\bigcup_{\alpha\in A}T^{k,\alpha}$ (each $\alpha\in A$ corresponds to a k-subset of the set of factors $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla Y$ or $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla\chi_1$). We will then show that for each $\alpha\in A$ there exists a linear combination $\sum_{b\in B^\alpha} a_b C^{b,i_1\dots i_{k+1}}_g$ of partial contractions in the form (2-64) or (2-65) with the first k free indices belonging to the factors in the set α, and the free index i_{k+1} *not* belonging to $\nabla\psi_1$, so that

$$\sum_{t\in T^{k,\alpha}} a_t C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon - X_*\operatorname{div}_{i_{k+1}}\sum_{b\in B^\alpha} a_b C^{b,i_1\dots i_{k+1}}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) \times\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon = \sum_{j\in J} a_j C^{j,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) \times\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon, \tag{2-90}$$

and

$$\sum_{t\in T^{k,\alpha}} a_t C^{t,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \times\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon - X_*\operatorname{div}_{i_{k+1}}\sum_{b\in B^\alpha} a_b C^{b,i_1\dots i_{k+1}}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \times\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon = \sum_{j\in J} a_j C^{j,i_1\dots i_k}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \times\nabla_{i_1}\upsilon\dots\nabla_{i_k}\upsilon. \tag{2-91}$$

If we can show the above for every $\alpha\in A$, then replacing the factor $\nabla\upsilon$ by $X_+\operatorname{div}$ we can derive our claim (2-86), (2-87). $\square$

Proof of (2-90) *and* (2-91). Refer to (2-88) and (2-89). Denote Y or χ_1 by $\psi_{\tau+1}$ for uniformity. We pick out any $\alpha\in A$; assume that $\alpha=\{\nabla\psi_{x_1},\dots,\nabla\psi_{x_k}\}$.

Pick out the sublinear combination where the factors $\nabla\psi_{x_1},\dots,\nabla\psi_{x_k}$ which belong to α contract against the same factor as $\nabla\psi_1$. This sublinear combination Z_g vanishes separately (that is, $Z_g=0$). We then apply the eraser to the factors $\nabla\psi_2,\dots,\nabla Y\in A$ (notice this is *well-defined*, since all the above factors *and* the factor $\nabla\psi_1$ contract against nonspecial indices). We obtain a new true equation, which we denote by $\mathrm{Erase}[Z_g]=0$. It then follows that $\mathrm{Erase}[Z_g]\cdot(\nabla_{i_1}\psi_{x_1}\nabla^{i_1}\upsilon\dots\nabla_{i_k}\psi_{x_k}\nabla^{i_k}\upsilon)=0$ is our desired conclusion (2-90), (2-91). □

Sketch of proof of (2-78), (2-79), (2-82), (2-83). These equations can be proven by only a slight modification of the idea above. We again subdivide the index sets T^k, V^k according to the set of factors $\nabla\psi_2,\dots,\nabla\psi_\tau$ or $\nabla\psi_2,\dots,\nabla\psi_\tau,\nabla\omega_1$ that contain the k free indices (so we write $T^k=\bigcup_{\alpha\in A}T^{k,\alpha}$ and $V^k=\bigcup_{\alpha\in A}V^{k,\alpha}$) and we prove the claims above separately for those sublinear combinations.

To prove this, we pick out the sublinear combination in our hypotheses with the factors $\nabla\psi_h$, $h\in\alpha$ contracting against the same factor against which $\nabla\psi_1$ and ∇Y (or $\nabla\psi_1$ and $\nabla\omega_1$) are contracting. Say $\alpha=\{h_1,\dots,h_k\}$; we then formally replace the expressions

$$S_*\nabla^{(\nu)}_{r_1\dots r_\mu l_1\dots l_k}R_{ijkl}\nabla_{l_1}\psi_{h_1}\dots\nabla^{l_k}\psi_{h_k}\nabla^i\tilde{\phi}_1\nabla^j\psi_1\nabla^kY \quad\text{or}$$
$$\nabla^{(A)}_{r_1\dots r_\mu l_1\dots l_k st}\Omega_1\nabla^{l_1}\psi_{h_1}\dots\nabla^{l_k}\psi_{h_k}\nabla^s\psi_1\nabla^tY$$

and so on by expressions

$$S_*\nabla^{(\nu-k)}_{r_1\dots r_\mu}R_{ijkl}\nabla^i\tilde{\phi}_1\nabla^j\psi_1\nabla^kY \quad\text{or}\quad \nabla^{(A-k)}_{r_1\dots r_\mu st}\Omega_1\nabla^s\psi_1\nabla^tY$$

and derive our claims (2-78), (2-79), (2-82), (2-83) as above. □

Proof of the claims of our general statement: Equations (2-68) and (2-69). We will prove these claims by an induction. Our inductive assumptions are that (2-68), (2-69) follow from (2-66), (2-67) for any weight $-W'$, $W'<K$ and when $W'=K$ they hold for any length $\upsilon+b\geq\gamma\geq 2$. We will then show the claim when the weight is $-K$, and $\upsilon+b=\gamma+1$. In the end, we will check our claims for the base case $\upsilon+b=2$.

Proof of the inductive step. Refer back to (2-66), (2-67). We will prove this claim in four steps.

Step 1: First, we will denote by Z^{spec}_a, $\bar{Z}^{spec}_a$, Z^{spec}_b the index sets of the tensor fields for which ∇Y or one of the factors $\nabla\chi_1$, $\nabla\chi_2$ (respectively) contracts against a special index. Then *using the inductive assumptions of our general claim*, we will show that there exists a linear combination of 2-tensor fields (indexed in W below) which satisfies all the requirements of (2-66), (2-68) so that

$$\sum_{\zeta\in Z_a^{spec}} a_\zeta C_g^{\zeta,i_1}\nabla_{i_1}\upsilon - X_*\operatorname{div}_{i_2}\sum_{w\in W} a_w C_g^{w,i_1i_2}\nabla_{i_1}\upsilon = \sum_{\zeta\in Z_a^{OK}} a_\zeta C_g^{\zeta,i_1}\nabla_{i_1}\upsilon + \sum_{j\in J} a_j C_g^{j,i_1}\nabla_{i_1}\upsilon, \tag{2-92}$$

where the tensor fields n Z_a^{OK} are *generic* linear combinations of tensor fields of the same general type as the ones indexed in Z_a in (2-66), (2-68) and where in addition none of the factors ∇Y or $\nabla\chi_1, \nabla\chi_2$ contract against a special index.

Thus, if we can show the above, by replacing $\nabla\upsilon$ by an $X_*\operatorname{div}_i$, and substituting back into (2-66), (2-68), we are reduced to showing (2-67), (2-69) under the additional assumption that $Z_a^{spec} = \varnothing$.

Step 2: Then, under the assumption that $Z_a^{spec} = \varnothing$, we will show that we can write

$$\sum_{\zeta\in Z_b^{spec}} a_\zeta X_*\operatorname{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in \bar{Z}_a^{spec}} a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_c} C_g^{\zeta,i_1\dots i_c} = X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_b}\sum_{c\in C} a_c C_g^{c,i_1\dots i_b} + \sum_{j\in J} a_j C_g^{j,i_1}, \tag{2-93}$$

where the tensor fields on the right-hand side are of the general form as the ones indexed in Z_b, $\bar{Z}_a$ in our hypothesis, and moreover the factors ∇Y (or the factors $\nabla\chi_1, \nabla\chi_2$) are not contracting against special indices.

Notice that if we can show (2-92), (2-93) then we are reduced to showing our claim under the additional assumption that for each $\zeta\in Z_a\cup\bar{Z}_a\cup Z_b$ the factors ∇Y (or $\nabla\chi_1, \nabla\chi_2$) are not contracting against special indices. We will show (2-92), (2-93) below.

Proof of (2-67), (2-69) *under the assumption that for each* $\zeta\in Z_a\cup\bar{Z}_a\cup Z_b$ *the factors* ∇Y *or* $(\nabla\chi_1, \nabla\chi_2)$ *do not contract against special indices.*

Step 3: Proof of (2-94) *below.* We note that for all the tensor fields in the rest of this proof will *not* have the factor ∇Y (or any of the factors $\nabla\chi_1, \nabla\chi_2$) contracting against a special index in any factor $S_*\nabla^{(\nu)}R_{ijkl}$ or $\nabla^{(B)}\Omega_h$. Now, we arbitrarily pick out one factor $T = S_*\nabla^{(\nu)}R_{ijkl}$ or $T = \nabla^{(B)}\Omega_x$ in $\overline{\kappa}_{\text{simp}}$ and call it the "chosen factor" for the rest of this subsection.

We will say that the factor ∇Y (or $\nabla\omega_2$) contracts against a good index in T, if it contracts against a nonspecial index in T when T is of the form $S_*\nabla^{(\nu)}R_{ijkl}$ with $\nu > 0$; when T is of the form $\nabla^{(B)}\Omega_x$, then it contracts against a good index provided $B\geq 3$.

We will say that the factor ∇Y (or $\nabla\omega_2$) contracts against a bad index if it contracts against the index $_j$ in a factor $T = S_*R_{ijkl}$ or an index in a factor $T = \nabla^{(2)}\Omega_x$. We denote by $Z_a^{BAD}\subset Z_a$ the index set of tensor fields for which ∇Y (or $\nabla\omega_2$)

contracts against a bad index. We also denote by $Z_b^{BAD} \subset Z_b$ the index set of the vector fields for which ∇Y contracts against a bad index in T and T also contains a free index. We will show that we can write

$$\sum_{\zeta \in Z_a^{BAD} \cup Z_b^{BAD}} a_\zeta C_g^{\zeta, i_1} \nabla_{i_1} \upsilon - X_* \operatorname{div}_{i_2} \sum_{h \in H} a_h C_g^{i_1 i_2} \nabla_{i_1} \upsilon = \sum_{\zeta \in Z_a^{\prime GOOD} \cup Z_b^{\prime GOOD}} a_\zeta C_g^{\zeta, i_1} \nabla_{i_1} \upsilon + \sum_{j \in J} a_j C_g^j, \tag{2-94}$$

where all the tensor fields indexed in $Z_a^{\prime GOOD} \cup Z_b^{\prime GOOD}$ are *generic* vector fields of the forms indexed in Z_a, Z_b, only with the factors ∇Y or $\nabla \omega_2$ contracting against a *good* index in the factor T. The tensor fields indexed in H are as required in the claim of our general statement (they correspond to the index set W in our general statement).

Step 4: Proof that (2-94) implies our claims (2-68), (2-69). We start by proving (2-94) (that is, we prove Step 3). Then, we will show how we can derive our claim from (2-94) (that is, we then prove Step 4).

Proof of Step 3: Proof of (2-94). We can prove this equation by virtue of our inductive assumption on our general claim. First, we define $\bar{Z}_a^{BAD} \subset \bar{Z}_a$ to stand for the index set of tensor fields where the factor ∇Y (or $\nabla \omega_2$) is contracting against a bad index in the chosen factor. We pick out the sublinear combination in our lemma assumption where ∇Y (or $\nabla \omega_2$) are contracting against the chosen factor $T = S_* R_{ijkl}$ or $T = \nabla^{(2)} \Omega_x$). This sublinear combination must vanish separately, and we thus derive that

$$\sum_{\zeta \in Z_a^{BAD} \cup Z_b^{BAD}} a_\zeta X_{**} \operatorname{div}_{i_1} C_g^{\zeta, i_1} + \sum_{\zeta \in \bar{Z}_a^{BAD}} a_\zeta X_{**} \operatorname{div}_{i_1} \dots X_{**} \operatorname{div}_{i_c} C_g^{\zeta, i_1 \dots i_c} + \sum_{\zeta \in Z_b^{nvBAD}} a_f C_g^{f, i_1} = \sum_{j \in J} a_j C_g^j, \tag{2-95}$$

where $X_{**} \operatorname{div}_{i_1}$ stands for the sublinear combination for which ∇_{i_1} is not allowed to hit the chosen factor T. $Z_b^{nvBAD} \subset Z_b$ stands for the index set of tensor fields indexed in Z_b with the free index ${}_{i_1}$ *not* belonging to the chosen factor and also with the factor ∇Y (or $\nabla \omega_2$) contracting against a bad index.

Now, define an operation Op[...] that acts on the complete contractions above by formally replacing any expression $\nabla^{(2)}_{ij} \Omega_x \nabla^i Y$ (or $\nabla^{(2)}_{ij} \Omega_x \nabla^i \chi_2$) by $\nabla_j D$ (D is a scalar function), or any expression $S_* R_{ijkl} \nabla^i \tilde{\phi}_1 \nabla^j Y$ (or $S_* R_{ijkl} \nabla^i \tilde{\phi}_1 \nabla^j \chi_2$) by $\nabla_{[k} \theta_1 \nabla_{l]} \theta_2$. (Denote by $\tilde{\kappa}_{\text{simp}}$ the simple character of these resulting vector fields.)

Acting on (2-95) by Op[...] produces a true equation, which we may write out as

$$\sum_{\zeta\in Z_a^{BAD}\cup Z_b^{BAD}} a_\zeta X_{**}\operatorname{div}_{i_1}\mathrm{Op}[C]_g^{\zeta,i_1}+X_{**}\operatorname{div}_{i_1}\sum_{f\in F}a_f C_g^{f,i_1} + \sum_{\zeta\in \bar{Z}_a^{BAD}} a_\zeta X_{**}\operatorname{div}_{i_1}\dots X_{**}\operatorname{div}_{i_c} C_g^{\zeta,i_1\dots i_c}=\sum_{j\in J}a_j C_g^j. \tag{2-96}$$

Here $X_{**}\operatorname{div}_i$ stands for the sublinear combination in div_i where ∇_i is not allowed to hit the factor to which ∇_i belongs, nor any of the factors $\nabla\phi_1,\dots,\nabla\phi_u$, $\nabla\psi_1,\dots,\nabla\psi_\tau$, nor any of the factors ∇D, $\nabla\theta_1$, $\nabla\theta_2$. The vector fields indexed in F are generic vector fields with a simple character $\tilde{\kappa}_{\text{simp}}$, for which the free index $_{i_1}$ *does not* belong to any of the factors $\nabla\psi_1,\dots,\nabla\psi_\tau$ or any of the factors ∇D, $(\nabla\chi_1)$, $\nabla\theta_1$, $\nabla\theta_2$.

Now, *observe that the above equation falls under our inductive assumption of the general statement we are proving*: We now either have factors

$$\begin{aligned}&\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla D, \quad \text{or}\\ &\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla\chi_1,\nabla D, \quad \text{or}\\ &\nabla\psi_1,\dots,\nabla\psi_\tau,[\nabla\theta_1,\nabla\theta_2], \quad \text{or}\\ &\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla\chi_1,[\nabla\theta_1,\nabla\theta_2].\end{aligned}$$

Notice that the tensor fields indexed in H_a^{BAD}, H_b^{BAD} are precisely the ones that contain a free index in one of these factors. Therefore, by our inductive assumption of the "general claim" we derive that there exists a linear combination of 2-tensor fields, $\sum_{\upsilon\in V}\dots$, (with factors $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla D$ and so on, and which satisfy the $\mathfrak{p}$-property for the factors $\nabla\psi_1,\dots,\nabla\psi_\tau$) so that

$$\sum_{\zeta\in Z_a^{BAD}\cup Z_b^{BAD}} a_\zeta\,\mathrm{Op}[C]_g^{\zeta,i_1}\nabla_{i_1}\upsilon - X_{**}\operatorname{div}_{i_2}\sum_{\upsilon\in V}a_\upsilon C_g^{\upsilon,i_1i_2}\nabla_{i_1}\upsilon=\sum_{j\in J}a_jC_g^{j,i_1}\nabla_{i_1}\upsilon. \tag{2-97}$$

Now, we define an operation $\mathrm{Op}^{-1}[\dots]$ that acts on the complete contractions in the above equation by replacing the factor $\nabla_j D$ by an expression $\nabla_{ij}\Omega_x\nabla^jY$ (or $\nabla_{ij}\Omega_x\nabla^j\omega_2$), or by replacing the expression $\nabla_{[a}\theta_1\nabla_{b]}\theta_2$ by $S_*R_{ijab}\nabla^i\tilde{\phi}_1\nabla^jY$ (or $S_*R_{ijab}\nabla^i\tilde{\phi}_1\nabla^j\omega_2$). The operation Op^{-1} clearly produces a true equation, which is our desired conclusion, (2-94). □

Proof of Step 4. We derive our conclusions (2-68), (2-69) in pieces. First, we show these equations with the sublinear combinations Z_a replaced by the index set $Z_{a,spec}$, which index the terms with the free index $_{i_1}$ belonging to the factor ∇Y or $\nabla\omega_1$ (this will be Substep A). After proving this claim, we will show (2-68), (2-69) under the additional assumption that $Z_{a,spec}=\varnothing$ (this will be Substep B).

Proof of Substep A. We make the $\nabla\upsilon$s into X_* divs in (2-94) and insert the resulting equations into our lemma hypothesis. We thus derive a new equation

$$\sum_{\zeta\in Z_a} a_\zeta X_*\,\mathrm{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in Z_b^1} a_\zeta X_*\,\mathrm{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in Z_b^2} a_\zeta X_*\,\mathrm{div}_{i_1}\ldots X_*\,\mathrm{div}_{i_a} C_g^{\zeta,i_1\ldots i_a} + \sum_{j\in J} a_j C_g^j = 0, \tag{2-98}$$

where we now have that the tensor fields indexed in Z_a have a free index among the factors $\nabla\psi_1,\ldots,\nabla\psi_\tau,\nabla Y$ (or $\nabla\psi_1,\ldots,\nabla\psi_\tau,\nabla\chi_1,\nabla\chi_2$), and furthermore the factors ∇Y (or the factors $\nabla\omega_1,\nabla\omega_2$) are not contracting against a bad index in the chosen factor T. The tensor fields indexed in Z_b^1 have a free index that *does not* belong to one of the factors $\nabla\psi_1,\ldots,\nabla\psi_\tau,\nabla Y$ (or $\nabla\psi_1,\ldots,\nabla\psi_\tau,\nabla\chi_1,\nabla\chi_2$), and furthermore *if* the factor ∇Y (or one of the factors $\nabla\omega_1,\nabla\omega_2$) is contracting against a bad index in the chosen factor T, then T does not contain the free index i_1. Finally the tensor fields indexed in Z_b^2 each have rank $a\geq 2$ and all free indices belong to the factors $\nabla\psi_1,\ldots,\nabla\psi_\tau,\nabla Y,(\nabla\omega_1,\nabla\omega_2)$. We may then rewrite (2-98) in the form

$$\sum_{\zeta\in Z_a} a_\zeta X_*\,\mathrm{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in Z_b^1} a_\zeta X_*\,\mathrm{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in Z_b^{2'}} a_\zeta X_*\,\mathrm{div}_{i_1}\ldots X_*\,\mathrm{div}_{i_a} C_g^{\zeta,i_1\ldots i_a} + \sum_{j\in J} a_j C_g^j = 0, \tag{2-99}$$

where now for the tensor fields indexed in $Z_b^{2'}$, each $a\geq 1$ and the factor $\nabla\psi_1$ does not contain a free index for any of the tensor fields for which ∇Y (or one of $\nabla\omega_1,\nabla\omega_2$) contracts against a bad index in the chosen factor.

We will denote by $Z_{b,\sharp}^1\subset Z_b^1$ and ${Z_{b,\sharp}^2}'\subset Z_b^{2'}$ the index sets of tensor fields where ∇Y (or one of $\nabla\omega_1,\nabla\omega_2$) contracts against a bad index in the chosen factor T.

From (2-99) we derive

$$\sum_{\zeta\in Z_{b,\sharp}^1} a_\zeta X_{**}\,\mathrm{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in {Z_{b,\sharp}^2}'} a_\zeta X_{**}\,\mathrm{div}_{i_1}\ldots X_{**}\,\mathrm{div}_{i_a} C_g^{\zeta,i_1\ldots i_a} + \sum_{j\in J} a_j C_g^j = 0, \tag{2-100}$$

where $X_{**}\,\mathrm{div}_i$ stands for the sublinear combination in $X_*\,\mathrm{div}_i$ for which ∇_i is in addition no allowed to hit the chosen factor T.

Then, applying operation Op as in Step 3 and the inductive assumption of the general claim we are proving,[67] and then using the operation $\mathrm{Op}^{-1}[\ldots]$ as in the

[67] The resulting equation falls under the inductive assumption, as in Step 3.

proof of Step 3, we derive

$$\sum_{\zeta\in Z^1_{b,\sharp}} a_\zeta X_*\operatorname{div}_{i_1} C^{\zeta,i_1}_g+\sum_{\zeta\in Z^2_{b,\sharp}{}'} a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_a} C^{\zeta,i_1\dots i_a}_g = \sum_{\zeta\in Z_{OK}} a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_a} C^{\zeta,i_1\dots i_a}_g+\sum_{j\in J} a_j C^j_g=0, \tag{2-101}$$

where the tensor fields indexed in Z_{OK} have rank $a\ge 1$ (no free indices belonging to factors $\nabla\psi_1,\dots,\nabla Y$ or $\nabla\psi_1,\dots,\nabla\chi_2$) and furthermore have the property that the one index in ∇Y or $\nabla\omega_1$ does not contract against a bad index in the chosen factor (and it is also not free). Thus, replacing the above back into (2-99), we derive

$$\sum_{\zeta\in Z_a} a_\zeta X_*\operatorname{div}_{i_1} C^{\zeta,i_1}_g+\sum_{\zeta\in Z^{1'}_b} a_\zeta X_*\operatorname{div}_{i_1} C^{\zeta,i_1}_g + \sum_{\zeta\in Z^{2''}_b} a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_a} C^{\zeta,i_1\dots i_a}_g+\sum_{j\in J} a_j C^j_g=0, \tag{2-102}$$

where the tensor fields indexed in $Z^{1'}_b$, $Z^{2''}_b$ have the additional restriction that if the factor ∇Y (or $\nabla\omega_1,\nabla\omega_2$) is contracting against the chosen factor T then it is not contracting against a bad index in T.

We are now in a position to derive Substep A from the above: To see this claim, we just apply $\mathrm{Erase}_{\nabla Y}$ or $\mathrm{Erase}_{\nabla\omega_1}$ to (2-102) and multiply the resulting equation by $\nabla_{i_1}Y\nabla^{i_1}\upsilon$.

Substep B: Now, we are reduced to showing our claim when $Z_{a,spec}=\varnothing$. In that setting, we denote by $Z_{a,s}\subset Z_a$ the index set of vector fields in Z_a for which the free index $_{i_1}$ belongs to the factor $\nabla\psi_s$; we prove our claim separately for each of the sublinear combinations $\sum_{\zeta\in Z_{a,s}}\dots$. This claim is proven by picking out the sublinear combinations in (2-66), (2-67) where the factors $\nabla\psi_s$ and ∇Y (or $\nabla\chi_1$) contract against the same factor;[68] we then apply the eraser to $\nabla\psi_s$ (this is well-defined and produces a true equation), and multiply by $\nabla_{i_1}\psi_s\nabla^{i_1}\upsilon$. The resulting equation is precisely our claim for the sublinear combination $\sum_{\zeta\in Z_{a,s}}\dots$.

Sketch of the proof of Steps 1 and 2: Equations (2-92) *and* (2-93). We will sketch the proof of these claims for the sublinear combinations in $Z^{spec}_a\cup Z^{spec}_b\cup\bar Z^a_{spec}$ where one of the special indices in C^{ζ,i_1} is an index $_k$ or $_l$ that belongs to a factor $S_*\nabla^{(\nu)}R_{ijkl}$. The remaining case (where the special indices belong to factors $\nabla^{(a)}\Omega_h$) can be seen by a similar (simpler) argument.[69]

[68] These sublinear combinations vanish separately.

[69] The only extra feature in this setting is that one must prove the claim by a separate induction on the *number* of factors $\nabla\psi_z$ that are contracting against $\nabla^{(a)}\Omega_h$.

For each $\zeta \in Z_a^{spec} \cup Z_b^{spec} \cup \bar{Z}^a_{spec}$, we denote by $\bar{C}_g^{\zeta,i_1}$, $\bar{C}_g^{\zeta,i_1\dots i_\gamma}$ the tensor fields that arise from C^{ζ,i_1}, $C_g^{\zeta,i_1\dots i_\gamma}$ in (2-66), (2-68) by replacing the expressions

$$S_*\nabla^{(\nu)}_{r_1\dots r_\nu} R_{ijkl}\nabla^i\tilde{\phi}_1\nabla^k Y \quad \text{and} \quad S_*\nabla^{(\nu)}_{r_1\dots r_\nu} R_{ijkl}\nabla^i\tilde{\phi}_1\nabla^k \chi_2$$

with a factor $\nabla^{(\nu+2)}_{r_1\dots r_\nu jl}\Omega_{b+1}$. We denote by $\tilde{\kappa}_{\text{simp}}$ the resulting simple character. We derive

$$\sum_{\zeta\in Z_a^{spec}\cup Z_b^{spec}} a_\zeta X_*\operatorname{div}_{i_1}\bar{C}_g^{\zeta,i_1} + \sum_{\zeta\in\bar{Z}_a^{spec}} a_\zeta X_*\operatorname{div}_{i_1}\dots X_*\operatorname{div}_{i_\gamma}\bar{C}_g^{\zeta,i_1} + \sum_{j\in J} a_j\bar{C}_g^j = 0. \tag{2-103}$$

Now, again applying the inductive assumption of our general statement to the above, we derive that there is a linear combination of tensor fields (indexed in W below) with a free index ${}_{i_1}$ belonging to one of the factors $\nabla\psi_1,\dots,\nabla\psi_\tau$ or $\nabla\psi_1,\dots,\nabla\psi_\tau,\nabla\chi_1$ so that

$$\sum_{\zeta\in Z_a^{spec}} a_\zeta\bar{C}_g^{\zeta,i_1}\nabla_{i_1}\upsilon - X_*\operatorname{div}_{i_2}\sum_{w\in W} a_w C_g^{w,i_1i_2}\nabla_{i_1}\upsilon = \sum_{j\in J} a_j\bar{C}_g^j. \tag{2-104}$$

Now, by applying an operation Op^* to the above which formally replaces the factor $\nabla^{(A)}_{r_1\dots r_A}\Omega_x$ with a factor

$$S_*\nabla^{(A-2)}_{r_1\dots r_{A-2}} R_{ir_{A-1}kr_A}\nabla^i\tilde{\phi}_1\nabla^k Y \quad \text{or} \quad S_*\nabla^{(A-2)}_{r_1\dots r_{A-2}} R_{ir_{A-1}kr_A}\nabla^i\tilde{\phi}_1\nabla^k\chi_2,$$

we derive (2-92) (since we can repeat the permutations by which (2-104) is made to hold formally, modulo introducing correction terms that allowed in the right-hand side of (2-92)).

We will now prove (2-93) (that is, Step 2) by repeating the induction performed in the "Consequence" we derived above (where we showed that inductively assuming (2-76), (2-77) we can derive (2-78), (2-79)):

We will show the claim of Step 2 in pieces: First consider the tensor fields indexed in $\bar{Z}_{a,\mathfrak{p}}$ of minimum rank 2 (denote the corresponding index set by $\bar{Z}^2_{a,\mathfrak{p}}$); we then show that we can write

$$\begin{aligned}\sum_{\zeta\in\bar{Z}^2_{a,\mathfrak{p}}} a_\zeta X\operatorname{div}_{i_1} X\operatorname{div}_{i_2} C_g^{\zeta,i_1i_2} &= \sum_{\zeta\in\bar{Z}^3_{a,\mathfrak{p}}} a_\zeta X\operatorname{div}_{i_1}\dots X\operatorname{div}_{i_3} C_g^{\zeta,i_1\dots i_3} \\ &\quad + \sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X\operatorname{div}_{i_1} C_g^{\zeta,i_1} + \sum_{\zeta\in Z_{OK}} a_\zeta X\operatorname{div}_{i_1}\dots X\operatorname{div}_{i_a} C_g^{\zeta,i_1\dots i_a} + \sum_{j\in} a_j C^j.\end{aligned} \tag{2-105}$$

The tensor fields indexed in $\bar{Z}^3_{a,\mathfrak{p}}$, $Z_{b,\mathfrak{p}}$ on the right-hand side are generic linear combinations in those forms (the first with rank 3). The tensor fields indexed in Z_{OK} are generic linear combinations as allowed in the right-hand side of (2-93). Assuming we can prove (2-105), we are then reduced to showing our claim when the minimum rank among the tensor fields indexed in $Z_{a,\mathfrak{p}}$ is 3. We may then "forget" about any $X\,\mathrm{div}_{i_h}$ where $_{i_h}$ belongs to the factor $\nabla\psi_1$. Therefore, we are reduced to showing our claim when the minimum rank is 2 and the factor $\nabla\psi_1$ does not contain a free index. We then claim our claim by an induction (for the rest of this derivation, all tensor fields will *not* have a free index in the factor $\nabla\psi_1$): Assume that the minimum rank of the tensor fields indexed in $\bar{Z}_{a,\mathfrak{p}}$ is k, and they are indexed in $\bar{Z}^k_{a,\mathfrak{p}}$. We then show that we can write

$$\sum_{\zeta\in\bar{Z}^k_{a,\mathfrak{p}}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_k} C^{\zeta,i_1\dots i_k}_g = \sum_{\zeta\in\bar{Z}^{k+1}_{a,\mathfrak{p}}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_{k+1}} C^{\zeta,i_1\dots i_{k+1}}_g + \sum_{\zeta\in Z_{b,\mathfrak{p}}} a_\zeta X\,\mathrm{div}_{i_1} C^{\zeta,i_1}_g + \sum_{\zeta\in Z_{OK}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_a} C^{\zeta,i_1\dots i_a}_g + \sum_{j\in} a_j C^j. \tag{2-106}$$

The tensor fields indexed in $\bar{Z}^3_{a,\mathfrak{p}}$, $Z_{b,\mathfrak{p}}$ on the right-hand side are generic linear combinations in those forms (the first with rank $k+1$). The tensor fields indexed in Z_{OK} are generic linear combinations as allowed in the right-hand side of our Step 2.

Iteratively repeating this step we are reduced to proving Step 2 when $Z_{a,\mathfrak{p}}=\varnothing$. In that case we then assume that the tensor fields indexed in $Z_{b,\mathfrak{p}}$ have minimum rank k (and the corresponding index set is $Z^k_{b,\mathfrak{p}}$) and we show that we can write

$$\sum_{\zeta\in Z^k_{b,\mathfrak{p}}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_k} C^{\zeta,i_1\dots i_k}_g = \sum_{\zeta\in Z^{k+1}_{b,\mathfrak{p}}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_{k+1}} C^{\zeta,i_1\dots i_{k+1}}_g + \sum_{\zeta\in Z_{OK}} a_\zeta X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_a} C^{\zeta,i_1\dots i_a}_g + \sum_{j\in} a_j C^j, \tag{2-107}$$

(with the same conventions as in the above equation).

If we can prove (2-105) and (2-107) we will have shown our Step 2.

Proof of (2-105), (2-106), (2-107). We start with a small remark: If the chosen factor is of the form $S_*\nabla^{(\nu)}R_{ijkl}$, our assumption implies a more convenient equation: Consider the tensor fields $C^{\zeta,i_1\dots i_a}_g$, $\zeta\in\bar{Z}_{a,\mathfrak{p}}\cup Z_{b,\mathfrak{p}}$; we denote by $\tilde{C}^{\zeta,i_1\dots i_a}_g$ the tensor fields that arise from $C^{\zeta,i_1\dots i_a}_g$ by replacing the expression

$$\nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}\nabla^i\tilde{\phi}_1\nabla^kY \quad (\text{or } \nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}\nabla^i\tilde{\phi}_1\nabla^k\chi_2)$$

by a factor $\nabla^{(\nu+2)}_{r_1\dots r_\nu jl}\Omega_{p+1}$. We then derive

$$\sum_{\zeta\in\bar{Z}_a\cup Z_b} a_\zeta X_*\,\mathrm{div}_{i_1}\dots X_*\,\mathrm{div}_{i_a} \times\tilde{C}^{\zeta,i_1\dots i_a}_g(\Omega_1,\dots,\Omega_{p+1},\phi_2,\dots,\phi_u,(\chi_1),\psi_1,\dots,\psi_\tau) = \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_{p+1},\phi_2,\dots,\phi_u,(\chi_1),\psi_1,\dots,\psi_\tau). \tag{2-108}$$

Now we can derive our claims.

Proof of (2-106). We split the index set $Z_{\bar{Z}^2_{a,\mathfrak{p}}}$ according to the two factors that contain the two free indices and we show our claim for each of those tensor fields separately. The proof goes as follows: We pick out the sublinear combination in our hypothesis (or in (2-108)) where the factors $\nabla\psi_h, \nabla\psi_{h'}$ (or $\nabla\psi_h, \nabla\chi_2$) contract against the same factor. Clearly, this sublinear combination, X_g, vanishes separately. We then formally erase the factor $\nabla\psi_h$. Then, we apply the inductive assumption of our general claim to the resulting equation (the minimum rank of the tensor fields will be 1), and (in case our assumption is (2-108) we also apply an operation Op^{-1} which replaces the factor $\nabla^{(y)}_{r_1\dots r_y}\Omega_{p+1}$ by

$$S_*\nabla^{(y-2)}_{r_1\dots r_{y-2}}R_{ir_{y-1}kr_y}\nabla^i\tilde{\phi}_1\nabla^k Y(\nabla^k\chi_1).$$

This is our desired conclusion. □

Proof of (2-105), (2-107). Now we show (2-105) for the subset $Z^{k,\alpha}_{a,\mathfrak{p}}$ (which indexes the k-tensor fields for which the free indices ${}_{i_1}, \dots, {}_{i_k}$ belong to a chosen subset of the factors $\nabla\psi_1, \dots, \nabla\psi_\tau, (\nabla\chi_1)$ (hence the label α designates the chosen subset). To prove this equation, we pick out the sublinear combination in (2-108) where the factors $\nabla\psi_2, \dots, \nabla\psi_\tau, (\nabla\chi_1)$ (indexed in α) contract against the same factor as $\nabla\psi_1$. Then we apply the eraser to these factors and the indices they contract against. This is our desired conclusion. To show (2-107), we only have to treat the factors $\nabla\psi_h$ as factors $\nabla\phi_h$. The claim then follows by applying Corollary 1 in [A 2010] and making the factors $\nabla\upsilon$ into X divs.[70] □

Proof of the base case ($v+b=2$) *of the general claim.* We first prove our claim when our hypothesis is (2-67) (as opposed to (2-66)).

Proof of the base case under the hypothesis (2-67). We observe that the weight $-K$ in our assumption must satisfy $K \geq 2\tau+8$ if $v>0$ and $K \geq 2\tau+6$ if $v=0$.

First consider the case where we have the strict inequalities $K > 2\tau+8$ if $v>0$ and $K > 2\tau+6$ if $v=0$. In that case our first claim of the base case can be proven straightforwardly, by picking out a removable index in each C^{ζ,i_a}_g, $\zeta \in Z_a$ and treating it as an X_* div (which can be done when we only have two real factors). Thus, in this setting we only have to show our second claims (2-70), (2-71).

In this setting, by using the "manual" constructions as in [A 2011], we can construct explicit tensor fields which satisfy all the assumptions of our claim in the

[70] Observe that by virtue of the factor $\nabla\psi_1$, we must have at least one nonsimple factor $S_*\nabla^{(\nu)}R_{ijkl}$ or $\nabla^{(B)}\Omega_h$ in (2-108)–hence (2-108) does not fall under any of the "forbidden cases" of Corollary 1 in [A 2010], by inspection.

base case (each with rank ≥ 2), so that

$$X_+ \operatorname{div}_{i_1} \sum_{\zeta \in Z'_a \cup Z_b} a_\zeta C_g^{\zeta, i_1}(\Omega_1, \dots, \Omega_b, \phi_1, \dots, \phi_v, [\chi_1, \chi_2], \psi_1, \dots, \psi_\tau) \tag{2-109}$$
$$= \sum_{q \in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q, i_q}(\Omega_1, \dots, \Omega_b, \phi_1, \dots, \phi_v, [\chi_1, \chi_2], \psi_1, \dots, \psi_\tau)$$
$$+ \sum_{p \in P} a_p X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_{c+1}} \times C_g^{p, i_1 \dots i_{c+1}}(\Omega_1, \dots, \Omega_b, \phi_1, \dots, \phi_v, [\chi_1, \chi_2], \psi_1, \dots, \psi_\tau)$$
$$+ \sum_{j \in J} a_j C_g^j.$$

Here the tensor field $C_g^{p, i_1 \dots i_{c+1}}$ will be in one of three forms:

• If $v = 2$ then each $C_g^{p, i_1 \dots i_{c+1}}$ will be in the form

$$pcontr\big(S_* \nabla^{(\nu_1) f_{b_1} \dots f_{b_h}}_{\quad i_1 \dots i_{c-1}} R_{x_1 j i_c l} \otimes S_* \nabla^{(\nu_2) f_{d_1} \dots f_{d_y}} R_{x_v}{}^{j'}{}_{i_{c+1}}{}^{l} \otimes [\nabla^j \chi_1 \otimes \nabla_{j'} \chi_2] \otimes \nabla_{f_1} \psi_1 \dots \otimes \nabla_{f_\tau} \psi_\tau \otimes \nabla^{x_1} \tilde{\phi}_1 \otimes \nabla^{x_2} \tilde{\phi}_2\big), \tag{2-110}$$

where $\{b_1, \dots, b_h, d_1, \dots, d_y\} = \{1, \dots, \tau\}$.

• If $v = 1$ then $\sum_{p \in P} \dots = 0$ (this can be arranged because of the two antisymmetric indices $_{k}, _{l}$ in the one factor $S_* \nabla^{(\nu)} R_{ijkl}$).

• If $v = 0$ then each $C_g^{p, i_1 \dots i_{c+1}}$ will be in the form

$$pcontr\big(\nabla^{(A_1) f_{b_1} \dots f_{b_h}}_{\quad i_1 \dots i_{c-1} j i_c} \Omega_1 \otimes \nabla^{(A_2) f_{d_1} \dots f_{d_y} j' i_{c+1}} \Omega_2 \otimes [\nabla^j \chi_1 \otimes \nabla_{j'} \chi_2] \otimes \nabla_{f_1} \psi_1 \dots \otimes \nabla_{f_\tau} \psi_\tau \otimes \nabla^{x_1} \tilde{\phi}_1 \otimes \nabla^{x_2} \tilde{\phi}_2\big), \tag{2-111}$$

where $\{b_1, \dots, b_h, d_1, \dots, d_y\} = \{1, \dots, \tau\}$.

Then, picking out the sublinear combination in (2-110), (2-111) with factors $\nabla \psi_1, \dots, \nabla \psi_\tau, \nabla \chi_1, \nabla \chi_2$, we derive that $\sum_{p \in P} \dots = 0$. This is precisely our desired conclusion in this case.

Now, the case where we have the equalities in our lemma hypothesis, $K = 2\tau + 8$ if $v > 0$ and $K = 2\tau + 6$ if $v = 0$. In this case we note that in our hypothesis $Z_b = \varnothing$ if $v \neq 1$, while $Z_a = \bar{Z}_a = \varnothing$ if $v = 1$.

Then, if $v \neq 1$, by the "manual" constructions as in [A 2011], it follows that we can construct tensor fields (as required in the claim of our "general claim"), so that:

$$\sum_{\zeta\in Z_a} a_\zeta C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon \\ \times X_*\operatorname{div}_{i_2} a_\zeta C^{\zeta,i_1i_2}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon \\ = a_* C^{*,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon \\ + \sum_{j\in J} a_j C^{j,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau)\nabla_{i_1}\upsilon, \tag{2-112}$$

where the tensor field C^{*,i_1}_g is in the form

$$pcontr\big(S_*\nabla^{(\nu_1)f_1\dots f_{\tau-1}}R_{x_1}{}^{f_\tau}{}_{kl}\otimes R_{x_2}{}^{j'kl} \\ \otimes[\nabla_{i_1}\chi_1\otimes\nabla_{j'}\chi_2]\otimes\nabla_{f_1}\psi_1\cdots\otimes\nabla_{f_\tau}\psi_\tau\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\nabla^{x_2}\tilde{\phi}_2\big), \tag{2-113}$$

if $v=2$, and in the form

$$pcontr\big(\nabla^{(\tau+1)f_1\dots f_\tau}{}_s\Omega_1\otimes\nabla^{j's}\Omega_2 \\ \otimes[\nabla_{i_1}\chi_1\otimes\nabla_{j'}\chi_2]\otimes\nabla_{f_1}\psi_1\cdots\otimes\nabla_{f_\tau}\psi_\tau\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\nabla^{x_2}\tilde{\phi}_2\big), \tag{2-114}$$

if $v=0$.

Thus, we are reduced to the case where Z_a only consists of the vector field (2-113) or (2-114), and all other tensor fields in our lemma hypothesis have rank ≥ 2 (we have denoted their index set by Z'_a). We then show that we can write

$$X_+\operatorname{div}_{i_1}\sum_{\zeta\in Z'_a} a_\zeta C^{\zeta,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\ = \sum_{q\in Q} a_q X_+\operatorname{div}_{i_1} C^{q,i_1}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\ + \sum_{p\in P} a_p X_+\operatorname{div}_{i_1}\dots X_+\operatorname{div}_{i_{c+1}} \\ \times C^{p,i_1\dots i_{c+1}}_g(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\ + \sum_{j\in J} a_j C^j_g, \tag{2-115}$$

where the tensor fields indexed in P here each have rank ≥ 2 and are all in one of the forms

$$pcontr\big(S_*\nabla^{(\nu_1)f_1\dots f_{\tau-1}}R_{x_1}{}^{f_\tau}{}_{i_kl}\otimes S_*R_{x_v}{}^{j'kl} \\ \otimes[\nabla_{i_1}\chi_1\otimes\nabla_{j'}\chi_2]\otimes\nabla_{y_1}\psi_1\cdots\otimes\nabla_{y_\tau}\psi_\tau\otimes\nabla^{x_1}\tilde{\phi}_1\otimes\nabla^{x_2}\tilde{\phi}_2\big), \tag{2-116}$$

or

$$pcontr\big(\nabla^{(\nu_1)f_1\dots f_\tau}{}_s\Omega_1\otimes\nabla^{j's}\Omega_2\otimes[\nabla_{i_1}\chi_1\otimes\nabla_{j'}\chi_2]\otimes\nabla_{y_1}\psi_1\cdots\otimes\nabla_{y_\tau}\psi_\tau\big), \tag{2-117}$$

where each of the indices ^{f_h} contracts against one of the indices ${}_{y_q}$. The indices ${}_{y_q}$ that do *not* contract against an index ^{f_h} are free indices.

Then, replacing the above into our lemma hypothesis (and making all the $\nabla\upsilon$s into X_+ divs), we derive that $a_p = 0$ for every $p \in P$ and $a_* = 0$. This concludes the proof of the base case when $v + b = 2$, $v \neq 1$. In the case $v = 1$ we show our claim by just observing that we can write

$$
\begin{aligned}
\text{(2-118)}\quad & X_+ \operatorname{div}_{i_1} \sum_{\zeta \in Z_b} a_\zeta C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\
&= \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\
&\quad + \sum_{j\in J} a_j C_g^j(\Omega_1,\dots,\Omega_b,\phi_1,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau);
\end{aligned}
$$

this concludes the proof of the base case, when the tensor fields in our lemma hypothesis are in the form (2-67). □

Now, we consider the setting where our hypothesis is (2-66). We again observe that if $v = 0$ then the weight $-K$ in our hypothesis must satisfy $K \geq 2\tau + 4$. If $v > 0$ it must satisfy $K \geq 2\tau + 6$. We then again first consider the case where we have the strict inequalities in the hypothesis of our general claim.

In this case (where we have the strict inequalities $K > 2\tau + 4$ if $v = 0$ and $K > 2\tau + 6$ if $v \neq 0$) our first claim follows straightforwardly (as above, we just pick out one removable index in each C_g^{ζ,i_1}, $\zeta \in Z_a$ and treat it as an X_* div). To show the second claim we proceed much as before:

We can "manually" construct tensor fields in order to write

$$
\begin{aligned}
\text{(2-119)}\quad & X_+ \operatorname{div}_{i_1} \sum_{\zeta \in Z'_a \cup Z_b} a_\zeta C_g^{\zeta,i_1}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,Y,\psi_1,\dots,\psi_\tau) \\
&= \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} C_g^{q,i_q}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\
&\quad + \sum_{p\in P} a_p X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_{c+1}} \\
&\qquad\qquad \times C_g^{p,i_1\dots i_{c+1}}(\Omega_1,\dots,\Omega_b,\phi_1,\dots,\phi_v,[\chi_1,\chi_2],\psi_1,\dots,\psi_\tau) \\
&\quad + \sum_{j\in J} a_j C_g^j.
\end{aligned}
$$

Here the tensor field $C_g^{p,i_1\dots i_{c+1}}$ will be in one of three forms:

• If $v = 2$ then each $C^{p,i_1\dots i_{c+1}}_g$ will be

$$pcontr\big(S_*\nabla^{(\nu_1)f_{b_1}\dots f_{b_h}}_{i_1\dots i_{c-1}} R_{x_1}{}^{f_{b_{h+1}}}{}_{i_c l} \otimes S_*\nabla^{(\nu_2)f_{d_1}\dots f_{d_y}} R_{x_v}{}^{f_{d_{y+1}}}{}_{i_{c+1}}{}^{l} \otimes\nabla_{f_{\tau+1}} Y \otimes \nabla_{f_1}\psi_1 \dots \otimes \nabla_{f_\tau}\psi_\tau \otimes \nabla^{x_1}\tilde{\phi}_1 \otimes \nabla^{x_2}\tilde{\phi}_2\big), \tag{2-120}$$

where $\{b_1, \dots, b_{h+1}, d_1, \dots, d_{y+1}\} = \{1, \dots, \tau+1\}$.

• If $v = 1$ then $\sum_{p\in P}\dots = 0$ (this is because of the two antisymmetric indices $_k, _l$ in the one factor $S_*\nabla^{(\nu)} R_{ijkl}$).

• If $v = 0$ then each $C^{p,i_1\dots i_{c+1}}_g$ will be in the form

$$pcontr\big(\nabla^{(A_1)f_{b_1}\dots f_{b_h}}_{i_1\dots i_{c-1} i_c}\Omega_1 \otimes \nabla^{(A_2)f_{d_1}\dots f_{d_y} i_{c+1}}\Omega_2 \otimes\nabla_{f_{\tau+1}} Y \otimes \nabla_{f_1}\psi_1 \dots \otimes \nabla_{f_\tau}\psi_\tau\big), \tag{2-121}$$

where $\{b_1, \dots, b_h, d_1, \dots, d_y\} = \{1, \dots, \tau+1\}$.

Then, picking out the sublinear combination in (2-120), (2-121) with factors $\nabla\psi_1, \dots, \nabla\psi_\tau, \nabla Y$ we derive that $\sum_{p\in P}\dots = 0$. This is precisely our desired conclusion in this case.

Finally, we prove our claim when we have the equalities $K = 2\tau + 4$ if $v < 2$ and $K = 2\tau + 6$ if $v = 2$) in the hypothesis of our general claim.

In this case by "manually" constructing X_+ divs so that we can write

$$\begin{aligned} \sum_{\zeta\in Z'_a\cup Z_b\cup \bar{Z}_a} a_\zeta X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_a} C^{\zeta,i_1\dots i_a}_g(\Omega_1, \dots \Omega_b, Y, \psi_1, \dots, \psi_\tau) \\ = \sum_{q\in Q} a_q X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_a} C^{q,i_1\dots i_a}_g(\Omega_1, \dots \Omega_b, Y, \psi_1, \dots, \psi_\tau) \\ + \sum_{p\in P} a_p X_+ \operatorname{div}_{i_1} \dots X_+ \operatorname{div}_{i_a} C^{p,i_1\dots i_a}_g(\Omega_1, \dots \Omega_b, Y, \psi_1, \dots, \psi_\tau) \\ + \sum_{j\in J} a_j C^j_g(\Omega_1, \dots \Omega_b, Y, \psi_1, \dots, \psi_\tau). \end{aligned} \tag{2-122}$$

Here the tensor fields indexed in P are in specific forms:

• If $v = 0$ then they will either be in the form

$$pcontr\big(\nabla_{i_*} Y \otimes \nabla^{(A)f_{x_1}\dots f_{x_a} s}\Omega_1 \otimes \nabla^{(B)f_{x_{a+1}}\dots f_{x_\tau}}_{s}\Omega_2 \otimes \nabla_{f_1}\psi_1 \otimes \dots \otimes \nabla_{f_\tau}\phi_\tau\big), \tag{2-123}$$

where $\{x_1, \dots, x_\tau\} = \{1, \dots, \tau\}$, or in the form

$$pcontr\big(\nabla_q Y \otimes \nabla^{(A)f_{x_1}\dots f_{x_a}}_{i_*}\Omega_1 \otimes \nabla^{(B)f_{x_{a+1}}\dots f_{x_\tau} q}\Omega_2 \otimes \nabla_{f_1}\psi_1 \otimes \dots \otimes \nabla_{f_\tau}\phi_\tau\big), \tag{2-124}$$

where $\{x_1, \dots, x_\tau\} = \{1, \dots, \tau\}$.

• If $v=2$ they will be in the form

$$pcontr\big(\nabla_{i_*}Y\otimes\nabla^{(A)f_{x_1}\dots f_{x_{a-1}}}S_*R^{if_{x_a}kl}\otimes\nabla^{(B)f_{x_{a+1}}\dots f_{x_{\tau-1}}}R^{i'f_{x_\tau}}{}_{kl}\otimes\nabla_{f_1}\psi_1\otimes\dots\otimes\nabla_{f_\tau}\phi_\tau\nabla_i\tilde{\phi}_1\otimes\nabla_{i'}\tilde{\phi}_2\big),\tag{2-125}$$

where $\{x_1,\dots,x_\tau\}=\{1,\dots,\tau\}$, or in the form

$$pcontr\big(\nabla_{q}Y\otimes\nabla^{(A)f_{x_1}\dots f_{x_{a-1}}}S_*R^{if_{x_a}ql}\otimes\nabla^{(B)f_{x_{a+1}}\dots f_{x_{\tau-1}}}R^{i'f_{x_\tau}}{}_{i_*l}\otimes\nabla_{y_1}\psi_1\otimes\dots\otimes\nabla_{y_\tau}\phi_\tau\nabla_i\tilde{\phi}_1\otimes\nabla_{i'}\tilde{\phi}_2\big).\tag{2-126}$$

• If $v=1$, Equation (2-122) will hold with $P=\varnothing$.

Then, picking out the sublinear combination in (2-122) which consists of terms with a factor ∇Y and replacing into our hypothesis, we derive that the coefficient of each of the tensor fields indexed in P must be zero. This completes the proof of our claim. □

2E. *Proof of Lemmas 2.2 and 2.4.*

Proof of Lemma 2.2. The first claim follows immediately, since each tensor field has a removable index (thus each tensor field separately can be written as an X_* div).

The proof of the second claim essentially follows the "manual" construction of divergences, as in [A 2011]. By "manually" constructing explicit divergences out of each $C^{h,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$, $h\in H_2$, we derive that we can write

$$\begin{aligned}\sum_{h\in H_2}a_hX\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha}C^{h,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\=(Const)_1X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\xi}C^{1,i_1\dots i_\xi}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\+(Const)_2X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\zeta}C^{2,i_1\dots i_\zeta}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\+\sum_{q\in Q}a_qX\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha}C^{q,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\+\sum_{j\in J}a_jC^j_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u),\end{aligned}\tag{2-127}$$

where the tensor fields indexed in Q are as required by our lemma hypothesis, while the tensor fields C^1, C^2 are explicit tensor fields which we will write out below (their precise form depends on the values p,σ_1,σ_2).[71]

We will then show that in (2-127) we will have $(Const)_1=(Const)_2=0$. That will complete the proof of Lemma 2.2. We distinguish cases based on the value of p: Either $p=2$ or $p=1$ or $p=0$.

[71] In some cases there will be no tensor fields C^1, C^2 (in which case we will just say that in (2-127) we have $(Const)_1=0$, $(Const)_2=0$).

The case $p = 2$. With no loss of generality we assume that $\nabla^{(A)}\Omega_1$ contracts against $\nabla\phi_1, \dots, \nabla\phi_x$ and that $\nabla^{(B)}\Omega_2$ contracts against $\nabla\phi_{x+1}, \dots, \nabla\phi_{x+t}$; we may also assume without loss of generality that $x \le t$. By manually constructing divergences, it follows that we can derive (2-127), where each of the tensor fields C^1, C^2 will be in the forms, respectively,

$$pcontr\big(\nabla_{i_*}Y\otimes\nabla^{(A)}_{v_1\dots v_x i_1\dots i_\gamma}\Omega_1\otimes\nabla^{(B)}_{y_1\dots y_t i_{\gamma+1}\dots i_{\gamma+\delta}}\Omega_2\otimes\nabla^{v_1}\phi_1\otimes\dots\otimes\nabla^{y_t}\phi_u\big), \tag{2-128}$$

where if $t \ge 2$ then $\delta = 0$, otherwise $t + \delta = 2$; or

$$pcontr\big(\nabla_q Y \otimes \nabla^q\nabla^{(A)}_{v_1\dots v_x i_1\dots i_\gamma}\Omega_1 \otimes \nabla^{(B)}_{y_1\dots y_t i_{\gamma+1}\dots i_{\gamma+\delta}}\Omega_2 \otimes\nabla^{v_1}\phi_1 \otimes \dots \otimes \nabla^{y_t}\phi_u\big), \tag{2-129}$$

where if $t \ge 2$ then $\delta = 0$, otherwise $t + \delta = 2$.

The case $p = 1$. We "manually" construct divergences to derive (2-127), where if $\sigma_1 = 1$ then there are no tensor fields C^1, C^2 (and hence (2-127) is our desired conclusion); if $\sigma_1 = 0$, $\sigma_2 = 1$ then there is only the tensor field C^1 in (2-127) and it is in the form

$$pcontr\big(\nabla^q Y \otimes S_*\nabla^{(\nu)}_{v_2\dots v_x i_1\dots i_\gamma}R_{i i_{\gamma+1}i_{\gamma+2}q} \otimes \nabla^{(B)}_{y_1\dots y_t i_{\gamma+1}\dots i_{\gamma+\delta}}\Omega_2 \otimes\nabla^i\tilde\phi_1 \otimes \nabla^{v_1}\phi_2 \otimes \dots \otimes \nabla^{y_t}\phi_u\big), \tag{2-130}$$

where if $t \ge 2$ then $\delta = 0$, otherwise $\delta = 2 - t$.

The case $p = 0$. We have three subcases: First $\sigma_2 = 2$, second $\sigma_2 = 1$ and $\sigma_1 = 1$, and third $\sigma_1 = 2$.

In the case $\sigma_2 = 2$, the tensor fields C^1, C^2 must be in the forms, respectively,

$$pcontr\big(\nabla_{i_*}Y \otimes S_*\nabla^{(\nu)}_{v_2\dots v_x i_1\dots i_\gamma}R_{i i_{\gamma+1}i_{\gamma+2}l} \otimes S_*\nabla^{(t-1)}_{y_1\dots y_t}R_{i' i_{\gamma+3}i_{\gamma+4}}{}^{l} \otimes\nabla^i\tilde\phi_1 \otimes \nabla^{i'}\tilde\phi_2 \otimes \nabla^{v_1}\phi_3 \otimes \dots \otimes \nabla^{y_t}\phi_u\big), \tag{2-131}$$

or

$$pcontr\big(\nabla^q Y \otimes S_*\nabla^{(\nu)}_{q v_2\dots v_x i_1\dots i_\gamma}R_{i i_{\gamma+1}i_{\gamma+2}l} \otimes S_*\nabla^{(t-1)}_{y_1\dots y_t}R_{i' i_{\gamma+3}i_{\gamma+4}}{}^{l} \otimes\nabla^i\tilde\phi_1 \otimes \nabla^{i'}\tilde\phi_2 \otimes \nabla^{v_1}\phi_3 \otimes \dots \otimes \nabla^{y_t}\phi_u\big), \tag{2-132}$$

(if $x = t = 0$ then the tensor field C^1 above will not be present).

In the case $\sigma_1 = 2$, the tensor fields C^1, C^2 must be in one of the two forms

$$pcontr\big(\nabla_{i_*}Y \otimes \nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma}R_{i i_{\gamma+1}i_{\gamma+2}l} \otimes \nabla^{(t-1)}_{y_1\dots y_t}R^{i}{}_{i_{\gamma+3}i_{\gamma+4}}{}^{l} \otimes\nabla^{v_1}\phi_1 \otimes \dots \otimes \nabla^{y_t}\phi_u\big), \tag{2-133}$$

or

$$pcontr\big(\nabla_q Y\otimes\nabla^q\nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma}R_{ii_{\gamma+1}i_{\gamma+2}l}\otimes\nabla^{(t-1)}_{y_1\dots y_t}R^{i}{}_{i_{\gamma+3}i_{\gamma+4}}{}^{l}\otimes\nabla^{v_1}\phi_1\otimes\dots\otimes\nabla^{y_t}\phi_u\big).\tag{2-134}$$

In the case $\sigma_1=1$ and $\sigma_2=1$, there will be only one tensor field C^1, in the form

$$pcontr\big(\nabla^q Y\otimes S_*\nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma}R_{ii_{\gamma+1}i_{\gamma+2}l}\otimes\nabla^{(t-1)}_{y_1\dots y_t}R_{qi_{\gamma+3}i_{\gamma+4}}{}^{l}\otimes\nabla^i\tilde\phi_1\otimes\nabla^{v_1}\phi_2\otimes\dots\otimes\nabla^{y_t}\phi_u\big).\tag{2-135}$$

We then derive that $(Const)_1=(Const)_2=0$ as in [A 2011] (by picking out the sublinear combination in (2-127) that consists of complete contractions with a factor ∇Y — differentiated only once). □

Proof of Lemma 2.4. We again "manually" construct explicit X div to write

$$\begin{aligned}\sum_{h\in H_2}a_h X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha}C^{h,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ =(Const)_1X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\xi}C^{1,i_1\dots i_\xi}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ +(Const)_2X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\zeta}C^{2,i_1\dots i_\zeta}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ +\sum_{q\in Q}a_q X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\alpha}C^{q,i_1\dots i_\alpha}_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u)\\ +\sum_{j\in J}a_jC^j_g(\Omega_1,\dots,\Omega_p,Y,\phi_1,\dots,\phi_u),\end{aligned}\tag{2-136}$$

where the tensor fields indexed in Q are as required by our lemma hypothesis, while the tensor fields C^1, C^2 are explicit tensor fields which we will write out below (they depend on the values p,σ_1,σ_2). In some cases there will be no tensor fields C^1, C^2 (in which case we will just say that in (2-127) we have $(Const)_1=0$, $(Const)_2=0$).

The case $p=2$. With no loss of generality we assume that $\nabla^{(A)}\Omega_1$ contracts against $\nabla\phi_1,\dots,\nabla\phi_x$ and that $\nabla^{(B)}\Omega_2$ contracts against $\nabla\phi_{x+1},\dots,\nabla_{\phi_{x+t}}$; we may also assume without loss of generality that $x\le t$. By manual construction of divergences, it follows that we can derive (2-127), where there is only the tensor field C^1 and it is in the form

$$pcontr\big(\nabla_{[i_*}\chi_1\otimes\nabla^{q]}\chi_2\otimes\nabla^{(A)}_{v_1\dots v_x i_1\dots i_\gamma}\Omega_1\otimes\nabla^{(B)}_{qy_1\dots y_t i_{\gamma+1}\dots i_{\gamma+\delta}}\Omega_2\otimes\nabla^{v_1}\phi_1\otimes\dots\otimes\nabla^{y_t}\phi_u\big),\tag{2-137}$$

where if $t\ge1$ then $\delta=0$, otherwise $\delta=1$.

The case $p=1$. We "manually" construct divergences to derive (2-136), where if $\sigma_1=1$ then there are no tensor fields C^1, C^2 in the right-hand side of (2-136) (and this is our desired conclusion); and if $\sigma_1 = 0$ and $\sigma_2 = 1$ then there is only the tensor field C^1 in the right-hand side of (2-136) and it is of the form

$$pcontr\big(\nabla_{[i_*}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes S_*\nabla^{(\nu)}_{v_2\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} q} \otimes \nabla^{(B)}_{y_1\dots y_t i_{\gamma+1}\dots i_{\gamma+\delta}}\Omega_2 \otimes\nabla^i\tilde{\phi}_1 \otimes \nabla^{v_1}\phi_2 \otimes\cdots\otimes \nabla^{y_t}\phi_u\big), \tag{2-138}$$

where if $t \geq 2$ then $\delta = 0$, otherwise $\delta = 2 - t$.

The case $p = 0$. We have three subcases: First $\sigma_2 = 2$, second $\sigma_2 = 1$ and $\sigma_1 = 1$, and third $\sigma_1 = 2$.

In the case $\sigma_2 = 2$, the tensor fields C^1, C^2 in the right-hand side of (2-127) will be in the two forms, respectively,

$$pcontr\big(\nabla_{[i_*}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes S_*\nabla^{(\nu)}_{q v_2\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} l} \otimes S_*\nabla^{(t-1)}_{y_1\dots y_t} R_{i' i_{\gamma+3} i_{\gamma+4}}{}^{l} \otimes \nabla^i\tilde{\phi}_1 \otimes \nabla^{i'}\tilde{\phi}_2 \otimes \nabla^{v_1}\phi_3 \otimes\cdots\otimes \nabla^{y_t}\phi_u\big), \tag{2-139}$$

or

$$pcontr(\nabla_{[p}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes S_*\nabla^{(\nu)}_{q v_2\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} p} \otimes S_*\nabla^{(t-1)}_{y_1\dots y_t} R_{i' i_{\gamma+3} i_{\gamma+4} q} \otimes \nabla^i\tilde{\phi}_1 \otimes \nabla^{i'}\tilde{\phi}_2 \otimes \nabla^{v_1}\phi_3 \otimes\cdots\otimes \nabla^{y_t}\phi_u). \tag{2-140}$$

In the case $\sigma_1 = 2$, the tensor fields C^1, C^2 will be the forms, respectively,

$$pcontr\big(\nabla_{[i_*}\omega_1 \otimes \nabla^{q]} \otimes \nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} l} \otimes \nabla^{(t-1)}_{q y_1\dots y_t} R^{i}{}_{i_{\gamma+3} i_{\gamma+4}}{}^{l} \otimes\nabla^{v_1}\phi_1 \otimes\cdots\otimes \nabla^{y_t}\phi_u\big), \tag{2-141}$$

or

$$pcontr\big(\nabla_{[p}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes \nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} p} \otimes \nabla^{(t-1)}_{y_1\dots y_t} R^{i}{}_{i_{\gamma+3} i_{\gamma+4} q} \otimes\nabla^{v_1}\phi_1 \otimes\cdots\otimes \nabla^{y_t}\phi_u\big), \tag{2-142}$$

if at least one of the two factors $\nabla^{(m)}R_{ijkl}$ contracts against a factor $\nabla\phi_h$, otherwise we can prove (2-136) with no tensor fields C^1, C^2 on the right-hand side.

In the case $\sigma_1 = 1$, $\sigma_2 = 1$, the tensor fields C^1, C^2 must be in the forms, respectively,

$$pcontr\big(\nabla_{[i_*}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes S_*\nabla^{(\nu)}_{v_1\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} l} \otimes \nabla^{(t-1)}_{y_1\dots y_t} R_{q i_{\gamma+3} i_{\gamma+4}}{}^{l} \otimes\nabla^i\tilde{\phi}_1 \otimes \nabla^{v_1}\phi_2 \otimes\cdots\otimes \nabla^{y_t}\phi_u\big), \tag{2-143}$$

or

$$pcontr(\nabla_{[p}\omega_1 \otimes \nabla^{q]}\omega_2 \otimes S_*\nabla^{(m_1)}_{v_1\dots v_x i_1\dots i_\gamma} R_{i i_{\gamma+1} i_{\gamma+2} l} \otimes \nabla^{(t-1)}_{y_1\dots y_t} R_{p q i_{\gamma+3}}{}^{l} \otimes\nabla^i\tilde{\phi}_1 \otimes \nabla^{v_1}\phi_2 \otimes\cdots\otimes \nabla^{y_t}\phi_u). \tag{2-144}$$

We then derive that $(Const)_1 = (Const)_2 = 0$ by picking out the sublinear combination in (2-136) that consists of complete contractions with two factors ∇Y, $\nabla\omega_2$ — each factor differentiated only once. □

3. The proof of Proposition 1.1 in the special cases

3A. ***The direct proof of Proposition 1.1 (in Case II) in the special cases.*** We now prove Proposition 1.1 directly in the special subcases of Case II. We recall that the settings of the special subcases of Proposition 1.1 in Case II are as follows: In Subcase IIA for each μ-tensor field (in (1-7)) of maximal refined double character, $C^{l,i_1\dots i_\mu}_g$ there is a unique factor in the form $T = \nabla^{(m)}R_{ijkl}$ for which two internal indices are free, and each derivative index is either free or contracting against a factor $\nabla\phi_h$. For Subcase IIB there is a unique factor in the form $T = \nabla^{(m)}R_{ijkl}$ for which one internal index is free, and each derivative index is either free or contracting against a factor $\nabla\phi_h$. In both Subcases IIA, IIB there is at least one free derivative index in the factor T.

Moreover, in both Subcases IIA and IIB, all *other* factors in one of the forms $\nabla^{(m)}R_{ijkl}$, $S_*\nabla^{(\nu)}R_{ijkl}$, $\nabla^{(p)}\Omega_h$ in $C^{l,i_1\dots i_\mu}_g$ are either in the form S_*R_{ijkl} or $\nabla^{(2)}\Omega_h$, or they are in the form $\nabla^{(m)}R_{ijkl}$, where all the m derivative indices contract against factors $\nabla\phi_h$.[72]

In order to prove Proposition 1.1 directly in the special subcases of Subcases IIA, IIB we will rely on a new lemma. It deals with two different settings, which we will label Setting A and Setting B below.

In Setting A, we let

$$\sum_{l\in\bar{L}} a_l C^{l,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

stand for a linear combination of μ-tensor fields with one factor $\nabla^{(m)}R_{ijkl}$ containing $\alpha \geq 2$ free indices, distributed according to the pattern $\nabla^{(m)}_{(free)\dots(free)}R_{(free)j(free)l}$, and all other factors being in one of the forms R_{ijkl}, S_*R_{ijkl}, $\nabla^{(2)}\Omega_h$. (In particular they have no removable indices.)

In Setting B we let

$$\sum_{l\in\bar{L}} a_l C^{l,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u)$$

stand for a linear combination of μ-tensor fields with one factor $\nabla^{(m)}R_{ijkl}$ containing $\alpha \geq 2$ free indices, distributed according to the pattern $\nabla^{(m)}_{(free)\dots(free)}R_{(free)j(free)l}$, and all *but one* of the other factors being in one of the forms R_{ijkl}, S_*R_{ijkl}, $\nabla^{(2)}\Omega_h$;

[72] For the rest of this subsection, we will slightly abuse notation and *not* write out the derivative indices that contract against factors $\nabla\phi_h$ — we will thus refer to factors R_{ijkl}, setting $m = 0$.

one of the other factors (which we label T') will be in the form ∇R_{ijkl}, $S_*\nabla R_{ijkl}$, $\nabla^{(3)}\Omega_h$. We will call this other factor "the factor with the extra derivative". Moreover, in Setting B we impose the additional restriction that if both the indices ${}_j, {}_l$ in the factor $\nabla^{(m)}_{(free)\dots(free)} R_{(free)\,j\,(free)l}$ contract against the same other factor T', then either T' is *not* the factor with the extra derivative, or if it is, then T' is in the form $\nabla_s R_{abcd}$, and furthermore the indices ${}_j, {}_l$ contract against the indices ${}_b, {}_c$ *and* we assume that the indices ${}_s, {}_a, {}_c$ are symmetrized over.[73]

Lemma 3.1. *Let $\sum_{l\in\bar{L}} a_l C^{l,i_1\dots i_\mu}_g$ be a linear combination of μ-tensor fields as described above. We assume the following special case of (1-7),*

$$\sum_{l\in\bar{L}\cup L'} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\mu} C^{l,i_1\dots i_\mu}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u) + \sum_{h\in H} a_h X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\beta} C^{h,i_1\dots i_\beta}_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u) = \sum_{j\in J} a_j C^j_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u), \tag{3-1}$$

holds; here, in both Cases A and B the terms indexed in $\bar{L}$ will be as described above; the μ-tensor fields indexed in L' will have fewer than α free indices in any given factor of the form $\nabla^{(m)}R_{ijkl}$. The tensor fields indexed in H each have rank $>\mu$ and also each of them has fewer than α free indices in any given factor of the form $\nabla^{(m)}R_{ijkl}$. Finally, the terms indexed in J are simply subsequent to $\vec{\kappa}_{\mathrm{simp}}$.

We claim that

$$\sum_{l\in\bar{L}} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon\dots\nabla_{i_\mu}\upsilon = 0. \tag{3-2}$$

We will prove this lemma shortly. Let us now prove that the above lemma directly implies Proposition 1.1 in the special Subcases IIA (directly) and IIB (after some manipulation).

Lemma 3.1 implies Proposition 1.1 in the special subcases of Case II. We first start with Subcase IIA: Consider the sublinear combination of μ-tensor fields of maximal refined double character in (1-7). Denote their index set by $L_{Max}\subset L$. Recall that since we are considering the subcase where (1-7) falls under the special case of Proposition 1.1 in Case IIA, it follows that for each $C^{l,i_1\dots i_\mu}_g$ there is a unique factor in the form $\nabla^{(m)}R_{ijkl}$ for which two internal indices are free, and each derivative index is either free or contracting against a factor $\nabla\phi_h$; denote by $M+2$ the number of free indices in that factor.[74]

[73] To put it in other words, in that case the two factors T, T' contract according to the pattern $\nabla^{(m)}_{(free)\dots(free)} R_{(free)\,j\,(free)l}\nabla_{(s}R_a{}^{jk}{}_{d)}$, where the indices ${}_s, {}_a, {}_d$ are symmetrized over.

[74] So we set $\alpha = M+2$.

By weight considerations (since we are in a special subcase of Proposition 1.1 in Case IIA), any tensor field of rank $> \mu$ in (1-7) must have strictly fewer than $M+2$ free indices in any given factor $\nabla^{(m)} R_{ijkl}$. Therefore in Subcase IA, (1-7) is of the form (3-1), with $L_{Max} \subset \bar{L}$. Therefore, we apply Lemma 3.1 to (1-7) and pick out the sublinear combination of terms with a refined double character $Doub(\vec{L}^z)$, $z \in Z'_{Max}$,[75] we thus obtain a *new* true equation, since (3-2) holds formally, and the double character is invariant under the formal permutations of indices that make (3-2) formally zero. This proves our claim in Subcase IIA.

Now we deal with Subcase IIB: We consider the μ-tensor fields of maximal refined double character in (1-7). By definition (since we now fall under a special case), they will each have a factor in the form $\nabla^{(m)}_{(free)\dots(free)} R_{(free)jkl}$, with a total of $M+1>1$ free indices.[76] Each of the other factors will be in the form R_{ijkl} or be simple factors in the form $S_* R_{ijkl}$, or in the form $\nabla^{(2)}\Omega_h$.

We denote by $\bar{L} \subset L$ the index set of μ-tensor fields with $M+1$ free indices in a factor $\nabla^{(m)} R_{ijkl}$. It follows by weight considerations that the factor in question will be unique for each $C^{l,i_1\dots i_\mu}_g$, $l \in \bar{L}$. We then start out with some explicit manipulation of the terms indexed in $\bar{L}$:

We will prove that there exists a linear combination of $(\mu+1)$-tensor fields, $\sum_{h\in H} a_h C^{h,i_1\dots i_{\mu+1}}_g$, as allowed in the statement of Proposition 1.1, so that

$$\sum_{l\in\bar{L}} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon = \sum_{h\in H} a_h X \,\mathrm{div}_{i_{\mu+1}} C^{h,i_1\dots i_{\mu+1}}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon + \sum_{l\in\bar{L}_{new}} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon \sum_{j\in J} a_j C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon. \tag{3-3}$$

Here the μ-tensor fields indexed in $\bar{L}_{new}$ have a factor $T = \nabla^{(M-1)}_{(free)\dots(free)} R_{(free)j(free)l}$, and *one other factor* T' has an extra derivative (meaning that T' is either in the form ∇R_{ijkl} or $S_*\nabla R_{ijkl}$, or $\nabla^{(3)}\Omega_h$). Moreover if both indices ${}_j, {}_l$ in T contract against indices ${}^j, {}^l$ in the same factor T'' and at least one of ${}^j, {}^l$ is removable, then $T' \neq T''$. Clearly, (3-3) in conjunction with Lemma 3.1 implies Proposition 1.1 in the "special cases" of Case II. So, matters are reduced to showing (3-3) (and then deriving Lemma 3.1).

Proof of (3-3). Apply the second Bianchi identity to the factor T to move one of the derivative free indices into the position ${}_k, {}_l$ in the factor $\nabla^{(M-1)}_{(free)\dots(free)} R_{(free)j(free)l}$. Thus, we derive that modulo terms of length $\geq \sigma+u+1$,

$$C^{l,i_1\dots i_\mu}_g = -C^{l,1,i_1\dots,i_\mu}_g + C^{l,2,i_1\dots,i_\mu}_g,$$

[75] Recall that $\vec{L}^z$, $z \in Z'_{Max}$, is the collection of maximal refined double characters that Proposition 1.1 deals with.

[76] So, we set $\alpha = M+1$.

where the partial contractions $C^{l,1,i_1\dots i_\mu}_g$ and $C^{l,2,i_1\dots i_\mu}_g$ have the factor T replaced by a factor in the form

$$\nabla^{(m)}_{k(free)\dots(free)}R_{(free)j(free)l} \quad \text{and} \quad \nabla^{(m)}_{l(free)\dots(free)}R_{(free)jk(free)},$$

respectively. We then erase the indices ${}_k, {}_l$ in these two factors (thus creating new tensor fields $C^{l,1,i_1\dots,i_\mu i_{\mu+1}}_g$ and $C^{l,2,i_1\dots,i_\mu i_{\mu+1}}_g$) by creating a free index ${}_{i_{\mu+1}}$, and subtract the $X\,\mathrm{div}_{i_{\mu+1}}[\dots]$ of the corresponding $(\mu+1)$-tensor field. We then derive

$$C^{l,1,i_1\dots,i_\mu}_g = X\,\mathrm{div}_{i_{\mu+1}}\,C^{l,1,i_1\dots i_{\mu+1}}_g + \sum_{l\in L_{new}} C^{l,i_1\dots i_\mu}_g, \tag{3-4}$$

where all the fields indexed in L_{new} satisfy the required property of Lemma 3.1, except for that one could have both indices ${}_j, {}_l$ in the factor $\nabla^{(M-1)}_{(free)\dots(free)}R_{(free)j(free)l}$ contracting against indices ${}^j, {}^l$ in a factor T' which has an additional derivative index. If $C^{l,i_1\dots i_\mu}_g$, $l\in L_{new}$, is not in the form allowed in the claim of Lemma 3.1, then (after possibly applying the second Bianchi identity and possibly introducing simply subsequent complete contractions) we may arrange that one of the indices ${}^j, {}^l$ is a derivative index.

In that case we construct another $(\mu+1)$-tensor field by erasing the derivative index ^{j} or ^{l} and making the index ${}_j$ or ${}_l$ in a free index ${}_{i_{\mu+1}}$. Then, subtracting the corresponding $X\,\mathrm{div}_{i_{\mu+1}}$ of this new $(\mu+1)$-tensor field, we derive our claim. □

Therefore, matters are reduced to proving Lemma 3.1.

Proof of Lemma 3.1. Let us start with some notational conventions.

Recall the first variation law of the curvature tensor under variations by a symmetric 2-tensor by v_{ij}: For any complete or partial contraction $T(g_{ij})$ (which is a function of the metric g_{ij}), we define

$$Image^1_{v_{ij}} = \frac{d}{dt}\Big|_{t=0}\big[T(g_{ij}+tv_{ij})\big].$$

(We write $Image^1_{v_{ij}}[\dots]$ or $Image^1_{v_{ab}}[\dots]$ below to stress that we are varying by a 2-tensor, rather than just by a scalar.)

We consider the equation $Image^1_{v_{ij}}[L_g]=0$ (which corresponds to the first *metric* variation of our lemma hypothesis (that is, of (1-7)). This equation holds modulo complete contractions with at least $\sigma+u+1$ factors.

Thus, we derive a new local equation,

$$\begin{aligned}\sum_{l\in L_\mu} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_\mu}\,Image^1_{v_{ab}}[C^{l,i_1\dots i_\mu}_g]\\ + \sum_{l\in L\setminus L_\mu} a_l X\,\mathrm{div}_{i_1}\dots X\,\mathrm{div}_{i_a}\,Image^1_{v_{ab}}[C^{l,i_1\dots i_a}_g] = \sum_{j\in J} a_j Image^1_{v_{ab}}[C^j_g],\end{aligned} \tag{3-5}$$

which holds modulo terms of length $\geq \sigma+u+1$.

Now, we wish to pass from the local equation above to an integral equation, and then to apply the *silly divergence formula* from [A 2009] to that integral equation (thus deriving a *new* local equation).

In order to do this, we start by introducing some more notation: Let us write out

$$Image^1_{v_{ab}}[C^{l,i_1\dots i_\mu}_g] = \sum_{t\in T^l} a_t C^{t,i_1\dots i_a}_g,$$

where each $C^{t,i_1\dots i_a}_g$ is in the form

$$pcontr\big(\nabla^{(A+2)}_{r_1\dots r_{A+2}} v_{ab}\otimes\nabla^{(m_1)}R_{ijkl}\otimes\cdots\otimes\nabla^{(m_{\sigma-1})}R \otimes\nabla^{(b_1)}\Omega_1\otimes\cdots\otimes\nabla^{(b_p)}\Omega_p\otimes\nabla\phi_1\otimes\cdots\otimes\nabla\phi_u\big). \tag{3-6}$$

For our next technical tool we introduce some notation: For each tensor field $C^{l,i_1\dots i_a}_g$ in the form above, we denote by C^l_g the complete contraction that arises by hitting each factor T_i ($i=1,2,3$) by m derivative indices $\nabla^{u_1\dots u_m}$, where $_{u_1},\dots,{}_{u_m}$ are the free indices that belong to T_i in $C^{l,i_1\dots i_a}_g$ (thus we obtain a factor with m internal contraction, each involving a derivative index). Notice there is a one-to-one correspondence between the tensor fields and the complete contractions we are constructing. We can then easily observe that there are two linear combinations

$$\sum_{r\in R_1} a_r C^r_g(\Omega_1,\dots\Omega_p,\phi_1,\dots,\phi_u),$$
$$\sum_{r\in R_2} a_r C^r_g(\Omega_1,\dots\Omega_p,\phi_1,\dots,\phi_u),$$

where each $C^r_g, r\in R_1$ has at least $\sigma+u+1$ factors, while each $C^r_g, r\in R_2$ has $\sigma+u$ factors but at least one factor $\nabla^{(p)}\phi_h\neq\Delta\phi_h$ with $p\geq 2$, so that for any compact orientable (M,g),

$$\int_M \sum_{l\in L} a_l\sum_{t\in T^l} a_t C^{t,*}_g(v_{ab}) + \sum_{r\in R_1} a_r C^r_g(v_{ab}) + \sum_{r\in R_2} a_r C^r_g(v_{ab})\,dV_g = 0 \tag{3-7}$$

(denote the integrand of the above by $Z_g(v_{ab})$). Here again each C^j_g has $\sigma+u$ factors and all factors $\nabla\phi_h$ have only one derivative but its simple character is subsequent to $\vec{\kappa}$. We call this technique (of going from the local equation (3-5) to the integral equation (3-7)) the "'inverse integration by parts".

Now, we derive a "silly divergence formula" from the above by performing integrations by parts with respect to the factor $\nabla^{(B)}v_{ab}$ (until we are left with a factor v_{ab} — without derivatives). This produces a new local equation which we denote by $silly[Z_g(v_{ab})]=0$. We will be using this equation in our derivation of Lemma 3.1.

Now, for each $C^{l,i_1\dots i_\mu}_g$, $l\in\bar{L}$, we consider the factor

$$T=\nabla^{(M)}_{(free)\dots(free)}R_{(free)j(free)l}$$

with the $M+2$ free indices. We define T^j to be the factor in $C^{l,i_1\dots i_\mu}_g$ that contracts against the index ${}_j$ in T and by T^l to be the factor in $C^{l,i_1\dots i_\mu}_g$ that contracts against the index ${}_l$ in T. We define $\bar{L}_{same}\subset\bar{L}$ to be the index set of tensor fields for which $T^j=T^l$; we define $\bar{L}_{not.same}\subset\bar{L}$ to be the index set of tensor fields for which $T^j\neq T^l$. We will then prove (3-2) separately for the two sublinear combinations indexed in $\bar{L}_{same}$, $\bar{L}_{not.same}$.

Proof of (3-2) for the index set $\bar{L}_{same}$. We first prove our claim for $\sigma>3$ and then note how to prove it when $\sigma=3$.

Consider $silly[L_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,\upsilon_{ab})]=0$. Pick out the sublinear combination $silly_+[L_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,\upsilon_{ab})]=0$ with $\mu-M-2$ internal contractions, and with the indices in the factor υ_{ab} contracting against a factor T' which either has *no* extra derivative indices, or if it does, then the contraction is according to the pattern $\upsilon^{ab}\otimes\nabla_s R_{ajbl}$; we also require that the two factors T'', T''' with an extra $M+2$ extra derivatives each. This sublinear combination must vanish separately, hence we derive

$$silly_+[Z_g(\Omega_1,\dots,\Omega_p,\phi_1,\dots,\phi_u,\upsilon_{ab})]=0. \tag{3-8}$$

We also observe that this sublinear combination can only arise (in the process of passing from the equation $L_g=0$ to deriving $silly_+[Z_g(\upsilon_{ab})]=0$) by replacing the factor $\nabla^{(M)}_{(free)\dots(free)}R_{(free)j(free)l}$ by $\nabla^{(M)}_{(free)\dots(free)}\upsilon_{jl}$ and then (in the inverse integration by parts) replacing all μ free indices by internal contractions,[77] and finally integrating by parts the $M+2$ pairs of derivative indices (∇^a,∇_a) and forcing all upper indices hit a factor $T''\neq T'$ and the lower indices to hit a factor $T'''\neq T'$, $T'''\neq T''$.[78]

Thus, we can prove our claim by starting from (3-8) and applying Sub_υ $\mu-M-2$ times,[79] just applying the eraser to the extra $M+2$ pairs of contracting derivatives,[80] and then replacing the factor υ_{ab} by $\nabla^{(M)}_{r_1\dots r_M}R_{iajb}\nabla^{r_1}\upsilon\dots\nabla^{r_M}\upsilon\nabla^a\upsilon\nabla^b\upsilon$. Finally we just divide by the combinatorial constant $\binom{\sigma-3}{2}$.

Let us now consider the case $\sigma=3$: In this case the terms of maximal refined double character can only arise in Subcase IIA,[81] and can only be in one of the

[77] Thus the factor $\nabla^{(M)}_{(free)\dots(free)}\upsilon_{jl}$ gets replaced by $\Delta^{M+2}\upsilon_{ij}$.

[78] The fact that $\sigma>3$ ensures the existence of two such factors.

[79] See the Appendix in [A 2012] for this operation, and just set $\omega=\upsilon$.

[80] This can be done by repeating the proof of the "Eraser" lemma in the Appendix in [A 2012].

[81] This follows by the symmetry of the indices ${}_s, {}_a, {}_d$ in any factor $\nabla_s R_{abcd}$ as discussed above.

forms

$$\nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l} \otimes R^{ijkl} \otimes \nabla^{(2)}_{ik} \Omega_1,$$
$$\nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l} \otimes R^{ijkl} \otimes R_{(free)i(free)k}.$$

Thus, in that case we define $silly_+[Z_g(\upsilon_{ab})]$ to stand for the terms

$$\upsilon_{jl} \otimes \nabla^{(M+2)}_{t_1\dots t_{M+2}} R^{ijkl} \otimes \nabla^{(M+4)}_{t_1\dots t_{M+4}ik} \Omega_1,$$
$$\upsilon_{jl} \otimes \nabla^{(M+2)}_{t_1\dots t_{M+2}} R^{ijkl} (\nabla^{(M+2)})^{t_1\dots t_{M+2}} \otimes R_{(free)i(free)k},$$

respectively, and then repeat the argument above. □

Proof of (3-2) for the index set $\bar{L}_{not.same}$. We prove our claim in steps: We first denote by $\bar{L}^{**}_{not.same} \subset \bar{L}_{not.same}$ the index set of tensor fields in $\bar{L}_{not.same}$ for which both indices ${}_j, {}_l$ in the factor $T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$ contract against special indices in factors T^j, T^l of the form S_*R_{ijkl}. We will first prove that

$$\sum_{l \in \bar{L}^{**}_{not.same}} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon = \sum_{l \in L'} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon. \tag{3-9}$$

Here the terms in the right-hand side have all the features of the terms in $\bar{L}_{not.same}$, but in addition at most one of the indices in the factor $T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$ contract against a special index in a factor of the form S_*R_{ijkl}. Thus, if we can prove (3-9), we are reduced to proving our claim under the additional assumption that $\bar{L}^{**}_{not.same} = \varnothing$.

For our next claim, we denote by $\bar{L}^{*}_{not.same} \subset \bar{L}_{not.same}$ the index set of tensor fields in $\bar{L}_{not.same}$ for which one of the indices ${}_j, {}_l$ in the factor

$$T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$$

contracts against a special index in factors T^j, T^l of the form S_*R_{ijkl}.

We will then prove that

$$\sum_{l \in \bar{L}^{*}_{not.same}} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon = \sum_{l \in L''} a_l C^{l,i_1\dots i_\mu}_g \nabla_{i_1}\upsilon \dots \nabla_{i_\mu}\upsilon. \tag{3-10}$$

Here the terms in the right-hand side have all the features of the terms in $\bar{L}_{not.same}$, but in addition none of the indices in the factor $T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$ contract against a special index in a factors of the form S_*R_{ijkl}. Thus, if we can prove (3-9), we are reduced to proving our claim under the further assumption that for each $C^{l,i_1\dots i_\mu}_g$, $l \in \bar{L}$, the two indices ${}_j, {}_l$ in the factor $T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$ contract against two different factors and none of the indices ${}^j, {}^l$ are special indices in a factor of the form S_*R_{ijkl}.

In our third step, we prove (3-2) under this additional assumption. We will indicate in the end how this proof can be easily modified to derive the first two steps.

For each $l \in \overline{L}_{not.same}$, let us denote by $link(l)$ the number of particular contractions between the factors T^j, T^l in the tensor fields $C^{l,i_1\dots i_\mu}_g$. (Note that by weight considerations, $0 \le link(l) \le 3$.) Let B be the maximum value of $link(l)$, $l \in \overline{L}_{not.same}$, and denote by $\overline{L}^B_{not.same} \subset \overline{L}_{not.same}$ the corresponding index set. We will then prove our claim for the tensor fields indexed in $\overline{L}^B_{not.same}$. By repeating this step at most four times, we will derive our third claim.

Consider $silly[L_g(\Omega_1, \dots, \Omega_p, \phi_1, \dots, \phi_u, \upsilon_{ab})] = 0$. Pick out the sublinear combination $silly_*[L_g(\Omega_1, \dots, \Omega_p, \phi_1, \dots, \phi_u, \upsilon_{ab})] = 0$ with $\mu - M - 2$ internal contractions, and with extra $M+2$ derivatives on the factors T^j, T^l against which the two indices of the factor υ_{ab} contract, *and* with $M+2+B$ particular contractions between the factors T^j, T^l. This sublinear combination must vanish separately,

$$silly_*[L_g(\Omega_1, \dots, \Omega_p, \phi_1, \dots, \phi_u, \upsilon_{ab})] = 0.$$

Moreover, we observe by following the "inverse integration by parts" and the silly divergence formula obtained from $\int_{M^n} Z_g(\upsilon_{ab})\, dV_g = 0$, that the left-hand side of the above can be described as follows:

For each $C^{l,i_1\dots i_\mu}_g$, $l \in \overline{L}^B_{not.same}$, we denote by $\tilde{C}^l_g(\upsilon_{ab})$ the complete contraction that arises by replacing the factor $T = \nabla^{(M)}_{(free)\dots(free)} R_{(free)j(free)l}$ by $\nabla^{(M+2)}_{(free)\dots(free)} \upsilon_{jl}$, and then replacing each free index that *does not* belong to the factor T by an internal contraction. We then denote by $\hat{C}^l_g(\upsilon_{ab})$ the complete contraction that arises from $\tilde{C}^l_g(\upsilon_{ab})$ by hitting the factor T^j (against which the index $_j$ in υ_{jl} contracts) by $(M+2)$ derivative indices $\nabla_{t_1}, \dots, \nabla_{t_{M+2}}$ and hitting the factor T^l (against which the index $_l$ in υ_{jl} contracts) by derivatives $\nabla^{t_1}, \dots, \nabla^{t_{M+2}}$.[82] It follows that

$$(0 =)\ silly_*[L_g(\Omega_1, \dots, \Omega_p, \phi_1, \dots, \phi_u, \upsilon_{ab})] = \sum_{\overline{L}^B_{not.same}} a_l 2^{M+1} [\hat{C}^l_g(\upsilon_{ab})].$$

Now, to derive our claim, we introduce a formal operation $\mathrm{Op}[\dots]$ which acts on the terms above by applying Sub_υ to each of the $\mu - M - 2$ internal contractions,[83] erasing $M+2$ particular contractions between the factors T^j, T^l and then replacing the factor υ_{jl} by $\nabla^{(m)}_{r_1\dots r_M} R_{ijkl} \nabla^{r_1}\upsilon \dots \nabla^{r_M}\upsilon \nabla^i \upsilon \nabla^k \upsilon$. This operation produces a new true equation; after we divide this new true equation by 2^{M+1}, we derive our claim. □

Note on the derivation of (3-9), (3-10). The equations can be derived by a straightforward modification of the ideas above: The only extra feature we add is that in

[82]These derivatives contract against the indices $\nabla_{t_1}, \dots, \nabla_{t_{M+2}}$ that have hit T^j.

[83]See the Appendix of [A 2012] for the definition of this operation.

the silly divergence formula we must pick out the terms for which (both/one of the) indices ${}_j, {}_l$ in v_{jl} contract against a special index in a factor $S_*\nabla^{(M+2)}R_{abcd}\nabla^a\tilde{\phi}_h$. This linear combination will vanish, modulo terms where one/none of the indices ${}_j, {}_l$ in v_{jl} contract against a special index in the factor S_*R_{ijkl}: This follows by the same argument that is used in [A 2010] to derive that Lemma 3.1 in [A 2010] implies Proposition 1.1 in Case I: We first replace the factor v_{jl} by an expression $y_{(j}y_{l)}$. We then just replace both/one of the expressions $\nabla_i\tilde{\phi}_h$, y_j by g_{ij} and apply Ricto Ω twice/once.[84] The only terms that survive this true equation are the ones indexed in $\bar{L}_{not.same}$, for which the expression(s) $S_*\nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}\nabla^i\tilde{\phi}_h\nabla^k y$ are replaced by $\nabla^{(\nu+2)}_{r_1\dots r_\nu jl}Y_f$. We then proceed as above, deriving that the sublinear combination of terms indexed in $\bar{L}_{not.same}$ must vanish, *after* we replace two/one expressions $S_*\nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}\nabla^i\tilde{\phi}_h\nabla^k y$ by $\nabla^{(\nu+2)}_{r_1\dots r_\nu jl}Y_f$. Then, repeating the permutations applied to any factors $\nabla^{(\nu+2)}_{r_1\dots r_\nu jl}Y_f$, to $S_*\nabla^{(\nu)}_{r_1\dots r_\nu}R_{ijkl}\nabla^i\tilde{\phi}_h\nabla^k y$ we derive our claim.

3B. *The remaining cases of Proposition 1.1 in Case III.* We recall that there are remaining cases only when $\sigma = 3$. In that case we have the remaining cases when $p = 3$ and $n - 2u - 2\mu \leq 2$, or when $p = 2$, $\sigma_2 = 1$ and $n = 2u + 2\mu$.

The case $p = 3$. Let us start with the subcase $n - 2u - 2\mu = 0$. In this case, all tensor fields in (1-7) will be in the form

$$pcontr\big(\nabla^{(A)}_{i_1\dots i_a j_1\dots j_b}\Omega_1 \otimes \nabla^{(B)}_{i_{a+1}\dots i_{a+a'} j_{b+1}\dots j_{b+b'}}\Omega_2 \otimes \nabla^{(C)}_{i_{a+a'+1}\dots i_{a+a'+a''} j_{b+b'+1}\dots j_{b+b'+b''}}\Omega_3 \otimes \nabla^{j_{x_1}}\phi_1 \cdots \otimes \nabla^{x_{j+j'+j''}}\phi_u\big), \tag{3-11}$$

where we make the following conventions: Each of the indices ${}_{i_f}$ is free; also, each of the indices ${}_{j_f}$ contracts against some factor $\nabla\phi_h$, and also $A, B, C \geq 2$.

Thus, we observe that in this subcase μ is also the *maximum* rank among the tensor fields appearing in (1-7). Now, assume that the μ-tensor fields in (1-7) of *maximal refined double character* have $a = \alpha$, $a' = \alpha'$, $a'' = \alpha''$. With no loss of generality (only up to renaming the factors $\Omega_1, \Omega_2, \Omega_3, \phi_1, \dots, \phi_u$) we may assume that $\alpha \geq \alpha' \geq \alpha''$ and that only the functions $\nabla\phi_1, \dots, \nabla\phi_{u_1}$ contract against $\nabla^{(A)}\Omega_1$ in $\vec{\kappa}_{\text{simp}}$. We will then show that the coefficient $a_{\alpha,\alpha',\alpha''}$ of this tensor field must be zero. This will prove Proposition 1.1 in this subcase.

We prove that $a_{\alpha,\alpha',\alpha''} = 0$ by considering the global equation $\int Z_g\, dV_g = 0$ and considering the silly divergence formula $silly[Z_g] = 0$. We then consider the sublinear combination $silly_+[Z_g]$ consisting of terms with α', α'' internal contractions in the factors $\nabla^{(D)}\Omega_2$, $\nabla^{(E)}\Omega_3$, with α particular contractions between those factors and with all factors $\nabla\phi_h$ that contracted against $\nabla^{(A)}\Omega_1$ in $\vec{\kappa}_{\text{simp}}$ being replaced by $\Delta\phi_h$, while all factors $\nabla\phi_h$ that contracted against $\nabla^{(B)}\Omega_2$, $\nabla^{(C)}\Omega_3$ still do so.

[84] Recall that this operation has been defined in the Appendix in [A 2012] and produces a true equation.

We easily observe that $silly_+[Z_g]=0$, and furthermore $silly_+[Z_g]$ consists of the complete contraction

$$contr\big(\Omega_1\otimes\nabla^{f_1\dots f_\alpha}{}_{j_{b+1}\dots j_{b+b'}}\Delta^{\alpha'}\Omega_2\otimes\nabla_{f_1\dots f_\alpha j_{b+b'+1}\dots j_{b+b'+b''}}\Delta^{\alpha''}\Omega_3 \otimes\Delta\phi_1\dots\nabla^{j_{x_{b+b'+b''}}}\phi_u\big) \tag{3-12}$$

times the constant $(-1)^{u_1}2^\alpha a_{\alpha,\alpha',\alpha''}$. Thus, we derive that $a_{\alpha,\alpha',\alpha''}=0$.

The second subcase. We now consider the setting where $\sigma=p=3$, $n-2u-2\mu=2$. In this setting, the maximum rank of the tensor fields appearing in (1-7) is $\mu+1$. In this case, all $(\mu+1)$-tensor fields in (1-7) will be in the form (3-11) (with $\alpha+\alpha'+\alpha''=\mu+1$, while all the μ-tensor fields will be in the form (3-11) but with $\alpha+\alpha'+\alpha''=\mu$, *and* with one particular contraction ${}_c,{}^c$ between two of the factors $\nabla^{(A)}\Omega_1$, $\nabla^{(B)}\Omega_2$, $\nabla^{(C)}\Omega_3$.

Now, if both the indices ${}_c,{}^c$ described above are removable, we can explicitly express $C^{l,i_1\dots i_\mu}_g$ as an X div of an acceptable $(\mu+1)$-tensor field. Therefore, we are reduced to showing our claim in this setting where for each μ-tensor field in (1-7) at least one of the indices ${}_c,{}^c$ is not removable. Now, let $z\in Z_{Max}$ stand for one of the index sets for which the sublinear combination $\sum_{l\in L^z}a_lC^{l,i_1\dots i_\mu}_g$ in (1-7) indexes tensor fields of maximal refined double character. We assume with no loss of generality that for each $l\in L^z$ the factors $\nabla^{(A)}\Omega_1$, $\nabla^{(B)}\Omega_2$, $\nabla^{(C)}\Omega_3$ have $\alpha\ge\alpha'\ge\alpha''$ free indices respectively.[85] Therefore, the tensor fields indexed in L^z can be in one of the two forms

$$pcontr\big(\nabla^c\nabla^{(A)}_{i_1\dots i_\alpha j_1\dots j_b}\Omega_1\otimes\nabla^{(B)}_{i_{\alpha+1}\dots i_{\alpha+\alpha'}j_{b+1}\dots j_{b+b'}}\Omega_2 \otimes\nabla^{(2)}_{ci_{\alpha+\alpha'+1}\dots i_{\alpha+\alpha'+\alpha''}j_{b+b'+1}\dots j_{b+b'+b''}}\Omega_3\otimes\nabla^{j_{x_1}}\phi_1\dots\otimes\nabla^{x_{j+j'+j''}}\phi_u\big), \tag{3-13}$$

or

$$pcontr\big(\nabla^{(A)}_{i_1\dots i_\alpha j_1\dots j_b}\Omega_1\otimes\nabla^c\nabla^{(B)}_{i_{\alpha+1}\dots i_{\alpha+\alpha'}j_{b+1}\dots j_{b+b'}}\Omega_2 \otimes\nabla^{(2)}_{ci_{\alpha+\alpha'+1}\dots i_{\alpha+\alpha'+\alpha''}j_{b+b'+1}\dots j_{b+b'+b''}}\Omega_3\otimes\nabla^{j_{x_1}}\phi_1\dots\otimes\nabla^{x_{j+j'+j''}}\phi_u\big), \tag{3-14}$$

where $A,B\ge 3$.

Now, by "manually subtracting" X divs from these μ-tensor fields, we can assume without loss of generality that the tensor fields indexed in our chosen L^z are in the form (3-14).

With that extra assumption, we can show that the coefficient of the tensor field (3-14) is zero. We see this by considering the (global) equation $\int_M Z_g\,dV_g=0$ and using the silly divergence formula $silly[Z_g]=0$ (which arises by integrations by parts with respect to the factor $\nabla^{(A)}\Omega_1$). Picking out the sublinear combination

[85] Recall that by our hypothesis $\alpha'\ge 2$.

$silly_+[Z_g]$ which consists of the complete contraction

$$contr\big(\Omega_1 \otimes \nabla^{cf_1\dots f_\alpha}_{j_{b+1}\dots j_{b+b'}} \Delta^{\alpha'}\Omega_2 \otimes \nabla_{cf_1\dots f_\alpha j_{b+b'+1}\dots j_{b+b'+b''}} \Delta^{\alpha''}\Omega_3 \otimes \Delta\phi_1 \dots \nabla^{j_{x_{b+b'+b''}}}\phi_u\big) \tag{3-15}$$

(notice that $silly_+[Z_g]=0$), we derive that the coefficient of (3-14) must vanish. Thus, we have shown our claim in this second subcase also. □

The case $p=2$, $\sigma_2=1$. Recall that in this case we fall under the special case when $n=2u+2\mu$. In this setting, we will have that in each index set L^z, $z\in Z'_{Max}$, (see the statement of Lemma 3.5 in [A 2010]) there is a *unique* μ-tensor field of maximal refined double character in (1-7), where the two indices ${}_k, {}_l$ in the factor $S_*\nabla^{(\nu)}R_{ijkl}$ will be contracting against one of the factors $\nabla^{(A)}\Omega_1$, $\nabla^{(B)}\Omega_2$ (without loss of generality we may assume that they are contracting against *different factors*). But now, recall that since we are considering Case A of Lemma 3.5 in [A 2010], one of the factors $\nabla^{(A)}\Omega_1$, $\nabla^{(B)}\Omega_2$ will have at least two free indices. Hence, in at least one of the factors $\nabla^{(A)}\Omega_1$, $\nabla^{(B)}\Omega_2$, the index ${}^k, {}^l$ is removable (meaning that it can be erased, and we will be left with an acceptable tensor field). We denote by $C^{l,i_1\dots i_\mu i_{\mu+1}}_g$ the tensor field that arises from $C^{l,i_1\dots i_\mu}_g$ by erasing the aforementioned ${}^k, {}^l$ and making ${}_k$ or ${}_l$ into a free index, we then observe that

$$C^{l,i_1\dots i_\mu}_g - X\,\mathrm{div}_{i_{\mu+1}} C^{l,i_1\dots i_\mu i_{\mu+1}}_g = 0 \tag{3-16}$$

(modulo complete contractions of length $\geq \sigma+u+1$). This completes the proof of our claim. □

Acknowledgement

This work has absorbed the best part of the author's energy over many years. This research was partially conducted during the period the author served as a Clay Research Fellow, an MSRI postdoctoral fellow, a Clay Liftoff fellow and a Procter Fellow.

The author is immensely indebted to Charles Fefferman for devoting twelve long months to the meticulous proof-reading of the present paper. He also wishes to express his gratitude to the Mathematics Department of Princeton University for its support during his work on this project.

References

[A 2009] S. Alexakis, "On the decomposition of global conformal invariants, I", *Ann. of Math.* (2) **170**:3 (2009), 1241–1306. MR 2011g:53063 Zbl 1190.53028

[A 2010] S. Alexakis, "The decomposition of global conformal invariants, IV: A proposition on local Riemannian invariants", *Adv. Math.* **225**:2 (2010), 515–597. MR 2012b:53049 Zbl 1198.53034

[A 2011] S. Alexakis, "The decomposition of global conformal invariants: some technical proofs, I", *Symmetry Integrability Geom. Methods Appl.* **7** (2011), Paper 019. MR 2804577 Zbl 1218.53038

[A 2012] S. Alexakis, *The decomposition of global conformal invariants*, Annals of Mathematics Studies **182**, Princeton University Press, 2012. MR 2918125 Zbl 06035836

[Bailey et al. 1994a] T. N. Bailey, M. G. Eastwood, and A. R. Gover, "Thomas's structure bundle for conformal, projective and related structures", *Rocky Mountain J. Math.* **24**:4 (1994), 1191–1217. MR 96e:53016 Zbl 0828.53012

[Bailey et al. 1994b] T. N. Bailey, M. G. Eastwood, and C. R. Graham, "Invariant theory for conformal and CR geometry", *Ann. of Math.* (2) **139**:3 (1994), 491–552. MR 95h:53016 Zbl 0814.53017

[Čap and Gover 2002] A. Čap and A. R. Gover, "Tractor calculi for parabolic geometries", *Trans. Amer. Math. Soc.* **354**:4 (2002), 1511–1548. MR 2003j:53033 Zbl 0997.53016

[Čap and Gover 2003] A. Čap and A. R. Gover, "Standard tractors and the conformal ambient metric construction", *Ann. Global Anal. Geom.* **24**:3 (2003), 231–259. MR 2004g:53016 Zbl 1039.53021

[Cartan 1896] E. Cartan, "Sur la reduction a sa forme canonique de la structure d'un groupe de transformations fini et continu", *Amer. J. Math.* **18**:1 (1896), 1–61. Reprinted as pp. 293–355 in his *Oeuvres Complètes*, **Partie 1**, Gauthier-Villars, Paris, 1952. MR 1505696 JFM 27.0288.01

[Deser and Schwimmer 1993] S. Deser and A. Schwimmer, "Geometric classification of conformal anomalies in arbitrary dimensions", *Phys. Lett. B* **309**:3-4 (1993), 279–284. MR 94g:81195

[Fefferman 1976] C. L. Fefferman, "Monge–Ampère equations, the Bergman kernel, and geometry of pseudoconvex domains", *Ann. of Math. (2)* **103**:2 (1976), 395–416. Correction in **104**:2 (1976), 393–394. MR 0407320 (53 #11097a) Zbl 0322.32012

[Fefferman and Graham 1985] C. Fefferman and C. R. Graham, "Conformal invariants", pp. 95–116 in *Élie Cartan et les mathématiques d'aujourd'hui* (Lyon, 1984), Astérisque, 1985. Numéro hors série. MR 87g:53060 Zbl 0602.53007

[Graham and Hirachi 2008] C. R. Graham and K. Hirachi, "Inhomogeneous ambient metrics", pp. 403–420 in *Symmetries and overdetermined systems of partial differential equations* (Minneapolis, MN, 2006), edited by M. Eastwood and W. Miller, Jr., IMA Vol. Math. Appl. **144**, Springer, New York, 2008. MR 2009f:53012 Zbl 1148.53023

[Hirachi 2000] K. Hirachi, "Construction of boundary invariants and the logarithmic singularity of the Bergman kernel", *Ann. of Math.* (2) **151**:1 (2000), 151–191. MR 2001f:32003 Zbl 0954.32002

[Thomas 1934] T. Y. Thomas, *The differential invariants of generalized spaces*, Cambridge University Press, 1934. JFM 60.0363.02

[Weyl 1939] H. Weyl, *The classical groups: Their invariants and representations*, Princeton University Press, 1939. MR 0000255 (1,42c) Zbl 0020.20601

Received December 21, 2009.

SPYROS ALEXAKIS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
40 ST. GEORGE ST
TORONTO M5S 2E4
CANADA

alexakis@math.toronto.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.89

ON DEFORMATION QUANTIZATIONS OF HYPERTORIC VARIETIES

GWYN BELLAMY AND TOSHIRO KUWABARA

Based on a construction by Kashiwara and Rouquier, we present an analogue of the Beilinson–Bernstein localization theorem for hypertoric varieties. In this case, sheaves of differential operators are replaced by sheaves of *W*-algebras. As a special case, our result gives a localization theorem for rational Cherednik algebras associated to cyclic groups.

1. Introduction

Kontsevich [2001] and Polesello and Schapira [2004] have shown that one can construct a stack of "W-algebroids" (or deformation-quantization algebroids) on any symplectic manifold. These stacks of W-algebroids provide a quantization of the sheaf of holomorphic functions on the manifold. In certain cases, these stacks of W-algebroids are the algebroids associated to a sheaf of noncommutative algebras called W-algebras. Locally this is always the case. When the symplectic manifold in question is the Hamiltonian reduction of a space equipped with a genuine sheaf of W-algebras, Kashiwara and Rouquier [2008] have shown that one can define a family of sheaves of W-algebras on the Hamiltonian reduction coming from the sheaf upstairs. This provides a large class of examples of sheaves of W-algebras on nontrivial symplectic manifolds. In this paper we study W-algebras on the simplest class of Hamiltonian reductions, those coming from the action of a torus T on a symplectic vector space V. These spaces $Y(A, \delta)$, where A is a matrix encoding the action of T on V and $\delta \in \mathbb{X}(T)$ is a character of T, are called hypertoric varieties. They were originally studied as hyperkähler manifolds by Bielawski and Dancer [2000]. Examples of hypertoric varieties include the cotangent space of projective n-space and resolutions of cyclic Kleinian singularities. More generally,

Kuwabara was partially supported by Grant-in-Aid for Young Scientists (B) 21740013, and by GCOE "Fostering top leaders in mathematics", Kyoto University. He was also partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST)(2010-0019516).
MSC2010: primary 14A22, 16S80; secondary 18E35.
Keywords: deformation quantization, hypertoric varieties.

the cotangent space of any smooth toric variety can be realized as a dense, open subvariety of the corresponding hypertoric variety.

One can also associate to the data of a reductive group G acting on a symplectic vector space a certain family of noncommutative algebras U_χ, where $\chi \in \mathbb{X}(\mathfrak{g})$ is a character of $\mathfrak{g} = \mathrm{Lie}(G)$, called quantum Hamiltonian reductions. In the case $G = T$ is a torus, these algebras have been extensively studied by Musson and Van den Bergh [1998]. The main goal of this paper is to prove a localization theorem, analogous to the celebrated Beilinson–Bernstein localization theorem [1981], giving an equivalence between the category of finitely generated modules for the quantum Hamiltonian reduction and a certain category of modules for a W-algebra. When the character δ is chosen to lie in the interior C of a G.I.T. chamber in $\mathbb{X}(T)$, the hypertoric variety $Y(A, \delta)$ is a symplectic manifold. Then each character $\chi \in \mathbb{X}(\mathfrak{t})$ gives a sheaf of W-algebras $\mathscr{A}_\chi$ on $Y(A, \delta)$. Associated to $\mathscr{A}_\chi$ is a category of "good" $\mathscr{A}_\chi$-modules, $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ and a subcategory $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ consisting (roughly) of those modules generated by their global section (the reader is referred to section 2 for the precise definition of these categories). Then we have natural localization and global section functors

$$\mathsf{Loc} : \mathsf{U}_\chi\text{-mod} \longrightarrow \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi), \quad \mathsf{Loc}(M) = \mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} M,$$
$$\mathsf{Sec} : \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi) \longrightarrow \mathsf{U}_\chi\text{-mod}, \quad \mathsf{Sec}(\mathscr{M}) = \mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \mathscr{M}).$$

Our main result can be stated as follows.

Theorem 1.1. *Let $\chi \in C_{\mathbb{Q}}$.*

(i) *The functor* Loc *defines an equivalence of categories* $\mathsf{U}_\chi\text{-mod} \xrightarrow{\sim} \underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ *with quasiinverse* Sec.

(ii) *There exists some $\theta \in C \cap \mathbb{X}(T)$ such that the functor* Loc *defines an equivalence of categories* $\mathsf{U}_{\chi+\theta}\text{-mod} \xrightarrow{\sim} \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\chi+\theta})$ *with quasiinverse* Sec.

The theorem shows that localization always gives an equivalence of categories, provided one is sufficiently far away from the G.I.T. walls.

Corollary 1.2. *Let $\chi \in C_{\mathbb{Q}}$. If the global dimension of U_χ is finite then the functor* Loc *defines an equivalence of categories* $\mathsf{U}_\chi\text{-mod} \xrightarrow{\sim} \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ *with quasiinverse* Sec.

A particular class of examples of hypertoric varieties are the minimal resolutions $(\mathbb{C}^2/\mathbb{Z}_m)^\sim$ of the Kleinian singularities of type A. Under mild restrictions on the parameters, the corresponding quantum Hamiltonian reductions are Morita equivalent to the rational Cherednik algebras $\boldsymbol{H_h}$ associated to cyclic groups. Then a corollary of our main result is a localization theorem for these rational Cherednik algebras.

Corollary 1.3. *For $\boldsymbol{h}$ not lying on a G.I.T. wall, the functor* $\mathsf{Loc}(e\cdot(\,\cdot\,))$ *defines an equivalence of categories*

$$H_{\boldsymbol{h}}\text{-mod} \xrightarrow{\sim} \operatorname{Mod}_F^{\text{good}}(\mathscr{A}_{\boldsymbol{h}})$$

with quasiinverse $H_{\boldsymbol{h}}e \otimes_{eH_{\boldsymbol{h}}e} \mathsf{Sec}(\,\cdot\,)$.

We summarize the content of each section. In Section 2 we introduce, following Kashiwara and Rouquier, W-algebras on symplectic manifolds in the equivariant setting. In Section 3 we give a criterion for the W-affinity of a class of W-algebras on those symplectic manifolds that are obtained by Hamiltonian reduction of a vector space acted upon by a reductive group. The W-algebras on hypertoric varieties that we will consider later are a special case of this more general setup. The main result of this section is Theorem 3.3.

Hypertoric varieties are introduced in Section 4 and we show that they possess the correct geometric properties that are required to apply the results of Section 3. Using the results of Musson and Van den Bergh, we prove our main results, Theorem 5.2 and Corollary 5.3. In the final section we consider the special case where the hypertoric variety is the resolution of a Kleinian singularity of type A and the global sections of the sheaf of W-algebras on this resolution can be identified with the spherical subalgebra of the rational Cherednik algebra associated to a cyclic group.

Convention. Throughout, a *variety* will always mean an integral, separated scheme of finite type over $\mathbb{C}$. A nonreduced space will be referred to as a scheme, again assumed to be over $\mathbb{C}$.

2. *W*-algebras

2A. In this section we recall the definition of W-algebras as given in [Kashiwara and Rouquier 2008]. We state results about the existence and "affinity" of W-algebras. Let X be a complex analytic manifold and let $\mathscr{O}_X$ denote the sheaf of regular, holomorphic functions on X. Denote by $\mathscr{D}_X$ the sheaf of differential operators on X with holomorphic coefficients. Denote by $\boldsymbol{k} = \mathbb{C}((\hbar))$ the field of formal Laurent series in $\hbar$ and by $\boldsymbol{k}(0)$ the subring $\mathbb{C}[[\hbar]]$ of formal functions on $\mathbb{C}$. Considering $\boldsymbol{k}$ and $\boldsymbol{k}(0)$ as abelian groups, the corresponding sheaves of locally constant functions on X will be denoted $\boldsymbol{k}_X$ and $\boldsymbol{k}(0)_X$ respectively. Given $m \in \mathbb{Z}$, we define $\mathscr{W}_{T^*\mathbb{C}^n}(m)$ to be the sheaf of formal power series $\sum_{i \geq -m} \hbar^i a_i$, $a_i \in \mathscr{O}_{T^*\mathbb{C}^n}$, on the cotangent bundle $T^*\mathbb{C}^n$ of $\mathbb{C}^n$. Let us fix coordinates $x_1, \dots, x_n$ on $\mathbb{C}^n$ and dual coordinates $\xi_1, \dots, \xi_n$ on $(\mathbb{C}^n)^*$, identifying $T^*\mathbb{C}^n$ with $\mathbb{C}^n \times (\mathbb{C}^n)^*$. Set $\mathscr{W}_{T^*\mathbb{C}^n} = \bigcup_{m \in \mathbb{Z}} \mathscr{W}_{T^*\mathbb{C}^n}(m)$. Then $\mathscr{W}_{T^*\mathbb{C}^n}$ is a sheaf of (noncommutative) $\boldsymbol{k}$-algebras on $T^*\mathbb{C}^n$. Multiplication is defined by

$$a \circ b = \sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} \frac{\hbar^{|\alpha|}}{\alpha!} \partial^\alpha_\xi a \cdot \partial^\alpha_x b, \tag{1}$$

where $|\alpha| = \sum_{i=1}^n \alpha_i$, $\alpha! = \alpha_1! \cdots \alpha_n!$ and $\partial^\alpha_\xi = \partial^{|\alpha|}/(\partial^{\alpha_1}\xi_1 \cdots \partial^{\alpha_n}\xi_n)$. There is a ring homomorphism $\mathscr{D}_{\mathbb{C}^n}(\mathbb{C}^n) \to \mathscr{W}_{T^*\mathbb{C}^n}(T^*\mathbb{C}^n)$ given by $x_i \mapsto x_i$ and $\partial/\partial x_i \mapsto \hbar^{-1}\xi_i$. Note that $\mathscr{W}_{T^*\mathbb{C}^n}(0)$ is a $\boldsymbol{k}(0)$-subalgebra. We denote the symbol map for $\mathscr{W}_{T^*\mathbb{C}^n}$ by

$$\sigma_m : \mathscr{W}_{T^*\mathbb{C}^n}(m) \longrightarrow \mathscr{W}_{T^*\mathbb{C}^n}(m)/\mathscr{W}_{T^*\mathbb{C}^n}(m-1) \simeq \hbar^{-m}\mathscr{O}_{T^*\mathbb{C}^n}.$$

The sheaf $\mathscr{O}_{T^*\mathbb{C}^n}$ is a sheaf of Poisson algebras with Poisson bracket given by

$$\{x_i, x_j\} = \{\xi_i, \xi_j\} = 0, \quad \{\xi_i, x_j\} = \delta_{ij} \quad \text{for all } i, j \in [1, n].$$

One sees from (1) that $\sigma_0(\hbar^{-1}[a, b]) = \{\sigma_0(a), \sigma_0(b)\}$ for all $a, b \in \mathscr{W}_{T^*\mathbb{C}^n}(0)$.

2B. Let us now assume that X is a complex symplectic manifold with holomorphic 2-form ω_X. A map f between open subsets $U \subset X$ and $V \subset Y$ of the symplectic manifolds (X, ω_1) and (Y, ω_2) is said to be a symplectic map if $f^*\omega_2 = \omega_1$. A symplectic map is always locally biholomorphic [Björk 1979, Lemma 5.5.2], therefore by symplectic map we will actually mean a biholomorphic symplectic map. Based on [Kontsevich 2001; Polesello and Schapira 2004], we have:

Definition 2.1. A *W-algebra* on X is a sheaf of $\boldsymbol{k}$-algebras $\mathscr{W}_X$ together with a $\boldsymbol{k}(0)$-subalgebra $\mathscr{W}_X(0)$ such that for each point $x \in X$ there exists an open neighborhood U of x in X, a symplectic map $f : U \longrightarrow V \subset T^*\mathbb{C}^n$ and a $\boldsymbol{k}$-algebra isomorphism $r : f^{-1}(\mathscr{W}_{T^*\mathbb{C}^n}|_V) \xrightarrow{\sim} \mathscr{W}_X|_U$ such that:

(i) The isomorphism r restricts to a $\boldsymbol{k}(0)$-isomorphism

$$f^{-1}(\mathscr{W}_{T^*\mathbb{C}^n}(0)|_V) \xrightarrow{\sim} \mathscr{W}_X(0)|_U.$$

(ii) Setting $\mathscr{W}_X(m) = \hbar^{-m}\mathscr{W}(0)$ for all $m \in \mathbb{Z}$, we have

$$\sigma_0 : \mathscr{W}_X(0) \longrightarrow \mathscr{W}_X(0)/\mathscr{W}_X(-1) \simeq \mathscr{O}_X,$$

and the following diagram commutes:

$$\begin{tikzcd} f^{-1}(\mathscr{W}_{T^*\mathbb{C}^n}(0)|_V) \arrow[r, "r"] \arrow[d, "\sigma_0"'] & \mathscr{W}_X(0)|_U \arrow[d, "\nu_0"] \\ f^{-1}(\mathscr{O}_{T^*\mathbb{C}^n}) \arrow[r, "f^\sharp"'] & \mathscr{O}_X \end{tikzcd}$$

2C. The first statement of property (ii) of Definition 2.1 is actually a consequence of property (i). Next, Definition 2.1(ii) implies that $\sigma_0(\hbar^{-1}[a,b]) = \{\sigma_0(a), \sigma_0(b)\}$ for all $a, b \in \mathcal{W}_X(0)$, where the Poisson bracket on $\mathcal{O}_X$ is the one induced from the symplectic form ω on X.

2D. ***Categories of W-modules.*** Unless explicitly stated, all modules will be left modules. Since $\mathcal{W}_X(0)$ is Noetherian (see [Kashiwara and Rouquier 2008, (2.2.2)]), a $\mathcal{W}_X(0)$-module $\mathcal{M}$ is said to be *coherent* if it is locally finitely generated. For a $\mathcal{W}_X$-module $\mathcal{M}$, a $\mathcal{W}_X(0)$-*lattice* of $\mathcal{M}$ is a $\mathcal{W}_X(0)$-submodule $\mathcal{N}$ of $\mathcal{M}$ such that the natural map $\mathcal{W} \otimes_{\mathcal{W}(0)} \mathcal{N} \to \mathcal{M}$ is an isomorphism. A $\mathcal{W}$-module $\mathcal{M}$ is said to be *good* if for every relatively compact open set U there exists a coherent $\mathcal{W}_X(0)|_U$-lattice for $\mathcal{M}|_U$. We will denote the category of left $\mathcal{W}_X$-modules as $\mathrm{Mod}(\mathcal{W}_X)$ and the full subcategory of good $\mathcal{W}_X$-modules as $\mathrm{Mod}^{\mathrm{good}}(\mathcal{W}_X)$. It is an abelian subcategory. If $\mathcal{M}(0)$ is a $\mathcal{W}_X(0)$-lattice of $\mathcal{M}$, set $\mathcal{M}(m) := \hbar^{-m}\mathcal{M}(0)$.

Lemma 2.2. *Let $\mathcal{M}$ be a coherent $\mathcal{W}_X$-module, equipped with a global $\mathcal{W}_X(0)$-lattice $\mathcal{M}(0)$. Then the filtration $\mathcal{M}(n)$, $n \in \mathbb{Z}$, is exhaustive, Hausdorff and complete; that is,*

(i) $\bigcup_{n\in\mathbb{Z}} \mathcal{M}(n) = \mathcal{M}$,

(ii) $\bigcap_{n\in\mathbb{Z}} \mathcal{M}(n) = 0$,

(iii) $\lim\limits_{-\infty\leftarrow n} \mathcal{M}/\mathcal{M}(n) = \mathcal{M}$.

(Our terminology is chosen to agree with that of [Weibel 1994, §5].)

Proof. The statement (i) is true if $\mathcal{M} = \mathcal{W}_X$. But, by the definition of a lattice, we have

$$\bigcup_{n\in\mathbb{Z}} \mathcal{M}(n) = \bigcup_{n\in\mathbb{N}} \mathcal{W}_X(n) \otimes_{\mathcal{W}_X(0)} \mathcal{M}(0) = \mathcal{W}_X \otimes_{\mathcal{W}_X(0)} \mathcal{M}(0) = \mathcal{M}.$$

Part (ii) follows from [Kashiwara and Rouquier 2008, Lemma 2.11]. Fix some open subset U of X and take a section $(f_n)_{n\in\mathbb{Z}} \in \lim_{-\infty\leftarrow n}(\mathcal{M}/\mathcal{M}(n))(U)$. Then, by part (i), there exists some integer $k > n$ such that the image f_n of f in $(\mathcal{M}/\mathcal{M}(n))(U)$ lies in $(\mathcal{M}(k)/\mathcal{M}(n))(U)$. Now by definition f_n is the image of f_{n-1} in the surjection

$$(\mathcal{M}/\mathcal{M}(n-1))(U) \longrightarrow (\mathcal{M}/\mathcal{M}(n))(U),$$

hence $f_{n-1} \in (\mathcal{M}(k)/\mathcal{M}(n-1))(U)$ too. Thus $(\hbar^{-k} f_n)_{n\in\mathbb{Z}}$ is in

$$\lim_{-\infty\leftarrow n} (\mathcal{M}(0)/\mathcal{M}(n))(U).$$

This implies that we have a surjective morphism

$$\boldsymbol{k}_X \otimes_{\boldsymbol{k}(0)_X} \lim_{-\infty\leftarrow n} \mathcal{M}(0)/\mathcal{M}(n) \longrightarrow \lim_{-\infty\leftarrow n} \mathcal{M}/\mathcal{M}(n).$$

But it follows once again from [Kashiwara and Rouquier 2008, Lemma 2.11] that

$$\mathcal{M} \simeq \boldsymbol{k}_X \otimes_{\boldsymbol{k}(0)_X} \mathcal{M}(0) \simeq \boldsymbol{k}_X \otimes_{\boldsymbol{k}(0)_X} \varprojlim_{-\infty \leftarrow n} \mathcal{M}(0)/\mathcal{M}(n).$$

Thus $\mathcal{M}$ surjects onto $\lim\limits_{-\infty \leftarrow n} \mathcal{M}/\mathcal{M}(n)$. Part (ii) implies this map is also injective. $\square$

2E. *G-equivariance.* Let G be a complex Lie group acting symplectically on X, via $T_g : X \xrightarrow{\sim} X$ for all $g \in G$. We assume that this action is Hamiltonian with moment map $\mu_X : X \to \mathfrak{g}^*$, where $\mathfrak{g}$ is the Lie algebra of G.

Definition 2.3. A G-action on the W-algebra $\mathscr{W}_X$ is a $\boldsymbol{k}_X$-algebra isomorphism $\rho_g : \mathscr{W}_X \xrightarrow{\sim} T_g^{-1}\mathscr{W}_X$ for every $g \in G$ such that $\rho_g(a)$ depends holomorphically on $g \in G$ for each section $a \in \mathscr{W}_X$ and $\rho_{g_1} \circ \rho_{g_2} = \rho_{g_1 g_2}$ for all $g_1, g_2 \in G$.

Definition 2.4. Suppose we have fixed a G-action on $\mathscr{W}_X$. A quasi-G-equivariant $\mathscr{W}_X$-module is a left $\mathscr{W}_X$-module $\mathcal{M}$, together with a $\boldsymbol{k}_X$-module isomorphism

$$\rho_g^{\mathcal{M}} : \mathcal{M} \xrightarrow{\sim} T_g^{-1}\mathcal{M}$$

for every $g \in G$ such that $\rho_g^{\mathcal{M}}(m)$ depends holomorphically on $g \in G$ for each section $m \in \mathcal{M}$, $\rho_g^{\mathcal{M}} \circ \rho_h^{\mathcal{M}} = \rho_{gh}^{\mathcal{M}}$ for all $g, h \in G$ and $\rho_g^{\mathcal{M}}(a \cdot m) = \rho_g(a) \cdot \rho_g^{\mathcal{M}}(m)$ for all $g \in G$, $a \in \mathscr{W}_X$ and $m \in \mathcal{M}$.

The category of quasi-G-equivariant $\mathscr{W}_X$-modules will be denoted $\mathrm{Mod}_G(\mathscr{W}_X)$. If $\mathcal{M}$ and $\mathcal{N}$ are elements in $\mathrm{Obj}(\mathrm{Mod}_G(\mathscr{W}_X))$, a morphism $\phi \in \mathrm{Hom}_{\mathrm{Mod}_G(\mathscr{W}_X)}(\mathcal{M}, \mathcal{N})$ is a collection of morphisms $\phi_U : \mathcal{M}(U) \to \mathcal{N}(U)$ of $\mathscr{W}_X(U)$-modules, one for each open set $U \subset X$, that satisfies the usual conditions of being a $\mathscr{W}_X$-homomorphism and is such that, for each $g \in G$, the diagram

$$\begin{tikzcd}
\mathcal{M}(U) \arrow[r, "\phi_U"] \arrow[d, "\rho_g^{\mathcal{M}}(U)"'] & \mathcal{N}(U) \arrow[d, "\rho_g^{\mathcal{N}}(U)"] \\
\mathcal{M}(T_g(U)) \arrow[r, "\phi_{T_g(U)}"'] & \mathcal{N}(T_g(U))
\end{tikzcd}$$

is commutative.

Definition 2.5. Let G act on the algebra $\mathscr{W}_X$. A map $\mu_{\mathscr{W}} : \mathfrak{g} \to \mathscr{W}_X(1)$ is said to be a *quantized moment map* for the G-action if $\mu_{\mathscr{W}}$ satisfies the following properties:

(i) $[\mu_{\mathscr{W}}(A), a] = \dfrac{\mathrm{d}}{\mathrm{d}t}\, \rho_{\exp(tA)}(a)|_{t=0}$,

(ii) $\sigma_0(\hbar \mu_{\mathscr{W}}(A)) = A \circ \mu_X$,

(iii) $\mu_{\mathscr{W}}(\mathrm{Ad}(g)A) = \rho_g(\mu_{\mathscr{W}}(A))$,

for every $A \in \mathfrak{g}$, $a \in \mathscr{W}_X$ and $g \in G$.

Let $\mathbb{X}(G) := \mathrm{Hom}_{\mathrm{gp}}(G, \mathbb{C}^*)$ be the lattice of G-characters. Note that if $a \in \mathcal{W}_X$ is a θ-semiinvariant of G (that is, $\rho_g(a) = \theta(g)a$ for all $g \in G$), where $\theta \in \mathbb{X}(G)$, then

(2) $$[\mu_{\mathcal{W}}(A), a] = d\theta(A)a,$$

where $d : \mathbb{X}(G) \to (\mathfrak{g}^*)^G$ is the differential sending a G-character to the corresponding $\mathfrak{g}$-character. From now on we omit the symbol d and think of $\theta \in \mathbb{X}(G)$ as a character for both G and $\mathfrak{g}$. For $\chi \in (\mathfrak{g}^*)^G$, we set

(3) $$\mathcal{L}_{X,\chi} = \mathcal{W}_X \Big/ \sum_{A \in \mathfrak{g}} \mathcal{W}_X(\mu_{\mathcal{W}}(A) - \chi(A)).$$

Note that $\mathcal{L}_{X,\chi}$ is a good quasi-G-equivariant $\mathcal{W}_X$-module, and has lattice

$$\mathcal{L}_{X,\chi}(0) := \mathcal{W}_X(0) \Big/ \sum_{A \in \mathfrak{g}} \mathcal{W}_X(-1)(\mu_{\mathcal{W}}(A) - \chi(A)).$$

We will require the following result, whose proof is based on Holland's result [1999, Proposition 2.4].

Proposition 2.6. *Assume that the moment map μ_X is flat. Then, on X we have an isomorphism of graded sheaves*

$$\mathsf{gr}(\mathcal{L}_{X,\chi}) \simeq \bigoplus_{n \in \mathbb{Z}} \mathcal{O}_{\mu_X^{-1}(0)} \hbar^{-n}.$$

Proof. The moment map $\mu_{\mathcal{W}}$ makes $\mathcal{W}_X$ into a right $U(\mathfrak{g})$-module. Let $\mathbb{C}_\chi$ be the one-dimensional $U(\mathfrak{g})$-module defined by the character χ so that

$$\mathcal{L}_{X,\chi} = \mathcal{W}_X \otimes_{U(\mathfrak{g})} \mathbb{C}_\chi.$$

As in [Holland 1999, Proposition 2.4], we denote by $B_\bullet$ the Chevalley–Eilenberg resolution of $\mathbb{C}_\chi$. Thus, $B_k = U(\mathfrak{g}) \otimes \bigwedge^k \mathfrak{g}$, and the differential is given by

$$\begin{aligned} d_k(f \otimes x_1 \wedge \cdots \wedge x_k) = &\sum_{i=1}^{k} (-1)^{i+1} f(x_i - \chi(x_i)) \otimes x_1 \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge x_k \\ &+ \sum_{1 \le i < j \le k} (-1)^{i+j} f \otimes [x_i, x_j] \wedge x_1 \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge \hat{x}_j \wedge \cdots \wedge x_k. \end{aligned}$$

Then $B_\bullet$ is a complex of free $U(\mathfrak{g})$-modules such that $H^0(B_\bullet) = \mathbb{C}_\chi$ and $H^k(B_\bullet) = 0$ for k nonzero. Let $C_\bullet = \mathcal{W}_X \otimes_{U(\mathfrak{g})} B_\bullet = \mathcal{W}_X \otimes \bigwedge^\bullet \mathfrak{g}$. The filtration on $\mathcal{W}_X$ induces a filtration $F_n C_k = \mathcal{W}_X(n-k) \otimes \bigwedge^k \mathfrak{g}$ on the complex $C_\bullet$ such that $d_k(F_n C_k) \subseteq F_n C_{k-1}$ (recall that $\mu_{\mathcal{W}}(\mathfrak{g}) \subset \mathcal{W}_X(1)$). Note that the filtration is not bounded above or below. However, by Lemma 2.2 the filtration on $C_\bullet$ is exhaustive, Hausdorff and complete. We denote by $E^r_{p,q}$ the spectral sequence corresponding to the filtration F_n on $C_\bullet$. Since the filtration is exhaustive, Hausdorff and complete, the proof of [Weibel

1994, Theorem 5.5.10] shows that the spectral sequence E converges to $H_\bullet(C)$ (that the sequence is regular follows from the fact, to be shown below, that it collapses at E^1). By construction, we have an isomorphism of filtered sheaves $H^0(C) \simeq \mathscr{L}_{X,\chi}$ and hence $\mathsf{gr}(H_0(C)) \simeq \mathsf{gr}(\mathscr{L}_{X,\chi})$. Denote by A the graded sheaf of algebras $\bigoplus_{n\in\mathbb{Z}} \mathbb{O}_X \hbar^{-n}$, where $\mathbb{O}_X$ is in degree zero and $\hbar$ has degree -1. The 0-th page of the spectral sequence is given by

$$E^0_{p,q} = \mathscr{W}_X(p-q) \otimes \bigwedge^{p+q} \mathfrak{g} / \mathscr{W}_X(p-1-q) \otimes \bigwedge^{p+q} \mathfrak{g} \simeq A_{p-q} \otimes \bigwedge^{p+q} \mathfrak{g}.$$

Since $\mathbb{C}[\mathfrak{g}^*]$ is a domain and μ_X is assumed to be flat, $\mu_X^* : \mu_X^{-1}\mathbb{O}_{\mathfrak{g}^*} \to \mathbb{O}_X$ is an embedding and we may think of $\mu_X^{-1}\mathbb{O}_{\mathfrak{g}^*}$ as a subsheaf of $\mathbb{O}_X$. Let $x_1, \dots, x_r$ be a basis of $\mathfrak{g}$. Then [Bruns and Herzog 1993, Proposition 1.1.2] implies that $\hbar^{-1}x_1, \dots, \hbar^{-1}x_r$ form a regular sequence in A at those points where they vanish. By Definition 2.5(ii), the symbol $\sigma_1(\mu_{\mathscr{W}}(x_i))$ equals $\hbar^{-1}x_i \in A$. Thus the differential on E^0 is given by

$$d_{p+q}(f \otimes x_1 \wedge \cdots \wedge x_{p+q}) = \sum_{i=1}^{p+q} f\hbar^{-1}x_i \otimes x_1 \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge x_{p+q}.$$

As is explained in [Holland 1999, Proposition 2.4], the only nonzero homology of E^0 is in the $(p, -p)$ position, where we have

$$E^1_{(p,-p)} = \frac{A_p}{A_{p-1} \cdot \hbar \mu_X^*(\mathfrak{g})} \simeq \mathbb{O}_{\mu_X^{-1}(0)} \hbar^{-p}.$$

Therefore the sequence collapses at E^1 and we have

$$\mathsf{gr}(\mathscr{L}_{X,\chi})_p \simeq \mathsf{gr}(H_0(C))_p \simeq \mathbb{O}_{\mu_X^{-1}(0)} \hbar^{-p},$$

as required. □

2F. *F-actions.* Here we repeat the definition of an F-action on $\mathscr{W}_X$-modules as defined in [Kashiwara and Rouquier 2008]. Let $\mathbb{C}^\times \ni t \mapsto T_t \in \mathrm{Aut}(X)$ denote an action of the torus $\mathbb{C}^\times$ on X such that the symplectic 2-form is a semiinvariant of positive weight: $T_t^*\omega_X = t^m\omega_X$ for some $m > 0$.

Definition 2.7. An F-action with exponent m on $\mathscr{W}_X$ is an action of the group $\mathbb{C}^\times$ on $\mathscr{W}_X$ as in Definition 2.3 except that $\mathbb{C}^\times$ also acts on $\hbar$: if $\mathscr{F}_t : \mathscr{W}_X \xrightarrow{\sim} T_t^{-1}\mathscr{W}_X$ denotes the action of $t \in \mathbb{C}^\times$ then we require that $\mathscr{F}_t(\hbar) = t^m\hbar$ for all $t \in \mathbb{C}^\times$.

It will be convenient to extend the F-action of $\mathbb{C}^\times$ to an action on

$$\mathscr{W}[\hbar^{1/m}] := \boldsymbol{k}(\hbar^{1/m}) \otimes_{\boldsymbol{k}} \mathscr{W}$$

by setting $\mathscr{F}_t(\hbar^{1/m}) = t\hbar^{1/m}$. The category of F-equivariant $\mathscr{W}_X$-modules will be denoted $\mathrm{Mod}_F(\mathscr{W}_X)$. As noted in [Kashiwara and Rouquier 2008, §2.3.1],

$\mathrm{Mod}_F(\mathcal{W}_X)$ is an abelian category. Moreover [ibid., §2.3], if there exists a relatively compact open subset U of X such that $\mathbb{C}^\times \cdot U = X$ then every good, F-equivariant $\mathcal{W}_X$-module admits globally a coherent $\mathcal{W}_X(0)$-lattice. Such an open set U will exist in the cases we consider. The following lemma will be used later.

Lemma 2.8. *Let $\mathcal{M}, \mathcal{N} \in \mathrm{Mod}^{\mathrm{good}}_{F,G}(\mathcal{W}_X)$. Assume that $\mathcal{M} \simeq \mathcal{W}_X/\mathcal{I}$ is a cyclic $\mathcal{W}_X$-module, generated by some G, F-invariant element, where $\mathcal{I}$ is a left ideal generated by finitely many global sections. Then*

$$\mathrm{Hom}_{\mathrm{Mod}^{\mathrm{good}}_{F,G}(\mathcal{W}_X)}(\mathcal{M}, \mathcal{N}) = \mathrm{Hom}_{\mathcal{W}_X(X)}(\mathcal{M}(X), \mathcal{N}(X))^{G,F}.$$

2G. *Example.* Let V be an n-dimensional vector space. We fix $X = T^*V$ with coordinates $x_1, \dots, x_n, \xi_1, \dots, \xi_n$ and define an action T_t of $\mathbb{C}^\times$ on X such that the corresponding action on coordinate functions is given by $T_t(x_i) = tx_i$ and $T_t(\xi_i) = t\xi_i$. Then $T_t^*\omega_X = t^2\omega_X$. We extend this to an F-action on $\mathcal{W}_{T^*V}$ by setting $\mathcal{F}_t(\hbar) = t^2\hbar$. Let $\mathfrak{D}(V)$ denote the ring of algebraic differential operators on V.

Lemma 2.9. *Taking F-invariants in $\mathcal{W}_{T^*V}(T^*V)$ gives*

$$\begin{aligned}\mathrm{End}_{\mathrm{Mod}_F(\mathcal{W}_{T^*V}[\hbar^{1/2}])}(\mathcal{W}_{T^*V}[\hbar^{1/2}])^{\mathrm{opp}} &= \mathbb{C}\big[\hbar^{-1/2}x_i, \hbar^{-1/2}\xi_i : i \in [1, n]\big] \\ &= \mathbb{C}\Big[\hbar^{-1/2}x_i, \hbar^{1/2}\frac{\partial}{\partial x_i} : i \in [1, n]\Big],\end{aligned}$$

where the second equality comes from

$$\mathfrak{D}(V) \hookrightarrow \mathcal{W}_{T^*V}(T^*V), \quad x_i \mapsto x_i \quad \textit{and} \quad \frac{\partial}{\partial x_i} \mapsto \hbar^{-1}\xi_i.$$

Proof. We can identify $\mathrm{End}_{\mathrm{Mod}_F(\mathcal{W}_{T^*V}[\hbar^{1/2}])}(\mathcal{W}_{T^*V}[\hbar^{1/2}])^{\mathrm{opp}}$ with the algebra of F-invariant global sections, $\mathcal{W}_{T^*V}[\hbar^{1/2}](T^*V)^F$. Since T^*V is connected, taking a power series expansion in a sufficiently small neighborhood of $0 \in T^*V$ defines an embedding $\mathcal{O}_{T^*V}(T^*V) \hookrightarrow \mathbb{C}[\![x_1, \dots, x_n, \xi_1, \dots, \xi_n]\!]$. As $\mathbb{C}^\times$-modules, we can identify $\mathcal{W}_{T^*V}[\hbar^{1/2}]$ with $\mathcal{O}_{T^*V} \mathbin{\widehat{\otimes}} \mathbb{C}(\!(\hbar^{1/2})\!)$ and we get a $\mathbb{C}^\times$-equivariant embedding

$$\mathcal{W}_{T^*V}[\hbar^{1/2}](T^*V) \hookrightarrow \mathbb{C}[\![x_1, \dots, x_n, \xi_1, \dots, \xi_n]\!] \mathbin{\widehat{\otimes}} \mathbb{C}(\!(\hbar^{1/2})\!),$$

where we denote by $\widehat{\otimes}$ the completed tensor product with respect to the linear topology. Taking invariants gives the desired result. □

A trivial application of Theorem 3.3 below, with $f = \mathrm{id}_{\mathbb{C}^n}$ and $G = \{1\}$, shows

$$\mathrm{Mod}_F(\mathcal{W}_{T^*V}[\hbar^{1/2}]) \simeq \mathbb{C}\Big[\hbar^{-1/2}x_i, \hbar^{1/2}\frac{\partial}{\partial x_i} : i \in [1, n]\Big]\text{-mod}.$$

3. *W*-affinity

In this section we give a criterion for the W-affinity of a class of W-algebras on those symplectic manifolds that are obtained by Hamiltonian reduction.

3A. ***The geometric setup.*** Let V be an n-dimensional vector space over $\mathbb{C}$. Its cotangent bundle T^*V has the structure of a complex symplectic manifold. Let G be a connected, reductive algebraic group acting algebraically on V. This action induces a Hamiltonian action on T^*V and we have a moment map

$$\mu_{T^*V} : T^*V \longrightarrow \mathfrak{g}^* := (\text{Lie } G)^*$$

such that $\mu_{T^*V}(0) = 0$. We fix a character $\vartheta \in \mathbb{X}(G)$. Let $\mathfrak{X}$ be the open subset of all ϑ-semistable points in T^*V and denote the restriction of μ_{T^*V} to $\mathfrak{X}$ by $\mu_{\mathfrak{X}}$. We assume that

(i) the set $\mu_{\mathfrak{X}}^{-1}(0)$ is nonempty,

(ii) G acts freely on $\mu_{\mathfrak{X}}^{-1}(0)$,

(iii) the moment map μ_{T^*V} is flat.

Set

$$Y_\vartheta := \mu_{\mathfrak{X}}^{-1}(0)/\!/G = \text{Proj} \bigoplus_{n \geq 0} \mathbb{C}[\mu_{T^*V}^{-1}(0)]^{n\vartheta}$$

and write $f : Y_\vartheta \to \mu_{T^*V}^{-1}(0)/\!/G =: Y_0$ for the corresponding projective morphism. Condition (i) implies that the categorical quotient Y_ϑ is nonempty. Condition (ii) implies that the morphism $\mu_{\mathfrak{X}}$ is regular at all points in $\mu_{\mathfrak{X}}^{-1}(0)$ and hence Y_ϑ is a nonsingular symplectic manifold. Condition (iii) will be used in Proposition 3.5. We add to our previous assumptions:

(iv) The morphism f is birational and Y_0 is a normal variety.

In the case of hypertoric varieties, it is shown in Section 4 that assumptions (i)–(iv) hold when the matrix A is unimodular.

Lemma 3.1. *Let $\mathbb{O}_{Y_\vartheta}^{\text{alg}}$ and $\mathbb{O}_{Y_0}^{\text{alg}}$ denote the sheaves of regular functions on Y_ϑ and Y_0, respectively. If Y_ϑ, Y_0, f satisfy assumption* (*iv*) *then $\Gamma(Y_\vartheta, \mathbb{O}_{Y_\vartheta}^{\text{alg}}) = \Gamma(Y_0, \mathbb{O}_{Y_0}^{\text{alg}})$.*

Proof. It is well-known that the condition implies the statement of the lemma, but we were unable to find any suitable reference, therefore we include a proof for the reader's convenience. For $s \geq 0$, fix $R_s = \mathbb{C}[\mu_{T^*V}^{-1}(0)]^{s\vartheta}$ and $R = \bigoplus_{s \geq 0} R_s$ so that $Y_\vartheta = \text{Proj } R$ and recall that f is the canonical projective morphism from Y_ϑ to Y_0. By Hilbert's Theorem (see [Kraft 1984, Zusatz 3.2]), R is finitely generated as an R_0-algebra. Let $x_1, \dots, x_n \in R$ be homogeneous generators (of degree at least one) of R as an R_0-algebra. Then the affine open sets $D_+(x_i) = \text{Spec } R_{(x_i)}$ form an open cover of Y_ϑ and

$$\Gamma(Y_\vartheta, \mathbb{O}_{Y_\vartheta}^{\text{alg}}) = \bigcap_{i=1}^{n} R_{(x_i)} \subseteq \bigcap_{i=1}^{n} R_{x_i},$$

Let $r \in \Gamma(Y_\vartheta, \mathbb{O}_{Y_\vartheta}^{\text{alg}})$. Then, for each i, there exists an m such that $x_i^m \cdot r \in R$. We choose one m sufficiently large so that $x_i^m \cdot r \in R$ for all i. Since the x_i generate

R, we actually have $y \cdot r \in R$ for all $y \in R_s$ and $s \geq m_0 := nmd$, where d is the maximum of the degrees of $x_1, \dots, x_n$. Therefore $y \cdot r \in R$ for all $y \in \bigoplus_{s \geq m_0} R_s$. Since r has degree zero,

$$y \cdot r \in \bigoplus_{s \geq m_0} R_s \quad \text{for all } \; y \in \bigoplus_{s \geq m_0} R_s.$$

Inductively, $y \cdot r^q \in \bigoplus_{s \geq m_0} R_s$ for all $q \geq 1$. Take $y = x_1^{m_0}$, then $r^q \in (1/x_1^{m_0})R$ for all $q \geq 1$ and hence $R[r] \subset (1/x_1^{m_0})R$. But, by Hilbert's basis theorem, R is Noetherian and the R-module $(1/x_1^{m_0})R$ is finitely generated, hence the algebra $R[r]$ is finite over R. This means r satisfies some monic polynomial $u^t + r_1 u^{t-1} + \cdots + r_t$ with coefficients in R. However R has degree zero so without loss of generality $r_i \in R_0$. Thus r is in the integral closure of R_0 in the degree zero part of the field of fractions of R. Now [Hartshorne 1977, Theorem 7.17] says that, since the map f is projective and birational, there exists an ideal I in R_0 such that $R_k \simeq I^k$ as R_0-modules and we have an isomorphism of graded rings $R \simeq \bigoplus_{k \geq 0} I^k$. That is, Y_ϑ is isomorphic to the blowup of Y_0 along $V(I)$. Therefore we can identify the degree zero part of the field of fractions of R with the field of fractions of R_0. Since R_0 is assumed to be normal, $r \in R_0$ as required. $\square$

3B. The quotient morphism will be written $p : \mu_{\mathfrak{X}}^{-1}(0) \to Y_\vartheta$. For each character $\theta \in \mathbb{X}(G)$ and vector space M on which G acts, we denote by M^θ the set of elements $m \in M$ such that $g \cdot m = \theta(g)m$ for all $g \in G$. We can define a coherent sheaf L_θ on the quotient Y_ϑ by $L_\theta(U) := \left[\mathbb{O}_{\mu_{\mathfrak{X}}^{-1}(0)}(p^{-1}(U))\right]^\theta$. Since G acts freely on $\mu_{\mathfrak{X}}^{-1}(0)$, L_θ is a line bundle on Y_ϑ.

3C. ***Quantum Hamiltonian reduction.*** Differentiating the action of G on V produces a morphism of Lie algebras $\mu_D : \mathfrak{g} \to \mathrm{Vect}(V)$, from $\mathfrak{g}$ into the Lie algebra of algebraic vector fields on V:

$$\mu_D(A)(r) := \frac{\mathrm{d}}{\mathrm{d}t}\, a^*_{\exp(tA)}(r)|_{t=0},$$

where $a : G \times V \to V$ is the action map and $a^* : G \times \mathbb{O}(V) \to \mathbb{O}(V)$ the induced action on functions. We write $\mathfrak{D}(V)$ for the ring of *algebraic* differential operators on V. Since $\mathrm{Vect}(V) \subset \mathfrak{D}(V)$ we get a map $\mu_D : \mathfrak{g} \to \mathfrak{D}(V)$ which extends to an algebra morphism $U(\mathfrak{g}) \to \mathfrak{D}(V)$. For $\chi \in (\mathfrak{g}^*)^G$, $\theta \in \mathbb{X}$, we define the left $\mathfrak{D}(V)$-module

$$\mathscr{L}_{D,\chi} := \mathfrak{D}(V) \Big/ \sum_{A \in \mathfrak{g}} \mathfrak{D}(V)(\mu_D(A) - \chi(A)),$$

and the algebra and $(\mathsf{U}_\chi, \mathsf{U}_{\chi+\theta})$-bimodule, respectively:

$$\mathsf{U}_\chi = \left(\mathrm{End}_{\mathfrak{D}(V)}(\mathscr{L}_{D,\chi})^G\right)^{\mathrm{opp}}, \qquad \mathsf{U}_\chi^\theta = \mathrm{Hom}_{\mathfrak{D}(V)}(\mathscr{L}_{D,\chi}, \mathscr{L}_{D,\chi+\theta} \otimes \mathbb{C}_\theta)^G.$$

Fix $\chi \in (\mathfrak{g}^*)^G$ and $\theta \in \mathbb{X}$. We consider the following natural homomorphisms:

$$\mathsf{U}^{-\theta}_{\chi+\theta} \otimes_{\mathbb{C}} \mathsf{U}^{\theta}_{\chi} \longrightarrow \mathsf{U}_{\chi+\theta}, \qquad \phi \otimes \psi \mapsto (\mathsf{id}_{\mathsf{U}_{\chi+\theta}} \otimes \mathsf{ev}) \circ (\psi \otimes \mathsf{id}_{-\theta}) \circ \phi, \tag{4}$$

$$\mathsf{U}^{\theta}_{\chi} \otimes_{\mathbb{C}} \mathsf{U}^{-\theta}_{\chi+\theta} \longrightarrow \mathsf{U}_{\chi}, \qquad \phi \otimes \psi \mapsto (\mathsf{id}_{\mathsf{U}_{\chi}} \otimes \mathsf{ev}) \circ (\psi \otimes \mathsf{id}_{\theta}) \circ \phi, \tag{5}$$

where $\circ$ is composition of morphisms and $\mathsf{ev} : \mathbb{C}_{-\theta} \otimes \mathbb{C}_{\theta} \to \mathbb{C}$ is the natural map. We write $\chi \to \chi + \theta$ if the map (4) is surjective and similarly $\chi + \theta \to \chi$ if the map (5) is surjective. Note that if $\chi + \theta \leftrightarrows \chi$ then, as shown in [McConnell and Robson 2001, Corollary 3.5.4], the algebras U_{χ} and $\mathsf{U}_{\chi+\theta}$ are Morita equivalent.

3D. ***The sheaf of W-algebras.*** Denote by $\mathscr{W}_{\mathfrak{X}}$ the restriction of the canonical W-algebra $\mathscr{W}_{T^*V}$ to $\mathfrak{X}$. We define an action of the torus $\mathbb{C}^{\times}$ on T^*V by $T_t(v) = t^{-1}v$ for all $v \in T^*V$; $\mathfrak{X}$ is a $\mathbb{C}^{\times}$-stable open set. The algebra $\mathscr{W}_{\mathfrak{X}}$ is then equipped with an F-action of weight 2 as defined in the setup of Lemma 2.9. Define $\widetilde{\mathscr{W}}_{T^*V} := \mathscr{W}_{T^*V}[\hbar^{1/2}]$ and write $\widetilde{\mathscr{W}}_{\mathfrak{X}}$ for its restriction to $\mathfrak{X}$. As noted in Section 2A, we have an embedding $j : \mathfrak{D}(V) \hookrightarrow \mathscr{W}_{T^*V}$, $x_i \mapsto x_i$ and $\partial/\partial x_i \mapsto \hbar^{-1}\xi_i$. Composing this morphism with the map $\mu_D : \mathfrak{g} \to \mathfrak{D}(V)$ gives us a map $\mu_{\mathscr{W}} = j \circ \mu_D : \mathfrak{g} \to \mathscr{W}_{T^*V}$. It is a quantized moment map in the sense of Definition 2.5. Then, as in (3), for each $\chi \in (\mathfrak{g}^*)^G$, we have defined the $\widetilde{\mathscr{W}}_{T^*V}$-module $\mathscr{L}_{T^*V,\chi}$. Its restriction to $\mathfrak{X}$ is denoted $\mathscr{L}_{\chi}$. Recall that $\mathscr{L}_{\chi}$ is a good quasi-G-equivariant $\widetilde{\mathscr{W}}_{\mathfrak{X}}$-module. If we let $\mathbb{C}^{\times}$ act trivially on $\mathfrak{g}$ then the morphism $\mu_{\mathscr{W}}$ is F-equivariant and hence $\mathscr{L}_{\chi}$ is equipped with an F-action. The image of 1 in $\mathscr{L}_{\chi}$ will be denoted by u_{χ}.

3E. Kashiwara and Rouquier [2008] show that one can quantize the process of Hamiltonian reduction to get a family of sheaves of W-algebras on Y_{ϑ} beginning from a W-algebra on T^*V. Set

$$\mathscr{A}_{\chi} = \left(\left(p_*\mathscr{E}nd_{\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_{\chi})\right)^G\right)^{\mathrm{opp}} \quad \text{and} \quad \mathscr{A}_{\chi,\theta} = \left(p_*\mathscr{H}om_{\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_{\chi}, \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_{\theta})\right)^G,$$

where $\theta \in \mathbb{X}(G)$ and $\mathbb{C}_{\theta}$ denotes the corresponding one dimensional G-module. By [Kashiwara and Rouquier 2008, Proposition 2.8], $\mathscr{A}_{\chi}$ is a W-algebra on Y_{ϑ} and $\mathscr{A}_{\chi,\theta}$ is a $(\mathscr{A}_{\chi}, \mathscr{A}_{\chi+\theta})$-bimodule. Let

$$\mathscr{A}_{\chi}(0) = \left(\left(p_*\mathscr{E}nd_{\widetilde{\mathscr{W}}_{\mathfrak{X}}(0)}(\mathscr{L}_{\chi}(0))\right)^G\right)^{\mathrm{opp}},$$
$$\mathscr{A}_{\chi,\theta}(0) = (p_*\mathscr{H}om_{\widetilde{\mathscr{W}}_{\mathfrak{X}}(0)}(\mathscr{L}_{\chi}(0), \mathscr{L}_{\chi+\theta}(0) \otimes \mathbb{C}_{\theta}))^G,$$

so that $\mathscr{A}_{\chi,\theta}(0)$ is a $\mathscr{A}_{\chi}(0)$-lattice of $\mathscr{A}_{\chi,\theta}$. We have $\mathscr{A}_{\chi}(0)/\mathscr{A}_{\chi}(-1/2) \simeq \mathscr{O}_{Y_{\vartheta}}$ and, as noted in [ibid., Proposition 2.8(iii)], $\mathscr{A}_{\chi,\theta}(0)/\mathscr{A}_{\chi,\theta}(-1/2) \simeq L_{-\theta}$, where L_{θ} is the line bundle as defined above. We say that a good $\mathscr{A}_{\chi}$-module $\mathscr{M}$ is generated, locally on Y_0, by its global sections if for each $y \in Y_0$ there exists some open neighborhood (in the complex analytic topology) $U \subset Y_0$ of y such that the natural map of left $(\mathscr{A}_{\chi})_{|f^{-1}(U)}$-modules $(\mathscr{A}_{\chi})_{|f^{-1}(U)} \otimes \mathscr{M}(f^{-1}(U)) \to \mathscr{M}_{|f^{-1}(U)}$ is surjective.

Definition 3.2. We denote by $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ the full subcategory of $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ consisting of all good, F-equivariant $\mathscr{A}_\chi$-modules $\mathscr{M}$ such that:

(i) $\mathscr{M}$ is generated, locally on Y_0, by its global sections.

(ii) For any nonzero submodule $\mathscr{N}$ of $\mathscr{M}$ in $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ we have

$$\mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \mathscr{N}) \neq 0.$$

3F. *W-affinity.* We can now state the main result relating the sheaf of W-algebras $\mathscr{A}_\chi$ on Y_ϑ and the algebra of quantum Hamiltonian reduction U_χ.

Theorem 3.3. *Let $\mathscr{A}_\chi$ and U_χ be as above and choose some $\theta \in \mathbb{X}(G)$ such that L_θ is ample.*

(i) *There is an isomorphism of algebras $\Gamma(Y_\vartheta, \mathscr{A}_\chi)^F \simeq \mathsf{U}_\chi$.*

(ii) *Assume that we have $\chi \leftarrow \chi + n\theta$ for all $n \in \mathbb{Z}_{\geq 0}$. Then the functor*

$$\mathscr{M} \mapsto \mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \mathscr{M})$$

defines an equivalence of categories $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi$-mod with quasiinverse $\mathscr{M} \mapsto \mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} \mathscr{M}$.

(iii) *Assume that we have $\chi \leftrightarrows \chi + n\theta$ for all $n \in \mathbb{Z}_{\geq 0}$. Then the functor*

$$\mathscr{M} \mapsto \mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \mathscr{M})$$

defines an equivalence of categories $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi$-mod with quasiinverse $\mathscr{M} \mapsto \mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} \mathscr{M}$.

The proof of Theorem 3.3 will occupy the remainder of Section 3.

3G. *Proof of the theorem.* We fix $\mathscr{A}_\chi$, U_χ and L_θ as in Theorem 3.3. First we require some preparatory lemmata. Denote by ι the embedding $\mathfrak{D}(V) \hookrightarrow \widetilde{\mathscr{W}}_{T^*V}(T^*V)$ given by $x_i \mapsto \hbar^{-1/2} x_i$ and $\partial_i \mapsto \hbar^{-1/2}\xi_i$. Equip $\mathfrak{D}(V)$ with a $\frac{1}{2}\mathbb{Z}$-filtration $F_\bullet\,\mathfrak{D}(V)$ by placing x_i and ∂_i in degree $\frac{1}{2}$ (this is the Bernstein filtration). Then ι is a strictly filtered embedding in the sense that

$$\iota(F_k\mathfrak{D}(V)) = \iota(\mathfrak{D}(V)) \cap \widetilde{\mathscr{W}}_{T^*V}(T^*V)(k), \quad \text{for all} \quad k \in \tfrac{1}{2}\mathbb{Z}.$$

By Lemma 2.9, the image of $\mathfrak{D}(V)$ in $\widetilde{\mathscr{W}}_{T^*V}(T^*V)$ is $\widetilde{\mathscr{W}}_{T^*V}(T^*V)^F$. This implies, since $\mathbb{C}^\times$ is reductive and $\mu_{\mathscr{W}}$ is equivariant, that

$$(6) \quad \sum_{A\in\mathfrak{g}} \widetilde{\mathscr{W}}_{T^*V}(T^*V)(\mu_{\mathscr{W}}(A) - \chi(A)) \cap \iota(\mathfrak{D}(V)) = \sum_{A\in\mathfrak{g}} \widetilde{\mathscr{W}}_{T^*V}(T^*V)^F(\mu_{\mathscr{W}}(A) - \chi(A)),$$

which in turn equals

$$\sum_{A \in \mathfrak{g}} \mathfrak{D}(V)(\mu_D(A) - \chi(A)). \tag{7}$$

Lemma 3.4 [Ginzburg et al. 2009, Lemma 2.2].

(i) *Multiplication in $\mathfrak{D}(V)$ defines an algebra structure on $(\mathscr{L}_{D,\chi})^G$ such that there is isomorphism of algebras $\mathsf{U}_\chi \xrightarrow{\sim} (\mathscr{L}_{D,\chi})^G$ given by $\phi \mapsto \phi(u_\chi)$ with inverse $f \mapsto r_f$, where $r_f = \cdot f$ is right multiplication by f.*

(ii) *We have an isomorphism of $(\mathsf{U}_\chi, \mathsf{U}_{\chi+\theta})$-bimodules $\mathsf{U}_\chi^\theta \xrightarrow{\sim} (\mathscr{L}_{D,\chi+\theta})^{-\theta}$ given by $\phi \mapsto f$, where $\phi(u_\chi) = f u_{\chi+\theta} \otimes \theta$, with inverse $f u_{\chi+\theta} \mapsto r_f \otimes \theta$.*

Let us introduce

$$\mathsf{E}_\chi = \big(\mathrm{End}_{\mathrm{Mod}_{G,F}\,\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_\chi)\big)^{\mathrm{opp}} \quad \text{and} \quad \mathsf{E}_\chi^\theta = \mathrm{Hom}_{\mathrm{Mod}_{F,G}\,\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_\chi, \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta),$$

so that E_χ^θ is a $(\mathsf{E}_\chi, \mathsf{E}_{\chi+\theta})$-bimodule and $\mathscr{L}_\chi$ is a $(\widetilde{\mathscr{W}}_{\mathfrak{X}}, \mathsf{E}_\chi)$-bimodule. By Lemma 2.8, we can identify

$$\mathsf{E}_\chi = \big(\mathrm{End}_{\widetilde{\mathscr{W}}_{\mathfrak{X}}(\mathfrak{X})}(\mathscr{L}_\chi)^{G,F}\big)^{\mathrm{opp}} \quad \text{and} \quad \mathsf{E}_\chi^\theta = \mathrm{Hom}_{\widetilde{\mathscr{W}}_{\mathfrak{X}}(\mathfrak{X})}\big(\mathscr{L}_\chi, \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta\big)^{G,F}.$$

Note that (6) implies that the map ι induces an embedding $\iota : \mathscr{L}_{D,\chi} \hookrightarrow \mathscr{L}_{T^*V,\chi}(T^*V)$, and after taking G, F-invariants,

$$\iota : \mathsf{U}_\chi \xrightarrow{\sim} \big(\mathrm{End}_{\widetilde{\mathscr{W}}_{T^*V}(T^*V)}(\mathscr{L}_{T^*V,\chi})^{G,F}\big)^{\mathrm{opp}}, \tag{8}$$

and $\mathsf{U}_\chi^\theta \simeq \mathrm{Hom}_{\widetilde{\mathscr{W}}_{T^*V}(T^*V)}(\mathscr{L}_{T^*V,\chi}, \mathscr{L}_{T^*V,\chi+\theta} \otimes \mathbb{C}_\theta)^{G,F}$.

Proposition 3.5. *We have a filtered isomorphism $\Psi_\chi : \mathsf{U}_\chi \xrightarrow{\sim} \mathsf{E}_\chi$ in the sense that $\Psi_\chi(F_k \mathsf{U}_\chi) = F_k \mathsf{E}_\chi$ for all $k \in \frac{1}{2}\mathbb{Z}$.*

Proof. The isomorphism (8) induced by the embedding ι is filtered in the same sense as Ψ_χ above. Therefore it suffices to show that the natural map

$$\begin{aligned}(\mathscr{L}_{T^*V,\chi}(T^*V))^{G,F} &= \big(\mathrm{End}_{\widetilde{\mathscr{W}}_{T^*V}(T^*V)}(\mathscr{L}_{T^*V,\chi})^{G,F}\big)^{\mathrm{opp}} \\ &\longrightarrow \big(\mathrm{End}_{\widetilde{\mathscr{W}}_{\mathfrak{X}}(\mathfrak{X})}(\mathscr{L}_{\mathfrak{X},\chi})^{G,F}\big)^{\mathrm{opp}} = (\mathscr{L}_{\mathfrak{X},\chi}(\mathfrak{X}))^{G,F}\end{aligned}$$

is a filtered isomorphism. The localization morphism $\mathscr{L}_{T^*V,\chi}(T^*V) \to \mathscr{L}_{T^*V,\chi}(\mathfrak{X})$ is clearly filtered in the weaker sense that it restricts to a map

$$\mathscr{L}_{T^*V,\chi}(T^*V)(k) \to \mathscr{L}_{T^*V,\chi}(\mathfrak{X})(k)$$

for each $k \in \frac{1}{2}\mathbb{Z}$. Since the moment map μ_{T^*V} is assumed to be flat, Proposition 2.6 says that the morphism of associated graded spaces is the natural localization map

$$\bigoplus_{k \in \frac{1}{2}\mathbb{Z}} \mathbb{O}_{\mu_{T^*V}^{-1}(0)}(T^*V)\hbar^{-k} \longrightarrow \bigoplus_{k \in \frac{1}{2}\mathbb{Z}} \mathbb{O}_{\mu_{T^*V}^{-1}(0)}(\mathfrak{X})\hbar^{-k}.$$

Note that the filtration on $\mathscr{L}_{T^*V,\chi}$ is stable with respect to both G and F. Lemma 2.2 says that the globally defined good filtration on $\mathscr{L}_{T^*V,\chi}$ is exhaustive and Hausdorff. Therefore, taking invariants with respect to G and F, it suffices to show that

$$\bigoplus_{k\in\frac{1}{2}\mathbb{Z}} \big(\mathbb{O}_{\mu^{-1}_{T^*V}(0)}(T^*V)\hbar^{-k}\big)^{G,F} \longrightarrow \bigoplus_{k\in\frac{1}{2}\mathbb{Z}} \big(\mathbb{O}_{\mu^{-1}_{T^*V}(0)}(\mathfrak{X})\hbar^{-k}\big)^{G,F}$$

is an isomorphism. But, since the F-action is contracting,

$$\big(\mathbb{O}_{\mu^{-1}_{T^*V}(0)}(T^*V)\hbar^{-k}\big)^{G,F} = \mathbb{C}\big[\mu^{-1}_{T^*V}(0)\big]^G_{-2k},$$

which is the space of G-invariant homogeneous polynomials on $\mu^{-1}_{T^*V}(0)$ of degree $-2k$. Similarly,

$$\big(\mathbb{O}_{\mu^{-1}_{T^*V}(0)}(\mathfrak{X})\hbar^{-k}\big)^{G,F} = \mathbb{C}\big[\mu^{-1}_{\mathfrak{X}}(0)\big]^G_{-2k}.$$

Therefore the result follows from Lemma 3.1, which says that

$$\mathbb{C}[\mu^{-1}_{\mathfrak{X}}(0)]^G = \Gamma(Y_\vartheta, \mathbb{O}^{\mathrm{alg}}_{Y_\vartheta}) = \Gamma(Y_0, \mathbb{O}^{\mathrm{alg}}_{Y_0}) = \mathbb{C}[\mu^{-1}_{T^*V}(0)]^G. \qquad \square$$

Remark 3.6. In general, it is *not* true that $\mathsf{U}^\theta_\chi \simeq \mathsf{E}^\theta_\chi$ when $\theta \neq 0$.

3H. *Shifting*. The localization theorem relies on the following result by Kashiwara and Rouquier:

Theorem 3.7 [Kashiwara and Rouquier 2008, Theorem 2.9]. *Let $\mathscr{A}_{\chi,\theta}$ and L_θ be as above such that L_θ is ample.*

(i) *Assume that for all $n \gg 0$, there exists a finite dimensional vector space W_n and a split epimorphism of left $\mathscr{A}_\chi$-modules $\mathscr{A}_{\chi,n\theta} \otimes W_n \twoheadrightarrow \mathscr{A}_\chi$. Then, for every good $\mathscr{A}_\chi$-module $\mathscr{M}$, we have $\mathbb{R}^i f_*(\mathscr{M}) = 0$ for $i \neq 0$.*

(ii) *Assume that for all $n \gg 0$ there exists a finite dimensional vector space U_n and a split epimorphism of left $\mathscr{A}_\chi$-modules $\mathscr{A}_\chi \otimes U_n \twoheadrightarrow \mathscr{A}_{\chi,n\theta}$. Then every good $\mathscr{A}_\chi$-module is generated, locally on Y_0, by its global sections.*

Lemma 3.8. *Let $\mathscr{A}_\chi$ and U_χ be as above and choose $\theta \in \mathbb{X}(G)$.*

(i) *If $\chi \leftarrow \chi + \theta$ then there exists a finite dimensional vector space W and a split epimorphism $\mathscr{A}_{\chi,\theta} \otimes W \twoheadrightarrow \mathscr{A}_\chi$.*

(ii) *If $\chi \rightarrow \chi + \theta$ then there exists a finite dimensional vector space U and a split epimorphism $\mathscr{A}_\chi \otimes U \twoheadrightarrow \mathscr{A}_{\chi,\theta}$.*

Proof. We begin with (i). Equation (8) implies that we have a morphism $\mathsf{U}_\chi^\theta \to \mathsf{E}_\chi^\theta$, which a direct calculation shows is a morphism of $(\mathsf{U}_\chi, \mathsf{U}_{\chi+\theta}) = (\mathsf{E}_\chi, \mathsf{E}_{\chi+\theta})$-bimodules (here we identify U_χ with E_χ via the isomorphism of Proposition 3.5). Thus $\chi \leftarrow \chi+\theta$ implies that $\mathsf{E}_\chi^\theta \otimes \mathsf{E}_{\chi+\theta}^{-\theta} \twoheadrightarrow \mathsf{E}_\chi$. Therefore there exists some k and $\phi_i \in \mathsf{E}_\chi^\theta$, $\psi_i \in \mathsf{E}_{\chi+\theta}^{-\theta}$ for $i \in [1,k]$ such that

$$(\mathsf{id}_{\mathscr{L}_\chi} \otimes \mathsf{ev}) \circ \left(\sum_{i=1}^k (\psi_i \otimes \mathsf{id}_{\mathbb{C}_\theta}) \circ \phi_i \right) = \mathsf{id}_{\mathscr{L}_\chi}.$$

Let $W = \mathrm{Span}_{\mathbb{C}}\{\psi_i : i \in [1,k]\}$ and define $\Psi : \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta \otimes W \to \mathscr{L}_\chi$ by

$$\Psi(u \otimes \theta \otimes \psi) = (\mathsf{id}_{\mathscr{L}_\chi} \otimes \mathsf{ev})(\psi(u) \otimes \theta).$$

The map $\tilde{\Psi} : \mathscr{L}_\chi \to \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta \otimes W$ defined by $v \mapsto \sum_{i=1}^k \phi_i(v) \otimes \psi_i$ is a right inverse to Ψ. Hence Ψ is a split epimorphism. Since Ψ and $\tilde{\Psi}$ are $(G, \mathbb{C}^\times)$-equivariant we can apply the functor $p_* \mathscr{H}om_{\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_\chi, -)^G$, which by [Kashiwara and Rouquier 2008, Proposition 2.8(ii)] is an equivalence, to the morphism $\mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta \otimes W \to \mathscr{L}_\chi$ to get the required (necessarily split, epic) morphism.

Part (ii) is similar. Again using Proposition 3.5, $\chi \to \chi+\theta$ implies that

$$\mathsf{E}_{\chi+\theta}^{-\theta} \otimes \mathsf{E}_\chi^\theta \twoheadrightarrow \mathsf{E}_{\chi+\theta}.$$

Therefore there exists some k and $\phi_i \in \mathsf{E}_{\chi+\theta}^{-\theta}$, $\psi_i \in \mathsf{E}_\chi^\theta$ for $i \in [1,k]$ such that

$$(\mathsf{id}_{\mathscr{L}_{\chi+\theta}} \otimes \mathsf{ev}) \circ \left(\sum_{i=1}^k (\psi_i \otimes \mathsf{id}_{\mathbb{C}_{-\theta}}) \circ \phi_i \right) = \mathsf{id}_{\mathscr{L}_{\chi+\theta}}.$$

Let $U = \mathrm{Span}_{\mathbb{C}}\{\psi_i : i \in [1,k]\}$ and define $\Phi : \mathscr{L}_\chi \otimes U \to \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta$ by $\Phi(u \otimes \psi) = \psi(u)$. The map $\tilde{\Phi} : \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta \to \mathscr{L}_\chi \otimes U$ defined by

$$v \mapsto (\mathsf{id}_{\mathscr{L}_\chi} \otimes \mathsf{id}_U \otimes \mathsf{ev}) \left(\sum_{i=1}^k \phi_i(v) \otimes \psi_i \right)$$

is a right inverse to Φ. Hence Φ is a split epimorphism. Since Φ and $\tilde{\Phi}$ are $(G, \mathbb{C}^\times)$-equivariant we can apply $p_* \mathscr{H}om_{\widetilde{\mathscr{W}}_{\mathfrak{X}}}(\mathscr{L}_\chi, -)^G$ to the morphism

$$\mathscr{L}_\chi \otimes U \to \mathscr{L}_{\chi+\theta} \otimes \mathbb{C}_\theta$$

to get the required (necessarily split, epic) morphism. $\square$

Proof of Theorem 3.3. It follows from the equivalence in [Kashiwara and Rouquier 2008, Proposition 2.8(iv)] that $\Gamma(Y_\vartheta, \mathscr{A}_\chi)^F = \mathsf{E}_\chi$. Therefore part (i) follows from Proposition 3.5. Lemma 3.8 and Theorem 3.7 show that $\chi \leftarrow \chi + n\theta$ for all $n \in \mathbb{Z}_{\geq 0}$ implies that $\mathbb{R}^i f_*(\mathscr{M}) = 0$ for all $i > 0$ and all $\mathscr{M} \in \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$. Similarly,

$\chi \to \chi + n\theta$ for all $n \in \mathbb{Z}_{\geq 0}$ implies that every good $\mathscr{A}_\chi$-module is generated, locally on Y_0, by its global sections. Let o denote the image of the origin of T^*V in Y_0. The $\mathbb{C}^\times$-action we have defined on Y_0 (via the $\mathbb{C}^\times$-action on T^*V) shrinks every point to o, in the sense that $\lim_{t\to\infty} T_t(y) = o$ for all $y \in Y_0$. In such a situation, [ibid., Lemma 2.13] says that $\mathbb{R}^i f_*(\mathscr{M}) = 0$ for all $i > 0$ and all $\mathscr{M} \in \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ implies that $\mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, -)$ is an exact functor. Similarly, [ibid., Lemma 2.14] says that if every good $\mathscr{A}_\chi$-module $\mathscr{M}$ is generated, locally on Y_0, by its global sections then every $\mathscr{M}$ is generated by its F-invariant global sections. That is,

$$\mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} \mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \mathscr{M}) \twoheadrightarrow \mathscr{M}.$$

With these facts, one can follow the proof of [Hotta et al. 2008, Corollary 11.2.6], more or less word for word. □

4. Hypertoric varieties

4A. As we have seen in the previous section, when one has a reductive group G acting on a vector space V, there exists a family of W-algebras on the Hamiltonian reduction of the cotangent bundle of V. The simplest such situation is where $G = \mathbb{T}$, a d-dimensional torus. In this case the corresponding Hamiltonian reduction is called a hypertoric variety. In this section we recall the definition of, and basic facts about, hypertoric varieties. The reader is advised to consult [Proudfoot 2008] for an excellent introduction to hypertoric varieties. Here we will follow the algebraic presentation given in [Hausel and Sturmfels 2002]. Thus, in this section only, spaces will be algebraic varieties over $\mathbb{C}$ in the Zariski topology.

4B. ***Torus actions.*** Fix $1 \leq d < n \in \mathbb{N}$ and let $\mathbb{T} := (\mathbb{C}^\times)^d$. We consider $\mathbb{T}$ acting algebraically on the n-dimensional vector space V. If we fix coordinates on V such that the corresponding coordinate functions $x_1, \dots, x_n$ are eigenvectors for $\mathbb{T}$ then the action of $\mathbb{T}$ is encoded by a $d \times n$ integer valued matrix

$$A = [a_1, \dots, a_n] = (a_{ij})_{i\in[1,d],\, j\in[1,n]},$$

and is given by $(\xi_1, \dots, \xi_d) \cdot x_i = \xi_1^{a_{1i}} \cdots \xi_d^{a_{di}} x_i$ for all $(\xi_1, \dots, \xi_d) \in \mathbb{T}$. We fix the coordinate ring of V to be $R := \mathbb{C}[x_1, \dots, x_n]$. The algebra R is graded by the action of $\mathbb{T}$, $\deg(x_i) = a_i$. We make the assumption that the $d \times d$ minors of A are relatively prime. This ensures that the map $\mathbb{Z}^n \xrightarrow{A} \mathbb{Z}^d$ is surjective and hence the stabilizer of a generic point is trivial.

4C. Since $\mathbb{Z}^d$ is a free $\mathbb{Z}$-module, the above assumption implies that we can choose an $n \times (n-d)$ integer valued matrix $B = [b_1, \dots, b_n]^T$ so that the following sequence is exact:

$$0 \longrightarrow \mathbb{Z}^{n-d} \xrightarrow{B} \mathbb{Z}^n \xrightarrow{A} \mathbb{Z}^d = \mathbb{X} \longrightarrow 0, \tag{9}$$

where, as before, $\mathbb{X} := \mathrm{Hom}_{gp}(\mathbb{T}, \mathbb{C}^\times)$ is the character lattice of $\mathbb{T}$ and $\mathbb{Z}^n$ is identified with the character lattice of $(\mathbb{C}^\times)^n \subset \mathrm{GL}(\mathbb{C}^n)$. The dual $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{X}, \mathbb{Z})$ of $\mathbb{X}$, which parametrizes one-parameter subgroups of $\mathbb{T}$, will be denoted $\mathbb{Y}$. Applying the functor $\mathrm{Hom}(\,\cdot\,, \mathbb{C}^\times)$ to the sequence (9) gives a short exact sequence of abelian groups

$$1 \longrightarrow \mathbb{T} \xrightarrow{A^T} (\mathbb{C}^\times)^n \xrightarrow{B^T} (\mathbb{C}^\times)^{n-d} \longrightarrow 1. \tag{10}$$

Let $\mathfrak{t}$ denote the Lie algebra of $\mathbb{T}$ and $\mathfrak{g}$ the Lie algebra of $(\mathbb{C}^\times)^n$. Differentiating the sequence (10) produces the short exact sequence

$$0 \longrightarrow \mathfrak{t} \xrightarrow{A^T} \mathfrak{g} \xrightarrow{B^T} \mathrm{Lie}(\mathbb{C}^\times)^{n-d} \longrightarrow 0 \tag{11}$$

of abelian Lie algebras.

4D. *Geometric invariant theory.* The standard approach to defining "sensible" algebraic quotients of V by $\mathbb{T}$ is to use geometric invariant theory. We recall here the basic construction that will be used. Let $\mathbb{X}_{\mathbb{Q}} := \mathbb{X} \otimes_{\mathbb{Z}} \mathbb{Q}$ be the space of fractional characters. We fix a stability parameter $\delta \in \mathbb{X}_{\mathbb{Q}}$. For $\underline{k} = (k_1, \dots, k_n) \in \mathbb{N}^n$, the monomial $x_1^{k_1} \cdots x_n^{k_n}$ will be written $x^{\underline{k}}$. Then $\lambda \cdot x^{\underline{k}} = \lambda^{A \cdot \underline{k}} x^{\underline{k}}$ and we define

$$R^\delta := \mathrm{Span}_{\mathbb{C}}(x^{\underline{k}} \mid A \cdot \underline{k} = \delta)$$

to be the space of $\mathbb{T}$-semiinvariants of weight δ. Note that $R^\delta = 0$ if $\delta \notin \mathbb{X}$. A point $p \in V$ is said to be δ-*semistable* if there exists an $n > 0$ such that $n\delta \in \mathbb{X}$ and $f \in R^{n\delta}$ with $f(p) \neq 0$. A point p is called δ-*stable* if it is δ-semistable and in addition its stabilizer under $\mathbb{T}$ is finite. The set of δ-semistable points in V will be denoted V_δ^{ss}. The parameter δ is said to be *effective* if $R^{n\delta} \neq 0$ for some $n > 0$ (by the Nullstellensatz this is equivalent to $V_\delta^{\mathrm{ss}} \neq \varnothing$).

Definition 4.1. Let $\delta \in \mathbb{X}_{\mathbb{Q}}$ be an effective stability condition. The *G.I.T quotient* of V by $\mathbb{T}$ with respect to δ is the variety

$$X(A, \delta) := \mathrm{Proj} \bigoplus_{k \geq 0} R^{k\delta};$$

it is projective over the affine quotient $X(A, 0) := \mathrm{Spec}(R^{\mathbb{T}})$.

If a point $p \in V$ is not δ-semistable it is called δ-*unstable*. Using the one-parameter subgroups of $\mathbb{T}$ one can describe the set V_δ^{us} of δ-unstable points. We denote by $\langle\,\cdot\,,\cdot\,\rangle$ the natural pairing between $\mathbb{Y}$ and $\mathbb{X}$ (and by extension between $\mathfrak{t}$ and $\mathfrak{t}^*$). Let $V(f_1, \dots, f_k)$ denote the set of common zeros of the polynomials $f_1, \dots, f_k \in R$.

Lemma 4.2. *Let $\delta \in \mathbb{X}_{\mathbb{Q}}$ be an effective stability parameter. The δ-unstable locus is*

$$V_\delta^{\mathrm{us}} = \bigcup_{\substack{\lambda \in \mathbb{Y} \\ \langle \lambda, \delta \rangle < 0}} V(x_i \mid \langle \lambda, a_i \rangle < 0). \tag{12}$$

Moreover, there exists a finite set $\mathscr{F}(\delta) = \{\lambda_1, \dots, \lambda_k\} \subset \mathbb{Y}$, $\langle \lambda_i, \delta \rangle < 0$ such that

$$\bigcup_{\substack{\lambda \in \mathbb{Y} \\ \langle \lambda, \delta \rangle < 0}} V(x_i \mid \langle \lambda, a_i \rangle < 0) = \bigcup_{\lambda \in \mathscr{F}(\delta)} V(x_i \mid \langle \lambda, a_i \rangle < 0).$$

Proof. Let $S := R[t]$ and extend the action of $\mathbb{T}$ from R to S by setting $g \cdot t = \delta(g)^{-1} t$ for all $g \in \mathbb{T}$. Then $(S)^{\mathbb{T}} = \bigoplus_{n \geq 0} R^{n\delta} \cdot t^n$. Now

$$\begin{aligned} u \in V_\delta^{\mathrm{us}} &\iff f(u) = 0 \quad \text{for all } f \in R^{n\delta}, n > 0, \\ &\iff F(u, 1) = 0 \quad \text{for all } F \in (S^{\mathbb{T}})_+ = (S_+)^{\mathbb{T}}, \\ &\iff \overline{\mathbb{T} \cdot (u, 1)} \cap V \times \{0\} \neq \varnothing, \end{aligned}$$

where $(S^{\mathbb{T}})_+ = (S_+)^{\mathbb{T}}$ follows from the fact that $\mathbb{T}$ is reductive. Then [Kempf 1978, Theorem 1.4] says that there exists a one-parameter subgroup $\lambda \in \mathbb{Y}$ such that $\lim_{t \to 0} \lambda(t) \cdot (u, 1) \in V \times \{0\}$. Writing $u = u_1 + \dots + u_n$ such that $x_i(u) = u_i$, we have

$$\lambda(t) \cdot (u, 1) = \left(\sum_{i=1}^n t^{-\langle \lambda, a_i \rangle} u_i, t^{\langle \lambda, \delta \rangle} \right),$$

which implies that $u_i = 0$ for all $i \in [1, n]$ such that $\langle \lambda, a_i \rangle > 0$ and $\langle \lambda, \delta \rangle > 0$. This shows that the left hand side of (12) is contained in the right hand side. Conversely, if u is δ-semistable then it is also ϕ-semistable with respect to the action of the one dimensional torus $\lambda : \mathbb{T} \hookrightarrow \mathbb{T}$ on V, where ϕ is the character of $\mathbb{T}$ defined by $t \mapsto t^{\langle \lambda, \delta \rangle}$. □

4E. The variety $X(A, \delta)$ is a toric variety and, as shown in [Hausel and Sturmfels 2002, Corollary 2.7], any semiprojective toric variety equipped with a fixed point is isomorphic to $X(A, \delta)$ for suitable A and δ. Fix $S \subset V$ and let $\delta_1, \delta_2 \in \mathbb{X}_{\mathbb{Q}}$ be two stability parameters such that $S_{\delta_1}^{\mathrm{ss}}, S_{\delta_2}^{\mathrm{ss}} \neq 0$. Then δ_1 and δ_2 are said to be equivalent if $S_{\delta_1}^{\mathrm{ss}} = S_{\delta_2}^{\mathrm{ss}}$. The set of all ρ equivalent to a fixed δ will be denoted $C(\delta)$. These equivalence classes form the relative interiors of the cones of a rational polyhedral fan $\Delta(\mathbb{T}, S)$, called the G.I.T. fan, in $\mathbb{X}_{\mathbb{Q}}$. The support of $\Delta(\mathbb{T}, S)$ is the set of all effective $\delta \in \mathbb{X}_{\mathbb{Q}}$ such that $S_\delta^{\mathrm{ss}} \neq 0$ and is denoted $|\Delta(\mathbb{T}, S)|$. We will mainly be concerned with $S = V$. The cones in $\Delta(\mathbb{T}, V)$ having the property that the stable locus is properly contained in the semistable locus are called the walls of $\Delta(\mathbb{T}, V)$. The G.I.T. fan is quite difficult to describe explicitly; see [Oda and Park 1991]. However one has the following explicit description of the walls of $\Delta(\mathbb{T}, V)$.

Lemma 4.3. *Let $\mathbb{T}$ act on V via A as in Section 4B. Then $|\Delta(\mathbb{T}, V)| = \sum_{i=1}^{n} \mathbb{Q}_{\geq 0} \cdot a_i$ and the walls of the fan are $\sum_{i \in J} \mathbb{Q}_{\geq 0} \cdot a_i$, where $J \subset [1, n]$ is any subset such that* $\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(a_i \mid i \in J)) = d - 1$.

Proof. Let $0 \neq \delta \in \sum_{i=1}^{n} \mathbb{Z}_{\geq 0} \cdot a_i$ and write $\delta = \sum_{i \in I} n_i a_i$ where $I \subset [1, n]$ and $n_i > 0$ for all $i \in I$. Then $0 \neq f = \prod_{i \in I} x_i^{n_i} \in R^{\delta}$ implies that δ is effective. Now let $\delta \in \mathbb{X}$ be any effective stability parameter and choose $0 \neq p \in V_{\delta}^{\mathrm{ss}}$. Write $p = p_1 + \cdots + p_n$ so that $x_i(p) = p_i$ and let $I = \{i \in [1, n] \mid p_i \neq 0\}$. Then Lemma 4.2 shows that

$$\begin{aligned} p \in V_{\delta}^{\mathrm{ss}} &\iff (\langle \lambda, \delta \rangle < 0 \Rightarrow \text{there exists } i \in I \text{ such that } \langle \lambda, a_i \rangle < 0) \\ &\iff \{\lambda \in \mathbb{Y} \mid \langle \lambda, \delta \rangle < 0\} \cap \Big(\sum_{i \in I} \mathbb{Z}_{\geq 0} \cdot a_i\Big)^{\vee} = \varnothing \\ &\iff \Big(\sum_{i \in I} \mathbb{Z}_{\geq 0} \cdot a_i\Big)^{\vee} \subset \{\lambda \in \mathbb{Y} \mid \langle \lambda, \delta \rangle \geq 0\} \\ &\iff \delta \in \sum_{i \in I} \mathbb{Z}_{\geq 0} \cdot a_i. \end{aligned}$$

Now choose $\delta \in \mathbb{X}$ to lie on a wall. By definition, there exists a δ-semistable point p such that $\dim \mathrm{Stab}_{\mathbb{T}}(p) \geq 1$. Let I be as above. Then $\dim \mathrm{Stab}_{\mathbb{T}}(p) \geq 1$ implies that the subspace $\sum_{i \in I} \mathbb{Q} \cdot a_i$ must be a proper subspace of $\mathbb{X}_{\mathbb{Q}}$. The above reasoning shows that $\delta \in \sum_{i \in I} \mathbb{Z}_{\geq 0} \cdot a_i$ as required. □

Lemma 4.3 shows that, under our assumption on A, the maximal cones of $\Delta(\mathbb{T}, V)$ are all d-dimensional. We will refer to these maximal cones as the *d-cones* of $\Delta(\mathbb{T}, V)$. The integer valued matrix A is said to be *unimodular*[1] if every $d \times d$ minor of A takes values in $\{-1, 0, 1\}$. Combining [Hausel and Sturmfels 2002, Corollary 2.7 and Corollary 2.9] gives the following theorem:

Theorem 4.4. *The variety $X(A, \delta)$ is an orbifold if and only if δ belongs to the interior of a d-cone of $\Delta(\mathbb{T}, V)$. It is a smooth variety if and only if δ belongs to the interior of a d-cone of $\Delta(\mathbb{T}, V)$ and A is unimodular.*

4F. *Hypertoric varieties.* Define $A^{\pm} := [A, -A]$, a $d \times 2n$ matrix. It defines a grading on the ring $R := \mathbb{C}[T^*V] = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. For δ in the interior of a d-cone of $\Delta(\mathbb{T}, T^*V)$, the corresponding toric variety $X(A^{\pm}, \delta)$ is called a Lawrence toric variety associated to A. It is a G.I.T. quotient of the symplectic vector space T^*V with canonical symplectic form

$$\omega = dx_1 \wedge dy_1 + \cdots + dx_n \wedge dy_n.$$

[1] In [Hausel and Sturmfels 2002], the authors define A to be unimodular if every nonzero $d \times d$ minor of A has the same absolute value. However they also, as do we, make the assumption that the $d \times d$ minors of A are relatively prime. Thus, their definition agrees with ours.

The action of $\mathbb{T}$ is Hamiltonian and the moment map is given by

$$\mu : T^*V \longrightarrow \mathfrak{t}^*, \quad \mu(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{j=1}^{n} a_{ij} x_j y_j \right)_{i \in [1,d]}.$$

Consider the ideal

$$I := I(\mu^{-1}(0)) = \left\langle \sum_{j=1}^{n} a_{ij} x_j y_j \;\middle|\; i \in [1, d] \right\rangle \subset R;$$

it is homogeneous and generated by $\mathbb{T}$-invariant polynomials.

Definition 4.5. The hypertoric variety associated to A and δ is defined to be

$$Y(A, \delta) := \mu^{-1}(0) /\!/_{\delta} \mathbb{T} = \operatorname{Proj} \bigoplus_{k \geq 0} (R/I)^{k\delta};$$

it is projective over the affine quotient $Y(A, 0) := \operatorname{Spec}((R/I)^{\mathbb{T}})$.

The basic properties of hypertoric varieties can be summarized as follows:

Proposition 4.6 [Hausel and Sturmfels 2002, Proposition 6.2]. *If δ is in the interior of a d-cone of $\Delta(\mathbb{T}, \mu^{-1}(0))$ then the hypertoric variety $Y(A, \delta)$ is an orbifold. It is smooth if and only if δ is in the interior of a d-cone of $\Delta(\mathbb{T}, \mu^{-1}(0))$ and A is unimodular.*

4G. In this subsection we show that the assumptions of Section 3A are valid for hypertoric varieties. Let $f : Y(A, \delta) \to Y(A, 0)$ be the projective morphism from $Y(A, \delta)$ to $Y(A, 0)$. Lemma 4.9 below together with Proposition 4.6 implies that f is birational and hence a resolution of singularities when $Y(A, \delta)$ is smooth. The symplectic form ω on T^*V induces a symplectic 2-form on the smooth locus of $Y(A, \delta)$. In particular, when $Y(A, \delta)$ is smooth it is a symplectic manifold. Proposition 4.11 below shows that $Y(A, 0)$ is a symplectic variety and the resolution f is a symplectic resolution.[2] This implies that $Y(A, 0)$ is also normal.

Lemma 4.7. *The moment map is flat and $\mu^{-1}(0)$ is a reduced complete intersection in T^*V. If no row of the matrix B is zero then $\mu^{-1}(0)$ is irreducible.*

Proof. The graded lexicographic ordering on a monomial $\underline{x}^{\alpha}$, $\alpha \in \mathbb{Z}^{2n}$, is defined by saying that $\underline{x}^{\alpha} > \underline{x}^{\beta}$ if and only if $|\alpha| > |\beta|$, or $|\alpha| = |\beta|$ and the leftmost nonzero entry of $\alpha - \beta$ is positive; see [Cox et al. 2007, page 58]. After permuting the variables $x_1, \dots, x_n$, we may assume that the first d columns of A are linearly independent. Applying an automorphism of $\mathbb{T}$ and then letting $\mathbb{T}$ act is the same as multiplying A on the right by some unimodular $d \times d$-matrix. Using this fact we

[2]We refer the reader to [Fu 2006] for the definition of symplectic variety and symplectic resolution.

may assume that the leftmost $d \times d$-block of A is the identity matrix. This allows us to rewrite the generators of I as

$$\left\{ x_i y_i - \sum_{j=d+1}^{n} c_{i,j} x_j y_j \;\middle|\; i = 1, \dots, d \right\}, \quad \text{where } c_{i,j} \in \mathbb{Z}.$$

By [Cox et al. 2007, Theorem 8, page 461], $\dim \mu^{-1}(0) = \dim V(\mathrm{in}(I))$, where $\mathrm{in}(I)$ denotes the initial ideal of I with respect to the ordering $x_1 > x_2 > \cdots > y_1 > y_2 \cdots$. Now by [Cox et al. 2007, Theorem 3 (Division algorithm), page 64], we have $\mathrm{in}(I) = \langle x_1 y_1, \dots, x_d y_d \rangle$. This is the zero set of a union of 2^d linear subspaces of T^*V of dimension $2n - d$. Therefore $\dim \mu^{-1}(0) = 2n - d$ and it follows from [Holland 1999, Lemma 2.3] that the moment map is flat. To prove that it is a complete intersection we must show that the generators of I given above form a regular sequence in the polynomial ring R. Once again, it suffices to note that $x_1 y_1, \dots, x_d y_d$ is a regular sequence. Also, since the ideal $\mathrm{in}(I)$ is radical, the ideal I is itself radical.

Now note that, since the sequence (9) is exact, the matrix B contains a row of zeros if and only if there exists an $i \in [1, .., d]$ such that $c_{i,j} = 0$ for all $j > d$. So, when B contains no rows equal to zero we can write

$$y_i = x_i^{-1} \sum_{j=d+1}^{n} c_{i,j} x_j y_j \mod I$$

on the open set $\mu^{-1}(0) \backslash V(x_1 \cdots x_d)$. This shows that $\mu^{-1}(0)$ contains an open set isomorphic to $\mathbb{A}^{2n-d}$. We just need to show that this open set is dense. Since $\mu^{-1}(0)$ is a complete intersection, it is pure dimensional. Therefore it suffices to show that the dimension of $\mu^{-1}(0) \cap V(x_1 \cdots x_d)$ is at most $2n - d - 1$. Consider $Y = \mu^{-1}(0) \cap V(x_1)$ and let $J = I(Y)$. We may assume without loss of generality that $c_{1,d+1} \neq 0$. Then J is generated by $x_1, x_i y_i - \sum_{j=d+2}^{n} c_{i,j} x_j y_j$ for $j = 2, \dots, d$ and $x_{d+1} y_{d+1} + \sum_{j=d+2}^{n} c_{1,j} x_j y_j$. Hence $\mathrm{in}(J) = \langle x_1, x_2 y_2, \dots, x_{d+1} y_{d+1} \rangle$, which defines a variety of dimension $2n - d - 1$ as required. □

From now on we assume that no row of the matrix B is zero.

Lemma 4.8. *For any A, we have* $\dim X(A^{\pm}, 0) = 2n - d$.

Proof. Let $U = V \backslash V(x_1 \cdots x_n)$ and let $S_1 = \mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]^{\mathbb{T}}$ denote the coordinate ring of the quotient $U/\mathbb{T}$. Let F_1 be the field of fractions of S_1. Let $S_2 = \mathbb{C}[X(A^{\pm}, 0)]$ and F_2 its field of fractions. We claim that $F_1 \subset F_2$. An element in F_1 is a fraction $f(x_1, \dots, x_n)/g(x_1, \dots, x_n)$, where f and g are homogeneous of the same weight with respect to $\mathbb{T}$. Then $f(\boldsymbol{x}) f(\boldsymbol{y}), g(\boldsymbol{x}) f(\boldsymbol{y}) \in S_2$ and $f(\boldsymbol{x}) f(\boldsymbol{y})/g(\boldsymbol{x}) f(\boldsymbol{y}) = f(\boldsymbol{x})/g(\boldsymbol{x})$ as required. Since $\dim \mathbb{T}^n/\mathbb{T} = n - d$, to prove the lemma it suffices to show that the field extension $F_1 \subset F_2$ has transcendental degree n. Consider the field $K = F_1\langle x_1 y_1, \dots, x_n y_n\rangle$. Then $F_1 \subset K \subset F_2$ and K is a purely transcendental extension of F_1 of degree n. We claim that $K = F_2$. To show this it is sufficient to show that if $f \in S_2$ is a polynomial in the x_i and y_j then $f \in K$. We show more generally that if $f = f_1/g$, where $f_1, g \in S_2$ and g a monomial, then $f \in K$. We prove the claim by induction on the number of terms in f (note that even though there is some choice in the exact form of each of the terms in f, the number of terms is unique). Let $u = \alpha x^{\boldsymbol{i}} y^{\boldsymbol{j}}$, $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{Z}^n$, be some nonzero term of f. Then $(xy)^{-\boldsymbol{j}} u \in F_1$ and $(xy)^{-\boldsymbol{j}} f - (xy)^{-\boldsymbol{j}} u \in K$ by induction. Since $(xy)^{-\boldsymbol{j}} \in K$, this implies that $f \in K$. □

Note that, unlike $X(A^{\pm}, 0)$, the dimension of $X(A, 0)$ can vary greatly depending on the specific entries of A.

Lemma 4.9. *For any A, we have $\dim Y(A, 0) = 2(n - d)$ and $Y(A, 0)$ is Cohen–Macaulay.*

Proof. By Hochster's Theorem [Bruns and Herzog 1993, Theorem 6.4.2], the ring $\mathbb{C}[X(A^{\pm}, 0)]$ is Cohen–Macaulay. As noted in Lemma 4.7, the generators $u_1, \dots, u_d$ of I form a regular sequence in R. Since $\mathbb{T}$ is reductive,

$$R = \mathbb{C}[X(A^{\pm}, 0)] \oplus E$$

as a $\mathbb{C}[X(A^{\pm}, 0)]$-module. Therefore projection from R to $\mathbb{C}[X(A^{\pm}, 0)]$ is a Reynolds operator in the sense of [ibid., page 270]. Since $u_1, \dots, u_d$ are $\mathbb{T}$-invariant, [ibid., Proposition 6.4.4] now says that they form a regular sequence in $X(A^{\pm}, 0)$. Therefore [ibid., Theorem 2.1.3] says that $Y(A, 0)$ is Cohen–Macaulay with $\dim Y(A, 0) = \dim X(A^{\pm}, 0) - d$. The lemma follows from Lemma 4.8. □

Lemma 4.10. *Let $\delta \in \mathbb{X}_{\mathbb{Q}}$. The graded ring $\bigoplus_{k \geq 0} (R/I)^{k\delta}$ is Cohen–Macaulay, that is, $Y(A, \delta)$ is arithmetically Cohen–Macaulay.*

Proof. Consider $S = R[t]$ with $\mathbb{T}$ acting on t via $g \cdot t = \delta(g)^{-1} t$. Replacing R with S in the proof of Lemma 4.9 gives a proof of the statement. □

Proposition 4.11. *Let A be unimodular and choose δ in the interior of a d-cone of $\Delta(\mathbb{T}, \mu^{-1}(0))$. Then $Y(A, 0)$ is a symplectic variety and the morphism $f : Y(A, \delta) \to Y(A, 0)$ is a symplectic resolution.*

Proof. The construction of $Y(A, \delta)$ and $Y(A, 0)$ as Hamiltonian reductions means that they are Poisson varieties and f preserves the Poisson structure. Therefore the smooth locus of $Y(A, 0)$ is a symplectic manifold since $Y(A, \delta)$ is a symplectic manifold. In [Proudfoot and Webster 2007, §2], a stratification of $Y(A, 0)$ into smooth locally closed subvarieties of even dimensions is constructed. This stratification shows that $Y(A, 0)$ is smooth in codimension one. Therefore the fact (Lemma 4.9) that $Y(A, 0)$ is Cohen–Macaulay together with Serre's normality criterion [Bruns and Herzog 1993, Theorem 2.2.22] implies that $Y(A, 0)$ is normal. Also, the fact that $Y(A,\delta)$ is a symplectic manifold implies that its canonical bundle is trivial. Therefore the Grauert–Riemenschneider vanishing theorem implies that $Y(A,0)$ has rational Gorenstein singularities. Then [Namikawa 2001, Theorem 6] says that $Y(A,0)$ is a symplectic variety. $\square$

4H. ***G.I.T. chambers for hypertoric varieties.*** Define the subvariety $\mathscr{E}$ of T^*V by $\mathscr{E} = \{(x,y) \in T^*V \mid x_i \cdot y_i = 0 \text{ for all } i \in [1,n]\}$. We decompose $\mathscr{E}$ into its n-dimensional irreducible components

$$\mathscr{E} = \bigcup_{I \subset [1,n]} \mathscr{E}_I, \quad \mathscr{E}_I := \{(x,y) \in T^*V \mid x_i = 0 \text{ for all } i \in I,\ y_i = 0 \text{ for all } i \in [1,n]\backslash I\}.$$

The subvariety $\mathscr{E}$ is preserved under the $\mathbb{T}$-action. Therefore we may consider the corresponding G.I.T. quotients. The G.I.T. quotient $\mathscr{E}/\!/_\delta \mathbb{T}$ is a closed subvariety of $Y(A,\delta)$, called the *extended core* of $Y(A,\delta)$; see [Proudfoot 2008] for details.

Lemma 4.12. *In $\mathbb{X}_{\mathbb{Q}}$ we have equalities of G.I.T. fans*

$$\Delta(\mathbb{T}, T^*V) = \Delta(\mathbb{T}, \mathscr{E}) = \Delta(\mathbb{T}, \mu^{-1}(0)).$$

Proof. For a fixed $I \subset [1,n]$, denote by $\pi_I : T^*V \twoheadrightarrow \mathscr{E}_I$ the projection that sends x_i to zero if $i \in I$ and y_j to zero if $j \in [1,n]\backslash I$. The restriction of π_I to $\mu^{-1}(0)$ will be denoted $\widetilde{\pi}_I$. The statement of lemma follows from the claim

$$(T^*V)^{\mathrm{ss}}_\delta = \bigcup_{I \subset [1,n]} \pi_I^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta) \quad \text{and} \quad (\mu^{-1}(0))^{\mathrm{ss}}_\delta = \bigcup_{I \subset [1,n]} (\widetilde{\pi}_I)^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta) \tag{13}$$

for each $\delta \in \mathbb{X}$. Let $p \in (\mathscr{E}_I)^{\mathrm{ss}}_\delta$. Then, without loss of generality, we may assume that there exists a monomial $f \in R^{N\delta}$, $N \geq 1$, such that $f(p) \neq 0$. Then $f(q) \neq 0$ for all $q \in \pi_I^{-1}(p)$. Hence $(T^*V)^{\mathrm{ss}}_\delta \supset \pi_I^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta)$ and $(\mu^{-1}(0))^{\mathrm{ss}}_\delta \supset (\widetilde{\pi}_I)^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta)$ for all $I \subset [1,n]$. Now choose $p \in (T^*V)^{\mathrm{ss}}_\delta$. Then there exist $m \in \mathbb{N}$ and $g \in R^{m\delta}$ such that $g(p) \neq 0$. We may assume without loss of generality that

$$g = \prod_i x_i^{u_i} \prod_i y_i^{v_i} \quad \text{for some } u_i, v_i \geq 0.$$

By definition, $\sum_i (u_i - v_i) a_i = m\delta$. For each i, define s_i and t_i by

(1) $u_i - v_i > 0 \implies s_i = u_i - v_i,\ t_i = 0;$

(2) $u_i - v_i < 0 \implies t_i = v_i - u_i,\ s_i = 0;$

(3) $u_i - v_i = 0 \implies s_i = t_i = 0.$

Set $I = \{i \in [1,n] \mid t_i \neq 0\}$. Then $g(p) \neq 0$ implies that $\pi_I(p) \neq 0$. Define $\tilde{g} = \prod_i x_i^{s_i} \prod_i y_i^{t_i} \in R^{m\delta}$. Then $\tilde{g}(\pi_I(p)) \neq 0$ implies that $p \in \pi_I^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta)$ and hence $(T^*V)^{\mathrm{ss}}_\delta = \bigcup_{I \subset [1,n]} \pi_I^{-1}((\mathscr{E}_I)^{\mathrm{ss}}_\delta)$ as required. The second equality in (13) follows from the first one. $\square$

Corollary 4.13. *Let $\mathbb{T}$ act on V via A as in Section 4B. Then $|\Delta(\mathbb{T},\mu^{-1}(0))| = \mathbb{X}_{\mathbb{Q}}$ and the walls of the fan $\Delta(\mathbb{T},\mu^{-1}(0))$ are $\sum_{i\in J}\mathbb{Q}\cdot a_i$, where $J\subset[1,n]$ is any subset such that* $\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(a_i \mid i\in J)) = d-1$.

Assume now that A is unimodular and choose $\delta\in\mathbb{X}$ to lie in the interior, denoted $C(\delta)$, of a d-cone of $\Delta(\mathbb{T},\mu^{-1}(0))$. If $\zeta\in C(\delta)\cap\mathbb{X}$ then $Y(A,\delta) = Y(A,\zeta)$. Recall from Section 3A that ζ also defines a line bundle L_ζ on $Y(A,\delta)$. From the definition of $Y(A,\delta)$ as proj of a graded ring, we see that L_ζ is an ample line bundle on $Y(A,\delta)$. Summarizing:

Lemma 4.14. *Let A be unimodular and let $C(\delta)$ denote the interior of a d-cone of $\Delta(\mathbb{T},\mu^{-1}(0))$. Then the line bundle L_ζ on $Y(A,\delta)$ is ample for all $\zeta\in C(\delta)\cap\mathbb{X}$.*

5. Quantum Hamiltonian reduction

5A. Recall that $\mathfrak{D}(V)$ denotes the ring of algebraic differential operators on the n-dimensional space V. Let $\mathbb{T}$ act on V with weights described by the matrix A (as in Sections 4B and 4C) and choose an element χ of the dual $\mathfrak{t}^*$ of the Lie algebra $\mathfrak{t}$ of $\mathbb{T}$. As explained in Section 3D, by differentiating the action of $\mathbb{T}$ we get a quantum moment map $\mu_D:\mathfrak{t}\to\mathfrak{D}(V)$, $t_i\mapsto\sum_{j=1}^n a_{ij}x_j\partial_j$. As in Section 3C, the quantum Hamiltonian reduction of V with respect to χ is defined to be the noncommutative algebra

$$\mathsf{U}_\chi := \big(\mathfrak{D}(V)\,\big/\,\mathfrak{D}(V)(\mu_D-\chi)(\mathfrak{t})\big)^{\mathbb{T}}.$$

We also have bimodules

$$\mathsf{U}_\chi^\theta := \big(\mathfrak{D}(V)\,\big/\,\mathfrak{D}(V)(\mu_D-(\chi+\theta))(\mathfrak{t})\big)^{-\theta}.$$

We say that χ and $\chi+\theta$ are comparable if the multiplication map $\mathsf{U}_\chi^\theta\otimes\mathsf{U}_{\chi+\theta}^{-\theta}\to\mathsf{U}_\chi$ is nonzero. By [Musson and Van den Bergh 1998, Theorem 7.3.1], the ring U_χ is a domain. Then [ibid., Proposition 4.4.2] says that this implies that comparability is an equivalence relation. As in Section 3C, write $\chi\to\chi+\theta$ if $\mathsf{U}_{\chi+\theta}^{-\theta}\otimes\mathsf{U}_\chi^\theta\twoheadrightarrow\mathsf{U}_{\chi+\theta}$. As noted in [ibid., Remark 4.4.3], the relation $\to$ is transitive. Therefore it defines a preorder on the set of elements in $\mathfrak{t}^*$ comparable to χ. We say that χ is *maximal* if χ is maximal in this preordering, that is, $\chi'\to\chi$ implies $\chi\to\chi'$.

5B. ***The main results.*** Write $\mathsf{pr}:\mathbb{C}\to\mathbb{Q}$ for the $\mathbb{Q}$-linear projection onto $\mathbb{Q}$ and denote by the same symbol the corresponding extension to $\mathfrak{t}^*$:

$$\mathsf{pr}:\mathfrak{t}^* = \mathbb{X}\otimes_{\mathbb{Z}}\mathbb{C}\longrightarrow\mathbb{X}_{\mathbb{Q}}.$$

We also write pr for the map $\mathbb{C}^n = \mathbb{C}\otimes_{\mathbb{Z}}\mathbb{Z}^n\to\mathbb{Q}^n$. Then $\mathsf{pr}(A\cdot v) = A\cdot\mathsf{pr}(v)$ for all $v\in\mathbb{C}^n$. The following proposition is the key to proving our main result. Its proof is given in Section 5D.

Proposition 5.1. *Let $C \subset \mathbb{X}_{\mathbb{Q}}$ be the interior of a d-cone in the fan $\Delta(\mathbb{T}, \mu^{-1}(0))$. Choose $\chi \in \mathfrak{t}^*$ such that $\mathsf{pr}(\chi) \in C$. Then there exists a nonempty d-dimensional integral cone $C(\chi) \subset C \cap \mathbb{X} \cup \{0\}$ such that for all $\theta \in C(\chi)$, $\chi \leftarrow \chi + p\theta$ for all $p \in \mathbb{Z}_{\geq 0}$.*

Recall from Sections 4F and 3D that for each $\chi \in \mathfrak{t}^*$ and $\delta \in C$, where C is the interior of a d-cone of $\Delta(\mathbb{T}, \mu^{-1}(0))$, we have defined the sheaf of algebras $\mathscr{A}_\chi$ on the smooth symplectic manifold $Y(A, \delta)$.

Theorem 5.2. *Let $C \subset \mathbb{X}_{\mathbb{Q}}$ be the interior of a d-cone of $\Delta(\mathbb{T}, \mu^{-1}(0))$. Choose $\chi \in \mathfrak{t}^*$ such that $\mathsf{pr}(\chi) \in C$ and choose $\delta \in C$. Let $\mathscr{A}_\chi$ be the corresponding W-algebra on $Y(A, \delta)$.*

(i) *The functor* $\mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \cdot)$ *defines an equivalence of categories*

$$\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi\text{-mod}$$

with quasiinverse $\mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} (\cdot)$.

(ii) *For any $0 \neq \theta \in C(\chi)$, there exists some $N > 0$ such that the functor* $\mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \cdot)$ *defines an equivalence of categories*

$$\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\chi+N\theta}) \xrightarrow{\sim} \mathsf{U}_{\chi+N\theta}\text{-mod}$$

with quasiinverse $\mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} (\cdot)$.

Proof. By Proposition 5.1 we can choose $0 \neq \theta \in C(\chi)$ such that $\chi \leftarrow \chi + p\theta$ for all $p \in \mathbb{Z}_{\geq 0}$. Since $C(\chi) \backslash \{0\} \subset C$, Lemma 4.14 says that θ defines an ample line bundle L_θ on $Y(A, \delta)$ and we have $\mathscr{A}_{\chi,\theta}(0)/\mathscr{A}_{\chi,\theta}(-1/2) \simeq L_{-\theta}$. Then part (i) of the theorem is a particular case of Theorem 3.3 (ii). The proof of Proposition 5.1 shows that we actually have

$$\chi \leftarrow \chi + \theta \leftarrow \chi + 2\theta \leftarrow \cdots.$$

Since the set of all covectors of the oriented matroid defined by A is finite and each $\mathfrak{Q}_\chi$ (which will be defined in Definition 5.9) is a subset of this set, we see that there are only finitely many different $\mathfrak{Q}_\chi$. Therefore we eventually get

$$\chi + N\theta \leftrightarrows \chi + (N+1)\theta \leftrightarrows \cdots$$

for some sufficiently large N. Then part (ii) of the theorem is a particular case of Theorem 3.3(iii). □

Corollary 5.3. *Let $Y(A, \delta), \mathscr{A}_\chi, \mathsf{U}_\chi, \ldots$ be as in Theorem 5.2 (with $\mathsf{pr}(\chi) \in C$). If the global dimension of U_χ is finite then the functor* $\mathrm{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)}(\mathscr{A}_\chi, \cdot)$ *defines an equivalence of categories* $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi\text{-mod}$ *with quasiinverse* $\mathscr{A}_\chi \otimes_{\mathsf{U}_\chi} (\cdot)$.

Proof. By Proposition 5.1 we can choose $0 \neq \theta \in C(\chi)$ such that $\chi \leftarrow \chi + p\theta$ for all $p \in \mathbb{Z}_{\geq 0}$. However, [Musson and Van den Bergh 1998, Theorem 9.1.1] says that χ is maximal if and only if the global dimension of U_χ is finite. Therefore $\chi \leftrightarrows \chi + \theta$ for all $\theta \in C(\chi)$. Then, as in the proof of Theorem 5.2, Theorem 3.3 implies the statement of the corollary. □

It seems natural to conjecture that, for any $\chi \in \mathfrak{t}^*$, U_χ has finite global dimension if and only if $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi\text{-mod}$, where $\mathscr{A}_\chi$ is the corresponding W-algebra, defined on some $Y(A, \delta)$.

5C. ***The case $d = 1$.*** In specific cases it is possible to strengthen Proposition 5.1. One such case is when the torus $\mathbb{T}$ is one-dimensional. Here the sets $\mathfrak{Q}_\chi$ (which will be defined in Definition 5.9) can be explicitly described, as was done in [Van den Bergh 1991]. Since A is assumed to be unimodular and $a_i \neq 0$ for all i we see that $a_i = \pm 1$ for all i. After reordering we may assume that $a_1, \dots, a_k = 1$ and $a_{k+1}, \dots, a_n = -1$. For simplicity let us assume that $n > 1$. Then

$$\begin{aligned}
\mathfrak{Q}_\chi = \{0\} \quad &\Longleftrightarrow \chi \in (\mathbb{C}\backslash\mathbb{Z}) \cup \{k-n+1, k-n+2, \dots, n-k-2, n-k-1\}, \\
\mathfrak{Q}_\chi = \{0, +\} &\Longleftrightarrow \chi \in \mathbb{Z}_{\geq n-k}, \\
\mathfrak{Q}_\chi = \{0, -\} &\Longleftrightarrow \chi \in \mathbb{Z}_{\leq k-n}.
\end{aligned}$$

In this situation $\mathbb{X}_{\mathbb{Q}} = \mathbb{Q}$ and there are two 1-cones with respect to the action of $\mathbb{T}$ on $\mu^{-1}(0)$; they are $\mathbb{Q}_{\geq 0}$ and $\mathbb{Q}_{\leq 0}$. Applying Theorem 3.3 gives:

Proposition 5.4. *Let* $\dim \mathbb{T} = 1$ *and* $n > 1$ *and choose* $\chi \in \mathfrak{t}^*$. *For* $\delta \neq 0$, *let* $\mathscr{A}_\chi$ *denote the corresponding W-algebra on* $Y(A, \delta)$.

(i) *When* $\delta = 1$ *we have an equivalence*

$$\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi\text{-mod}$$

if and only if $\chi \in (\mathbb{C}\backslash\mathbb{Z}) \cup \mathbb{Z}_{\geq 0}$ *and* $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) = \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ *if and only if* $\chi \in (\mathbb{C}\backslash\mathbb{Z}) \cup \mathbb{Z}_{\geq n-k}$.

(ii) *When* $\delta = -1$ *we have an equivalence* $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) \xrightarrow{\sim} \mathsf{U}_\chi\text{-mod}$ *if and only if* $\chi \in (\mathbb{C}\backslash\mathbb{Z}) \cup \mathbb{Z}_{<0}$ *and* $\underline{\mathrm{Mod}}_F^{\mathrm{good}}(\mathscr{A}_\chi) = \mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_\chi)$ *if and only if* $\chi \in (\mathbb{C}\backslash\mathbb{Z}) \cup \mathbb{Z}_{\leq k-n}$.

This result can be viewed as a variant of [Van den Bergh 1991, Theorem 6.1.3], where sufficient conditions for the D-affinity of weighted projective spaces are stated.

5D. The remainder of this section is devoted to the proof of Proposition 5.1. Since $\mathbb{T}$ can be considered as a subgroup of $\mathbb{T}^n$, $\mathfrak{t}$ is a Lie subalgebra of $\mathfrak{g} = \mathrm{Lie}(\mathbb{T}^n)$ and we may regard elements of $\mathfrak{t}$ as linear functionals on $\mathfrak{g}^*$. Let $\rho : \mathfrak{g}^* \twoheadrightarrow \mathfrak{t}^*$ be the natural map.

Definition 5.5. Let $\lambda \in \mathbb{Y}$ and $\theta \in \left(\sum_{\langle\lambda,a_i\rangle=0} \mathbb{C}\cdot a_i\right) / \left(\sum_{\langle\lambda,a_i\rangle=0} \mathbb{Z}\cdot a_i\right)$. We say that the pair (λ,θ) is attached to χ if there exists $\alpha \in \rho^{-1}(\chi)$ such that

$$\sum_{\langle\lambda,a_i\rangle=0} \alpha_i a_i \equiv \theta \mod \sum_{\langle\lambda,a_i\rangle=0} \mathbb{Z}\cdot a_i,$$

and

$$\begin{aligned} \langle\lambda,a_i\rangle > 0 &\implies \alpha_i \in \mathbb{Z}, \alpha_i \geq 0,\\ \langle\lambda,a_i\rangle < 0 &\implies \alpha_i \in \mathbb{Z}, \alpha_i < 0,\\ \langle\lambda,a_i\rangle = 0 &\implies \alpha_i \in \mathbb{C}\backslash\mathbb{Z}. \end{aligned}$$

Remark 5.6. The above definition is based on [Musson and Van den Bergh 1998, Definition 7.2.1]. There it is stipulated that $\lambda \in \mathfrak{t}\cap\mathbb{Q}^n$, but we only care about whether $\langle\lambda,a_i\rangle$ is greater than, less than or equal to 0; therefore we can assume $\lambda \in \mathbb{Y}$. Also our sign convention in Definition 5.5 is opposite to that given in [ibid., Definition 7.2.1] so that it agrees with the conventions of Section 4.

5E. Let us define an equivalence relation on the set of pairs (λ,θ) by saying that (λ_1,θ_1) is equivalent to (λ_2,θ_2) if $\{i \mid \langle\lambda_1,a_i\rangle > 0\} = \{i \mid \langle\lambda_2,a_i\rangle > 0\}$, $\{i \mid \langle\lambda_1,a_i\rangle < 0\} = \{i \mid \langle\lambda_2,a_i\rangle < 0\}$ and $\theta_2 \equiv \theta_1 \mod \sum_{\langle\lambda_1,a_i\rangle=0} \mathbb{Z}\cdot a_i$. Denote by $\mathcal{P}_\chi$ the set of equivalence classes of pairs (λ,θ) that are attached to χ. The set of all possible λ up to equivalence consist of the (finitely many) covectors of the oriented matroid defined by A. It will be convenient to parametrize each λ (again up to equivalence) as an element in $\{+,0,-\}^n$, $\lambda \leftrightarrow (e_i)_{i\in[1,n]}$ with $e_i = +$ if $\langle\lambda,a_i\rangle > 0$ and so forth. Note, however, that not every element of $\{+,0,-\}^n$ can be realized as some λ.

Proposition 5.7 [Musson and Van den Bergh 1998, Proposition 7.7.1]. *Choose $\chi,\chi' \in \mathfrak{t}^*$. Then, the set $\mathcal{P}_\chi$ parametrizes the primitive ideals in U_χ and $\chi \to \chi'$ if and only if $\mathcal{P}_{\chi'} \subseteq \mathcal{P}_\chi$.*

Since we are interested in sheaves of W-algebras on smooth hypertoric varieties we may assume that A is unimodular. This allows us to remove θ from the description of $\mathcal{P}_\chi$.

Lemma 5.8. *Assume that A is unimodular and let (λ,θ) and (λ,ϑ), be attached to χ via $\alpha \in \rho^{-1}(\chi)$ and $\beta \in \rho^{-1}(\chi)$, respectively. Then (λ,θ) is equivalent to (λ,ϑ).*

Proof. By definition, θ is the equivalence class of $\sum_{\langle\lambda,a_i\rangle=0} \alpha_i a_i$ in the quotient $\left(\sum_{\langle\lambda,a_i\rangle=0} \mathbb{C}\cdot a_i\right) / \left(\sum_{\langle\lambda,a_i\rangle=0} \mathbb{Z}\cdot a_i\right)$, and similarly for ϑ. Therefore we must show that $\sum_{i=1}^n \alpha_i a_i = \sum_{i=1}^n \beta_i a_i$ implies

$$\sum_{\langle\lambda,a_i\rangle=0}^{n} \alpha_i a_i \equiv \sum_{\langle\lambda,a_i\rangle=0}^{n} \beta_i a_i \mod \sum_{\langle\lambda,a_i\rangle=0} \mathbb{Z}\cdot a_i. \tag{14}$$

Choose $\{a_{i_1}, \dots, a_{i_k}\} \subset \{a_i \mid \langle \lambda, a_i \rangle = 0\}$ to be a basis of the space spanned by the set $\{a_i \mid \langle \lambda, a_i \rangle = 0\}$. We can extend this to a basis $a_{i_1}, \dots, a_{i_k}, a_{i_{k+1}}, \dots, a_d$ of $\mathfrak{t}^*$. Since A is unimodular the determinant of this basis is ± 1. Hence $\{a_{i_1}, \dots, a_{i_k}\}$ span a direct summand of the lattice $\mathbb{X}$. This implies

$$\Big(\sum_{\langle \lambda, a_i \rangle = 0} \mathbb{C} \cdot a_i \Big) \cap \mathbb{X} = \sum_{\langle \lambda, a_i \rangle = 0} \mathbb{Z} \cdot a_i, \tag{15}$$

which in turn implies (14). □

It is shown in [ibid., Example 7.2.7] that θ is not defined up to equivalence by χ and λ if A is not unimodular.

5F. Based on Lemma 5.8, we make the following definition.

Definition 5.9. Let $\lambda \in \mathbb{Y}$ and $\chi \in \mathfrak{t}^*$. We say that λ is attached to χ if there exists $\alpha \in \rho^{-1}(\chi)$ such that

$$\begin{aligned} \langle \lambda, a_i \rangle > 0 &\implies \alpha_i \in \mathbb{Z}, \alpha_i \geq 0, \\ \langle \lambda, a_i \rangle < 0 &\implies \alpha_i \in \mathbb{Z}, \alpha_i < 0, \\ \langle \lambda, a_i \rangle = 0 &\implies \alpha_i \in \mathbb{C} \backslash \mathbb{Z}. \end{aligned} \tag{16}$$

If $\lambda_1, \lambda_2 \in \mathbb{Y}$ are attached to χ then we say that λ_1 is equivalent to λ_2 if

$$\{i \mid \langle \lambda_1, a_i \rangle > 0\} = \{i \mid \langle \lambda_2, a_i \rangle > 0\} \quad \text{and} \quad \{i \mid \langle \lambda_1, a_i \rangle < 0\} = \{i \mid \langle \lambda_2, a_i \rangle < 0\}.$$

Let $\mathcal{Q}_\chi$ denote the set of equivalence classes of elements in $\mathbb{Y}$ that are attached to χ.

Lemma 5.10. *Assume that A is unimodular. Then $\chi \to \chi'$ if and only if $\mathcal{Q}_{\chi'} \subseteq \mathcal{Q}_\chi$ and $\chi - \chi' \in \mathbb{X}$.*

Proof. If $\chi \to \chi'$ then clearly $\chi - \chi' \in \mathbb{X}$ and Proposition 5.7 implies that $\mathcal{P}_{\chi'} \subseteq \mathcal{P}_\chi$. This implies that $\mathcal{Q}_{\chi'} \subseteq \mathcal{Q}_\chi$.

Now assume that $\mathcal{Q}_{\chi'} \subseteq \mathcal{Q}_\chi$ and $\chi - \chi' \in \mathbb{X}$. Let $\lambda \in \mathcal{Q}_{\chi'}$ and choose $\alpha \in \rho^{-1}(\chi')$, respectively $\beta \in \rho^{-1}(\chi)$, satisfying the conditions of Definition 5.9 for λ with respect to χ', respectively χ. Write $\alpha = \alpha^{(1)} + \alpha^{(2)}$, where $(\alpha^{(1)})_i = \alpha_i$ if $\langle \lambda, a_i \rangle \neq 0$ and $(\alpha^{(1)})_i = 0$ if $\langle \lambda, a_i \rangle = 0$. Decompose $\beta = \beta^{(1)} + \beta^{(2)}$ in a similar fashion. Then

$$(\chi - \chi') - \rho(\beta^{(1)} - \alpha^{(1)}) \in \Big(\sum_{\langle \lambda, a_i \rangle = 0} \mathbb{C} \cdot a_i \Big) \cap \mathbb{X},$$

which, by (15), equals $\sum_{\langle \lambda, a_i \rangle = 0} \mathbb{Z} \cdot a_i$. Therefore we can choose $u \in \mathbb{Z}^n$ such that $u_i = 0$ for all i such that $\langle \lambda, a_i \rangle = 0$ and $\rho(u) = (\chi - \chi') - \rho(\beta^{(1)} - \alpha^{(1)})$. Define $\delta^{(2)} = \alpha^{(2)} + u$ and $\delta = \beta^{(1)} + \delta^{(2)}$ so that $\rho(\delta) = \chi$. We have

$$\bar{\delta}^{(2)} = \bar{\alpha}^{(2)} \in \Big(\sum_{\langle \lambda, a_i \rangle = 0} \mathbb{C} \cdot a_i \Big) \Big/ \Big(\sum_{\langle \lambda, a_i \rangle = 0} \mathbb{Z} \cdot a_i \Big)$$

and $(\lambda, \bar{\alpha}^{(2)})$ and $(\lambda, \bar{\delta}^{(2)})$ are attached to χ' and χ, respectively, in the sense of Definition 5.5. Therefore $(\lambda, \bar{\delta}^{(2)}) = (\lambda, \bar{\alpha}^{(2)}) \in \mathcal{P}_\chi$ implies that $\mathcal{P}_{\chi'} \subseteq \mathcal{P}_\chi$. Hence Proposition 5.7 implies that $\chi \to \chi'$. □

Proof of Proposition 5.1. As was stated in Section 4D, the cone $\overline{C}$ is a rational cone. Therefore we can choose $\mu_1, \dots, \mu_k$ in $\mathbb{Y}$ such that

$$\overline{C} = \{\chi \in \mathbb{X}_\mathbb{R} \mid \langle \mu_i, \chi\rangle \geq 0 \text{ for all } i \in [1,k]\}$$
$$\supset \{\chi \in \mathbb{X}_\mathbb{R} \mid \langle \mu_i, \chi\rangle > 0 \text{ for all } i \in [1,k]\} = C.$$

We will construct $C(\chi)$ in three stages.

Claim 1. There exists an integer $N_0 \gg 0$ such that $pN_0 \cdot \mathsf{pr}(\chi) \in \mathbb{X} \cap C$ and $\chi + pN_0 \cdot \mathsf{pr}(\chi) \to \chi$ for all $p \in \mathbb{N}$.

For each $\lambda \in \mathcal{Q}_\chi$ fix an element $\beta^\lambda \in \rho^{-1}(\chi)$ such that β^λ satisfies the properties listed in Definition 5.9 with respect to λ. Then $\mathsf{pr}(\chi) = \sum_{i=1}^n \mathsf{pr}(\beta_i^\lambda) a_i$ and we choose N_0 such that $N_0 \cdot \mathsf{pr}(\beta_i^\lambda) \in \mathbb{Z}$ for all $\lambda \in \mathcal{Q}_\chi$ and all i. The element $(\beta_i^\lambda + pN_0\beta_i^\lambda)_{i\in[1,n]}$ in $\mathfrak{g}^*$ satisfies the properties of Definition 5.9 with respect to λ hence $\mathcal{Q}_\chi \subseteq \mathcal{Q}_{\chi+pN_0\cdot\mathsf{pr}(\chi)}$. Since $pN_0 \cdot \mathsf{pr}(\chi) \in \mathbb{X}$, Lemma 5.10 says that $\chi + pN_0 \cdot \mathsf{pr}(\chi) \to \chi$ for all $p \in \mathbb{N}$. Note also that

$$\langle \mu_i, \mathsf{pr}(\chi + pN_0 \cdot \mathsf{pr}(\chi))\rangle = (1 + pN_0)\langle \mu_i, \mathsf{pr}(\chi)\rangle > 0$$

for all i shows that $\mathsf{pr}(\chi + pN_0 \cdot \mathsf{pr}(\chi)) \in C$.

Claim 2. Fix $\delta = \sum_{i=1}^n a_i \in \mathbb{X}$. There exists an integer $N_1 \gg 0$ such that

$$N_1 \cdot \mathsf{pr}(\chi) + \delta \in \mathbb{X} \cap C \quad \text{and} \quad \chi + p(N_1 \cdot \mathsf{pr}(\chi) + \delta) \to \chi$$

for all $p \in \mathbb{N}$. Moreover, for all $\lambda \in \mathcal{P}_\chi$, there exists β^λ as before except that $\beta_i^\lambda \neq 0$ for all i.

Choose $N_1 = pN_0$ such that

$$(N_1/d) \cdot \langle \mu_i, \mathsf{pr}(\chi)\rangle > |\langle \mu_i, a_j\rangle| \tag{17}$$

for all $i \in [1,k]$ and $j \in [1,n]$. Let $\beta^\lambda \in \rho^{-1}(\chi + N_1 \cdot \mathsf{pr}(\chi))$ satisfy Definition 5.9 with respect to λ for $\lambda \in \mathcal{Q}_\chi$. By choosing a larger p if necessary we may assume that $\beta_i^\lambda \in \mathbb{Z}\backslash\{0\}$ implies that $|\beta_i^\lambda| > 1$. Then $\beta_i^\lambda + 1 < 0$ if $\beta_i^\lambda \in \mathbb{Z}_{<0}$ and $\beta_i^\lambda + 1 > 0$ if $\beta_i^\lambda \in \mathbb{Z}_{\geq 0}$. Moreover $(\beta_i^\lambda + 1)_{i\in[1,n]}$ satisfies (16) with respect to λ,

$$\sum_{i=1}^n (\beta_i^\lambda + 1)a_i = \chi + (N_1 \cdot \mathsf{pr}(\chi) + \delta),$$

and hence $\chi + (N_1 \cdot \mathsf{pr}(\chi) + \delta) \to \chi$. The same holds for all $\chi + p(N_1 \cdot \mathsf{pr}(\chi) + \delta)$. Finally (17) implies that $\mathsf{pr}(\chi + q(N_1 \cdot \mathsf{pr}(\chi) + \delta)) \in C$ for all $q \in \mathbb{Z}_{\geq 0}$.

Proof of the proposition. Note that (17) implies $p(N_1 \cdot \mathsf{pr}(\chi) + \delta) \in C$ for all p as well. Let

$$I_\chi = \{\epsilon \in (-1, 1)^n \subset \mathbb{Q}^n \mid -\langle \mu_i, (N_1 \cdot \mathsf{pr}(\chi) + \delta)\rangle < \langle \mu_i, \epsilon\rangle \text{ for all } i\}.$$

Since C is d-dimensional, there exists some $0 < c < 1$ such that $[-c, c]^n \subset I_\chi$. Let $\{v_j \mid j \in [1, 2^n]\} \subset [-c, c]^n$ be the vertices of the box. Choose $p \in \mathbb{N}$ such that $p \cdot v_j \in \mathbb{Z}^n$ for all j. The same argument as in Claims 1 and 2 shows

$$\chi \leftarrow \chi + q(p(N_1 \cdot \mathsf{pr}(\chi) + \delta + A \cdot v_j)) \quad \text{for all } q \in \mathbb{Z}_{\geq 0}.$$

We set $u_j = p(N_1 \cdot \mathsf{pr}(\chi) + \delta + A \cdot v_j)$. One can check as above $\chi \leftarrow \chi + \sum_{i=1}^{2^n} k_i \cdot u_i$ for all $k_i \in \mathbb{Z}_{\geq 0}$. $\square$

Remark 5.11. We conclude with a couple of remarks regarding Proposition 5.1.

(i) Note that in the proof of Proposition 5.1 we only used the fact that C is the interior of some d-dimensional rational cone.

(ii) In general, the proposition is false when $\mathsf{pr}(\chi) \in C$ is replaced by $\mathsf{pr}(\chi) \in \overline{C}$.

(iii) It would be very interesting to directly relate the sets $\mathfrak{Q}_\chi$ to the G.I.T. fan.

6. The rational Cherednik algebra associated to cyclic groups

6A. As explained in the introduction, the original motivation for this article was to reproduce the results of [Kashiwara and Rouquier 2008] for the rational Cherednik algebra $H_{\boldsymbol{h}}(\mathbb{Z}_m)$ associated to the cyclic group $\mathbb{Z}_m$. These rational Cherednik algebras are parametrized[3] by an m-tuple $\boldsymbol{h} = (h_i)_{i \in [0, m-1]} \in \mathbb{C}^m$, where the indices are taken modulo m. We fix a one-dimensional space $\mathfrak{h} = \mathbb{C} \cdot y$ and $\mathfrak{h}^* = \mathbb{C} \cdot x$ such that $\langle x, y\rangle = 1$. The cyclic group $\mathbb{Z}_m = \langle \varepsilon \rangle$ acts on $\mathfrak{h}$ and $\mathfrak{h}^*$ via $\varepsilon \cdot y = \zeta^{-1} y$ and $\varepsilon \cdot x = \zeta x$, where ζ is a fixed primitive m-th root of unity. The idempotents in $\mathbb{C}\,\mathbb{Z}_m$ corresponding to the simple $\mathbb{Z}_m$-modules are

$$e_i = \frac{1}{m} \sum_{j=0}^{m-1} \zeta^{-ij} \varepsilon^j, \quad i \in [0, m-1],$$

so that $\varepsilon \cdot e_i = \zeta^i e_i$. Then $e_{i+1} \cdot x = x \cdot e_i$ and $e_{i-1} \cdot y = y \cdot e_i$. If we fix $\alpha_{\varepsilon^i} = \sqrt{2} \cdot x$ and $\alpha_{\varepsilon^i}^\vee = (-1/\sqrt{2}) \cdot y$ then the commutation relations defining $H_{\boldsymbol{h}}(\mathbb{Z}_m)$, as stated in [Rouquier 2008], become

$$\varepsilon \cdot x = \zeta x \cdot \varepsilon, \qquad \varepsilon \cdot y = \zeta^{-1} y \cdot \varepsilon, \qquad [y, x] = 1 + m \sum_{i=0}^{m-1} (h_{i+1} - h_i) e_i,$$

where indices are taken modulo m.

[3] In this paper the parameters (h_i) and (χ_i) are used. However the paper [Kuwabara 2010] uses the parameters (κ_i) and (c_i). The different parametrizations are related by $h_i \leftrightarrow \kappa_i$ and $c_i \leftrightarrow \chi_i - \chi_{i+1}$.

6B. The category $\mathcal{O} \subset H_{\boldsymbol{h}}$-mod is defined to be the subcategory of all finitely generated $H_{\boldsymbol{h}}$-modules such that the action of $y \in \mathbb{C}[\mathfrak{h}^*]$ is locally nilpotent. It is a highest weight category. To each simple $\mathbb{Z}_m$-module $\mathbb{C} \cdot e_i$, one can associate a *standard module* in the category $\mathcal{O}$ defined by

$$\Delta(e_i) := H_{\boldsymbol{h}} \otimes_{\mathbb{C}[\mathfrak{h}^*] \rtimes \mathbb{Z}_m} \mathbb{C} \cdot e_i,$$

where $y \in \mathbb{C}[\mathfrak{h}^*]$ acts as zero on e_i. Each $\Delta(e_i)$ has a simple head $L(e_i)$ and $L(e_i) \not\simeq L(e_j)$ for $i \neq j$. The set of simple modules $\{L(e_i)\}_{i \in [0,m-1]}$ is, up to isomorphism, all simple modules in $\mathcal{O}$. Fix $i \in [0, m-1]$ and let c_i be the smallest element in $\mathbb{Z}_{\geq 1} \cup \{\infty\}$ such that $c_i + mh_{i+c_i} - mh_i = 0$. The identity

$$[y, x^j] = x^{j-1}\left(j + m \sum_{i=0}^{m-1} (h_{i+j} - h_i) e_i\right), \text{ for all } j \geq 0,$$

shows that $L(e_i) = (\mathbb{C}[x]/(x^{c_i})) \otimes e_i$. Fix $e := e_0$, the trivial idempotent. The algebra $eH_{\boldsymbol{h}}e$ is called the spherical subalgebra of $H_{\boldsymbol{h}}$. Multiplication by e defines a functor $e : H_{\boldsymbol{h}}\text{-mod} \to eH_{\boldsymbol{h}}e\text{-mod}$ with left adjoint $H_{\boldsymbol{h}}e \otimes_{eH_{\boldsymbol{h}}e} (\,\cdot\,)$. Let $\mathscr{C} \subset \mathbb{C}^m$ be the union of the finitely many hyperplanes defined by the equations $j + mh_{i+j} - mh_i = 0$, where $i \in [1, m-1]$ and $j \in [0, m-i]$.

Lemma 6.1. *The functor* $e : H_{\boldsymbol{h}}\text{-mod} \to eH_{\boldsymbol{h}}e\text{-mod}$ *is an equivalence if and only if* $\boldsymbol{h} \notin \mathscr{C}$. *This implies that* $eH_{\boldsymbol{h}}e$ *has finite global dimension when* $\boldsymbol{h} \notin \mathscr{C}$.

Proof. The functor e will be an equivalence if and only if $H_{\boldsymbol{h}}eH_{\boldsymbol{h}} = H_{\boldsymbol{h}}$. By Ginzburg's generalized Duflo theorem [Ginzburg 2003, Theorem 2.3], $H_{\boldsymbol{h}}eH_{\boldsymbol{h}} \neq H_{\boldsymbol{h}}$ implies that there is some simple module in the category $\mathcal{O}$ that is annihilated by e. This happens if and only if $\boldsymbol{h} \in \mathscr{C}$. The second statement follows from the fact that $H_{\boldsymbol{h}}$ has finite global dimension. □

6C. *The minimal resolution of* $\mathbb{C}^2/\mathbb{Z}_m$. In order to relate the spherical subalgebra of $H_{\boldsymbol{h}}$ to a W-algebra on the resolution of the corresponding Kleinian singularity $\mathbb{C}^2/\mathbb{Z}_m$, we must describe $eH_{\boldsymbol{h}}e$ as a quantum Hamiltonian reduction. Such an isomorphism is well known and is a particular case of a more general construction by Holland [1999]. First we describe the minimal resolution of $\mathbb{C}^2/\mathbb{Z}_m$ as a hypertoric variety. Let Q be the cyclic quiver with vertices $V = \{v_0, \dots, v_{m-1}\}$ and arrows $u_i : v_{i-1} \to v_i$ for $i \in [1, m]$ (where v_m is identified with v_0). Let ν be the dimension vector with 1 at each vertex. Then the space of representations for Q with dimension vector ν is the affine space

$$\mathsf{Rep}(Q, \nu) = \{(u_i)_{i \in [1,m]} \mid u_i \in \mathbb{C}\} \simeq \mathbb{C}^m$$

and we write $\mathbb{C}[\mathsf{Rep}(Q, \nu)] = \mathbb{C}[x_1, \dots, x_m]$. There is an action of

$$\mathbb{T}^m = \{(\lambda_i)_{i \in [1,m]} \mid \lambda_i \in \mathbb{C}^\times\}$$

on $\mathsf{Rep}(Q,\nu)$ given by $\lambda\cdot u_i=\lambda_i\lambda_{i-1}^{-1}u_i$, and hence $\lambda\cdot x_i=\lambda_i^{-1}\lambda_{i-1}x_i$. The one-dimensional torus $\mathbb{T}$ embedded diagonally in $\mathbb{T}^m$ acts trivially on $\mathsf{Rep}(Q,\nu)$. Therefore $\mathbb{T}^{m-1}:=\mathbb{T}^m/\mathbb{T}$ acts on $\mathsf{Rep}(Q,\nu)$. The lattice of characters $\mathbb{X}(\mathbb{T}^{m-1})$ is the sublattice of $\mathbb{X}(\mathbb{T}^m)=\bigoplus_{i=1}^m\mathbb{Z}\cdot v_i$ consisting of points $\phi=\sum_{i=1}^m\phi_i v_i$ such that $\sum_{i=1}^m\phi_i=0$. We fix the basis $\{w_i=v_i-v_{i+1}\mid i\in[0,m-2]\}$ of $\mathbb{X}(\mathbb{T}^{m-1})$ so that $\phi=(\phi_i)_{i\in[1,n]}=\sum_{i=1}^{m-1}\chi_i w_i$, where $\chi_i=\sum_{j=1}^i\phi_i$. Then the $(m-1)\times m$ matrix encoding the action of $\mathbb{T}^{m-1}$ is given by

$$A=(a_1,\dots,a_m)=\begin{pmatrix}1&0&\dots&0&-1\\0&1&&\vdots&-1\\\vdots&&\ddots&0&\vdots\\0&\dots&0&1&-1\end{pmatrix}.$$

The G.I.T walls in $\mathfrak{t}^*$, where $\mathfrak{t}=\mathrm{Lie}(\mathbb{T}^{m-1})$, are given by the hyperplanes $H_i=(\chi_i=0)$, $i\in[1,m-1]$, and $H_{ij}=(\chi_i=\chi_j)$, $i\neq j\in[1,m-1]$. Hence the m-cones are the connected components of the complement to this union of hyperplanes. As was shown originally in terms of hyperkähler manifolds by Kronheimer [1989] and then by Cassens and Slodowy [1998] in the algebraic setting, we have:

Proposition 6.2. *Let δ belong to the interior of an m-cone. Then the hypertoric variety $Y(A,\delta)$ is isomorphic to the minimal resolution $(\mathbb{C}^2/\mathbb{Z}_m)^\sim$ of the Kleinian singularity $\mathbb{C}^2/\mathbb{Z}_m$.*

As is well-known, the hypertoric variety $Y(A,\delta)$ is a toric variety. It is shown in [Hausel and Sturmfels 2002, Theorem 10.1] that a hypertoric variety is toric if and only if it is a product of varieties of the form $(\mathbb{C}^2/\mathbb{Z}_m)^\sim$ for various m. Let us now consider the corresponding quantum Hamiltonian reduction

$$\mathsf{U}_\chi=\big(\mathfrak{D}(\mathrm{Rep}(Q,\nu))\;\big/\;\mathfrak{D}(\mathrm{Rep}(Q,\nu))(\mu_D-\chi)(\mathfrak{t})\big)^{\mathbb{T}^{m-1}}.$$

The quantum moment map in this case is given by

$$\mu_D:\mathfrak{t}\longrightarrow\mathfrak{D}(\mathrm{Rep}(Q,\nu)),\qquad t_i\mapsto x_i\partial_i-x_m\partial_m\quad\text{for all }\ i\in[1,m-1].$$

Since

$$\mathfrak{D}(\mathrm{Rep}(Q,\nu))^{\mathbb{T}^{m-1}}=\langle\partial_1\cdots\partial_m,x_1\cdots x_m,x_1\partial_1,\dots,x_m\partial_m\rangle,$$
$$\langle(\mu_D-\chi)(\mathfrak{t})\rangle=\langle x_i\partial_i-x_m\partial_m-\chi_i\mid i\in[1,m-1]\rangle,$$

where we set $\partial_i:=\partial/\partial x_i$, U_χ is generated by $\partial_1\cdots\partial_m$, $x_1\cdots x_m$ and $x_m\partial_m$.

6D. ***The Dunkl embedding.*** Let $\mathfrak{h}_{\mathrm{reg}}:=\mathfrak{h}\backslash\{0\}$ and denote by $\mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})$ the ring of algebraic differential operators on $\mathfrak{h}_{\mathrm{reg}}$. In order to show that the spherical subalgebra of $H_{\boldsymbol{h}}$ is isomorphic to a suitable quantum Hamiltonian reduction, we realize $eH_{\boldsymbol{h}}e$ as a subalgebra of $\mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})$ using the Dunkl embedding. Similarly, using the "radial parts map”, we will also realize U_χ as the same subalgebra of $\mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})$. The Dunkl embedding is the map $\Theta_{\boldsymbol{h}} : H_{\boldsymbol{h}} \to \mathfrak{D}(\mathfrak{h}_{\mathrm{reg}}) \rtimes \mathbb{Z}_m$ defined by

$$\Theta_{\boldsymbol{h}}(y) = \frac{\mathrm{d}}{\mathrm{d}x} + \frac{m}{x}\sum_{i=0}^{m-1} h_i e_i, \quad \Theta_{\boldsymbol{h}}(x) = x \quad \text{and} \quad \Theta_{\boldsymbol{h}}(\varepsilon) = \varepsilon.$$

The algebra $\mathfrak{D}(\mathfrak{h}_{\mathrm{reg}}) \rtimes \mathbb{Z}_m$ is filtered by order of differential operators, that is, $\deg(\mathrm{d}/\mathrm{d}x) = 1$ and $\deg(x) = \deg(\varepsilon) = 0$. If we define a filtration on $H_{\boldsymbol{h}}$ by setting $\deg(y) = 1$ and $\deg(x) = \deg(\varepsilon) = 0$, then the map $\Theta_{\boldsymbol{h}}$ is filter preserving. Localizing $H_{\boldsymbol{h}}$ at the regular element x provides an isomorphism

$$\Theta_{\boldsymbol{h}} : H_{\boldsymbol{h}}[x^{-1}] \xrightarrow{\sim} \mathfrak{D}(\mathfrak{h}_{\mathrm{reg}}) \rtimes \mathbb{Z}_m.$$

Therefore $\Theta_{\boldsymbol{h}}$ is injective. Applying the trivial idempotent produces

$$\Theta_{\boldsymbol{h}} : eH_{\boldsymbol{h}}e \longrightarrow e\mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})e \simeq \mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})^{\mathbb{Z}_m}.$$

Let us note that $\mathsf{gr}(H_{\boldsymbol{h}}) \simeq \mathbb{C}[x, y] \rtimes \mathbb{Z}_m$ and $\mathsf{gr}(eH_{\boldsymbol{h}}e) \simeq \mathbb{C}[x, y]^{\mathbb{Z}_m}$. Therefore $eH_{\boldsymbol{h}}e$ is generated by $x^m e$, xye and $y^m e$. Since $\Theta_{\boldsymbol{h}}(ye_i) = (\mathrm{d}/\mathrm{d}x + (m/x)h_i)e_i$ we get

$$\Theta_{\boldsymbol{h}}(y^m e) = \prod_{i=1}^{m}\Big(\frac{\mathrm{d}}{\mathrm{d}x} + \frac{m}{x}h_i\Big) \quad \text{and} \quad \Theta_{\boldsymbol{h}}(xye) = x\frac{\mathrm{d}}{\mathrm{d}x} + mh_m.$$

We note that

$$\Theta_{\boldsymbol{h}}(y^m e)(x^r) = \prod_{i=1}^{m}(r - m + i + mh_i)x^{r-m}. \tag{18}$$

6E. ***The radial parts map.*** In this subsection we show that $\mathsf{U}_\chi \simeq \Theta_{\boldsymbol{h}}(eH_{\boldsymbol{h}}e)$. The isomorphism we describe is not new, it was first constructed by Holland [1999] (see also [Kuwabara 2008]), but we give it in order to fix parameters. There is a natural embedding $\mathfrak{h} \hookrightarrow \mathrm{Rep}(Q, \nu)$ given by $x \mapsto (x, \dots, x)$. This defines a surjective morphism $\mathbb{C}[\mathrm{Rep}(Q, \nu)] \to \mathbb{C}[\mathfrak{h}]$, $x_i \mapsto x$, which descends to a “Chevalley isomorphism”

$$\rho : \mathbb{C}[\mathrm{Rep}(Q, \nu)]^{\mathbb{T}^{m-1}} \xrightarrow{\sim} \mathbb{C}[\mathfrak{h}]^{\mathbb{Z}_m}, \quad x_1 \cdots x_m \mapsto x^m.$$

Define a section

$$\rho^{-1} : \mathbb{C}[\mathfrak{h}] \longrightarrow \mathbb{C}[\mathrm{Rep}(Q, \nu)][x_i^{1/m} \mid i \in [1, m]] \quad \text{by} \quad x^r \mapsto x_1^{r/m} \cdots x_m^{r/m}.$$

This can be extended to a twisted Harish-Chandra morphism

$$\hat{\mathfrak{R}}_{\boldsymbol{h}} : \mathscr{D}(\mathrm{Rep}(Q, \nu))^{\mathbb{T}} \longrightarrow \mathfrak{D}(\mathfrak{h}_{\mathrm{reg}})^{\mathbb{Z}_m}$$

given by

$$\hat{\mathfrak{R}}_{\boldsymbol{h}}(D)(f)=\rho(\delta_{\boldsymbol{h}}^{-1}D(\rho^{-1}(f)\delta_{\boldsymbol{h}}))\quad\text{for all } f\in\mathbb{C}[\mathfrak{h}],\quad\text{where } \delta_{\boldsymbol{h}}=\prod_{i=1}^{m}x_i^{h_i+\frac{i-m}{m}}.$$

Calculating the action of $\hat{\mathfrak{R}}_{\boldsymbol{h}}(\partial_1\cdots\partial_m)$ on x^r and comparing with (18) shows that

$$\hat{\mathfrak{R}}_{\boldsymbol{h}}(m^m\cdot\partial_1\cdots\partial_m)=\Theta_{\boldsymbol{h}}(y^m e).$$

Similarly,

$$\hat{\mathfrak{R}}_{\boldsymbol{h}}(x_1\cdots x_m)=\Theta_{\boldsymbol{h}}(x^m e)\quad\text{and}\quad\hat{\mathfrak{R}}_{\boldsymbol{h}}(x_i\partial_i)=\frac{1}{m}\Bigl(x\frac{\mathrm{d}}{\mathrm{d}x}+mh_i+i-m\Bigr).$$

This implies that $\hat{\mathfrak{R}}_{\boldsymbol{h}}$ defines a surjection $\mathfrak{D}(\operatorname{Rep}(Q,\nu))^{\mathbb{T}^{m-1}}\twoheadrightarrow\Theta_{\boldsymbol{h}}(eH_{\boldsymbol{h}}e)$. We fix

$$\chi_i=h_i-h_m+\frac{i-m}{m},\quad i\in[1,m-1].\tag{19}$$

Then $\hat{\mathfrak{R}}_{\boldsymbol{h}}(x_i\partial_i-x_m\partial_m-\chi_i)=0$ and $\hat{\mathfrak{R}}_{\boldsymbol{h}}$ descends to a surjective morphism

$$\mathfrak{R}_{\boldsymbol{h}}:\mathsf{U}_\chi\longrightarrow\Theta_{\boldsymbol{h}}(eH_{\boldsymbol{h}}e).$$

As above, $\mathfrak{D}(\operatorname{Rep}(Q,\nu))$ is a filtered algebra by setting $\deg(\partial_i)=1$ and $\deg(x_i)=0$ for $i\in[1,m]$. This induces a filtration on U_χ and we see from the definitions that $\mathfrak{R}_{\boldsymbol{h}}$ is filter preserving. Therefore we get a morphism of associated graded algebras

$$\operatorname{gr}\mathfrak{R}_{\boldsymbol{h}}:\operatorname{gr}(\mathsf{U}_\chi)\longrightarrow\operatorname{gr}(eH_{\boldsymbol{h}}e).$$

Now [Holland 1999, Proposition 2.4] says that

$$\operatorname{gr}(\mathsf{U}_\chi)=\mathbb{C}[\mu^{-1}(0)]^{\mathbb{T}^{m-1}}\simeq\mathbb{C}[x,y]^{\mathbb{Z}_m}=\operatorname{gr}(eH_{\boldsymbol{h}}e).$$

This isomorphism is realized by $x_1\cdots x_m\mapsto x^m$, $m^m\cdot y_1\cdots y_m\mapsto y^m$ and $x_1y_1\mapsto(1/m)\cdot xy$. But we see from above that this is precisely what $\operatorname{gr}\mathfrak{R}_{\boldsymbol{h}}$ does to the principal symbols of the generators $m^m\cdot\partial_1\cdots\partial_m$, $x_1\cdots x_m$ and $x_1\partial_1$ of U_χ. Therefore $\operatorname{gr}\mathfrak{R}_{\boldsymbol{h}}$ is an isomorphism and hence $\Theta_{\boldsymbol{h}}^{-1}\circ\mathfrak{R}_{\boldsymbol{h}}:\mathsf{U}_\chi\xrightarrow{\sim}eH_{\boldsymbol{h}}e$ is a filtration-preserving isomorphism.

6F. ***Localization of $H_{\boldsymbol{h}}(\mathbb{Z}_m)$.*** As noted in Proposition 6.2, the hypertoric varieties $Y(A,\delta)$ are all isomorphic provided δ does not belong to a wall in $\mathbb{X}_{\mathbb{Q}}$. Therefore, for any $\chi\in\mathfrak{t}^*$, we may refer to the sheaf $\mathscr{A}_\chi$ on the minimal resolution $(\mathbb{C}^2/\mathbb{Z}_m)^\sim$, but the reader should be aware that in doing so we have implicitly fixed an identification $(\mathbb{C}^2/\mathbb{Z}_m)^\sim=Y(A,\delta)$. Recall the union of hyperplanes $\mathscr{C}\subset\mathbb{C}^m$ defined in Lemma 6.1.

Theorem 6.3. *Choose $\boldsymbol{h} \in \mathbb{C}^m \setminus \mathscr{C}$ and let χ be defined by (19). Write $\mathscr{A}_{\boldsymbol{h}} := \mathscr{A}_\chi$ for the sheaf of W-algebras on $(\mathbb{C}^2/\mathbb{Z}_m)^\sim$. Then the functor*

$$\operatorname{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\boldsymbol{h}})}(\mathscr{A}_{\boldsymbol{h}}, \cdot\,)$$

defines an equivalence of categories $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\boldsymbol{h}}) \xrightarrow{\sim} eH_{\boldsymbol{h}}e\text{-mod}$ with quasiinverse $\mathscr{A}_{\boldsymbol{h}} \otimes_{eH_{\boldsymbol{h}}e} (\,\cdot\,)$. Moreover, the functor

$$H_{\boldsymbol{h}}e \otimes_{eH_{\boldsymbol{h}}e} \operatorname{Hom}_{\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\boldsymbol{h}})}(\mathscr{A}_{\boldsymbol{h}}, \cdot\,)$$

defines an equivalence of categories $\mathrm{Mod}_F^{\mathrm{good}}(\mathscr{A}_{\boldsymbol{h}}) \xrightarrow{\sim} H_{\boldsymbol{h}}\text{-mod}$ with quasiinverse $\mathscr{A}_{\boldsymbol{h}} \otimes_{eH_{\boldsymbol{h}}e} eH_{\boldsymbol{h}} \otimes_{H_{\boldsymbol{h}}} (\,\cdot\,)$.

Proof. The condition $\chi_i \neq \chi_j$ for $i \neq j \in [1, m-1]$ translates, via (19), into $h_i - h_j + (i-j)/m \neq 0$ for all $i \neq j \in [1, m-1]$. Similarly, the condition $\chi_i \neq 0$ for all $i \in [1, m-1]$ translates into $h_i - h_m + \frac{i-m}{m} \neq 0$ for all $i \in [1, m-1]$. Therefore the linear map $\mathbb{C}^m \to \mathbb{C}^{m-1}$ defined by (19) maps the union of hyperplanes $\mathscr{C}$ onto

$$\{\chi \in \mathbb{C}^{m-1} \mid \chi_i = \chi_j \text{ for } i \neq j \in [1, m-1] \text{ or } \chi_i = 0 \text{ for } i \in [1, m-1])\},$$

which is precisely the union of the G.I.T. walls in $\mathbb{C}^{m-1}$. Therefore Lemma 6.1 implies that U_χ has finite global dimension when χ lies in the interior of some G.I.T. cone C. Now the theorem follows from Corollary 5.3. □

Remark 6.4. In the above situation it is possible to explicitly calculate the sets $\mathfrak{Q}_\chi$ and hence describe the partial ordering on comparability classes as defined in Section 5A. However the answer is not very illuminating.

Finally, we would just like to note the various forms in which the rational Cherednik algebra $H_{\boldsymbol{h}}(\mathbb{Z}_m)$ appears in the literature. It is isomorphic to the deformed preprojective algebra of type A as studied in [Crawley-Boevey and Holland 1998]. It is well-known that its spherical subalgebra $eH_{\boldsymbol{h}}(\mathbb{Z}_m)e$ coincides with a "generalized $U(\mathfrak{sl}_2)$-algebra", as studied by Hodges [1993] and Smith [1990]. Combining this fact with Premet's results [2002] shows that $eH_{\boldsymbol{h}}(\mathbb{Z}_m)e$ is also isomorphic to the finite $\mathscr{W}$-algebra associated to $\mathfrak{gl}_m(\mathbb{C})$ at a subregular nilpotent element. Recently, Losev [2012] has constructed explicit isomorphisms between the spherical subalgebra of certain rational Cherednik algebras and their related finite $\mathscr{W}$-algebras, which as a special case gives the above mentioned isomorphism.

Musson [2005] and Boyarchenko [2007] have studied a certain localization of $eH_{\boldsymbol{h}}(\mathbb{Z}_m)e$ by using the formalism of directed algebras (or $\mathbb{Z}$-algebras). Analogous localizations for finite $\mathscr{W}$-algebras were established by Ginzburg [2009]. Recently, Dodd and Kremnizer [2009] described a localization theorem for finite $\mathscr{W}$-algebras in the spirit of Kashiwara–Rouquier, and in particular for the finite W-algebra isomorphic to $eH_{\boldsymbol{h}}(\mathbb{Z}_m)e$. However, their result is via a different quantum Hamiltonian reduction than the one used in Theorem 6.3.

In [Kuwabara 2010], the second author gives an explicit description of the standard modules $\Delta(e_i)$ and simple modules $L(e_i)$ as sheaves of $\mathscr{A}_{\boldsymbol{h}}$-modules on the minimal resolution.

Acknowledgments

The authors would like to express their gratitude to Professors Iain Gordon, Masaki Kashiwara and Raphael Rouquier for their helpful comments and many fruitful discussions. They would also like to thank Professors Bernd Sturmfels and Tamás Hausel for outlining the proof of Lemma 4.7. The first author thanks Maurizio Martino and Olaf Schnürer for stimulating discussions.

References

[Beĭlinson and Bernstein 1981] A. Beĭlinson and J. Bernstein, "Localisation de g-modules", *C. R. Acad. Sci. Paris Sér. I Math.* **292**:1 (1981), 15–18. MR 82k:14015 Zbl 0476.14019

[Van den Bergh 1991] M. Van den Bergh, "Differential operators on semi-invariants for tori and weighted projective spaces", pp. 255–272 in *Topics in invariant theory* (Paris, 1989/1990), edited by M.-P. Malliavin-Brameret, Lecture Notes in Math. **1478**, Springer, Berlin, 1991. MR 93h:16046 Zbl 0802.13005

[Bielawski and Dancer 2000] R. Bielawski and A. S. Dancer, "The geometry and topology of toric hyperkähler manifolds", *Comm. Anal. Geom.* **8**:4 (2000), 727–760. MR 2002c:53078 Zbl 0992.53034

[Björk 1979] J.-E. Björk, *Rings of differential operators*, North-Holland Mathematical Library **21**, North-Holland Publishing Co., Amsterdam, 1979. MR 82g:32013 Zbl 0499.13009

[Boyarchenko 2007] M. Boyarchenko, "Quantization of minimal resolutions of Kleinian singularities", *Adv. Math.* **211**:1 (2007), 244–265. MR 2008f:14001 Zbl 1138.16012

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1993. MR 95h:13020 Zbl 0788.13005

[Cassens and Slodowy 1998] H. Cassens and P. Slodowy, "On Kleinian singularities and quivers", pp. 263–288 in *Singularities* (Oberwolfach, 1996), edited by V. I. Arnold et al., Progr. Math. **162**, Birkhäuser, Basel, 1998. MR 99k:14003 Zbl 0957.14004

[Cox et al. 2007] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra*, 3rd ed., Undergraduate Texts in Mathematics, Springer, New York, 2007. MR 2007h:13036 Zbl 1118.13001

[Crawley-Boevey and Holland 1998] W. Crawley-Boevey and M. P. Holland, "Noncommutative deformations of Kleinian singularities", *Duke Math. J.* **92**:3 (1998), 605–635. MR 99f:14003 Zbl 0974.16007

[Dodd and Kremnizer 2009] C. Dodd and K. Kremnizer, "A localization theorem for finite W-algebras", preprint, 2009. arXiv 0911.2210v1

[Fu 2006] B. Fu, "A survey on symplectic singularities and symplectic resolutions", *Ann. Math. Blaise Pascal* **13**:2 (2006), 209–236. MR 2007j:14019 Zbl 1116.14008

[Ginzburg 2003] V. Ginzburg, "On primitive ideals", *Selecta Math. (N.S.)* **9**:3 (2003), 379–407. MR 2005f:16039 Zbl 1041.16016

[Ginzburg 2009] V. Ginzburg, "Harish-Chandra bimodules for quantized Slodowy slices", *Represent. Theory* **13** (2009), 236–271. MR 2010g:17009 Zbl 05706397

[Ginzburg et al. 2009] V. Ginzburg, I. Gordon, and J. T. Stafford, "Differential operators and Cherednik algebras", *Selecta Math. (N.S.)* **14**:3-4 (2009), 629–666. MR 2011b:16089 Zbl 1193.14007

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Hausel and Sturmfels 2002] T. Hausel and B. Sturmfels, "Toric hyperKähler varieties", *Doc. Math.* **7** (2002), 495–534. MR 2004i:53054 Zbl 1029.53054

[Hodges 1993] T. J. Hodges, "Noncommutative deformations of type-A Kleinian singularities", *J. Algebra* **161**:2 (1993), 271–290. MR 94i:14038 Zbl 0807.16029

[Holland 1999] M. P. Holland, "Quantization of the Marsden–Weinstein reduction for extended Dynkin quivers", *Ann. Sci. École Norm. Sup.* (4) **32**:6 (1999), 813–834. MR 2001a:16042 Zbl 1036.16024

[Hotta et al. 2008] R. Hotta, K. Takeuchi, and T. Tanisaki, *D-modules, perverse sheaves, and representation theory*, Progress in Mathematics **236**, Birkhäuser, Boston, MA, 2008. MR 2008k:32022 Zbl 1136.14009

[Kashiwara and Rouquier 2008] M. Kashiwara and R. Rouquier, "Microlocalization of rational Cherednik algebras", *Duke Math. J.* **144**:3 (2008), 525–573. MR 2010a:16053 Zbl 1147.14002

[Kempf 1978] G. R. Kempf, "Instability in invariant theory", *Ann. of Math.* (2) **108**:2 (1978), 299–316. MR 80c:20057 Zbl 0406.14031

[Kontsevich 2001] M. Kontsevich, "Deformation quantization of algebraic varieties", *Lett. Math. Phys.* **56**:3 (2001), 271–294. MR 2002j:53117 Zbl 1081.14500

[Kraft 1984] H. Kraft, *Geometrische Methoden in der Invariantentheorie*, Aspects of Mathematics **D1**, Vieweg, Braunschweig, 1984. MR 86j:14006 Zbl 0569.14003

[Kronheimer 1989] P. B. Kronheimer, "The construction of ALE spaces as hyper-Kähler quotients", *J. Differential Geom.* **29**:3 (1989), 665–683. MR 90d:53055 Zbl 0671.53045

[Kuwabara 2008] T. Kuwabara, "Characteristic cycles of standard modules for the rational Cherednik algebra of type $\mathbb{Z}/l\mathbb{Z}$", *J. Math. Kyoto Univ.* **48**:1 (2008), 167–217. MR 2009k:16055 Zbl 1190.16042

[Kuwabara 2010] T. Kuwabara, "Representation theory of rational cherednik algebras of type $\mathbb{Z}/l\mathbb{Z}$ via microlocal analysis", preprint, 2010. To appear in *Publ. Res. Int. Math. Sci.* arXiv 1003.3407

[Losev 2012] I. Losev, "Isomorphisms of quantizations via quantization of resolutions", *Adv. Math.* **231**:3-4 (2012), 1216–1270. MR 2964603

[McConnell and Robson 2001] J. C. McConnell and J. C. Robson, *Noncommutative Noetherian rings*, Revised ed., Graduate Studies in Mathematics **30**, American Mathematical Society, Providence, RI, 2001. MR 2001i:16039 Zbl 0980.16019

[Musson 2005] I. M. Musson, "Hilbert schemes and noncommutative deformations of type A Kleinian singularities", *J. Algebra* **293**:1 (2005), 102–129. MR 2006g:16066 Zbl 1082.14008

[Musson and Van den Bergh 1998] I. M. Musson and M. Van den Bergh, "Invariants under tori of rings of differential operators and related topics", *Mem. Amer. Math. Soc.* **136**:650 (1998), viii+85. MR 99i:16051 Zbl 0928.16019

[Namikawa 2001] Y. Namikawa, "Extension of 2-forms and symplectic varieties", *J. Reine Angew. Math.* **539** (2001), 123–147. MR 2002i:32011 Zbl 0996.53050

[Oda and Park 1991] T. Oda and H. S. Park, "Linear Gale transforms and Gel′fand–Kapranov–Zelevinskij decompositions", *Tohoku Math. J.* (2) **43**:3 (1991), 375–399. MR 92d:14042 Zbl 0782 52006

[Polesello and Schapira 2004] P. Polesello and P. Schapira, "Stacks of quantization-deformation modules on complex symplectic manifolds", *Int. Math. Res. Not.* **2004**:49 (2004), 2637–2664. MR 2005e:32018 Zbl 1086.53107

[Premet 2002] A. Premet, "Special transverse slices and their enveloping algebras", *Adv. Math.* **170**:1 (2002), 1–55. MR 2003k:17014 Zbl 1005.17007

[Proudfoot 2008] N. J. Proudfoot, "A survey of hypertoric geometry and topology", pp. 323–338 in *Toric topology*, edited by M. Harada et al., Contemp. Math. **460**, Amer. Math. Soc., Providence, RI, 2008. MR 2010a:14087 Zbl 1151.53042

[Proudfoot and Webster 2007] N. Proudfoot and B. Webster, "Intersection cohomology of hypertoric varieties", *J. Algebraic Geom.* **16**:1 (2007), 39–63. MR 2008m:14037 Zbl 1119.14016

[Rouquier 2008] R. Rouquier, "q-Schur algebras and complex reflection groups", *Mosc. Math. J.* **8**:1 (2008), 119–158, 184. MR 2010b:20081 Zbl 1213.20007

[Smith 1990] S. P. Smith, "A class of algebras similar to the enveloping algebra of sl(2)", *Trans. Amer. Math. Soc.* **322**:1 (1990), 285–314. MR 91b:17013 Zbl 0732.16019

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

Received July 22, 2011.

GWYN BELLAMY
SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF GLASGOW
UNIVERSITY GARDENS
GLASGOW
G12 8QW
UNITED KINGDOM

gwyn.bellamy@glasgow.ac.uk

TOSHIRO KUWABARA
DEPARTMENT OF MATHEMATICS
NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS
7 VAVILOVA STREET
MOSCOW
117312
RUSSIA

toshiro.kuwa@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.129

ALMOST FACTORIALITY OF INTEGRAL DOMAINS AND KRULL-LIKE DOMAINS

GYU WHAN CHANG, HWANKOO KIM AND JUNG WOOK LIM

Let D be an integral domain, $\bar{D}$ be the integral closure of D, and Γ be a numerical semigroup with $\Gamma \subsetneq \mathbb{N}_0$. Let t be the so-called t-operation on D. We will say that D is an AK-domain (resp., AUF-domain) if for each nonzero ideal $(\{a_\alpha\})$ of D, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})_t$ is t-invertible (resp., principal). In this paper, we study several properties of AK-domains and AUF-domains. Among other things, we show that if $D \subseteq \bar{D}$ is a bounded root extension, then D is an AK-domain (resp., AUF-domain) if and only if $\bar{D}$ is a Krull domain (resp., Krull domain with torsion t-class group) and D is t-linked under $\bar{D}$. We also prove that if D is a Krull domain (resp., UFD) with $\operatorname{char}(D) \neq 0$, then the (numerical) semigroup ring $D[\Gamma]$ is a nonintegrally closed AK-domain (resp., AUF-domain).

Introduction

Throughout this paper, D is an integral domain with quotient field K, $\bar{D}$ denotes the integral closure of D in K, X is an indeterminate over D, $D[X]$ is the polynomial ring over D, $\mathbb{N}_0$ (resp., $\mathbb{Z}$) is the set of nonnegative integers (resp., integers), Γ is a numerical semigroup with $\Gamma \subsetneq \mathbb{N}_0$ and $D[\Gamma]$ means the numerical semigroup ring of Γ over D.

We say that D is a *GCD-domain* if $aD \cap bD$ is principal for all $0 \neq a, b \in D$. In [Zafrullah 1985], Zafrullah introduced the notion of an almost GCD-domain (AGCD-domain). He called D an *AGCD-domain* if for each $0 \neq a, b \in D$, there exists an integer $n = n(a, b) \geq 1$ such that $a^n D \cap b^n D$ is principal. After Zafrullah's paper [1985], several types of almost divisibility of integral domains have been studied, for example, AB-domains, AP-domains, APvMDs, API-domains and AD-domains (see Section 1). Recall from [Kang 1989a] that D is a Krull domain (resp., UFD) if and only if for every nonzero ideal I of D, I_t is t-invertible (resp.,

The second author was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0011996). The third author was supported by the Brain Korea 21 Project Team to Nurture the Next Generation of First-class Mathematical Scientists by the Korean Government.
MSC2010: primary 13F05; secondary 13A15, 13B25, 13G05.
Keywords: AK-domain, AUF-domain, numerical semigroup, bounded root extension.

principal). In this paper, we define D to be an *AK-domain* (resp., *AUF-domain*) if for each nonzero ideal $(\{a_\alpha\})$ of D, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})_t$ is t-invertible (resp., principal). (For the sake of convenience, we will use the notation $\{a_\alpha\}$ instead of $\{a_\alpha\}_{\alpha\in\Lambda}$, where Λ is an indexed set. Also, for a nonempty subset $\{a_\alpha\}$ of $D \setminus \{0\}$, we mean by $(\{a_\alpha\}D)$ the ideal of D generated by the set $\{a_\alpha\}$.)

In Section 1, we review multifarious integral domains related to the theory of almost divisibility and some results on them.

We devote Section 2 to the study of AK-domains. Precisely, we show that if $D \subseteq \overline{D}$ is a bounded root extension, then D is an AK-domain if and only if $\overline{D}$ is a Krull domain and D is t-linked under $\overline{D}$, if and only if (i) t-$\dim(D) = 1$, (ii) D_P is an API-domain for each $P \in t$-$\mathrm{Max}(D)$ and (iii) $D = \bigcap_{P\in t\text{-}\mathrm{Max}(D)} D_P$ and this intersection has finite character. We prove that if D is a Krull domain, then $D[\Gamma]$ is an AK-domain if and only if $\mathrm{char}(D) \neq 0$. This result can be used to construct a simple example of nonintegrally closed AK-domains. We also prove that the (numerical) semigroup ring $D[\Gamma]$ is an AK-domain if and only if $D[X]$ is an AK-domain and $\mathrm{char}(D) \neq 0$.

In Section 3, we introduce the notions of an AUF-domain and an almost π-domain. We show that D is an AUF-domain (resp., almost π-domain) if and only if D is an AK-domain and $Cl(D)$ is torsion (resp., $G(D)$ is torsion). We prove that $D[\Gamma]$ is an AUF-domain if and only if $D[X]$ is an AUF-domain and $\mathrm{char}(D) \neq 0$. Also, we show that if D is a UFD (resp., π-domain), then $D[\Gamma]$ is an AUF-domain (resp., almost π-domain) if and only if $\mathrm{char}(D) \neq 0$. Finally, we give an example of an AUF-domain that is neither integrally closed nor an API-domain.

Now, we review some definitions and notation. Let $\mathbf{F}(D)$ be the set of nonzero fractional ideals of D. For an $I \in \mathbf{F}(D)$, we denote by I^{-1} the fractional ideal $\{x \in K \mid xI \subseteq D\}$ of D. Recall that the *v-operation* on D is the mapping on $\mathbf{F}(D)$ defined by $I \mapsto I_v = (I^{-1})^{-1}$, and the *$t$-operation* on D is the mapping on $\mathbf{F}(D)$ defined by $I \mapsto I_t = \bigcup\{J_v \mid J$ is a nonzero finitely generated fractional subideal of $I\}$. Clearly, if an $I \in \mathbf{F}(D)$ is finitely generated, then $I_t = I_v$. An $I \in \mathbf{F}(D)$ is called a *t-ideal* (resp., *v-ideal*) if $I_t = I$ (resp., $I_v = I$). An $I \in \mathbf{F}(D)$ is said to be *v-finite type* if $I_v = J_v$ for some finitely generated ideal J of D. A t-ideal M of D is called a *maximal t-ideal* if M is maximal among proper integral t-ideals of D. Let t-$\mathrm{Max}(D)$ be the set of maximal t-ideals of D. It is known that t-$\mathrm{Max}(D) \neq \varnothing$ if D is not a field; a prime ideal minimal over a t-ideal is a t-ideal; a maximal t-ideal is a prime ideal; and each proper integral t-ideal is contained in a maximal t-ideal. We say that D has *t-dimension one*, denoted by t-$\dim(D) = 1$, if each maximal t-ideal is of height one. An $I \in \mathbf{F}(D)$ is said to be *t-invertible* if $(II^{-1})_t = D$; equivalently, $II^{-1} \nsubseteq M$ for all $M \in t$-$\mathrm{Max}(D)$. It was shown that an $I \in \mathbf{F}(D)$ is t-invertible if and only if $I_v = J_v$ for some finitely generated fractional ideal J of

D and ID_M is principal for each $M \in t\text{-Max}(D)$ [Kang 1989b, Corollary 2.7]. We say that D is a *Prüfer v-multiplication domain* (PvMD) if each nonzero finitely generated ideal of D is t-invertible. A nonzero prime ideal Q of $D[X]$ is called an *upper to zero* in $D[X]$ if $Q \cap D = (0)$. As in [Houston and Zafrullah 1989], we say that D is a *UMT-domain* if every upper to zero in $D[X]$ is t-invertible. It is known that D is an integrally closed UMT-domain if and only if D is a PvMD [Houston and Zafrullah 1989, Proposition 3.2]

Let $T(D)$ be the abelian group of t-invertible fractional t-ideals of D under the t-multiplication $I * J = (IJ)_t$ and $\text{Inv}(D)$ (resp., $\text{Prin}(D)$) be the subgroup of $T(D)$ of invertible (resp., principal) fractional ideals of D. Then it is obvious that $\text{Prin}(D) \subseteq \text{Inv}(D) \subseteq T(D)$. The *$t$-class group* of D is an abelian group $Cl(D) = T(D)/\text{Prin}(D)$ and the *Picard group* $\text{Pic}(D) = \text{Inv}(D)/\text{Prin}(D)$ of D is a subgroup of $Cl(D)$. The *local t-class group* $G(D)$ of D is defined by $G(D) = Cl(D)/\text{Pic}(D)$.

Let $D \subseteq E$ be an extension of integral domains. Then E is said to be a *root extension* of D if for each $z \in E$, $z^n \in D$ for some $n \geq 1$. We say that $D \subseteq E$ is a *bounded root extension* if there exists a fixed positive integer n such that $z^n \in D$ for all $z \in E$. The domain D is said to be *root closed* if for $z \in K$, $z^n \in D$ for some integer $n \geq 1$ implies $z \in D$.

A *numerical semigroup* is a nonempty subset Γ of $\mathbb{N}_0$ that is closed under addition, contains 0 and generates $\mathbb{Z}$ as a group. It is known that the set $\mathbb{N}_0 \setminus \Gamma$ is finite, and Γ has a unique numerical system of generators. Also, there always exists the largest nonnegative integer that is not contained in Γ. Such an integer is called the *Frobenius number* of Γ and is denoted by $F(\Gamma)$. If $\Gamma \subsetneq \mathbb{N}_0$, then $D + X^{F(\Gamma)+1}D[X] \subseteq D[\Gamma] \subsetneq D[X]$.

Our general reference for results from multiplicative ideal theory will be [Gilmer 1992]. For any undefined terms, readers are referred to [Kaplansky 1994].

1. Almost divisibility of integral domains

In multiplicative ideal theory, one of the important topics during the past few decades was the theory of factorizations in integral domains. Among various kinds of integral domains, many mathematicians have studied Bézout domains, Prüfer domains, principal ideal domains (PID) and Dedekind domains. As for the t-operation analogues, they have also investigated GCD-domains, PvMDs, generalized GCD-domains (GGCD-domain), unique factorization domains (UFD), π-domains and Krull domains.

In [Storch 1967], almost factorial domains were studied as Krull domains with torsion divisor class groups. Motivated by this, Zafrullah first began to study a general theory of almost factoriality. To do this, he first defined an *almost GCD-domain* (AGCD-domain) that is an integral domain D in which for each $0 \neq a, b \in D$, there exists an integer $n = n(a, b) \geq 1$ such that $(a^n, b^n)_t$, equivalently $a^n D \cap b^n D$, is principal [Zafrullah 1985]. He also showed that D is an integrally closed AGCD-domain if and only if D is a PvMD with torsion t-class group [Zafrullah 1985, Corollary 3.8 and Theorem 3.9].

Anderson and Zafrullah continued the investigation of AGCD-domains and introduced several closely related domains. They proved that D is an AGCD-domain if and only if $\bar{D}$ is an AGCD-domain [Anderson and Zafrullah 1991, Section 4], $D \subseteq \bar{D}$ is a root extension and D is t-linked under $\bar{D}$ [ibid., Theorem 5.9]. (See the remark after Example 2.4 for the definition of "t-linked under".) They also defined AB-domains and AP-domains as follows: D is an *almost Bézout domain* (AB-domain) (resp., *almost Prüfer domain* (AP-domain)) if for each $0 \neq a, b \in D$, there exists a positive integer $n = n(a, b)$ such that (a^n, b^n) is principal (resp., invertible). They proved that D is an AB-domain if and only if D is an AP-domain with torsion t-class group [ibid., Lemma 4.4]. They also showed that D is an AP-domain (resp., AB-domain) if and only if $\bar{D}$ is a Prüfer domain (resp., with torsion (t-)class group) and $D \subseteq \bar{D}$ is a root extension [ibid., Corollary 4.8].

Recently, Li introduced the concept of almost Prüfer v-multiplication domains. She defined D to be an *almost Prüfer v-multiplication domain* (APvMD) if for each $0 \neq a, b \in D$, there exists a positive integer $n = n(a, b)$ such that $(a^n, b^n)_v$, equivalently $a^n D \cap b^n D$, is t-invertible [Li 2012, Definition 2.1 and Theorem 2.3]. It was shown that D is an AGCD-domain if and only if D is an APvMD with torsion t-class group [Li 2012, Theorem 3.1].

The notion of an almost generalized GCD-domain was first introduced by Anderson and Zafrullah [1991, Section 3], and was also investigated by Lewin [1997]. (Recall that D is an *almost generalized GCD-domain* (AGGCD-domain) if for each $0 \neq a, b \in D$, there is an integer $n = n(a, b) \geq 1$ such that $(a^n, b^n)_v$ is invertible.) In [Chang et al. 2012], we studied AGGCD-domains further. We showed that D is an AGGCD-domain if and only if D is an APvMD and $G(D)$ is torsion [Chang et al. 2012, Theorem 2.11]. This result corrects an error in [Lewin 1997, Theorem 5.2], which incorrectly states that D is an integrally closed AGGCD-domain if and only if D is a PvMD and $G(D) = 0$. (See the review of [Lewin 1997] in Mathematical Reviews database for more details.)

Assume that D is integrally closed. If D is an AB-domain or an AP-domain, then D is a Prüfer domain [Anderson and Zafrullah 1991, Theorem 4.7]. Also, if D is an APvMD, an AGCD-domain or an AGGCD-domain, then D is a PvMD [Li 2012, Theorem 2.4; Lewin 1997, Theorem 5.2]. As mentioned above, Prüfer domains and PvMDs have been much studied in the context of factorization theory, and there are many well-known results about them. So, from this point of view, integrally closed domains are no longer of interest in the theory of almost factoriality. This is why we need to investigate the almost divisibility of nonintegrally closed domains. One of the requisites to study them is to find some examples of such domains. Among several methods, in [Chang et al. 2012], the authors gave simple examples of nonintegrally closed APvMDs, AGCD-domains, AGGCD-domains, AP-domains and AB-domains via the $D+X^nK[X]$ constructions. In fact, they proved that for an integer $n \geq 2$, $D+X^nK[X]$ is an APvMD (resp., AGCD-domain, AGGCD-domain, AP-domain, AB-domain) if and only if D is an APvMD (resp., AGCD-domain, AGGCD-domain, AP-domain, AB-domain) and $\operatorname{char}(D) \neq 0$ [Chang et al. 2012, Theorem 2.6 and Corollaries 2.10 and 2.13].

Anderson and Zafrullah [1991] defined an *AD-domain* (resp., *almost principal ideal domain* (API-domain)) to be a domain D such that for any nonempty subset $\{a_\alpha\} \subseteq D \setminus \{0\}$ there exists a positive integer $n = n(\{a_\alpha\})$ with $(\{a_\alpha^n\})$ invertible (resp., principal). They showed that D is an API-domain if and only if D is an AD-domain with torsion t-class group [Anderson and Zafrullah 1991, Lemma 4.4]. Moreover, if $D \subseteq \bar{D}$ is a bounded root extension, then D is an AD-domain (resp., API-domain) if and only if $\bar{D}$ is a Dedekind domain (resp., $\bar{D}$ is a Dedekind domain with torsion class group) [ibid., Corollary 4.13].

As “almost” versions of Krull domains and UFDs, we will define and investigate AK-domains and AUF-domains in Sections 2 and 3, respectively. We also provide an example of AK-domains (resp., AUF-domains) which is not a Krull domain (resp., UFD).

2. AK-domains

Let D be an integral domain with quotient field K, $\bar{D}$ be the integral closure of D in K, $X^1(D)$ be the set of height-one prime ideals of D, Γ be a numerical semigroup with $\Gamma \subsetneq \mathbb{N}_0$, and $D[\Gamma]$ be the (numerical) semigroup ring of Γ over D.

We say that D is an *AK-domain* if for each nonzero ideal $(\{a_\alpha\})$ of D, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})$, equivalently, $(\{a_\alpha^n\})_t$ is t-invertible. We have avoided using the terminology “almost Krull” for an AK-domain because the term “almost Krull” is already used to mean an integral domain that is locally Krull in the literature [Pirtle 1968]. Clearly, AK-domains are APvMDs; so if D is an AK-domain, then D is a UMT-domain and $D \subseteq \bar{D}$ is a root extension [Li 2012, Theorem 3.8]. Also, a Krull domain is an AK-domain, but not vice versa. For example, if $m \equiv 5 \pmod 8$, then $\mathbb{Z}[\sqrt{m}]$ is a nonintegrally closed API-domain [Anderson and Zafrullah 1991, Theorem 4.17]; so $\mathbb{Z}[\sqrt{m}]$ is an AK-domain that is not a Krull domain. Note that if D is integrally closed, then $(\{a_\alpha^n\})_t = ((\{a_\alpha\})^n)_t$ for any integer $n \geq 1$ and any nonempty subset $\{a_\alpha\}$ of $D \setminus \{0\}$ [Anderson and Zafrullah 1991, Corollary 6.4]. Thus, an integrally closed AK-domain is a Krull domain.

Our first result is the AK-domain analogue of the well-known fact that D is a Dedekind domain if and only if D is a Krull domain and each maximal ideal is a t-ideal.

Proposition 2.1. *D is an AD-domain if and only if D is an AK-domain and each maximal ideal of D is a t-ideal.*

Proof. ($\Rightarrow$) Since an AD-domain is an AP-domain, by [Chang et al. 2012, Remark before Theorem 2.6], each maximal ideal of D is a t-ideal. Clearly, D is an AK-domain.

($\Leftarrow$) Let $(\{a_\alpha\})$ be a nonzero ideal of D. Since D is an AK-domain, there is a positive integer $m = m(\{a_\alpha\})$ such that $(\{a_\alpha^m\})$ is t-invertible. So $(\{a_\alpha^m\})(\{a_\alpha^m\})^{-1} \nsubseteq P$ for all $P \in t\text{-Max}(D)$, and since each maximal ideal is a t-ideal, $(\{a_\alpha^m\})(\{a_\alpha^m\})^{-1} \nsubseteq M$ for all maximal ideals M of D. Thus $(\{a_\alpha^m\})(\{a_\alpha^m\})^{-1} = D$. □

Let D be a Noetherian domain. It is known that D is a UMT-domain if and only if $t\text{-dim}(D) = 1$ [Houston and Zafrullah 1989, Theorem 3.7]; so if D is an AK-domain, then $t\text{-dim}(D) = 1$. But, we do not know whether an AK-domain is generally of t-dimension one.

We next give a characterization of AK-domains under the "t-dimension one" assumption, which turns out to be very similar to that of a Krull domain as described in its definition: D is called a *Krull domain* if (i) $D = \bigcap_{P \in X^1(D)} D_P$, (ii) D_P is a local PID for each $P \in X^1(D)$ and (iii) the intersection $D = \bigcap_{P \in X^1(D)} D_P$ has finite character, i.e., each nonzero element $d \in D$ is a unit in D_P for all but a finite number of P's in $X^1(D)$.

Theorem 2.2. *If $t\text{-dim}(D) = 1$, then D is an AK-domain if and only if*

(1) *D_P is an API-domain for each $P \in t\text{-Max}(D)$, and*

(2) *$D = \bigcap_{P \in t\text{-Max}(D)} D_P$ and this intersection has finite character.*

Proof. ($\Rightarrow$) (1) Let $\{a_\alpha\}$ be a nonempty subset of $D \setminus \{0\}$. Since D is an AK-domain, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})$ is t-invertible. Hence $(\{a_\alpha^n\})D_P$ is principal for each $P \in t\text{-Max}(D)$ [Kang 1989b, Corollary 2.7]. Thus D_P is an API-domain.

(2) Note that $D = \bigcap_{P \in t\text{-Max}(D)} D_P$ [Kang 1989b, Proposition 2.9]. If $P = (\{a_\alpha\}) \in t\text{-Max}(D)$, then there is an integer $n = n(P) \geq 1$ such that $(\{a_\alpha^n\})$ is t-invertible, and hence $(\{a_\alpha^n\})_t = (a_1, \dots, a_k)_t$ for some $a_1, \dots, a_k \in P$. Clearly, $P = \sqrt{(a_1, \dots, a_k)_t}$. If $0 \neq d \in D$, then each maximal t-ideal of D containing d is minimal over dD because $t\text{-dim}(D) = 1$. Thus D has a finite number of maximal t-ideals that contain d [Sahandi 2010, Corollary 2.5], since each maximal t-ideal is the radical of a finite type t-ideal.

($\Leftarrow$) Let $(\{a_\alpha\})$ be a nonzero ideal of D. By (2), the number of maximal t-ideals of D containing $\{a_\alpha\}$ is finite, say $P_1, \dots, P_k$. Since each D_{P_i} is an API-domain by (1), there exists a positive integer m_i such that $(\{a_\alpha^{m_i}\})D_{P_i}$ is principal for each $i = 1, \dots, k$, say $(\{a_\alpha^{m_i}\})D_{P_i} = (a_i)D_{P_i}$ for some $a_i \in \{a_\alpha^{m_i}\}$. Let $m = m_1 \cdots m_k$ and set $\hat{m}_i = m/m_i$ for each $i = 1, \dots, k$. Then $(\{a_\alpha^m\})D_{P_i} = (a_i^{\hat{m}_i})D_{P_i}$ for all $i = 1, \dots, k$ [Gilmer 1992, Theorem 6.5(c)]; so $(\{a_\alpha^m\})$ is t-locally principal and $(\{a_\alpha^m\})$ is contained in only a finite number of maximal t-ideals of D. Choose any nonzero element $a \in (\{a_\alpha^m\})$. By (2), there exist only finitely many maximal t-ideals $M_1, \dots, M_n$ of D containing a. Let P be a maximal t-ideal of D which is distinct from all P_i, $i = 1, \dots, k$. If $P \notin \{M_1, \dots, M_n\}$, then $a \notin P$, and hence $(a)D_P = D_P$. Suppose that $P \in \{M_1, \dots, M_n\}$, say $P = M_j$ for some $1 \leq j \leq n$. Then, for some $b_j \in (\{a_\alpha^m\}) \setminus M_j$, we have $(b_j)D_{M_j} = D_{M_j}$. Let I be the ideal of D generated by $a, a_1^{\hat{m}_1}, \dots, a_k^{\hat{m}_k}$ and the b_j's (if needed). Then I is contained in $(\{a_\alpha^m\})$ and I is a finitely generated ideal of D, because the number of b_j's cannot exceed n. Let P be a maximal t-ideal such that $P \neq P_i$ for all $1 \leq i \leq k$. If $a \notin P$, then $ID_P = D_P = (\{a_\alpha^m\})D_P$. If $a \in P$, then $P = M_j$ for some $1 \leq j \leq n$; so $ID_P = D_P = (\{a_\alpha^m\})D_P$ by the choice of a suitable b_j. Note that $ID_{P_i} = (\{a_\alpha^m\})D_{P_i}$ for each $1 \leq i \leq k$. Thus $ID_M = (\{a_\alpha^m\})D_M$ for all maximal t-ideals M of D. It follows from [Kang 1989b, Proposition 2.8(3)] that $(\{a_\alpha^m\})_t = I_t$, i.e., $(\{a_\alpha^m\})_t$ is a finite type t-ideal of D. Hence $(\{a_\alpha^m\})$ is t-invertible [Kang 1989b, Corollary 2.7], and thus D is an AK-domain. □

Recall that D is a *weakly Krull domain* if $D = \bigcap_{P \in X^1(D)} D_P$ and this intersection has finite character. Note that if t-$\dim(D) = 1$, then $X^1(D) = t$-$\mathrm{Max}(D)$. Thus, by Theorem 2.2, we have

Corollary 2.3. *An AK-domain with t-dimension one is a weakly Krull domain.*

However, the converse of Corollary 2.3 does not hold, and we provide such an example.

Example 2.4. Let $\mathbb{Q}$ (resp., $\mathbb{R}$) be the field of rational (resp., real) numbers, X be an indeterminate over $\mathbb{R}$, and $D = \mathbb{Q} + X\mathbb{R}[X]$.

(1) Note that $\mathbb{R}[X]$ is a Krull domain, and hence a weakly Krull domain. Thus D is a weakly Krull domain [Anderson et al. 2006, Theorem 3.4], whence t-$\dim(D) = 1$ [Anderson et al. 1992, Lemma 2.1].

(2) Let $F = \{a \in \mathbb{R} \mid a \text{ is integral over } \mathbb{Q}\}$. Clearly, F is an integral domain. Let $0 \neq a \in F$. Then $a^n + q_{n-1}a^{n-1} + \cdots + q_0 = 0$ for some integer $n \geq 1$ and $q_0, \dots, q_{n-1} \in \mathbb{Q}$ with $q_0 \neq 0$; so

$$\left(\frac{1}{a}\right)^n + \frac{q_1}{q_0}\left(\frac{1}{a}\right)^{n-1} + \cdots + \frac{1}{q_0} = 0.$$

Hence $\frac{1}{a}$ is integral over $\mathbb{Q}$; so F is a field. Also, it is easy to see that $\overline{D} = F + X\mathbb{R}[X]$, because $\mathbb{R}[X]$ is integrally closed and the quotient field of D contains $\mathbb{R}[X]$. It is well known that $e \in \mathbb{R} \setminus F$, where $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$ is the base of the natural logarithm. Hence $\bar{D}$ is not a PvMD [Anderson and El Abidine 2001, Lemma 2.1]; so D is not an APvMD [Li 2012, Theorem 3.6]. Thus D is not an AK-domain.

Let $D \subseteq E$ be an extension of integral domains. Following [Dobbs et al. 1989], we say that E is *t-linked over D* if $I^{-1} = D$ for I a nonzero finitely generated ideal of D implies $(IE)^{-1} = E$; equivalently, if M is a maximal t-ideal of E with $M \cap D \neq (0)$, then $(M \cap D)_t \subsetneq D$ [Anderson et al. 1993, Proposition 2.1]. Anderson and Zafrullah introduced the concept of "t-linked under" which is the opposite notion of "t-linked over" [Anderson and Zafrullah 1991]. They defined that D is *t-linked under E* if for each nonzero finitely generated ideal I of D, $(IE)^{-1} = E$ implies $I^{-1} = D$. Clearly, D is t-linked under E if and only if $(JE)_t = E$ implies $J_t = D$ for all nonzero ideals J of D.

In [Anderson and Zafrullah 1991, Theorem 4.11], the authors showed that if $D \subseteq E$ is a bounded extension with $E \subseteq \bar{D}$, then D is an API-domain (resp., AD-domain) if and only if E is an API-domain (resp., AD-domain). Now, we give the AK-domain version of this result.

Theorem 2.5. *Let $D \subseteq E$ be a bounded root extension with $E \subseteq \bar{D}$. Assume that E is t-linked over D. Then D is an AK-domain if and only if E is an AK-domain and D is t-linked under E.*

Proof. ($\Rightarrow$) Let $\{a_\alpha\}$ be a nonempty subset of $E \setminus \{0\}$. Since $D \subseteq E$ is a bounded root extension, there exists a positive integer n such that $z^n \in D$ for all $z \in E$; so $\{a_\alpha^n\} \subseteq D$. Since D is an AK-domain, there exists a positive integer $m = m(\{a_\alpha^n\})$ such that $(\{a_\alpha^{nm}\})$ is t-invertible, and hence $((\{a_\alpha^{nm}\})J)_t = D$ for some v-finite type ideal J of D. Since E is a t-linked overring of D, $(((\{a_\alpha^{nm}\})J)E)_t = E$. Thus E is an AK-domain.

Let I be a finitely generated ideal of D such that $(IE)^{-1} = E$. Suppose to the contrary that $I_v \subsetneq D$. Then there exists a maximal t-ideal M of D such that $I_t \subseteq M$. Since $E \subseteq \bar{D}$, $M = M_0 \cap D$ for some prime ideal M_0 of E [Kaplansky 1994, Theorem 44]. Note that D is a UMT-domain [Li 2012, Theorem 3.8]. Hence M_0 is a t-ideal of E [Fontana et al. 1998, Proposition 1.4]. Therefore we have

$$E = (IE)_t \subseteq (ME)_t \subseteq (M_0)_t = M_0,$$

which is impossible. Thus $I_v = D$ (or equivalently, $I^{-1} = D$), which says that D is t-linked under E.

($\Leftarrow$) Assume that E is an AK-domain and let $(\{b_\beta\})$ be a nonzero ideal of D. Then there exists a positive integer $m = m(\{b_\beta\})$ such that $(\{b_\beta^m\}E)$ is t-invertible, and hence $((\{b_\beta^m\}E)J)_t = E$ for some v-finite ideal J of E. Let $J = (\{j_1, \dots, j_l\}E)_t$, where $0 \neq j_1, \dots, j_l \in K$. Then we have

$$(\{b_\beta^m j_i\}E)_t = ((\{b_\beta^m\}E)(\{j_1, \dots, j_l\}E))_t = ((\{b_\beta^m\}E)J)_t = E.$$

Since $D \subseteq E$ is a bounded root extension, there is an integer $n \geq 1$ such that $b_\beta^{mn} j_i^n = (b_\beta^m j_i)^n \in D$ for all $b_\beta^m j_i$. Clearly, $(\{b_\beta^{mn} j_i^n\}E)_t = E$; so

$$((\{b_\beta^{mn}\}D)(\{j_i^n\}D))_t = (\{b_\beta^{mn} j_i^n\}D)_t = D,$$

because D is t-linked under E. Thus D is an AK-domain. $\square$

Corollary 2.6. *If $D \subseteq \bar{D}$ is a bounded root extension, then the following statements are equivalent.*

(1) *D is an AK-domain.*

(2) *$\bar{D}$ is a Krull domain and D is t-linked under $\bar{D}$.*

(3) (a) t-$\dim(D) = 1$,

(b) *D_P is an API-domain for each $P \in t$-$\mathrm{Max}(D)$, and*

(c) *$D = \bigcap_{P \in t\text{-}\mathrm{Max}(D)} D_P$ and this intersection has finite character.*

Proof. (1) $\Leftrightarrow$ (2) Note that $\bar{D}$ is t-linked over D [Anderson et al. 1993, Proposition 2.4], because $D \subseteq \bar{D}$ is a (bounded) root extension. Also, recall that an integrally closed AK-domain is a Krull domain. Thus the result follows from Theorem 2.5.

(2) $\Rightarrow$ (3) By Theorem 2.2, it suffices to show that t-$\dim(D) = 1$. If t-$\dim(D) \geq 2$, then there are prime t-ideals P_1, P_2 of D with $P_1 \subsetneq P_2$. Let Q_1, Q_2 be prime ideals of $\bar{D}$ such that $Q_i \cap D = P_i$ for $i = 1, 2$. Note that D is t-linked under $\bar{D}$; so Q_2 is a t-ideal, which means that t-$\dim(\bar{D}) \geq 2$, a contradiction. Thus t-$\dim(D) = 1$.

(3) $\Rightarrow$ (1) Theorem 2.2. $\square$

Lemma 2.7. *$D[\Gamma] \subsetneq D[X]$ is a (bounded) root extension if and only if* $\mathrm{char}(D) \neq 0$.

Proof. If $D[\Gamma] \subsetneq D[X]$ is a root extension, then there is a positive integer n such that $(1 + X)^n \in D[\Gamma]$; so $nX \in D[\Gamma]$. Thus $\mathrm{char}(D) \neq 0$. Conversely, assume that $\mathrm{char}(D) = p \neq 0$, and let m be a positive integer such that $p^m \geq F(\Gamma) + 1$. If $a + Xg \in D[X]$, where $a \in D$ and $g \in D[X]$, then $(a + Xg)^{p^m} = a^{p^m} + X^{p^m} g^{p^m}$ lies in $D[\Gamma]$. Thus $D[\Gamma] \subsetneq D[X]$ is a (bounded) root extension. $\square$

Let t-$\mathrm{Spec}(A)$ be the set of prime t-ideals of an integral domain A. In [Chang et al. 2012, Theorem 1.5], we showed that the map $\varphi : t\text{-}\mathrm{Spec}(D[X]) \to t\text{-}\mathrm{Spec}(D[\Gamma])$, given by $Q \mapsto Q \cap D[\Gamma]$, is an order-preserving bijection. This shows that $D[X]$ is t-linked over $D[\Gamma]$, and at the same time $D[\Gamma]$ is t-linked under $D[X]$.

Corollary 2.8. *If D is a Krull domain, then* $\mathrm{char}(D) \neq 0$ *if and only if $D[\Gamma]$ is an AK-domain.*

Proof. ($\Rightarrow$) Note that $\overline{D[\Gamma]} = \bar{D}[X] = D[X]$; so $\overline{D[\Gamma]}$ is a Krull domain [Gilmer 1992, Corollary 43.11((3)] and $D[\Gamma]$ is t-linked under $\overline{D[\Gamma]}$. Thus $D[\Gamma]$ is an AK-domain by Corollary 2.6 and Lemma 2.7.

($\Leftarrow$) Since an AK-domain is an APvMD, $D[\Gamma]$ is an APvMD, and thus $\mathrm{char}(D) \neq 0$ [Chang et al. 2012, Theorem 2.2]. $\square$

Corollary 2.9. *If $D[X] \subseteq \bar{D}[X]$ is a bounded root extension, then D is an AK-domain if and only if $D[X]$ is an AK-domain.*

Proof. This follows from Corollary 2.6, because D is t-linked under $\bar{D}$ if and only if $D[X]$ is t-linked under $\bar{D}[X]$ [Anderson et al. 2004, Proposition 3.3]. □

Example 2.10. Let $F \subsetneq L$ be a pair of finite fields, and let X, Y be indeterminates over L. Then $D = F + YL[\![Y]\!]$ is an API-domain [Anderson et al. 1994, Example 3.8]; so D is an AK-domain. Since $F[X] \subsetneq L[X]$ is not a root extension, $D[X] \subsetneq \bar{D}[X]$ is not a (bounded) root extension [Anderson et al. 1994, Proposition 2.4]; so $D[X]$ is not an APvMD [Li 2012, Theorem 3.8]. Thus $D[X]$ is not an AK-domain. This shows that the assumption "$D[X] \subsetneq \bar{D}[X]$ is a bounded root extension" is essential in Corollary 2.9.

The next theorem characterizes when the numerical semigroup ring $D[\Gamma]$ is an AK-domain in terms of polynomial rings.

Theorem 2.11. *$D[\Gamma]$ is an AK-domain if and only if $D[X]$ is an AK-domain and* $\mathrm{char}(D) \neq 0$.

Proof. ($\Rightarrow$) Assume that $D[\Gamma]$ is an AK-domain. Since an AK-domain is an APvMD, $\mathrm{char}(D) \neq 0$ [Chang et al. 2012, Theorem 2.2]. Let $(\{f_\alpha\})$ be a nonzero ideal of $D[X]$. Then there exists a positive integer m such that $f_\alpha^m \in D[\Gamma]$ for all f_α by Lemma 2.7. Since $D[\Gamma]$ is an AK-domain, $((\{f_\alpha^{rm}\})(\{f_\alpha^{rm}\}D[\Gamma])^{-1})_t = D[\Gamma]$ for some integer $r = r(\{f_\alpha^m\}) \geq 1$. Note that $D[X]$ is t-linked over $D[\Gamma]$; so

$$((\{f_\alpha^{rm}\})(\{f_\alpha^{rm}\}D[\Gamma])^{-1}D[X])_t = D[X].$$

Thus $D[X]$ is an AK-domain.

($\Leftarrow$) Let $(\{g_\alpha\})$ be a nonzero ideal of $D[\Gamma]$. Since $D[X]$ is an AK-domain, there exists an integer $n = n(\{g_\alpha\}) \geq 1$ such that $((\{g_\alpha^n\})JD[X])_t = D[X]$ for some v-finite type ideal J of $D[X]$. Let $J = ((f_1, \dots, f_m)D[X])_t$, where $0 \neq f_1, \dots, f_m \in K(X)$, the quotient field of $D[X]$. Then we have

$$((\{g_\alpha^n f_i\})D[X])_t = ((\{g_\alpha^n\})(f_1, \dots, f_m)D[X])_t = D[X].$$

Since $\mathrm{char}(D) \neq 0$, $D[\Gamma] \subsetneq D[X]$ is a bounded root extension by Lemma 2.7; so there is an integer $k \geq 1$ such that $g_\alpha^{kn} f_i^k \in D[\Gamma]$ for all $g_\alpha^n f_i$. Also, it is obvious that $((\{g_\alpha^{kn} f_i^k\})D[X])_t = D[X]$, because $((\{g_\alpha^n f_i\})D[X])_t = D[X]$. Finally, since $D[\Gamma]$ is t-linked under $D[X]$, we obtain

$$((\{g_\alpha^{nk}\})(\{f_i^k\})D[\Gamma])_t = ((\{g_\alpha^{nk} f_i^k\})D[\Gamma])_t = D[\Gamma].$$

Thus $D[\Gamma]$ is an AK-domain. □

It is known that D is a Krull domain if and only if each nonzero prime ideal of D contains a t-invertible prime ideal [Kang 1989a, Theorem 3.6]. Also, we know that D is a Krull domain if and only if D is a Mori PvMD [Kang 1989a, Theorem 3.2]. So we have the following natural questions on AK-domains:

(1) Assume that each nonzero prime ideal of D contains a nonzero prime ideal $P = (\{a_\alpha\})$ such that $(\{a_\alpha^n\})$ is t-invertible for some integer $n = n(P) \geq 1$. Is D an AK-domain?

(2) Is a Mori APvMD an AK-domain?

We end this section by giving negative answers to these two questions.

Example 2.12. (1) Let D be the integral domain as in Example 2.10. Then D is a UMT-domain [Li 2012, Theorem 3.8], because an API-domain is an APvMD; so $t\text{-dim}(D[X]) = 1$. Hence we have

$$t\text{-Max}(D[X]) = \{Q \mid Q \text{ is an upper to zero in } D[X]\} \cup \{P[X] \mid P \in X^1(D)\}.$$

Note that if $Q \in t\text{-Max}(D[X])$ with $Q \cap D = (0)$, then Q is t-invertible. Next, if $P = (\{a_\alpha\}) \in X^1(D)$, then there exists a positive integer $n = n(P)$ such that $(\{a_\alpha^n\})[X]$ is principal, and hence t-invertible.

(2) Let p be a prime, $F = \bigcup_{n\geq 1} GF(p^{2^n})$, $L = F(GF(p^3))$, and let $D = F + XL[\![X]\!]$. Then D is a one-dimensional quasilocal Noetherian AB-domain, but not an API-domain [Anderson et al. 1994, Example 3.6]. Thus D is a Mori APvMD, but not an AK-domain by Theorem 2.2.

3. AUF-domains

We say that D is an *AUF-domain* if for every nonzero ideal $(\{a_\alpha\})$ of D, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})_t$ is principal. (Here, an AUF-domain stands for an almost unique factorization domain.)

We start this section with the AUF-domain analogue of the fact that D is a UFD if and only if D is a Krull domain and $Cl(D) = 0$.

Theorem 3.1. *D is an AUF-domain if and only if D is an AK-domain and $Cl(D)$ is torsion.*

Proof. ($\Rightarrow$) Let I be a t-invertible t-ideal of D. Then $I = (a_1, \ldots, a_k)_t$ for some $a_1, \ldots, a_k \in D$. Since D is an AUF-domain, $(I^n)_t = (a_1^n, \ldots, a_k^n)_t$ [Anderson and Zafrullah 1991, Lemma 3.3] is principal for some integer $n = n(a_1, \ldots, a_k) \geq 1$. Thus $Cl(D)$ is torsion. Clearly, D is an AK-domain.

($\Leftarrow$) Let $(\{b_\alpha\})$ be a nonzero ideal of D. Since D is an AK-domain, there exists a positive integer $n = n(\{b_\alpha\})$ such that $(\{b_\alpha^n\})_t$ is t-invertible. Also, since $Cl(D)$ is torsion, $(\{b_\alpha^{nm}\})_t = ((\{b_\alpha^n\})^m)_t$ [Anderson and Zafrullah 1991, Lemma 3.3] is principal for some integer $m \geq 1$. Thus D is an AUF-domain. $\square$

Unlike the AK-domain case, an integrally closed AUF-domain need not be a UFD. For example, $\mathbb{Z}[\sqrt{-5}]$ is an integrally closed AUF-domain [Anderson and Zafrullah 1991, Theorem 4.17] which is not a UFD [Balcerzyk and Józefiak 1989, Example 3.4.1]. We give a characterization of integrally closed AUF-domains, which is also the analogue of the fact that D is an integrally closed API-domain if and only if D is a Dedekind domain with torsion class group [Anderson and Zafrullah 1991, Theorem 4.12].

Corollary 3.2. *D is an integrally closed AUF-domain if and only if D is a Krull domain and $Cl(D)$ is torsion.*

Proof. Recall that an integrally closed AK-domain is a Krull domain. Thus this result comes directly from Theorem 3.1. □

The next corollary explains the relationship between API-domains and AUF-domains.

Corollary 3.3. *D is an API-domain if and only if D is an AUF-domain and each maximal ideal is a t-ideal.*

Proof. ($\Rightarrow$) Since an API-domain is an AD-domain, by Proposition 2.1, each maximal ideal of D is a t-ideal. Clearly, D is an AUF-domain.

($\Leftarrow$) Recall that D is an API-domain if and only if D is an AD-domain with torsion t-class group [Anderson and Zafrullah 1991, Lemma 4.4]. Thus the result is an immediate consequence of Proposition 2.1 and Theorem 3.1. □

As in [Anderson et al. 1992], we say that D is an *almost weakly factorial domain* (AWFD) if for each nonzero nonunit $d \in D$, there exists a positive integer $n = n(d)$ such that d^n is a product of primary elements of D. It is known that D is an AWFD if and only if D is a weakly Krull domain with $Cl(D)$ torsion [Anderson et al. 1992, Theorem 3.4]. Thus by Corollary 2.3 and Theorem 3.1, we have

Corollary 3.4. *An AUF-domain with t-dimension one is an AWFD.*

Remark 3.5. Let $D = \mathbb{Q} + X\mathbb{R}[X]$. Then D is an AWFD [Anderson et al. 2006, Theorem 3.5], so $t\text{-dim}(D) = 1$. Note that D is not an AK-domain by Example 2.4. Hence by Theorem 3.1, D is not an AUF-domain. Thus the converse of Corollary 3.4 is also not true.

The next two corollaries are analogues of Theorem 2.11 and Corollary 2.8, respectively.

Corollary 3.6. *$D[\Gamma]$ is an AUF-domain if and only if $D[X]$ is an AUF-domain and* $\mathrm{char}(D) \neq 0$.

Proof. Recall that $\mathrm{Pic}(K[\Gamma])$ is torsion if and only if $\mathrm{char}(D) \neq 0$ [Chang et al. 2012, Lemma 2.7]. Since $Cl(D[\Gamma]) = Cl(D[X]) \oplus \mathrm{Pic}(K[\Gamma])$ [Anderson and

Chang 2004, Theorem 5], $Cl(D[\Gamma])$ is torsion if and only if $Cl(D[X])$ is torsion and $\mathrm{char}(D) \neq 0$. Thus this equivalence follows from Theorems 2.11 and 3.1. □

Corollary 3.7. *If D is a UFD, then $\mathrm{char}(D) \neq 0$ if and only if $D[\Gamma]$ is an AUF-domain.*

Proof. The assertion follows from Corollary 3.6, because D is a UFD if and only if $D[X]$ is a UFD [Zariski and Samuel 1975, Theorem 10, §17, Chapter 1]. □

The AUF-domain version of Theorem 2.5 also carries over.

Theorem 3.8. *Let $D \subseteq E$ be a bounded root extension with $E \subseteq \overline{D}$. Assume that E is t-linked over D. Then D is an AUF-domain if and only if E is an AUF-domain and D is t-linked under E.*

Proof. ($\Rightarrow$) Let $\{a_\alpha\}$ be a nonempty subset of $E \setminus \{0\}$. By the assumption, there exists a positive integer n such that $\{a_\alpha^n\} \subseteq D$. Since D is an AUF-domain, there exists a positive integer $m = m(\{a_\alpha^n\})$ such that $(\{a_\alpha^{nm}\})_t$ is principal, say $(\{a_\alpha^{nm}\})_t = (d)$ for some $d \in D$. Hence $\big((1/d)\,(\{a_\alpha^{nm}\})D\big)_t = D$. Since E is t-linked over D, $\big((1/d)\,(\{a_\alpha^{nm}\})E\big)_t = E$ [Anderson et al. 1993, Proposition 2.1]; so $(\{a_\alpha^{nm}\}E)_t = dE$. Thus E is an AUF-domain. Since an AUF-domain is an AK-domain, by Theorem 2.5, D is t-linked under E.

($\Leftarrow$) Assume that E is an AUF-domain and let $(\{b_\beta\})$ be a nonzero ideal of D. Then there exists a positive integer $m = m(\{b_\beta\})$ such that $(\{b_\beta^m\}E)_t = xE$ for some $x \in E$; so $\big(\{b_\beta^m/x\}E\big)_t = E$. Since $D \subseteq E$ is a bounded root extension, there is an integer $n \geq 1$ such that $b_\beta^{mn}/x^n = \big(b_\beta^m/x\big)^n \in D$ for all b_β^m/x. Clearly, $\big(\{b_\beta^{mn}/x^n\}E\big)_t = E$, and since D is t-linked under E, $\big(\{b_\beta^{mn}/x^n\}D\big)_t = D$. Therefore $(\{b_\beta^{mn}\})_t = x^n D$, and thus D is an AUF-domain. □

By combining Theorem 3.8 with Corollary 3.2, we have

Corollary 3.9. *If $D \subseteq \overline{D}$ is a bounded root extension, then D is an AUF-domain if and only if $\overline{D}$ is a Krull domain, $Cl(\overline{D})$ is torsion and D is t-linked under $\overline{D}$.*

Corollary 3.10. *If $D[X] \subseteq \overline{D}[X]$ is a bounded root extension, then D is an AUF-domain if and only if $D[X]$ is an AUF-domain.*

Proof. Note that $Cl(\overline{D}[X]) = Cl(\overline{D})$ [Gabelli 1987, Theorem 3.6] (or [El Baghdadi et al. 2002, Corollary 2.13]), $\overline{D}$ and $\overline{D}[X]$ are Krull domains simultaneously [Gilmer 1992, Corollary 43.11(3)], and D is t-linked under $\overline{D}$ if and only if $D[X]$ is t-linked under $\overline{D}[X]$ [Anderson et al. 2004, Proposition 3.3]. Thus the result is an immediate consequence of Corollary 3.9. □

Remark 3.11. Example 2.10 also shows that the assumption "$D[X] \subsetneq \overline{D}[X]$ is a bounded root extension" is essential in Corollary 3.10.

As an application of Corollary 3.10, we give an example of AUF-domains that is not an API-domain.

Corollary 3.12. *Let $F \subsetneq L$ be a field extension, $0 \neq p = \mathrm{char}(F)$, X, Y be indeterminates over L and let $D = F + YL[Y]$ or $F + YL[\![Y]\!]$. Assume that $L^{p^n} \subseteq F$ for some positive integer n. Then*

(1) *D is an API-domain.*

(2) *$D[X]$ is an AUF-domain that is not an API-domain.*

Proof. (1) This appears in [Anderson et al. 1994, Corollary 2.2 and Theorem 3.5(3)].

(2) We do the case where $D = F + YL[Y]$. Note that if $f \in D$, then $f^{p^n} \in F[Y]$. Also, we note that $\overline{D} = L[Y]$. Let $f = a + g \in \overline{D}[X]$, where $a \in L$ and $g \in (X, Y)L[Y][X]$. Then $a^{p^n} \in F$; so f^{p^n} belongs to $F + (X, Y)L[Y][X]$. Say $f^{p^n} = h_0 + h_1X + \cdots + h_mX^m + g_1Y + \cdots + g_kY^k$, where $h_0 \in F$, $h_i \in L$ $(1 \leq i \leq m)$ and $g_j \in L[X]$ $(1 \leq j \leq k)$. Since $L^{p^n} \subseteq F$ and $\mathrm{char}(F) = p$, we have

$$\begin{aligned} f^{p^{2n}} &= (f^{p^n})^{p^n} \\ &= h_0^{p^n} + h_1^{p^n} X^{p^n} + \cdots + h_m^{p^n} X^{p^n} + \sum_{j \geq 1} (g_j Y^j)^{p^n} \\ &\in F[X] + YL[X][Y] = D[X]. \end{aligned}$$

Hence $D[X] \subseteq \overline{D}[X]$ is a bounded root extension, and thus $D[X]$ is an AUF-domain by Corollary 3.10 and (1). For the sake of contradiction, assume that $D[X]$ is an API-domain. Then $\overline{D[X]} = \overline{D}[X]$ is a Prüfer domain [Anderson and Zafrullah 1991, Corollary 4.8]. Therefore $\overline{D}$, and hence D is a field [Gilmer 1992, Exercise 15, Section 22], a contradiction. Thus $D[X]$ is not an API-domain.

A similar proof holds for the case where $D = F + YL[\![Y]\!]$. (Note that if $D = F + YL[\![Y]\!]$, then $\overline{D} = L[\![Y]\!]$.) □

Example 3.13. (1) Let F be a field with $\mathrm{char}(F) = p > 0$, X be an indeterminate over F and n be a positive integer. Then $F(X^{p^n}) \subsetneq F(X)$ is a field extension such that $F(X)^{p^n} \subseteq F(X^{p^n})$.

(2) Let D be an integral domain as in Corollary 3.12. Note that $\mathrm{char}(D) \neq 0$ and $D[X]$ is an AUF-domain. Hence by Corollary 3.6, $D[\Gamma]$ is a nonintegrally closed AUF-domain. This gives another example of an AUF-domain that is not a UFD.

Recall that a nonzero nonunit $p \in D$ is a *prime block* if for all $x, y \in D$ with $(x, p)_v \subsetneq D$ and $(y, p)_v \subsetneq D$, there exist an integer $n \geq 1$ and an element $d \in D$ such that $(x^n, y^n) \subseteq dD$ with $(x^n/d, p)_v = D$ or $(y^n/d, p)_v = D$. Following [Zafrullah 1985, Definition 1.10 and Remark 1.11], we say that D is an *almost unique factorization domain* (AUFD) if for every nonzero nonunit $d \in D$, there is a positive integer $n = n(d)$ such that d^n is expressible as a product of finitely many mutually v-coprime prime blocks; equivalently, every nonzero nonunit of D

is expressible as a product of finitely many prime blocks. (This explains why we use the name "AUF-domain" instead of AUFD.) It is obvious that each nonzero nonunit of a valuation domain V is a prime block. Hence a valuation domain is an AUFD, but if V has Krull dimension at least 2 or V is a rank one nondiscrete valuation domain, then V is not an AUF-domain since an integrally closed AUF-domain is a Krull domain by Corollary 3.2. This shows that the converse of the next theorem does not hold.

Theorem 3.14. *An AUF-domain with t-dimension one is an AUFD.*

Proof. Let D be an AUF-domain of t-dimension one. If $P = (\{a_\alpha\}) \in X^1(D)$, then there exists an integer $m \geq 1$ such that $(\{a_\alpha^m\})_t = pD$ for some $p \in D$. We claim that p is a prime block. For $x, y \in D$ with $(x, p)_v \subsetneq D$ and $(y, p)_v \subsetneq D$, there exists an integer $n \geq 1$ such that $(x^n, y^n) \subseteq (x^n, y^n)_v = dD$ for some $d \in D$. So $(x^n/d, y^n/d)_v = D$. Suppose to the contrary that $(x^n/d, p)_v \subsetneq D$ and $(y^n/d, p)_v \subsetneq D$. Since t-dim$(D) = 1$ and $P = \sqrt{pD} \in X^1(D)$, we have that $(x^n/d, y^n/d)_v \subseteq P \subsetneq D$, a contradiction. Thus $(x^n/d, p)_v = D$ or $(y^n/d, p)_v = D$, which indicates that p is a prime block. Again, since t-dim$(D) = 1$, each nonzero prime ideal of D contains a prime block. Thus D is an AUFD [Zafrullah 1985, Theorem 2.2]. □

We say that D is an *almost π-domain* if for every nonzero ideal $(\{a_\alpha\})$ of D, there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})_t$ is invertible. (To prevent the reader's confusion, we should point out that our almost π-domain is different from Anderson's almost π-domain. He called D an *almost π-domain* if D is a Krull domain with torsion local t-class group [Anderson 1978, page 202]. Therefore, Anderson's almost π-domain is a special case of our almost π-domain. (See Theorem 3.15 or Corollary 3.16.)) It is clear that API-domain $\Rightarrow$ AUF-domain $\Rightarrow$ almost π-domain $\Rightarrow$ AK-domain.

The next theorem is the almost π-domain analogue of the result that D is a π-domain if and only if D is a Krull domain with trivial local t-class group [Anderson 1978, Theorem 1].

Theorem 3.15. *D is an almost π-domain if and only if D is an AK-domain and $G(D)$ is torsion.*

Proof. ($\Rightarrow$) Let I be a t-invertible t-ideal of D. Then $I = (a_1, \dots, a_m)_t$ for some $a_1, \dots, a_m \in D$. Since D is an almost π-domain, there exists an integer $n = n(a_1, \dots, a_m) \geq 1$ such that $(I^n)_t = (a_1^n, \dots, a_m^n)_t$ [Anderson and Zafrullah 1991, Lemma 3.3] is invertible. Thus $G(D)$ is torsion. Clearly, D is an AK-domain.

($\Leftarrow$) If $(\{a_\alpha\})$ is a nonzero ideal of an AK-domain D, then there exists a positive integer $n = n(\{a_\alpha\})$ such that $(\{a_\alpha^n\})_t$ is t-invertible. Since $G(D)$ is torsion, there is an integer $m = m((\{a_\alpha^n\})_t) \geq 1$ such that $(\{a_\alpha^{nm}\})_t = ((\{a_\alpha^n\})^m)_t$ [Anderson and Zafrullah 1991, Lemma 3.3] is invertible. Thus D is an almost π-domain. □

Our remaining corollaries are the almost π-domain analogues of Corollaries 3.2 and 3.7, respectively.

Corollary 3.16. *D is an integrally closed almost π-domain if and only if D is a Krull domain and $G(D)$ is torsion.*

Proof. Recall again that an integrally closed AK-domain is a Krull domain. Thus the result comes directly from Theorem 3.15. □

Example 3.17. Let X, Y be analytic indeterminates over $\mathbb{Z}[\sqrt{-5}]$ and consider $D = \mathbb{Z}[\sqrt{-5}][X^2, XY, Y^2]$. Then D is a Krull domain with $\mathrm{Pic}(D) = \mathbb{Z}/2\mathbb{Z} \subsetneq \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} = Cl(D)$ [Anderson and Ryckaert 1988, Example 4.7(1)]; so $G(D)$ is torsion. Hence by Corollary 3.16, D is an integrally closed almost π-domain. Note that D is not a π-domain. Thus, an integrally closed almost π-domain need not be a π-domain.

Corollary 3.18. *If D is a π-domain, then* $\mathrm{char}(D) \neq 0$ *if and only if $D[\Gamma]$ is an almost π-domain.*

Proof. Recall that a π-domain D is a Krull domain with $Cl(D) = \mathrm{Pic}(D)$. Therefore

$$\begin{aligned} Cl(D[\Gamma]) &= Cl(D[X]) \oplus \mathrm{Pic}(K[\Gamma]) \\ &= Cl(D) \oplus \mathrm{Pic}(K[\Gamma]) \\ &\supseteq \mathrm{Pic}(D[\Gamma]) \\ &\supseteq Cl(D) \oplus \{0\}, \end{aligned}$$

where the first equality comes from [Anderson and Chang 2004, Theorem 5] and the second equality follows from [Gabelli 1987, Theorem 3.6] (or [El Baghdadi et al. 2002, Corollary 2.13]). We also note that $\mathrm{char}(D) \neq 0$ if and only if $\mathrm{Pic}(K[\Gamma])$ is torsion [Chang et al. 2012, Lemma 2.7]. Hence if D is a π-domain with nonzero characteristic, then $G(D[\Gamma])$ is torsion. Thus the result is an immediate consequence of Corollary 2.8 and Theorem 3.15. □

It is worth remarking at this point that the assumption "bounded root extension" has a significant role in proving many results in our paper (for example, Theorems 2.5 and 3.8). In [Anderson and Zafrullah 1991], Anderson and Zafrullah also utilized this hypothesis to show some theorems about AB-domains, AP-domains, API-domains and AD-domains. Unfortunately, we are unable to prove Corollary 2.6 without this assumption, i.e., we do not know whether the t-dimension of an AK-domain is generally 1. (Note that the t-dimension of a Krull domain is always 1.) We are closing this article with a couple of questions.

Question 3.19. Let D be an AK-domain or AUF-domain.

(1) What is the t-dimension of D? Is it true that $t\text{-}\dim(D) = 1$?

(2) Is $\overline{D}$ necessarily a Krull domain?

Appendix

In this appendix, we give three diagrams of various integral domains related to (almost) factorization theory in order to help the readers better understand the correlation between the theory of almost factoriality and factorization theory. Each of the corresponding vertices in Figures 1 and 2 represent the correlation of integral domains in the relationship of the theory of almost factoriality and factorization theory. (For example, an AK-domain in the theory of almost factoriality is the corresponding notion to Krull domains in factorization theory.) In Figure 3, each vertex represents the integrally closed domain version of corresponding vertex in Figure 1. (For instance, an integrally closed API-domain is a Dedekind domain with torsion class group.) We also cite some well-known results about these domains.

For more on integral domains and almost factoriality see [Anderson et al. 1994; Anderson and Zafrullah 1991; Lewin 1997; Li 2012; Zafrullah 1985]. For more on integral domains in factorization theory, see [Anderson and Anderson 1980; Gilmer 1992; Kaplansky 1994; Mott and Zafrullah 1981].

Remark 1. Let D be one of domains in Figure 1.

(1) In general, D is not integrally closed.

(2) The arrows always hold, but none of the converse is true.

(3) Each implication of type $\leftarrow$ holds (for example, an AK-domain is an AD-domain) if and only if $\mathrm{Max}(D) = t\text{-}\mathrm{Max}(D)$.

(4) Each implication of type $\nearrow$ holds (for example, an AK-domain is an AUF-domain) if and only if $Cl(D)$ is torsion.

(5) Each implication of type $\dashrightarrow$ holds (for example, an AK-domain is an almost π-domain) if and only if $G(D)$ is torsion.

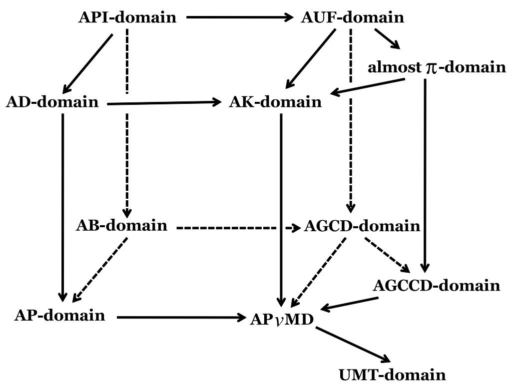

Figure 1. Nonintegrally closed domains in the theory of almost factoriality.

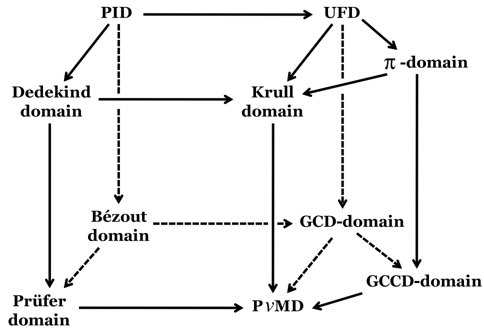

Figure 2. Integrally closed domains in factorization theory.

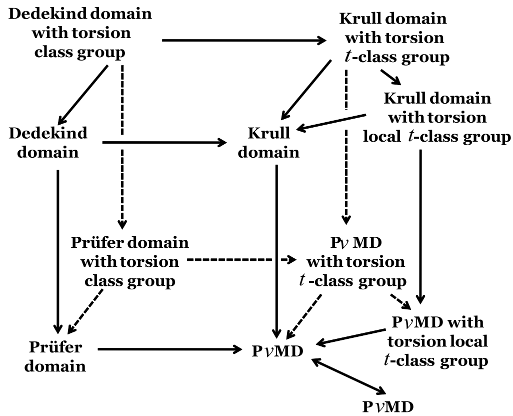

Figure 3. Integrally closed domain version of domains in the theory of almost factoriality.

Remark 2. Let D denote one of domains in Figure 2.

(1) D is integrally closed.

(2) The arrows always hold, but none of the converse holds.

(3) Each implication of type $\leftarrow$ holds (for instance, a Krull domain is a Dedekind domain) if and only if $\mathrm{Max}(D) = t\text{-}\mathrm{Max}(D)$.

(4) Each implication of type $\nearrow$ holds (for instance, a Krull domain is a UFD) if and only if $Cl(D) = 0$.

(5) Each implication of type $\nearrow$ holds (for instance, a Krull domain is a π-domain) if and only if $G(D) = 0$.

Acknowledgements

We would like to thank the referee for several valuable comments and suggestions, which resulted in an improved version of the paper.

References

[Anderson 1978] D. D. Anderson, "π-domains, overrings, and divisorial ideals", *Glasgow Math. J.* **19**:2 (1978), 199–203. MR 58 #5647 Zbl 0379.13011

[Anderson and Anderson 1980] D. D. Anderson and D. F. Anderson, "Generalized GCD domains", *Comment. Math. Univ. St. Paul.* **28**:2 (1980), 215–221. MR 81j:13017 Zbl 0434.13001

[Anderson and Chang 2004] D. F. Anderson and G. W. Chang, "The class group of $D[\Gamma]$ for D an integral domain and Γ a numerical semigroup", *Comm. Algebra* **32**:2 (2004), 787–792. MR 2005g:13021 Zbl 1092.13017

[Anderson and El Abidine 2001] D. F. Anderson and D. N. El Abidine, "The $A + XB[X]$ and $A + XB[\![X]\!]$ constructions from GCD-domains", *J. Pure Appl. Algebra* **159**:1 (2001), 15–24. MR 2002f:13037 Zbl 1004.13004

[Anderson and Ryckaert 1988] D. F. Anderson and A. Ryckaert, "The class group of $D + M$", *J. Pure Appl. Algebra* **52**:3 (1988), 199–212. MR 89h:13008 Zbl 0668.13013

[Anderson and Zafrullah 1991] D. D. Anderson and M. Zafrullah, "Almost Bézout domains", *J. Algebra* **142**:2 (1991), 285–309. MR 92g:13026 Zbl 0749.13013

[Anderson et al. 1992] D. D. Anderson, J. L. Mott, and M. Zafrullah, "Finite character representations for integral domains", *Boll. Un. Mat. Ital. B* (7) **6**:3 (1992), 613–630. MR 93k:13001 Zbl 0773.13004

[Anderson et al. 1993] D. D. Anderson, E. G. Houston, and M. Zafrullah, "t-linked extensions, the t-class group, and Nagata's theorem", *J. Pure Appl. Algebra* **86**:2 (1993), 109–124. MR 94e:13036 Zbl 0777.13002

[Anderson et al. 1994] D. D. Anderson, K. R. Knopp, and R. L. Lewin, "Almost Bézout domains, II", *J. Algebra* **167**:3 (1994), 547–556. MR 95f:13026 Zbl 0821.13006

[Anderson et al. 2004] D. D. Anderson, T. Dumitrescu, and M. Zafrullah, "Almost splitting sets and AGCD domains", *Comm. Algebra* **32**:1 (2004), 147–158. MR 2005d:13002 Zbl 1063.13002

[Anderson et al. 2006] D. F. Anderson, G. W. Chang, and J. Park, "Weakly Krull and related domains of the form $D+M$, $A+XB[X]$ and $A+X^2B[X]$", *Rocky Mountain J. Math.* **36**:1 (2006), 1–22. MR 2007h:13027 Zbl 1133.13022

[Balcerzyk and Józefiak 1989] S. Balcerzyk and T. Józefiak, *Commutative Noetherian and Krull rings*, Ellis Horwood, Chichester, 1989. MR 92f:13001 Zbl 0685.13001

[Chang et al. 2012] G. W. Chang, H. Kim, and J. W. Lim, "Numerical semigroup rings and almost Prüfer v-multiplication domains", *Comm. Algebra* **40**:7 (2012), 2385–2399.

[Dobbs et al. 1989] D. E. Dobbs, E. G. Houston, T. G. Lucas, and M. Zafrullah, "t-linked overrings and Prüfer v-multiplication domains", *Comm. Algebra* **17**:11 (1989), 2835–2852. MR 90j:13016 Zbl 0691.13015

[El Baghdadi et al. 2002] S. El Baghdadi, L. Izelgue, and S. Kabbaj, "On the class group of a graded domain", *J. Pure Appl. Algebra* **171**:2-3 (2002), 171–184. MR 2003d:13012 Zbl 1058.13006

[Fontana et al. 1998] M. Fontana, S. Gabelli, and E. Houston, "UMT-domains and domains with Prüfer integral closure", *Comm. Algebra* **26**:4 (1998), 1017–1039. MR 99d:13024 Zbl 0930.13014

[Gabelli 1987] S. Gabelli, "On divisorial ideals in polynomial rings over Mori domains", *Comm. Algebra* **15**:11 (1987), 2349–2370. MR 88m:13006 Zbl 0642.13017

[Gilmer 1992] R. Gilmer, *Multiplicative ideal theory*, Queen's Papers in Pure and Applied Mathematics **90**, Queen's University, Kingston, ON, 1992. MR 93j:13001 Zbl 0804.13001

[Houston and Zafrullah 1989] E. Houston and M. Zafrullah, "On t-invertibility, II", *Comm. Algebra* **17**:8 (1989), 1955–1969. MR 90i:13019 Zbl 0717.13002

[Kang 1989a] B. G. Kang, "On the converse of a well-known fact about Krull domains", *J. Algebra* **124**:2 (1989), 284–299. MR 90g:13033 Zbl 0694.13011

[Kang 1989b] B. G. Kang, "Prüfer v-multiplication domains and the ring $R[X]_{N_v}$", *J. Algebra* **123**:1 (1989), 151–170. MR 90e:13017 Zbl 0668.13002

[Kaplansky 1994] I. Kaplansky, *Commutative rings*, Polygonal, Washington, NJ, 1994.

[Lewin 1997] R. L. Lewin, "Almost generalized GCD-domains", pp. 371–382 in *Factorization in integral domains* (Iowa City, IA, 1996), edited by D. D. Anderson, Lecture Notes in Pure and Appl. Math. **189**, Marcel Dekker, New York, 1997. MR 98i:13043 Zbl 0909.13009

[Li 2012] Q. Li, "On almost Prüfer v-multiplication domains", *Algebra Colloq.* **19**:3 (2012), 493–500.

[Mott and Zafrullah 1981] J. L. Mott and M. Zafrullah, "On Prüfer v-multiplication domains", *Manuscripta Math.* **35**:1-2 (1981), 1–26. MR 83d:13026 Zbl 0477.13007

[Pirtle 1968] E. M. Pirtle, Jr., "Integral domains which are almost-Krull", *J. Sci. Hiroshima Univ. Ser. A-I Math.* **32** (1968), 441–447. MR 39 #5538 Zbl 0181.04903

[Sahandi 2010] P. Sahandi, "Minimal prime ideals and semistar operations", *Rocky Mountain J. Math.* **40**:3 (2010), 1039–1049. MR 2011h:13003 Zbl 1200.13008

[Storch 1967] U. Storch, *Fastfaktorielle Ringe*, Schr. Math. Inst. Univ. Münster **36**, Universität Münster, Münster, 1967. MR 35 #5437 Zbl 0168.28901

[Zafrullah 1985] M. Zafrullah, "A general theory of almost factoriality", *Manuscripta Math.* **51**:1-3 (1985), 29–62. MR 86m:13023 Zbl 0587.13010

[Zariski and Samuel 1975] O. Zariski and P. Samuel, *Commutative algebra*, vol. 1, Graduate Texts in Mathematics **28**, Springer, New York, 1975. MR 52 #5641 Zbl 0313.13001

Received January 14, 2012. Revised March 15, 2012.

GYU WHAN CHANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF INCHEON
INCHEON 406-772
SOUTH KOREA
whan@incheon.ac.kr

HWANKOO KIM
DEPARTMENT OF INFORMATION SECURITY
HOSEO UNIVERSITY
ASAN 336-795
SOUTH KOREA
hkkim@hoseo.edu

JUNG WOOK LIM
DEPARTMENT OF MATHEMATICS
SOGANG UNIVERSITY
SEOUL 121-742
SOUTH KOREA
lovemath@postech.ac.kr
jwlim@sogang.ac.kr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.149

SINGULARITIES OF FREE GROUP CHARACTER VARIETIES

CARLOS FLORENTINO AND SEAN LAWTON

Let $\mathfrak{X}_r$ be the moduli space of SL_n, SU_n, GL_n, or U_n-valued representations of a rank r free group. We classify the algebraic singular stratification of $\mathfrak{X}_r$. This comes down to showing that the singular locus corresponds exactly to reducible representations if there exist singularities at all. Then by relating algebraic singularities to topological singularities, we show the moduli spaces $\mathfrak{X}_r$ generally are not topological manifolds, except for a few examples we explicitly describe.

1. Introduction

During the last few decades, character varieties have played important roles in knot theory, hyperbolic geometry, Higgs and vector bundle theory, and quantum field theory. However, many of their fundamental properties and structure are not completely understood.

In this article, we first classify the (algebraic) singular locus of SL_n and GL_n-character varieties of free groups by relating the existence of a singularity with the reducibility of the corresponding representation. We then classify all such character varieties that arise as manifolds by explicitly describing the topological neighborhoods of generic singularities. The results we obtain do not necessarily extend to general G-character varieties of finitely generated groups Γ, if G is not one of SL_n, SU_n, GL_n, or U_n and Γ is not free; explicit counterexamples can be obtained via methods different from those considered in this paper (see Section 3I). Our first main theorem generalizes results in [Heusener and Porti 2004], and our second main theorem generalizes results in [Bratholdt and Cooper 2001]. They may be described more precisely as follows.

Let $\mathtt{F}_r$ be a rank r free group and let G be a reductive complex algebraic group with K a maximal compact subgroup (see Section 2). Let $\mathfrak{R}_r(G) = \mathrm{Hom}(\mathtt{F}_r, G)$ and $\mathfrak{R}_r(K) = \mathrm{Hom}(\mathtt{F}_r, K)$ be varieties of representations, and let G, respectively

This work was partially supported by the Center for Mathematical Analysis, Geometry and Dynamical Systems at I.S.T., and by the Fundação para a Ciência e a Tecnologia through the programs Praxis XXI, POCI/ MAT/ 58549/ 2004 and FEDER..

MSC2010: primary 14B05, 32S05; secondary 14L30, 20CXX, 14P10.

Keywords: character variety, singularities, local topology.

K, act by conjugation on these representation spaces.

Consider the space $\mathfrak{X}_r(K) := \mathfrak{R}_r(K)/K$ which is the conjugation orbit space of $\mathfrak{R}_r(K)$ where $\rho \sim \psi$ if and only if there exists $k \in K$ so $\rho = k\psi k^{-1}$. Let $\mathbb{C}[\mathfrak{R}_r(G)]$ be the affine coordinate ring of $\mathfrak{R}_r(G)$ and let $\mathbb{C}[\mathfrak{R}_r(G)]^G$ be the subring of G-conjugation invariants. Then define

$$\mathfrak{X}_r(G) := \mathrm{Spec}_{\max}\left(\mathbb{C}[\mathfrak{R}_r(G)]^G\right),$$

which parametrizes unions of conjugation orbits where two orbits are in the same union if and only if their closures have a nonempty intersection.

The space $\mathfrak{X}_r(G)$, called the G-character variety of $\mathtt{F}_r$, is a complex affine variety and so has a well-defined (algebraic) singular locus (a proper subvariety) which we denote by $\mathfrak{X}_r(G)^{\mathrm{sing}}$. Similarly, $\mathfrak{X}_r(K)$ is a semialgebraic set and so has a real algebraic coordinate ring which likewise determines an algebraic singular locus $\mathfrak{X}_r(K)^{\mathrm{sing}}$. For simplicity, despite the fact it is generally not an algebraic set, we will also refer to $\mathfrak{X}_r(K)$ as a character variety.

We will be mainly concerned with the cases when G is the general linear group GL_n or the special linear group SL_n (over $\mathbb{C}$), for which K is the unitary group U_n or the special unitary group SU_n, respectively. In these cases a representation ρ is called *irreducible* if with respect to the standard action of G, respectively K, on $\mathbb{C}^n$ the induced action of $\rho(\mathtt{F}_r)$ does not have any nontrivial proper invariant subspaces. Otherwise ρ is called *reducible*. This allows one to define the sets $\mathfrak{X}_r(G)^{\mathrm{red}}$ and $\mathfrak{X}_r(K)^{\mathrm{red}}$ which correspond to the spaces of equivalence classes in $\mathfrak{X}_r(G)$, respectively $\mathfrak{X}_r(K)$, that have a representative which is reducible.

In Section 2, we show that the (algebraic) singular locus of $\mathfrak{X}_r(\mathsf{SL}_n)$ and $\mathfrak{X}_r(\mathsf{GL}_n)$ respectively determines the (algebraic) singular locus of $\mathfrak{X}_r(\mathsf{SU}_n)$ and $\mathfrak{X}_r(\mathsf{U}_n)$. We then show $\mathfrak{X}_r(\mathsf{SL}_n) \subset \mathfrak{X}_r(\mathsf{GL}_n)$ has its singular locus determined by the singular locus of $\mathfrak{X}_r(\mathsf{GL}_n)$. This reduces the classification of the singular loci of these four families of moduli spaces to $\mathfrak{X}_r(\mathsf{GL}_n)$ alone. We end Section 2 with examples of $\mathfrak{X}_r(G)$ that are homeomorphic to manifolds with boundary; we conjectured in [Florentino and Lawton 2009] that these were the *only* examples.

It is straightforward to establish that $\mathfrak{X}_1(\mathsf{SL}_n) \cong \mathbb{C}^{n-1}$ and $\mathfrak{X}_2(\mathsf{SL}_2) \cong \mathbb{C}^3$ are affine spaces and so smooth, and $\mathfrak{X}_1(\mathsf{SL}_n)^{\mathrm{red}} = \mathfrak{X}_1(\mathsf{SL}_n)$. In [Heusener and Porti 2004], it is shown that $\mathfrak{X}_r(\mathsf{SL}_2)^{\mathrm{sing}} = \mathfrak{X}_r(\mathsf{SL}_2)^{\mathrm{red}}$ for $r \geq 3$. More generally, one can establish that all irreducible representations in SL_n-character varieties of free groups are in fact smooth; that is $\mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{sing}} \subset \mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{red}}$. In [Lawton 2007] it is shown that the singular locus of $\mathfrak{X}_2(\mathsf{SL}_3)$ corresponds exactly to the set of equivalence classes of reducible representations; that is, $\mathfrak{X}_2(\mathsf{SL}_3)^{\mathrm{red}} = \mathfrak{X}_2(\mathsf{SL}_3)^{\mathrm{sing}}$. These examples generalize to our first main result:

Theorem 1.1. *Let $r, n \geq 2$. Let G be SL_n or GL_n and K be SU_n or U_n. Then $\mathfrak{X}_r(G)^{\mathrm{red}} = \mathfrak{X}_r(G)^{\mathrm{sing}}$ and $\mathfrak{X}_r(K)^{\mathrm{red}} = \mathfrak{X}_r(K)^{\mathrm{sing}}$ if and only if $(r, n) \neq (2, 2)$.*

In fact we are able to use an induction argument to completely classify the singular stratification of these semialgebraic spaces. The proof and development of this result constitutes Section 3, including a brief review of a weak version of the celebrated Luna slice theorem.

Theorem 1.1 is sharper than it might appear at first. Replacing $\mathtt{F}_r$ by a general finitely presented group Γ one can find examples where irreducibles are singular and examples where reducibles are smooth. On the other hand, changing G to a general reductive complex algebraic group, we find there are examples where irreducibles are singular. In Section 3I, we discuss this in further detail.

A locally Euclidean Hausdorff space M with a countable basis is called a *topological manifold*. More generally, if the neighborhoods are permitted to be Euclidean half-spaces then M is said to be a *topological manifold with boundary*. In [Florentino and Lawton 2009] we determined the homeomorphism type of $\mathfrak{X}_r(\mathsf{SU}_n)$ in the cases $(r, n) = (r, 1)$, $(1, n)$, $(2, 2)$, $(2, 3)$, and $(3, 2)$ where we showed all were topological manifolds with boundary; this is reviewed in Section 2B. In [Bratholdt and Cooper 2001] it is established that the $\mathfrak{X}_r(\mathsf{SU}_2)$ are not topological manifolds when $r \geq 4$.

Motivated by this we conjectured in [Florentino and Lawton 2009] that the examples computed there are the only cases where a topological manifold with boundary arise. Our second main theorem in this paper establishes that conjecture:

Theorem 1.2. *Let $r, n \geq 2$. Let G be SL_n or GL_n and K be SU_n or U_n. $\mathfrak{X}_r(G)$ is a topological manifold with boundary if and only if $(r, n) = (2, 2)$. $\mathfrak{X}_r(K)$ is a topological manifold with boundary if and only if $(r, n) = (2, 2)$, $(2, 3)$, or $(3, 2)$.*

Theorem 1.1 and the observation that the reducible locus is nonempty for $n \geq 2$, does not immediately imply Theorem 1.2 since algebraic singularities may or may not be an obstruction to the existence of a Euclidean neighborhood (topological singularities). For example, both the varieties given by $xy = 0$ and $y^2 = x^3$ in $\mathbb{C}^2$ (or $\mathbb{R}^2$) are (algebraically) singular at the point $(0, 0)$ but only the latter has a Euclidean neighborhood at the origin. So, only the former is topologically singular. The variety $xy = 0$ is reducible; an example of an irreducible variety that has an algebraic singularity that is also a topological singularity is the affine cone over $\mathbb{C}\mathsf{P}^1 \times \mathbb{C}\mathsf{P}^1$ discussed in Section 3E.

The proof of Theorem 1.2 constitutes Section 4. To prove our main theorems we use slice theorems and explicitly describe the homeomorphism type of neighborhoods (showing them to be non-Euclidean) for a family of examples. It is interesting to note that since $\mathfrak{X}_r(\mathsf{SL}_n)$ deformation retracts to $\mathfrak{X}_r(\mathsf{SU}_n)$, by [Florentino and Lawton 2009], it must be the case that for $(r, n) = (2, 3)$ and $(3, 2)$ the non-Euclidean neighborhoods deformation retract to Euclidean neighborhoods. Curiously, these are the only cases $(n \geq 2)$ where $\mathfrak{X}_r(\mathsf{SU}_n)$ is a topological manifold, and both are homeomorphic to spheres (see [Florentino and Lawton 2009] or Section 2B).

2. Character varieties

Let G be a complex affine reductive algebraic group and let K be a maximal compact subgroup. Then, $G = K_{\mathbb{C}}$ is the complexification of K (the set complex zeros of K as a real algebraic set). For instance, $K_{\mathbb{C}} = \mathsf{SL}_n$ is the complexification of $K = \mathsf{SU}_n$, and $K_{\mathbb{C}} = \mathsf{GL}_n$ is the complexification of $K = \mathsf{U}_n$.

Let Γ be a finitely generated group and let $\mathfrak{R}_\Gamma(G) = \mathrm{Hom}(\Gamma, G)$ be the G-valued representations of Γ. We call $\mathfrak{R}_\Gamma(G)$ the *G-representation variety of* Γ, although it is generally only an affine algebraic set.

In the category of affine varieties, $\mathfrak{R}_\Gamma(G)$ has a quotient by the conjugation action of G, a regular action, given by $\rho \mapsto g\rho g^{-1}$. This quotient is realized as $\mathfrak{X}_\Gamma(G) = \mathrm{Spec}_{\max}(\mathbb{C}[\mathfrak{R}_\Gamma(G)]^G)$, where $\mathbb{C}[\mathfrak{R}_\Gamma(G)]^G$ is the subring of invariant polynomials in the affine coordinate ring $\mathbb{C}[\mathfrak{R}_\Gamma(G)]$. We call $\mathfrak{X}_\Gamma(G)$ the *G-character variety of* Γ. Concretely, it parametrizes unions of conjugation orbits where two orbits are in the same union if and only if their closures intersect nontrivially. Within each union of orbits, denoted $[\rho]$ and called an extended orbit equivalence class, there is a unique closed orbit (having minimal dimension). Any representative from this closed orbit is called a *polystable point*. For SL_n and GL_n the polystable points will have the property that with respect to the action of $\rho(\Gamma)$ on $\mathbb{C}^n$, they are completely reducible; that is, each decomposes into a finite direct sum of irreducible subactions (on nonzero subspaces).

Let $\mathtt{F}_r = \langle x_1, \dots, x_r\rangle$ be a rank r free group. The *G-representation variety of* $\mathtt{F}_r$, and the *G-character variety of* $\mathtt{F}_r$ will simply be denoted by $\mathfrak{R}_r(G)$ and $\mathfrak{X}_r(G)$, respectively. The evaluation mapping $\mathfrak{R}_r(G) \to G^r$ defined by sending $\rho \mapsto (\rho(x_1), \dots, \rho(x_r))$ is a bijection and since G is a smooth affine variety, $\mathfrak{R}_r(G)$ naturally inherits the structure of a smooth affine variety as well. Note that we are not assuming that an algebraic variety is irreducible. Whenever G is an irreducible algebraic set however, $\mathfrak{R}_r(G)$ is irreducible, and consequently $\mathfrak{X}_r(G)$ is irreducible as well.

Since an algebraic reductive group over $\mathbb{C}$ is always linear, we can assume that G is a subgroup of GL_N, for some N, and hence $\mathfrak{R}_r(G) \subset \mathbb{C}^{rN^2}$. So, $\mathfrak{R}_r(G)$ inherits the *ball topology*. Given a set of generators $f_1, \dots, f_k$ of the ring of invariants $\mathbb{C}[\mathfrak{R}_r(G)]^G$, $\mathfrak{X}_r(G)$ also inherits the ball topology from the embedding of $\mathfrak{X}_r(G)$ into $\mathbb{C}^k$ given by $[\rho] \mapsto (f_1(\rho), \dots, f_k(\rho))$. In this topology $\mathfrak{X}_r(G)$ is Hausdorff and has a countable basis. Although the ball topology is dependent on an embedding a priori, an affine embedding corresponds exactly to a set of generators for the associated ring, but all choices result in the same homeomorphism type, so the ball topology is intrinsic. Also, in the ball topology, at each point in $\mathfrak{X}_r(G)$ there is a neighborhood homeomorphic to a real cone over a space with Euler characteristic 0 [Sullivan 1971].

Given a compact Lie group K, for brevity we also call the orbit space $\mathfrak{X}_r(K) = \mathfrak{R}_r(K)/K$ a *K-character variety* of $\mathtt{F}_r$. Note however that $\mathfrak{X}_r(K)$ is generally only a semialgebraic set, and does not equal, in general, the set of real points of a complex variety. In this case, the topology, also Hausdorff with a countable basis, is the quotient topology. $\mathfrak{X}_r(K)$ is compact since K is compact. Likewise, it is path-connected whenever K is path-connected.

Definition 2.1. Let $\rho : \Gamma \to G$ be a representation into a reductive complex algebraic group. If the image of ρ does not lie in a parabolic subgroup of G, then ρ is called *irreducible*. If, for every parabolic P containing $\rho(\Gamma)$ there is a Levi factor $L \subset P$ such that $\rho(\Gamma) \subset L$, then ρ is called *completely reducible*.

For SL_n and GL_n the irreducible representations are exactly those that, with respect to their actions on $\mathbb{C}^n$, do not admit any proper (nontrivial) invariant subspaces. Any representation that is not irreducible is called *reducible*. Denote the set of reducible representations by $\mathfrak{R}_\Gamma(G)^{\mathrm{red}}$. A point is called *stable* if the stabilizer is finite and if the orbit is closed.

The following theorem can be found in [Sikora 2012], building on earlier work in [Johnson and Millson 1987, pp. 54–57]. Let $\mathsf{P}G = G/Z(G)$ where $Z(G)$ is the center. Note that the action of $\mathsf{P}G$ and G define the same GIT quotients and the same orbit spaces and thus, since the $\mathsf{P}G$ action is effective, we will sometimes consider this action.

Theorem 2.2 [Johnson and Millson 1987; Sikora 2012]. *Let G be reductive. The irreducibles are exactly the stable points under the action of $\mathsf{P}G$ on $\mathfrak{R}_\Gamma(G)$. The completely reducibles are the polystable points.*

Definition 2.3. The reducibles $\mathfrak{X}_\Gamma(G)^{\mathrm{red}}$ are the image of the projection

$$\mathfrak{R}_\Gamma(G)^{\mathrm{red}} \subset \mathfrak{R}_\Gamma(G) \longrightarrow \mathfrak{X}_\Gamma(G).$$

Since $\mathfrak{R}_r(G) \cong G^r$ all points are smooth, and since $\mathfrak{X}_r(G)$ is an affine quotient of a reductive group, there exists $\rho^{\mathrm{ss}} \in [\rho]$ which has a closed orbit and corresponds to a completely reducible representation. Thus, for G either SL_n or GL_n we can assume it is in block diagonal form. In other words, $\rho^{\mathrm{ss}} \leftrightarrow (X_1, \dots, X_r)$ where X_i all have the same block diagonal form (if they are irreducible then there would be only one block). These representations induce a semisimple module structure on $\mathbb{C}^n$. We denote the set of semisimple representations by $\mathfrak{R}_r(G)^{\mathrm{ss}}$. We note that $\mathfrak{R}_r(G)^{\mathrm{ss}}/G \cong \mathfrak{X}_r(G)$ since all extended orbits have a semisimple representative, and that the semisimple representations are also the completely reducible representations which are also the polystable representations. Likewise, we denote the irreducible representations (those giving simple actions on $\mathbb{C}^n$) by $\mathfrak{R}_r(G)^{\mathrm{s}}$ and their quotient by $\mathfrak{X}_r(G)^{\mathrm{s}}$.

2A. ***The determinant fibration.*** In order to compare SL_n-character varieties to GL_n-character varieties, the following setup will be useful. The usual exact sequence of groups given by the determinant of an invertible matrix

$$\mathsf{SL}_n \to \mathsf{GL}_n \overset{\det}{\to} \mathbb{C}^* \tag{1}$$

induces (by fixing generators of $\mathtt{F}_r$, as before) what we will call the *determinant map*:

$$\det : \mathfrak{X}_r(\mathsf{GL}_n) \to \mathrm{Hom}(\mathtt{F}_r, \mathbb{C}^*) \cong \left(\mathbb{C}^*\right)^r, \quad [\rho] \mapsto \det(\rho),$$

where $\det(\rho) = (\det(X_1), \dots, \det(X_r))$, for $\rho = (X_1, \dots, X_r) \in \mathfrak{R}_r(\mathsf{GL}_n)$. Note that the map is clearly well-defined on conjugation classes. Considering the algebraic torus $(\mathbb{C}^*)^r = \mathrm{Hom}(\mathtt{F}_r, \mathbb{C}^*) = \mathfrak{X}_r(\mathbb{C}^*)$ as an algebraic group (with identity $\mathbf{1} = (1, \dots, 1)$ and componentwise multiplication) it is immediate that the SL_n-character variety is the "kernel" of the determinant map, $\mathfrak{X}_r(\mathsf{SL}_n) = \det^{-1}(\mathbf{1})$. Therefore, the sequence (1) induces another exact sequence

$$\mathfrak{X}_r(\mathsf{SL}_n) \to \mathfrak{X}_r(\mathsf{GL}_n) \overset{\det}{\to} \left(\mathbb{C}^*\right)^r. \tag{2}$$

In this way, SL_n-character varieties appear naturally as subvarieties of GL_n-character varieties.

Note also that $\mathfrak{X}_r(\mathsf{GL}_n)$ can be viewed as a $\mathfrak{X}_r(\mathbb{C}^*)$-space, as it admits a well-defined action of this torus. That is, we can naturally define $\rho \cdot \lambda \in \mathfrak{X}_r(\mathsf{GL}_n)$, given $\rho \in \mathfrak{X}_r(\mathsf{GL}_n)$ and $\lambda \in \mathfrak{X}_r(\mathbb{C}^*)$. Given that $\mathsf{PSL}_n = \mathsf{GL}_n /\!\!/ \mathbb{C}^*$, it is easy to see that the corresponding quotient is the PSL_n-character variety:

$$\mathfrak{X}_r(\mathsf{PSL}_n) = \mathfrak{X}_r(\mathsf{GL}_n) /\!\!/ \mathfrak{X}_r(\mathbb{C}^*).$$

Also, GL_n^r is a quasiaffine subvariety of $\mathfrak{gl}(n, \mathbb{C})^r$. In fact, it is the principal open set defined by the product of the determinants of generic matrices. Since the determinant is an invariant function and taking invariants commutes with localizing at those invariants, we have

$$\mathbb{C}[\mathsf{GL}_n^r]^{\mathsf{GL}_n} \approx \mathbb{C}[\mathfrak{gl}(n, \mathbb{C})^r /\!\!/ \mathsf{GL}_n] \left[\frac{1}{\det(X_1) \cdots \det(X_r)} \right],$$

where the expression on the right is the localization at the product of determinants.

We now prove how the fixed determinant character varieties, complex and compact, relate to the nonfixed determinant character varieties. Identify the cyclic group of order n, $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$, with $Z(\mathsf{SL}_n) \cong Z(\mathsf{SU}_n)$.

Theorem 2.4. *The following are isomorphisms of semialgebraic sets*:

(i) $\mathfrak{X}_r(\mathsf{GL}_n) \cong \mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathfrak{X}_r(\mathbb{Z}_n)} \mathfrak{X}_r(\mathsf{GL}_1)$.

(ii) $\mathfrak{X}_r(\mathsf{U}_n) \cong \mathfrak{X}_r(\mathsf{SU}_n) \times_{\mathfrak{X}_r(\mathbb{Z}_n)} \mathfrak{X}_r(\mathsf{U}_1)$.

Proof. We first note that $\mathfrak{X}_r(\mathsf{U}_1) \cong (S^1)^r$ and $\mathfrak{X}_r(\mathsf{GL}_1) \cong (\mathbb{C}^*)^r$, and thus $\mathfrak{X}_r(\mathbb{Z}_n) \cong \mathbb{Z}_n^r$, as the groups involved are abelian.

The determinant map (1) defines a principal SL_n-bundle $\mathsf{SL}_n \hookrightarrow \mathsf{GL}_n \to \mathbb{C}^*$, which also expresses $\mathsf{GL}_n \cong \mathsf{SL}_n \rtimes \mathbb{C}^*$ as a semidirect product since there exists a homomorphic section (SL_n is a normal subgroup).

Let $\mathbb{Z}_n$ correspond to n-th roots of unity $\omega_k = e^{\frac{2\pi i k}{n}}$. As algebraic sets one can show directly, by the mapping $(A, \lambda) \mapsto \lambda A$, that $\mathsf{GL}_n \cong (\mathsf{SL}_n \times \mathbb{C}^*) /\!\!/ \mathbb{Z}_n := \mathsf{SL}_n \times_{\mathbb{Z}_n} \mathbb{C}^*$ where $\mathbb{Z}_n$ acts by $\omega_k \cdot (g, \lambda) = (g\omega_k, \omega_k^{-1}\lambda)$ and $\mathbb{C}^*$ is the center of GL_n. This implies that, as algebraic sets,

$$
\begin{aligned}
(3) \qquad \mathfrak{X}_r(\mathsf{GL}_n) &\cong ((\mathsf{SL}_n \times \mathbb{C}^*) /\!\!/ \mathbb{Z}_n)^r /\!\!/ \mathsf{SL}_n \\
&\cong \big((\mathsf{SL}_n^r \times (\mathbb{C}^*)^r) /\!\!/ \mathbb{Z}_n^r\big) /\!\!/ \mathsf{SL}_n \\
&\cong \big((\mathsf{SL}_n^r \times (\mathbb{C}^*)^r) /\!\!/ \mathsf{SL}_n\big) /\!\!/ \mathbb{Z}_n^r \cong \mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathbb{Z}_n^r} (\mathbb{C}^*)^r,
\end{aligned}
$$

since the action of $\mathbb{Z}_n^r$ commutes with the action of SL_n which is trivial on $(\mathbb{C}^*)^r$.

In the same way we obtain the other "twisted product" isomorphism $\mathfrak{X}_r(\mathsf{U}_n) \cong \mathfrak{X}_r(\mathsf{SU}_n) \times_{\mathbb{Z}_n^r} (S^1)^r$. □

This result provides an explicit way to write $\mathfrak{X}_r(\mathsf{GL}_n)$ as a $\mathfrak{X}_r(\mathsf{SL}_n)$-bundle over the algebraic r-torus $(\mathbb{C}^*)^r$ and $\mathfrak{X}_r(\mathsf{U}_n)$ as a $\mathfrak{X}_r(\mathsf{SU}_n)$-bundle over the geometric r-torus $(S^1)^r$.

There are a number of consequences to Theorem 2.4.

Corollary 2.5. *$\mathfrak{X}_r(\mathsf{U}_n)$, respectively $\mathfrak{X}_r(\mathsf{GL}_n)$, is a manifold whenever $\mathfrak{X}_r(\mathsf{SU}_n)$, respectively $\mathfrak{X}_r(\mathsf{SL}_n)$, is a manifold.*

Proof. The action of $\mathbb{Z}_n^r$ is free and proper. □

Corollary 2.6. *$\mathfrak{X}_r(\mathsf{GL}_n)$ and $\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r$ are étale equivalent.*

Proof. First note that $\mathsf{SL}_n^r \times (\mathbb{C}^*)^r$ is smooth and hence a normal variety. This implies (see [Drézet 2004]) that $(\mathsf{SL}_n^r \times (\mathbb{C}^*)^r) /\!\!/ \mathsf{SL}_n = \mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r$ is also normal. However, the GIT projection

$$\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r \to \mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathbb{Z}_n^r} (\mathbb{C}^*)^r$$

is then étale because $\mathbb{Z}_n^r$ is finite and acts freely [ibid.]. Then by Theorem 2.4 $\mathfrak{X}_r(\mathsf{GL}_n) \cong \mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathbb{Z}_n^r} (\mathbb{C}^*)^r$ which establishes the result. □

Corollary 2.7. *Let $[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)$ and let $[\psi] \in \mathfrak{X}_r(\mathsf{SU}_n)$. Then:*

(i) $[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{sing}}$ *if and only if* $[\rho] \in \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{sing}}$.

(ii) $[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{sm}}$ *if and only if* $[\rho] \in \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{sm}}$.

(iii) $[\psi] \in \mathfrak{X}_r(\mathsf{SU}_n)^{\mathrm{sing}}$ *if and only if* $[\psi] \in \mathfrak{X}_r(\mathsf{U}_n)^{\mathrm{sing}}$.

(iv) $[\psi] \in \mathfrak{X}_r(\mathsf{SU}_n)^{\mathrm{sm}}$ *if and only if* $[\psi] \in \mathfrak{X}_r(\mathsf{U}_n)^{\mathrm{sm}}$.

Proof. First let $[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)$. Corollary 2.6 tells that $\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r \to \mathfrak{X}_r(\mathsf{GL}_n)$ is an étale equivalence and such mappings preserve tangent spaces, we conclude

$$T_{[\rho]}(\mathfrak{X}_r(\mathsf{GL}_n)) \cong T_{[\rho]}(\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r) \cong T_{[\rho]}(\mathfrak{X}_r(\mathsf{SL}_n)) \oplus \mathbb{C}^r.$$

By counting dimensions and noticing that

$$\dim_{\mathbb{C}}(\mathfrak{X}_r(\mathsf{GL}_n)) = \dim_{\mathbb{C}}(\mathfrak{X}_r(\mathsf{SL}_n)) + r,$$

results (i) and (ii) follow.

Results (iii) and (iv) follow from (i) and (ii) and the additional observations that $\dim_{\mathbb{C}}(\mathfrak{X}_r(K_{\mathbb{C}})) = \dim_{\mathbb{R}}(\mathfrak{X}_r(K))$ and $\dim_{\mathbb{C}}\big(T_{[\psi]}(\mathfrak{X}_r(K_{\mathbb{C}}))\big) = \dim_{\mathbb{R}}\big(T_{[\psi]}(\mathfrak{X}_r(K))\big)$. □

Corollary 2.8. *We have the following isomorphisms of character varieties*:

(i) $\mathfrak{X}_r(\mathsf{PSL}_n) \cong \mathfrak{X}_r(\mathsf{SL}_n)/\!\!/\mathbb{Z}_n^r$ *in the category of algebraic varieties.*

(ii) $\mathfrak{X}_r(\mathsf{PU}_n) \cong \mathfrak{X}_r(\mathsf{SU}_n)/\mathbb{Z}_n^r$ *in the category of semialgebraic sets.*

Proof. From the previous theorem we have

$$\mathfrak{X}_r(\mathsf{GL}_n) \cong \mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathbb{Z}_n^r} (\mathbb{C}^*)^r.$$

Taking the quotient of both sides by $(\mathbb{C}^*)^r$ we can conclude $\mathfrak{X}_r(\mathsf{PSL}_n) \cong \mathfrak{X}_r(\mathsf{SL}_n)/\!\!/\mathbb{Z}_n^r$. More precisely letting $\mu = (\mu_1, \dots, \mu_r) \in (\mathbb{C}^*)^r$ act only on the second factor of $\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r$,

$$\mu \cdot \big([(A_1, \dots, A_r)], (\lambda_1, \dots, \lambda_r)\big) = \big([(A_1, \dots, A_r)], (\mu_1\lambda_1, \dots, \mu_r\lambda_r)\big),$$

and going through the isomorphisms in (3), one gets that the action on $\mathfrak{X}_r(\mathsf{GL}_n)$ corresponds to scalar multiplication of each entry, so we obtain

$$\begin{aligned}\mathfrak{X}_r(\mathsf{PSL}_n) \cong \mathfrak{X}_r(\mathsf{GL}_n)/\!\!/(\mathbb{C}^*)^r &\cong \big(\mathfrak{X}_r(\mathsf{SL}_n) \times_{\mathbb{Z}_n^r} (\mathbb{C}^*)^r\big)/\!\!/(\mathbb{C}^*)^r \\ &\cong \big((\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r)/\!\!/\mathbb{Z}_n^r\big)/\!\!/(\mathbb{C}^*)^r \\ &\cong \big((\mathfrak{X}_r(\mathsf{SL}_n) \times (\mathbb{C}^*)^r)/\!\!/(\mathbb{C}^*)^r\big)/\!\!/\mathbb{Z}_n^r \cong \mathfrak{X}_r(\mathsf{SL}_n)/\!\!/\mathbb{Z}_n^r,\end{aligned}$$

as wanted. The other statement is analogous. □

2B. ***Examples.*** We use the results in Section 2A and the theorems from [Florentino and Lawton 2009] to describe the homeomorphism types of the examples of G-character varieties of $\mathtt{F}_r$ known to be manifolds with boundary. Let $\overline{B}_n$ denote a closed real ball of indicated dimension, and let $\{*\}$ denote the space consisting of one point.

One can show that whenever $\phi : \mathfrak{X}_r(\mathsf{SL}_n) \to M$ is an isomorphism (as affine varieties), then $\mathfrak{X}_r(\mathsf{SU}_n) \cong \phi(\mathfrak{X}_r(\mathsf{SU}_n)) \subset M$ (as semialgebraic sets) by restricting ϕ to $\mathfrak{X}_r(\mathsf{SU}_n) \subset \mathfrak{X}_r(\mathsf{SL}_n)$ [Procesi and Schwarz 1985].

We first consider the trivial case $(r, n) = (r, 1)$. In this case the conjugation action is trivial, and thus we deduce the following table of moduli of $(r, 1)$-representations:

	fixed determinant	nonfixed determinant
complex	$\mathfrak{X}_r(\mathsf{SL}_1) \cong \{*\}$	$\mathfrak{X}_r(\mathsf{GL}_1) \cong (\mathbb{C}^*)^r$
compact	$\mathfrak{X}_r(\mathsf{SU}_1) \cong \{*\}$	$\mathfrak{X}_r(\mathsf{U}_1) \cong (S^1)^r$

We next consider the case $r = 1$. The coefficients of the characteristic polynomial of a matrix X, $\{c_1(X), \dots, c_{n-1}(X), \det(X)\}$, define conjugate invariant regular mappings $\mathfrak{X}_1(\mathsf{SL}_n) \to \mathbb{C}^{n-1}$ and $\mathfrak{X}_1(\mathsf{GL}_n) \to \mathbb{C}^{n-1} \times \mathbb{C}^*$ which are isomorphisms. Thus we get this table of moduli of $(1, n)$-representations:

	fixed determinant	nonfixed determinant
complex	$\mathfrak{X}_1(\mathsf{SL}_n) \cong \mathbb{C}^{n-1}$	$\mathfrak{X}_1(\mathsf{GL}_n) \cong \mathbb{C}^{n-1} \times \mathbb{C}^*$
compact	$\mathfrak{X}_1(\mathsf{SU}_n) \cong \overline{B}_{n-1}$	$\mathfrak{X}_1(\mathsf{U}_n) \cong \overline{B}_{n-1} \times S^1$

Remark 2.9. With respect to the second table, each of the four families of moduli spaces contains no irreducible representations, yet each space is smooth. For this reason these moduli spaces should perhaps be regarded as everywhere singular, since we will see that irreducibles will generally be smooth points for $r \geq 2$.

In the $r = 2$ case we have a well-known isomorphism $\mathfrak{X}_2(\mathsf{SL}_2) \to \mathbb{C}^3$ given by $[(A, B)] \mapsto (\operatorname{tr}(A), \operatorname{tr}(B), \operatorname{tr}(AB))$; see [Goldman 2009; Vogt 1889; Fricke and Klein 1912]. More generally there is an isomorphism $\mathfrak{gl}(2, \mathbb{C})^2 /\!\!/ \mathsf{PGL}_2 \to \mathbb{C}^5$ given by

$$[(A, B)] \mapsto (\operatorname{tr}(A), \operatorname{tr}(B), \operatorname{tr}(AB), \det(A), \det(B)).$$

Therefore we get the following moduli of $(2, 2)$-representations:

	fixed determinant	nonfixed determinant
complex	$\mathfrak{X}_2(\mathsf{SL}_2) \cong \mathbb{C}^3$	$\mathfrak{X}_2(\mathsf{GL}_2) \cong \mathbb{C}^3 \times (\mathbb{C}^* \times \mathbb{C}^*)$
compact	$\mathfrak{X}_2(\mathsf{SU}_2) \cong \overline{B}_3$	$\mathfrak{X}_2(\mathsf{U}_2) \cong \overline{B}_3 \times (S^1 \times S^1)$

In [Florentino and Lawton 2009] the following fixed determinant cases are established:

	fixed determinant	nonfixed determinant
compact $(3, 2)$	$\mathfrak{X}_3(\mathsf{SU}_2) \cong S^6$	$\mathfrak{X}_3(\mathsf{U}_2) \cong S^6 \times_{\mathbb{Z}_2^3} (S^1 \times S^1 \times S^1)$
compact $(2, 3)$	$\mathfrak{X}_2(\mathsf{SU}_3) \cong S^8$	$\mathfrak{X}_2(\mathsf{U}_3) \cong S^8 \times_{\mathbb{Z}_3^2} (S^1 \times S^1)$

Remark 2.10. The complex $(3, 2)$ and $(2, 3)$ cases are left out in this last table since we will show they are not manifolds. In each of these cases, the complex

moduli space of fixed determinant is a branched double cover of complex affine space which deformation retract to a sphere. The explicit scheme structures are known as well. See [Florentino and Lawton 2009; Lawton 2007].

We conjectured in [Florentino and Lawton 2009] that this covers all the cases where a topological manifold with boundary can arise. We will prove this conjecture in Section 4.

3. Singularities

3A. ***Algebro-geometric singularities.*** There are a number of equivalent ways to describe smoothness for affine varieties.

Let $X = V(f_1, \ldots, f_k) \subset \mathbb{C}^n$ be an affine variety. Then its tangent space at the point $\boldsymbol{p} = (p_1, \ldots, p_n) \in X$ is the vector space

$$T_{\boldsymbol{p}}(X) = \left\{ (v_1, \ldots, v_n) \in \mathbb{C}^n \;\middle|\; \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\bigg|_{\boldsymbol{p}} (v_j - p_j) = 0 \text{ for all } i \right\}.$$

This coincides with the more general definition $T_{\boldsymbol{p}}(X) = (\mathfrak{m}_{\boldsymbol{p}}/\mathfrak{m}_{\boldsymbol{p}}^2)^*$, which is the dual to the cotangent space $\mathfrak{m}_{\boldsymbol{p}}/\mathfrak{m}_{\boldsymbol{p}}^2$, where $\mathfrak{m}_{\boldsymbol{p}}$ is a maximal ideal in $\mathbb{C}[X]$ corresponding to $\boldsymbol{p}$ by Hilbert's Nullstellensatz.

Definition 3.1. The singular locus of X is defined to be

$$X^{\text{sing}} = \{\boldsymbol{p} \in X \mid \dim_{\mathbb{C}} T_{\boldsymbol{p}}(X) > \dim_{\text{Krull}} X\}.$$

The complement of this set, $X - X^{\text{sing}}$, is a complex manifold. If X is irreducible, then X is path-connected and furthermore $X - X^{\text{sing}}$ is likewise path-connected. See [Shafarevich 1994].

Let $c = n - \dim_{\text{Krull}} X$. And let J be the $k \times n$ Jacobian matrix of partial derivatives of the k relations defining $X \subset \mathbb{C}^n$. We can assume n is minimal. Then X^{sing} is concretely realized as the affine variety determined by the determinant of the $c \times c$ minors of J. This ideal is referred to as the Jacobian ideal. In this way, X^{sing} is seen to be a proper subvariety of X.

For example, in [Heusener and Porti 2004] it is shown (for $r \geq 3$) that

$$\mathfrak{X}_r(\mathsf{SL}_2)^{\text{sing}} = \mathfrak{X}_r(\mathsf{SL}_2)^{\text{red}}.$$

In [Lawton 2007], explicitly computing the Jacobian ideal, a similar result is also shown: $\mathfrak{X}_2(\mathsf{SL}_3)^{\text{red}} = \mathfrak{X}_2(\mathsf{SL}_3)^{\text{sing}}$.

3B. ***Tangent spaces.*** Let $\mathfrak{g}$ be the Lie algebra of G. Having addressed the $r = 1$ and $n = 1$ cases, we now assume that $r, n \geq 2$.

The following two lemmas are classical, and in fact are true for any algebraic Lie group over $\mathbb{R}$ or $\mathbb{C}$. See [Weil 1964]. For a representation $\rho : \mathtt{F}_r \to G$, let us denote by $\mathfrak{g}_{\mathrm{Ad}_\rho}$ the $\mathtt{F}_r$-module $\mathfrak{g}$ with the adjoint action via ρ. That is, any word $w \in \mathtt{F}_r$ acts as $w \cdot X = \mathrm{Ad}_{\rho(w)} X = \rho(w) X \rho(w)^{-1}$, for $X \in \mathfrak{g}$. Consider the cocycles, coboundaries and cohomology of $\mathtt{F}_r$ with coefficients in this module. Explicitly:

$$Z^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}) := \{u : \mathtt{F}_r \to \mathfrak{g} \mid u(xy) = u(x) + \mathrm{Ad}_{\rho(x)} u(y)\},$$
$$B^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}) := \{u : \mathtt{F}_r \to \mathfrak{g} \mid u(x) = \mathrm{Ad}_{\rho(x)} X - X \text{ for some } X \in \mathfrak{g}\},$$
$$H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}) := Z^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}) / B^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}).$$

Lemma 3.2. *Let G be any algebraic Lie group over $\mathbb{R}$ or $\mathbb{C}$.*

$$T_\rho\left(\mathfrak{R}_r(G)\right) \cong \mathfrak{g}^r \cong Z^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}).$$

Let $\mathsf{Orb}_\rho = \{g\rho g^{-1} \mid g \in G\}$ be the G-orbit of ρ, and let $\mathsf{Stab}_\rho = \{g \in G \mid g\rho g^{-1} = \rho\}$ be the G-stabilizer (or isotropy subgroup).

Lemma 3.3. *Let G be any algebraic Lie group over $\mathbb{R}$ or $\mathbb{C}$.*

$$T_\rho(\mathsf{Orb}_\rho) \cong \mathfrak{g}/\{X \in \mathfrak{g} \mid \mathrm{Ad}_{\rho(x)} X = X\} \cong B^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}).$$

It is not always the case that the tangent space to the quotient is the quotient of tangent spaces. Consider representations from the free group of rank 1 into SL_3. The ring of invariants is two dimensional and the ring is generated by $\mathrm{tr}(X)$ and $\mathrm{tr}(X^{-1})$. So the ideal is zero and the ring is free. Consequently it is smooth and the representation sending everything to the identity (having maximal stabilizer) is a nonsingular point. This illustrates that there can be smooth points in the quotient that have positive-dimensional stabilizer. At these points, $T_\rho(\mathfrak{R}_r(G)/\!\!/G) \ncong T_\rho(\mathfrak{R}_r(G))/T_\rho(\mathsf{Orb}_\rho)$, seen by simply comparing dimensions.

We also note that if we replace free groups by finitely generated groups Γ then the above isomorphisms require a more careful treatment due to the possible existence of nilpotents in the coordinate ring of the scheme associated to $\mathfrak{R}_\Gamma(G)$ [Sikora 2012].

Recall that $\mathfrak{R}_r(G)^s$ is the set of irreducible representations, and $\mathfrak{X}_r(G)^s = \mathfrak{R}_r(G)^s/G$. An action is called *locally free* if the stabilizer is finite, and is called proper if the action $G \times X \to X \times X$ is a proper mapping. In general, the quotient by a proper locally free action of a reductive group on a smooth manifold is an orbifold (a space locally modeled on finite quotients of $\mathbb{R}^n$).

The following lemma can be found in [Johnson and Millson 1987, pp. 54–57]. See also [Goldman 1990; 1984].

Lemma 3.4. *Let G be reductive and $r, n \geq 2$. The $\mathsf{P}G$ action on $\mathfrak{R}_r(G)^s$ is locally free and proper.*

Therefore, $\mathfrak{R}_r(G)^s/G = \mathfrak{R}_r(G)^s/\mathsf{P}G$ are orbifolds.

Lemma 3.5. *For G equal to SL_n, GL_n, SU_n, or U_n and $r, n \geq 2$, the associated $\mathsf{P}G$ action on $\mathfrak{R}_r(G)^s$ is free. Therefore, in these cases $\mathfrak{R}_r(G)^s/G$ is a smooth manifold.*

Proof. Let $\rho = (X_1, \dots, X_r) \in \mathfrak{R}_r(G)^s$. Then by Burnside's theorem [Lang 2002] the collection $\{X_1, \dots, X_r\}$ generates all of $n \times n$ matrices $M_{n\times n}$ as an algebra, since $r > 1$ and they form an irreducible set of matrices. Suppose there exists $g \in G$ so that for all $1 \leq k \leq r$ we have $gX_kg^{-1} = X_k$. Then g stabilizes all of $M_{n\times n}$.

Consider $M = \mathbb{C}^n$ as a module over $R = M_{n\times n}$. Clearly, M is a simple module since no nontrivial proper subspaces are left invariant by all matrices. Let f_g be the automorphism of $\mathbb{C}^n$ defined by mapping $v \mapsto gv$. Then f_g defines an R-module automorphism of M since g stabilizes all of R. Thus by Schur's lemma the action of g is equal to the action of a scalar; that is, g is central. □

Lemma 3.5 and Lemma 3.16 (see Section 3D) together immediately imply the following corollary.

Corollary 3.6. *Let $G = \mathsf{SL}_n, \mathsf{GL}_n, \mathsf{SU}_n$, or U_n. If $[\rho] \in \mathfrak{X}_r(G)^s$ and $r, n \geq 2$, then*

$$T_{[\rho]}(\mathfrak{X}_r(G)) \cong H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}).$$

For $G = \mathsf{SL}_n$ we can calculate that $\dim_{\mathbb{C}} \mathfrak{X}_r(G)^s = (n^2-1)(r-1)$ and for $K = \mathsf{SU}_n$, we have $\dim_{\mathbb{R}} \mathfrak{X}_r(K)^s = (n^2-1)(r-1)$. Likewise, for $G = \mathsf{GL}_n$ we calculate $\dim_{\mathbb{C}} \mathfrak{X}_r(G)^s = n^2(r-1)+1$ and for $K = \mathsf{U}_n$, $\dim_{\mathbb{R}} \mathfrak{X}_r(K)^s = n^2(r-1)+1$.

Let $\mathfrak{X}_r(G)^{\mathrm{sm}} = \mathfrak{X}_r(G) - \mathfrak{X}_r(G)^{\mathrm{sing}}$ be the smooth stratum, which is a complex manifold, open and dense as a subspace of $\mathfrak{X}_r(G)$. The calculation of dimensions above and Corollary 3.6 imply the following lemma which expresses the fact that the irreducibles not only form a smooth manifold but are naturally contained in the smooth stratum of the variety.

Lemma 3.7. *Let $r, n \geq 2$ and G be one of SL_n or GL_n. Then the following equivalent statements hold*:

(i) $\mathfrak{X}_r(G)^s \subset \mathfrak{X}_r(G)^{\mathrm{sm}}$.

(ii) $\mathfrak{X}_r(G)^{\mathrm{sing}} \subset \mathfrak{X}_r(G)^{\mathrm{red}}$.

The next lemmas address important technical points that we will need in our proofs.

Lemma 3.8. *$\mathfrak{X}_r(G)^{\mathrm{red}}$ is an algebraic set*; *that is*, *a subvariety of $\mathfrak{X}_r(G)$.*

Proof. The irreducibles are exactly the GIT stable points (zero dimensional stabilizer and closed orbits) and in general these are Zariski open, which implies the complement is an algebraic set [Dolgachev 2003]. □

Lemma 3.9. *Suppose there exists a set $\mathbb{O} \subset \mathfrak{X}_r(G)^{\mathrm{sing}} \cap \mathfrak{X}_r(G)^{\mathrm{red}}$ that is dense with respect to the ball topology in $\mathfrak{X}_r(G)^{\mathrm{red}}$. Then $\mathfrak{X}_r(G)^{\mathrm{sing}} = \mathfrak{X}_r(G)^{\mathrm{red}}$.*

Proof. Since both $\mathfrak{X}_r(G)^{\mathrm{sing}} \subset \mathfrak{X}_r(G)^{\mathrm{red}}$ are subvarieties (by Lemmas 3.7 and 3.8), $\mathbb{O}$ is dense in both with respect to the ball topology. This follows since $\mathbb{O}$ is dense in $\mathfrak{X}_r(G)^{\mathrm{red}}$ with respect to the ball topology and $\mathbb{O} \subset \mathfrak{X}_r(G)^{\mathrm{sing}} \cap \mathfrak{X}_r(G)^{\mathrm{red}}$. Thus $\mathfrak{X}_r(G)^{\mathrm{sing}} = \overline{\mathbb{O}} = \mathfrak{X}_r(G)^{\mathrm{red}}$, where $\overline{\mathbb{O}}$ is the closure of $\mathbb{O}$ in $\mathfrak{X}_r(G)$ with respect to the (metric) ball topology. □

A set as in Lemma 3.9 was called an *adherence set* in [Heusener and Porti 2004].

3C. ***Denseness of reducibles with minimal stabilizer.*** Now consider the following subvarieties of reducibles. Recall that the 0 vector space is not considered to be an irreducible subrepresentation.

Definition 3.10. Define $U_{r,n} \subset \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$ and $W_{r,n} \subset \mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{red}}$ by:

$$U_{r,n} = \left\{[\rho_1 \oplus \rho_2] \in \mathfrak{X}_r(\mathsf{GL}_n) : \ \rho_i : \mathsf{F}_r \to \mathsf{GL}_{n_i} \text{ are irreducible}\right\}$$
$$W_{r,n} = \left\{[\rho_1 \oplus \rho_2] \in \mathfrak{X}_r(\mathsf{SL}_n) : \ \rho_i : \mathsf{F}_r \to \mathsf{GL}_{n_i} \text{ are irreducible}\right\},$$

where we consider all possible decompositions $n = n_1 + n_2$, with $n_i > 0$.

Note that a given $\rho \in U_{r,n}$ uniquely determines the integers n_1 and n_2, up to permutation. We will refer to this situation by saying that ρ is of *reduced type* $[n_1, n_2]$. Similar remarks and terminology apply to $W_{r,n}$.

It is clear that

$$\mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{red}} = \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}} \cap \mathfrak{X}_r(\mathsf{SL}_n) \tag{4}$$

and that

$$W_{r,n} = U_{r,n} \cap \mathfrak{X}_r(\mathsf{SL}_n).$$

The following lemma is likewise clear by the proof of Lemma 3.5.

Lemma 3.11. *A representation ρ is in $U_{r,n}$ if and only if* $\mathsf{Stab}_\rho \cong (\mathbb{C}^*)^2$. *Also, $\rho \in W_{r,n}$ if and only if* $\mathsf{Stab}_\rho \cong \mathbb{C}^*$.

The strategy is now to show that $U_{r,n}$ and $W_{r,n}$ contain only singularities. However, we must first establish the following lemma.

Lemma 3.12. *Let $r, n \geq 2$. $U_{r,n}$ is dense in $\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$ with respect to the ball topology.*

Proof. When $n = 2$, $U_{r,n}$ coincides with $\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$, since any completely reducible representation is of reduced type $[1, 1]$. So we assume $n \geq 3$. Let $\rho \in [\rho] \in \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$ have at least three irreducible blocks; that is, $\rho = \rho_1 \oplus \rho_2 \oplus \rho_3$ where ρ_1 and ρ_2 are irreducible and ρ_3 is semisimple. In other words, $[\rho] \in \mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}} - U_{r,n}$.

Then $\rho_2 \oplus \rho_3$ is a semisimple representation into GL_k for some k. Since the irreducible representations $\mathsf{F}_r \to \mathsf{GL}_k$ are dense (here we use $r > 1$), there exists an irreducible sequence $\sigma_j \in \operatorname{Hom}(\mathtt{F}_r, \mathsf{GL}_k)$ satisfying

$$\lim_{j\to\infty} \sigma_j = \rho_2 \oplus \rho_3,$$

which in turn implies

$$\lim_{j\to\infty} \rho_1 \oplus \sigma_j = \rho_1 \oplus \rho_2 \oplus \rho_3 = \rho,$$

where $\rho_1 \oplus \sigma_j$ is in $U_{r,n}$. Thus we have a sequence $[\rho_1 \oplus \sigma_j] \in U_{r,n} \subset \mathfrak{X}_r(\mathsf{GL}_n)^{\text{red}}$ whose limit is $[\rho_1 \oplus \rho_2 \oplus \rho_3]$. This shows that $U_{r,n}$ is dense in $\mathfrak{X}_r(\mathsf{GL}_n)^{\text{red}}$ and proves the lemma. □

Corollary 3.13. *Let $r, n \geq 2$. Then $W_{r,n}$ is dense in $\mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}}$ with respect to the ball topology.*

Proof. First we show that $\mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}} \subset \overline{W_{r,n}}$. Using the previous lemma and Equation (4), let

$$\begin{aligned}[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}} &= \mathfrak{X}_r(\mathsf{SL}_n) \cap \mathfrak{X}_r(\mathsf{GL}_n)^{\text{red}} \\ &= \mathfrak{X}_r(\mathsf{SL}_n) \cap \overline{U_{r,n}}.\end{aligned}$$

Then, we can write $\rho = \lim \sigma_j$, where $\sigma_j = \rho_1^{(j)} \oplus \rho_2^{(j)} \in U_{r,n}$ is of reduced type $[n_1, n_2]$. Let us write $\lambda_j := \det \rho_1^{(j)} \det \rho_2^{(j)}$. Since the limit is a well-defined point $[\rho] \in \mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}}$, we can arrange for the sequence to be in $W_{r,n}$ as follows. Letting $\alpha_j = (1/\lambda_j)^{1/n_1}$ (for any choice of branch cut), we can also write $\rho = \lim \eta_j$ where $\eta_j = (\rho_1^{(j)}\alpha_j) \oplus \rho_2^{(j)} \in W_{r,n}$, (since now η_j has unit determinant), from which one sees that $\rho \in \overline{W_{r,n}}$, as wanted. Finally, we get:

$$\begin{aligned}\mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}} \subset \overline{W_{r,n}} &= \overline{\mathfrak{X}_r(\mathsf{SL}_n) \cap U_{r,n}} \\ &\subset \mathfrak{X}_r(\mathsf{SL}_n) \cap \overline{U_{r,n}} \\ &= \mathfrak{X}_r(\mathsf{SL}_n)^{\text{red}},\end{aligned}$$

which implies all these sets coincide, finishing the proof. Here, we used the standard fact that the closure of an intersection is contained in the intersection of the closures, and that $\mathfrak{X}_r(\mathsf{SL}_n)$ is closed in $\mathfrak{X}_r(\mathsf{GL}_n)$. □

3D. *The Luna slice theorem and the Zariski tangent space.* We now prove a strong lemma, first proved in [Heusener and Porti 2004] and later in more generality in [Sikora 2012], which tells exactly how to understand the Zariski tangent space at a general free group representation. For a similar result see also [Drézet 2004, p. 45]. To that end, we review the Luna slice theorem [1973]. We recommend [Drézet 2004] for a good exposition.

Following [Schwarz 2004], we define an étale map between complex affine varieties as a local analytic isomorphism in the ball topology.

Theorem 3.14 (weak Luna slice theorem at smooth points). *Let G be a reductive algebraic group acting on an affine variety X. Let $x \in X$ be a smooth point with* Orb_x *closed. Then there exists a subvariety $x \in V \subset X$, and* Stab_x*-invariant étale morphism $\phi : V \to T_x V$ satisfying*:

(i) *V is locally closed, affine, smooth, and* Stab_x*-stable.*

(ii) *$V \hookrightarrow X \to X/\!\!/G$ induces $T_{[x]}(V/\!\!/\mathsf{Stab}_x) \cong T_{[x]}(X/\!\!/G)$.*

(iii) *$\phi(x) = 0$ and $d\phi_x = \mathrm{Id}$.*

(iv) *$T_x X = T_x(\mathsf{Orb}_x) \oplus T_x V$ with respect to the* Stab_x*-action.*

(v) *ϕ induces $T_{[x]}(V/\!\!/\mathsf{Stab}_x) \cong T_0(T_x V/\!\!/\mathsf{Stab}_x)$.*

Remark 3.15. The reader familiar with Luna's slice theorem may be wondering how Theorem 3.14 is implied. Firstly, note that ψ is an étale mapping if and only if the completion of the local rings satisfy $\widehat{\mathcal{O}}_x \cong \widehat{\mathcal{O}}_{\psi(x)}$ which implies the subset of derivations are isomorphic, the latter being isomorphic to the Zariski tangent spaces. The usual Luna slice theorem implies $\phi : V/\!\!/\mathsf{Stab}_x \to \phi(V)/\!\!/\mathsf{Stab}_x$ is étale, $(G \times V)/\!\!/\mathsf{Stab}_x \cong U \subset X$ is saturated and open, and $V/\!\!/\mathsf{Stab}_x \to U/\!\!/G$ is étale. We thus respectively conclude lines (v), (iv), and (ii) in the above theorem.

Lemma 3.16. *Let G be a complex algebraic reductive Lie group. For any $[\rho] \in \mathfrak{X}_r(G)$,*

$$T_{[\rho]}\mathfrak{X}_r(G) \cong T_0\big(H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})/\!\!/\mathsf{Stab}_{\rho^{ss}}\big),$$

where ρ^{ss} is a polystable representative from the extended orbit $[\rho]$.

Proof. Any $\rho^{ss} \in [\rho]$ has a closed orbit and is a smooth point of $\mathfrak{R}_r(G)$, and every point $[\rho] \in \mathfrak{X}_r(G)$ contains such a ρ^{ss}.

By the Luna slice theorem, there exists an algebraic set $\rho^{ss} \in V_{\rho^{ss}} \subset \mathfrak{R}_r(G)$ such that:

(i) $\mathsf{Stab}_{\rho^{ss}}(V_{\rho^{ss}}) \subset V_{\rho^{ss}}$

(ii) With respect to the reductive action of $\mathsf{Stab}_{\rho^{ss}}$,

$$\begin{aligned} Z^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}}) &\cong T_{\rho^{ss}}(\mathfrak{R}_r(G)) \cong T_{\rho^{ss}}(\mathsf{Orb}_{\rho^{ss}}) \oplus T_{\rho^{ss}}(V_{\rho^{ss}}) \\ &\cong B^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}}) \oplus T_{\rho^{ss}}(V_{\rho^{ss}}), \end{aligned}$$

since ρ^{ss} is smooth.

(iii) Thus, $H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}}) \cong T_{\rho^{ss}}(V_{\rho^{ss}})$, as $\mathsf{Stab}_{\rho^{ss}}$-spaces.

(iv) $V_{\rho^{ss}} \hookrightarrow \mathfrak{R}_r(G) \to \mathfrak{X}_r(G)$ induces $T_{[\rho^{ss}]}(V_{\rho^{ss}}/\!\!/\mathsf{Stab}_{\rho^{ss}}) \cong T_{[\rho]}\mathfrak{X}_r(G)$.

(v) $T_{[\rho^{ss}]}(V_{\rho^{ss}}/\!\!/\mathsf{Stab}_{\rho^{ss}}) \cong T_0\big(T_{\rho^{ss}}(V_{\rho^{ss}})/\!\!/\mathsf{Stab}_{\rho^{ss}}\big)$, since ρ^{ss} is smooth.

Putting these steps together we conclude

$$T_{[\rho]}\mathfrak{X}_r(G) \cong T_0\big(T_{\rho^{ss}}(V_{\rho^{ss}})/\!\!/\mathsf{Stab}_{\rho^{ss}}\big) \cong T_0\big(H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})/\!\!/\mathsf{Stab}_{\rho^{ss}}\big). \qquad \square$$

Remark 3.17. Upon closer examination we find $H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})/\!\!/\mathsf{Stab}_{\rho^{ss}}$ to be an étale neighborhood; that is, an algebraic set that maps, via an étale mapping, to an open set (in the ball topology) of $\mathfrak{X}_r(G)$ [Schwarz 2004, p. 223].

3E. ***The $\mathbb{C}^*$-action on cohomology.*** As we saw in Corollary 3.13, the generic singularity will occur when Stab_ρ is the smallest possible torus group, namely $\mathbb{C}^*$ or $\mathbb{C}^* \times \mathbb{C}^*$, for the cases $G = \mathsf{SL}_n$ or $G = \mathsf{GL}_n$, respectively.

To study the $\mathbb{C}^*$-action on cohomology, the following setup will be relevant.

Fix two integers $n, k \geq 1$. Consider the vector space $\mathbb{C}^{2n} = \mathbb{C}^n \times \mathbb{C}^n$ with variables $(\boldsymbol{z}, \boldsymbol{w}) = (z_1, \ldots, z_n, w_1, \ldots, w_n)$ and the action of $\mathbb{C}^*$ given by

$$\lambda \cdot (\boldsymbol{z}, \boldsymbol{w}) = (\lambda^k \boldsymbol{z}, \lambda^{-k} \boldsymbol{w}). \tag{5}$$

Let us denote by $\mathbb{C}^{2n}/\!\!/_k\mathbb{C}^*$ the corresponding affine GIT quotient. It is the spectrum of the ring $\mathbb{C}[\boldsymbol{z}, \boldsymbol{w}]^{\mathbb{C}^*}$ of polynomial invariants under this action. To describe this ring, let

$$p(\boldsymbol{z}, \boldsymbol{w}) = z_1^{a_1} \cdots z_n^{a_n} w_1^{b_1} \cdots w_n^{b_n}$$

be a monomial, with $a_i, b_i \in \mathbb{N}$, and define

$$\partial p := \sum_{j=1}^{n} a_j - b_j .$$

Any polynomial invariant under the action is a sum of monomials p such that $\partial p = 0$. Considering the monomials with smallest degree, we are led to conclude that

$$\mathbb{C}[\boldsymbol{z}, \boldsymbol{w}]^{\mathbb{C}^*} = \mathbb{C}[z_1 w_1, \ldots, z_1 w_n, \ldots, z_n w_1, \ldots, z_n w_n].$$

Note that this shows that the quotient is independent of k. By viewing these n^2 generators as elements of an $n \times n$ matrix $X = (x_{ij}),\ x_{ij} = z_i w_j$, which necessarily has rank at most one, we conclude that this is the ring of polynomial functions on the variety $V \subset M_{n\times n}(\mathbb{C})$ of matrices of rank ≤ 1:

$$\mathbb{C}[\boldsymbol{z}, \boldsymbol{w}]^{\mathbb{C}^*} = \mathbb{C}[V].$$

The variety V is called a determinantal variety [Harris 1992] and one can show that $\mathbb{C}[V] = \mathbb{C}\left[x_{ij}\right]/I$ where I is the ideal of 2×2 minors of X. By simple computations, V has a unique singularity, the zero matrix, which corresponds to the orbit of zero in $\mathbb{C}^{2n}$.

Now, observe that all orbits of the action (5) are closed except those contained in

$$Z := \{0\} \times \mathbb{C}^n \cup \mathbb{C}^n \times \{0\},$$

and moreover there is only one closed orbit in Z, which is easily seen to be the only singular point of $\mathbb{C}^{2n} /\!\!/_k \mathbb{C}^*$. Therefore, by GIT, the quotient

$$(\mathbb{C}^{2n} \setminus Z)/\mathbb{C}^*$$

is a geometric quotient. We summarize these results as follows.

Lemma 3.18. *Let $n \geq 2$.*

(i) *$\mathbb{C}^{2n} /\!\!/_k \mathbb{C}^*$ is isomorphic to the determinantal variety of $n \times n$ square matrices of rank ≤ 1. Its unique singularity is the orbit of the origin.*

(ii) *$(\mathbb{C}^{2n} \setminus Z)/\mathbb{C}^*$ is isomorphic to $\mathbb{C}^* \times \mathbb{CP}^{n-1} \times \mathbb{CP}^{n-1}$.*

Because of the fact that the GIT quotient is obtained from $(\mathbb{C}^{2n} \setminus Z)/\mathbb{C}^*$ by adding just one point, the singular point, and because of (ii) above, we will refer to $\mathbb{C}^{2n} /\!\!/_k \mathbb{C}^*$ as an affine cone over $\mathbb{CP}^{n-1} \times \mathbb{CP}^{n-1}$, and denote it by $\mathscr{C}_{\mathbb{C}}(\mathbb{CP}^{n-1} \times \mathbb{CP}^{n-1})$. It is called the affine cone over the Segre variety in [Mukai 2003].

Now consider the following antiholomorphic involution of $\mathbb{C}^{2n} = \mathbb{C}^n \times \mathbb{C}^n$:

$$j : (\boldsymbol{z}, \boldsymbol{w}) \mapsto -(\bar{\boldsymbol{w}}, \bar{\boldsymbol{z}}),$$

and consider the same action as above, but restrict it to $S^1 \subset \mathbb{C}^*$. This will be relevant in the study of the compact quotients. The fixed point set of the involution j is the set

$$F := \{(\boldsymbol{z}, -\bar{\boldsymbol{z}}) : \boldsymbol{z} \in \mathbb{C}\} \subset \mathbb{C}^n \times \mathbb{C}^n,$$

which is canonically identified with the first copy of $\mathbb{C}^n$ (as real vector spaces).

Lemma 3.19. (i) *The S^1-action on $\mathbb{C}^{2n}$ commutes with j.*

(ii) *The quotient F/S^1 of its restriction to F is homeomorphic to a real open cone over $\mathbb{CP}^{n-1}$ denoted by $\mathscr{C}(\mathbb{CP}^{n-1})$.*

Proof. Proving (i) is straightforward, and we leave it to the reader.

To prove (ii) first observe that on the fixed point set, the S^1-action just gives

$$\lambda \cdot (\boldsymbol{z}, -\bar{\boldsymbol{z}}) = (\lambda \boldsymbol{z}, -\bar{\lambda}\bar{\boldsymbol{z}}), \quad \lambda \in S^1$$

so we can describe it as an action of S^1 on the first copy of $\mathbb{C}^n$. Since the action is free except for the origin, all orbits are circles and the quotient $\mathbb{C}^n/S^1$ is the union of $\mathbb{C}^n \setminus \{0\}/S^1$ with a single point. Since $\mathbb{C}^n \setminus \{0\}/S^1$ is homeomorphic to $\left(S^{2n-1}/S^1\right) \times \mathbb{R}$, we obtain that F/S^1 is the real cone over S^{2n-1}/S^1, the latter being well known to be $\mathbb{CP}^{n-1}$. $\square$

These singularity types will be encountered in SL_n and SU_n- character varieties. In fact, the same singularities will also appear in GL_n and U_n-character varieties, because the actions in these cases are very similar.

Indeed one can easily show the following:

Proposition 3.20. *Let $n \geq 2$. Let $T = \mathbb{C}^* \times \mathbb{C}^*$ act on a vector space $V = \mathbb{C}^{2n} = \mathbb{C}^n \times \mathbb{C}^n$ as follows*:

$$(\lambda, \mu) \cdot (z, \boldsymbol{w}) = (\lambda\mu^{-1} z, \mu\lambda^{-1} \boldsymbol{w}).$$

Then, $\mathbb{C}^{2n} /\!\!/ T$ is isomorphic to $\mathbb{C}^{2n} /\!\!/_2 \mathbb{C}^*$. In particular, as before, this quotient is the determinantal variety of $n \times n$ square matrices of rank ≤ 1, which has dimension $2n - 1$. Its unique singularity is the orbit of the origin.

Proof. We just need to argue, as before, that the invariant polynomials are generated by the same monomials, those of the form $z_j w_k$, for any indices $j, k \in \{1, \ldots, n\}$, so they form an $n \times n$ matrix with rank one. $\square$

Finally, note that for $n = 1$ we get a smooth variety: $\mathbb{C}^2 /\!\!/_2 \mathbb{C}^* \cong \mathbb{C}$.

3F. *Proof of Theorem 1.1, Case 1:* GL_n *or* SL_n.

Theorem 3.21. *Let $r, n \geq 2$ and $G = \mathsf{GL}_n$ or SL_n. Then $\mathfrak{X}_r(G)^{\text{sing}} = \mathfrak{X}_r(G)^{\text{red}}$ if and only if $(r, n) \neq (2, 2)$.*

Remark 3.22. If $n = 1$ the statement is vacuously true since in these cases there are no reducibles, nor are there singularities. We have already seen that there are smooth reducibles in the cases $r = 1, n \geq 2$, and $(r, n) = (2, 2)$ since there always exist reducibles in these cases and the entire moduli spaces are smooth.

Proof. Let $G = \mathsf{GL}_n$. By Lemma 3.7 it is enough to show $\mathfrak{X}_r(G)^{\text{red}} \subset \mathfrak{X}_r(G)^{\text{sing}}$.

Let $\rho \in U_{r,n} \subset \mathfrak{R}_r(G)^{\text{red}}$ be of reduced type $[n_1, n_2]$ with $n_1, n_2 > 0$ and $n = n_1 + n_2$ (see Definition 3.10) and write it in the form

$$\rho = \rho_1 \oplus \rho_2 = \begin{pmatrix} \vec{X} & \vec{0}_{n_1 \times n_2} \\ \vec{0}_{n_2 \times n_1} & \vec{Y} \end{pmatrix},$$

where $\vec{X} = (X_1, \ldots, X_r) \in M^r_{n_1 \times n_1}$ and $\vec{Y} = (Y_1, \ldots, Y_r) \in M^r_{n_2 \times n_2}$ and $\vec{0}_{k \times l} = (0_{k \times l}, \ldots, 0_{k \times l})$ where $0_{k \times l}$ is the k by l matrix of zeros and the vector has r entries. Recall that these representations form a dense set in $\mathfrak{X}_r(G)^{\text{red}}$, by Lemma 3.12.

Let $\operatorname{diag}(a_1, \ldots, a_n)$ be an $n \times n$ matrix whose (i, j)-entry is 0 if $i \neq j$ and is equal to a_i otherwise. Then $\mathsf{Stab}_\rho = \mathbb{C}^* \times \mathbb{C}^*$ is given by

$$\operatorname{diag}(\underbrace{\lambda, \ldots, \lambda}_{n_1}, \overbrace{\mu, \ldots, \mu}^{n_2}).$$

We note that the action of the center is trivial so we often consider the stabilizer with respect to the action of G modulo its center.

Then the cocycles satisfy

$$Z^1(\mathtt{F}_r;\mathrm{Ad}_\rho)\cong\mathfrak{g}^r$$
$$=\left\{\begin{pmatrix}\vec{A}&\vec{B}\\\vec{C}&\vec{D}\end{pmatrix}\;\middle|\;\vec{A}\in M^r_{n_1\times n_1},\vec{B}\in M^r_{n_1\times n_2},\vec{C}\in M^r_{n_2\times n_1},\vec{D}\in M^r_{n_2\times n_2}\right\},$$

which has dimension n^2r since it is the tangent space to a representation and the representation variety is smooth. The coboundaries are given by

$$\begin{aligned}B^1(\mathtt{F}_r;\mathrm{Ad}_\rho)&\cong\left\{\begin{pmatrix}A&B\\C&D\end{pmatrix}-\begin{pmatrix}\vec{X}&\vec{0}_{n_1\times n_2}\\\vec{0}_{n_2\times n_1}&\vec{Y}\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}\vec{X}^{-1}&\vec{0}_{n_1\times n_2}\\\vec{0}_{n_2\times n_1}&\vec{Y}^{-1}\end{pmatrix}\right\}\\&\cong\left\{\begin{pmatrix}A&B\\C&D\end{pmatrix}-\begin{pmatrix}\vec{X}A\vec{X}^{-1}&\vec{X}B\vec{Y}^{-1}\\\vec{Y}C\vec{X}^{-1}&\vec{Y}D\vec{Y}^{-1}\end{pmatrix}\right\},\end{aligned}$$

for a fixed element $\left(\begin{smallmatrix}A&B\\C&D\end{smallmatrix}\right)\in\mathfrak{g}$. This has dimension n^2-2 since it is the tangent space to the G-orbit of ρ which has dimension equal to that of the group minus its stabilizer.

Thus with respect to the torus action,

$$H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)\cong H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_1})\oplus H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_2})\oplus W,\tag{6}$$

where W exists since the torus action is reductive. By considering the Euler characteristic, one has that

$$\dim_{\mathbb{C}}H^0(\mathtt{F}_r;\mathrm{Ad}_\rho)-\dim_{\mathbb{C}}H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)=(1-r)\dim_{\mathbb{C}}\mathfrak{gl}(n,\mathbb{C}).$$

Then since $H^0(\mathtt{F}_r;\mathrm{Ad}_\rho)=Z^0(\mathtt{F}_r;\mathrm{Ad}_\rho)$ is the centralizer in $\mathfrak{g}$ of the image of ρ, we calculate

$$\begin{aligned}\dim_{\mathbb{C}}H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)&=n^2(r-1)+2,\\\dim_{\mathbb{C}}H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_i})&=n_i^2(r-1)+1,\quad i=1,2.\end{aligned}$$

This then implies $\dim_{\mathbb{C}}H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)/\!\!/(\mathbb{C}^*\times\mathbb{C}^*)=n^2(r-1)+1=\dim_{\mathbb{C}}\mathfrak{X}_r(G)$, since the diagonal of the $\mathbb{C}^*\times\mathbb{C}^*$-action is the center which acts trivially. We conclude that

$$\dim_{\mathbb{C}}W=(n^2-n_1^2-n_2^2)(r-1)=2n_1n_2(r-1).$$

Explicitly, the Stab_ρ action on $H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)$ is given by

$$\mathrm{diag}(\underbrace{\lambda,\ldots,\lambda}_{n_1},\overbrace{\mu,\ldots,\mu}^{n_2})\cdot\left[\begin{pmatrix}\vec{A}&\vec{B}\\\vec{C}&\vec{D}\end{pmatrix}\right]\mapsto\left[\begin{pmatrix}\vec{A}&\lambda\vec{B}\mu^{-1}\\\mu\vec{C}\lambda^{-1}&\vec{D}\end{pmatrix}\right]$$

which respects representatives up to coboundary.

So, the action on $H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_1})\oplus H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_2})$ is trivial (but not so on W) and we conclude

$$H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)/\!\!/(\mathbb{C}^*\times\mathbb{C}^*)\cong H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_1})\oplus H^1(\mathtt{F}_r;\mathrm{Ad}_{\rho_2})\oplus\left(W/\!\!/(\mathbb{C}^*\times\mathbb{C}^*)\right).$$

Therefore, by Proposition 3.20, we have established that 0 is a singularity (solution to the generators of the singular locus) of $W/\!\!/(\mathbb{C}^*\times\mathbb{C}^*)$ which then implies it is a singularity to $H^1(\mathtt{F}_r;\mathrm{Ad}_\rho)/\!\!/(\mathbb{C}^*\times\mathbb{C}^*)$ (whenever $\dim_{\mathbb{C}} W>2$) which then in turn implies any $\rho\in U_{r,n}$ is a singularity in $\mathfrak{X}_r(G)$ by Lemma 3.16 (note $\rho=\rho^{\mathrm{ss}}$ here). $U_{r,n}$ is dense in $\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$ by Lemma 3.12. Then Lemma 3.9 applies to show that $\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{sing}}=\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$ whenever $\dim_{\mathbb{C}} W=2n_1n_2(r-1)>2$; that is, whenever $(r,n)\neq(2,2)$.

Now let $[\rho]\in\mathfrak{X}_r(\mathsf{SL}_n)$. Then it is easy to see that $[\rho]\in\mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{red}}$ if and only if $[\rho]\in\mathfrak{X}_r(\mathsf{GL}_n)^{\mathrm{red}}$. Then Corollary 2.7 and the previously established case together imply $\mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{red}}=\mathfrak{X}_r(\mathsf{SL}_n)^{\mathrm{sing}}$.

This finishes the proof of Theorem 1.1 for the groups SL_n and GL_n. □

Remark 3.23. We note that the cohomology decomposition used in the proof depends on the decomposition of ρ. For instance, in the 2×2 determinant 1 case, the reducible representation takes values in $\mathsf{SL}_1\times\mathsf{GL}_1=\mathbb{C}^0\times\mathbb{C}^*$, where $\mathbb{C}^0$ is a point. Then by Lemma 3.18:

$$\begin{aligned}H^1(\mathtt{F}_r,\mathrm{Ad}_\rho)/\!\!/\mathbb{C}^*&\cong H^1(\mathtt{F}_r,\mathrm{Ad}_{\rho_1})\oplus H^1(\mathtt{F}_r,\mathrm{Ad}_{\rho_2})\oplus(W/\!\!/\mathbb{C}^*)\\&\cong\mathbb{C}^0\times\mathbb{C}^r\times\left((\mathbb{C}^{2r}/\mathbb{C}^2)/\!\!/\mathbb{C}^*\right)\\&\cong\mathbb{C}^r\times\mathbb{C}^{2r-2}/\!\!/_2\mathbb{C}^*\\&\cong\mathbb{C}^r\times\mathscr{C}_{\mathbb{C}}(\mathbb{C}\mathsf{P}^{r-2}\times\mathbb{C}\mathsf{P}^{r-2}).\end{aligned}$$

Remark 3.24. The proof above works directly, with suitable modifications for the case $G=\mathsf{SL}_n$. For instance the action of the stabilizer in this case is $\mathtt{Stab}_\rho=\mathbb{C}^*$ given by

$$\mathrm{diag}(\underbrace{\lambda,\dots,\lambda,}_{n_1}\overbrace{\mu,\dots,\mu}^{n_2}),$$

where $\lambda^{n_1}\mu^{n_2}=1$ which is equivalent to $\mu=\lambda^{\frac{-n_1}{n_2}}$. The cocycles satisfy

$$\begin{aligned}&Z^1(\mathtt{F}_r;\mathrm{Ad}_\rho)\\&\quad\cong\mathfrak{g}^r=\left\{\begin{pmatrix}\vec{A}&\vec{B}\\\vec{C}&\vec{D}\end{pmatrix}\;\middle|\;\begin{array}{r}\vec{A}\in M^r_{n_1\times n_1},\,\vec{B}\in M^r_{n_1\times n_2},\,\vec{C}\in M^r_{n_2\times n_1},\,\vec{D}\in M^r_{n_2\times n_2},\\\mathrm{tr}(A_i)=-\mathrm{tr}(D_i),\,1\le i\le r\end{array}\right\},\end{aligned}$$

which has dimension $(n^2-1)r$. The rest carries over without significant change.

Remark 3.25. Similar results for the moduli of tuples of generic matrices have been obtained in [Le Bruyn and Procesi 1987], and with respect to the moduli of vector bundles similar results have been obtained in [Laszlo 1996].

3G. ***Proof of Theorem 1.1, Case 2:*** SU_n ***or*** U_n**.** Let $K = \mathsf{SU}_n$ or U_n and let $\mathfrak{k}$ be its Lie algebra in either case.

The tangent space at a point $[\rho] \in \mathfrak{X}_r(K)$ is defined from the semialgebraic structure; that is, any real semialgebraic set has a well-defined coordinate ring which allows one to define the Zariski tangent space as we did at the start of this section [Bochnak et al. 1998]. At smooth points this corresponds to the usual tangent space defined by differentials. It is not hard to see that the semialgebraic set $\mathfrak{X}_r(K)$ is a subset of the real points of $\mathfrak{X}_r(K_{\mathbb{C}})$. Then, the Zariski tangent space of $\mathfrak{X}_r(K)$ at $[\rho]$, $T_{[\rho]}(\mathfrak{X}_r(K))$, consists of the real points of the complex Zariski tangent space $T_{[\rho]}(\mathfrak{X}_r(K_{\mathbb{C}}))$.

As is true for $K_{\mathbb{C}}$-representations, we define a K-representation to be irreducible if it does not admit any proper (nontrivial) invariant subspaces with respect to the standard action on $\mathbb{C}^n$. As with $K_{\mathbb{C}}$-valued representations, we call a K-valued representation reducible if it is not irreducible.

Lemma 3.26. $\mathfrak{X}_r(K_{\mathbb{C}})^{\mathrm{red}} \cap \mathfrak{X}_r(K) = \mathfrak{X}_r(K)^{\mathrm{red}}$.

Proof. First note that $\mathfrak{X}_r(K) \subset \mathfrak{X}_r(K_{\mathbb{C}})$ (see [Florentino and Lawton 2009]). So it suffices to prove that every K-valued representation is K-conjugate to a reducible representation if and only if it is $K_{\mathbb{C}}$-conjugate to a reducible representation.

Let ρ be a K-representation and suppose that it is K-conjugate to a representation that admits a nontrivial proper invariant subspace of $\mathbb{C}^n$, then since $K \subset K_{\mathbb{C}}$ it is true that ρ is $K_{\mathbb{C}}$-conjugate to a reducible representation. Conversely, suppose that a K-representation ρ is $K_{\mathbb{C}}$-conjugate to a reducible representation. However, conjugating by $K_{\mathbb{C}}$ is simply a change-of-basis, and such a change-of-basis is always possible by conjugating by K by using the Gram–Schmidt algorithm. □

Lemma 3.27. $\mathfrak{X}_r(K_{\mathbb{C}})^{\mathrm{sing}} \cap \mathfrak{X}_r(K) = \mathfrak{X}_r(K)^{\mathrm{sing}}$

Proof. Let

$$[\rho] \in \mathfrak{X}_r(K) \subset \mathfrak{X}_r(K_{\mathbb{C}}).$$

Then $[\rho] \in \mathfrak{X}_r(K)^{\mathrm{sing}}$ if and only if $\dim_{\mathbb{R}} T_{[\rho]}\mathfrak{X}_r(K) = \dim_{\mathbb{C}} T_{[\rho]}\mathfrak{X}_r(K_{\mathbb{C}})$ exceeds $\dim_{\mathbb{R}} \mathfrak{X}_r(K) = \dim_{\mathbb{C}} \mathfrak{X}_r(K_{\mathbb{C}})$, the latter occurring if and only if $[\rho] \in \mathfrak{X}_r(K_{\mathbb{C}})^{\mathrm{sing}}$. □

The last case to consider to finish the proof of Theorem 1.1 is $\mathfrak{X}_r(K)$ in terms of SU_n and U_n.

Theorem 3.28. *Let K be either U_n or SU_n. Then $\mathfrak{X}_r(K)^{\mathrm{red}} = \mathfrak{X}_r(K)^{\mathrm{sing}}$ if*

$$\mathfrak{X}_r(K_{\mathbb{C}})^{\mathrm{red}} = \mathfrak{X}_r(K_{\mathbb{C}})^{\mathrm{sing}}.$$

Proof. This follows directly by Lemmas 3.26 and 3.27. □

Since we have already established in Theorem 3.21 that, for $r, n \geq 2$ and $K \in \{\mathsf{U}_n, \mathsf{SU}_n\}$, we have $\mathfrak{X}_r(K_{\mathbb{C}})^{\text{red}} = \mathfrak{X}_r(K_{\mathbb{C}})^{\text{sing}}$ if and only if $(r, n) \neq (2, 2)$, Theorem 3.28 is enough to finish the proof of Theorem 1.1.

3H. ***Iterative reducibles and the singular stratification.*** As above let K be either U_n or SU_n and $G = K_{\mathbb{C}}$, and let the N-th singular stratum be defined by

$$\mathsf{Sing}_N(\mathfrak{X}_r(G)) = (\cdots((\mathfrak{X}_r(G))^{\text{sing}})^{\text{sing}\cdots})^{\text{sing}},$$

which is well-defined since each singular locus is a variety and as such has a singular locus itself.

The N-th level reducibles

$$\mathsf{Red}_N(\mathfrak{X}_r(G)) = (\cdots((\mathfrak{X}_r(G))^{\text{red}})^{\text{red}\cdots})^{\text{red}}$$

is defined inductively in the following way.

Let $\mathsf{Red}_1(\mathfrak{X}_r(G)) = \mathfrak{X}_r(G)^{\text{red}}$. For $k \geq 1$ define $\mathsf{Red}_k(\mathfrak{X}_r(G))^{(k+1)}$ to be the set of $\rho \in \mathsf{Red}_k(\mathfrak{X}_r(G))$ which is minimally reducible, that is has a decomposition into irreducible subrepresentations that has minimal summands. We define $\mathsf{Red}_{k+1}(\mathfrak{X}_r(G)) = \mathsf{Red}_k(\mathfrak{X}_r(G)) - \mathsf{Red}_k(\mathfrak{X}_r(G))^{(k+1)}$ to be the complement of $\mathsf{Red}_k(\mathfrak{X}_r(G))^{(k+1)}$ in $\mathsf{Red}_k(\mathfrak{X}_r(G))$. Thus, $\mathsf{Red}_1(\mathfrak{X}_r(G))^{(2)}$ is always the reducibles that have exactly 2 irreducible subrepresentations — these are exactly the ones we considered in the proof of Theorem 3.21. More generally, $\mathsf{Red}_k(\mathfrak{X}_r(G))^{(k+1)}$ are the representations which decompose into exactly $k+1$ irreducible subrepresentations. For example, $\mathsf{Red}_2(\mathfrak{X}_r(\mathsf{SL}_3))$ are the representations conjugate to a representation that has its semisimplification diagonal, and $\mathsf{Red}_3(\mathfrak{X}_r(\mathsf{SL}_3)) = \varnothing$.

Likewise we have $\mathsf{Red}_N(\mathfrak{X}_r(K))$ and $\mathsf{Sing}_N(\mathfrak{X}_r(K))$.

Theorem 3.29. *Let $r, n \geq 2$ and $(r, n) \neq (2, 2)$. If $N \geq 1$, then*

$$\mathsf{Sing}_N(\mathfrak{X}_r(G)) \cong \mathsf{Red}_N(\mathfrak{X}_r(G)) \quad \textit{and} \quad \mathsf{Sing}_N(\mathfrak{X}_r(K)) \cong \mathsf{Red}_N(\mathfrak{X}_r(K)).$$

The result follows by induction on the irreducible block forms and observing that each block form now corresponds to GL_k, or U_k in the compact cases.

3I. ***Remarks about other groups.***

3I.1. *General reductive groups.* Let G be a reductive complex algebraic group. It can be shown [Sikora 2012] that the definition given before of an irreducible representation $\rho : \Gamma \to G$ corresponds exactly to the quotient group $\mathsf{Stab}_\rho / Z(G)$ being finite.

Proposition 3.30. *If the adjoint action of ρ is irreducible on $\mathfrak{g}$, then $[\rho]$ is smooth in $\mathfrak{X}_r(G)$.*

Proof. If Ad_ρ is irreducible, then $\mathsf{Stab}_{\rho^{ss}}$ is central and so $\mathsf{Stab}_{\rho^{ss}}$ acts trivially on $H^1(\mathtt{F}_r;\mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})$. Hence 0 is not in the Jacobian ideal of $H^1(\mathtt{F}_r;\mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})/\!\!/\mathsf{Stab}_{\rho^{ss}}$. So, by Lemma 3.16, $[\rho]$ is smooth in $\mathfrak{X}_r(G)$. □

From the proof of Proposition 3.30, we obtain:

Corollary 3.31. *Let G be a complex reductive algebraic group and $\rho \in \mathfrak{R}_r(G)$ is irreducible with central stabilizer. Then $[\rho]$ is smooth in $\mathfrak{X}_r(G)$.*

A representation satisfying the conditions of this corollary is called *good*. In other words, $\rho \in \mathfrak{R}_r(G)^s$ is good if and only if $\mathsf{Stab}_\rho/Z(G)$ is trivial. Letting $\mathfrak{R}_r(G)^{\text{good}}$ be the open subset of good representations, it easily follows that

$$\mathfrak{X}_r(G)^{\text{good}} := \mathfrak{R}_r(G)^{\text{good}}/G \subset \mathfrak{X}_r(G)^s \subset \mathfrak{X}_r(G)$$

is always a smooth manifold.

[Heusener and Porti 2004] shows that our main theorem, i.e.,

$$\mathfrak{X}_r(G)^{\text{red}} = \mathfrak{X}_r(G)^{\text{sing}},$$

is not true for all reductive Lie groups G and free groups $\mathtt{F}_r$ since for PSL_2 there are irreducible representations which are singular. The issue is that the stabilizer of an irreducible representation, modulo the center of G, may not be trivial in general. This is not an issue for GL_n or SL_n since Lemma 3.5 shows the action is free on the set of irreducibles; that is, in these cases a representation is good if and only if it is irreducible.

Let O_n be the group of $n \times n$ complex orthogonal matrices, and let Sp_{2n} be the group of $2n \times 2n$ complex symplectic matrices.

Proposition 3.32. *There exists irreducible representations $\rho : \mathtt{F}_r \to G$ for G any of O_n, PSL_n, and Sp_{2n} such that ρ is not good.*

Proof. It is sufficient in each case to find, for some n, a nonparabolic subgroup of G whose centralizer contains a noncentral element.

First consider a SL_2-representation ρ contained in the subgroup of diagonal and antidiagonal matrices (containing at least one nondiagonal element and one noncentral element). Then $\mathsf{Stab}_\rho/Z(\mathsf{SL}_2)$ is trivial, and so such a representation is irreducible. However ρ also determines an irreducible PSL_2-valued representation consisting of diagonal and antidiagonal matrices. However, its stabilizer now contains $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, since this element acts as scalar multiplication by -1 on the antidiagonal components and trivially on the diagonal components, so the action is trivial for PSL_2-representations but nontrivial for SL_2-representations. This element is not central in SL_2. Thus ρ defines an irreducible representation into PSL_2 that has finite noncentral stabilizer, and thus is not good.

For O_n representations consider any representation whose image consists of all matrices of the form

$$\left\{\begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix}\right\}.$$

One easily computes that the stabilizer is finite and not trivial and thus they are irreducible with $\mathsf{Stab}_\rho/Z(\mathsf{O}_n)$ not trivial and thus are not good.

For Sp_{2n} representations we can likewise find examples like the following for $n=2$. Let the representation have its image generated by

$$\left\{\pm\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \pm\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}, \pm\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\right\}.$$

We get a subgroup of order 16 with finite stabilizer; as such, this group is an irreducible with finite noncentral stabilizer. Again we see that $\mathsf{Stab}_\rho/Z(\mathsf{Sp}_{2n})$ is not trivial and thus this representation is not good. □

Remark 3.33. In the case of PSL_2 representations (and consequently for SL_2-valued representations) there are irreducible representations that act reducibly on $\mathfrak{g}$. However, for PSL_2 these are singular points, but for SL_2 they are smooth. This shows that Ad-reducibility does not imply nonsmoothness in general. In fact, in $\mathfrak{X}_2(\mathsf{PSL}_2)$ there are simultaneously reducibles that are smooth points and irreducibles that are singular. See [Heusener and Porti 2004].

Conjecture 3.34. Let G be a complex reductive algebraic group, and suppose $r \geq 3$. Then $\mathfrak{X}_r(G)^{\mathrm{red}} \subset \mathfrak{X}_r(G)^{\mathrm{sing}}$, and if G is semisimple equality holds if and only if G is a Cartesian product of SL_n's.

We leave the exploration of this interesting conjecture and the description of singular irreducibles to future work.

3I.2. *What if Γ is not free?* One may wonder what the relationships exist, if any, between reducible representations and singular points in $\mathfrak{X}_\Gamma(G)$ for a general finitely generated group Γ.

With a given presentation of Γ as $\Gamma = \langle x_1, \ldots, x_r \mid r_1, \ldots, r_k\rangle$ we can naturally associate the canonical epimorphism $\mathtt{F}_r \to \Gamma = F_r/\langle r_1, \ldots, r_k\rangle$ which induces the inclusion $\mathfrak{X}_\Gamma(G) \subset \mathfrak{X}_{\mathtt{F}_r}(G)$ providing $\mathfrak{X}_\Gamma(G)$ with the structure of an affine subvariety. As such, ρ is irreducible (resp. completely reducible) in $\mathfrak{X}_\Gamma(G)$ if and only if ρ is irreducible (resp. completely reducible) in $\mathfrak{X}_{\mathtt{F}_r}(G)$.

However, the notion of singularity is very far from being well behaved:

(i) If Γ is free abelian then all representations are reducible and thus the singularities cannot equal the reducibles since the singularities are a proper subset. So reducibles can be smooth; in fact this example shows all smooth points can be reducible.

(ii) The irreducibles are not generally all smooth in the representation variety let alone in the quotient variety; see [Sikora 2012, Example 38]. Such representations can project to singular points in the quotient (as one might hope is the general situation). Therefore, there can be representations in $\mathfrak{X}_\Gamma(G)^{\text{sing}} \subset \mathfrak{X}_\Gamma(G) \subset \mathfrak{X}_{\mathtt{F}_r}(G)$ which are smooth in $\mathfrak{X}_{\mathtt{F}_r}(G)$.

(iii) Singularities in the quotient do not necessarily arise from singularities in the representation space. For example, if Γ is the fundamental group of a genus 2 surface there exist representations in $\mathfrak{R}_\Gamma(\mathsf{SU}_2)$ that are singular but the quotient $\mathfrak{X}_\Gamma(\mathsf{SU}_2) \approx \mathbb{C}\mathsf{P}^3$ is smooth. See [Narasimhan and Seshadri 1965; Narasimhan and Ramanan 1969].

(iv) Lemma 3.16 and its generalizations [Sikora 2012] do not necessarily apply in general.

Therefore, when Γ is not free there is little one can say in general.

4. Local structure and classification of manifold cases

Having completed the proof of Theorem 1.1, we now move on to prove Theorem 1.2. As stated earlier, in [Bratholdt and Cooper 2001] it is established that $\mathfrak{X}_r(\mathsf{SU}_2)$ are not topological manifolds when $r \geq 4$. They compute explicit examples where the representations (abelian, nontrivial) are contained in a neighborhood homeomorphic to $\mathscr{C}(\mathbb{C}\mathsf{P}^{r-2}) \times \mathbb{R}^r$, where $\mathscr{C}(X) = (X \times [0,1))/(X \times \{0\})$ is the real open cone over a topological space X. From this characterization, simple arguments imply that $\mathfrak{X}_r(\mathsf{SU}_2)$ is not a manifold for $r \geq 4$. It is also a consequence of the following criterion, which will be useful later.

Lemma 4.1. *Let X be a manifold of dimension n and let $d \geq 0$. If $\mathscr{C}(X) \times \mathbb{R}^d$ is Euclidean (i.e, homeomorphic to $\mathbb{R}^{d+n+1}$) then X is homotopically equivalent to S^n (a sphere of dimension n). Also, if $\mathscr{C}(X) \times \mathbb{R}^d$ is half-Euclidean (i.e, homeomorphic to a closed half-space in $\mathbb{R}^{d+n+1}$) then X is homotopically equivalent to either a point or S^n.*

Proof. Let p be the cone point of $\mathscr{C}(X)$. Using the natural deformation retraction from $\mathscr{C}(X) - \{p\}$ to X, we see that

$$\mathscr{C}(X) \times \mathbb{R}^d - (\{p\} \times \mathbb{R}^d) = X \times (0,1) \times \mathbb{R}^d \simeq X,$$

where $Y \simeq X$ symbolizes Y being homotopic to X. On the other hand, if $\mathscr{C}(X) \times \mathbb{R}^d = \mathbb{R}^{n+d+1}$ then $\mathscr{C}(X) \times \mathbb{R}^d - (\{p\} \times \mathbb{R}^d) = \mathbb{R}^{n+d+1} - \mathbb{R}^d \simeq S^n$.

The other statement follows in a similar fashion if the cone point is not on the boundary of the half-space. Otherwise, $\{p\} \times \mathbb{R}^d$ is contained in the boundary so extracting it results in a contractible space. $\square$

4A. $\mathfrak{X}_r(\mathsf{SU}_n)$ *and* $\mathfrak{X}_r(\mathsf{U}_n)$. In this subsection we establish the compact cases of Theorem 1.2.

Let $K = \mathsf{SU}_n$ and let $\mathfrak{k}$ be its Lie algebra. Let $d_{r,n} = (n^2-1)(r-1) = \dim_{\mathbb{C}} \mathfrak{X}_r(G) = \dim_{\mathbb{R}} \mathfrak{X}_r(K)$. Whenever $\mathfrak{X}_r(K)$ is not a topological manifold, there exists a point $[\rho] \in \mathfrak{X}_r(K)$ and a neighborhood $\mathcal{N}$ containing $[\rho]$ that is not locally homeomorphic to $\mathbb{R}^{d_{r,n}}$, or $\mathbb{R}_+^{d_{r,n}}$ in the case of a boundary point.

We need a smooth version of Mostow's slice theorem [Mostow 1957; Bredon 1972]. Let $\mathcal{N}_x$ denote a neighborhood at x.

Lemma 4.2. *For any $[\rho] \in \mathfrak{X}_r(K)$, there is a neighborhood $\mathcal{N}_{[\rho]}$ homeomorphic to $H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho$. Moreover,*

$$T_{[\rho]}\mathfrak{X}_r(K) \cong T_0\left(H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho\right).$$

Proof. Let $\mathfrak{R}_r(K) = \mathrm{Hom}(\mathtt{F}_r, K)$. Since $\rho \in \mathfrak{R}_r(K)$ is a smooth point, $T_\rho\mathfrak{R}_r(K) \cong Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$. Moreover, $T_\rho\mathsf{Orb}_\rho \cong B^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) \subset Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$. Since Stab_ρ is compact and acts on $B^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$, there exists a Stab_ρ-invariant complement W. Thus $Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) \cong T_\rho\mathfrak{R}_r(K) \cong B^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) \oplus W$, which respects the action of the stabilizer. Since $\mathfrak{R}_r(K)$ is a smooth compact Riemannian manifold we can invariantly exponentiate W to obtain a slice $\exp(W) = S \subset \mathfrak{R}_r(K)$ such that $T_\rho S = W$. Therefore, $T_\rho S \cong H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$ as Stab_ρ-spaces.

Saturating S by K we obtain an open K-invariant space, which contains the orbit of ρ since $\rho \in S$; namely $U = K(S)$. Since U is open $T_\rho U = T_\rho\mathfrak{R}_r(K)$, and since it is saturated $U/K \cong S/\mathsf{Stab}_\rho$ is an open subset of $\mathfrak{X}_r(K)$.

Putting these observations together we conclude S is locally diffeomorphic to $T_\rho S$ which implies the neighborhood $U/K \cong H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho$, which establishes our first claim.

Then S/Stab_ρ is locally homeomorphic to $T_\rho S/\mathsf{Stab}_\rho$, which then implies

$$T_{[\rho]}(S/\mathsf{Stab}_\rho) \cong T_0(T_\rho S/\mathsf{Stab}_\rho). \tag{7}$$

But

$$T_{[\rho]}\mathfrak{X}_r(K) = T_{[\rho]}(U/K) \cong T_{[\rho]}(S/\mathsf{Stab}_\rho) \tag{8}$$

and

$$T_0(T_\rho S/\mathsf{Stab}_\rho) \cong T_0(H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho). \tag{9}$$

Equations (7), (8), and (9) together complete the proof. $\square$

Remark 4.3. The above lemma holds for all compact Lie groups K.

Theorem 4.4. *Let $r, n \geq 2$ and let $\rho \in \mathfrak{R}_r(\mathsf{SU}_n)$ be of reduced type $[n_1, n_2]$. Then, there exists a neighborhood $[\rho] \in \mathcal{N} \subset \mathfrak{X}_r(\mathsf{SU}_n)$ that is homeomorphic to $\mathbb{R}^{d_{\mathsf{SU}_n}} \times \mathscr{C}\left(\mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1}\right)$, where $d_{\mathsf{SU}_n} = (r-1)(n_1^2+n_2^2-1)+1$. Also, if $\rho \in \mathfrak{R}_r(\mathsf{U}_n)$ is of reduced type $[n_1, n_2]$, there exists a neighborhood $[\rho] \in \mathcal{N} \subset \mathfrak{X}_r(\mathsf{U}_n)$ that is homeomorphic to $\mathbb{R}^{d_{\mathsf{U}_n}} \times \mathscr{C}\left(\mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1}\right)$, where $d_{\mathsf{U}_n} = (r-1)(n_1^2+n_2^2)+2$.*

Corollary 4.5. *If $K = \mathsf{U}_n$ or $K = \mathsf{SU}_n$, both $r, n \geq 2$, and $(r, n) \neq (2, 2)$, $(2, 3)$, or $(3, 2)$, then $\mathfrak{X}_r(K)$ is not a manifold with boundary.*

Proof. Theorem 4.4 implies that $\mathfrak{X}_r(\mathsf{U}_n)$ and $\mathfrak{X}_r(\mathsf{SU}_n)$ are manifolds only if $\mathbb{R}^{d_{\mathsf{U}_n}} \times \mathscr{C}(\mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1})$ and $\mathbb{R}^{d_{\mathsf{SU}_n}} \times \mathscr{C}(\mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1})$, respectively, are locally Euclidean. By Lemma 4.1, this can only be the case if $n_1n_2(r-1)-1 \in \{0, 1\}$, with $n = n_1+n_2$ and $n_1, n_2 > 0$. In the first case, $n_1n_2(r-1) = 1$, which implies $n_1 = n_2 = 1$ and $r = 2$, so $(r, n) = (2, 2)$. From Section 2B we know $\mathfrak{X}_2(\mathsf{U}_2)$ and $\mathfrak{X}_2(\mathsf{SU}_2)$ are manifolds with boundary, and we conclude the neighborhood in this case is half-Euclidean since $\mathcal{N} = \mathbb{R}^d \times [0, 1)$, for appropriate d.

The other possibility is $n_1n_2(r-1) = 2$ so that $n_1 = 2$ and $n_2 = 1$, or $n_1 = 1$ and $n_2 = 2$, and $r = 2$. This is the case $(r, n) = (2, 3)$. Otherwise, $r = 3$ and $n_1 = n_2 = 1$, which is the case $(r, n) = (3, 2)$. Moreover, from Section 2B these two are the only cases which are manifolds.

Having exhausted all possibilities, the proof is complete. $\square$

We now prove Theorem 4.4.

Proof of Theorem 4.4. Similar to Theorem 3.21, there is a direct computational proof of Theorem 4.4. However, using Theorem 3.21, Lemma 3.19 and the relation between K and its complexification, we can provide a shorter argument.

Let τ be the Cartan involution on $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$, the Lie algebra of GL_n, which is just the linear map $A \mapsto -\bar{A}^T$, acting on a matrix $A \in \mathfrak{gl}(n, \mathbb{C})$. By definition, the fixed point subspace of τ is $\mathfrak{k}$, the Lie algebra $\mathfrak{u}_n$ of U_n. One easily checks that τ induces an involution on $Z^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho}) \cong \mathfrak{g}^r$, whose fixed subspace is $Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) \cong \mathfrak{k}^r$, and similarly $B^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho})^\tau = B^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$. This implies that τ induces an involution, also denoted τ, on the first cohomology, and that $H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})$ is naturally isomorphic to $H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho})^\tau$.

Now, assume that $\rho = \rho_1 \oplus \rho_2 \in U_{r,n} \cap \mathfrak{R}_r(\mathsf{U}_n)$, is of reduced type $[n_1, n_2]$ $(n_1, n_2 > 0,\ n_1 + n_2 = n)$. Note that ρ_1 and ρ_2 are irreducible representations in $\mathfrak{R}_r(\mathsf{U}_{n_1})$ and $\mathfrak{R}_r(\mathsf{U}_{n_2})$, respectively, and with respect to the PU_n conjugation action $\mathsf{Stab}_\rho \cong S^1$. Then a cocycle $\phi \in Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) \cong \mathfrak{k}^r$ has the form

$$\phi = \begin{pmatrix} \phi_1 & A \\ -\bar{A}^T & \phi_2 \end{pmatrix},$$

where $\phi_i \in Z^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_i}})$, and as in Theorem 3.21, A is now an arbitrary r-tuple of $n_1 \times n_2$ matrices. This shows that τ respects the decomposition in Equation (6) , so we get

$$\begin{aligned} H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho}) = H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_\rho})^\tau &= H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho_1}})^\tau \oplus H^1(\mathtt{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho_2}})^\tau \oplus W^\tau \\ &= H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_1}}) \oplus H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_2}}) \oplus F \end{aligned}$$

where, by the form of the cocycles above, we can write

$$F := W^\tau = \{(z, -\bar{z}) : z \in \mathbb{C}^{n_1 n_2 (r-1)}\};$$

using also $\dim_{\mathbb{C}} W = 2n_1 n_2 (r-1)$.

It follows from Lemma 4.2 that a neighborhood of ρ is locally homeomorphic to $H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho$. As in the proof of Theorem 3.21, the action of $\mathsf{Stab}_\rho = S^1$ does not affect $H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_i}})$, $i = 1, 2$, and we conclude that

$$\begin{aligned} H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_\rho})/\mathsf{Stab}_\rho &= H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_1}}) \oplus H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_2}}) \oplus F/S^1 \\ &\cong \mathbb{R}^{d_{\mathsf{U}_n}} \oplus \mathscr{C}(\mathbb{C}\mathsf{P}^{n_1 n_2 (r-1)-1}), \end{aligned}$$

by using Lemma 3.19. The dimension d_{U_n} is computed by:

$$d_{\mathsf{U}_n} = \sum_{i=1}^{2} \dim_{\mathbb{R}} H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\rho_i}}) = (n_1^2 + n_2^2)(r-1) + 2.$$

The case of $K = \mathsf{SU}_n$ is similar. □

Remark 4.6. In [Le Bruyn and Teranishi 1990] it is shown that the cases (2, 2), (2, 3), and (3, 2) are also the only examples which are complete intersections.

Remark 4.7. Note that using the identity representation (maximal stabilizer) results in $H^1(\mathtt{F}_r; \mathfrak{k}_{\mathrm{Ad}_{\mathrm{id}}})/\mathsf{Stab}_{\mathrm{id}} = \mathfrak{k}^r/\mathsf{SU}_n$ since the coboundaries are trivial. Removing a point results in a homological sphere quotient $S^{(n^2-1)(r-1)-1}/\mathsf{SU}_n$. If there was a Euclidean neighborhood about the identity, then this sphere quotient would be a homology sphere $S^{(n^2-1)(r-2)-1}$. We find this quite likely to give a different obstruction. At the other extreme (central stabilizer) the points are smooth and thus admit Euclidean neighborhoods.

Conjecture 4.8. If K is equal to SU_n or U_n, $[\rho] \in \mathfrak{X}_r(K)^{\mathrm{red}}$, $r, n \geq 2$, and $(r, n) \neq (2, 2), (2, 3)$ or $(3, 2)$, then there does not exists a neighborhood of $[\rho]$ that is Euclidean.

We proved this conjecture for representations of reduced type $[n_1, n_2]$ above. In fact, it seems likely that the neighborhoods around most singularities do not even admit an orbifold structure (not homeomorphic to a finite quotient of a Euclidean ball).

4B. $\mathfrak{X}_r(\mathsf{SL}_n)$ *and* $\mathfrak{X}_r(\mathsf{GL}_n)$. In this last subsection, we complete the proof of Theorem 1.2 by proving the following result.

Theorem 4.9. *Let $r, n \geq 2$ and let G be SL_n or GL_n. $\mathfrak{X}_r(G)$ is a topological manifold with boundary if and only if $(r, n) = (2, 2)$.*

Proof. By Remark 3.17, $H^1(\mathsf{F}_r; \mathfrak{g}_{\mathrm{Ad}_{\rho^{ss}}})/\!\!/\mathsf{Stab}_{\rho^{ss}}$ is an étale neighborhood; that is, an algebraic set that maps, via an étale mapping, to an open set (in the ball topology) of $\mathfrak{X}_r(G)$. Thus we see that at a reducible representation with minimal stabilizer ($\mathbb{C}^*$ for SL_n and $\mathbb{C}^* \times \mathbb{C}^*$ for GL_n), that this neighborhood is étale equivalent to $\mathbb{C}^{(n_1^2+n_2^2)(r-1)+2} \times \mathscr{C}(\mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1} \times \mathbb{C}\mathsf{P}^{(r-1)n_1n_2-1})$ in $\mathfrak{X}_r(\mathsf{GL}_n)$, where the cone here is the affine cone defined over $\mathbb{C}^*$. In $\mathfrak{X}_r(\mathsf{SL}_n)$ we have a similar neighborhood. Either way, these spaces are not locally Euclidean neighborhoods for $r, n \geq 2$ unless $n = 2 = r$ which implies that $n_1 = 1 = n_2$. This is seen by similar arguments given above in the compact cases. □

Acknowledgements

The second author thanks W. Goldman for helpful discussions, the Instituto Superior Técnico for its support during the 2007–2009 academic years, and A. Sikora for useful conversations. We also thank M. Reineke and O. Serman for pointing out relevant references.

References

[Bochnak et al. 1998] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **36**, Springer, Berlin, 1998. MR 2000a:14067 Zbl 0912.14023

[Bratholdt and Cooper 2001] S. Bratholdt and D. Cooper, "On the topology of the character variety of a free group", *Rend. Istit. Mat. Univ. Trieste* **32**:Suppl. 1 (2001), 45–53. MR 2003d:14072 Zbl 1062.14505

[Bredon 1972] G. E. Bredon, *Introduction to compact transformation groups*, Pure and Applied Mathematics **46**, Academic Press, New York, 1972. MR 54 #1265 Zbl 0246.57017

[Dolgachev 2003] I. Dolgachev, *Lectures on invariant theory*, London Mathematical Society Lecture Note Series **296**, Cambridge University Press, Cambridge, 2003. MR 2004g:14051 Zbl 1023.13006

[Drézet 2004] J.-M. Drézet, "Luna's slice theorem and applications", pp. 39–89 in *Algebraic group actions and quotients* (Wykno, 2000), edited by J. Wiśniewski, Hindawi, Cairo, 2004. MR 2006k:14082 Zbl 1109.14307

[Florentino and Lawton 2009] C. Florentino and S. Lawton, "The topology of moduli spaces of free group representations", *Math. Ann.* **345**:2 (2009), 453–489. MR 2010h:14075 Zbl 1200.14093

[Fricke and Klein 1912] R. Fricke and F. Klein, *Vorlesungen über die Theorie der automorphen Funktionen, II: Die funktionentheoretischen Ausführungen und die Andwendungen, 3: Direkte Beweismethoden der Fundamentaltheoreme und Anhang*, Teubner, Leipzig, 1912. Reprinted in Bibliotheca Mathematica Teubneriana, vol. 4, by Johnson Reprint, New York, 1965. MR 32 #1348 JFM 43.0529.08

[Goldman 1984] W. M. Goldman, "The symplectic nature of fundamental groups of surfaces", *Adv. in Math.* **54**:2 (1984), 200–225. MR 86i:32042 Zbl 0574.32032

[Goldman 1990] W. M. Goldman, "Convex real projective structures on compact surfaces", *J. Differential Geom.* **31**:3 (1990), 791–845. MR 91b:57001 Zbl 0711.53033

[Goldman 2009] W. M. Goldman, "Trace coordinates on Fricke spaces of some simple hyperbolic surfaces", pp. 611–684 in *Handbook of Teichmüller theory, II*, edited by A. Papadopoulos, IRMA Lect. Math. Theor. Phys. **13**, Eur. Math. Soc., Zürich, 2009. MR 2010j:30093 Zbl 1175.30043

[Harris 1992] J. Harris, *Algebraic geometry: a first course*, Graduate Texts in Mathematics **133**, Springer, New York, 1992. MR 93j:14001 Zbl 0779.14001

[Heusener and Porti 2004] M. Heusener and J. Porti, "The variety of characters in $\mathrm{PSL}_2(\mathbb{C})$", *Bol. Soc. Mat. Mexicana* (3) **10**:Special Issue (2004), 221–237. MR 2006m:57020 Zbl 1100.57014

[Johnson and Millson 1987] D. Johnson and J. J. Millson, "Deformation spaces associated to compact hyperbolic manifolds", pp. 48–106 in *Discrete groups in geometry and analysis* (New Haven, CT, 1986), edited by R. Howe, Progr. Math. **67**, Birkhäuser, Boston, 1987. MR 88j:22010 Zbl 0664.53023

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics **211**, Springer, New York, 2002. MR 2003e:00003 Zbl 0984.00001

[Laszlo 1996] Y. Laszlo, "Local structure of the moduli space of vector bundles over curves", *Comment. Math. Helv.* **71**:3 (1996), 373–401. MR 97j:14012 Zbl 0949.14015

[Lawton 2007] S. Lawton, "Generators, relations and symmetries in pairs of 3×3 unimodular matrices", *J. Algebra* **313**:2 (2007), 782–801. MR 2008k:16039 Zbl 1119.13004

[Le Bruyn and Procesi 1987] L. Le Bruyn and C. Procesi, "Étale local structure of matrix invariants and concomitants", pp. 143–175 in *Algebraic groups* (Utrecht, 1986), edited by A. M. Cohen et al., Lecture Notes in Math. **1271**, Springer, Berlin, 1987. MR 89b:16042 Zbl 0634.14034

[Le Bruyn and Teranishi 1990] L. Le Bruyn and Y. Teranishi, "Matrix invariants and complete intersections", *Glasgow Math. J.* **32**:2 (1990), 227–229. MR 91d:15059 Zbl 0714.16016

[Luna 1973] D. Luna, "Slices étales", pp. 81–105 in *Sur les groupes algébriques*, Bull. Soc. Math. France Mém. **33**, Soc. Math. France, Paris, 1973. MR 49 #7269 Zbl 0286.14014

[Mostow 1957] G. D. Mostow, "Equivariant embeddings in Euclidean space", *Ann. of Math.* (2) **65** (1957), 432–446. MR 19,291c Zbl 0080.16701

[Mukai 2003] S. Mukai, *An introduction to invariants and moduli*, Cambridge Studies in Advanced Mathematics **81**, Cambridge University Press, Cambridge, 2003. MR 2004g:14002 Zbl 1033.14008

[Narasimhan and Ramanan 1969] M. S. Narasimhan and S. Ramanan, "Moduli of vector bundles on a compact Riemann surface", *Ann. of Math.* (2) **89** (1969), 14–51. MR 39 #3518 Zbl 0186.54902

[Narasimhan and Seshadri 1965] M. S. Narasimhan and C. S. Seshadri, "Stable and unitary vector bundles on a compact Riemann surface", *Ann. of Math.* (2) **82** (1965), 540–567. MR 32 #1725 Zbl 0171.04803

[Procesi and Schwarz 1985] C. Procesi and G. Schwarz, "Inequalities defining orbit spaces", *Invent. Math.* **81**:3 (1985), 539–554. MR 87h:20078 Zbl 0578.14010

[Schwarz 2004] G. W. Schwarz, "Group actions and quotients for compact Lie groups and algebraic groups", pp. 209–227 in *Invariant theory in all characteristics* (Kingston, ON, 2002), edited by H. E. A. E. Campbell and D. L. Wehlau, CRM Proc. Lecture Notes **35**, Amer. Math. Soc., Providence, RI, 2004. MR 2005d:20083 Zbl 1131.14312

[Shafarevich 1994] I. R. Shafarevich, *Basic algebraic geometry, 2: Schemes and complex manifolds*, 2nd ed., Springer, Berlin, 1994. MR 95m:14002 Zbl 0797.14002

[Sikora 2012] A. S. Sikora, "Character varieties", *Trans. Amer. Math. Soc.* **364**:10 (2012), 5173–5208. MR 2931326 arXiv 0902.2589

[Sullivan 1971] D. Sullivan, "Combinatorial invariants of analytic spaces", pp. 165–168 in *Singularities — Symposium I* (Liverpool, 1969–1970), edited by C. T. C. Hall, Lecture Notes in Math. **192**, Springer, Berlin, 1971. MR 43 #4063 Zbl 0227.32005

[Vogt 1889] H. Vogt, "Sur les invariants fondamentaux des équations différentielles linéaires du second ordre", *Ann. Sci. École Norm. Sup.* (3) **6** (1889), 3–71. MR 1508833 JFM 21.0314.01

[Weil 1964] A. Weil, "Remarks on the cohomology of groups", *Ann. of Math.* (2) **80** (1964), 149–157. MR 30 #199 Zbl 0192.12802

Received May 30, 2012.

CARLOS FLORENTINO
DEPARTAMENTO DE MATEMÁTICA
INSTITUTO SUPERIOR TÉCNICO
1049-001 LISBOA
PORTUGAL
carlos.florentino@math.ist.utl.pt

SEAN LAWTON
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS-PAN AMERICAN
1201 WEST UNIVERSITY DRIVE
EDINBURG, TX 78539
UNITED STATES
lawtonsd@utpa.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.181

ENERGY IDENTITY FOR THE MAPS FROM A SURFACE WITH TENSION FIELD BOUNDED IN L^p

LI JIAYU AND ZHU XIANGRONG

Let M be a closed Riemannian surface and u_n a sequence of maps from M to Riemannian manifold N satisfying

$$\sup_n \big(\|\nabla u_n\|_{L^2(M)} + \|\tau(u_n)\|_{L^p(M)} \big) \le \Lambda$$

for some $p > 1$, where $\tau(u_n)$ is the tension field of the mapping u_n.

For a general target manifold N, if $p \ge \frac{6}{5}$, we prove the energy identity and the neckless property during blowing up.

1. Introduction

Let (M, g) be a closed Riemannian manifold and (N, h) be a Riemannian manifold without boundary. For a mapping u from M to N in $W^{1,2}(M, N)$, the energy density of u is defined by

$$e(u) = \tfrac{1}{2}|du|^2 = \mathrm{Trace}_g u^* h,$$

where $u^* h$ is the pull-back of the metric tensor h.

The energy of the mapping u is defined as

$$E(u) = \int_M e(u)\, dV,$$

where dV is the volume element of (M, g).

A map $u \in C^1(M, N)$ is called harmonic if it is a critical point of the energy E.

By the Nash embedding theorem we know that (N, h) can be isometrically into a Euclidean space $\mathbb{R}^K$ with some positive integer K. Then (N, h) may be considered as a submanifold of $\mathbb{R}^K$ with the metric induced from the Euclidean metric. Thus a map $u \in C^1(M, N)$ can be considered as a map of $C^1(M, \mathbb{R}^K)$ whose image lies in N. In this sense we can get the Euler–Lagrange equation

$$\triangle u = A(u)(du, du).$$

The research was supported by NSFC grants number 11071236, 11131007 and 11101372 and by PCSIRT.

MSC2010: 58E20.

Keywords: energy identity, tension field, neckless.

The tension field $\tau(u)$ is defined by

$$\tau(u) = \triangle_M u - A(u)(du, du),$$

where $A(u)(du, du)$ is the second fundamental form of N in $\mathbb{R}^K$. So u being harmonic means that $\tau(u) = 0$.

The harmonic mappings are of special interest when M is a Riemann surface. Consider a sequence of mappings u_n from Riemann surface M to N with bounded energies. It is clear that u_n converges weakly to u in $W^{1,2}(M, N)$ for some u in $W^{1,2}(M, N)$. But in general, it may not converge strongly in $W^{1,2}(M, N)$. When $\tau(u_n) = 0$, that is, when u_n are all harmonic, Parker [1996] proved that the lost energy is exactly the sum of some harmonic spheres, which are defined as harmonic mappings from S^2 to N. This result is called the energy identity. Also he proved that the images of these harmonic spheres and $u(M)$ are connected, that is, there is no neck during blowing up.

When $\tau(u_n)$ is bounded in L^2, the energy identity was proved in [Qing 1995] for the sphere, and in [Ding and Tian 1995] and [Wang 1996] for a general target manifold. Qing and Tian [1997] proved there is no neck during blowing up. For the heat flow of harmonic mappings, the results can also be found in [Topping 2004a; 2004b]. When the target manifold is a sphere, we proved the energy identity in [Li and Zhu 2011] for a sequence of mappings with tension fields bounded in $L \ln^+ L$, using good observations from [Lin and Wang 2002]. On the other hand, in the same paper we constructed a sequence of mappings with tension fields bounded in $L \ln^+ L$ such that there is a positive neck during blowing up. In [Zhu 2012] the neckless property during blowing up was proved for a sequence of maps u_n with

$$\lim_{\delta \to 0} \sup_n \sup_{B(x,\delta) \subset D_1} \|\tau(u_n)\|_{L \ln^+ L(B(x,\delta))} = 0.$$

In this paper we prove the energy identity and neckless property during blowing up of a sequence of maps u_n with $\tau(u_n)$ bounded in L^p for some $p \geq \frac{6}{5}$, for a general target manifold.

When $\tau(u_n)$ is bounded in L^p for some $p > 1$, the small energy regularity proved in [Ding and Tian 1995] implies that u_n converges strongly in $W^{1,2}(M, N)$ outside a finite set of points. For simplicity of exposition, it is no matter to assume that M is the unit disk $D_1 = D(0, 1)$ and there is only one singular point at 0.

In this paper we prove the following theorem.

Theorem 1. *Let $\{u_n\}$ be a sequence of mappings from D_1 to N in $W^{1,2}(D_1, N)$ with tension field $\tau(u_n)$. If*

(a) $\|u_n\|_{W^{1,2}(D_1)} + \|\tau(u_n)\|_{L^p(D_1)} \leq \Lambda$ *for some* $p \geq \frac{6}{5}$,

(b) $u_n \to u$ *strongly in* $W^{1,2}(D_1 \setminus \{0\}, \mathbb{R}^K)$ *as* $n \to \infty$, *then there exists a subsequence of* $\{u_n\}$ (*we still denote it by* $\{u_n\}$) *and some nonnegative integer* k *so that for any* $i = 1, \dots, k$, *there exist points* x_n^i, *positive numbers* r_n^i *and a nonconstant harmonic sphere* w^i (*which we view as a map from* $\mathbb{R}^2 \cup \{\infty\} \to N$) *such that*:

(1) $x_n^i \to 0, r_n^i \to 0$ *as* $n \to \infty$.

(2) $\lim_{n\to\infty}\left(\frac{r_n^i}{r_n^j} + \frac{r_n^j}{r_n^i} + \frac{|x_n^i - x_n^j|}{r_n^i + r_n^j}\right) = \infty$ *for any* $i \neq j$.

(3) w^i *is the weak limit or strong limit of* $u_n(x_n^i + r_n^i x)$ *in* $W^{1,2}_{Loc}(\mathbb{R}^2, N)$.

(4) ***Energy identity***: *We have*

$$\lim_{n\to\infty} E(u_n, D_1) = E(u, D_1) + \sum_{i=1}^{k} E(w^i). \tag{1-1}$$

(5) ***Neckless property***: *The image* $u(D_1) \cup \bigcup_{i=1}^k w^i(\mathbb{R}^2)$ *is a connected set.*

This paper is organized as follows. In Section 2 we state some basic lemmas and some standard arguments in the blow-up analysis.

In Section 3 and Section 4 we prove Theorem 1. In the proof, we use delicate analysis on the difference between normal energy and tangential energy. The energy identity is proved in Section 3 and the neckless property is proved in Section 4.

Throughout this paper, the letter C denotes a positive constant that depends only on p, Λ and the target manifold N and may vary in different places. We also don't distinguish between a sequence and one of its subsequences.

2. Some basic lemmas and standard arguments

We recall the regular theory for a mapping with small energy on the unit disk and tension field in L^p $(p > 1)$.

Lemma 2. *Let* $\bar{u}$ *be the mean value of* u *on the disk* $D_{1/2}$. *There exists a positive constant* ϵ_N *that depends only on the target manifold such that if* $E(u, D_1) \le \epsilon_N^2$ *then*

$$\|u - \bar{u}\|_{W^{2,p}(D_{1/2})} \le C\big(\|\nabla u\|_{L^2(D_1)} + \|\tau(u)\|_p\big), \tag{2-1}$$

where $p > 1$.

As a consequence of (2-1) *and the Sobolev embedding* $W^{2,p}(\mathbb{R}^2) \subset C^0(\mathbb{R}^2)$, *we have*

$$\|u\|_{Osc(D_{1/2})} = \sup_{x,y\in D_{1/2}} |u(x) - u(y)| \le C\big(\|\nabla u\|_{L^2(D_1)} + \|\tau(u)\|_p\big). \tag{2-2}$$

Remarks.

- In [Ding and Tian 1995] this lemma is proved for the mean value of u on the unit disk. Note that
$$\left| \frac{\int_{D_1} u(x)\,dx}{|D_1|} - \frac{\int_{D_{1/2}} u(x)\,dx}{|D_{1/2}|} \right| \le C\|\nabla u\|_{L^2(D_1)}.$$
So we can use the mean value of u on $D_{1/2}$ in this lemma.

- Suppose we have a sequence of mappings u_n from the unit disk D_1 to N with $\|u_n\|_{W^{1,2}(D_1)} + \|\tau(u_n)\|_{L^p(D_1)} \le \Lambda$ for some $p > 1$.

  A point $x \in D_1$ is called an energy concentration point (blow-up point) if for any r such that $D(x,r) \subset D_1$, we have
$$\sup_n E(u_n, D(x,r)) > \epsilon_N^2,$$
where ϵ_N is given in this lemma. If $x \in D_1$ isn't an energy concentration point, we can find a positive number δ such that
$$E(u_n, D(x,\delta)) \le \epsilon_N^2 \quad \text{for all } n.$$
Then it follows from Lemma 2 that we have a uniformly $W^{2,p}(D(x,\delta/2))$-bound for u_n. Because $W^{2,p}$ is compactly embedded in $W^{1,2}$, there is a subsequence of u_n (still denoted by u_n) and $u \in W^{2,p}(D(x,\delta/2))$ such that
$$\lim_{n\to\infty} u_n = u \quad \text{in } W^{1,2}(D(x,\delta/2)).$$
So u_n converges to u strongly in $W^{1,2}(D_1)$ outside a finite set of points.

Under the assumptions of our theorem, by the standard blow-up argument, that is by repeatedly rescaling u_n in a suitable way, we can obtain some nonnegative integer k so that for any $i = 1, \dots, k$, there exist a point x_n^i, a positive number r_n^i and a nonconstant harmonic sphere w^i satisfying (1), (2) and (3) of Theorem 1. By the standard induction argument in [Ding and Tian 1995] we only need to prove the theorem in the case where there is only one bubble.

In that case we can assume that w is the strong limit of the sequence $u_n(x_n+r_nx)$ in $W^{1,2}_{Loc}(\mathbb{R}^2)$. We may assume that $x_n = 0$. Set $w_n(x) = u_n(r_nx)$.

As
$$\lim_{\delta\to 0}\lim_{n\to\infty} E(u_n, D_1 \setminus D_\delta) = E(u, D_1),$$
the energy identity is equivalent to

$$\lim_{\delta\to 0}\lim_{n\to\infty}\lim_{R\to\infty} E(u_n, D_\delta \setminus D_{r_nR}) = 0. \tag{2-3}$$

To prove the sets $u(D_1)$ and $w(\mathbb{R}^2 \cup \infty)$ are connected, it is enough to show that

$$\lim_{\delta\to 0}\lim_{n\to\infty}\lim_{R\to\infty}\sup_{x,y\in D_\delta\setminus D_{r_nR}} |u_n(x) - u_n(y)| = 0. \tag{2-4}$$

3. Energy identity

In this section, we prove the energy identity for a general target manifold when $p \geq \frac{6}{5}$.

Assume that there is only one bubble w which is the strong limit of $u_n(r_n \cdot)$ in $W^{1,2}_{Loc}(\mathbb{R}^2)$. Let ϵ_N be the constant in Lemma 2. By the standard argument of blow-up analysis we can assume that, for any n,

$$E(u_n, D_{r_n}) = \sup_{\substack{r \leq r_n \\ D(x,r) \subseteq D_1}} E(u_n, D(x,r)) = \tfrac{1}{4}\epsilon_N^2. \tag{3-1}$$

Lemma 3 [Ding and Tian 1995]. *If $\tau(u_n)$ is bounded in L^p for some $p > 1$, then the tangential energy on the neck domain is zero, that is,*

$$\lim_{\delta \to 0} \lim_{R \to \infty} \lim_{n \to \infty} \int_{D_\delta \setminus D_{r_n R}} |x|^{-2} |\partial_\theta u|^2 \, dx = 0. \tag{3-2}$$

Proof. The proof is the same as in [Ding and Tian 1995], so we only sketch it.

For any $\epsilon > 0$, take δ, R such that, for any n,

$$E(u, D_{4\delta}) + E(w, \mathbb{R}^2 \setminus D_R) + \delta^{4(p-1)/p} < \epsilon^2.$$

We may suppose that $r_n R = 2^{-j_n}, \delta = 2^{-j_0}$. When n is big enough we have, for any $j_0 \leq j \leq j_n$,

$$E(u_n, D_{2^{1-j}} \setminus D_{2^{-j}}) < \epsilon^2.$$

For any j, set

$$h_n(2^{-j}) = \frac{1}{2\pi} \int_{S^1} u_n(2^{-j}, \theta) \, d\theta$$

and

$$h_n(t) = h_n(2^{-j}) + \big(h_n(2^{1-j}) - h_n(2^{-j})\big) \frac{\ln(2^j t)}{\ln 2}, \quad t \in [2^{-j}, 2^{1-j}].$$

It is easy to check that

$$\frac{d^2 h_n(t)}{dt^2} + \frac{1}{t} \frac{d h_n(t)}{dt} = 0, \quad t \in [2^{-j}, 2^{1-j}].$$

Consider $h_n(x) = h_n(|x|)$ as a map from $\mathbb{R}^2$ to $\mathbb{R}^K$, then $\triangle h_n = 0$ in $\mathbb{R}^2$. Setting $P_j = D_{2^{1-j}} \setminus D_{2^{-j}}$ we have

$$\triangle(u_n - h_n) = \triangle u_n - \triangle h_n = \triangle u_n = A(u_n) + \tau(u_n), \quad x \in P_j. \tag{3-3}$$

Taking the inner product of this equation with $u_n - h_n$ and integrating over P_j, we get that

$$\int_{P_j} |\nabla(u_n - h_n)|^2 dx = -\int_{P_j} (u_n - h_n)(A(u_n) + \tau(u_n)) \, dx + \int_{\partial P_j} (u_n - h_n)(u_n - h_n)_r \, ds.$$

Note that by definition, $h_n(2^{-j})$ is the mean value of $\{2^{-j}\}\times S^1$ and $(h_n)_r$ is independent of θ. So the integral of $(u_n-h_n)(h_n)_r$ on ∂P_j vanishes.

When $j_0<j<j_n$, by Lemma 2 we have

$$\begin{aligned}\|u_n-h_n\|_{C^0(P_j)} &\le \|u_n-h_n(2^{-j})\|_{C^0(P_j)}+\|u_n-h_n(2^{1-j})\|_{C^0(P_j)}\\ &\le 2\|u_n\|_{Osc(P_j)}\\ &\le C\big(\|\nabla u_n\|_{L^2(P_{j-1}\cup P_j\cup P_{j+1})}+2^{2(1-p)j/p}\|\tau(u_n)\|_p\big)\\ &\le C\big(\epsilon+2^{-2(p-1)j/p}\big)\\ &\le C\big(\epsilon+\delta^{2(p-1)/p}\big)\le C\epsilon.\end{aligned}$$

Summing over j for $j_0<j<j_n$ gives

$$\begin{aligned}&\int_{D_\delta\setminus D_{2r_nR}}|\nabla(u_n-h_n)|^2dx\\ &\quad=\sum_{j_0<j<j_n}\int_{P_j}|\nabla(u_n-h_n)|^2\,dx\\ &\quad\le\sum_{j_0<j<j_n}\int_{P_j}|u_n-h_n|\big(|A(u_n)|+|\tau(u_n)|\big)\,dx\\ &\qquad\qquad+\sum_{j_0<j<j_n}\int_{\partial P_j}(u_n-h_n)(u_n-h_n)_r\,ds\\ &\quad\le C\epsilon\bigg(\int_{D_{2\delta}\setminus D_{2r_nR}}\big(|\nabla u_n|^2+|\tau(u_n)|\big)\,dx+\int_{\partial D_{2\delta}\cup\partial D_{2r_nR}}|\nabla u_n|\,ds\bigg)\\ &\quad\le C\epsilon\bigg(\int_{D_{2\delta}\setminus D_{2r_nR}}|\nabla u_n|^2\,dx+\delta^{2(p-1)/p}+\epsilon\bigg)\le C\epsilon.\end{aligned}\tag{3-4}$$

Here we use the inequality

$$\int_{\partial D_{2\delta}\cup\partial D_{2r_nR}}|\nabla u_n|\,ds\le C\epsilon,$$

which can be derived from the Sobolev trace embedding theorem.

As $h_n(x)$ is independent of θ, it can be shown that

$$\int_{D_{2\delta}\setminus D_{2r_nR}}|x|^{-2}|\partial_\theta u_n|^2\,dx\le\int_{D_{2\delta}\setminus D_{2r_nR}}|\nabla(u_n-h_n)|^2\,dx\le C\epsilon,$$

so this lemma is proved. □

It is left to show that the normal energy on the neck domain also equals to zero. We need the following equality.

Lemma 4 (Pohozaev equality [Lin and Wang 1998, Lemma 2.4, page 374]). *Let u be a solution to*

$$\Delta u+A(u)(du,du)=\tau(u).$$

Then

$$\int_{\partial D_t}\big(|\partial_r u|^2 - r^{-2}|\partial_\theta u|^2\big)\, ds = \frac{2}{t}\int_{D_t} \tau\cdot(x\nabla u)\, dx. \tag{3-5}$$

As a direct corollary, by integrating over $[0,\delta]$, *we have*

$$\int_{D_\delta}\big(|\partial_r u|^2 - r^{-2}|\partial_\theta u|^2\big)\, dx = \int_0^\delta \frac{2}{t}\int_{D_t} \tau\cdot(x\nabla u)\, dx dt. \tag{3-6}$$

Proof. Multiplying both sides of the equation by $x\nabla u$ and integrating over D_t, we get

$$\int_{D_t} |\nabla u|^2\, dx - t\int_{\partial D_t} |\partial_r u|^2\, ds + \frac{1}{2}\int_{D_t} x\nabla|\nabla u|^2\, dx = -\int_{D_t}\tau\cdot(x\nabla u)\, dx.$$

Note that

$$\frac{1}{2}\int_{D_t} x\nabla|\nabla u|^2\, dx = -\int_{D_t}|\nabla u|^2\, dx + \frac{t}{2}\int_{\partial D_t}|\nabla u|^2\, ds.$$

Hence,

$$\int_{\partial D_t}\big(|\partial_r u|^2 - \tfrac{1}{2}|\nabla u|^2\big)\, ds = \frac{1}{t}\int_{D_t}\tau\cdot(x\nabla u)\, dx.$$

As $|\nabla u|^2 = |\partial_r u|^2 + r^{-2}|\partial_\theta u|^2$, we have proved this lemma. □

Now we use this equality to estimate the normal energy on the neck domain. We prove the following lemma.

Lemma 5. *If* $\tau(u_n)$ *is bounded in* L^p *for some* $p \ge \frac{6}{5}$, *then for* δ *small enough we have*

$$\left|\int_{D_\delta}\big(|\partial_r u_n|^2 - |x|^{-2}|\partial_\theta u|^2\big)\, dx\right| \le C\delta^{(p-1)/p},$$

where C *depends on* p, Λ, *the target manifold* N *and the bubble* w.

Proof. Take $\psi \in C_0^\infty(D_2)$ satisfying $\psi = 1$ in D_1, then

$$\triangle(\psi u_n) = \psi A(u_n)(du_n, du_n) + \psi\tau_n + 2\nabla\psi\nabla u_n + u_n\triangle\psi.$$

Set $g_n = \psi A(u_n)(du_n, du_n) + \psi\tau_n + 2\nabla\psi\nabla u_n + u_n\triangle\psi$. When $|x| < 1$,

$$\partial_i u_n(x) = R_i * g_n(x) = \int \frac{x_i - y_i}{|x-y|^2} g_n(y)\, dy.$$

Let Φ_n be the Newtonian potential of $\psi\tau_n$, then $\triangle\Phi_n = \psi\tau_n$. The corresponding Pohozaev equality is

$$\int_{D_\delta}\big(|\partial_r\Phi_n|^2 - r^{-2}|\partial_\theta\Phi_n|^2\big)\, dx = \int_0^\delta \frac{2}{t}\int_{D_t}\psi\tau_n\cdot(x\nabla\Phi_n)\, dx dt. \tag{3-7}$$

Here

$$\partial_i \Phi_n(x) = R_i * (\psi \tau_n)(x) = \int \frac{x_i - y_i}{|x-y|^2} (\psi \tau_n)(y)\, dy.$$

As τ_n is bounded in L^p $(p > 1)$, we have

$$\int_{D_\delta} |\nabla \Phi_n|^2\, dx \le C\delta^{4(p-1)/p} \|\nabla \Phi_n\|^2_{2p/(2-p)} \le C\delta^{4(p-1)/p} \|\tau_n\|_p^2 \le C\delta^{4(p-1)/p}.$$

By (3-7), it can be shown that for any $\delta > 0$,

$$\left| \int_0^\delta \frac{1}{t} \int_{D_t} \psi \tau_n \cdot (x \nabla \Phi_n)\, dx dt \right| \le \int_{D_\delta} |\nabla \Phi_n|^2\, dx \le C\delta^{4(p-1)/p}. \tag{3-8}$$

For δ small enough, we have

$$\begin{aligned} &\left| \int_{D_\delta} \left(|\partial_r u_n|^2 - r^{-2} |\partial_\theta u_n|^2 \right) dx \right| \\ &= \left| \int_0^\delta \frac{2}{t} \int_{D_t} \tau_n \cdot (x \nabla u_n)\, dx dt \right| \\ &\le 2 \left| \int_0^\delta \frac{1}{t} \int_{D_t} \tau_n \cdot (x \nabla \Phi_n)\, dx dt \right| + 2 \int_0^\delta \frac{1}{t} \int_{D_t} |x \tau_n|\, |\nabla (u_n - \Phi_n)(x)|\, dx dt \\ &\le C\delta^{4(p-1)/p} + 2 \int_{D_\delta} |x \tau_n| \left| \nabla (u_n - \Phi_n)(x) \right| \left(\int_{|x|}^\delta \frac{1}{t} dt \right) dx \\ &\le C\delta^{4(p-1)/p} + 2 \int_{D_\delta} |\tau_n|\, |\nabla (u_n - \Phi_n)(x)|\, |x| \ln \frac{1}{|x|}\, dx. \end{aligned} \tag{3-9}$$

For any $j > 0$, set $\varphi_j(x) = \psi\left(\frac{x}{2^{2-j}\delta}\right) - \psi\left(\frac{x}{2^{-2-j}\delta}\right)$. When $2^{-j}\delta \le |x| < 2^{1-j}\delta$, we obtain

$$\begin{aligned} |\partial_i (u_n - \Phi_n)(x)| &= \left| \int \frac{x_i - y_i}{|x-y|^2} (g_n(y) - \psi \tau_n(y))\, dy \right| \\ &\le \int \frac{|\psi A(u_n)(du_n, du_n) + 2 \nabla \psi \nabla u_n + u_n \triangle \psi|(y)}{|x-y|} dy \\ &\le \int \frac{|\psi A(u_n)(y)|}{|x-y|} dy + C \int_{1<|y|<2} (|\nabla u_n| + |u_n|)(y)\, dy \\ &\le \int \frac{|\varphi_j A(u_n)(y)|}{|x-y|} dy + \int \frac{|(\psi - \varphi_j) A(u_n)(y)|}{|x-y|} dy + C \\ &\le \int \frac{|\varphi_j A(u_n)(y)|}{|x-y|} dy + \frac{\int |A(u_n)(y)|\, dy}{|x|} + C \\ &\le \int \frac{|\varphi_j A(u_n)(y)|}{|x-y|} dy + \frac{C}{|x|}. \end{aligned} \tag{3-10}$$

When $\delta > 0$ is small enough and n is big enough, for any $j > 0$, we claim that

$$\|\varphi_j A(u_n)\|_{p/(2-p)} \le C(2^{-j}\delta)^{-4(p-1)/p}, \tag{3-11}$$

where the constant C depends only on p, Λ, the bubble w and the target manifold N.

Take $\delta > 0$ and $R(w)$ that depends on w such that

$$E(u, D_{8\delta}) \le \tfrac{1}{8}\epsilon_N^2 \quad \text{and} \quad E(w, \mathbb{R}^2 \setminus D_{R(w)}) \le \tfrac{1}{8}\epsilon_N^2.$$

The standard blow-up analysis (see [Ding and Tian 1995]) shows that for any j with $8r_n R(w) \le 2^{-j}\delta$ and n big enough, we have

$$E(u_n, D_{2^{4-j}\delta} \setminus D_{2^{-3-j}\delta}) \le \tfrac{1}{3}\epsilon_N^2.$$

By (3-1), when $2^{-j}\delta < r_n/16$, we get

$$E(u_n, D_{2^{4-j}\delta} \setminus D_{2^{-3-j}\delta}) \le \tfrac{1}{4}\epsilon_N^2.$$

So when $2^{-j}\delta < r_n/16$ or $2^{-j}\delta \ge 8r_n R(w)$, by Lemma 2, we see that

$$\begin{aligned}
\|\varphi_j A(u_n)\|_{p/(2-p)} &\le C\|\nabla u_n\|^2_{L^{2p/(2-p)}(D_{2^{3-j}\delta}\setminus D_{2^{-2-j}\delta})} \\
&\le C\|u_n - \bar{u}_{n,j}\|^2_{W^{2,p}(D_{2^{3-j}\delta}\setminus D_{2^{-2-j}\delta})} \\
&\le C\big[(2^{-j}\delta)^{-4\frac{p-1}{p}}\|\nabla u_n\|^2_{L^2(D_{2^{4-j}\delta}\setminus D_{2^{-4-j}\delta})} + \|\tau(u_n)\|_p^2\big] \\
&\le C(2^{-j}\delta)^{-4\frac{p-1}{p}},
\end{aligned}$$

where $\bar{u}_{n,j}$ is the mean of u_n on $D_{2^{3-j}\delta} \setminus D_{2^{-2-j}\delta}$.

On the other hand, when $r_n/16 \le 2^{-j}\delta \le 8r_n R(w)$, we can find no more than $CR(w)^2$ balls with radius $r_n/2$ to cover $D_{2^{3-j}\delta} \setminus D_{2^{-2-j}\delta}$, that is,

$$D_{2^{3-j}\delta} \setminus D_{2^{-2-j}\delta} \subset \bigcup_{i=1}^{m} D(y_i, \tfrac{1}{2}r_n).$$

Set $B_i = D(y_i, \frac{1}{2}r_n)$ and $2B_i = D(y_i, r_n)$. By (3-1), for any i with $i \le m$ we have

$$E(u_n, 2B_i) \le \tfrac{1}{4}\epsilon_N^2.$$

Using Lemma 2 we obtain

$$\begin{aligned}
\|\varphi_j A(u_n)\|_{p/(2-p)} &\le C\|\nabla u_n\|^2_{L^{2p/(2-p)}(D_{2^{3-j}\delta}\setminus D_{2^{-2-j}\delta})} \\
&\le C\bigg(\sum_{i=1}^{m}\|\nabla u_n\|^{2p/(2-p)}_{L^{2p/(2-p)}(B_i)}\bigg)^{(2-p)/p} \\
&\le C\sum_{i=1}^{m}\|\nabla u_n\|^2_{L^{2p/(2-p)}(B_i)}
\end{aligned}$$

$$
\begin{aligned}
&\le C\sum_{i=1}^{m}\|u_n-\bar{u}_{n,i}\|^2_{W^{2,p}(B_i)}\\
&\le C\sum_{i=1}^{m}\big((r_n)^{-4(p-1)/p}\|\nabla u_n\|^2_{L^2(2B_i)}+\|\tau(u_n)\|_p^2\big)\\
&\le Cm\big((2^{-j}\delta)^{-4(p-1)/p}+1\big)\\
&\le C(2^{-j}\delta)^{-4(p-1)/p},
\end{aligned}
$$

where $\bar{u}_{n,i}$ is the mean of u_n over B_i and the constant C depends only on p, Λ, the bubble w and the target manifold N. So we have proved (3-11).

By (3-10) and (3-11), when $p>1$ we get

$$
\begin{aligned}
\text{(3-12)}\quad &\int_{D_\delta}|\tau_n|\,|\nabla(u_n-\Phi_n)(x)|\,|x|\ln\frac{1}{|x|}\,dx\\
&\le\sum_{j=1}^{\infty}\int_{2^{-j}\delta<|x|<2^{1-j}\delta}|\tau_n|\,|\nabla(u_n-\Phi_n)(x)|\,|x|\ln\frac{1}{|x|}\,dx\\
&\le C\sum_{j=1}^{\infty}\int_{2^{-j}\delta<|x|<2^{1-j}\delta}|\tau_n|\bigg(\frac{1}{|x|}+\int\frac{|\varphi_jA(u_n)(y)|}{|x-y|}dy\bigg)|x|\ln\frac{1}{|x|}\,dx\\
&\le C\bigg(\int_{D_\delta}|\tau_n|\ln\frac{1}{|x|}\,dx\\
&\qquad\qquad+\sum_{j=1}^{\infty}\int_{2^{-j}\delta<|x|<2^{1-j}\delta}|\tau_n|\bigg(\int\frac{|\varphi_jA(u_n)(y)|}{|x-y|}dy\bigg)|x|\ln\frac{1}{|x|}\,dx\bigg)\\
&\le C\bigg(\Big\|\ln\frac{1}{|\cdot|}\Big\|_{L^{p/(p-1)}(D_\delta)}+\sum_{j=1}^{\infty}2^{-j}\delta\ln\frac{2^j}{\delta}\Big\|\int\frac{|\varphi_jA(u_n)(y)|}{|\cdot-y|}dy\Big\|_{\frac{p}{p-1}}\bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\times\|\tau_n\|_p\\
&\le C\bigg(\delta^2\Big(\ln\frac{1}{\delta}\Big)^{1/(p-1)}+\sum_{j=1}^{\infty}2^{-j}\delta\ln\frac{2^j}{\delta}\|\varphi_jA(u_n)\|_{2p/(3p-2)}\bigg).
\end{aligned}
$$

Here we use the fact that the fraction integral operator $I(f)=\frac{1}{|\cdot|}*f$ is bounded from $L^q(\mathbb{R}^2)$ to $L^{2q/(2-q)}(\mathbb{R}^2)$ for $1<q<2$.

When $p\ge\frac{6}{5}$, that is, when $2p/(3p-2)\le p/(2-p)$, by (3-11) we have

$$
\begin{aligned}
\text{(3-13)}\qquad \|\varphi_jA(u_n)\|_{\frac{2p}{3p-2}}&\le C(2^{-j}\delta)^{\frac{5p-6}{p}}\|\varphi_jA(u_n)\|_{\frac{p}{2-p}}\\
&\le C(2^{-j}\delta)^{\frac{5p-6}{p}-\frac{4(p-1)}{p}}\le C(2^{-j}\delta)^{-\frac{2-p}{p}}.
\end{aligned}
$$

From (3-12) and (3-13) we get

$$
\begin{aligned}
(3\text{-}14)\quad \int_{D_\delta} |\tau_n|\,|\nabla(u_n-\Phi_n)(x)|\,|x|\ln\frac{1}{|x|}\,dx
&\le C\Bigl(\delta^2\Bigl(\ln\frac{1}{\delta}\Bigr)^{\frac{1}{p-1}}+\sum_{j=1}^{\infty}2^{-j}\delta\ln\frac{2^j}{\delta}\|\varphi_j A(u_n)\|_{\frac{2p}{3p-2}}\Bigr)\\
&\le C\Bigl(\delta+\sum_{j=1}^{\infty}2^{-j}\delta\ln\frac{2^j}{\delta}(2^{-j}\delta)^{-\frac{2-p}{p}}\Bigr)\\
&\le C\Bigl(\delta+\delta^{\frac{2(p-1)}{p}}\ln\frac{1}{\delta}\Bigr)\le C\delta^{\frac{p-1}{p}}.
\end{aligned}
$$

It is clear that (3-9) and (3-14) imply that

$$
(3\text{-}15)\qquad \Bigl|\int_{D_\delta}\bigl(|\partial_r u_n|^2-r^{-2}|\partial_\theta u_n|^2\bigr)\,dx\Bigr|\le C\delta^{(p-1)/p}.
$$

This concludes the proof. □

Now we use these lemmas to prove the energy identity. Note that w is harmonic. From Lemma 4 we see that $\int_{D_R}(|\partial_r w|^2-r^{-2}|\partial_\theta w|^2)dx=0$ for any $R>0$. It is easy to see that

$$
\begin{aligned}
\lim_{R\to\infty}\lim_{n\to\infty}\Bigl|\int_{D_{r_nR}}(|\partial_r u_n|^2-r^{-2}|\partial_\theta u_n|^2)dx\Bigr| &= \lim_{R\to\infty}\Bigl|\int_{D_R}(|\partial_r w|^2-r^{-2}|\partial_\theta w|^2)dx\Bigr|\\
&=0.
\end{aligned}
$$

Letting $\delta\to 0$ in (3-15), we obtain

$$
\begin{aligned}
(3\text{-}16)\quad \lim_{\delta\to0}\lim_{R\to\infty}\lim_{n\to\infty}\Bigl|\int_{D_\delta\setminus D_{r_nR}}\bigl(|\partial_r u_n|^2-r^{-2}|\partial_\theta u_n|^2\bigr)\,dx\Bigr|
&\le \lim_{\delta\to0}\lim_{n\to\infty}\Bigl|\int_{D_\delta}\bigl(|\partial_r u_n|^2-r^{-2}|\partial_\theta u_n|^2\bigr)\,dx\Bigr|\\
&\quad+\lim_{R\to\infty}\lim_{n\to\infty}\Bigl|\int_{D_{r_nR}}\bigl(|\partial_r u_n|^2-r^{-2}|\partial_\theta u_n|^2\bigr)\,dx\Bigr|\\
&=0.
\end{aligned}
$$

Using Lemma 3 we obtain that the normal energy also vanishes on the neck domain, so the energy identity is proved.

4. Neckless property

In this section we use the method in [Qing and Tian 1997] to prove the neckless property during blowing up.

For any $\epsilon > 0$, take δ, R such that

$$E(u, D_{4\delta}) + E(w, \mathbb{R}^2 \setminus D_R) + \delta^{4(p-1)/p} < \epsilon^2.$$

Suppose $r_n R = 2^{-j_n}, \delta = 2^{-j_0}$. When n is big enough, the standard blow-up analysis shows that for any $j_0 \le j \le j_n$,

$$E(u_n, D_{2^{1-j}} \setminus D_{2^{-j}}) < \epsilon^2.$$

For any $j_0 < j < j_n$, set $L_j = \min\{j - j_0, j_n - j\}$. Now we estimate the norm $\|\nabla u_n\|_{L^2(P_j)}$. Set $P_{j,t} = D_{2^{t-j}} \setminus D_{2^{-t-j}}$ and take $h_{n,j,t}$ similar to h_n in the last section, but

$$h_{n,j,t}(2^{\pm t-j}) = \frac{1}{2\pi} \int_{S^1} u_n(2^{\pm t-j}, \theta)\, d\theta.$$

By an argument similar to the one used in deriving (3-4), we have, for $0 < t \le L_j$,

$$\begin{aligned} &\int_{P_{j,t}} r^{-2} |\partial_\theta u_n|^2\, dx \\ &\qquad \le C\epsilon \left(\int_{P_{j,t}} |\nabla u_n|^2\, dx + (2^{t-j})^{\frac{2(p-1)}{p}} \right) + \int_{\partial P_{j,t}} |u_n - h_{n,j,t}|\, |\nabla u_n|\, ds. \end{aligned} \tag{4-1}$$

Set $f_j(t) = \int_{P_{j,t}} |\nabla u_n|^2 dx$, a simple computation shows that

$$f_j'(t) = \ln 2 \left(2^{t-j} \int_{\{2^{t-j}\}\times S^1} |\nabla u_n|^2\, ds + 2^{-t-j} \int_{\{2^{-t-j}\}\times S^1} |\nabla u_n|^2\, ds \right).$$

Combining that $h_{n,j,t}$ is independent of θ and $h_{n,j,t}$ is the mean value of u_n at the two components of $\partial P_{j,t}$ with the Poincaré inequality yields that

$$\begin{aligned} &\int_{\partial P_{j,t}} |u_n - h_{n,j,t}|\, |\nabla u_n|\, ds \\ &\qquad = \int_{\{2^{t-j}\}\times S^1} |u_n - h_{n,j,t}|\, |\nabla u_n|\, ds + \int_{\{2^{-t-j}\}\times S^1} |u_n - h_{n,j,t}|\, |\nabla u_n|\, ds \\ &\qquad \le \left(\int_{\{2^{t-j}\}\times S^1} |u_n - h_{n,j,t}|^2\, ds \right)^{\frac12} \left(\int_{\{2^{t-j}\}\times S^1} |\nabla u_n|^2\, ds \right)^{\frac12} \\ &\qquad\qquad + \left(\int_{\{2^{-t-j}\}\times S^1} |u_n - h_{n,j,t}|^2\, ds \right)^{\frac12} \left(\int_{\{2^{-t-j}\}\times S^1} |\nabla u_n|^2\, ds \right)^{\frac12} \\ &\qquad \le C \left(2^{t-j} \int_{\{2^{t-j}\}\times S^1} |\nabla u_n|^2\, ds + 2^{-t-j} \int_{\{2^{-t-j}\}\times S^1} |\nabla u_n|^2\, ds \right) \\ &\qquad \le C f_j'(t). \end{aligned}$$

On the other hand, by a similar argument as we made to obtain (3-15), we get

$$\left|\int_{P_{j,t}} \left(|\partial_r u_n|^2 - r^{-2}|\partial_\theta u_n|^2\right) dx\right| \le C\left((2^{t-j})^{\frac{p-1}{p}} + (2^{-t-j})^{\frac{p-1}{p}}\right) \le C(2^{t-j})^{\frac{p-1}{p}}. \tag{4-2}$$

Since $|\nabla u|^2 = |\partial_r u|^2 + r^{-2}|\partial_\theta u|^2 = 2r^{-2}|\partial_\theta u|^2 + (|\partial_r u|^2 - r^{-2}|\partial_\theta u|^2)$, by (4-1) and (4-2) we have

$$\begin{aligned} f_j(t) &\le 2\int_{P_{j,t}} r^{-2}|\partial_\theta u_n|\,dx + \left|\int_{P_{j,t}} \left(|\partial_r u_n|^2 - r^{-2}|\partial_\theta u_n|^2\right) dx\right| \\ &\le C\epsilon\left(f_j(t) + (2^{t-j})^{\frac{2(p-1)}{p}}\right) + Cf_j'(t) + C(2^{t-j})^{\frac{p-1}{p}} \\ &\le C\left(\epsilon f_j(t) + 2^{-\frac{(p-1)j}{p}} 2^{\frac{(p-1)t}{p}} + f_j'(t)\right). \end{aligned}$$

Take ϵ small enough and set $\epsilon_p = \frac{p-1}{p}\ln 2$, then for some positive constant C big enough we get

$$f_j'(t) - \frac{1}{C} f_j(t) + Ce^{-\epsilon_p j} e^{\epsilon_p t} \ge 0.$$

We may assume that $\epsilon_p > 1/C$, then we have

$$(e^{-t/C} f_j(t))' + Ce^{-\epsilon_p j} e^{(\epsilon_p - 1/C)t} \ge 0.$$

Integrating this inequality over $[2, L_j]$ gives

$$\begin{aligned} f_j(2) &\le C\left(e^{-L_j/C} f_j(L_j) + e^{-\epsilon_p j}\int_1^{L_j} e^{(\epsilon_p - 1/C)t}\,dt\right) \\ &\le C\left(e^{-L_j/C} f_j(L_j) + e^{-\epsilon_p j} e^{(\epsilon_p - 1/C)L_j}\right). \end{aligned}$$

Note that $j \ge L_j$, so

$$f_j(2) \le C(e^{-L_j/C} f_j(L_j) + e^{-j/C}).$$

Since the energy identity was proved in the last section, we can take δ small such that the energy on the neck domain is less than ϵ^2, which implies that $f_j(L_j) < \epsilon^2$. So we get

$$f_j(2) \le C\left(e^{-L_j/C}\epsilon^2 + e^{-j/C}\right).$$

Using Lemma 2 on the domain $P_j = D_{2^{1-j}} \setminus D_{2^{-j}}$ when $j < j_n$, we obtain

$$\begin{aligned} \|u_n\|_{Osc(P_j)} &\le C\left(\|\nabla u_n\|_{L^2(P_{j-1}\cup P_j \cup P_{j+1})} + 2^{\frac{2(1-p)j}{p}}\|\tau(u_n)\|_p\right) \\ &\le C\left(f_j(2) + e^{-2\epsilon_p j}\right). \end{aligned}$$

Summing over j from j_0 to j_n yields

$$\begin{aligned}
\|u_n\|_{Osc(D_\delta\setminus D_{2r_nR})} &\le \sum_{j=j_0}^{j_n}\|u_n\|_{Osc(P_j)}\\
&\le C\sum_{j=j_0}^{j_n}\bigl(f_j(2)+e^{-2\epsilon_p j}\bigr)\\
&\le C\sum_{j=j_0}^{j_n}\bigl(e^{-L_j/C}\epsilon^2+e^{-j/C}+e^{-2\epsilon_p j}\bigr)\\
&\le C\Biggl(\sum_{i=0}^{\infty}e^{-i/C}\epsilon^2+\sum_{j=j_0}^{\infty}e^{-j/C}\Biggr)\\
&\le C\bigl(\epsilon^2+e^{-j_0/C}\bigr)\le C\bigl(\epsilon^2+\delta^{1/C}\bigr).
\end{aligned}$$

Here we used the assumption that $\epsilon_p > 1/C$. So we have proved that there is no neck during the blowing up.

References

[Ding and Tian 1995] W. Ding and G. Tian, "Energy identity for a class of approximate harmonic maps from surfaces", *Comm. Anal. Geom.* **3**:3-4 (1995), 543–554. MR 97e:58055 Zbl 0855.58016

[Li and Zhu 2011] J. Li and X. Zhu, "Small energy compactness for approximate harmonic mappings", *Commun. Contemp. Math.* **13**:5 (2011), 741–763. MR 2847227 Zbl 1245.58008

[Lin and Wang 1998] F. Lin and C. Wang, "Energy identity of harmonic map flows from surfaces at finite singular time", *Calc. Var. Partial Differential Equations* **6**:4 (1998), 369–380. MR 99k:58047 Zbl 0908.58008

[Lin and Wang 2002] F. Lin and C. Wang, "Harmonic and quasi-harmonic spheres, II", *Comm. Anal. Geom.* **10**:2 (2002), 341–375. MR 2003d:58029 Zbl 1042.58005

[Parker 1996] T. H. Parker, "Bubble tree convergence for harmonic maps", *J. Differential Geom.* **44**:3 (1996), 595–633. MR 98k:58069 Zbl 0874.58012

[Qing 1995] J. Qing, "On singularities of the heat flow for harmonic maps from surfaces into spheres", *Comm. Anal. Geom.* **3**:1-2 (1995), 297–315. MR 97c:58154 Zbl 0868.58021

[Qing and Tian 1997] J. Qing and G. Tian, "Bubbling of the heat flows for harmonic maps from surfaces", *Comm. Pure Appl. Math.* **50**:4 (1997), 295–310. MR 98k:58070 Zbl 0879.58017

[Topping 2004a] P. Topping, "Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow", *Ann. of Math.* (2) **159**:2 (2004), 465–534. MR 2005g:58029 Zbl 1065.58007

[Topping 2004b] P. Topping, "Winding behaviour of finite-time singularities of the harmonic map heat flow", *Math. Z.* **247**:2 (2004), 279–302. MR 2004m:53120 Zbl 1067.53055

[Wang 1996] C. Wang, "Bubble phenomena of certain Palais-Smale sequences from surfaces to general targets", *Houston J. Math.* **22**:3 (1996), 559–590. MR 98h:58053 Zbl 0879.58019

[Zhu 2012] X. Zhu, "No neck for approximate harmonic maps to the sphere", *Nonlinear Anal.* **75**:11 (2012), 4339–4345. MR 2921993 Zbl 1243.58011

Received October 1, 2011. Revised May 3, 2012.

LI JIAYU
WU WENJUN KEY LABORATORY
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
96 JINZHAI ROAD
HEFEI, 230026
CHINA

and

AMSS, CHINESE ACADEMY OF SCIENCES
BEIJING, 100190
CHINA

jiayuli@ustc.edu.cn

ZHU XIANGRONG
DEPARTMENT OF MATHEMATICS
ZHEJIANG NORMAL UNIVERSITY
688 YINGBIN ROAD
JINHUA, 321004
CHINA

zxr@zjnu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.197

REMARKS ON SOME ISOPERIMETRIC PROPERTIES OF THE $k-1$ FLOW

YU-CHU LIN AND DONG-HO TSAI

We consider the evolution of a convex closed plane curve γ_0 along its inward normal direction with speed $k-1$, where k is the curvature. This flow has the feature that it is the gradient flow of the (length − area) functional and has been previously studied by Chou and Zhu, and Yagisita. We revisit the flow and point out some interesting isoperimetric properties not discussed before.

We first prove that if the curve γ_t converges to the unit circle S^1 (without rescaling), its length $L(t)$ and area $A(t)$ must satisfy certain monotonicity properties and inequalities.

On the other hand, if the curve γ_t (assume γ_0 is not a circle) expands to infinity as $t\to\infty$ and we interpret Yagisita's result in the right way, the isoperimetric difference $L^2(t)-4\pi A(t)$ of γ_t will decrease to a *positive* constant as $t\to\infty$. Hence, without rescaling, the expanding curve γ_t will not become circular. It is asymptotically close to some expanding curve C_t, where C_0 is not a circle and each C_t is *parallel* to C_0. The asymptotic speed of C_t is given by the constant 1.

1. Introduction

Let γ_0 be a smooth embedded convex closed curve in $\mathbb{R}^2$ (with positive curvature everywhere) parametrized by $X_0 := X_0(\varphi) : S^1 \to \mathbb{R}^2$, where S^1 is the unit circle. We study the geometric behavior of γ_0 driven by the equation

$$\frac{\partial X}{\partial t}(\varphi,t) = (k(\varphi,t)-1)N_{\text{in}}(\varphi,t),\quad X(\varphi,0)=X_0(\varphi),\quad \varphi\in S^1 \tag{1}$$

where $k(\varphi,t)$ is the curvature of the curve γ_t (parametrized by $X(\varphi,t)$), $N_{\text{in}}(\varphi,t)$ is the unit inward normal vector of γ_t.

Without the constant term, (1) is the well-known *curve shortening flow*. See [Gage and Hamilton 1986] for the case when γ_0 is convex and [Grayson 1987] for the case when γ_0 is a simple closed curve. Also see [Andrews 1998] for more

Both authors are supported by NSC and NCTS of Taiwan.
MSC2010: 51K05, 52A40.
Keywords: curve shortening flow, isoperimetric inequality, plane curve evolution.

general types of curvature flows. Unlike the situation in [Grayson 1987], a simple closed curve γ_0 may develop self-intersections under the flow (1) due to the constant term -1. This will make the problem hard to manage. Thus we focus only on the case when γ_0 is convex.

According to [Gage and Hamilton 1986], (1) is a parabolic flow and there exists a unique smooth solution $X(\varphi, t) : S^1 \times [0, T) \to \mathbb{R}^2$ to the flow for a short time $T > 0$, $T < \infty$. We want to study its long-time convergence behavior. The flow (1) has the interesting property that it is the *gradient flow* of the functional

$$E(\gamma) = \text{ length} - \text{area} = \int_\gamma ds - \tfrac{1}{2}\int_\gamma (x\,dy - y\,dx) \tag{2}$$

with respect to the L^2 inner product $\langle u, v\rangle = \int_\gamma uv\,ds$ on the space of all normal variations of γ. One can also view it as a competition between the curve shortening flow $\partial X/\partial t = k\boldsymbol{N}_{\text{in}}$ (contraction) and the unit-speed outward normal flow $\partial X/\partial t = -\boldsymbol{N}_{\text{in}} = \boldsymbol{N}_{out}$ (expansion). See [Gage and Hamilton 1986; Green and Osher 1999].

It is expected intuitively that, depending on the convex initial curve, the flow (1) will either converge to a point, converge to a round circle S^1, or expand to infinity, with each γ_t remaining smooth and convex. This is indeed true due to Theorem 3.12 (see also Remark 3.14) of [Chou and Zhu 2001]. Moreover, for given initial data $X_0(\varphi) : S^1 \to \mathbb{R}^2$, if we consider its homothetic class

$$\mathcal{H}(X_0) = \{\lambda X_0(\varphi) : \lambda > 0,\ \varphi \in S^1\}, \tag{3}$$

there exists a unique number $\Lambda > 0$ (for convenience we call it the *critical number* of X_0) such that under the flow (1) with initial data $\lambda X_0(\varphi)$, $\lambda = \Lambda$, γ_t will converge to the unit circle S^1 (without rescaling) smoothly as $t \to \infty$. For $0 < \lambda < \Lambda$, the flow exists on a maximal finite time interval $[0, T_{\max})$, $T_{\max} < \infty$, and γ_t converges to a point $p \in \mathbb{R}^2$ as $t \to T_{\max}$; and for $\lambda > \Lambda$, the flow expands to infinity as $t \to \infty$. Thus the generic behavior of the $k - 1$ flow is either converging to a point or expanding to infinity.

The asymptotic behavior of γ_t as $t \to T_{\max}$ (or $t \to \infty$) in the above three cases are known due to [Chou and Zhu 2001, Theorem 3.12; Gage 1984; Gage and Hamilton 1986; Chow and Tsai 1996]. Also see [Yagisita 2005] for a more refined estimate in the expanding case.

Our purpose is to give an estimate of the number Λ and to point out some monotonicity properties of length $L(t)$ and area $A(t)$ not observed before. See Theorem 2.1. We also reinterpret Yagisita's estimate in terms of the asymptotic behavior of the isoperimetric difference $D(t) := L^2(t) - 4\pi A(t)$. See Lemmas 3.7 and 3.9.

Remark 1.1. This is to explain how to convert Chou and Zhu's results for $k - \lambda$ flow into results for $k - 1$ flow. Chou and Zhu [2001, Theorem 3.12] considered a general flow which includes the following as a special case:

$$\frac{\partial X}{\partial t}(\varphi,t)=(k(\varphi,t)-\lambda)\boldsymbol{N}_{\text{in}}(\varphi,t),\quad X(\varphi,0)=X_0(\varphi),\quad \varphi\in S^1. \tag{4}$$

Here $\lambda\in\mathbb{R}$ is a number serving as a parameter. For a given initial curve $\gamma_0:=X_0(\varphi)$, there exists a unique number Λ such that the flow (4) with $\lambda=\Lambda$ will evolve γ_0 smoothly into a circle with radius $1/\Lambda$ as $t\to\infty$, $t\in[0,\infty)$. We assert that if we replace γ_0 by $\tilde{\gamma}_0:=\Lambda\gamma_0$ and rescale time (denote the new time as τ), then, under the $k-1$ flow (1), $\tilde{\gamma}_0$ will converge to the unit circle S^1 as $\tau\to\infty$. More precisely, let $X(\varphi,t)$ be the solution to the $k-\Lambda$ flow with initial condition $X_0(\varphi)$ and set

$$\widetilde{X}(\varphi,\tau)=\Lambda X(\varphi,t),\quad \tau=\Lambda^2 t\in[0,\infty).$$

Then by $\tilde{k}(\varphi,\tau)=(1/\Lambda)k(\varphi,t)$, $\widetilde{\boldsymbol{N}}_{\text{in}}(\varphi,\tau)=\boldsymbol{N}_{\text{in}}(\varphi,t)$, we have

$$\widetilde{X}(\varphi,0)=\Lambda X_0(\varphi)=\tilde{\gamma}_0,$$

$$\frac{\partial\widetilde{X}}{\partial\tau}(\varphi,\tau)=\Lambda\frac{dt}{d\tau}\frac{\partial X}{\partial t}(\varphi,t)=\frac{1}{\Lambda}(k(\varphi,t)-\Lambda)\boldsymbol{N}_{\text{in}}(\varphi,t)=(\tilde{k}(\varphi,\tau)-1)\widetilde{\boldsymbol{N}}_{\text{in}}(\varphi,\tau)$$

for all $(\varphi,\tau)\in S^1\times[0,\infty)$. That is, $\widetilde{X}(\varphi,\tau)$ satisfies the $k-1$ flow (with initial condition $\tilde{\gamma}_0=\Lambda\gamma_0$) and converges to the unit circle S^1 as $\tau\to\infty$.

2. Estimate of the critical number Λ

According to [Chou and Zhu 2001], the critical number Λ is obtained via a contradiction argument and for a given curve $X_0(\varphi)$ we do not know what it is. However, we can use the following theorem to give an estimate of Λ (see Corollary 2.7).

Theorem 2.1. *Let γ_0 be a convex closed curve (which is not a unit circle) and consider (1) with initial data γ_0. If the flow is defined on time interval $[0,\infty)$ and γ_t converges (without rescaling) to the unit circle S^1 as $t\to\infty$, its length $L(t)$ and enclosed area $A(t)$ must satisfy the estimate*

$$L(t)>2\pi,\quad A(t)<\pi,\quad L(t)-2\pi>\pi-A(t) \tag{5}$$

for all $t\in[0,\infty)$. Moreover, $L(t)$ is strictly decreasing, $A(t)$ is strictly increasing, and $(L(t)-2\pi)-(\pi-A(t))$ is strictly decreasing on $[0,\infty)$.

The proof consists of several simple lemmas. Recall that for a family of time-dependent simple closed curves $\gamma_t=\gamma(\cdot,t)$ in the plane its length $L(t)$ and enclosed area $A(t)$ satisfy the equations

$$\frac{dL}{dt}(t)=-\int_{\gamma(\cdot,t)}\left\langle\frac{\partial\gamma}{\partial t},k\boldsymbol{N}_{\text{in}}\right\rangle ds,\quad \frac{dA}{dt}(t)=-\int_{\gamma(\cdot,t)}\left\langle\frac{\partial\gamma}{\partial t},\boldsymbol{N}_{\text{in}}\right\rangle ds, \tag{6}$$

where $\boldsymbol{N}_{\text{in}}$ is the unit inward normal of γ_t and k is its signed curvature with respect to $\boldsymbol{N}_{\text{in}}$. Therefore

$$\frac{d}{dt}(L(t)-A(t)) = -\int_{\gamma(\cdot,t)} \left\langle \frac{\partial\gamma}{\partial t}, (k-1)\boldsymbol{N}_{\text{in}} \right\rangle ds, \tag{7}$$

which explains why (1) is the gradient flow of the functional $E(\gamma)$ in (2). In particular, under the flow (1), we have

$$\frac{dL}{dt}(t) = -\int_{\gamma(\cdot,t)} k^2 ds + 2\pi, \quad \frac{dA}{dt}(t) = L(t) - 2\pi. \tag{8}$$

As γ_0 is strictly convex, γ_t will remain so for a short time (we may assume γ_t is convex on $[0, T)$ for some $T > 0$). Thus one can use the outward normal angle $\theta \in S^1$ of γ_t as a parametrization variable. In terms of $(\theta, t) \in S^1 \times [0, T)$ we have

$$\frac{\partial k}{\partial t}(\theta, t) = k^2(\theta, t)[k_{\theta\theta}(\theta, t) + k(\theta, t)] - k^2(\theta, t) \tag{9}$$

and

$$\frac{\partial u}{\partial t}(\theta, t) = 1 - k(\theta, t) = 1 - \frac{1}{u_{\theta\theta}(\theta, t) + u(\theta, t)}, \tag{10}$$

where $u(\theta, t)$ is the support function of γ_t. We also have

$$L(t) = \int_0^{2\pi} u(\theta, t)\, d\theta, \quad A(t) = \tfrac{1}{2}\int_0^{2\pi} [u^2(\theta, t) - u_\theta^2(\theta, t)]\, d\theta. \tag{11}$$

Let $w(\theta, t) = k(\theta, t)e^{t/4}$ and compute

$$\frac{\partial w}{\partial t}(\theta, t) = k^2(\theta, t) w_{\theta\theta}(\theta, t) + \big(k(\theta, t) - \tfrac{1}{2}\big)^2 w(\theta, t).$$

By the maximum principle we can obtain a lower bound of the curvature:

$$k(\theta, t) \geq k_{\min}(0) e^{-t/4} > 0 \tag{12}$$

for all $(\theta, t) \in S^1 \times [0, T)$, where $k_{\min}(0) = \min_{\theta\in[0,2\pi]} k(\theta, 0)$. By Theorem 3.12 of [Chou and Zhu 2001], the flow γ_t (each γ_t remains smooth and convex) is either defined on a finite maximal time interval $[0, T_{\max})$ with $\lim_{t\to T_{\max}} k_{\max}(t) = \infty$ or on an infinite time interval $[0, \infty)$ with $\lim_{t\to\infty} k(\theta, t) = 1$ or $\lim_{t\to\infty} k(\theta, t) = 0$ uniformly on S^1.

Note that for any simple closed curve γ in the plane, we have $\int_\gamma k\, ds = 2\pi$ and

$$\int_\gamma k^2\, ds \geq \frac{4\pi^2}{L} \quad \text{(Hölder inequality)}, \tag{13}$$

and, by Gage's isoperimetric inequality [1983], we have

$$\int_\gamma k^2\,ds \geq \frac{\pi L}{A} \tag{14}$$

for any convex closed curve γ in $\mathbb{R}^2$. We also need the fact that the equality holds in (13) or (14) if and only if γ is a circle.

As a consequence of (14), the isoperimetric difference and ratio of γ_t, under the $k-1$ flow, are both decreasing (strictly decreasing if γ_0 is not a circle) in time due to

$$\frac{d}{dt}(L^2 - 4\pi A) \leq -\frac{2\pi}{A}(L^2 - 4\pi A) \leq 0 \tag{15}$$

and

$$\frac{d}{dt}\Big(\frac{L^2}{4\pi A} - 1\Big) \leq \frac{-L}{A}\Big(\frac{L^2}{4\pi A} - 1\Big) \leq 0. \tag{16}$$

Thus, in any case of convergence, γ_t is getting more and more circular.

Let $L(0)$ and $A(0)$ be the length and area of γ_0 (γ_0 is not a unit circle).

Lemma 2.2. *If $L(0) \leq 2\pi$, the flow (1) strictly decreases $L(t)$ and $A(t)$, and γ_t converges to a point $p \in \mathbb{R}^2$ in finite time $T_{\max}$.*

Remark 2.3. If $L(0) \leq 2\pi$, γ_0 may not be enclosed by a circle with radius less than 1. Otherwise the result is trivial due to the maximum principle.

Proof. Since γ_0 is not a unit circle, if $L(0) \leq 2\pi$, we must have $A(0) < \pi$ due to $L^2(0) > 4\pi A(0)$. We also have strict inequality in (13). By (8), we have

$$\frac{dL}{dt} = -\int_{\gamma(\cdot,t)} k^2\,ds + 2\pi \leq \frac{2\pi}{L(t)}(L(t) - 2\pi),$$

and at $t = 0$ we have $(dL/dt)(0) < 0$. Thus the flow (1) strictly decreases $L(t)$. By $dA/dt = L(t) - 2\pi$, it also strictly decreases $A(t)$. As a consequence of Theorem 3.12 of [Chou and Zhu 2001], γ_t will converge to a point $p \in \mathbb{R}^2$ in finite time $T_{\max}$. $\square$

Lemma 2.4. *If $A(0) \geq \pi$, the flow (1) strictly increases $A(t)$, and γ_t expands to infinity as $t \to \infty$.*

Remark 2.5. If $A(0) \geq \pi$, γ_0 may not enclose a circle with radius larger than 1. Otherwise the result is trivial due to the maximum principle.

Proof. We now have $L(0) > 2\pi$ and $(dA/dt)(0) = L(0) - 2\pi > 0$. By continuity $L(t)$ remains $L(t) > 2\pi$ for a short time $[0, T)$ and $A(t)$ is strictly increasing with $A(t) > \pi$ on $(0, T)$. As time proceeds, the inequality $L^2 \geq 4\pi A$ forces $L(t)$ to remain $L(t) > 2\pi$ and $A(t)$ keeps strictly increasing. Again by Theorem 3.12 of

[Chou and Zhu 2001], the flow is defined on $[0,\infty)$ and γ_t expands to infinity as $t \to \infty$. □

Lemma 2.6. *If $L(0) > 2\pi$, $A(0) < \pi$, and $L(0) - 2\pi \leq \pi - A(0)$, γ_t converges to a point $p \in \mathbb{R}^2$ in finite time $T_{\max}$.*

Proof. By continuity $L(t) > 2\pi$ and $A(t) < \pi$ for a short time $[0, T)$, and during this time interval we have

$$\begin{aligned} (17)\qquad \frac{d}{dt}(L(t) - 2\pi) &= -\int_{\gamma(\cdot,t)} k^2\, ds + 2\pi \\ &< -\frac{\pi L(t)}{A(t)} + 2\pi < -(L(t) - 2\pi) = \frac{d}{dt}(\pi - A(t)) < 0. \end{aligned}$$

Thus $L(t)$ is strictly decreasing and $A(t)$ is strictly increasing.

Equation (17) says that $L(t) - 2\pi$ decreases more rapidly than $\pi - A(t)$ (as long as $L(t) > 2\pi$ and $A(t) < \pi$). Since $L(0) - 2\pi$ is closer to 0 than $\pi - A(0)$, $L(t)$ must touch 2π earlier than $A(t)$ touches π. More precisely, let $t_* > 0$ be the first time at which $L(t) > 2\pi$ and $A(t) < \pi$ on $[0, t_*)$ and $L(t_*) = 2\pi$. Such a t_* must exist and is finite. Otherwise we would have $L(t) > 2\pi$ and $A(t) < \pi$ on $[0,\infty)$, and, by [Chou and Zhu 2001, Theorem 3.12], the flow would have to converge to the unit circle S^1 (without rescaling), which is impossible due to the inequality

$$\begin{aligned} (L(t) - 2\pi) - (\pi - A(t)) &< (L(t_1) - 2\pi) - (\pi - A(t_1)) \\ &< [(L(0) - 2\pi) - (\pi - A(0))] \leq 0 \end{aligned}$$

for all $t > t_1 > 0$ in $[0,\infty)$. (Note that now we have $L(t) \to 2\pi$ and $A(t) \to \pi$ as $t \to \infty$.) Therefore $t_* > 0$ is finite and $L(t_*) = 2\pi$. By Lemma 2.2, γ_t must converge to a point $p \in \mathbb{R}^2$ in finite time $T_{\max}$. □

Proof of Theorem 2.1. Combining these three lemmas, the proof of Theorem 2.1 is now clear. Since we assume that γ_t converges to the unit circle S^1 (without rescaling) as $t \to \infty$, if at some time $t_0 \in [0,\infty)$ we have $L(t_0) \leq 2\pi$ or $A(t_0) \geq \pi$, the curve will either converge to a point or expand to infinity. Hence we must have $L(t) > 2\pi$ and $A(t) < \pi$ for all time. By Lemma 2.6 we also have $L(t) - 2\pi > \pi - A(t)$ for all time. The monotonicity of $L(t)$, $A(t)$, and $(L(t) - 2\pi) - (\pi - A(t))$ can all be seen from (17). □

As a consequence of Theorem 2.1, we can give an estimate of the number Λ in Theorem 3.12 of [Chou and Zhu 2001].

Corollary 2.7. *Let γ_0 be a convex closed curve (which is not a unit circle) with length $L(0)$ and area $A(0)$. Then its critical number Λ satisfies*

$$(18)\qquad \Lambda L(0) > 2\pi, \quad \Lambda^2 A(0) < \pi, \quad \Lambda L(0) - 2\pi > \pi - \Lambda^2 A(0),$$

which implies

$$\max\left\{\frac{2\pi}{L(0)}, \frac{-L(0)+\sqrt{L^2(0)+12\pi A(0)}}{2A(0)}\right\} < \Lambda < \sqrt{\frac{\pi}{A(0)}}. \tag{19}$$

Remark 2.8. Let $k_0(\theta)$ be the curvature of γ_0. As $\Lambda\gamma_0$ converges to the unit circle S^1 as $t\to\infty$, its curvature $(1/\Lambda)k_0(\theta)$ must satisfy $(1/\Lambda)\max_{\theta\in S^1} k_0(\theta) > 1$ and $(1/\Lambda)\min_{\theta\in S^1} k_0(\theta) < 1$. This gives a rough estimate of Λ in terms of the curvature of γ_0, that is,

$$\min_{\theta\in S^1} k_0(\theta) < \Lambda < \max_{\theta\in S^1} k_0(\theta). \tag{20}$$

We explain that (19) is better than (20). To see this, by the identity

$$L(0) = \int_0^{2\pi} \frac{1}{k_0(\theta)}\, d\theta$$

and Gage's inequality (14), we have

$$\min_{\theta\in S^1} k_0(\theta) < \frac{2\pi}{L(0)} \quad \text{and} \quad \frac{L(0)}{2A(0)} < \max_{\theta\in S^1} k_0(\theta).$$

Combined with the classical isoperimetric inequality

$$\frac{2\pi}{L(0)} < \sqrt{\frac{\pi}{A(0)}} < \frac{L(0)}{2A(0)},$$

we conclude

$$\min_{\theta\in S^1} k_0(\theta) < \frac{2\pi}{L(0)} < \sqrt{\frac{\pi}{A(0)}} < \frac{L(0)}{2A(0)} < \max_{\theta\in S^1} k_0(\theta). \tag{21}$$

Hence (19) is better than the curvature estimate (20).

Remark 2.9. Under the assumption of Theorem 2.1, the curvature $k(\theta,t)\to 1$ uniformly as $t\to\infty$. One can follow a similar proof to that of [Gage and Hamilton 1986, Theorem 5.7.1] to conclude the following curvature estimate: for any $m\in\mathbb{N}$ and any $\alpha\in(0,1)$, there exists a constant C depending only on m and γ_0 such that

$$\left\|\frac{\partial^m k}{\partial\theta^m}(\theta,t)\right\|_{L^\infty(S^1)} \le C(m)e^{-2\alpha t} \tag{22}$$

for time t large enough.

Given initial curve γ_0, if we replace the $k-1$ flow by $ck-d$ flow, the critical number $\widetilde{\Lambda}$ for the $ck-d$ flow and the critical number Λ for the $k-1$ flow are related by the following.

Corollary 2.10. *Let c, d be two positive constants and let γ_0 be a convex closed curve parametrized by $X_0(\varphi)$, $\varphi \in S^1$. Then the critical number Λ in the $k-1$ flow and the critical number $\widetilde{\Lambda}$ in the $ck-d$ flow are related by*

$$\widetilde{\Lambda} = \frac{c}{d}\Lambda. \tag{23}$$

Proof. This is a consequence of scaling. By definition, the solution $\widetilde{X}(\varphi, t)$ to the initial value problem

$$\frac{\partial \widetilde{X}}{\partial t}(\varphi, t) = (c\tilde{k}(\varphi, t) - d)\widetilde{\boldsymbol{N}}_{\text{in}}(\varphi, t), \quad \widetilde{X}(\varphi, 0) = \widetilde{\Lambda} X_0(\varphi)$$

will converge to the circle with radius $R = c/d$ as $t \to \infty$. Let

$$Y(\varphi, t) = \frac{d}{c}\widetilde{X}\Big(\varphi, \frac{c}{d^2}t\Big).$$

Then it satisfies

$$\frac{\partial Y}{\partial t}(\varphi, t) = (k^{(Y)}(\varphi, t) - 1)\boldsymbol{N}_{\text{in}}^{(Y)}(\varphi, t), \quad Y(\varphi, 0) = \frac{d}{c}\widetilde{\Lambda} X_0(\varphi),$$

where $k^{(Y)}(\varphi, t)$ and $\boldsymbol{N}_{\text{in}}^{(Y)}(\varphi, t)$ are the curvature and normal at $Y(\varphi, t)$, respectively. Since $Y(\varphi, t)$ will converge to the unit circle, we have $(d/c)\widetilde{\Lambda} = \Lambda$. □

3. The asymptotic behavior of $L^2 - 4\pi A$ in the expanding case

There is another interesting property of the flow (1) not discussed before when, given an initial curve γ_0, it expands to infinity. It is about the value of $D(t) := L^2(t) - 4\pi A(t)$ as $t \to \infty$. From (8) it is easy to see that $L(t)$ has scale t and $A(t)$ has scale t^2 as $t \to \infty$ (since $k(\theta, t) \to 0$ uniformly on S^1). If we integrate (16) with respect to time, we get

$$0 \le \frac{L^2(t)}{4\pi A(t)} - 1 \le \Big(\frac{L^2(0)}{4\pi A(0)} - 1\Big)e^{-\int_0^t (L(z)/A(z))\,dz}, \quad \lim_{t\to\infty}\int_0^t \frac{L(z)}{A(z)}\,dz = \infty.$$

This implies $L^2(t)/(4\pi A(t)) \to 1$ exponentially as $t \to \infty$. But if we integrate (15), we only get

$$0 \le D(t) \le D(0)e^{-\int_0^t (2\pi/A(z))\,dz}, \quad \lim_{t\to\infty}\int_0^t \frac{2\pi}{A(z)}dz < \infty,$$

which implies that $L^2(t) - 4\pi A(t)$ decreases to some number bounded above by

$$D(0)e^{-\int_0^\infty (2\pi/A(z))\,dz}.$$

We shall see that, unless γ_0 is a circle, $D(t)$ will not decrease to zero as $t \to \infty$.

Remark 3.1. Note that if we have a family of curves γ_t, $t \in [0, T)$, $T \le \infty$, so that $A(t)$ has uniform positive upper and lower bounds, then $\lim_{t\to T} D(t) = 0$ is equivalent to $\lim_{t\to T} L^2(t)/(4\pi A(t)) = 1$. But if $\lim_{t\to T} A(t) = 0$ or ∞, then they may not be equivalent.

Recall that if we evolve a convex closed curve γ_0 by the *unit-speed* (that is, the constant 1) *outward normal flow*, we get a family of *parallel curves* γ_t expanding to infinity. We get a similar result if we replace the speed constant 1 by a positive time function $a(t)$, where $\lim_{t\to\infty} a(t) = \infty$. Moreover, any two parallel convex closed curves (or simple closed curves) γ_{t_1} and γ_{t_2} have the *same* isoperimetric difference.

The intuitive observation is that when γ_0 expands to infinity under the $k-1$ flow, its asymptotic behavior is given by the unit-speed outward normal flow. As the unit-speed outward normal flow preserves the isoperimetric difference, we expect that $D(t)$ will not decrease to zero as $t \to \infty$. This is indeed the case based on results in [Yagisita 2005], which are explained below.

From now on we use $\theta \in S^1$ to denote the outward normal angle of γ_t (γ_t is convex) and use $\sigma \in S^1$ to denote the *polar angle* of γ_t (we may assume that γ_t encloses the origin of $\mathbb{R}^2$). In Theorem 2 of [Yagisita 2005], he looked at the radial function $r(\sigma, t)$ of γ_t and proved that there exists a smooth function $\ell(\sigma)$ defined on S^1 such that

$$\lim_{t\to\infty} \|r(\sigma, t) - (R(t) + \ell(\sigma))\|_{C^k(S^1)} = 0 \tag{24}$$

for any $k \in \mathbb{N}$, where $R(t)$ is the solution to the ODE

$$\frac{dR}{dt}(t) = 1 - \frac{1}{R(t)}, \qquad R(0) = 2. \tag{25}$$

Note that $R(t)$ is strictly increasing on $[0, \infty)$ with $\lim_{t\to\infty} R'(t) = 1$.

To see that $D(t)$ decreases to a *positive* constant asymptotically, we need to see what (24) implies in terms of the support function $u(\theta, t)$ of γ_t. By (11), it suffices to look at the asymptotic behavior of $u(\theta, t)$ and $u_\theta(\theta, t)$.

Remark 3.2. Yagisita [2005] used the radial function $r(\sigma, t)$ to study the flow (1) instead of the support function $u(\theta, t)$. The advantage is that one can get a *quasilinear* uniformly parabolic equation for the difference $A(\sigma, \tau) := r(\sigma, t) - R(t)$ (see pages 227–230 of [Yagisita 2005]) if we also rescale time. More precisely, let

$$\tau(t) = \log\Big(1 - \frac{1}{R(t)}\Big) : [0, \infty) \to [-\log 2, 0), \qquad \frac{d\tau}{dt} = \frac{1}{R^2(t)}.$$

Then we have

$$\begin{aligned} \frac{\partial A}{\partial \tau}(\sigma, \tau) & \\ &= \frac{R^2(t(\tau))}{[A(\sigma, \tau) + R(t(\tau))]^2 + A_\sigma^2(\sigma, \tau)} A_{\sigma\sigma}(\sigma, \tau) + \text{lower order terms} \end{aligned} \tag{26}$$

for all $(\sigma, \tau) \in S^1 \times [-\log 2, 0)$, where $R(t(\tau)) = 1/(1-e^\tau)$. On the other hand, the evolution equation for $B(\theta, \tau) := u(\theta, t) - R(t)$ is also uniformly parabolic but *fully nonlinear*, that is,

$$\frac{\partial B}{\partial \tau}(\theta, \tau) = \frac{R(t(\tau))}{B_{\theta\theta}(\theta, \tau) + B(\theta, \tau) + R(t(\tau))}(B_{\theta\theta}(\theta, \tau) + B(\theta, \tau)) \tag{27}$$

for all $(\theta, \tau) \in S^1 \times [-\log 2, 0)$. Equation (26) is easier to handle than (27). However, the disadvantage is that it is very awkward to use $r(\sigma, t)$ to study the isoperimetric difference $D(t)$.

Lemma 3.3. *Let γ_0 be a convex closed curve and consider the flow (1) with initial data γ_0. Then (24) implies*

$$\lim_{t\to\infty} \|u(\theta, t) - (R(t) + \ell(\theta))\|_{C^1(S^1)} = 0. \tag{28}$$

In particular, we have

$$\lim_{t\to\infty} D(t) = \left(\int_0^{2\pi} \ell(\theta)\, d\theta\right)^2 - 2\pi \int_0^{2\pi} \left(\ell^2(\theta) - (\ell'(\theta))^2\right) d\theta \geq 0. \tag{29}$$

Remark 3.4. We may get higher order convergence of $u(\theta, t)$. But (28) is sufficient.

Proof. For a point $p \in \gamma_t$ with position vector P, its support function $u(\theta, t)$ and radial function $r(\sigma, t)$ are related by

$$P = u(\theta, t)(\cos\theta, \sin\theta) + u_\theta(\theta, t)(-\sin\theta, \cos\theta) = r(\sigma, t)(\cos\sigma, \sin\sigma). \tag{30}$$

From this we get

$$u(\theta, t) = r(\sigma, t)\cos(\sigma - \theta), \quad u_\theta(\theta, t) = r(\sigma, t)\sin(\sigma - \theta), \tag{31}$$

and

$$\sigma = \sigma(\theta, t) = \tan^{-1}\left(\frac{u(\theta, t)\sin\theta + u_\theta(\theta, t)\cos\theta}{u(\theta, t)\cos\theta - u_\theta(\theta, t)\sin\theta}\right), \quad \theta \in S^1. \tag{32}$$

In particular, at any point p where $\theta = \sigma$, we have $u(\theta, t) = r(\sigma, t)$ and $u_\theta(\theta, t) = 0$.

Since we know that $|u_\theta(\theta, t)|$ and $|u_{\theta\theta}(\theta, t)|$ are both uniformly bounded on $S^1 \times [0, \infty)$ (see [Chow and Tsai 1996]) and $u(\theta, t) \to \infty$ uniformly, we have $\lim_{t\to\infty} \sigma(\theta, t) = \theta$ uniformly on S^1 and

$$\lim_{t\to\infty} \frac{\partial\sigma}{\partial\theta}(\theta, t) = \lim_{t\to\infty} \frac{u(\theta, t)(u_{\theta\theta}(\theta, t) + u(\theta, t))}{u^2(\theta, t) + u_\theta^2(\theta, t)} = 1 \tag{33}$$

uniformly on S^1. Since $u_\theta(\theta, t) = r(\sigma, t)\sin(\sigma - \theta)$, we have

$$(34)\quad R(t)\sin(\sigma-\theta) = u_\theta(\theta,t) - (r(\sigma,t)-R(t)-\ell(\sigma))\sin(\sigma-\theta) - \ell(\sigma)\sin(\sigma-\theta).$$

This implies that $|R(t)\sin(\sigma-\theta)|$ is also uniformly bounded on $S^1\times[0,\infty)$. Now, by (24), we conclude

$$(35)\quad \begin{aligned} &\lim_{t\to\infty}(u(\theta,t)-R(t)) \\ &= \lim_{t\to\infty}\big((r(\sigma,t)-R(t)-\ell(\sigma))\cos(\sigma-\theta)+(R(t)+\ell(\sigma))\cos(\sigma-\theta)-R(t)\big) \\ &= \lim_{t\to\infty}\left((R(t)\sin(\sigma-\theta))\frac{\sigma-\theta}{\sin(\sigma-\theta)}\frac{\cos(\sigma-\theta)-1}{\sigma-\theta}\right)+\ell(\theta) \\ &= \ell(\theta), \end{aligned}$$

uniformly on S^1. We next claim that $u_\theta(\theta,t)\to\ell'(\theta)$ uniformly on S^1 as $t\to\infty$. Apply $\partial/\partial\theta$ to $u(\theta,t)$ and use the chain rule to get

$$u_\theta(\theta,t)=r_\sigma(\sigma,t)\frac{\partial\sigma}{\partial\theta}\cos(\sigma-\theta)-r(\sigma,t)\Big(\frac{\partial\sigma}{\partial\theta}(\theta,t)-1\Big)\sin(\sigma-\theta),\quad \sigma=\sigma(\theta,t),$$

and hence

$$(36)\qquad \lim_{t\to\infty}u_\theta(\theta,t)=\ell'(\theta)$$

uniformly on S^1 due to (33) and (24). Since

$$\begin{aligned}\|u(\theta,t)-r(\sigma,t)\|_{C^1(S^1)} \le \|u(\theta,t)-(R(t)+\ell(\theta))\|_{C^1(S^1)}+\|\ell(\theta)-\ell(\sigma)\|_{C^1(S^1)} \\ +\|(\ell(\sigma)+R(t))-r(\sigma,t)\|_{C^1(S^1)}\end{aligned}$$

we have

$$(37)\qquad \lim_{t\to\infty}\|u(\theta,t)-r(\sigma,t)\|_{C^1(S^1)}=0,\quad \sigma=\sigma(\theta,t),$$

where the C^1 norm is taken with respect to $\theta\in S^1$. By (36) and (34), we also have

$$(38)\qquad \begin{aligned}\lim_{t\to\infty}R(t)(\sigma-\theta)&=\lim_{t\to\infty}R(t)\sin(\sigma-\theta)\\ &=\ell'(\theta)\end{aligned}$$

uniformly on S^1.

As a consequence of (11) we have

$$(39)\quad \begin{aligned}&\lim_{t\to\infty}D(t)\\ &=\lim_{t\to\infty}\left[\left(\int_0^{2\pi}u(\theta,t)\,d\theta\right)^2-2\pi\int_0^{2\pi}u^2(\theta,t)\,d\theta\right]+2\pi\int_0^{2\pi}(\ell'(\theta))^2\,d\theta.\end{aligned}$$

Since we have the identity

$$\left(\int_0^{2\pi} u(\theta,t)\,d\theta\right)^2 - 2\pi\int_0^{2\pi} u^2(\theta,t)\,d\theta = \left(\int_0^{2\pi}(u(\theta,t)-R(t))\,d\theta\right)^2 - 2\pi\int_0^{2\pi}(u(\theta,t)-R(t))^2\,d\theta, \tag{40}$$

which is due to the fact that any two parallel convex closed curves have the same isoperimetric difference, we conclude

$$\lim_{t\to\infty} D(t) = \left(\int_0^{2\pi}\ell(\theta)\,d\theta\right)^2 - 2\pi\int_0^{2\pi}\left(\ell^2(\theta)-(\ell'(\theta))^2\right)d\theta \geq 0. \tag{41}$$

□

Remark 3.5. By (30) and (28) the position vector $P(\theta,t)$ of γ_t satisfies

$$\lim_{t\to\infty}|P(\theta,t)-Q(\theta,t)| = 0$$

uniformly on S^1, where

$$Q(\theta,t) = R(t)(\cos\theta,\sin\theta) + \ell(\theta)(\cos\theta,\sin\theta) + \ell_\theta(\theta)(-\sin\theta,\cos\theta) \tag{42}$$

and it represents a family of expanding circles centered at $(a,b)\in\mathbb{R}^2$ if and only if $\ell(\theta)$ is given by

$$\ell(\theta) = c + a\cos\theta + b\sin\theta, \quad \theta\in S^1 \tag{43}$$

for some constants a, b, c. Also note that by the classical Minkowski inequality the right side of (41) is zero if and only if $\ell(\theta)$ has the form (43). So $\lim_{t\to\infty} D(t) = 0$ if and only if $P(\theta,t)$ is asymptotically close to a family of expanding circles centered at some $(a,b)\in\mathbb{R}^2$, which can be evaluated by the integral

$$(a,b) = \lim_{t\to\infty}\frac{1}{2\pi}\int_0^{2\pi}P(\theta,t)\,d\theta = \frac{1}{\pi}\int_0^{2\pi}\ell(\theta)(\cos\theta,\sin\theta)\,d\theta. \tag{44}$$

On the other hand, $\lim_{t\to\infty} D(t) = d > 0$ if and only if $P(\theta,t)$ is asymptotically close to a family of *noncircular* parallel curves (described by $Q(\theta,t)$) expanding to infinity. This family of parallel curves have the same fixed *center* (now "*center*" means "*average position vector*") given by (44). The speed of this family of parallel curves is

$$\frac{dR}{dt}(t) = 1 - \frac{1}{R(t)} \to 1$$

as $t\to\infty$. Therefore, asymptotically, it is the *unit-speed outward normal flow*.

To go further, we need the following ODE result.

Lemma 3.6. *For any constant $c \in \mathbb{R}$ there exists a positive solution $s(t)$ to the ODE*

$$\frac{ds}{dt} = 1 - \frac{1}{s}$$

defined on some interval $[T, \infty)$, $T \geq 0$, $s(T) \geq 2$, such that

$$\lim_{t\to\infty} (s(t) - R(t)) = c.$$

Proof. Assume first that $c > 0$. Let $s(t) = R(t+c)$, $t \in [0, \infty)$. It satisfies the same ODE with $s(0) = R(c) > R(0) = 2$. Now

$$\lim_{t\to\infty} (s(t) - R(t)) = \lim_{t\to\infty} \int_t^{t+c} R'(z)\, dz = \lim_{t\to\infty} \int_t^{t+c} \Big(1 - \frac{1}{R(z)}\Big)\, dz = c.$$

For $c < 0$, let $s(t) = R(t+c)$, $t \in [-c, \infty)$. Then $s(t)$ is a positive solution to the ODE on $[-c, \infty)$ with $s(-c) = R(0) = 2$ and

$$\lim_{t\to\infty} (R(t) - s(t)) = \lim_{t\to\infty} \int_{t+c}^{t} R'(z)\, dz = \lim_{t\to\infty} \int_{t+c}^{t} \Big(1 - \frac{1}{R(z)}\Big)\, dz = -c. \qquad \square$$

Our next result is a property about uniqueness.

Lemma 3.7. *If γ_t expands to infinity under the flow (1), $\lim_{t\to\infty} D(t) = 0$ if and only if γ_0 is a circle. Therefore if γ_0 is not a circle, $D(t)$ will decrease to a positive constant as $t \to \infty$.*

Proof. Assume that $\lim_{t\to\infty} D(t) = 0$. Then $\ell(\theta) = c + a\cos\theta + b\sin\theta$ in (28), and by Lemma 3.6 there exists a positive solution $s(t)$ to the ODE on some interval $[T, \infty)$, $T \geq 0$, $s(T) \geq 2$, such that

$$\lim_{t\to\infty} \|u(\theta, t) - (s(t) + a\cos\theta + b\sin\theta)\|_{C^1(S^1)} = 0.$$

Now if at time T we consider a circle C_T centered at (a, b) with radius $s(T)$ and evolve it under the flow (1), its support function $U(\theta, t)$ will satisfy

$$U(\theta, t) = s(t) + a\cos\theta + b\sin\theta,$$

$(\theta, t) \in S^1 \times [T, \infty)$ (note that this $U(\theta, t)$ satisfies Equation (10)). By previous discussions, the radial function $r_1(\sigma, t)$ of the evolving curve γ_t (with support function $u(\theta, t)$) and the radial function $r_2(\sigma, t)$ of the evolving circle C_t (with support function $U(\theta, t)$) on the domain $S^1 \times [T, \infty)$ will satisfy the estimate

$$\lim_{t\to\infty} \|r_1(\sigma, t) - r_2(\sigma, t)\|_{C^1(S^1)} = 0, \tag{45}$$

and we can apply Theorem 3 of [Yagisita 2005] to conclude that $\gamma_t \equiv C_t$ for all $t \in [T, \infty)$. In particular γ_T is also a circle. But this is impossible unless γ_0 is a circle. $\square$

Remark 3.8. One can also apply *Andrews' backward uniqueness result* [2002], which follows, to prove the above lemma: Assume $v(\theta, t)$ is a smooth solution to the *uniformly parabolic* equation

$$\frac{\partial v}{\partial t} = a(\theta, t) v_{\theta\theta} + b(\theta, t) v_\theta + c(\theta, t) v, \quad (\theta, t) \in S^1 \times [0, T], \tag{46}$$

where a, b, c are smooth functions on $S^1 \times [0, T]$ with

$$\frac{1}{C} \le a(\theta, t) \le C, \quad (\theta, t) \in S^1 \times [0, T],$$

for some positive constant $C > 0$. If $v(\theta, T) = 0$ for all $\theta \in S^1$, we have $v(\theta, 0) = 0$ for all $\theta \in S^1$ (in particular, $v(\theta, t) \equiv 0$ for all (θ, t)). Now by (24), in terms of the variable $(\sigma, \tau) \in S^1 \times [-\log 2, 0)$ in Remark 3.2, the bounded function $A(\sigma, \tau) = r_1(\sigma, t) - R(t)$ can be smoothly extended to $S^1 \times [-\log 2, 0]$ even though the function $R(t(\tau)) = 1/(1 - e^\tau)$ is undefined at $\tau = 0$. The full equation for (26) is (see the top equation on page 229 of [Yagisita 2005])

$$\begin{aligned} \frac{\partial A}{\partial \tau}(\sigma, \tau) = {} & \frac{1}{C(\sigma, \tau)} A_{\sigma\sigma}(\sigma, \tau) + \frac{1}{D^2(\sigma, \tau) + D(\sigma, \tau)\sqrt{C(\sigma, \tau)}} A_\sigma^2(\sigma, \tau) \\ & + \frac{1}{C(\sigma, \tau)} \Big[D(\sigma, \tau) A(\sigma, \tau) - (1 - e^\tau)\Big(\frac{2}{D(\sigma, \tau)} - 1\Big) A_\sigma^2(\sigma, \tau) \Big], \end{aligned} \tag{47}$$

where

$$\begin{aligned} C(\sigma, \tau) &= [A(\sigma, \tau)(1 - e^\tau) + 1]^2 + A_\sigma^2(\sigma, \tau)(1 - e^\tau)^2, \\ D(\sigma, \tau) &= A(\sigma, \tau)(1 - e^\tau) + 1. \end{aligned}$$

At $\tau = 0$, we have $C(\sigma, \tau) = D(\sigma, \tau) = 1$, which implies that (47) is a *uniformly parabolic* equation with *smooth* coefficients on $S^1 \times [-\delta, 0]$ for some small $\delta > 0$. Moreover, the smooth function

$$w(\sigma, \tau) := (r_1(\sigma, t) - R(t)) - (r_2(\sigma, t) - R(t)), \quad (\sigma, \tau) \in S^1 \times [-\delta, 0],$$

satisfies a uniformly parabolic equation of the form

$$\frac{\partial w}{\partial \tau} = a(\sigma, \tau) w_{\sigma\sigma} + b(\sigma, \tau) w_\sigma + c(\sigma, \tau) w$$

with coefficients smooth on $S^1 \times [-\delta, 0]$, and, by (45), $w(\sigma, 0) = 0$ for all $\sigma \in S^1$. Andrews' result implies $w(\sigma, \tau) \equiv 0$ and the initial curve γ_0 must be a circle.

Lemma 3.9. *For any number $d \ge 0$ and any small $\varepsilon > 0$, one can construct an expanding $k-1$ flow so that its isoperimetric difference $D(t)$ satisfies $|D(t) - d| < \varepsilon$ as long as t is large enough.*

Proof. For any $d \geq 0$ one can find a convex closed curve γ with support function $\ell(\theta)$ such that

$$d = L^2 - 4\pi A = \left(\int_0^{2\pi} \ell(\theta)\, d\theta\right)^2 - 2\pi \int_0^{2\pi} \left(\ell^2(\theta) - (\ell'(\theta))^2\right) d\theta \geq 0.$$

Following the proof of Theorem 4 of [Yagisita 2005], we can obtain the following: For any smooth function $\ell(\theta)$ defined on S^1 and any small $\delta > 0$, there exists a large time $M > 0$ such that if $T > M$ and γ_T is a convex closed curve with support function $R(T) + \ell(\theta)$, the support function $u(\theta, t)$ of γ_t (γ_t is the evolution of γ_T under (1) on time interval $[T, \infty)$) will satisfy

$$\sup_{t\in[T,\infty)} \|u(\theta, t) - (R(t) + \ell(\theta))\|_{C^0(S^1)} < \delta. \tag{48}$$

This says that γ_t is close to a parallel curve of γ_T, which is intuitively correct.

For the isoperimetric difference of γ_t we need to be careful, because now the norm in (48) is only in C^0 norm [Yagisita 2005, Proof of Theorem 4], and by (11) we need to know the behavior of $u_\theta(\theta, t)$ in order to control the area. However, there is a result on page 53 of [Schneider 1993], which says that if two compact convex sets K_1, K_2 in $\mathbb{R}^2$ have their support functions $u_1(\theta)$, $u_2(\theta)$ close to each other, their *Hausdorff distance* is also close to each other. In particular, their lengths and areas are also close to each other. But be careful again that the two families of curves γ_t and p_t (p_t is the parallel curve of γ_T with support function $R(t)+\ell(\theta)$ for $t \in [T, \infty)$) are expanding to infinity as $t \to \infty$, so even if their Hausdorff distance is less than δ, $|D(t) - d|$ may not be small as $t \to \infty$ (however, $|D(t) - d| \leq d$ since $D(t) \geq 0$ is decreasing on $[T, \infty)$ with $D(T) = d$). To overcome this we can write (48) as

$$C + \ell(\theta) - \delta \leq u(\theta, t) - R(t) + C \leq C + \ell(\theta) + \delta, \quad (\theta, t) \in S^1 \times [T, \infty),$$

where $C > 0$ is a constant with (recall that $|u_{\theta\theta}(\theta, t)|$ is uniformly bounded on $S^1 \times [T, \infty)$ by [Chow and Tsai 1996])

$$\begin{gathered}(C + \ell(\theta))_{\theta\theta} + (C + \ell(\theta)) > 0,\\ (u(\theta, t) - R(t) + C)_{\theta\theta} + (u(\theta, t) - R(t) + C) > 0.\end{gathered} \tag{49}$$

for all $(\theta, t) \in S^1 \times [T, \infty)$. Equation (49) implies the existence of a convex closed curve C_1 with support function $C + \ell(\theta)$ and a convex closed curve $C_2(t)$ with support function $u(\theta, t) - R(t) + C$, where the support function of $C_2(t)$ is close to the support function of C_1 for all $t \in [T, \infty)$. Moreover, the curves C_1 and $C_2(t)$ are both enclosed by two parallel convex curves $C_\pm$ with support functions $C + \ell(\theta) + \delta$ and $C + \ell(\theta) - \delta$, respectively. However, we worry about the situation where, when δ is getting smaller and smaller, the constant C may be getting larger and larger.

We claim that the constant C in (49) can be chosen to be independent of δ. If we let δ tends to zero, the time T in (48) will tend to infinity and the initial value $u(\theta, T) = R(T) + \ell(\theta)$ will also tend to infinity. However, $u_\theta(\theta, T) = \ell'(\theta)$ and $u_{\theta\theta}(\theta, T) = \ell''(\theta)$ are unaffected by T. From the proofs of Proposition 1 and Lemma 4 in [Chow and Tsai 1996], one can see that $|u_\theta(\theta, t)|$ and $|u_{\theta\theta}(\theta, t)|$ are both uniformly bounded on $S^1 \times [T, \infty)$ and the bounds are independent of T. This, together with (48), implies that $C > -u_{\theta\theta}(\theta, t) + R(t) - u(\theta, t)$ is independent of T (and δ).

As $\delta \to 0$, the curve C_1 is unchanged and the Hausdorff distance between $C_2(t)$ and C_1 is getting smaller. We note that the isoperimetric difference of C_1 is given by d and the isoperimetric difference of $C_2(t)$ is the same as $D(t)$ of γ_t. Therefore, for any small $\varepsilon > 0$, by making $\delta > 0$ as small as possible, one can construct an expanding $k-1$ flow satisfying $|D(t) - d| < \varepsilon$ as long as t is large enough. □

Remark 3.10. In the above proof we use the fact that any smooth function $h(\theta)$ defined on S^1, satisfying $h''(\theta) + h(\theta) > 0$ for all $\theta \in S^1$, is the support function of some convex closed curve.

Acknowledgments

We thank the National Science Council and the National Center for Theoretical Sciences of Taiwan for their support over the past several years. Tsai thanks Professor Ben Andrews of the Australian National University for his longstanding mentoring.

References

[Andrews 1998] B. Andrews, "Evolving convex curves", *Calc. Var. Partial Differential Equations* **7**:4 (1998), 315–371. MR 99k:58038 Zbl 0931.53030

[Andrews 2002] B. Andrews, "Non-convergence and instability in the asymptotic behaviour of curves evolving by curvature", *Comm. Anal. Geom.* **10**:2 (2002), 409–449. MR 2003e:53086 Zbl 1029.53079

[Chou and Zhu 2001] K.-S. Chou and X.-P. Zhu, *The curve shortening problem*, CRC, Boca Raton, FL, 2001. MR 2003e:53088 Zbl 1061.53045

[Chow and Tsai 1996] B. Chow and D.-H. Tsai, "Geometric expansion of convex plane curves", *J. Differential Geom.* **44**:2 (1996), 312–330. MR 97m:58041 Zbl 0872.58052

[Gage 1983] M. E. Gage, "An isoperimetric inequality with applications to curve shortening", *Duke Math. J.* **50**:4 (1983), 1225–1229. MR 85d:52007 Zbl 0534.52008

[Gage 1984] M. E. Gage, "Curve shortening makes convex curves circular", *Invent. Math.* **76**:2 (1984), 357–364. MR 85i:52004 Zbl 0542.53004

[Gage and Hamilton 1986] M. Gage and R. S. Hamilton, "The heat equation shrinking convex plane curves", *J. Differential Geom.* **23**:1 (1986), 69–96. MR 87m:53003 Zbl 0621.53001

[Grayson 1987] M. A. Grayson, "The heat equation shrinks embedded plane curves to round points", *J. Differential Geom.* **26**:2 (1987), 285–314. MR 89b:53005 Zbl 0667.53001

[Green and Osher 1999] M. Green and S. Osher, "Steiner polynomials, Wulff flows, and some new isoperimetric inequalities for convex plane curves", *Asian J. Math.* **3**:3 (1999), 659–676. MR 2001j:53089 Zbl 0969.53040

[Schneider 1993] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, 1993. MR 94d:52007 Zbl 0798.52001

[Yagisita 2005] H. Yagisita, "Asymptotic behaviors of star-shaped curves expanding by $V = 1 - K$", *Differential Integral Equations* **18**:2 (2005), 225–232. MR 2005m:53128 Zbl 1212.53094

Received November 29, 2011. Revised June 24, 2012.

YU-CHU LIN
DEPARTMENT OF MATHEMATICS
NATIONAL CHENG KUNG UNIVERSITY
1 DASYUE ROAD
TAINAN 70101
TAIWAN

yclin@math.ncku.edu.tw

DONG-HO TSAI
DEPARTMENT OF MATHEMATICS AND NATIONAL CENTER FOR THEORETICAL SCIENCES
NATIONAL TSING HUA UNIVERSITY
101, SECTION 2, KUANG-FU ROAD
HSINCHU 30013
TAIWAN

dhtsai@math.nthu.edu.tw

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.215

DEMYSTIFYING A DIVISIBILITY PROPERTY OF THE KOSTANT PARTITION FUNCTION

KAROLA MÉSZÁROS

We prove a family of identities connected to a divisibility property of the Kostant partition function. A special case of these identities first appeared in a paper of Baldoni and Vergne. To prove their identities, Baldoni and Vergne used residue techniques, and called the resulting divisibility property "mysterious." Our proofs are entirely combinatorial and provide a natural explanation for why divisibility occurs, both in the Baldoni and Vergne identities and in their generalizations.

1. Introduction

The objective of this paper is to provide a natural combinatorial explanation of a divisibility property of the Kostant partition function. The question of evaluating Kostant partition functions has been the subject of much interest, without a satisfactory combinatorial answer. To mention perhaps the most famous such case: it is known that

$$K_{A_n^+}\left(1, 2, \ldots, n, -\binom{n+1}{2}\right) = \prod_{k=1}^{n} C_k, \quad \text{where } C_k = \frac{1}{k+1}\binom{2k}{k}$$

denotes the Catalan numbers, yet there is no combinatorial proof of this identity! While endowed with combinatorial meaning, Kostant partition functions were introduced in and are a vital part of representation theory: weight multiplicities and tensor product multiplicities can be expressed in terms of the Kostant partition function. Kostant partition functions also come up in toric geometry and analytic residue theory.

Given the lack of understanding of the evaluation of the Kostant partition function, it seems a worthy proposition to provide a simple explanation for certain of its divisibility properties. We explore divisibility properties of Kostant partition functions of types A_n and C_{n+1}, noting that such properties in types B_{n+1} and D_{n+1} are easy consequences of the type C_{n+1} case. A significant part of the type A_n

The author is supported by a National Science Foundation Postdoctoral Research Fellowship.
MSC2010: 05E10.
Keywords: Kostant partition function, divisibility, flow.

family of identities we study first appeared in [Baldoni and Vergne 2008], where the authors prove the identities using residues, and where they call the divisibility property "mysterious." It is our hope that the combinatorial argument we provide successfully demystifies the divisibility property of the Kostant partition function and provides a natural explanation for why things happen the way they do.

The outline of the paper is as follows. In Section 2 we define Kostant partition functions of type A_n and prove a family of identities, including the Baldoni–Vergne identities, combinatorially. Our proof is bijective, and as such it also yields an affirmative answer to a question of Stanley [2000] regarding a possible bijective proof of a special case of the Baldoni–Vergne identities. In Section 3 we define Kostant partition functions of type C_{n+1}, relate them to flows, and show how to modify our proof of the identities from Section 2 to obtain their analogues for type C_{n+1}.

2. A family of Kostant partition function identities

In this section we prove a family of Kostant partition function identities exhibiting divisibility properties. We start by proving Baldoni–Vergne identities, our proof of which clearly points to several generalizations of these identities. We provide some of these generalizations in this section, and some in the next.

The Baldoni–Vergne identities. Before stating the Baldoni–Vergne identities, we need a few definitions. Throughout this section the graphs G we consider are on the vertex set $[n+1]$, possibly with multiple edges, but no loops. Denote by m_{ij} the multiplicity of edge (i, j), $i < j$, in G. To each edge (i, j), $i < j$, of G, associate the positive type A_n root $e_i - e_j$, where e_i is the i-th standard basis vector. Let $\{\{\alpha_1, \dots, \alpha_N\}\}$ be the multiset of vectors corresponding to the multiset of edges of G as described above. Note that $N = \sum_{1 \le i < j \le n+1} m_{ij}$.

The *Kostant partition function* K_G evaluated at the vector $\boldsymbol{a} \in \mathbb{Z}^{n+1}$ is defined as

$$K_G(\boldsymbol{a}) = \#\left\{ (b_i)_{i \in [N]} \;\middle|\; \sum_{i \in [N]} b_i \alpha_i = \boldsymbol{a} \text{ and } b_i \in \mathbb{Z}_{\ge 0} \right\}. \tag{1}$$

That is, $K_G(\boldsymbol{a})$ is the number of ways to write the vector $\boldsymbol{a}$ as a nonnegative linear combination of the positive type A_n roots corresponding to the edges of G, without regard to order. Note that in order for $K_G(\boldsymbol{a})$ to be nonzero, the partial sums of the coordinates of $\boldsymbol{a}$ have to satisfy $a_1 + \cdots + a_i \ge 0$, $i \in [n]$, and $a_1 + \cdots + a_{n+1} = 0$.

We now proceed to state and prove Theorem 1, which first appeared in [Baldoni and Vergne 2008]. Baldoni and Vergne gave a proof of it using residues, and called the result "mysterious". We provide a natural combinatorial explanation of the result. Our explanation also answers a question of Stanley [2000] in the affirmative, regarding a possible bijective proof of a special case of the Baldoni–Vergne identities.

For brevity, we write $G-e$, or $G-\{e_1,\dots,e_k\}$, to mean a graph obtained from G with the edge e, or the edges $e_1,\dots,e_k$, deleted.

Theorem 1 [Baldoni and Vergne 2008]. *Given a connected graph G on the vertex set $[n+1]$ with $m_{n-1,n}=m_{n-1,n+1}=m_{n,n+1}=1$, and such that*

$$\frac{m_{j,n-1}+m_{j,n}+m_{j,n+1}}{m_{j,n-1}}=c \quad \text{for all } j\in[n-2],$$

for some constant c independent of j, we have

$$K_G(\boldsymbol{a})=\left(\frac{a_1+\cdots+a_{n-2}}{c}+a_{n-1}+1\right)K_{G-(n-1,n)}(\boldsymbol{a}), \tag{2}$$

for any $\boldsymbol{a}=\left(a_1,\dots,a_n,-\sum_{i=1}^n a_i\right)\in\mathbb{Z}^{n+1}$.

Before proceeding to the formal proof of Theorem 1 we outline it, to fully expose the underlying combinatorics. We introduce the notation

$$Q(\boldsymbol{a}):=\frac{a_1+\cdots+a_{n-2}}{c}+a_{n-1}+1 \tag{3}$$

for the factor in (2). Rephrasing Equation (1), $K_G(\boldsymbol{a})$ counts the number of *flows* $\boldsymbol{f}_G=(b_i)_{i\in N}$ on G satisfying

$$\sum_{i\in[N]} b_i\alpha_i=\boldsymbol{a} \quad \text{and} \quad b_i\in\mathbb{Z}_{\geq 0}.$$

In the proof of Theorem 1, we introduce the concept of *partial flows* $\boldsymbol{f}_H$, about which we prove two key statements:

- The elements of the set of partial flows are in bijection with the flows on $G-(n-1,n)$ that the Kostant partition function $K_{G-(n-1,n)}(\boldsymbol{a})$ counts. That is,
$$\#\text{ partial flows}=K_{G-(n-1,n)}(\boldsymbol{a}).$$
- The elements of the multiset of partial flows $\boldsymbol{f}_H$ — whose cardinality is $Q(\boldsymbol{a})$ times the cardinality of the set of partial flows — are in bijection with the flows on G that the Kostant partition function $K_G(\boldsymbol{a})$ counts. That is,
$$Q(\boldsymbol{a})\ (\#\text{ partial flows})=K_G(\boldsymbol{a}).$$

From these two statements we see that the two Kostant partition functions $K_G(\boldsymbol{a})$ and $K_{G-(n-1,n)}(\boldsymbol{a})$ are connected by

$$K_G(\boldsymbol{a})=Q(\boldsymbol{a})K_{G-(n-1,n)}(\boldsymbol{a}),$$

which is Equation (2).

Proof of Theorem 1. Let $\{\{\alpha_1, \dots, \alpha_N\}\}$ be the multiset of vectors corresponding to the edges of G. Let $\alpha_N = e_{n-1} - e_n$, $\alpha_{N-1} = e_{n-1} - e_{n+1}$, and $\alpha_{N-2} = e_n - e_{n+1}$. Then (2) can be rewritten as

$$\#\left\{(b_i)_{i\in[N]} \;\middle|\; \sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}\right\} = Q(\boldsymbol{a})\ \#\left\{(b_i)_{i\in[N-1]} \;\middle|\; \sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}\right\}, \tag{4}$$

where $Q(\boldsymbol{a})$ is defined in (3).

We proceed to prove the equality (4) bijectively. The key concept we use is that of a partial flow, which we now define.

Consider a flow $\boldsymbol{f}_H = (b_i)_{i\in[N-3]}$ $(b_i \in \mathbb{Z}_{\geq 0})$ on the edges of the graph

$$H := G - \{(n-1, n),\ (n-1, n+1),\ (n, n+1)\}.$$

We call $\boldsymbol{f}_H$ *partial* if

$$\sum_{i=1}^{N-3} b_i\alpha_i = (a_1, \dots, a_{n-2}, x_{n-1}, x_n, x_{n+1}), \tag{5}$$

for some $x_{n-1}, x_n, x_{n+1} \in \mathbb{Z}$ satisfying $x_{n-1} \leq a_{n-1}$ and $x_n \leq a_n$.

Note that a given partial flow $\boldsymbol{f}_H = (b_i)_{i\in[N-3]}$ can be extended uniquely to a flow $\boldsymbol{f}_{G-\{(n-1,n)\}} = (b_i)_{i\in[N-1]}$ $(b_i \in \mathbb{Z}_{\geq 0})$ on $G - \{(n-1, n)\}$ such that $\sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}$. Furthermore, each such flow arises from a uniquely determined partial flow. Denote by f the map that takes a partial flow $\boldsymbol{f}_H$ into a flow $\boldsymbol{f}_{G-\{(n-1,n)\}}$ as just described. See Figure 1 for an example. The observations above imply:

Claim 1. *The map f establishes a bijection between the set of partial flows $\boldsymbol{f}_H$ on H and the set of flows on $G - \{(n-1, n)\}$ such that $\sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}$. In symbols,*

$$\#\left\{(b_i)_{i\in[N-1]} \;\middle|\; \sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}\right\} = \sum_{\boldsymbol{f}_H} 1, \tag{6}$$

where the summation runs over all partial flows $\boldsymbol{f}_H$.

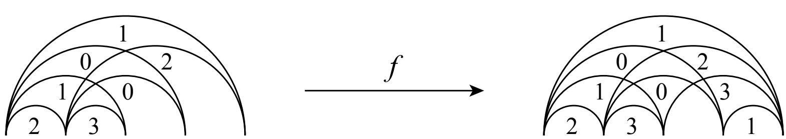

Figure 1. Here G is the complete graph on 5 vertices and $\boldsymbol{a} = (4, 3, -1, 1, -7)$. The flows are written immediately below the corresponding edges. On the left is a partial flow $\boldsymbol{f}_H$ and on the right is its image under f.

Note that the left-hand side of (6) is the same as the cardinality on the right-hand side of (4).

Given a partial flow $\boldsymbol{f}_H$, denote by $Y_i(\boldsymbol{f}_H)$, for $i \in \{n-1, n, n+1\}$, the total *inflow* into vertex $i \in \{n-1, n, n+1\}$ in H, that is, the sum of all the flows b_i on edges of H incident to $i \in \{n-1, n, n+1\}$. Note that a partial flow $\boldsymbol{f}_H$ can be extended in $Y_{n-1}(\boldsymbol{f}_H) + a_{n-1} + 1$ ways to a flow $\boldsymbol{f}_G = (b_i)_{i \in [N]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) of G such that $\sum_{i=1}^N b_i \alpha_i = \boldsymbol{a}$. Furthermore, given a flow $\boldsymbol{f}_G = (b_i)_{i \in [N]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) such that $\sum_{i=1}^N b_i \alpha_i = \boldsymbol{a}$, there is a unique partial flow $\boldsymbol{f}_H = (b_i)_{i \in [N-3]}$ from which it can be obtained. Therefore, the above establishes a map g which is a bijection between the multiset of partial flows $\mathcal{M}$ such that each partial flow $\boldsymbol{f}_H$ appears exactly $Y_{n-1}(\boldsymbol{f}_H) + a_{n-1} + 1$ times in $\mathcal{M}$, and the (multi)set of flows $\boldsymbol{f}_G = (b_i)_{i \in [N]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) on G such that $\sum_{i=1}^N b_i \alpha_i = \boldsymbol{a}$. See Figure 2 for an example. We thus obtain

$$(7) \qquad \#\left\{ (b_i)_{i \in [N]} \;\middle|\; \sum_{i=1}^{N} b_i \alpha_i = \boldsymbol{a} \right\} = \sum_{\boldsymbol{f}_H} \big(Y_{n-1}(\boldsymbol{f}_H) + a_{n-1} + 1 \big) = \#\mathcal{M},$$

where the second summation runs over the set of partial flows $\boldsymbol{f}_H$.

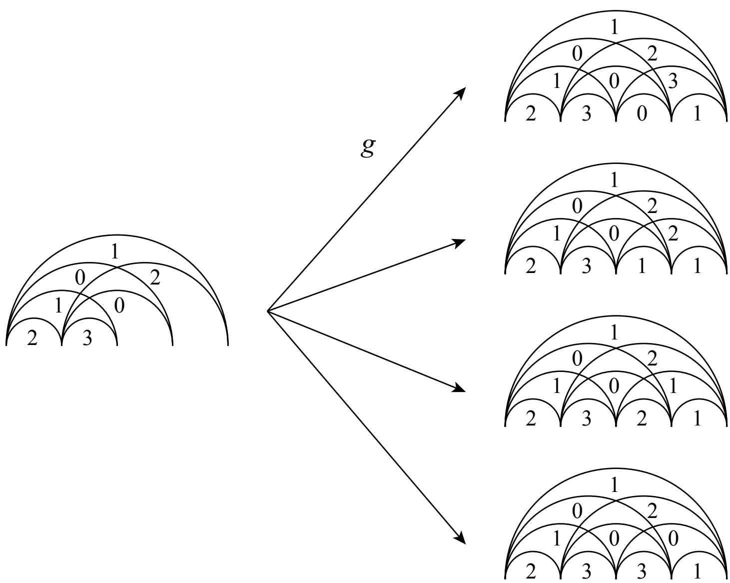

Figure 2. Again G is the complete graph on 5 vertices and $\boldsymbol{a} = (4, 3, -1, 1, -7)$. The flows are written immediately below the corresponding edges. On the left is a partial flow $\boldsymbol{f}_H$ and on the right are the $Y_{n-1}(\boldsymbol{f}_H) + a_{n-1} + 1 = 4$ images under g of the 4 copies of $\boldsymbol{f}_H$ appearing in $\mathcal{M}$.

Claim 2. *The map g just described establishes a bijection between the multiset $\mathcal{M}$ of partial flows $\boldsymbol{f}_H = (b_i)_{i\in[N-3]}$ and the (multi)set of flows $\boldsymbol{f}_G = (b_i)_{i\in[N]}$ $(b_i \in \mathbb{Z}_{\geq 0})$ on G such that $\sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}$. Moreover,*

$$\#\mathcal{M} = Q(\boldsymbol{a}) \sum_{\boldsymbol{f}_H} 1, \tag{8}$$

where the summation runs over all partial flows $\boldsymbol{f}_H$.

We already showed that g is a bijection; there remains to show (8). Now, by assumption we have

$$\frac{m_{j,n-1} + m_{j,r} + m_{j,n+1}}{m_{j,n-1}} = c, \tag{9}$$

where c is independent of $j \in [n-2]$; hence

$$\begin{aligned} c\sum_{\boldsymbol{f}_H} Y_{n-1}(\boldsymbol{f}_H) &= \sum_{\boldsymbol{f}_H}\bigl(Y_{n-1}(\boldsymbol{f}_H) + Y_n(\boldsymbol{f}_H) + Y_{n+1}(\boldsymbol{f}_H)\bigr) \\ &= \sum_{\boldsymbol{f}_H}(a_1 + \cdots + a_{n-2}), \end{aligned} \tag{10}$$

that is,

$$\sum_{\boldsymbol{f}_H} Y_{n-1}(\boldsymbol{f}_H) = \sum_{\boldsymbol{f}_H} \frac{a_1 + \cdots + a_{n-2}}{c}. \tag{11}$$

Adding $a_{n-1}+1$ to both summands and recalling (3) and the second equality in (7) yields (8), completing the proof of Claim 2.

In light of (4) it is clear that Claims 1 and 2 together provide a bijective proof of Theorem 1. In terms of equations this can also be seen as follows. Rewrite (7) as

$$\begin{aligned} \#\left\{(b_i)_{i\in[N]} \;\middle|\; \sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}\right\} &= Q(\boldsymbol{a})\sum_{\boldsymbol{f}_H} 1 \\ &= Q(\boldsymbol{a})\,\#\left\{(b_i)_{i\in[N-1]} \;\middle|\; \sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}\right\}, \end{aligned} \tag{12}$$

where the first equality uses (7) and (8), and the second equality uses (6). □

Further Kostant partition function identities. Several generalizations are suggested by our proof of the Baldoni–Vergne identities. Here we present an immediate one, Theorem 2; to ensure clarity we include its proof, which follows that of Theorem 1. For brevity, we do not explicitly state the analogues of Claims 1 and 2 from the proof of Theorem 1, though it is easy to obtain them from our proof. In the next section we present further generalizations of other types, but we omit the proofs.

Theorem 2. *Let G be a connected graph on the vertex set $[n+1]$. Given $k \le n$, suppose the graph $S = (V(S), E(S))$ defined by*

$$V(S) = \{k, k+1, \dots, n, n+1\} \subset [n+1]$$

and

$$E(S) = \{(i, j) \in E(G) \mid i < j,\ i \in V(S)\}$$

satisfies these conditions: that

- $\mathrm{outdeg}_S(j) = 2$ *and* $\mathrm{indeg}_S(j) = 0$, *for some $k \le j \le n$, and* $\mathrm{outdeg}_S(i) = 1$ *for $k \le i \le n$, $i \ne j$.*
- *For some constant c independent of l,*

$$\frac{\sum_{i=k}^{n+1} m_{l,i}}{m_{l,j}} = c \quad \text{for all } l \in [k-1].$$

Let (j, z) be one of the outgoing edges from j in S. Then

$$K_G(\boldsymbol{a}) = Q'(\boldsymbol{a}) K_{G-(j,z)}(\boldsymbol{a}), \tag{13}$$

for any $\boldsymbol{a} = \left(a_1, \dots, a_n, -\sum_{i=1}^n a_i\right) \in \mathbb{Z}^{n+1}$, where $Q'(\boldsymbol{a}) = \dfrac{a_1 + \dots + a_{k-1}}{c} + a_j + 1$.

The most important cases of Theorem 2 are for $k \ge n-1$, since for $k < n-1$ several of the edges of G can be "contracted" and reduced to the case $k \ge n-1$. For $k \ge n-1$ we obtain four interesting cases from Theorem 2 depending on the form of the graph S and the edge (j, z). If we take

$$V(S) = \{n-1, n, n+1\} \quad \text{and} \quad E(S) = \{(n-1, n),\ (n-1, n+1),\ (n, n+1)\}$$

and $(j, z) = (n-1, n)$, then Theorem 2 specializes to the original Baldoni–Vergne identities, while other choices of S and (j, z) lead to new identities.

Proof of Theorem 2. Let $\{\{\alpha_1, \dots, \alpha_N\}\}$ be the multiset of vectors corresponding to the edges of G such that the multiset of vectors corresponding to the edges of S are $\{\{\alpha_N - (n+1-k), \dots, \alpha_N\}\}$. Also, let α_N correspond to the edge (j, z).

Consider a flow $\boldsymbol{f}_H = (b_i)_{i \in [N-(n+2-k)]}$ $(b_i \in \mathbb{Z}_{\ge 0})$ of the edges of the graph $H := ([n+1], E(G) \backslash E(S))$. We call $\boldsymbol{f}_H$ *partial* if

$$\sum_{i=1}^{N-(n+2-k)} b_i \alpha_i = (a_1, \dots, a_{k-1}, x_k, x_{k+1}, \dots, x_{n+1}),$$

for some $x_i \in \mathbb{Z}$, $k \le i \le n+1$ satisfying $x_i \le a_i$, for $k \le i \le n$.

A given partial flow $\boldsymbol{f}_H = (b_i)_{i\in[N-(n+2-k)]}$, it can be extended uniquely to a flow $\boldsymbol{f}_{G-\{(j,z)\}} = (b_i)_{i\in[N-1]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) on $G-\{(j,z)\}$ such that $\sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}$. Furthermore, this correspondence is a bijection. Therefore,

$$\#\left\{(b_i)_{i\in[N-1]} \;\middle|\; \sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}\right\} = \sum_{f_H} 1, \tag{14}$$

where the summation runs over all partial flows $\boldsymbol{f}_H$.

Given a partial flow $\boldsymbol{f}_H$ in H, we denote by $Y_i(\boldsymbol{f}_H)$, where $i\in\{k, k+1, \dots, n+1\}$, the total *inflow* into vertex i, that is, the sum of all the flows b_i on edges of H incident to i. The partial flow $\boldsymbol{f}_H$ can be extended in $Y_j(\boldsymbol{f}_H)+a_j+1$ ways to a flow $\boldsymbol{f}_G = (b_i)_{i\in[N]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) of G such that $\sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}$. Furthermore, given a flow $\boldsymbol{f}_G = (b_i)_{i\in[N]}$ ($b_i \in \mathbb{Z}_{\geq 0}$) such that $\sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}$, there is a unique partial flow $\boldsymbol{f}_H = (b_i)_{i\in[N-(n+2-k)]}$ from which it can be obtained. Therefore,

$$\#\left\{(b_i)_{i\in[N]} \;\middle|\; \sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}\right\} = \sum_{f_H}\big(Y_j(\boldsymbol{f}_H)+a_j+1\big), \tag{15}$$

where the summation runs over all partial flows $\boldsymbol{f}_H$.

By assumption, we have

$$\frac{\sum\limits_{i=k}^{n+1} m_{l,i}}{m_{l,j}} = c,$$

where c is independent of $l \in [k-1]$; hence

$$c\sum_{f_H} Y_j(\boldsymbol{f}_H) = \sum_{f_H}\sum_{i=k}^{n+1} Y_i(\boldsymbol{f}_H) = \sum_{f_H}(a_1+\cdots+a_{k-1}),$$

that is,

$$\sum_{f_H} Y_j(\boldsymbol{f}_H) = \sum_{f_H}\frac{a_1+\cdots+a_{k-1}}{c}. \tag{16}$$

Thus, (15) can be rewritten as

$$\begin{aligned}\#\left\{(b_i)_{i\in[N]} \;\middle|\; \sum_{i=1}^{N} b_i\alpha_i = \boldsymbol{a}\right\} &= \sum_{f_H}\left(\frac{a_1+\cdots+a_{k-1}}{c}+a_j+1\right)\\ &= Q'(\boldsymbol{a})\sum_{f_H} 1 = Q'(\boldsymbol{a})\,\#\left\{(b_i)_{i\in[N-1]} \;\middle|\; \sum_{i=1}^{N-1} b_i\alpha_i = \boldsymbol{a}\right\},\end{aligned}$$

where the first equality uses (15) and (16), and the last equality uses (14). □

3. Type C_{n+1} Kostant partition functions and the Baldoni–Vergne identities

We now show two generalizations of Theorem 1 in the type C_{n+1} case. We first give the necessary definitions and explain the notion of flow in the context of signed graphs. Throughout this section, the graphs G on the vertex set $[n+1]$ we consider are signed, that is there is a sign $\epsilon \in \{+, -\}$ assigned to each of its edges, with possible multiple edges, and all loops labeled positive. Denote by $(i, j, -)$ and $(i, j, +)$, $i \leq j$, a negative and a positive edge, respectively. Denote by m_{ij}^{ϵ} the multiplicity of edge (i, j, ϵ) in G, $i \leq j$, $\epsilon \in \{+, -\}$. To each edge (i, j, ϵ), $i \leq j$, of G, associate the positive type C_{n+1} root $\mathrm{v}(i, j, \epsilon)$, where $\mathrm{v}(i, j, -) = e_i - e_j$ and $\mathrm{v}(i, j, +) = e_i + e_j$. Let $\{\{\alpha_1, \ldots, \alpha_N\}\}$ be the multiset of vectors corresponding to the multiset of edges of G as described above. Note that $N = \sum_{1 \leq i < j \leq n+1} (m_{ij}^{-} + m_{ij}^{+})$. The *Kostant partition function* K_G evaluated at the vector $\boldsymbol{a} \in \mathbb{Z}^{n+1}$ is defined as

$$K_G(\boldsymbol{a}) = \#\left\{ (b_i)_{i \in [N]} \;\middle|\; \sum_{i \in [N]} b_i \alpha_i = \boldsymbol{a} \text{ and } b_i \in \mathbb{Z}_{\geq 0} \right\}.$$

That is, $K_G(\boldsymbol{a})$ is the number of ways to write the vector $\boldsymbol{a}$ as a nonnegative linear combination of the positive type C_{n+1} roots corresponding to the edges of G, without regard to order.

Just like in the type A_n case, we would like to think of the vector $(b_i)_{i \in [N]}$ as a *flow*. For this we here give a precise definition of flows in the type C_{n+1} case, of which type A_n is of course a special case.

Let G be a signed graph on the vertex set $[n+1]$. Let $\{\{e_1, \ldots, e_N\}\}$ be the multiset of edges of G, and $\{\{\alpha_1, \ldots, \alpha_N\}\}$ the multiset of vectors corresponding to the multiset of edges of G. Fix an integer vector $\boldsymbol{a} = (a_1, \ldots, a_n, a_{n+1}) \in \mathbb{Z}^{n+1}$. A *nonnegative integer $\boldsymbol{a}$-flow $\boldsymbol{f}_G$* on G is a vector $\boldsymbol{f}_G = (b_i)_{i \in [N]}$ $(b_i \in \mathbb{Z}_{\geq 0})$ such that for all $1 \leq i \leq n+1$, we have

$$\sum_{\substack{e \in E \\ \mathrm{inc}(e,v) = -}} b(e) + a_v = \sum_{\substack{e \in E \\ \mathrm{inc}(e,v) = +}} b(e) + \sum_{e = (v,v,+)} b(e), \tag{17}$$

where $b(e_i) = b_i$, $\mathrm{inc}(e, v) = -$ if edge $e = (g, v, -)$, $g < v$, and $\mathrm{inc}(e, v) = +$ if $e = (g, v, +)$, $g < v$, or $e = (v, j, \epsilon)$, $v < j$, and $\epsilon \in \{+, -\}$.

Call $b(e)$ the *flow* assigned to edge e of G. If the edge e is negative, one can think of $b(e)$ units of fluid flowing on e from its smaller to its bigger vertex. If the edge e is positive, then one can think of $b(e)$ units of fluid flowing away both from e's smaller and bigger vertex to infinity. Edge e is then a "leak" taking away $2b(e)$ units of fluid.

From the above explanation it is clear that if we are given an $\boldsymbol{a}$-flow $\boldsymbol{f}_G$ such that $\sum_{e=(i,j,+), i \leq j} b(e) = y$, then $\boldsymbol{a} = \big(a_1, \ldots, a_n, 2y - \sum_{i=1}^{n} a_i\big)$. It is then a matter of checking the definitions to see that for a signed graph G on the vertex set $[n+1]$ and vector $\boldsymbol{a} = \big(a_1, \dots, a_n, 2y - \sum_{i=1}^{n} a_i\big) \in \mathbb{Z}^{n+1}$, the number of nonnegative integer $\boldsymbol{a}$-flows on G is equal to $K_G(\boldsymbol{a})$.

Thinking of $K_G(\boldsymbol{a})$ as the number of nonnegative integer $\boldsymbol{a}$-flows on G, there is a straightforward generalization of Theorem 1 in the type C_{n+1} case:

Theorem 3. *Given a connected signed graph G on the vertex set $[n+1]$ with $m^-_{n-1,n} = m^-_{n-1,n+1} = m^-_{n,n+1} = 1$, $m^+_{j,n-1} = m^+_{j,n} = m^+_{j,n+1} = 0$, for $j \in [n+1]$, and such that*

$$\frac{m^-_{j,n-1} + m^-_{j,n} + m^-_{j,n+1}}{m^-_{j,n-1}} = c \quad \textit{for all } \ j \in [n-2],$$

for some constant c independent of j, we have

$$K_G(\boldsymbol{a}) = Q''(\boldsymbol{a}) K_{G-(n-1,n)}(\boldsymbol{a}) \tag{18}$$

for any $\boldsymbol{a} = \big(a_1, \dots, a_n, 2y - \sum_{i=1}^{n} a_i\big) \in \mathbb{Z}^{n+1}$, where

$$Q''(\boldsymbol{a}) = \frac{a_1 + \dots + a_{n-2} - 2y}{c} + a_{n-1} + 1.$$

The proof of Theorem 3 proceeds analogously to that of Theorem 1. Namely, define *partial flows* $\boldsymbol{f}_H = (b_i)_{i \in [N-3]}$ on

$$H := G - \{(n-1, n, -), (n-1, n+1, -), (n, n+1, -)\}$$

such that

$$\sum_{i=1}^{N-3} b_i \alpha_i = (a_1, \dots, a_{n-2}, x_{n-1}, x_n, x_{n+1}),$$

for some $x_{n-1}, x_n, x_{n+1} \in \mathbb{Z}$, such that $x_{n-1} \le a_{n-1}$, $x_n \le a_n$ and the sum of flows on positive edges is y.

Then, one can prove:

- The elements of the set partial flows are in bijection with the nonnegative integer $\boldsymbol{a}$-flows on $G - (n-1, n)$. That is,

$$\#\text{ partial flows} = K_{G-(n-1,n)}(\boldsymbol{a}).$$

- The elements of the multiset of partial flows $\boldsymbol{f}_H$, whose cardinality is $Q''(\boldsymbol{a})$ times the cardinality of the set of partial flows, are in bijection with the nonnegative integer $\boldsymbol{a}$-flows on G. That is,

$$Q''(\boldsymbol{a})\ (\#\text{ partial flows}) = K_G(\boldsymbol{a}).$$

Thus,

$$K_G(\boldsymbol{a}) = Q''(\boldsymbol{a}) K_{G-(n-1,n)}(\boldsymbol{a}).$$

We do not have to require that only negative edges are incident to the vertices $n-1, n, n+1$ in G, as the following theorem shows. The proof is analogous to earlier ones.

Theorem 4. *Given a connected signed graph G on the vertex set $[n+1]$ with $m^-_{n-1,n} = m^-_{n-1,n+1} = m^-_{n,n+1} = 1$, $m^+_{i,j} = 0$, for $i, j \in \{n-1, n, n+1\}$, and such that*

$$\frac{m^{\epsilon}_{j,n-1} + m^{\epsilon}_{j,r} + m^{\epsilon}_{j,n+1}}{m^{\epsilon}_{j,n-1}} = c \quad \text{for all } j \in [n-2] \text{ and } \epsilon \in \{+,-\},$$

for some constant c independent of j, we have

$$K_G(\boldsymbol{a}) = \left(\frac{a_1 + \cdots + a_{n-2} - 2y}{c} + a_{n-1} + 1\right) K_{G-(n-1,n)}(\boldsymbol{a}) \tag{19}$$

for any $\boldsymbol{a} = \left(a_1, \ldots, a_n, 2y - \sum_{i=1}^n a_i\right) \in \mathbb{Z}^{n+1}$.

Of course, Theorem 2 also has type C_{n+1} generalizations and variations. We invite the reader to write these out and check each step of the proof of Theorem 1 and see how they can be adapted.

Acknowledgement

I thank Richard Stanley for his intriguing slides and for pointing out the work of Baldoni and Vergne. I also thank Alejandro Morales and Alex Postnikov for numerous conversations about flows and the Kostant partition function.

References

[Baldoni and Vergne 2008] W. Baldoni and M. Vergne, "Kostant partitions functions and flow polytopes," *Transform. Groups* **13**:3-4 (2008), 447–469. MR 2009k:52029 Zbl 1200.52008

[Stanley 2000] R. P. Stanley, "Acyclic flow polytopes and Kostant's partition function," conference transparencies, 2000, Available at http://www-math.mit.edu/~rstan/transparencies/kostant.ps.

Received November 27, 2011. Revised April 4, 2012.

KAROLA MÉSZÁROS
DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NY 14853
UNITED STATES

karola@math.cornell.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.227

EXCEPTIONAL LIE ALGEBRAS, SU(3), AND JORDAN PAIRS

PIERO TRUINI

A simple unifying view of the exceptional Lie algebras is presented. The underlying Jordan pair content and role are exhibited. Each algebra contains three Jordan pairs sharing the same Lie algebra of automorphisms and the same external su(3) symmetry. Eventual physical applications and implications of the theory are outlined.

1. Introduction

The main purpose of this paper is to exhibit a unifying view of all exceptional Lie algebras, which is also very intuitive from the point of view of elementary particle physics. The result is represented by the root diagram in Figure 1.

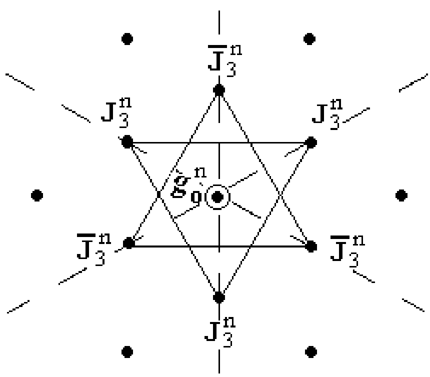

Figure 1. A unifying view of the exceptional Lie algebra roots.

It is a very simple, highly intuitive unifying view of all exceptional Lie algebras and we will use it repeatedly to unfold the largest algebra $\mathbf{e_8}$. The picture shows the projection of the roots of the exceptional Lie algebras on a $\mathbf{su(3)}$ plane, recognizable by the dots forming the external hexagon, and it exhibits the Jordan pair content of

MSC2010: primary 17B25; secondary 17C40.
Keywords: exceptional Lie algebras, e_8, Jordan pairs, magic square.

each exceptional Lie algebra. There are three Jordan pairs $(\mathbf{J_3^n}, \mathbf{\bar{J}_3^n})$, each of which lies on an axis symmetrically with respect to the center of the diagram. Each pair doubles a Jordan algebra $\mathbf{J_3^n}$ with involution (the conjugate representation $\mathbf{\bar{J}_3^n}$), which is the algebra of 3×3 Hermitian matrices over $\mathbf{H}$, where $\mathbf{H} = \mathbf{R}, \mathbf{C}, \mathbf{Q}, \mathfrak{C}$ (real, complex, quaternion, and octonion numbers) for $\mathbf{n} = 1, 2, 4, 8$. We get $\mathbf{f_4}$, $\mathbf{e_6}$, $\mathbf{e_7}$, and $\mathbf{e_8}$ for $\mathbf{n} = 1, 2, 4, 8$, respectively ($\mathbf{g_2}$ can be also represented in the same way, with the Jordan algebra reduced to a single element). The three Jordan algebras (and their conjugates) globally behave like a $\mathbf{3}$ (and a $\mathbf{\bar{3}}$)-dimensional representation of the outer $\mathbf{su(3)}$. The algebra denoted by $\mathbf{g_0^n}$ in the center (plus the Cartan generator associated with the axis along which the pair lies) is the algebra of the automorphism group of the Jordan pair (the structure group of the corresponding Jordan algebra). In the case of $\mathbf{e_8}$, the algebra $\mathbf{g_0^8}$ is $\mathbf{e_6}$, described by a similar diagram, and we can iterate the process. What we eventually end up with is a decomposition of $\mathbf{e_8}$ entirely given in terms of $\mathbf{su(3)}$'s and Jordan pairs (that we associate to particle–antiparticle pairs): three pairs $(\mathbf{J_3^8}, \mathbf{\bar{J}_3^8})$ for three colors of quark–antiquarks, plus three pairs $(\mathbf{J_3^2}, \mathbf{\bar{J}_3^2})$, in the colorless $\mathbf{g_0^8} = \mathbf{e_6}$, for three families of leptons–antileptons.

The interest of physicists in the exceptional Lie algebras, and $\mathbf{e_8}$ in particular, is a long-standing tradition, starting from the pioneering work of Gürsey [Frampton et al. 1980] on grand unification, and continuing with [Green and Schwarz 1984; Cremmer 1982; Truini and Biedenharn 1982; Candelas et al. 1985; Gross 1986; Ferrara and Kallosh 1996a; 1996b; Ferrara and Günaydin 1998]. In the effort of unifying all interactions in a consistent quantum theory that includes gravity, the most successful model of string theory is based on $\mathbf{e_8}$; an alternative theory known as *loop quantum gravity* (see [Rovelli 2004] for an excellent and comprehensive review) has also led towards the exceptional algebras, and $\mathbf{e_8}$ in particular [Lisi et al. 2010].

There is a wide consensus in both mathematics and physics on the appeal of the largest exceptional Lie algebra $\mathbf{e_8}$, considered beautiful by many in spite of its complexity. The best synthesis of this was stated by B. Kostant:[1] "It is easy to arrive at the feeling that a final understanding of the universe must somehow involve $E(8)$, or otherwise put, nature would be foolish not to utilize $E(8)$."

Kostant defines $\mathbf{e_8}$ as "a symphony of 2, 3, 5." In the more modest view of the exceptional algebras I present here the numbers 1, 2, and 3 play the central role: they govern the structure. Number 1 is the whole, the universe of the theory: a Lie algebra. Number 2 stands for pair, and we view it as a particle–antiparticle duality represented by Jordan pairs. Number 3 is the number of colors and the number of families: each Jordan pair occurs three times, in a $\mathbf{su(3)}$ symmetry. That is all you need in order to build $\mathbf{e_8}$, as we are going to show.

[1] Quoted by Benjamin Wallace-Wells in "Surfing the Universe", *The New Yorker*, 21 July 2008.

2. Jordan pairs

In this section we review the concept of a Jordan pair [Loos 1975] (see also [McCrimmon 2004] for an enlightening overview). Jordan algebras have traveled a long journey since their appearance in the 30s [Jordan et al. 1934]. The modern formulation [Jacobson 1966] involves a quadratic map $U_x y$ (like xyx for associative algebras) instead of the original symmetric product $x \circ y = xy + yx$. The quadratic map and its linearization $V_{x,y}z = (U_{x+z} - U_x - U_z)y$ (like $xyz + zyx$ in the associative case) reveal the mathematical structure of Jordan algebras much more clearly, through the notion of inverse, inner ideal, generic norm, etc. The axioms are:

$$U_1 = \mathrm{Id.}, \qquad U_x V_{y,x} = V_{x,y} U_x, \qquad U_{U_x y} = U_x U_y U_x. \tag{2-1}$$

The quadratic formulation led to the concept of Jordan triple systems [Meyberg 1970], an example of which is a pair of modules represented by rectangular matrices. There is no way of multiplying two matrices x and y, say $n \times m$ and $m \times n$, respectively, by means of a bilinear product. But one can do it using a product like xyx, quadratic in x and linear in y. Notice that, like in the case of rectangular matrices, there needs not be a unity in these structures. The axioms are in this case:

$$U_x V_{y,x} = V_{x,y} U_x, \qquad V_{U_x y, y} = V_{x, U_y x}, \qquad U_{U_x y} = U_x U_y U_x. \tag{2-2}$$

Finally, a Jordan pair is just a pair of modules (V^+, V^-) acting on each other (but not on themselves) like a Jordan triple:

$$\begin{aligned} U_{x^\sigma} V_{y^{-\sigma}, x^\sigma} &= V_{x^\sigma, y^{-\sigma}} U_{x^\sigma}, \\ V_{U_{x^\sigma} y^{-\sigma}, y^{-\sigma}} &= V_{x^\sigma, U_{y^{-\sigma}} x^\sigma}, \\ U_{U_{x^\sigma} y^{-\sigma}} &= U_{x^\sigma} U_{y^{-\sigma}} U_{x^\sigma}, \end{aligned} \tag{2-3}$$

where $\sigma = \pm$ and $x^\sigma \in V^{+\sigma}$, $y^{-\sigma} \in V^{-\sigma}$.

Jordan pairs are strongly related to the Tits–Kantor–Koecher construction of Lie algebras [Tits 1962; Kantor 1964; Koecher 1967] (see also the interesting relation to Hopf algebras [Faulkner 2000]):

$$\mathfrak{L} = J \oplus \mathrm{str}(J) \oplus \bar{J}, \tag{2-4}$$

where J is a Jordan algebra and $\mathrm{str}(J) = L(J) \oplus \mathrm{Der}(J)$ is the structure algebra of J [McCrimmon 2004]; $L(x)$ is the left multiplication in J: $L(x)y = x \circ y$; and $\mathrm{Der}(J) = [L(J), L(J)]$ is the algebra of derivations of J (the algebra of the automorphism group of J) [Schafer 1949; 1966].

In the case of (complex) exceptional Lie algebras this construction applies to $\mathbf{e_7}$, with $J = \mathbf{J_3^8}$, the 27-dimensional exceptional Jordan algebra of 3×3 Hermitian matrices over the octonions, and $\mathrm{str}(J) = \mathbf{e_6} \otimes \mathbf{C}$ (where $\mathbf{C}$ is the complex field). The algebra $\mathbf{e_6}$ is called the *reduced structure algebra* of J, $\mathrm{str}_0(J)$, which is namely the structure algebra with the generator corresponding to multiplication by a complex number taken away: $\mathbf{e_6} = L(J_0) \oplus \mathrm{Der}(J)$, with J_0 denoting the traceless elements of J.

The Tits–Kantor–Koecher construction can be generalized as follows: if $\mathfrak{L}$ is a three-graded Lie algebra,

$$\mathfrak{L} = \mathfrak{L}_{-1} \oplus L_0 \oplus \mathfrak{L}_1, \qquad [\mathfrak{L}_i, \mathfrak{L}_j] \subset \mathfrak{L}_{i+j}, \tag{2-5}$$

so that $[\mathfrak{L}_i, \mathfrak{L}_j] = 0$ whenever $|i + j| > 1$, then $(\mathfrak{L}_1, \mathfrak{L}_{-1})$ forms a Jordan pair, with the Jacobi identity forcing the elements of the pair to act on each other like in a Jordan triple system. The link with the Tits–Kantor–Koecher construction is obtained by letting $J = \mathfrak{L}_1$ and $\bar{J} = \mathfrak{L}_{-1}$. The structure group of J is the automorphism group of the Jordan pair $(J, \bar{J})$ and the trilinear product $V_{x^\sigma, y^{-\sigma}} z^\sigma$ is

$$V_{x^\sigma, y^{-\sigma}} z^\sigma = [[x^\sigma, y^{-\sigma}], z^\sigma].$$

3. The Freudenthal–Tits magic square

The theory of exceptional Lie algebras has had a major advance with the development of two related objects: the Tits construction and the Freudenthal–Tits magic square [Tits 1955; Freudenthal 1959].

The Freudenthal–Tits magic square is a table of Lie algebras related to both Jordan algebras and Hurwitz algebras $\mathbf{H}$, namely the algebras of real ($\mathbf{R}$), complex ($\mathbf{C}$), quaternion ($\mathbf{Q}$), and octonion or Cayley ($\mathfrak{C}$) numbers. In particular the Jordan algebras involved in the magic square are the algebras of 3×3 Hermitian matrices over $\mathbf{H}$:

$$\begin{pmatrix} \alpha & a & \bar{b} \\ \bar{a} & \beta & c \\ b & \bar{c} & \gamma \end{pmatrix}, \qquad \alpha, \beta, \gamma \in \mathbf{C}; \quad a, b, c \in \mathbf{H}. \tag{3-1}$$

We denote them by $\mathbf{J_3^n}$ where $\mathbf{n} = 1, 2, 4, 8$ for $\mathbf{H} = \mathbf{R}, \mathbf{C}, \mathbf{Q}, \mathfrak{C}$, respectively. In this paper only complex Lie algebras are considered. Therefore each algebra $\mathbf{H}$ is over the complex field and the $a \to \bar{a}$ conjugation in (3-1) changes the signs of the imaginary units of $\mathbf{H}$ but does not conjugate the imaginary unit of the underlying complex field. The Freudenthal–Tits magic square is shown in Table 1.

The way the magic square is constructed is due to Tits:

$$\mathfrak{L} = \mathrm{Der}(\mathbf{H}) \oplus (\mathbf{H}_0 \otimes \mathbf{J_0}) \oplus \mathrm{Der}(\mathbf{J}). \tag{3-2}$$

Here the subscript 0 stands for traceless. $\mathrm{Der}(\mathbf{H})$ is the algebra of derivations of $\mathbf{H}$, which is nothing for $\mathbf{H} = \mathbf{R}, \mathbf{C}$, whereas $\mathrm{Der}(\mathbf{Q}) = \mathbf{a_1}$ and $\mathrm{Der}(\mathfrak{C}) = \mathbf{g_2}$.

$\mathbf{H}\setminus\mathbf{J}$	$\mathbf{J}_3^1$	$\mathbf{J}_3^2$	$\mathbf{J}_3^4$	$\mathbf{J}_3^8$
$\mathbf{R}$	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{c_3}$	$\mathbf{f_4}$
$\mathbf{C}$	$\mathbf{a_2}$	$\mathbf{a_2}\oplus\mathbf{a_2}$	$\mathbf{a_5}$	$\mathbf{e_6}$
$\mathbf{Q}$	$\mathbf{c_3}$	$\mathbf{a_5}$	$\mathbf{d_6}$	$\mathbf{e_7}$
$\mathfrak{C}$	$\mathbf{f_4}$	$\mathbf{e_6}$	$\mathbf{e_7}$	$\mathbf{e_8}$

Table 1. The Freudenthal–Tits magic square.

We also have the following tight link between the entries of the magic square and Jordan structures.

The Lie algebras $\mathbf{g}_{\mathrm{I}}$ in the first row of the magic square are the algebras of derivations of the Jordan algebra in the same column (the corresponding group is the automorphism group of the Jordan algebra):

$$\mathbf{g}_{\mathrm{I}} = \mathrm{Der}(\mathbf{J}), \quad \text{namely:}$$

$$\mathbf{a_1} = \mathrm{Der}(\mathbf{J}_3^1), \quad \mathbf{a_2} = \mathrm{Der}(\mathbf{J}_3^2), \quad \mathbf{c_3} = \mathrm{Der}(\mathbf{J}_3^4), \quad \mathbf{f_4} = \mathrm{Der}(\mathbf{J}_3^8).$$

The Lie algebras $\mathbf{g}_{\mathrm{II}}$ in the second row are the reduced structure algebras of the Jordan algebra in the same column

$$\mathbf{g}_{\mathrm{II}} = \mathrm{str}_0(\mathbf{J}), \quad \text{namely:}$$

$$\mathbf{a_2} = \mathrm{str}_0(\mathbf{J}_3^1), \quad \mathbf{a_2}\oplus\mathbf{a_2} = \mathrm{str}_0(\mathbf{J}_3^2), \quad \mathbf{a_5} = \mathrm{str}_0(\mathbf{J}_3^4), \quad \mathbf{e_6} = \mathrm{str}_0(\mathbf{J}_3^8).$$

The Lie algebras $\mathbf{g}_{\mathrm{III}}$ in the third row are three graded and can be written via the Tits–Kantor–Koecher construction (2-4) or in terms of generalized 2×2 matrices [Truini et al. 1986]:

$$\mathbf{g}_{\mathrm{III}} = \mathbf{J} \oplus (\mathbf{g}_{\mathrm{II}} \otimes \mathbf{C}) \oplus \bar{\mathbf{J}}, \quad \text{namely:}$$

$$\mathbf{c_3} = \mathbf{J}_3^1 \oplus (\mathbf{a_2}\oplus\mathbf{C}) \oplus \bar{\mathbf{J}}_3^1,\ \mathbf{a_5} = \mathbf{J}_3^2 \oplus (\mathbf{a_2}\oplus\mathbf{a_2}\oplus\mathbf{C}) \oplus \bar{\mathbf{J}}_3^2,$$
$$\mathbf{d_6} = \mathbf{J}_3^4 \oplus (\mathbf{a_5}\oplus\mathbf{C}) \oplus \bar{\mathbf{J}}_3^4,\ \mathbf{e_7} = \mathbf{J}_3^8 \oplus (\mathbf{e_6}\oplus\mathbf{C}) \oplus \bar{\mathbf{J}}_3^8.$$

In our opinion the most natural way of extending a similar analysis to the fourth row is the one described in this paper and represented in Figure 1 or in the expression (4-1) in the next section.

4. The Jordan pairs inside the exceptional Lie algebras

In this section we work with the roots of the exceptional Lie algebras and postpone the discussion on explicit representations of the generators to the next section. The notation for the explicit set of roots we use, [Bourbaki 1968] is shown in Table 2 in the Appendix.

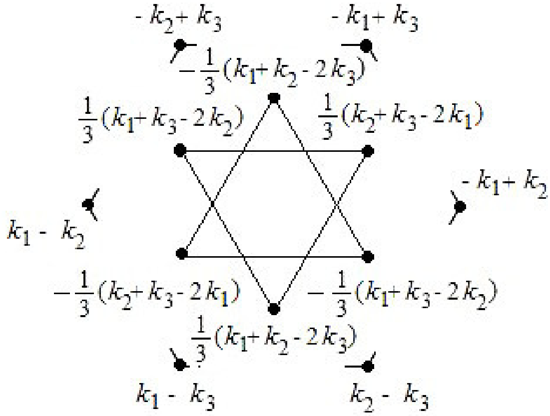

Figure 2. Roots of $\mathbf{g_2}$.

The roots can be placed, case by case, as in Figure 1, where they are shown by their projections on the plane of an $\mathbf{a_2}$ subalgebra (we use the standard notation $\mathbf{a_2}$ for the complexification of $\mathbf{su}(\mathbf{3})$). Notice that $\mathbf{g_2}$ itself, as shown in Figure 2, has a root diagram represented by the same dots appearing in Figure 1.

The root diagrams of $\mathbf{f_4}$, $\mathbf{e_6}$, $\mathbf{e_7}$, and $\mathbf{e_8}$ are as in Figure 3. The notation for the Jordan algebras in the figure is the same used in Table 1 for the Freudenthal–Tits magic square: $\mathbf{J_3^n}$, $\mathbf{n} = 1, 2, 4, 8$ is the Jordan algebra of 3×3 Hermitian matrices over $\mathbf{R}$, $\mathbf{C}$, $\mathbf{Q}$, and $\mathfrak{C}$ respectively. The algebra $\mathbf{g_0^n}$, $\mathbf{n} = 1, 2, 4, 8$, is $\mathbf{a_2}$, $\mathbf{a_2} \oplus \mathbf{a_2}$, $\mathbf{a_5}$, and $\mathbf{e_6}$, respectively; $\mathbf{g_0^n} \oplus \mathbf{C}$ is the algebra of the automorphism group of each Jordan pair $\mathbf{V^n} = (\mathbf{J_3^n}, \mathbf{\bar{J}_3^n})$. We associate roots to Jordan pairs and we check that

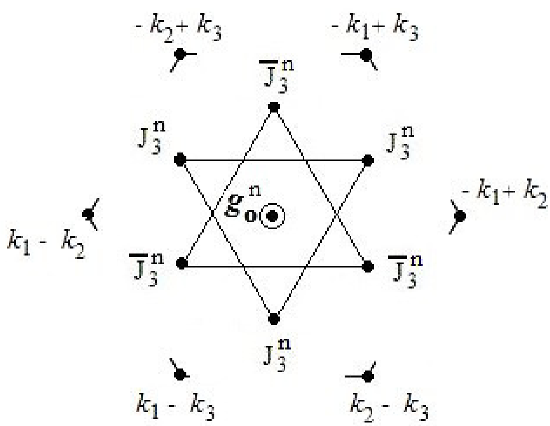

Figure 3. Roots of $\mathbf{f_4}$, $\mathbf{e_6}$, $\mathbf{e_7}$, and $\mathbf{e_8}$, for $n = 1, 2, 4, 8$, respectively, projected on the plane Π.

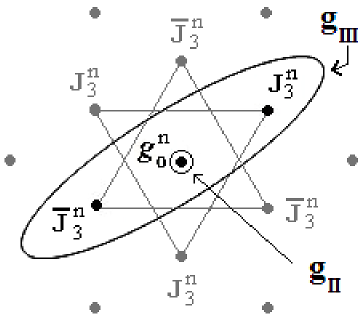

Figure 4. Roots of the exceptional Lie algebras with $\mathbf{g}_{\mathrm{II}}$ and $\mathbf{g}_{\mathrm{III}}$ highlighted.

the projection of these roots lie along an axis, symmetrically with respect to the center. The $\mathbf{C}$ in $\mathbf{g}_0^\mathbf{n} \oplus \mathbf{C}$ stands for the complex field and represents the action on $\mathbf{V}^\mathbf{n}$ (multiplication by a complex number) of the Cartan generator associated with that axis.

If we write $\mathfrak{L}^n$ for $\mathbf{f_4}$, $\mathbf{e_6}$, $\mathbf{e_7}$, and $\mathbf{e_8}$, $n = 1, 2, 4, 8$, we get the unifying expression

$$\mathfrak{L}^n = \mathbf{a_2} \oplus \mathbf{g}_0^\mathbf{n} \oplus 3 \times (\mathbf{J}_3^\mathbf{n}, \bar{\mathbf{J}}_3^\mathbf{n}), \quad \text{where} \quad \mathbf{g}_0^\mathbf{n} = \mathrm{str}_0(\mathbf{J}_3^\mathbf{n}). \tag{4-1}$$

This is not only a unifying view of the exceptional Lie algebras, but also, in our opinion, a natural way of looking at the fourth row of the magic square. Notice that $\mathbf{g}_0^\mathbf{n}$ is the Lie algebra in the second row ($\mathbf{g}_{\mathrm{II}}$), at the same column of $\mathfrak{L}^n$ and that $\mathbf{g}_0^\mathbf{n} \oplus \mathbf{C} \oplus \mathbf{V}^\mathbf{n}$ is the Lie algebra in the third row ($\mathbf{g}_{\mathrm{III}}$), same column, for any of the three Jordan pairs $\mathbf{V}^\mathbf{n}$ in $\mathfrak{L}^n$ (Figure 4).

We explicitly show in the Appendix the roots associated with a Jordan algebra in Figure 3. In particular we will pick the one whose projection on the plane Π is $\frac{1}{3}(k_2 + k_3 - 2k_1)$ (see Figure 2 for this vector), that is, the highest weight in the 3-dimensional representation of $\mathbf{su(3)} \sim \mathbf{a_2}$. We will refer to this Jordan algebra as the highest-weight (HW) $\mathbf{J}_3^\mathbf{n}$. The other Jordan algebras are obtained by a permutation of indexes and their conjugate ones by a change of sign.

Let us explain why we say that certain roots correspond to a Jordan pair. The reason lies in the Tits–Kantor–Koecher construction (2-4), which is related to the third row of the Freudenthal–Tits magic square. There is only one way of realizing the embedding $\mathbf{g}_{\mathrm{II}} \subset \mathbf{g}_{\mathrm{III}} \subset \mathfrak{L}^n$ so that the $(\mathbf{J}_3^\mathbf{n}, \bar{\mathbf{J}}_3^\mathbf{n})$ modules for $\mathbf{g}_{\mathrm{II}}$ lie on parallel spaces at the same distance along a fixed axis. This is precisely the way we will describe the Jordan pair content of the algebras and this shows the uniqueness of the construction. We know from the three grading structure of $\mathbf{g}_{\mathrm{III}}$ that the pair $(\mathbf{J}_3^\mathbf{n}, \bar{\mathbf{J}}_3^\mathbf{n})$ is indeed a Jordan pair and that $\mathrm{str}(\mathbf{J}_3^\mathbf{n}) = \mathbf{g}_{\mathrm{II}} \oplus \mathbf{C}$ is the Lie algebra of the automorphism group of the Jordan pair. This proves that the Jordan structures we have referred to so far are indeed so.

In the Appendix, the four exceptional algebras $\mathbf{f_4}$, $\mathbf{e_6}$, $\mathbf{e_7}$, and $\mathbf{e_8}$ are examined case by case. For each we show:

(1) the roots associated with the HW Jordan algebra $\mathbf{J_3^n}$,

(2) the roots associated with $\mathbf{g_0^n} = \mathbf{g}_{\mathrm{II}}$, and

(3) the nested Jordan pairs.

4.1. ***The geometry of the Jordan pair* V *and of* $\mathbf{g_{III}} = \mathbf{g_{II}} \oplus \mathbf{C} \oplus \mathbf{V}$.** The root vectors of the HW $\mathbf{J_3^n}$ all lie on a $(r-2)$-dimensional space Σ^+, where r is the rank of the exceptional Lie algebra. The space Σ^+ is parallel to the $(r-2)$-dimensional space Σ^0 on which the $\mathbf{g_0^n}$ roots lie, and to the $(r-2)$-dimensional space Σ^- on which the roots of the $\bar{\mathbf{J}}_\mathbf{3}^\mathbf{n}$ opposite to the HW $\mathbf{J_3^n}$ lie. Both spaces $\Sigma^\pm$ have the same distance $\frac{\sqrt{6}}{3}$ from $\Sigma^\mathbf{0}$, but lie on opposite sides with respect to it.

This is shown in Figure 5, in the case of $\mathbf{f_4}$. The two $\mathbf{J_3^1}$ form a Jordan pair of conjugate $\mathbf{a_2}$-representations $(6, \bar{6})$. The roots on the three planes form the root diagram of $\mathbf{c_3}$.

This Jordan pair is clearly visible in the figure. The Lie algebra of the automorphism group of this pair is $\mathbf{a_2} \oplus \mathbf{C}$ where $\mathbf{C}$ is the complex linear span of the Cartan generator associated with the axis along the vector $\frac{1}{3}(k_2 + k_3 - 2k_1)$ which is precisely the direction of the Jordan pair in Figure 3. All the points of the HW $\mathbf{J_3^1}$ (respectively, $\bar{\mathbf{J}}_\mathbf{3}^\mathbf{1}$, opposite to it with respect to the center of Figure 3) project on the point $\frac{1}{3}(k_2 + k_3 - 2k_1)$ (respectively, $-\frac{1}{3}(k_2 + k_3 - 2k_1)$) in the plane of Figure 3.

There is only one way of embedding a $\mathbf{c_3}$ subalgebra within $\mathbf{f_4}$ so that the $(6, \bar{6})$ modules for $\mathbf{a_2}$ lie on parallel planes at the same distance along a fixed axis. This is precisely the way we have described above and this shows its uniqueness. We know from the three grading structure of $\mathbf{c_3}$ that the pair $(\mathbf{J_3^1}, \bar{\mathbf{J}}_\mathbf{3}^\mathbf{1})$ is indeed a Jordan

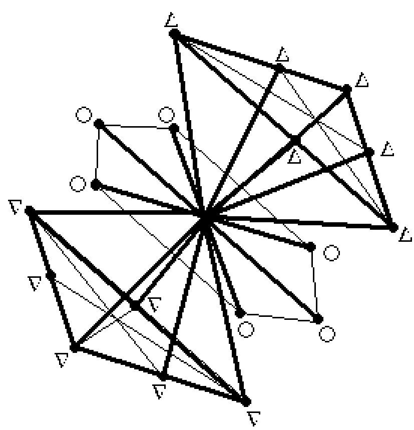

Figure 5. Root digram of $\mathbf{c_3}$ showing $\mathbf{a_2}$ and the Jordan pair $(6, \bar{6})$.

pair and that $\mathrm{str}(\mathbf{J})$ $(= \mathbf{a_2} \oplus \mathbf{C}$ in this case) is the Lie algebra of the automorphism group of the Jordan pair.

By a cyclic permutation of the indexes of k_1, k_2, and k_3 we obtain an analogous result for the other two Jordan pairs, all sharing the same $\mathbf{a_2}$ roots for the algebra $\mathbf{g_0^1}$, but with different orientations of the axis defining $\mathbf{C}$ along the vectors $\frac{1}{3}(k_1 + k_3 - 2k_2)$ and $\frac{1}{3}(k_1 + k_2 - 2k_3)$. We get in four dimensions three copies of $\mathbf{c_3}$ all sharing the same $\mathbf{a_2}$. All the spaces spanned by the three Jordan pairs are parallel to the space $\Sigma^{\mathbf{0}}$, and all at the same distance $\pm\frac{\sqrt{6}}{3}$ from it. Notice that in r dimensions there are an infinite number of $(r-2)$-dimensional spaces parallel to a given one, all at the same distance from it.

We get exactly the same feature for the other exceptional Lie algebras, with the Lie algebras of the second and third rows of the magic square playing the same role as for $\mathbf{f_4}$.

5. Representations

I briefly sketch in this section a possible representation of the $\mathbf{e_8}$ algebra which exhibits its Jordan pair content.

The way I would represent $\mathbf{e_8}$ is a development of the representation of $\mathbf{e_7}$ through generalized 2×2 matrices, shown in [Truini et al. 1986]. The starting point of that paper is the representation of the quaternion algebra through Pauli matrices, which leads directly to the three grading of $\mathbf{e_7}$. In the case of $\mathbf{e_8}$ a suitable representation of the octonions is via the Zorn matrices [Zorn 1933; Loos et al. 2008], which exhibit the $(3, \bar{3})$ structure that we can extend to the Jordan pair content of $\mathbf{e_8}$ and to the action on the $(3, \bar{3})$ modules of the *external* $\mathbf{a_2}$ in Figure 3.

The guidelines go as follows:

- Represent the octonions as Zorn matrices.
- Extend the Zorn matrices to represent $\mathrm{Der}(\mathfrak{C}) = \mathbf{g_2}$.
- Combine the extended Zorn matrices with the Tits construction (3-2).
- Decompose the representation of $\mathbf{e_6}$ to finally get $\mathbf{e_8}$ in terms of Jordan pairs and $\mathbf{a_2}$'s only.

If $a \in \mathfrak{C}$ we write $a = a_0 + \sum_{k=1}^{7} a_k u_k$ where $a_\ell \in \mathbf{C}$ for $\ell = 0, \ldots, 7$ and $u_1, \ldots, u_7$ are the octonionic imaginary units.

Let us denote by i the imaginary unit in $\mathbf{C}$. We introduce two idempotent elements

$$\rho_\pm = \tfrac{1}{2}(1 \pm i u_7)$$

and six nilpotent elements

$$\varepsilon_k^\pm = \rho^\pm u_k, \quad k = 1, 2, 3.$$

The Zorn representation of $a \in \mathfrak{C}$ is:

$$a = \alpha^+\rho^+ + \alpha^-\rho^- + \sum_k (\alpha_k^+\varepsilon_k^+ + \alpha_k^-\varepsilon_k^-) \leftrightarrow \begin{bmatrix} \alpha^+ & A^+ \\ A^- & \alpha^- \end{bmatrix}, \tag{5-1}$$

where $A^\pm \in \mathbf{C}^3$ have vector components $\alpha_k^\pm$, $k = 1, 2, 3$, and the octonionic multiplication is a generalization of matrix multiplication:

$$\begin{aligned} ab &\leftrightarrow \begin{bmatrix} \alpha^+ & A^+ \\ A^- & \alpha^- \end{bmatrix} \begin{bmatrix} \beta^+ & B^+ \\ B^- & \beta^- \end{bmatrix} \\ &= \begin{bmatrix} \alpha^+\beta^+ + A^+ \cdot B^- & \alpha^+B^+ + \beta^-A^+ + A^- \times B^- \\ \alpha^-B^- + \beta^+A^- + A^+ \times B^+ & \alpha^-\beta^- + A^- \cdot B^+ \end{bmatrix}, \end{aligned} \tag{5-2}$$

with $A^\pm \cdot B^\mp = -\alpha_k^\pm \beta_k^\mp$ and where $A, B \to A \times B$ is the standard vector product in $\mathbf{C}^3$.

The next step is to write the Lie algebra $\mathbf{g_2}$ using an extension of the Zorn matrices and their multiplication rule with an $\mathbf{a_2}$ matrix replacing α^+. This representation shows $\mathbf{g_2}$ as $\mathbf{a_2}$ plus its modules $(3, \bar{3})$.

Finally, let me outline how the Tits construction fits into this picture. The idea is to write

$$\mathbf{e_8} = \mathrm{Der}(\mathfrak{C}) \oplus \mathfrak{C}_0 \otimes \mathbf{J}_0^8 \oplus \mathrm{Der}(\mathbf{J}^8) = \mathfrak{L}_0 \oplus \sum_{\pm k} \mathfrak{L}_{\pm k}, \quad k = 1, 2, 3, \tag{5-3}$$

where

$$\mathfrak{L}_0 = D_7 \oplus iu_7 \otimes \mathbf{J}_0^8 \oplus \mathrm{Der}(\mathbf{J}^8) \quad \text{and} \quad \mathfrak{L}_{\pm k} = d_k^\pm D_k^\pm \oplus \alpha_k^\pm \varepsilon_k^\pm \otimes \mathbf{J}_0^8, \qquad d_k^\pm, \alpha_k^\pm \in \mathbf{C}.$$

Here $\mathbf{J}^8 \equiv \mathbf{J_3^8}$ and $\mathbf{J}_0^8$ is a traceless $\mathbf{J_3^8}$ matrix; $D_7 = \mathbf{a_2}$ is the subalgebra of derivations leaving the imaginary unit u_7 fixed; and $D_k^\pm = \pm\frac{3}{2} D_{iu_7, \varepsilon_k^\pm}$ is a derivation:

$$D_{a,b}c = \tfrac{1}{3}[[a, b], c] - (a, b, c), \qquad (a, b, c) = (ab)c - a(bc).$$

We identify $a \otimes x$ with $a_z \otimes x$, where a_z is the Zorn matrix representation of a and $\mathrm{Der}_k^\pm$ with the corresponding Zorn matrix representation of $\varepsilon_k^\pm$. We use the complex parameters $d_k^\pm$ in order to provide the trace to $\mathbf{J}_0$.

The chain of implications, starting from the Tits construction, would be like this:

$$\begin{aligned} \mathbf{e_8} &= \mathrm{Der}(\mathfrak{C}) \oplus \mathfrak{C}_0 \otimes \mathbf{J}_0^8 \oplus \mathrm{Der}(\mathbf{J}^8) \\ &= \mathbf{a_2^c} \oplus \alpha_k^\pm \varepsilon_k^\pm \otimes \mathbf{J}_0^8 \oplus d_k^\pm \,\mathrm{Der}_k^\pm \oplus (iu_7) \otimes \mathbf{J}_0^8 \oplus \mathrm{Der}(\mathbf{J}^8) \\ &= \mathbf{a_2^c} \oplus \alpha_k^\pm \varepsilon_k^\pm \otimes \mathbf{J}^8 \oplus \mathrm{Der}(\mathfrak{C}) \oplus \mathfrak{C}_0 \otimes \mathbf{J}_0^2 \oplus \mathrm{Der}(\mathbf{J}^2) \\ &= \mathbf{a_2^c} \oplus \alpha_k^\pm \varepsilon_k^\pm \otimes \mathbf{J}^8 \oplus \mathbf{a_2^f} \oplus \alpha_k^\pm \varepsilon_k^\pm \otimes \mathbf{J}^2 \oplus (iu_7) \otimes \mathbf{J}_0^2 \oplus \mathrm{Der}(\mathbf{J}^2) \\ &= \mathbf{a_2^c} \oplus \mathbf{a_2^f} \oplus \mathbf{a_2^{g_1}} \oplus \mathbf{a_2^{g_2}} \oplus 3 \times (\mathbf{J}^8, \bar{\mathbf{J}}^8) \oplus 3 \times (\mathbf{J}^2, \bar{\mathbf{J}}^2). \end{aligned} \tag{5-4}$$

Work is still in progress along these lines and will appear in a forthcoming paper.

6. Elementary particle physics

If we look at the decomposition (5-4) (see also (A.7) in the Appendix) we are led to interpret the labels $\mathbf{c}$ as color and $\mathbf{f}$ as flavor. In this interpretation the three pairs $(\mathbf{J}_3^8, \bar{\mathbf{J}}_3^8)$ accommodate the quarks in three colors of particles–antiparticles, whereas the three pairs $(\mathbf{J}_3^2, \bar{\mathbf{J}}_3^2)$ sitting in the colorless $\mathbf{g}_0^8$ accommodate the three families of leptons–antileptons. Including spin, each particle must appear with four different degrees of freedom: left (up and down) and right (up and down), except, possibly, for the neutrino, which could be a Majorana neutrino and be only left-handed. We can therefore put six (quarks, antiquarks) in a (say) blue $(\mathbf{J}_3^8, \bar{\mathbf{J}}_3^8)$. We can make them coincide with three octonions: one for blue up–down quarks, one for blue charm–strange quarks, one for blue top–bottom quarks. We are left with three extra degrees of freedom. In the same fashion, we can put a family of leptons–antileptons pairs in $(\mathbf{J}_3^2, \bar{\mathbf{J}}_3^2)$ by letting the six off-diagonal degrees of freedom of be the electron and a Majorana neutrino, and analogously for the families of the muon and τ leptons. Again we are left with three extra degrees of freedom, which reduce to only one in the case where right-handed neutrinos are included.

Let us review the explicit form of the roots (see the Appendix) according to this interpretation.

Quarks of color $\mathbf{c} = 1, 2, 3$ (corresponding antiquarks have reversed signs):

$$-k_c \pm k_j, \quad j = 4, \ldots, 8, \quad -k_c + k_1 + k_2 + k_3,$$
$$-k_c + \tfrac{1}{2}(k_1 + k_2 + k_3 \pm k_4 \pm k_5 \pm k_6 \pm k_7 \pm k_8) \qquad \text{(even \# of + signs)}.$$

Leptons in the family $\mathbf{f} = 4, 5, 6$ (corresponding antileptons have reversed signs):

$$-k_f \pm k_j, \quad j = 7, 8, \quad -k_f + k_4 + k_5 + k_6,$$
$$-k_f + \tfrac{1}{2}[\pm(k_1 + k_2 + k_3) + k_4 + k_5 + k_6 \pm k_7 \pm k_8] \qquad \text{(even \# of + signs)}.$$

$\mathbf{a}_2^{\mathbf{c}}$: $\pm(k_i - k_j)$, $i < j = 1, 2, 3$.

$\mathbf{a}_2^{\mathbf{f}}$: $\pm(k_i - k_j)$, $i < j = 4, 5, 6$.

$\mathbf{a}_2^{\mathbf{g}_1}$: $\pm(k_7 + k_8)$, $\pm\frac{1}{2}(k_1 + k_2 + k_3 + k_4 + k_5 + k_6 - k_7 - k_8) \pm \frac{1}{2}(k_1 + k_2 + k_3 + k_4 + k_5 + k_6 + k_7 + k_8)$.

$\mathbf{a}_2^{\mathbf{g}_2}$: $\pm(k_7 - k_8)$, $\pm\frac{1}{2}(-k_1 - k_2 - k_3 + k_4 + k_5 + k_6 - k_7 + k_8) \pm \frac{1}{2}(-k_1 - k_2 - k_3 + k_4 + k_5 + k_6 + k_7 - k_8)$.

What physics should a theory with an $\mathbf{e_8}$ symmetry describe? Certainly a very high-energy physics, far beyond our present experience and our experimental reach. It could relate to a string theory, like the heterotic one, since we are dealing with a complex Lie algebra hence an $\mathbf{e_8} \times \mathbf{e_8}$ algebra over $\mathbf{R}$. It could extend to supersymmetry, although the $\mathbf{e_8}$ symmetry is so beautiful as it stands that one

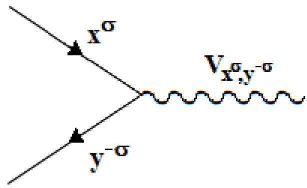

Figure 6. An elementary interaction, viewed as a Feynman diagram.

should force such an extension into the theory: $\mathbf{e_8}$, in the view I am presenting here, shows particle–antiparticle pairs, the Jordan pairs, in the right number of colors and families, plus their symmetries, which in turn are generated by the pairs themselves, through the trilinear map $z^\sigma \to V_{x^\sigma, y^{-\sigma}} z^\sigma$. Besides, another peculiarity contributes to the beauty of $\mathbf{e_8}$: its lowest-dimensional irreducible representation is the adjoint representation.

My personal point of view is that, at such a high energy, at or beyond the Planck scale, the picture of spacetime has to be radically changed. I can hardly make any sense of the fact that such an energetic particle is sitting on a background spacetime, if I think that general relativity taught us that spacetime is in fact dynamical. I would rather view that particle as feeling only (quantum) interactions, including one that leads to gravity, to be accommodated within $\mathbf{a_2^{g_1}} \oplus \mathbf{a_2^{g_2}}$. I would still view an *elementary* interaction being described by an elementary Feynman diagram involving the trilinear map, as depicted in Figure 6, but with no question of point or extended particle, simply because the underlying spacetime geometry is not there: there is only a, let us say, background independent *spectral theory*.

In this view the classical spacetime is a byproduct of the interactions, obtained by taking very rough approximations. It is as far from the interactions exchanged by elementary particles at the Planck scale, as the Planck scale is far from our experience.

The aim of developing along these lines a physical theory that could not possibly rely on any direct confirmation, is to find a consistent quantum theory of gravity together with the other known basic interactions. As Carlo Rovelli says [2004]: "the difficulty is not to discriminate among many complete and consistent quantum theories of gravity. We would be content with one."

This is, of course, far beyond the scope of the present paper, since no physics has been spoken here besides these mere speculations.

Appendix

The explicit set of roots we use is shown in Table 2 [Bourbaki 1968]; $\{k_i,\ i = 1, \dots, 8\}$ denotes an orthonormal basis in $\mathbf{R}^8$.

$\mathfrak{L}$	Roots $\{k_i, i = 1, \dots 8\}$ an orthonormal basis in $\mathbf{R}^8$	# of roots
$\mathbf{g_2}$		**12**
	$(k_i - k_j), \quad i \neq j = 1, 2, 3,$	6
	$\pm\frac{1}{3}(-2k_i + k_j + k_l), \quad i \neq j \neq l = 1, 2, 3,$	6
$\mathbf{f_4}$		**48**
	$\pm k_i, \quad i = 1, \dots, 4,$	8
	$\pm k_i \pm k_j, \quad i \neq j = 1, \dots, 4,$	$4 \times \binom{4}{2} = 24$
	$\frac{1}{2}(\pm k_1 \pm k_2 \pm k_3 \pm k_4)$	$2^4 = 16$
$\mathbf{e_6}$		**72**
	$\pm k_i \pm k_j, \quad i \neq j = 1, \dots, 5,$	$4 \times \binom{5}{2} = 40$
	$\frac{1}{2}(\pm k_1 \pm k_2 \pm k_3 \pm k_4 \pm k_5 \pm \sqrt{3}k_6)^*$	$2^5 = 32$
	* odd number of $+$ signs	
$\mathbf{e_7}$		**126**
	$\pm\sqrt{2}k_7,$	2
	$\pm k_i \pm k_j, \quad i \neq j = 1, \dots, 6,$	$4 \times \binom{6}{2} = 60$
	$\frac{1}{2}(\pm k_1 \pm k_2 \pm k_3 \pm k_4 \pm k_5 \pm k_6 \pm \sqrt{2}k_7)^*$	$2^6 = 64$
	* even number of $+\frac{1}{2}$	
$\mathbf{e_8}$		**240**
	$\pm k_i \pm k_j, \quad i \neq j = 1, \dots, 8,$	$4 \times \binom{8}{2} = 112$
	$\frac{1}{2}(\pm k_1 \pm k_2 \pm k_3 \pm k_4 \pm k_5 \pm k_6 \pm k_7 \pm k_8)^*$	$2^7 = 128$
	* even number of $+$ signs	

Table 2. The roots of the exceptional Lie algebras.

A.1 $\mathbf{f_4}$.

A.1.1 *The roots associated with the HW* $\mathbf{J_3^1}$.

(A.1) $$-k_1, \qquad -k_1 \pm k_4, \qquad \tfrac{1}{2}(-k_1 + k_2 + k_3 \pm k_4), \qquad k_2 + k_3.$$

A.1.2 *The roots associated with* $\mathbf{g}_{\mathrm{II}} = \mathbf{g_0^1}$.

(A.2) $$\pm k_4, \qquad \pm\tfrac{1}{2}(k_1 + k_2 + k_3 \pm k_4).$$

A.1.3 *Nested Jordan pairs.* If we dig inside $\mathbf{g_0^1}$ we find another Jordan pair plus the Lie algebra of its automorphism group: these are a $(2, \bar{2})$ of $\mathbf{a_1}$ plus $\mathbf{a_1} \oplus \mathbf{C}$ making up, all together, $\mathbf{a_2}$.

A.2 $\mathbf{e_6}$.

A.2.1 *The roots associated with the HW* $\mathbf{J_3^2}$.

$$\begin{aligned} & -k_1 \pm k_4, \qquad -k_1 \pm k_5, \qquad k_2+k_3, \\ & \tfrac{1}{2}(-k_1+k_2+k_3+k_4-k_5-\sqrt{3}k_6), \\ & \tfrac{1}{2}(-k_1+k_2+k_3+k_4+k_5+\sqrt{3}k_6), \\ & \tfrac{1}{2}(-k_1+k_2+k_3-k_4+k_5-\sqrt{3}k_6), \\ & \tfrac{1}{2}(-k_1+k_2+k_3-k_4-k_5+\sqrt{3}k_6). \end{aligned} \tag{A.3}$$

A.2.2 *The roots associated with* $\mathbf{g}_{\mathrm{II}} = \mathbf{g_0^2}$.

$$\begin{aligned} \mathbf{a_2^{(1)}} : & \\ & \pm(k_4+k_5), \\ & \pm\tfrac{1}{2}(k_1+k_2+k_3-k_4-k_5-\sqrt{3}k_6), \\ & \pm\tfrac{1}{2}(k_1+k_2+k_3+k_4+k_5-\sqrt{3}k_6). \end{aligned}$$

$$\begin{aligned} \mathbf{a_2^{(2)}} : & \\ & \pm(k_4-k_5), \\ & \pm\tfrac{1}{2}(k_1+k_2+k_3-k_4+k_5+\sqrt{3}k_6), \\ & \pm\tfrac{1}{2}(k_1+k_2+k_3+k_4-k_5+\sqrt{3}k_6). \end{aligned}$$

A.2.3 *Nested Jordan pairs.* If we dig inside $\mathbf{g_0^2}$ we find another Jordan pair plus the Lie algebra of its automorphism group: these are two replicas of a $(2, \bar{2})$ of $\mathbf{a_1}$ plus $\mathbf{a_1} \oplus \mathbf{C}$ making up, all together, $\mathbf{a_2} \oplus \mathbf{a_2}$.

A.3 $\mathbf{e_7}$.

A.3.1 *The roots associated with the HW* $\mathbf{J_3^4}$.

$$\begin{aligned} & -k_1 \pm k_4, \qquad -k_1 \pm k_5, \qquad -k_1 \pm k_6, \qquad k_2+k_3, \\ & \tfrac{1}{2}(-k_1+k_2+k_3-k_4-k_5-k_6 \pm \sqrt{2}k_7), \\ & \tfrac{1}{2}(-k_1+k_2+k_3-k_4+k_5+k_6 \pm \sqrt{2}k_7), \\ & \tfrac{1}{2}(-k_1+k_2+k_3+k_4-k_5+k_6 \pm \sqrt{2}k_7), \\ & \tfrac{1}{2}(-k_1+k_2+k_3+k_4+k_5-k_6 \pm \sqrt{2}k_7). \end{aligned} \tag{A.4}$$

A.3.2 *The roots associated with* $\mathbf{g}_{\mathrm{II}} = \mathbf{g_0^4}$.

$$\begin{gathered} \pm k_4 \pm k_5, \quad \pm k_4 \pm k_6, \quad \pm k_5 \pm k_6, \quad \pm\sqrt{2}k_7, \\ \pm\tfrac{1}{2}(k_1+k_2+k_3 \pm k_4 \pm k_5 \pm k_6 \pm \sqrt{2}k_7) \quad (\text{even number of } +\tfrac{1}{2}). \end{gathered}$$

A.3.3 *Nested Jordan pairs.* If we dig inside $\mathbf{g_0^4} = \mathbf{a_5}$ we find the Jordan pair $(\mathbf{J_3^2}, \mathbf{\bar{J}_3^2}) = (3\times 3, \bar{3}\times\bar{3})$ plus the Lie algebra of its automorphism group $\mathbf{a_2} \oplus \mathbf{a_2} \oplus \mathbf{C}$ described in the previous case of $\mathbf{e_6}$.

A.4 $\mathbf{e_8}$.

A.4.1 *The roots associated with the HW* $\mathbf{J_3^8}$.

$$-k_1 \pm k_j, \quad j = 4, \dots, 8, \qquad k_2 + k_3,$$
$$\tfrac{1}{2}(-k_1 + k_2 + k_3 \pm k_4 \pm k_5 \pm k_6 \pm k_7 \pm k_8) \quad \text{(even number of + signs)}. \tag{A.5}$$

A.4.2 *The roots associated with* $\mathbf{g}_{\mathrm{II}} = \mathbf{g_0^8}$. The 72 roots of $\mathbf{g_0^8} = \mathbf{e_6}$ are

$$\pm k_i \pm k_j, \quad i, j = 4, \dots, 8,$$
$$\pm\frac{1}{2}(k_1 + k_2 + k_3 \pm k_4 \pm k_5 \pm k_6 \pm k_7 \pm k_8) \quad \text{(even number of + signs)}. \tag{A.6}$$

A.4.3 *Nested Jordan pairs.* If we dig inside $\mathbf{g_0^8} = \mathbf{e_6}$ we find three Jordan pairs, each of the type $(\mathbf{J_3^2}, \mathbf{\bar{J}_3^2}) = (3 \times 3, \bar{3} \times \bar{3})$, plus the Lie algebra of the automorphism group of each of them $\mathbf{a_2} \oplus \mathbf{a_2} \oplus \mathbf{C}$ described in the previous case of $\mathbf{e_6}$.

We thus identify four different $\mathbf{a_2}$'s within $\mathbf{e_8}$ plus six Jordan pairs. Giving different superscripts to the four $\mathbf{a_2}$'s we have:

$$\begin{aligned}
\mathbf{e_8} &= \mathbf{a_2^c} \oplus 3 \times (\mathbf{J_3^8}, \mathbf{\bar{J}_3^8}) \oplus \mathbf{g_0^8} \\
&= \mathbf{a_2^c} \oplus 3 \times (\mathbf{J_3^8}, \mathbf{\bar{J}_3^8}) \oplus \mathbf{a_2^f} \oplus 3 \times (\mathbf{J_3^2}, \mathbf{\bar{J}_3^2}) \oplus \mathbf{g_0^2} \\
&= \mathbf{a_2^c} \oplus 3 \times (\mathbf{J_3^8}, \mathbf{\bar{J}_3^8}) \oplus \mathbf{a_2^f} \oplus 3 \times (\mathbf{J_3^2}, \mathbf{\bar{J}_3^2}) \oplus \mathbf{a_2^{g_1}} \oplus \mathbf{a_2^{g_2}}.
\end{aligned} \tag{A.7}$$

Acknowledgments

I am very grateful to Professor V. S. Varadarajan for scientific advice and financial support. The discussions we have had together have actually given life to the idea of writing this paper. I also thank Professor Varadarajan and the UCLA Department of Mathematics for their kind hospitality. The visit to the University of California at Los Angeles was made possible by the *Istituto Nazionale di Fisica Nucleare*, on grant In. Spec. GE 41.

References

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie, Chapitres IV–VI*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. MR 39 #1590 Zbl 0186.33001

[Candelas et al. 1985] P. Candelas, G. T. Horowitz, A. Strominger, and E. Witten, "Vacuum configurations for superstrings", *Nuclear Phys. B* **258**:1 (1985), 46–74. MR 87k:83091a

[Cremmer 1982] E. Cremmer, "Dimensional reduction in field theory and hidden symmetries in extended supergravity", pp. xi+489 in *Supergravity: Proceedings of the First School* (Trieste, 1981), edited by S. Ferrara and J. G. Taylor, Cambridge University Press, Cambridge, 1982.

[Faulkner 2000] J. R. Faulkner, "Jordan pairs and Hopf algebras", *J. Algebra* **232**:1 (2000), 152–196. MR 2001e:17042 Zbl 1018.17022

[Ferrara and Günaydin 1998] S. Ferrara and M. Günaydin, "Orbits of exceptional groups, duality and BPS states in string theory", *Internat. J. Modern Phys. A* **13**:13 (1998), 2075–2088. MR 99g:81160

[Ferrara and Kallosh 1996a] S. Ferrara and R. Kallosh, "Supersymmetry and attractors", *Phys. Rev. D* (3) **54**:2 (1996), 1514–1524. MR 97e:83101 Zbl 1171.83329

[Ferrara and Kallosh 1996b] S. Ferrara and R. Kallosh, "University of supersymmetry and attractors", *Phys. Rev. D* **54**:2 (1996), 1525–1534.

[Frampton et al. 1980] P. H. Frampton, S. L. Glashow, and A. Yildiz (editors), *First Workshop on Grand Unification* (Durham, NH, 1980), Lie Groups: History, Frontiers and Applications **11**, Math Sci Press, Brookline, MA, 1980. MR 81j:81066

[Freudenthal 1959] H. Freudenthal, "Beziehungen der $\mathfrak{E}_7$ und $\mathfrak{E}_8$ zur Oktavenebene, VIII-IX", *Nederl. Akad. Wetensch. Proc. Ser. A* **62** (1959), 447–465. Zbl 0128.15302

[Green and Schwarz 1984] M. B. Green and J. H. Schwarz, "Anomaly cancellations in supersymmetric $D = 10$ gauge theory and superstring theory", *Phys. Lett. B* **149**:1-3 (1984), 117–122. MR 85k:81173

[Gross 1986] D. J. Gross, "The heterotic string", pp. 357–399 in *Workshop on unified string theories* (Santa Barbara, CA, 1985), edited by M. B. Green and D. Gross, World Scientific, Singapore, 1986. MR 849110

[Jacobson 1966] N. Jacobson, "Structure theory for a class of Jordan algebras", *Proc. Nat. Acad. Sci. USA* **55** (1966), 243–251. MR 33 #1337 Zbl 0133.28903

[Jordan et al. 1934] P. Jordan, J. von Neumann, and E. Wigner, "On an algebraic generalization of the quantum mechanical formalism", *Ann. of Math.* (2) **35**:1 (1934), 29–64. MR 1503141 Zbl 0008.42103

[Kantor 1964] I. L. Kantor, "Classification of irreducible transitive differential groups", *Dokl. Akad. Nauk SSSR* **158** (1964), 1271–1274. In Russian; translated in *Sov. Math. Dokl.* **5** (1965), 1404–1407. MR 31 #217 Zbl 0286.17011

[Koecher 1967] M. Koecher, "Imbedding of Jordan algebras into Lie algebras, I", *Amer. J. Math.* **89** (1967), 787–816. MR 35 #5480 Zbl 0209.06801

[Lisi et al. 2010] A. G. Lisi, L. Smolin, and S. Speziale, "Unification of gravity, gauge fields and Higgs bosons", *J. Phys. A* **43**:44 (2010), 445401, 10. MR 2011k:81283 Zbl 1202.81226

[Loos 1975] O. Loos, *Jordan pairs*, Lecture Notes in Mathematics **460**, Springer, Berlin, 1975. MR 56 #3071 Zbl 0301.17003

[Loos et al. 2008] O. Loos, H. P. Petersson, and M. L. Racine, "Inner derivations of alternative algebras over commutative rings", *Algebra Number Theory* **2**:8 (2008), 927–968. MR 2009i:17048 Zbl 1191.17011

[McCrimmon 2004] K. McCrimmon, *A taste of Jordan algebras*, Springer, New York, 2004. MR 2004i:17001 Zbl 1044.17001

[Meyberg 1970] K. Meyberg, "Jordan-Tripelsysteme und die Koecher-Konstruktion von Lie-Algebren", *Math. Z.* **115** (1970), 58–78. MR 41 #8482 Zbl 0186.34501

[Rovelli 2004] C. Rovelli, *Quantum gravity*, Cambridge University Press, Cambridge, 2004. MR 2005m:83049 Zbl 1091.83001

[Schafer 1949] R. D. Schafer, "Inner derivations of non-associative algebras", *Bull. Amer. Math. Soc.* **55** (1949), 769–776. MR 11,77a Zbl 0033.34803

[Schafer 1966] R. D. Schafer, *An introduction to nonassociative algebras*, Pure and Applied Mathematics **22**, Academic Press, New York, 1966. MR 35 #1643 Zbl 0145.25601

[Tits 1955] J. Tits, "Sur certaines classes d'espaces homogènes de groupes de Lie", *Acad. Roy. Belg. Cl. Sci. Mém. Coll. in* 8° **29**:3 (1955), 268. MR 17,874f Zbl 0067.12301

[Tits 1962] J. Tits, "Une classe d'algèbres de Lie en relation avec les algèbres de Jordan", *Nederl. Akad. Wetensch. Proc. Ser. A* **65** (1962), 530–535. MR 26 #3753 Zbl 0104.26002

[Truini and Biedenharn 1982] P. Truini and L. C. Biedenharn, "An $\mathscr{E}_6 \otimes \mathscr{U}(1)$ invariant quantum mechanics for a Jordan pair", *J. Math. Phys.* **23**:7 (1982), 1327–1345. MR 84j:81021 Zbl 0495.22012

[Truini et al. 1986] P. Truini, G. Olivieri, and L. C. Biedenharn, "Three graded exceptional algebras and symmetric spaces", *Z. Phys. C* **33**:1 (1986), 47–65. MR 88c:17012 Zbl 0696.17005

[Zorn 1933] M. Zorn, "Alternativkörper und quadratische Systeme", *Abh. Math. Semin. Hamb. Univ.* **9** (1933), 395–402. Zbl 0007.05403

Received March 29, 2012.

PIERO TRUINI
DIPARTIMENTO DI FISICA
UNIVERSITÀ DI GENOVA
VIA DODECANESO 33
I-16146 GENOVA
ITALY
truini@ge.infn.it

PACIFIC JOURNAL OF MATHEMATICS
Vol. 260, No. 1, 2012
dx.doi.org/10.2140/pjm.2012.260.245

LOWER ESTIMATE OF MILNOR NUMBER AND CHARACTERIZATION OF ISOLATED HOMOGENEOUS HYPERSURFACE SINGULARITIES

STEPHEN S.-T. YAU AND HUAIQING ZUO

Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a germ of a complex analytic function with an isolated critical point at the origin. Let $V = \{ z \in \mathbb{C}^n : f(z) = 0 \}$. A beautiful theorem of Saito [1971] gives a necessary and sufficient condition for V to be defined by a weighted homogeneous polynomial. It is a natural and important question to characterize (up to a biholomorphic change of coordinates) a homogeneous polynomial with an isolated critical point at the origin. For a two-dimensional isolated hypersurface singularity V, Xu and Yau [1992; 1993] found a coordinate-free characterization for V to be defined by a homogeneous polynomial. Lin and Yau [2004] and Chen, Lin, Yau, and Zuo [2001] gave necessary and sufficient conditions for 3- and 4-dimensional isolated hypersurface singularities with $p_g \geq 0$ and $p_g > 0$, respectively. However, it is quite difficult to generalize their methods to give characterization of homogeneous polynomials. In 2005, Yau formulated the Yau Conjecture 1.1: (1) Let μ and ν be the Milnor number and multiplicity of $(V, 0)$, respectively. Then $\mu \geq (\nu - 1)^n$, and the equality holds if and only if f is a semihomogeneous function. (2) If f is a quasihomogeneous function, then $\mu = (\nu - 1)^n$ if and only if f is a homogeneous polynomial after change of coordinates. In this paper we solve part (1) of Yau Conjecture 1.1 for general n. We introduce a new method, which allows us to solve the part (2) of Yau Conjecture 1.1 for $n = 5$ and 6. As a result we have shown that for $n = 5$ or 6, f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau = (\nu - 1)^n$. As a by-product we have also proved Yau Conjecture 1.2 in some special cases.

1. Introduction

Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be the germ of a complex analytic function with an isolated critical point at the origin. Let $V = \{ z \in \mathbb{C}^n : f(z) = 0 \}$. It is a natural question to ask when V is defined by a weighted homogeneous polynomial or a homogeneous

Dedicated to Professor Banghe Li on the occasion of his 70th birthday.

MSC2010: 32S25.

Keywords: Homogeneous singularities.

polynomial up to biholomorphic change of coordinates. Recall that the multiplicity of the singularity V is defined to be the order of the lowest nonvanishing term in the power series Taylor expansion of f at 0, and the Milnor number μ and the Tjurina number τ of the singularity $(V, 0)$ are defined respectively by

$$\mu = \dim \mathbb{C}\{z_1, z_2, \ldots, z_n\}/(f_{z_1}, \ldots, f_{z_n}),$$
$$\tau = \dim \mathbb{C}\{z_1, z_2, \ldots, z_n\}/(f, f_{z_1}, \ldots, f_{z_n}).$$

The following theorem gives a necessary and sufficient condition for V to be defined by a weighted homogeneous polynomial:

Theorem 1.1 [Saito 1971]. *The function f is a weighted homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau$.*

Let $\pi : (M, A) \to (V, 0)$ be a resolution of singularity with exceptional set $A = \pi^{-1}(0)$. The geometric genus p_g of the singularity $(V, 0)$ is the dimension of $H^{n-2}(M, \mathbb{O})$ and is independent of the resolution M. Xu and Yau [1993] gave necessary and sufficient conditions for a 2-dimensional V to be defined by a homogeneous polynomial.

Theorem 1.2 [Xu and Yau 1993]. *Let $(V, 0)$ be a 2-dimensional isolated hypersurface singularity defined by a holomorphic function $f(z_1, z_2, z_3) = 0$. Let μ be the Milnor number, τ the Tjurina number, p_g the geometric genus, and ν the multiplicity of the singularity. Then f is a homogeneous polynomial after a biholomorphic change of variables if and only if $\mu = \tau$ and $\mu - \nu + 1 = 6p_g$.*

Based on above theorem, a conjecture was made by Yau in 2005 as follows:

Yau Conjecture 1.1 [Lin et al. 2006b]. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a holomorphic germ defining an isolated hypersurface singularity $V = \{ z : f(z) = 0 \}$ at the origin. Let μ and ν be the Milnor number and multiplicity of $(V, 0)$, respectively. Then*

$$\mu \geq (\nu - 1)^n, \tag{1-1}$$

and equality holds if and only if f is a semihomogeneous function (i.e., $f = f_\nu + g$, where f_ν is a nondegenerate homogeneous polynomial of degree ν and g consists of terms of degree at least $\nu + 1$) after a biholomorphic change of coordinates. Furthermore, if f is a quasihomogeneous function, i.e., $f \in (\partial f/\partial z_1, \ldots, \partial f/\partial z_n)$, then the equality in (1-1) holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates.

Yau Conjecture 1.2 [Chen et al. 2011]. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a weighted homogeneous polynomial with an isolated singularity at the origin. Let μ, p_g, and ν be the Milnor number, geometric genus, and multiplicity of the singularity $V = \{ z : f(z) = 0 \}$. Then*

$$\mu - p(\nu) \geq n!\, p_g,$$

where $p(\nu) = (\nu-1)^n - \nu(\nu-1)\dots(\nu-n+1)$, *and equality holds if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

These conjectures are sharp estimates and have some important applications in geometry. The Yau conjectures were proved only for very low dimensional singularities. For Yau Conjecture 1.1, Lin, Wu, Yau, and Luk proved the following two theorems:

Theorem 1.3 [Lin et al. 2006b]. *Let* $f : (\mathbb{C}^2, 0) \to (\mathbb{C}, 0)$ *be a germ of a holomorphic function defining an isolated plane curve singularity* $V = \{ z \in \mathbb{C}^2 : f(z) = 0 \}$ *at the origin. Let* μ *and* ν *be the Milnor number and multiplicity of* $(V, 0)$, *respectively. Then*

$$\mu \ge (\nu - 1)^2.$$

Furthermore, if V *has at most two irreducible branches at the origin, or if* f *is a quasihomogeneous function, then equality holds if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

Theorem 1.4 [Lin et al. 2006b]. *Let* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ *be a germ of a holomorphic function defining an isolated hypersurface singularity* $V = \{ z \in \mathbb{C}^n : f(z) = 0 \}$ *at the origin. Let* μ, ν, *and* $\tau = \dim \mathbb{C}\{z_1, \dots, z_n\}/(f, \partial f/\partial z_1, \dots, \partial f/\partial z_n)$ *be the Milnor number, multiplicity, and Tjurina number of* $(V, 0)$, *respectively. Suppose* $\mu = \tau$ *and* n *is either* 3 *or* 4. *Then*

$$\mu \ge (\nu - 1)^n,$$

and equality holds if and only if f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

For Yau Conjecture 1.2, Lin, Tu, and Yau have the following theorem:

Theorem 1.5 [Lin and Yau 2004; Lin et al. 2006a]. *Let* $(V, 0)$ *be a* 3*-dimensional isolated hypersurface singularity defined by a weighted homogeneous polynomial* $f(x, y, z, w) = 0$. *Let* μ *be the Milnor number,* p_g *the geometric genus, and* ν *the multiplicity of the singularity. Then*

$$\mu - (2\nu^3 - 5\nu^2 + 2\nu + 1) \ge 4!\, p_g,$$

and equality holds if and only if f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

Remark. The above theorem is proved in [Lin and Yau 2004] with $p_g > 0$. For $p_g = 0$, the theorem is proved in [Lin et al. 2006a].

Corollary 1.1. *Let* $(V, 0)$ *be a* 3*-dimensional isolated hypersurface singularity defined by a polynomial* $f(x, y, z, w) = 0$. *Let* μ, p_g, ν, *and* τ *be the Milnor number, geometric genus, multiplicity, and Tjurina number of the singularity,* *respectively. Then f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau$ and $\mu - (2\nu^3 - 5\nu^2 + 2\nu + 1) = 4!\, p_g$.*

Recently, Chen, Lin, Yau, and Zuo [Chen et al. 2011] generalized the above theorem to any 4-dimensional isolated hypersurface singularity with an additional assumption $p_g > 0$.

Theorem 1.6 [Chen et al. 2011]. *Let $(V, 0)$ be a 4-dimensional isolated hypersurface singularity defined by a weighted homogeneous polynomial $f(x, y, z, w, t) = 0$. Let μ be the Milnor number, p_g the geometric genus, and ν the multiplicity of the singularity. If $p_g > 0$, then*

$$\mu - [(\nu-1)^5 + \nu(\nu-1)(\nu-2)(\nu-3)(\nu-4)] \geq 5!\, p_g,$$

and equality holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates.

Corollary 1.2. *Let $(V, 0)$ be a 4-dimensional isolated hypersurface singularity defined by a polynomial $f(x, y, z, w, t) = 0$. Let μ, p_g, ν, and τ be the Milnor number, geometric genus, multiplicity, and Tjurina number of the singularity, respectively. Moreover, if $p_g > 0$, then f is a homogeneous polynomial after a biholomorphic change of coordinate if and only if $\mu = \tau$ and*

$$\mu - [(\nu-1)^5 + \nu(\nu-1)(\nu-2)(\nu-3)(\nu-4)] = 5!\, p_g.$$

The purpose of this paper is to prove the following results:

Proposition A. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a holomorphic germ defining an isolated hypersurface singularity $V = \{z : f(z) = 0\}$ at the origin. Let μ and ν be the Milnor number and multiplicity of $(V, 0)$, respectively. Then*

$$\mu \geq (\nu-1)^n,$$

and equality holds if and only if f is a semihomogeneous function (i.e., $f = f_\nu + g$, where f is a nondegenerate homogeneous polynomial of degree ν and g consists of terms of degree at least $\nu+1$) after a biholomorphic change of coordinates.

Theorem B. *Let $f : (\mathbb{C}^k, 0) \to (\mathbb{C}, 0)$, where k is either 5 or 6, be a weighted homogeneous polynomial with an isolated singularity at the origin. Let μ and ν be the Milnor number and multiplicity of the singularity $V = \{z : f(z) = 0\}$, respectively. Then*

$$\mu \geq (\nu-1)^k,$$

and equality holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates.

Theorem C. *Let* $f : (\mathbb{C}^5, 0) \to (\mathbb{C}, 0)$ *be a weighted homogeneous polynomial with an isolated singularity at the origin. Let* μ, p_g, *and* ν *be the Milnor number, geometric genus, and multiplicity of the singularity* $V = \{ z : f(z) = 0 \}$. *Then*

$$\mu - p(\nu) \geq 5!\, p_g,$$

where $p(\nu) = (\nu-1)^5 - \nu(\nu-1)(\nu-2)(\nu-3)(\nu-4)$, *and equality holds if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

Theorem D. *Let* $f : (\mathbb{C}^6, 0) \to (\mathbb{C}, 0)$ *be a weighted homogeneous polynomial with an isolated singularity at the origin. Let* μ, p_g, *and* ν *be the Milnor number, geometric genus, and multiplicity of the singularity* $V = \{ z : f(z) = 0 \}$. *If* $p_g = 0$, *then*

$$\mu - p(\nu) \geq 6!\, p_g,$$

where $p(\nu) = (\nu-1)^6 - \nu(\nu-1)(\nu-2)(\nu-3)(\nu-4)(\nu-5)$ (*which equals* 0), *and equality holds if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

Corollary E. *Let* $f : (\mathbb{C}^k, 0) \to (\mathbb{C}, 0)$, *where* k *is either* 5 *or* 6, *be a polynomial with an isolated singularity at the origin. Let* μ, τ, *and* ν *be the Milnor number, Tjurina number, and multiplicity of the singularity* $V = \{ z : f(z) = 0 \}$, *respectively. Then* f *is a homogeneous polynomial after a biholomorphic change of coordinates if and only if* $\mu = \tau = (\nu - 1)^k$.

In Section 2, we recall the necessary materials needed to prove the main theorems. In Section 3, we prove the main theorems.

2. Preliminary

In this section, we recall some known results that are needed to prove the main theorems. Let $f(z_1, \ldots, z_n)$ be a germ of an analytic function at the origin such that $f(0) = 0$. Suppose f has an isolated critical point at the origin. It can be developed in a convergent Taylor series $f(z_1, \ldots, z_n) = \sum a_\lambda z^\lambda$, where $z^\lambda = z_1^{\lambda_1} \ldots z_n^{\lambda_n}$. Recall that the Newton boundary $\Gamma(f)$ is the union of compact faces of $\Gamma_+(f)$, where $\Gamma_+(f)$ is the convex hull of the union of subsets $\{\lambda + \mathbb{R}^n_+\}$ for λ such that $a_\lambda \neq 0$. Let $\Gamma_-(f)$, the Newton polyhedron of f, be the cone over $\Gamma(f)$ with cone point at 0. For any closed face Δ of $\Gamma(f)$, we associate the polynomial $f_\Delta(z) = \sum_{\lambda \in \Delta} a_\lambda z^\lambda$. We say that f is nondegenerate if f_Δ has no critical point in $(\mathbb{C}^*)^n$ for any $\Delta \in \Gamma(f)$, where $\mathbb{C}^* = \mathbb{C} - \{0\}$. We say that a point p of the integral lattice $\mathbb{Z}^n$ in $\mathbb{R}^n$ is positive if all coordinates of p are positive. The following beautiful theorem holds:

Theorem 2.1 [Merle and Teissier 1980]. *Let* $(V, 0)$ *be an isolated hypersurface singularity defined by a nondegenerate holomorphic function* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$. *Then the geometric genus* $p_g = \#\{ p \in \mathbb{Z}^n \cap \Gamma_-(f) : p \text{ is positive} \}$.

A polynomial $f(z_1,\dots,z_n)$ is weighted homogeneous of type $(w_1,\dots,w_n)$, where $w_1,\dots,w_n$ are fixed positive rational numbers, if it can be expressed as a linear combination of monomials $z_1^{i_1}\dots z_n^{i_n}$ for which $i_1/w_1+\dots+i_n/w_n=1$. As a consequence of the theorem of Merle–Teissier, for isolated singularity defined by a weighted homogeneous polynomial, computing the geometric genus is equivalent to counting the number of positive integral points in the tetrahedron defined by $x_1/w_1+\dots+x_n/w_n\le 1$ and $x_1\ge 0,\dots,x_n\ge 0$. We also need the following result:

Theorem 2.2 [Milnor and Orlik 1970]. *Let $f(z_1,\dots,z_n)$ be a weighted homogeneous polynomial of type $(w_1,\dots,w_n)$ with isolated singularity at the origin. Then the Milnor number is $\mu=(w_1-1)\dots(w_n-1)$.*

The following theorem is about the relation of weight and multiplicity:

Theorem 2.3 [Sękalski 2008]. *If f is a quasihomogeneous isolated singularity of type $(\omega_1,\dots,\omega_n)$, then* $\operatorname{mult}(f)=\min\{\,m\in\mathbb{N}:m\ge\min\{\,\omega_i:i=1,\dots,n\,\}\}$.

There is a lower bound for the p_g of a hypersurface singularity:

Theorem 2.4 [Yau 1977]. *Let*

$$f(z_1,\dots,z_{n-1},z_n)=z_n^m+a_1(z_1,\dots,z_{n-1})z_n^{m-1}+\dots+a_m(z_1,\dots,z_{n-1})$$

be holomorphic near $(0,\dots,0)$. Let d_i be the order of the zero of $a_i(z_1,\dots,z_{n-1})$ at $(0,\dots,0)$ with $d_i\ge i$. Let $d=\min_{1\le i\le m}(d_i/i)$. Suppose that

$$V=\{\,(z_1,\dots,z_n):f(z_1,\dots,z_n)=0\,\},$$

defined in a suitably small polydisc, has $p=(0,\dots,0)$ as its only singularity. Let $\pi:M\to V$ be resolution of V. Then $\dim H^{n-2}(M,\mathbb{O})>(m-1)d-(n-1)$.

In the following theorem, it is convenient for us to use another definition of weight type. Let $f\in\mathbb{C}\{z_1,\dots,z_n\}$ define an isolated singularity at the origin. Let $w=(w_1,\dots,w_n)$ be a weight on the coordinates $(z_1,\dots,z_n)$ by positive integer numbers w_i for $i=1,\dots,n$. We have the weighted Taylor expansion $f=f_\rho+f_{\rho+1}+\cdots$ with respect to w and $f_\rho\ne 0$, where f_k is a weighted homogeneous of type $(w_1,\dots,w_n;k)$ for $k\ge\rho$, i.e., f_k is linear combination of monomials $z_1^{i_1}\dots z_n^{i_n}$ for which $i_1w_1+\dots+i_nw_n=k$. We only use this definition of weight for the following theorem as well as in the proof of Proposition A. For any other place we use the previous definition before Theorem 2.2 for weight type.

Theorem 2.5 [Furuya and Tomari 2004]. *Let the situation be as above, and let $f\in\mathbb{C}\{z_1,\dots,z_n\}$ define an isolated singularity at the origin. Then*

$$\mu(f)\ge\Bigl(\frac{\rho}{w_1}-1\Bigr)\dots\Bigl(\frac{\rho}{w_n}-1\Bigr), \tag{2-1}$$

and equality holds if and only if f_ρ defines an isolated singularity at the origin.

Here we recall that f is called a semiquasihomogeneous function if the initial term f_ρ defines an isolated singularity at the origin.

Definition 2.1. Let $f, g : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be germs of holomorphic functions defining the respective isolated hypersurface singularities $V_f = \{ z : f(z) = 0 \}$ and $V_g = \{ z : g(z) = 0 \}$. Let $\phi : (\mathbb{C}^n, 0) \to (\mathbb{C}^n, 0)$ be a germ of biholomorphic map.

(1) If $\phi(V_f) = V_g$, then f is *contact equivalent* to g.

(2) If $g = f \circ \phi$, then f is *right equivalent* to g.

The Milnor number is an invariant of contact equivalence [Teissier 1975].

3. Proof of the main theorems

Proof of Proposition A. Let $f(z_1, \dots, z_n) : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a holomorphic function with an isolated singularity at the origin. Let μ and ν be the Milnor number and multiplicity of the singularity $V = \{ z : f(z) = 0 \}$. By an analytic change of coordinates, one can assume that the z_n-axis is not contained in the tangent cones of V so that $f(0, \dots, 0, z_n) \neq 0$. By the Weierstrass preparation theorem, near 0, the germ f can be represented as a product

$$f(z_1, \dots, z_n) = u(z_1, \dots, z_n)\, g(z_1, \dots, z_n),$$

where $u(0, \dots, 0) \neq 0$, and

$$g(z_1, \dots, z_{n-1}, z_n) = z_n^\nu + a_1(z_1, \dots, z_{n-1}) z_n^{\nu-1} + \dots + a_\nu(z_1, \dots, z_{n-1}),$$

where ν is the multiplicity of $f(z_1, \dots, z_n)$ and $a_i \in (x_1, \dots, x_{n-1})^i$ for $i = 1, \dots, \nu$. Therefore, $f(z_1, \dots, z_n)$ is contact equivalent to $g(z_1, \dots, z_n)$.

Let d_i be the order of the zero of $a_i(z_1, \dots, z_{n-1})$ at $(0, \dots, 0)$, $d_i \geq i$. Let $d = \min_{1 \leq i \leq \nu} [d_i / i]$, so $d \geq 1$. We define a weight w on the new coordinate systems by $w(z_n) = d$ with $w(z_i) = 1$ for $1 \leq i \leq n-1$. Here the definition of weight type is the same as in Theorem 2.5. With respect to the new weights, z_n^ν has degree $d\nu$, and $a_i(z_1, \dots, z_{n-1}) z_n^{\nu-i}$ has degree at least $d(\nu - i) + d_i \geq d\nu - di + di = d\nu$. Thus, the initial term of $f(z_1, \dots, z_n)$ has the degree $\rho = d\nu$. Because the Milnor number is an invariant under contact equivalence, by (2-1) we have $\mu = \mu(g) \geq (d\nu/d - 1)(d\nu/1 - 1) \dots (d\nu/1 - 1) = (\nu - 1)(d\nu - 1)^{n-1} \geq (\nu - 1)^n$.

Suppose f is a semihomogeneous polynomial. Since the Milnor number of f is the same as its initial part (see [Arnold 1974]), $\mu = (\nu - 1)^n$ is obvious.

If $\mu = (\nu - 1)^n$, then by $\mu \geq (\nu - 1)(d\nu - 1)^{n-1} \geq (\nu - 1)^n$, we have $d = 1$, and by the last part of Theorem 2.5, $g_{d\nu}(z_1, \dots, z_n) = g_\nu(z_1, \dots, z_n)$ is a homogeneous polynomial of degree ν defining an isolated singularity. Hence, $f(z_1, \dots, z_n)$ is contact equivalent to a semihomogeneous singularity. $\square$

We prove a lemma that is useful in the proof of Theorem B.

Lemma 3.1. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a weighted homogeneous polynomial of weight type $(w_1, \dots, w_n)$ with an isolated singularity at the origin. If w_i is not an integer, $z_i^{a_i} z_{j_i} \in \operatorname{supp}(f)$, where a_i is a positive integer, $j_i \neq i$, and $[w_i] < \nu$, then $\nu = a_i + 1$ and $w_i/w_{j_i} \neq 1$.*

Proof. Since $z_i^{a_i} z_{j_i} \in \operatorname{supp}(f)$, $a_i/w_i + 1/w_{j_i} = 1$. It follows from the fact that w_i is not an integer that $w_i/w_{j_i} \neq 1$, and $a_i/w_i + 1/w_{j_i} = 1$ implies that $w_i > a_i$. Since $[w_i] < \nu$, by Theorem 2.3, we have $\nu = [w_i] + 1 \geq a_i + 1$. By the definition of multiplicity, we also have $\nu \leq a_i + 1$. Therefore, $\nu = a_i + 1$. □

Proof of Theorem B. We shall give a detailed proof for $k = 5$.

Let $f : (\mathbb{C}^5, 0) \to (\mathbb{C}, 0)$ be a weighted homogeneous polynomial with an isolated singularity at the origin. Let μ and ν be the Milnor number and multiplicity of the singularity $V = \{ z : f(z) = 0 \}$, respectively. We want to show $\mu \geq (\nu - 1)^5$ and that the equality holds if and only if f is a homogeneous polynomial. By Proposition A, it suffices to show that equality holds if and only if f is a homogeneous polynomial. Set $w(z_i) = w_i$ for $1 \leq i \leq 5$. We assume that $2 \leq w_1 \leq \min\{w_2, \dots, w_5\}$, where w_i for $i = 1, \dots, 5$ are positive rational numbers, without loss of generality.

If $\nu = 2$, then the theorem is trivial by the Milnor–Orlik formula (Theorem 2.2). In the following, we only consider $\nu \geq 3$ or, equivalently, $w_1 > 2$.

If w_1 is an integer, then by Theorem 2.3, $\nu = w_1$. Since $\mu = (w_1 - 1) \dots (w_5 - 1)$, $\mu = (\nu - 1)^5$ if and only if $w_1 = w_2 = \cdots = w_5$, i.e., f is an homogeneous polynomial.

If w_1 is not an integer, by Theorem 2.3, $\nu = [w_1] + 1$, where $[w_1]$ denotes the integer part of w_1. We want to show that $\mu > (\nu - 1)^5$. Since f is an isolated singularity, for every $i \in \{1, \dots, 5\}$, either $z_i^{a_i}$ or $z_i^{a_i} z_j$ is in the support of f, where $j \neq i$ and a_i is a positive integer. By assumption, w_1 is not an integer, so $z_1^{a_1} z_{j_1} \in \operatorname{supp}(f)$. By Lemma 3.1, we have $\nu = a_1 + 1$. We shall show that $(\nu - 1)^2 < (w_1 - 1)(w_{j_1} - 1)$. Since $a_1/w_1 + 1/w_{j_1} = 1$, $a_1 = w_1 - w_1/w_{j_1}$ and $\nu = w_1 - w_1/w_{j_1} + 1$. Therefore, the fraction part of w_1 is w_1/w_{j_1}. In order to make the notation simple, we set $x = [w_1]$, where $x \geq 2$, and $y = w_1/w_{j_1}$, where $0 < y < 1$, and then $x = \nu - 1$, $w_1 = x + y$, and $w_{j_1} = (x + y)/y$. By a simple calculation, $(\nu - 1)^2 < (w_1 - 1)(w_{j_1} - 1)$ is the same as $x^2 < (x + y - 1)((x + y)/y - 1)$, which is true for $x \geq 2$.

We consider $\{w_1, \dots, w_5\} \setminus \{w_1, w_{j_1}\}$, the set of three rational numbers obtained from $\{w_1, w_2, w_3, w_4, w_5\}$ by removing w_1 and w_{j_1}. Without loss of generality, we assume that $w_2 \in \{w_1, \dots, w_5\} \setminus \{w_1, w_{j_1}\}$ (which is not the empty set) is the minimal weight in this set.

If w_2 is a positive integer, then $\nu \leq w_2$; hence, $\nu - 1 \leq w_2 - 1$. Since w_2 is the minimal weight in the set $\{w_1, \dots, w_5\} \setminus \{w_1, w_{j_1}\}$, we have $\mu > (\nu - 1)^5$.

If w_2 is not a positive integer and $[w_2] > [w_1]$, then we have $\nu - 1 < w_2 - 1$. The same reason as before gives $\mu > (\nu - 1)^5$.

If w_2 is not a positive integer and $[w_2] = [w_1]$, then our goal is to prove that $(\nu-1)^2 < (w_2-1)(w_{j_2}-1)$, where w_{j_2} depends on w_2. Since w_2 is not an integer, there exists a_2, a positive integer number such that $z_2^{a_2} z_{j_2} \in \operatorname{supp} f$, where $j_2 \neq 2$. There are three cases to consider:

Case 1. If $j_2 = 1$, then $z_2^{a_2} z_1 \in \operatorname{supp} f$. Then $a_2/w_2 + 1/w_1 = 1$, and $w_2 \geq w_1$ implies $a_2 + 1/w_2 \leq 1$ and $w_2 \geq (a_2+1) \geq \nu$, which contradicts $\nu = [w_1]+1 = [w_2]+1$. This case cannot happen.

Case 2. If $j_2 \in \{1, \dots, 5\} \setminus \{1, 2, j_1\}$, then $a_2/w_2 + 1/w_{j_2} = 1$ since $z_2^{a_2} z_{j_2} \in \operatorname{supp} f$. We want to show that $(\nu-1)^2 < (w_2-1)(w_{j_2}-1)$. We have $\nu \leq a_2+1$, and $\nu = a_1+1$ implies $a_2 \geq a_1$. Furthermore, $a_2 = w_2 - w_2/w_{j_2} \geq a_1 \geq \nu - 1$. Let $x = w_2/w_{j_2}$, so by Lemma 3.1 we have $x \neq 1$. Then $0 < x < 1$, $w_2 \geq \nu - 1 + x$, and $w_{j_2} \geq (\nu-1+x)/x$. It suffices to show that $(\nu-1)^2 < (\nu-1+x-1)((\nu-1+x)/x-1)$, which is true for $\nu > 2$ and $0 < x < 1$.

Case 3. If $j_2 = j_1$, then $z_2^{a_2} z_{j_1} \in \operatorname{supp} f$. Since f has isolated singularity and both $z_1^{a_1} z_{j_1} \in \operatorname{supp} f$ and $z_2^{a_2} z_{j_1} \in \operatorname{supp} f$, then either $z_1^{b_1} z_2^{b_2} \in \operatorname{supp} f$, where $b_i > 0$ and $i = 1, 2$, or $z_1^{b_1} z_2^{b_2} z_{j_{12}} \in \operatorname{supp} f$, where $b_i \geq 0$ for $i = 1, 2$. However, in the latter case b_1 and b_2 cannot both equal 0 and $j_{12} \in \{1, \dots, 5\} \setminus \{1, 2, j_1\}$.

Subcase 1. We have $z_1^{b_1} z_2^{b_2} \in \operatorname{supp} f$, where $b_i > 0$ for $i = 1, 2$. In this case, we have $b_1/w_1 + b_2/w_2 = 1$. Then $b_1/w_2 + b_2/w_2 \leq b_1/w_1 + b_2/w_2 = 1$, which implies $w_2 \geq b_1 + b_2 \geq \nu$, contradicting $\nu - 1 = [w_1] = [w_2]$. This case cannot happen.

Subcase 2. Now we have $z_1^{b_1} z_2^{b_2} z_{j_{12}} \in \operatorname{supp} f$, where $b_i \geq 0$ for $i = 1, 2$ and $j_{12} \in \{1, \dots, 5\} \setminus \{1, 2, j_1\}$. In this case we divide it into three subcases:

(a) If $b_1 = 0$, then $z_2^{b_2} z_{j_{12}} \in \operatorname{supp} f$. This case is same as the previous Case 2.

(b) If $b_2 = 0$, then $z_1^{b_1} z_{j_{12}} \in \operatorname{supp} f$. By Lemma 3.1, we have $\nu = b_1 + 1$. Therefore, $a_1 = b_1$. Remember that we also have $a_1/w_1 + 1/w_{j_1} = 1$; thus, $w_{j_1} = w_{j_{12}}$. Since we have proved $(\nu-1)^2 < (w_1-1)(w_{j_1}-1)$, then we get $(\nu-1)^2 < (w_2-1)(w_{j_{12}}-1)$.

(c) If $b_1 \neq 0$ and $b_2 \neq 0$, then $b_1/w_1 + b_2/w_2 + 1/w_{j_{12}} = 1$, which implies that $(b_1+b_2)/w_2 + 1/w_{j_{12}} \leq 1$. Since $\nu \leq b_1 + b_2 + 1$ and $\nu = a_1 + 1$, $a_1 \leq b_1 + b_2$. Then $a_1/w_2 + 1/w_{j_{12}} \leq 1$, so $a_1 \leq w_2 - w_2/w_{j_{12}}$. Since $j_{12} \in \{1, \dots, 5\} \setminus \{1, 2, j_1\}$, then $w_{j_{12}} \geq w_2$. If $w_{j_{12}} = w_2$, then $w_2 \geq a_1 + w_2/w_{j_{12}} = a_1 + 1$, which contradicts $[w_1] = [w_2] = \nu - 1$, so $w_{j_{12}} > w_2$. Let $x = w_2/w_{j_{12}}$, so $0 < x < 1$. Since $w_2 \geq a_1 + x$, then $w_{j_{12}} \geq (a_1+x)/x$. We want to show that $(\nu-1)^2 < (w_2-1)(w_{j_{12}}-1)$. It suffices to show that $a_1^2 < (a_1+x-1)((a_1+x)/x-1)$, which follows from $0 < (a_1-1)(1-x)$, where $a_1 \geq 2$ and $0 < x < 1$.

After the above steps, either we finish the proof, or after reordering the subindex, we have proved $(\nu-1)^4 < (w_1-1)(w_2-1)(w_{j_1}-1)(w_{j_2}-1)$, where z_1, z_2, z_{j_1}, and z_{j_2} are different variables. There is only one variable left. Without loss of generality, we use z_3 to denote the remaining variable. We know $w_3 \geq w_2 \geq w_1$, and w_1 and w_2 are not positive integers by the previous arguments.

If w_3 is a positive integer, or w_3 is not a positive integer and $[w_3] > [w_1]$, then we have $\nu \leq w_3$ and $\nu - 1 \leq w_3 - 1$. Therefore, $\mu > (\nu-1)^5$ in this case. The proof ends.

Suppose that w_3 is not a positive integer and $[w_3] = [w_1]$. Since $w_3 \geq w_1$, we have $w_3 - 1 \geq w_1 - 1$. We have already proved $(\nu-1)^2 < (w_1-1)(w_{j_1}-1)$ and $(\nu-1)^2 < (w_2-1)(w_{j_2}-1)$. In order to prove $(\nu-1)^5 < \mu$, it suffices to show that $(\nu-1)^3 < (w_1-1)^2(w_{j_1}-1)$. In order to make the notation simple, we set $x = [w_1]$, where $x \geq 2$, and $y = w_1/w_{j_1}$, where $0 < y < 1$. Then $x = \nu - 1$, $w_1 = x + y$, and $w_{j_1} = (x+y)/y$. By simple calculation, $(\nu-1)^3 < (w_1-1)^2(w_{j_1}-1)$ is equivalent to $x^3 \leq (x+y-1)^2((x+y)/y-1)$, i.e., $x(x-2)(1-y)+(y-1)^2 > 0$, which follows from $x \geq 2$ and $0 < y < 1$.

In summary, we have proved $(\nu-1)^5 = \mu$ if and only if f is a homogeneous polynomial.

For $k=6$, using the same argument as $k=5$, we obtain $(\nu-1)^2 < (w_1-1)(w_{j_1}-1)$ and $(\nu-1)^2 < (w_2-1)(w_{j_2}-1)$. Without loss of generality we assume w_3 and w_4 are the remaining two weights. Then the same argument as above shows that $(\nu-1)^3 < (w_1-1)(w_{j_1}-1)(w_3-1)$ and $(\nu-1)^3 < (w_2-1)(w_{j_2}-1)(w_4-1)$. Thus, we have $(\nu-1)^6 < (w_1-1)\dots(w_6-1) = \mu$, which is what we want for proving $(\nu-1)^6 = \mu$ if and only if the f is homogeneous polynomial. $\square$

Proof of Theorem C. If $p_g > 0$, it follows from Theorem 1.6.

If $p_g = 0$, then by Theorem 2.4, $0 > (\nu-1)d - 4$, where $d = \min_{1 \leq i \leq \nu}(d_i/i)$, and d_i is the order of the zero of $a_i(x_1, \dots, x_4)$ at $(0, \dots, 0)$ with $d_i \geq i$. Then $\nu < 4/d + 1$. Since $d \geq 1$, ν is an integer of at least 2 for isolated hypersurface singularities, so $2 \leq \nu \leq 4$. Therefore, $p(\nu) = (\nu-1)^5 - \nu(\nu-1)\dots(\nu-4) = (\nu-1)^5$. The theorem is reduced to proving that

$$\mu \geq (\nu-1)^5,$$

and equality holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates. The proof follows from Theorem B. $\square$

Proof of Theorem D. It follows from the same argument in the proofs of Theorem C and Theorem B. $\square$

Proof of Corollary E. It follows from Theorem B and Theorem 1.1. $\square$

References

[Arnold 1974] V. I. Arnold, "Normal forms of functions in the neighborhood of degenerate critical points", *Uspehi Mat. Nauk* **29**:2(176) (1974), 11–49. In Russian; translated in *Russian Math. Surveys* **29**:2 (1974), 10–50. MR 58 #24324

[Chen et al. 2011] I. Chen, K.-P. Lin, S. Yau, and H. Zuo, "Coordinate-free characterization of homogeneous polynomials with isolated singularities", *Comm. Anal. Geom.* **19**:4 (2011), 661–704. MR 2880212 Zbl 06050014

[Furuya and Tomari 2004] M. Furuya and M. Tomari, "A characterization of semi-quasihomogeneous functions in terms of the Milnor number", *Proc. Amer. Math. Soc.* **132**:7 (2004), 1885–1890. MR 2005f:14007 Zbl 1052.32023

[Lin and Yau 2004] K.-P. Lin and S. S.-T. Yau, "Classification of affine varieties being cones over nonsingular projective varieties: hypersurface case", *Comm. Anal. Geom.* **12**:5 (2004), 1201–1219. MR 2006a:14102 Zbl 1072.32022

[Lin et al. 2006a] K.-P. Lin, Z. Tu, and S. S. T. Yau, "Characterization of isolated homogeneous hypersurface singularities in $\mathbb{C}^4$", *Sci. China Ser. A* **49**:11 (2006), 1576–1592. MR 2008c:32037 Zbl 1115.32017

[Lin et al. 2006b] K.-P. Lin, X. Wu, S. S.-T. Yau, and H.-S. Luk, "A remark on lower bound of Milnor number and characterization of homogeneous hypersurface singularities", *Comm. Anal. Geom.* **14**:4 (2006), 625–632. MR 2008a:32026 Zbl 1111.32025

[Merle and Teissier 1980] M. Merle and B. Teissier, "Conditions d'adjonction (D'apres Du Val)", pp. 229–245 in *Séminaire sur les singularités des surfaces* (Palaiseau, 1976–1977), edited by M. Demazure et al., Lecture Notes in Mathematics **777**, Springer, Berlin, 1980. Zbl 0461.14009

[Milnor and Orlik 1970] J. Milnor and P. Orlik, "Isolated singularities defined by weighted homogeneous polynomials", *Topology* **9** (1970), 385–393. MR 45 #2757 Zbl 0204.56503

[Saito 1971] K. Saito, "Quasihomogene isolierte Singularitäten von Hyperflächen", *Invent. Math.* **14** (1971), 123–142. MR 45 #3767 Zbl 0224.32011

[Sękalski 2008] M. Sękalski, "On the multiplicity of a quasi-homogeneous isolated singularity", *Univ. Iagell. Acta Math.* **46** (2008), 95–97. MR 2010h:32037 Zbl 1181.14008

[Teissier 1975] B. Teissier, "Introduction to equisingularity problems", pp. 593–632 in *Algebraic geometry* (Arcata, CA, 1974), edited by R. Hartshorne, Proc. Sympos. Pure Math. **29**, Amer. Math. Soc., Providence, RI, 1975. MR 54 #10247 Zbl 0322.14008

[Xu and Yau 1993] Y.-J. Xu and S. S.-T. Yau, "Durfee conjecture and coordinate free characterization of homogeneous singularities", *J. Differential Geom.* **37**:2 (1993), 375–396. MR 94f:32073 Zbl 0793.32016

[Yau 1977] S. S. T. Yau, "Two theorems on higher dimensional singularities", *Math. Ann.* **231**:1 (1977), 55–59. MR 58 #11511 Zbl 0343.32010

Received February 8, 2012.

STEPHEN S.-T. YAU
DEPARTMENT OF MATHEMATICAL SCIENCE
TSINGHUA UNIVERSITY
BEIJING 100084
CHINA
yau@uic.edu

HUAIQING ZUO
DEPARTMENT OF MATH, STATISTICS AND COMPUTER SCIENCE (M/C 249)
UNIVERSITY OF ILLINOIS AT CHICAGO
CHICAGO, IL 60607-7045
UNITED STATES
hqzuo@uic.edu

Guidelines for Authors

Authors may submit manuscripts at msp.berkeley.edu/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 260 No. 1 November 2012
