Pacific Journal of Mathematics

Volume 305 No. 1 March 2020

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

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

The subscription price for 2020 is US $520/year for the electronic version, and $705/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.1

THE POINCARÉ HOMOLOGY SPHERE, LENS SPACE SURGERIES, AND SOME KNOTS WITH TUNNEL NUMBER TWO

KENNETH L. BAKER

APPENDIX WRITTEN JOINTLY WITH NEIL R. HOFFMAN

We exhibit an infinite family of knots in the Poincaré homology sphere with tunnel number 2 that have a lens space surgery. Notably, these knots are not doubly primitive and provide counterexamples to a few conjectures. Additionally, we update (and correct) our earlier work on Hedden's almost-simple knots. In the appendix, it is shown that a hyperbolic knot in the Poincaré homology sphere with a lens space surgery has either no symmetries or just a single strong involution.

1. A family of knots in the Poincaré homology sphere with lens space surgeries

Figure 1 shows a one-parameter family of arcs κ_n, $n \in \mathbb{Z}$, on the pretzel knot $P(-2,3,5) = P(3,5,-2)$ for which a -1-framed banding produces the two bridge knot $B(3n^2+n+1, -3n+2)$. Indeed,

$$[n, -1, -n, 3] = -\frac{3n^2+n+1}{-3n+2}.$$

Through double branched coverings and the Montesinos trick, the arc κ_n lifts to a knot K_n in the Poincaré homology sphere on which an integral surgery yields the lens space $L(3n^2+n+1, -3n+2)$.

As presented, the arc κ_n lies in a bridge sphere that presents $P(-2,3,5)$ as a 3-bridge link. Thus the knot K_n lies in a genus 2 Heegaard surface for the Poincaré homology sphere $\mathcal{P}$. Nevertheless, as we shall see, generically K_n does not have tunnel number one.

Theorem 1. *For each integer n, the knot K_n in the Poincaré homology sphere $\mathcal{P}$ has an integral surgery to the lens space $L(3n^2+n+1, -3n+2)$. If $|n| \geq 4$, then K_n has tunnel number 2.*

MSC2010: 57M27.
Keywords: Poincaré homology sphere, Dehn surgery, lens space, tunnel number.

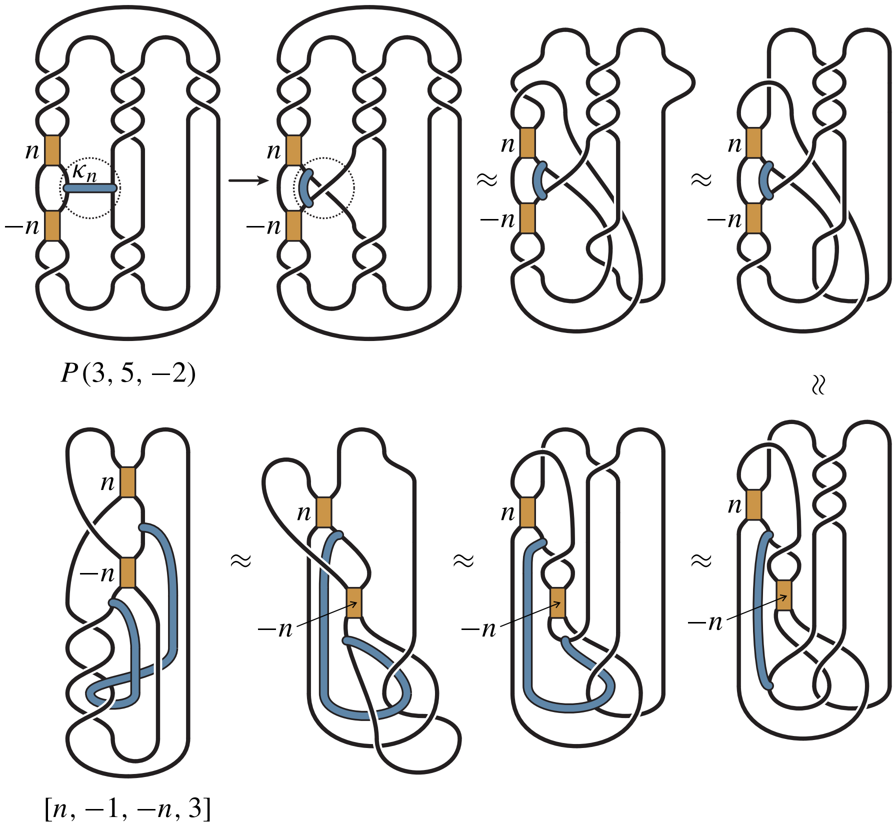

Figure 1. Top left shows a family of arcs κ_n with end points on the pretzel knot $P(-2, 3, 5)$. A numbered vertical box signifies a stack of half-twists where the sign informs the handedness. We perform a banding on κ_n (framed relative to the horizontal bridge sphere) followed by a sequence of isotopies to recognize the result as the two-bridge knot $B(3n^2 + n + 1, -3n + 2)$. The arc κ_n^* dual to the banding is also carried along.

These knots arose from an application of the "longitudinally jointly primitive" ideas presented in [Baker et al. 2019]. After observing that the negative Whitehead link is a two-component genus 1 fibered link for which -1 surgery on both components produces $\mathcal{P}$, we found a family of jointly primitive knots in the twice-punctured torus bundle that is the exterior of this Whitehead link. Our knots K_n are the images of these jointly primitive knots in the $\mathcal{P}$ filling. This is discussed further in [Baker et al. 2019, §2.3.3].

Previously, the only known knots in $\mathcal{P}$ admitting a surgery to a lens space are surgery duals to Tange knots [2009a] and certain Hedden knots [Hedden 2011;

Rasmussen 2007; Baker 2014b]. (See Section 5 for a discussion of the Hedden knots.) These knots in $\mathcal{P}$ are all *doubly primitive*, they may be presented as curves in a genus 2 Heegaard splitting that represent a generator in π_1 of each handlebody bounded by the Heegaard surface. We see this as Tange knots and Hedden knots all admit descriptions by doubly pointed genus 1 Heegaard diagrams. (Tange knots are all simple knots in lens spaces and Hedden knots are "almost simple".) Equivalently, they are all 1-bridge knots in their lens spaces. Any integral surgery on a 1-bridge knot in a lens space naturally produces a 3-manifold with a genus 2 Heegaard surface in which the dual knot sits as a doubly primitive knot.

The Berge conjecture posits that any knot in S^3 with a lens space surgery is a doubly primitive knot [Berge 1990]; see [Gordon 1991, Question 5.5]. It is also proposed that any knot in $S^1 \times S^2$ with a lens space surgery is a doubly primitive knot [Baker et al. 2016, Conjecture 1.1; Greene 2013, Conjecture 1.9]. By analogy and due to a sense of simplicity[1], one may suspect that any knot in $\mathcal{P}$ with a lens space surgery should also be doubly primitive. However this is not the case. Since our knots K_n have tunnel number 2 for $|n| \geq 4$, they cannot be doubly primitive.

Corollary 2. *There are knots in the Poincaré homology sphere with lens space surgeries that are not doubly primitive. That is, the analogy to the Berge conjecture fails for the Poincaré homology sphere.*

Corollary 3. *[Greene 2013, Conjecture* 1.10*] is false. For example, the lens space* $L(5, 1)$ *contains a knot* K_1^* *with a surgery to* $\mathcal{P}$ *but does not contain a Tange knot or a Hedden knot with a* $\mathcal{P}$ *surgery.*

Proof. For $n = 1$, surgery on K_1 produces $L(5, -1)$ as shown in Figure 1. A quick check of [Tange 2009a, Table 2] shows that $L(5, -1)$ contains no Tange knots. The Hedden knots with surgery to $\pm\mathcal{P}$ are homologous to Berge lens space knots of type VII [Rasmussen 2007; Greene 2013]: one further observes that $L(5, -1)$ does not contain any of these knots since $k^2 + k + 1 = 0 \pmod 5$ has no integral solution. □

Remark 4. The knot K_1 is actually the $(2, 3)$-cable of the knot J that is surgery dual to the trefoil in S^3.

As we shall see in Theorem 7(3), it is the $(2g(K_n) - 1)$-surgery on our knots K_n that produces a lens space; $g(K_n)$ is the Seifert genus of the knot K_n. Hence they also provide counterexamples to [Hedden 2011, Conjecture 1.7] (see [Baker et al. 2008, Conjecture 1.6]) and [Greene 2013, Conjecture 1.10]. It seems plausible that it is the largeness of the knot genus with respect to the lens space surgery

[1] The manifolds S^3, $\mathcal{P}$, and its mirror $-\mathcal{P}$ are the only homology 3-spheres with finite fundamental group (by Perelman) and the only known irreducible L-space homology 3-spheres, e.g., [Eftekhary 2009; Eftekhary 2018].

slope that enables the failure of our knots to be doubly primitive. Hedden and Rasmussen also observe a distinction at this slope for lens space surgeries on knots in L-space homology spheres [Hedden 2011; Rasmussen 2007]. Indeed, from the view of Heegaard Floer homology for integral slopes on knots in homology spheres, this slope is right at the threshold at which a knot could have an L-space surgery [Ozsváth and Szabó 2011, Proposition 9.6] (see also [Hedden 2009, Lemma 2.13])[2] and just below what implies it has simple knot Floer homology [Eftekhary 2011][3]. With this in mind, we adjust and update [Greene 2013, Conjecture 1.10].

Conjecture 5. *Suppose that p-surgery on a knot K in the Poincaré homology sphere produces a lens space. If $p > 2g(K)-1$ is an integer, then K is a doubly primitive knot. Furthermore it is surgery dual to one of the Tange knots.*

Conjecture 6. *If a knot in the Poincaré homology sphere is doubly primitive, then it is surgery dual to one of the Tange knots or one of the Hedden knots.*

As mentioned in the paragraph following its statement, [Hedden 2011, Theorem 3.1] (presented here in Theorem 14) applies to any L-space homology sphere, not just S^3. Thus, since doubly primitive knots are surgery dual to one-bridge knots in lens spaces, [Hedden 2011, Theorem 3.1] implies that such surgery duals are either simple knots or one of the Hedden knots T_L or T_R. In Section 5 we clarify and correct our work in [Baker 2014b] and further classify when Hedden's knots are dual to integral surgeries on knots homology spheres, $\mathcal{P}$ in particular. Tange has produced a list of simple knots in lens spaces that are surgery dual to knots in $\mathcal{P}$ [Tange 2009a], and it is expected that this list is complete. Verifying its completeness will affirm Conjecture 6.

Let us collect various properties of our knots.

Theorem 7. *Let $p = 3n^2+n+1$ and $q = -3n+2$. Then the following hold for our family of knots K_n in $\mathcal{P}$:*

(1) *Positive p-surgery on K_n produces a lens space $L(p,q)$.*

(2) *K_n is fibered and supports the tight contact structure on $\mathcal{P}$.*

(3) $g(K_n) = \frac{1}{2}(p+1)$

(4) $\operatorname{rk}\widehat{\mathrm{HFK}}(L(p,q), K_n^*) = p+2$

(5) *Let T_n be the $(3n+1,n)$-torus knot in S^3, and note that p-surgery on T_n also produces $L(p,q)$. The surgery dual to K_n is homologous to the surgery dual to T_n.*

(6) $\Delta_{K_n}(t) = \Delta_{T_n}(t) - (t^{(p-1)/2}+t^{-(p-1)/2}) + (t^{(p+1)/2}+t^{-(p+1)/2})$.

[2] This threshold may be lower for knots in homology spheres with $\tau(K) < g(K)$.

[3] A knot K^* in a rational homology sphere Y has *simple knot Floer homology* if $\operatorname{rk}\widehat{\mathrm{HFK}}(Y,K^*) = \operatorname{rk}\widehat{\mathrm{HF}}(Y)$. This definition does not require Y to be an L-space itself.

Theorem 7(5) is given as Lemma 12. The remainder of Theorem 7 follows from assembling the works [Hedden 2011; Rasmussen 2007; Tange 2011, 2009b]. In fact, appealing to Greene's proof of the lens space realization problem [Greene 2013], we tease out the following general theorem from which Theorem 7(3) and (6) follow.

Theorem 8. *Suppose that p-surgery on knot K in an L-space homology sphere Y with $d(Y) = 2$ produces the lens space $L(p, q)$. Then $p = 2g(K) - 1$ if and only if p-surgery on some Berge knot B in S^3 also produces the lens space $L(p, q)$ in which the surgery duals to K and B are homologous.*

When this holds, $\Delta_K(t) = \Delta_B(t) - (t^{(p-1)/2} + t^{-(p-1)/2}) + (t^{(p+1)/2} + t^{-(p+1)/2})$.

2. Questions about knots in the Poincaré homology sphere with lens space surgeries.

2.1. *Homology classes of knots.*

Question 9. (1) For which Berge knots B that have a positive, odd p-surgery producing the lens space $L(p, q)$, is there a knot K in $\mathcal{P}$ such that $2g(K)-1 = p$, p-surgery produces the lens space $L(p, q)$, and the surgery duals B^* and K^* are homologous?

(2) For a Berge knot B as above, can there be two distinct knots K_1 and K_2 in $\mathcal{P}$ with p-surgery to $L(p, q)$ such that the surgery duals B^*, K_1^*, and K_2^* are all homologous?

Presently the only Berge knots known to answer Question 9(1) are the Berge knots of type VII (those that embed in the fiber of the trefoil) due to [Hedden 2011; Baker 2014b] and the $(3n+1, n)$-torus knots with $p = 3n^2 + n + 1$ due to our knots introduced here. Indeed, what about the case of the $(3n + 1, n)$-torus knots when $p = 3n^2 + n - 1$?

To clarify the word "distinct" in Question 9(2), note that the mapping class group of the Poincaré homology sphere is trivial [Boileau and Otal 1991, Thèorém 3 and Corollaire 4]. Thus knots in $\mathcal{P}$ that are related by a diffeomorphism of $\mathcal{P}$ are also isotopic.

2.2. ***Symmetries.*** One may observe that, thus far, all the known examples of knots in $\mathcal{P}$ with a lens space surgery are strongly invertible. Hence there is an involution of $\mathcal{P}$ taking each of these knots with some orientation to its reverse. Wang and Zhou [1992] have shown that if a knot in S^3 other than a torus knots has a nontrivial lens space surgery, then its only possible symmetry is a strong involution.

Question 10. (1) Is every knot in $\mathcal{P}$ with a lens space surgery strongly invertible?

(2) What are the possible symmetries of a knot in $\mathcal{P}$ with a lens space surgery?

In the appendix to this article, the author and Hoffman provide Theorem A.1 which addresses Question 10(2): A hyperbolic knot in $\mathcal{P}$ with a lens space surgery either has no symmetries or just a single strong involution.

2.3. ***Hopf plumbings.*** Let us inquire about another potential analogy with the Berge knots in S^3. The Giroux correspondence says that any two fibered knots supporting the same contact structure are related by a sequence of plumbings and de-plumbings of positive Hopf bands [Giroux 2002]. Since the Berge knots (with positive surgeries to lens spaces) can all be expressed as closures of positive braids[4], each Berge knot can be obtained from the unknot by a sequence of plumbings of Hopf bands. No de-plumbings are necessary.

In $\mathcal{P}$ the unknot is not fibered, but the knot J that is -1-surgery dual to the (negative) trefoil is. As we shall see in Lemma 13, J is a reasonable surrogate for the unknot in $\mathcal{P}$: it is the unique genus one fibered knot in $\mathcal{P}$, and it supports the tight contact structure on $\mathcal{P}$. Since a knot in $\mathcal{P}$ with a lens space surgery is fibered and supports the tight contact structure (by Theorem 7(2)), the Giroux correspondence says that it may be obtained from J by a sequence of plumbings and de-plumbings of positive Hopf bands.

Question 11. If a knot in $\mathcal{P}$ has a lens space surgery, then can it be obtained from J by a sequence of plumbings of Hopf bands?

We have yet to systematically check this question for the knots produced here or for the duals to the Tange and Hedden knots. As noted in Remark 4, the knot K_1 is the $(2,3)$-cable of J. One may use this to show that K_1 is indeed obtained by plumbing two Hopf bands onto J.

3. Notation and conventions

Let K be a knot in an oriented 3-manifold M. Choose an orientation on K and let μ be a meridian of K in the torus $\partial\mathcal{N}(K)$ that positively links K. Let λ be an oriented curve in $\partial\mathcal{N}(K)$ that is isotopic to K in $\mathcal{N}(K)$; if K is null-homologous in M, we choose λ so that it is null-homologous in $M-\mathcal{N}(K)$. If γ is an essential simple closed curve in $\partial\mathcal{N}(K)$, then when it is oriented $[\gamma]=p[\mu]+q[\lambda]\in H_1(\partial\mathcal{N}(K))$ where p,q are coprime integers; changing the orientation of γ changes the signs of both p and q. We refer to both the unoriented isotopy class of γ in $\partial\mathcal{N}(K)$ and the number $p/q\in\mathbb{Q}\cup\{\infty\}$ as a *slope* and (when λ is null-homologous) say it is *positive* if $0<p/q<\infty$. If the slope is integral, we also say it is a *longitude* or a *framing* of K.

The result of γ-*Dehn surgery* on the knot $K\subset M$ is the manifold $M_K(\gamma)$ obtained by attaching $S^1\times D^2$ to $M-\mathcal{N}(K)$ along the torus $\partial\mathcal{N}(K)$ so that $\mathrm{pt}\times\partial D^2$ is

[4]This is remarked preceding the conjecture of Goda and Teragaito [2000], for example.

identified with the slope γ. The image of the curve $S^1 \times \mathrm{pt}$ in $M_K(\gamma)$ is a knot called the *surgery dual* to K and denoted K^*. Observe that γ is the meridian of K^* and μ-surgery on K^* returns M with surgery dual $K = (K^*)^*$.

If M is the double branched cover of S^3 over a link L, an arc κ such that $\kappa \cap L = \partial\kappa$ lifts to a knot K in the cover M. Via the Montesinos trick [1975], each integral surgery on K corresponds to a *banding* along κ. With $I = [-1, 1]$, if $R = I \times I$ is a disk embedded in S^3 such that $R \cap L = \partial I \times I$ and $\{0\} \times I = \kappa$, then the link $L' = (L - \partial I \times I) \cup I \times \partial I)$ is the result of a banding along κ. The arc $\kappa^* = I \times \{0\}$ is the dual arc of the banding. In the double branched cover over L', κ^* lifts to the dual knot K^* of the corresponding integral surgery on K.

The lens space $L(p, q)$ is defined to be the manifold that results from $-p/q$-surgery on the unknot in S^3. The two-bridge link $B(p, q)$ is the link in S^3 whose double branched cover is $L(p, q)$ [Hodgson and Rubinstein 1985]. Using the continued fraction $[x_1, x_2, \ldots, x_n] = x_1 - 1/(x_2 - 1/(\cdots - 1/x_n))$ to express $-p/q$ describes the two-bridge link $B(p, q)$ geometrically in plat presentation as in the lower left of Figure 1.

A nontrivial knot in a lens space $L(p, q)$ is *simple* (or *grid number one*) if, in the standard genus one Heegaard diagram of the lens space, it may be represented by two of the p intersection points. The simple knot is then the union of the arcs connecting those points in the two meridional disks whose boundaries are described by the diagram. Equivalently, a simple knot (including the trivial knot) is a knot represented by a doubly pointed genus 1 Heegaard diagram of $L(p, q)$ that has p intersection points. There is one for each homology class $k \in H_1(L(p, q))$ in $L(p, q)$, denoted $K(p, q, k)$. See for example [Rasmussen 2007; Hedden 2011; Baker et al. 2008].

4. Proofs

Proof of Theorem 1. Figure 1 exhibits arcs κ_n on the pretzel knot $P(-2, 3, 5)$ that have a banding to a two-bridge link. By passing to the double branched covers (and using the Montesinos trick), this describes knots K_n in $\mathcal{P}$ that have an integral surgery to a lens space.

Since K_n lies in a genus 2 Heegaard surface, its tunnel number is at most 2. If K_n were to have tunnel number one, then every surgery on K_n would have Heegaard genus at most 2. However we will see that surgery along the framing induced by the Heegaard surface produces a toroidal manifold which, according to Kobayashi's classification [1984], does not have Heegaard genus 2. (The same scheme was recently employed in [Eudave-Muñoz et al. 2015] to demonstrate a family of strongly invertible knots in S^3 with a Seifert fibered surgery and tunnel number 2.)

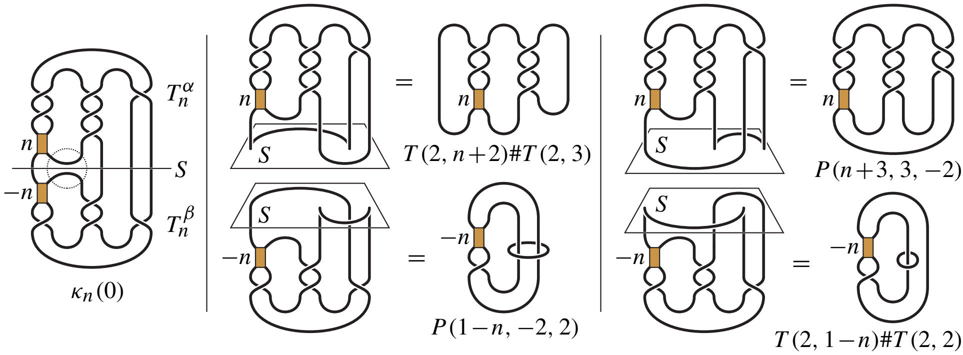

Figure 2. Left: The bridge sphere S for $P(-2, 3, 5)$ splits $\kappa_n(0)$ into two tangles, T_n^α above and T_n^β below. Middle: Filling T_n^α with a rational tangle to get $T(2, n+2)\#T(2, 3)$ makes T_n^β into $P(1-n, -2, 2)$. Right: Filling T_n^β with a rational tangle to get $T(2, 1-n)\#T(2, 2)$ makes T_n^α into $P(n+3, 3, -2)$.

Kobayashi shows that if $M = M^\alpha \cup_T M^\beta$ is a genus 2 manifold decomposed along a torus T into two atoroidal manifolds, then one of M^α or M^β admits a Seifert fibering over the disk with 2 or 3 exceptional fibers or over the Mobius band with up to 2 exceptional fibers such that filling the other along the slope induced by a regular fiber in T produces a lens space. (This is more general than Kobayashi's classification but suitable for our needs.) The slope of a regular fiber may be identified since filling one of these Seifert fibered spaces produces a reducible manifold if and only if the filling is done along the slope of a regular fiber.

Figure 2 (left) shows that the result of a 0-framed banding along κ_n has a sphere S dividing it into two 2-string tangles. Up to homeomorphism, these two tangles are the tangle sums $T_n^\alpha = 1/(n+2) + \frac{1}{3}$ and $T_n^\beta = 1/(1-n) - \frac{1}{2}$. Their corresponding double branched covers, M_n^α and M_n^β, are in general each Seifert fiber spaces over the disk with exactly two exceptional (and nondegenerate) fibers: M_n^α has type $D^2(|n+2|, 3)$ and M_n^β has type $D^2(|1-n|, 2)$. This fails for M_n^α when $n = -1, -2, -3$ and for M_n^β when $n = 0, 1, 2$; in each, the middle value yields a degenerate Seifert fibration while the other values yield solid tori. In these cases the torus T that is the double branched cover of S is compressible and so Kobayashi's classification does not apply; moreover, one may observe that K_n has tunnel number one in these cases. Hence we assume $n \notin \{-3, -2, -1, 0, 1, 2\}$.

When $n = -1$ or 3, M_n^β has type $D^2(2, 2)$ and thus is a twisted I-bundle over the Klein bottle. Therefore it has an alternative Seifert fibration over the Mobius band with no exceptional fibers. We already assume $n \neq -1$. We shall assume $n \neq 3$ as well.

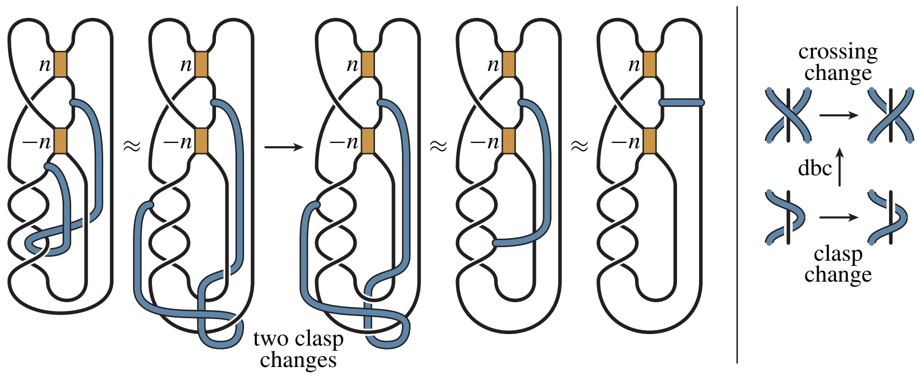

Figure 3. Left: An isotopy of the arc κ_n^* on the two bridge link $B(3n^2+n+1, -3n+2)$ from the lower left of Figure 1 has two clasp changes performed. The resulting arc τ_n^* is then isotoped to lie in a bridge sphere. Right: A clasp change of an arc around a segment of a branch locus lifts to a crossing change about a segment of the fixed set in the double branched cover.

Figure 2 (middle) shows the results of filling T_n^α and T_n^β with the rational tangle ρ^α defined by a pair of arcs in S so that $T_n^\alpha(\rho^\alpha)$ is composite. We then observe that the filling $T_n^\beta(\rho^\alpha)$ is the pretzel link $P(1-n, 2, 2)$, and this is a two-bridge link if and only if $n = 0, 2$. We have already omitted these values of n.

Figure 2 (right) shows the results of filling T_n^α and T_n^β with the rational tangle ρ^β defined by a pair of arcs in S so that $T_n^\beta(\rho^\beta)$ is composite. We then observe that the filling $T_n^\alpha(\rho^\beta)$ is the pretzel link $P(n+2, 3, -2)$, and this is a two-bridge link if and only if $n = -1, -3$. We have already omitted these values of n too.

Therefore, to conclude that the knot K_n has tunnel number 2, it is sufficient to require that $n \notin \{-3, -2, -1, 0, 1, 2, 3\}$. □

Lemma 12. *The knot K_n^* in $L(3n^2+n+1, -3n+2)$ that is surgery dual to K_n is homologous to the knot T_n^* that is surgery dual to the $(3n+1, n)$-torus knot in S^3.*

Proof. Starting from the lower left of Figure 1 where the dual arc κ_n^* is a presented on a standard form of the two bridge link $B(3n^2+n+1, -3n+2)$, Figure 3 (left) shows how two clasp changes (and isotopy) transforms κ_n^* into an arc τ_n^* in a bridge sphere. Since τ_n^* lies in the bridge sphere, its lift to the double branched cover is a knot T_n^* in the Heegaard torus of the lens space. Lifting the transformation of Figure 3 (left) to the double branched cover thus shows that K_n^* and T_n^* are related by two crossing changes. (The clasp changes lift to crossing changes as indicated in Figure 3 (right).) Hence these knots are homotopic, and therefore homologous, in the lens space.

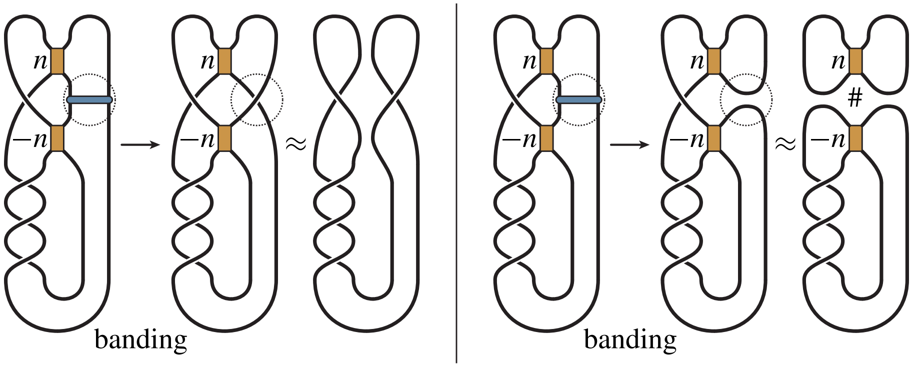

Figure 4. Left: A banding of $B(3n^2+n+1,-3n+2)$ along the arc τ_n^* produces the unknot. Right: A different banding of $B(3n^2+n+1,-3n+2)$ along the arc τ_n^* produces the connected sum of $B(n,-1)$ and $B(3n+1,3)$.

Finally, Figure 3 (left) shows a banding of $B(3n^2+n+1,-3n+2)$ along the arc τ_n^* to the unknot. Thus we may identify T_n^* as surgery dual to a torus knot T_n in S^3. Because Figure 3 (left) shows a banding of $B(3n^2+n+1,-3n+2)$ along the arc τ_n^* to the connected sum $B(n,-1)\#B(3n+1,3)$, the torus knot T_n has a reducible surgery to $L(n,-1)\#L(3n+1,3) \cong L(-n,3n+1)\#L(3n+1,-n)$ which is the mirror of $L(n,3n+1)\#L(3n+1,n)$. Thus T_n must be the $(3n+1,n)$-torus knot (for example, see [Moser 1971]). □

Proof of Theorem 7. (1) Theorem 1 shows that either $\pm p$-surgery on K_n is a lens space. Tange shows that any integral lens space surgery on a knot in $\mathcal{P}$ must be positive [Tange 2011, Theorem 1.1].

(2) In [Tange 2011, Theorem 3.1], the paragraph following, and its proof, Tange further shows that if an L-space homology sphere Y contains a knot K with irreducible exterior for which a positive integer surgery produces an L-space, then K is a fibered knot supporting a tight contact structure. (Tange notes that the fiberedness of K follows from the proof of [Ozsváth and Szabó 2005, Theorem 1.2] and [Ni 2007]. Tange then demonstrates that the Heegaard Floer contact invariant of the contact structure supported by K is nonzero.)

(5) If T_n is the $(3n+1,n)$-torus knot in S^3, then p-surgery on T_n produces the lens space $L(p,q)$. By Lemma 12, the surgery dual knot T_n^* is homologous to the knot K_n^* that is surgery dual to K_n.

(3) and (6) Since the dual knots K_n^* and T_n^* are homologous, Theorem 8 implies $2g(K^n)=p-1$ and the stated relationship of the Alexander polynomials of K_n and T_n.

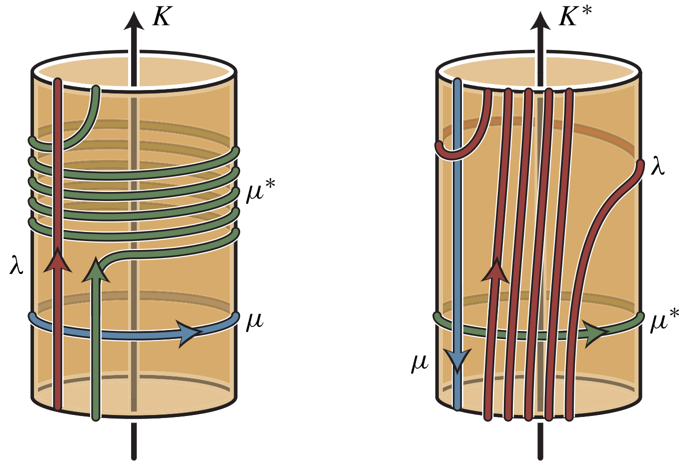

Figure 5. Left: A regular neighborhood of an oriented knot K in a homology sphere Y has meridian μ linking K positively and longitude λ oriented as the boundary of a Seifert surface. The surgery curve μ^* of positive slope $5/1$ shown here is oriented as $[\mu^*] = [\lambda] + 5[\mu]$ in $H_1(\partial\mathcal{N}(K))$. Right: Performing μ^*-surgery on K produces the rational homology sphere Z and surgery dual knot K^*. Orienting K^* to be linked positively by its meridian μ^*, the boundary of the (rational) Seifert surface λ is homologous to $5[K^*]$ in $\mathcal{N}(K^*)$ while the longitude μ is homologous to $-[K^*]$.

(4) Knowing that $g(K_n) = (p+1)/2$ implies $\operatorname{rk}\widehat{\mathrm{HFK}}(L(p,q), K_n^*) = p+2$ [Hedden 2011, Theorem 1.4] (see also [Rasmussen 2007, Proposition 4.5]). □

Proof of Theorem 8. Let K^* be the surgery dual to K in $L(p,q)$ and (for some choice of orientations) let J^* be the simple knot in $L(p,q)$ homologous to K^*. Then J^* has an integral surgery to an L-space homology sphere Y_J [Rasmussen 2007, Theorem 2].

Since *positive* p-surgery on K gives $L(p,q)$, then the self-linking number of K^* is $-1/p$ (mod 1); see [Rasmussen 2007, §2]. This is demonstrated in Figure 5: If Σ is an oriented Seifert surface for K giving the oriented longitude $\lambda = \partial\Sigma$, then the positively linking meridian μ of K is oriented as shown. In order for λ to run positively along K^*, we need $\mu^* \cdot \lambda > 0$. Thus we must orient μ^* so that $[\mu^*] = [\lambda] + p[\mu]$. This forces μ to be an anti-parallel longitude of K^*. Hence we may use $-\mu$ to calculate the self-linking number of K^* as $(-\mu \cdot \lambda)/p$ (mod 1).

With this set-up, μ-surgery on K^* may be regarded as -1-surgery on K^*. This means $Y = K^*_{-1}$ in the notation of [Rasmussen 2007], though not explicitly stated. Similarly we also have $Y_J = J^*_{-1}$.

If $2g(K)-1 \neq p$, then the knots J^* and K^* have isomorphic knot Floer homology [Rasmussen 2007, Theorem 2] (see also [Hedden 2011, Theorem 1.4]). Having isomorphic (and simple) knot Floer homology would imply that the L-space homology spheres Y_J and Y have the same d-invariants. Hence $d(Y_J)=2$ too. But now Y_J cannot be S^3 since $d(S^3)=0$.

If $2g(K)-1=p$ then width $\widehat{\mathrm{HFK}}(L(p,q), K^*)=2p$ [Ni 2009]; see [Rasmussen 2007, Theorem 4.3]. Thus $d(Y)=d(Y_J)+2$ by [Rasmussen 2007, Proposition 5.4] and so $d(Y_J)=0$. Greene's solution [2013] to the lens space realization problem proceeds by first identifying the pairs (lens space, homology class) that contain knots for which some integral surgery produces an L-space homology sphere with $d=0$, and then observing that each of these pairs contains the surgery dual to a Berge knot. Since the lens space surgery duals to Berge knots are simple knots, we have that in fact J^* is surgery dual to a Berge knot B and $Y_J=S^3$.

For the statement about Alexander polynomials, let us first summarize work of Tange [2009b] on Alexander polynomials of knots in L-space homology spheres for which a positive integral surgery yields a lens space. The symmetrized Alexander polynomial of a knot K in an L-space homology sphere with a nontrivial L-space surgery can be expressed as

$$\Delta_K(t)=\sum_{i\in\mathbb{Z}} a_i(K)t^i=(-1)^k+\sum_{j=1}^{k}(-1)^{k-j}(t^{n_j}+t^{-n_j})\in\mathbb{Z}[t^{\pm 1}]$$

following the arguments of [Ozsváth and Szabó 2005]. For a given positive integer p, pass to the quotient $\mathbb{Z}[t^{\pm1}]/(t^p-1)$ to obtain a polynomial $\tilde{\Delta}_K(t)=\sum_{i\in\mathbb{Z}/p\mathbb{Z}}\tilde{a}_i(K)t^i$ with coefficients $\tilde{a}_i(K)=\sum_{j=i \bmod p} a_j(K)$. Assuming that p-surgery gives a lens space Z, then $2g(K)-1\le p$ [Kronheimer et al. 2007]. Note that K is fibered [5] and hence the degree of $\Delta_K(t)$ equals $2g(K)$. Because of this and that the coefficients of $\Delta_K(t)$ are $a_i(K)=0$ or ±1 (and the nonzero ones alternate in sign), it follows that $\tilde{a}_i(K)=0$, ±1, or 2 where $\tilde{a}_i(K)=2$ implies that both p is even and $i=\frac{1}{2}p$ [Tange 2009b, Corollary 2]. Furthermore, one may determine that the coefficients $\tilde{a}_i(K)$ and the polynomial $\tilde{\Delta}_K(t)$ only depend on the lens space Z and the homology class of the surgery dual knot K^* [Brody 1960] (see [Rasmussen 2007, §5]). In particular, if $p=2g(K)-1$ then

$$\Delta_K(t)=\tilde{\Delta}_K(t)-(t^{(p-1)/2}+t^{-(p-1)/2})+(t^{(p+1)/2}+t^{-(p+1)/2})$$

where the indices for the coefficients of $\tilde{\Delta}_K(t)$ are taken to be the representatives of $\mathbb{Z}/p\mathbb{Z}$ from $-\frac{1}{2}(p-1)$ and $\frac{1}{2}(p-1)$. See also the discussion preceding [Tange 2009a, Proposition 3.3].

[5] Since lens spaces are irreducible, a knot in a homology sphere with a lens space surgery must have irreducible exterior. Thus K has irreducible exterior and so it is fibered by [Ozsváth and Szabó 2005, Theorem 1.2] and [Ni 2007]; see Theorem 7(2).

Now in our present situation, since the Berge knot B in S^3 has p-surgery to $L(p, q)$ in which the surgery dual is the simple knot J^*, we therefore have $2g(B) < p$ because p is odd [Hedden 2011, Theorem 3.1][6]. Hence $\tilde{a}_i(B) = a_i(B)$ for all $i \in \mathbb{Z}/p\mathbb{Z}$ and $\Delta_B(t) = \tilde{\Delta}_B(t)$. Because the surgery dual K^* is homologous to the simple knot B^* in $L(p, q)$, $\tilde{\Delta}_K(t) = \tilde{\Delta}_B(t)$. Therefore we have

$$\Delta_K(t) = \Delta_B(t) - (t^{(p-1)/2} + t^{-(p-1)/2}) + (t^{(p+1)/2} + t^{-(p+1)/2})$$

as claimed. □

Lemma 13. *There is a unique genus one fibered knot J in $\mathcal{P}$. As an open book, it supports the tight contact structure on $\mathcal{P}$. Furthermore, it is surgery dual to the negative trefoil knot in S^3.*

Proof. This can be seen in the spirit of [Baker 2014a] which relates genus one fibered knots to axes of closed 3-braids as follows: Noting that since $\mathcal{P}$ has a unique genus 2 Heegaard splitting [Boileau and Otal 1991], $P(-2, 3, 5)$ is the only 3-bridge link whose double branched cover gives $\mathcal{P}$. (In fact $\mathcal{P}$ is the double branched cover of no other link.) Since $P(-2, 3, 5)$ is isotopic to the $(3, 5)$-torus knot, its 3-braid axis A lifts to a genus one fibered knot $J \subset \mathcal{P}$. By [Birman and Menasco 1993, classification theorem] and [Baker 2014a, Lemma 3.8], for instance, A is the only 3-braid axis for $P(-2, 3, 5)$ up to isotopy of unoriented links. Hence J is the only genus one fibered knot up to homeomorphism in $\mathcal{P}$. Since the mapping class group of $\mathcal{P}$ is trivial [Boileau and Otal 1991], J is the only genus one fibered knot in $\mathcal{P}$ up to isotopy.

In the double branched cover, a braid presentation of the branch locus lifts to a Dehn twist presentation of the monodromy of the lift of the axis. Since the axis A presents $P(-2, 3, 5)$ as a positive braid (indeed, as the $(3, 5)$-torus knot), we obtain a presentation of the monodomy of J as a product of positive Dehn twists. Hence it supports the tight contact structure on $\mathcal{P}$ [Giroux 2002].

One may view J as surgery dual to the negative trefoil in several ways. Taking the route through branched covers, this is demonstrated in [Baker 2014b, Proof 2], though for the mirrored situation. □

5. Hedden's almost-simple knots.

In each lens space $L(p, q)$ with coprime $p > q > 0$, Hedden diagramattically describes two unoriented 1-bridge knots T_L and T_R (*the Hedden knots*) via the doubly pointed genus 1 Heeggaard diagrams of $L(p, q)$ with $p + 2$ intersection points [Hedden 2011, Figure 3]. An alternative presentation of local pictures of these diagrams near the two marked points are shown in the top row of Figure 6.

[6] See also [Greene 2013, Theorem 1.4].

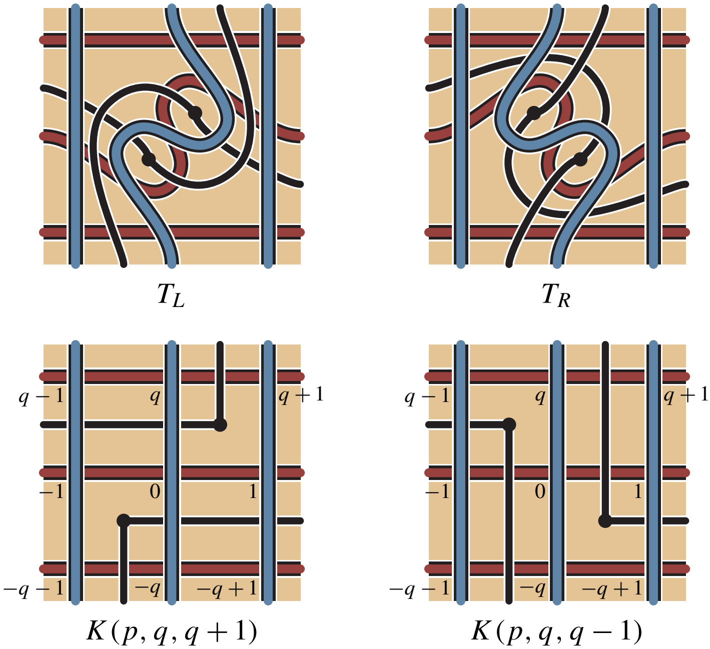

Figure 6. Local portions of the doubly pointed diagrams for the almost-simple knots T_L and T_R and their associated simple knots $K(p, q, q+1)$ and $K(p, q, q-1)$; see [Baker 2014b, Figure 1; Hedden 2011, Figure 3]. In the bottom two pictures, the intersection points of the red α-curve and blue β-curve are numbered in order mod p along the α-curve.

Since the *simple knots* in $L(p, q)$ are those that can be described via doubly pointed genus 1 Heeggaard diagrams of $L(p, q)$ with only p intersection points, we like to regard the Hedden knots as *almost-simple*. The bottom row of Figure 6 shows the same portion of the diagram for related simple knots. These knots are notable because of the following theorem.

Theorem 14 [Hedden 2011, Theorem 3.1]. *If K^* is a 1-bridge knot in $L(p, q)$ with an integral surgery to an L-space homology sphere, then either*

- *$p > 2g - 1$ and K^* is a simple knot, or*
- *$p = 2g - 1$ and K^* is either T_L or T_R,*

where g is the Seifert genus of the surgery dual knot.

Though the statement of this theorem in [Hedden 2011] is explicitly only for surgeries to S^3, Hedden notes that his proof applies for surgeries to any L-space homology sphere.

Here we would like to clarify a few points about these almost-simple knots, correct a couple of misstatements in [Baker 2014b], and discuss the homology spheres that may be obtained by integral surgery on them following the arguments of [Baker 2014b]. Afterwards, we summarize these results in Proposition 15.

The knots T_L and T_R are not homologous in general, contrary to what was stated in [Baker 2014b]. One can observe this directly from the descriptions in Figure 6 (or [Hedden 2011, Figure 3]), noting that the magnitude of their algebraic intersection numbers with the blue β-curve differ by 2. One may further observe that the mirror of the knot T_L in $L(p,q)$ is the knot T_R in $L(p,-q)$.

Furthermore, the proof in [Baker 2014b] that T_L differs from (and thus is homologous to) $K(p,q,q+1)$ by a crossing change also shows that T_R differs from $K(p,q,q-1)$ by a crossing change. Unfortunately, [Baker 2014b, Figure 2] should be mirrored through the Heegaard torus so that the red α-disks are below, the blue β-disks are above, and the twist in T_L is right-handed in order to be consistent with the orientation conventions. In particular, the crossing change from $K(p,q,q+1)$ to T_L is achieved by a -1-surgery on a loop C_L in the Heegaard that encircles the two base-points in the diagram for $K(p,q,q+1)$ of Figure 6 (bottom left). The crossing change from $K(p,q,q-1)$ to T_R may be similarly achieved by a $+1$-surgery on a loop C_R.

A knot homologous to $K(p,q,k)$ has an integral surgery to a homology sphere if and only if its self linking is $\pm 1/p \pmod 1$, or equivalently if and only if $k^2 = \pm q \pmod p$ [Rasmussen 2007, Lemma 2.6]; see [Fintushel and Stern 1980]. Here the choice of sign of $\pm$ is consistent and agrees with whether this integral surgery is a ± 1-surgery; see Figure 5 and the discussion in the proof of Theorem 8.

If T_L in $L(p,q)$ is dual to positive p-surgery on a knot K in a homology sphere Y, then so is $K(p,q,q+1)$ and hence $(q+1)^2 = -q \pmod p$. Making the substitution $-k = q+1$ gives the equation $k^2+k+1 = 0 \pmod p$. Thus $K(p,q,q+1)$ is dual to a type VII Berge knot B, one that lies in the fiber of a trefoil knot; see [Rasmussen 2007, §6.2; Greene 2013, §1.2]. Since its fiber positively frames B, this trefoil knot is the *negative* trefoil (not the positive trefoil as stated in [Baker 2014b]), and it can be identified with C_L under the S^3 surgery on $K(p,q,q+1)$. Performing -1-surgery on the negative trefoil C_L in S^3 produces $\mathcal{P}$ and takes $B \subset S^3$ to $K \subset \mathcal{P}$. Since the linking of B and C_L is 0, the positive p-surgery on B becomes the positive p-surgery on K. Consequently, if -1-surgery on T_L is a homology sphere, it is $\mathcal{P}$.

Similarly, if T_R in $L(p,q)$ is dual to positive p-surgery on a knot K' in a homology sphere, then $(q-1)^2 = -q \pmod p$. Making the substitution $k = q-1$ gives the equation $k^2-k-1 = 0 \pmod p$. Thus $K(p,q,q-1)$ is dual to a type VIII Berge knot B', one that lies in the fiber of the figure eight knot. This figure eight knot may be identified with C_R. Performing $+1$-surgery on C_R in S^3 produces the Brieskorn sphere $\Sigma(2,3,7)$ taking the knot B' to the knot K'. Consequently, if -1-surgery on T_R is a homology sphere, it is $\Sigma(2,3,7)$.

This coincides with the difference between the τ-invariants of these knots as noted by Rasmussen in the last two paragraphs of [Rasmussen 2007, §5]: $\tau(T_L, \mathfrak{s}_0) = -1$ and so $\tau(T_R, \mathfrak{s}_0) = +1$. Rasmussen further shows that if integral surgery on T_L or T_R produces a homology sphere, then it is an L-space homology sphere if and only if the surgery is a -1-surgery on T_L or a $+1$-surgery on T_R [Rasmussen 2007, Proposition 4.5].

In summary, the above discussion shows:

Proposition 15. (1) *In $L(p,q)$, T_L is homologous to the simple knot $K(p,q,q+1)$ and T_R is homologous to the simple knot $K(p,q,q-1)$.*

(2) *The mirror of $(L(p,q),T_L)$ is $(L(p,-q),T_R)$.*

(3) *If -1-surgery on T_L is a homology sphere, then it is $\mathcal{P} = \Sigma(2,3,5)$ and $K(p,q,q+1)$ is positive surgery dual to a type VII Berge knot.*

(4) *If -1-surgery on T_R is a homology sphere, then it is $\Sigma(2,3,7)$ and $K(p,q,q-1)$ is positive surgery dual to a type VIII Berge knot.*

Appendix: Symmetries of knots with lens space and Poincaré homology sphere surgeries

Kenneth L. Baker and Neil R. Hoffman

The aim of this appendix is to compute the symmetry group of a hyperbolic knot in a lens space with a surgery to the Poincaré homology sphere $\mathcal{P}$. Specifically, it will provide a proof of the following theorem:

Theorem A.1. *Let X be a one-cusped hyperbolic manifold admitting both a Poincaré homology sphere filling and a lens space filling. Then the symmetry group of X is trivial or generated by a single strong involution.*

Our method blends the classification of spherical orbifolds with results involving exceptional Dehn fillings. We refer the reader to [Boileau et al. 2012] for the most closely related paper on this topic and Thurston's notes [1979, Chapter 13] for a more general introduction to orbifolds. In preparation, we first extend the notation of Section 3 to address orbifolds and orbifold fillings and then discuss symmetries and quotients of hyperbolic manifolds. Thereafter, the argument in this appendix will gradually "whittle-down" the symmetry group of the exterior X of a hyperbolic knot K in a lens space with a Dehn surgery to $\mathcal{P}$. First, we give an argument which eliminates orientation reversing symmetries so that $\mathrm{Sym}(X) = \mathrm{Sym}^+(X)$. Then we proceed to analyze a subgroup $Z(X)$ of $\mathrm{Sym}^+(X)$, showing

(a) that it has index at most 2,

(b) if the index is 2 then $\mathrm{Sym}(X)$ contains a strong involution, and ultimately
(c) that $Z(X)$ is trivial.

This is pulled together in the proof of Theorem A.1 at the end of this appendix.

Background on orbifolds and symmetries. Let Q be a 3-dimensional manifold or orbifold with *singular set* $\Sigma(Q)$, the set of points fixed by some nontrivial element of $\pi_1^{orb}(Q)$. (So Q is a manifold if $\Sigma(Q) = \varnothing$.) If K is a knot in Q, then its complement is $Q - K$ while its exterior is $X = Q - \mathcal{N}(K)$. We further assume that K is either a component of $\Sigma(Q)$ or disjoint from $\Sigma(Q)$.

The torus $\partial\mathcal{N}(K)$ represents the *cusp* of $Q - K$ corresponding to K. In such a torus we consider essential closed curves up to free homotopy. A *primitive curve* is homotopic to an essential simple closed curve, and hence is a slope. Any nonprimitive essential curve is a multiple of a primitive curve. Given two closed curves α and β in the torus we say the *distance* between α and β, $\Delta(\alpha, \beta)$ is the minimal (unoriented) geometric intersection number between the two curves. If α is a curve in $\partial\mathcal{N}(K)$ that is an n-fold multiple of a slope γ, then the result of α-Dehn surgery on $K \subset Q$ is the orbifold $X(\alpha)$ with underlying manifold $|X(\alpha)|$ obtained as the γ-Dehn surgery on $K \subset |Q|$ and a new singular set $\Sigma(X(\alpha))$ consisting of $\Sigma(X)$ and the core of the attached $S^1 \times D^2$ with order n (unless $n = 1$ in which case the singular set remains only $\Sigma(X)$). The orbifold attached to X is a *orbi-solid torus* of order n. An orbi-solid torus of order 1 is just a solid torus.

Following [Boileau et al. 2012, §3], an *orbi-lens space* is the quotient of S^3 by a finite cyclic subgroup of $SO(4)$ and is consequentially a union of two orbi-solid tori with coprime orders along their common boundary. Its singular set is the union of the cores of these orbi-solid tori that have order at least 2.

A knot K in a closed 3-manifold M is *hyperbolic* if its complement $M - K$ is a hyperbolic 3-manifold with a single torus cusp. We also say the exterior $X = M - \mathcal{N}(K)$ of such a knot is *hyperbolic* if the interior of X is hyperbolic. For a hyperbolic knot K with exterior X, the symmetry group $\mathrm{Sym}(X)$ of homeomorphisms modulo isotopy is identified with the isometry group of the interior of X. The orientation preserving symmetry group is denoted $\mathrm{Sym}^+(X)$.

Within $\mathrm{Sym}^+(X)$ is the subgroup $Z(X)$, the maximal subgroup that acts freely on ∂X. In particular, $Z(X)$ is the maximal subgroup of $\mathrm{Sym}^+(X)$ for which the quotient orbifold (or manifold) $X/Z(X)$ has a torus cusp.

Since X is hyperbolic, the quotient $X/\mathrm{Sym}^+(X)$ is an orientable cusped hyperbolic 3-orbifold with $\mathrm{vol}(X/\mathrm{Sym}^+(X)) = \mathrm{vol}(X)/|\mathrm{Sym}^+(X)|$. Since the minimal volume of an orientable cusped hyperbolic 3-orbifold is positive (e.g., [Meyerhoff 1986]) and X has finite volume, it follows that $\mathrm{Sym}^+(X)$ is a finite group.

In a knot exterior X, the *homological longitude* is the slope λ in ∂X so that $[\lambda] = 0$ in $H_1(X; \mathbb{Q})$. This slope is unique up to sign as it is a generator of the kernel of the map $i_*\colon H_1(\partial X;\mathbb{Q}) \to H_1(X;\mathbb{Q})$ which has rank 1 by the half-lives, half-dies lemma. Since it is unique, elements of $\mathrm{Sym}(X)$ must preserve the slope λ up to sign. Elements of $\mathrm{Sym}^+(X)$ of order 2 that reverse the orientation of the homological longitude λ of the manifold X are called *strong involutions*.

If a group G acts on a hyperbolic knot exterior X, the cusp of the quotient orbifold X/G records important information about the action of G on the cusp of X. If the quotient X/G has a torus cusp, then G acts freely on ∂X. Furthermore, each element of G restricts to a translation on ∂X. If X/G has an $S^2(2,2,2,2)$ cusp, then G contains an order 2 element which fixes a discrete set of points on ∂X. In particular, if $X/\mathrm{Sym}^+(X)$ has an $S^2(2,2,2,2)$ boundary, then such an element of $\mathrm{Sym}^+(X)$ must preserve isotopy classes of slopes in ∂X but reverse their orientations. Hence $\mathrm{Sym}^+(X)$ contains a strong involution. Furthermore, this means that there is a degree 2 cover with a torus cusp, and hence $Z(X)$ has index 2 in $\mathrm{Sym}^+(X)$.

Whittling down the symmetry group.

Proposition A.2. *Let X be the exterior of a knot in a lens space $L(p,q)$ with $p \geq 2$ admitting a Dehn surgery to an integral homology sphere. Then either X is Seifert fibered or* $\mathrm{Sym}(X) = \mathrm{Sym}^+(X)$.

Proof. Assume X admits an orientation reversing symmetry τ, and let λ be the homological longitude of X. Then $\tau(\lambda)$ is isotopic to $\pm\lambda$ in ∂X. Since τ restricts to a homeomorphism on the boundary torus ∂X, there must be a slope δ with $\Delta(\delta,\lambda) = 1$ such that $\tau(\delta) = \pm\delta$. Because this restriction of τ to ∂X must also be orientation reversing, τ reverses the orientation on exactly one of δ and λ.

Let μ be the slope in ∂X such that $X(\mu) = L(p,q)$. Since X has a filling producing an integer homology sphere, X must be a homology solid torus. Thus, $\mu = p\delta + q'\lambda$ in $H_1(\partial X)$ for some integer q' coprime to p and $p \geq 2$. Since $\tau(\mu) = \pm(p\delta - q'\lambda)$ and $q' \neq 0$, we have that $\Delta(\mu,\tau(\mu)) = |2pq'| > 2$.

The symmetry τ implies $X(\mu) \cong X(\tau(\mu))$ so that $\pi_1(X(\mu))$ and $\pi_1(X(\tau(\mu)))$ are both cyclic. Thus, since $\Delta(\mu,\tau(\mu)) > 1$, X must be Seifert fibered by the cyclic surgery theorem [Culler et al. 1987]. □

We now further classify $Z(X)$.

Lemma A.3. *Let X be the exterior of a hyperbolic knot K in a lens space $L(p,q)$ with $p \geq 2$ admitting a Dehn surgery to an integral homology sphere. Then the following hold*:

(1) *$Z(X)$ is cyclic.*

(2) $[\mathrm{Sym}^+(X) : Z(X)] \leq 2$.

(3) *If* $[\mathrm{Sym}^+(X) : Z(X)] = 2$, *then X has a symmetry which is a strong involution.*

(4) *In the quotient orbifold $\mathcal{Z} = X/Z(X)$, the meridian μ of K maps to a primitive curve $\bar{\mu}$ in $\partial\mathcal{Z}$.*

(5) *The Dehn filling $\mathcal{Z}(\bar{\mu})$ is an orbi-lens space in which the core $\bar{K}$ of the filling is disjoint from the singular set of $\mathcal{Z}(\bar{\mu})$.*

Proof. We first utilize the fact that X is the exterior of a hyperbolic knot in an integral homology sphere to show that the boundary of $X/\operatorname{Sym}^+(X)$ is either a torus or $S^2(2, 2, 2, 2)$.

Since X admits an integer homology sphere filling, X is a homology solid torus. Furthermore, once oriented, the slope γ in ∂X of this filling generates $H_1(X)$. Thus together with a choice of orientation on the homological longitude λ it forms a basis (γ, λ) for $H_1(\partial X)$.

Now consider the action of $\operatorname{Sym}^+(X)$ on ∂X. Since the homological longitude is the unique slope in ∂X that is null homologous in $H_1(X; \mathbb{Q})$, any element of $\operatorname{Sym}^+(X)$ must send λ to $\pm\lambda$. Since $\operatorname{Sym}^+(X)$ is finite because X is hyperbolic, consideration of the action on $H_1(\partial X; \mathbb{Q})$ shows that any element of $\operatorname{Sym}^+(X)$ must send γ to $\pm\gamma$. Then, because $\operatorname{Sym}^+(X)$ is orientation preserving on ∂X too, any element of $\operatorname{Sym}^+(X)$ must send the basis (γ, λ) to $\pm(\gamma, \lambda)$. Hence all slopes in ∂X are preserved by the action of $\operatorname{Sym}^+(X)$. Consequently, the restriction to ∂X of a symmetry of X either is trivial, has no fixed point (and hence is a translation), or has order 2. Therefore, by consideration of the Euclidean orbifolds which are orientable quotients of the torus, the boundary of $X/\operatorname{Sym}^+(X)$ is either a torus or $S^2(2, 2, 2, 2)$.

We now further refine this analysis by incorporating the lens space filling. Since its exterior X is a homology solid torus, the knot K in $L(p, q)$ represents a generator of $H_1(L(p, q), \mathbb{Z})$. Hence it lifts to a knot $\widetilde{K} \subset S^3$ in the p-fold cyclic cover $S^3 \to L(p, q)$. This cover restricts to a p-fold cyclic cover of knot exteriors $\widetilde{X} \to X$ in which the meridian $\mu \subset \partial X$ of K lifts to p disjoint copies of the meridian of $\widetilde{K}$. (Here $\widetilde{X}$ is the exterior of $\widetilde{K}$.) The deck transformation group H of this cover is therefore isomorphic to the cyclic group $\mathbb{Z}/p\mathbb{Z}$ and extends to an action on S^3 that is free on $\partial\widetilde{X}$. Hence H is a subgroup of $Z(\widetilde{X})$.

Consider the composition $\alpha\colon \widetilde{X} \to X/\operatorname{Sym}^+(X)$ of the covers $\widetilde{X} \to X$ and $X \to X/\operatorname{Sym}^+(X)$ with deck transformation groups H and $\operatorname{Sym}^+(X)$ respectively. Let G be the deck transformation group of α. If α is regular, then G is a subgroup of $\operatorname{Sym}^+(\widetilde{X})$ and H must be a normal subgroup of G with $\operatorname{Sym}^+(X) = G/H$. As the boundary of $X/\operatorname{Sym}^+(X)$ is either a torus or $S^2(2, 2, 2, 2)$, we claim that α is indeed a regular cover. If this boundary is a torus, then this follows directly from [Reid 1991, Lemma 4]. However if this boundary is $S^2(2, 2, 2, 2)$, then the argument of [Reid 1991, Lemma 4] extends to also show that the covering $\alpha\colon \widetilde{X} \to X/\operatorname{Sym}^+(X)$ is a regular cover as noted in the proof of [Neumann and

Reid 1992, Proposition 9.1]. (Ultimately, this is because conjugation by an order 2 element of the peripheral subgroup $\pi_1(S^2(2,2,2,2))$ inverts elements of the maximal abelian subgroup. This maximal abelian subgroup contains the image of the peripheral subgroup of $\widetilde{X}$.)

Because $\widetilde{X}$ is the exterior of a hyperbolic knot in S^3, we know that $\mathrm{Sym}^+(\widetilde{X})$ must be cyclic or dihedral [Boileau et al. 2012, p. 627] (see also [Neumann and Reid 1992, Proposition 9.1]). More specifically $Z(\widetilde{X})$ is cyclic, $[\mathrm{Sym}^+(\widetilde{X}) : Z(\widetilde{X})] \le 2$, and $[\mathrm{Sym}^+(\widetilde{X}) : Z(\widetilde{X})] = 2$ only if $\mathrm{Sym}^+(\widetilde{X})$ contains a strong involution.

Since G is a subgroup of $\mathrm{Sym}^+(\widetilde{X})$, either it is a subgroup of $Z(\widetilde{X})$ and hence cyclic and acting freely on $\partial\widetilde{X}$ itself, or it contains a strong involution. In the former case, $\mathrm{Sym}^+(X) = G/H$ is cyclic as it is the quotient of a cyclic group G and $X/\mathrm{Sym}^+(X)$ has torus boundary so that $\mathrm{Sym}^+(X) = Z(X)$. In the latter case, $\mathrm{Sym}^+(X) = G/H$ must also contain a strong involution since H is a subgroup of $Z(\widetilde{X})$. Furthermore, because $H \cong \mathbb{Z}/p\mathbb{Z}$ is a nontrivial subgroup of $Z(\widetilde{X})$ (since $p \ge 2$), G cannot be generated by a single strong involution. Thus G must be a dihedral group (of order ≥ 4), and so $\mathrm{Sym}^+(X) = G/H$ is a dihedral group too (though possibly of order 2). Since $X/\mathrm{Sym}^+(X)$ has $S^2(2,2,2,2)$ boundary, $[\mathrm{Sym}^+(X) : Z(X)] = 2$ as noted in the introduction to this appendix. Since $Z(X)$ cannot contain a strong involution, is a cyclic group. Consequently, in both cases claims (1), (2), and (3) of the lemma hold.

Finally we examine the quotient orbifold $\mathcal{Z} = X/Z(X)$ and its filling $\mathcal{Z}(\bar{\mu})$. Since $\widetilde{X}$ is the exterior of the knot $\widetilde{K} \subset S^3$, $\mathrm{Sym}^+(\widetilde{X})$ naturally identifies with $\mathrm{Sym}^+(S^3, \widetilde{K})$. In particular, $Z(\widetilde{X})$ acts freely on $\widetilde{K}$. Therefore, if $\widetilde{\mu}$ is the lift of the slope μ to $\widetilde{X}$ and $\bar{\mu}$ is its image in the quotient orbifold $\mathcal{Z} = \widetilde{X}/Z(\widetilde{X}) = X/Z(X)$, then $\bar{\mu}$ must be a primitive slope and (4) holds.

Filling $\mathcal{Z}$ along the slope $\bar{\mu}$ produces

$$\mathcal{Z}(\bar{\mu}) = X(\mu)/Z(X),$$

which we recognize as an orbi-lens space since it is a cyclic quotient of $S^3 = \widetilde{X}(\widetilde{\mu})$ (see [Boileau et al. 2012, §3] for context). Because $\bar{\mu}$ is a primitive slope, the core of the solid torus in the filling $\mathcal{Z}(\bar{\mu})$ is a knot $\bar{K}$ that is disjoint from the singular set $\Sigma(\mathcal{Z}(\bar{\mu}))$ of the orbi-lens space. In particular, $\mathcal{Z}$ is the exterior of a knot in an orbi-lens space that is disjoint from the singular set. Thus (5) holds, which completes the proof. □

In light of the above lemma which employs a classification of cyclic quotients of S^3, it will also prove useful to have a list of manifolds and orbifolds that are cyclic quotients of $\mathcal{P}$. This is accomplished by Lemma A.4 below which essentially follows from combining the classification of elliptic manifolds with the corresponding classification of elliptic orbifolds in Dunbar [1988].

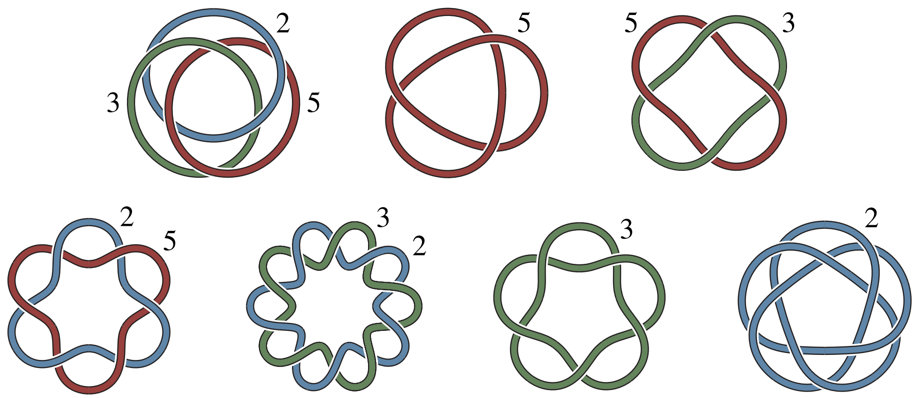

Figure 7. The possible singular sets of orbifolds cyclically covered by $\mathcal{P}$ with underlying space S^3. These are the relevant cases from a more general classification in [Dunbar 1988]. Following the notation of that paper, a link component labeled by n indicates a cone angle of $2\pi/n$ along that component in the singular set of the corresponding orbifold.

Lemma A.4. *If the finite cyclic quotient of* $\mathcal{P}$ *by symmetries is an orbifold* Q, *then* Q *is homeomorphic to one of the following*:

(1) *a manifold* M *with* $\pi_1(M) \cong \pi_1(\mathcal{P}) \times \mathbb{Z}/n\mathbb{Z}$ *for some integer* n *coprime to* 30,

(2) *an orbifold that fibers over* $S^2(2, 3, 5)$ *where the singular set of* Q *has* 1, 2, *or* 3 *components*, *or*

(3) *an orbifold with base space* S^3 *and singular set a knot or link as pictured in Figure 7*.

Moreover, *in the case that* Q *is an orbifold*, *the components of its singular set are fixed locally by finite cyclic groups of relatively prime orders.*

Proof. By the orbifold theorem (see [Boileau et al. 2003; Cooper et al. 2000, Chapter 7]), we know that Q is a spherical 3-orbifold.

Case 1 follows directly from the classification of elliptic manifolds (see for example [Thurston 1997, Theorem 4.4.14]), while Cases 2 and 3 follow from a careful reading of [Dunbar 1988]. Dunbar [1988] also classifies orbifolds covered by elliptic manifolds where the singular sets are trivalent graphs. However, no orbifold of this type can be cyclically covered by a manifold because the group of isometries that fixes a vertex in the trivalent graph is never cyclic. □

A simple case analysis shows that drilling the singular sets of orbifolds from the previous lemma results in several interesting manifolds.

Lemma A.5. *In Cases* 2 *and* 3 *of Lemma A.4*:

- *If* $|\Sigma(Q)| = 1$, *then the complement of* $\Sigma(Q)$ *is a Seifert fibered space over the disk with two exceptional fibers.*
- *If* $|\Sigma(Q)| = 2$, *then the complement of* $\Sigma(Q)$ *is a Seifert fibered space over the annulus with one exceptional fiber.*
- *If* $|\Sigma(Q)| = 3$, *then the complement of* $\Sigma(Q)$ *is* $F \times S^1$, *where* F *is a pair of pants.*

Lemma A.6. *Let* X *be the exterior of a hyperbolic knot* K *in a lens space* $L(p,q)$ *with* $p \geq 2$ *admitting a Dehn surgery to* $\mathcal{P}$. *Then every nontrivial subgroup of* $Z(X)$ *acts nonfreely on* X.

Proof. Assume $Z_f \subset Z(X)$ is a nontrivial subgroup of elements that act freely on X. Then the quotient X/Z_f is a hyperbolic manifold with a torus boundary. Since Z_f is nontrivial, $|Z_f| \geq 2$. Since $Z(X)$ is cyclic by Lemma A.3, the subgroup Z_f must be cyclic.

Assume μ is the meridian of the knot K in the lens space, and γ is the slope of is $\mathcal{P}$ surgery. Let $\bar{\mu}$ and $\bar{\gamma}$ be the images of μ and γ in the quotient X/Z_f. Then the orbifold fillings of X/Z_f are quotients of the Dehn fillings of X. In particular, $(X/Z_f)(\bar{\mu}) = X(\mu)/Z_f$ and $(X/Z_f)(\bar{\gamma}) = X(\gamma)/Z_f$ where the singular sets, if nonempty, are connected, consisting solely of the cores of the filling orbi-solid tori. Note that these quotients are manifolds exactly when their curves $\bar{\mu}$ and $\bar{\gamma}$ are primitive.

First we observe Lemma A.3 implies that $\bar{\mu}$ must be a primitive curve because X/Z_f is a quotient of X that covers the orbifold $\mathcal{Z} = X/Z(X)$. Thus we have two cases depending on whether $\bar{\gamma}$ is a primitive curve.

Case 1: $\bar{\gamma}$ *are primitive*. Observe that $\Delta(\bar{\mu}, \bar{\gamma}) = |Z_f|\Delta(\mu, \gamma)$, where $\Delta(\bar{\mu}, \bar{\gamma})$ is the distance in the boundary of the quotient and $\Delta(\mu, \gamma)$ is measured in ∂X.

Since μ and γ have primitive images in the quotient manifold X/Z_f, Z_f acts freely on the fillings $X(\mu)$ and $X(\gamma)$. Hence $(X/Z_f)(\bar{\mu}) = X(\mu)/Z_f$ is a manifold quotient of the lens space $X(\mu)$ and $(X/Z_f)(\bar{\gamma}) = X(\gamma)/Z_f$ is a manifold quotient of $\mathcal{P} = X(\gamma)$. Thus they are both covered by S^3, and therefore they are both finite manifolds (i.e., have finite π_1). Since $X(\gamma)/Z_f$ is covered by $\mathcal{P}$, $|Z_f|$ must be relatively prime to $|\pi_1(\mathcal{P})| = 120$ by Lemma A.4(1). Thus we must have $|Z_f| > 6$. Therefore $\Delta(\bar{\mu}, \bar{\gamma}) > 6$ which contradicts the bound on distance between the slopes of two finite surgeries on a hyperbolic knot provided by Boyer and Zhang [1996]

Case 2: $\bar{\gamma}$ *is not primitive*. In this case $(X/Z_f)(\bar{\gamma})$ is an orbifold that is a cyclic quotient of $\mathcal{P}$ whose singular set is the core of the filling of X/Z_f. However Lemma A.5 shows that the exterior of any such core would be the exterior of a torus knot in S^3 and hence Seifert fibered, contradicting that X/Z_f must be hyperbolic. □

The final lemma needed to establish the theorem can be proved in a similar manner to the lemma above.

Lemma A.7. *Let X be the exterior of a hyperbolic knot K in a lens space $L(p,q)$ with $p \geq 2$ admitting a Dehn surgery to $\mathcal{P}$. Then $|Z(X)| = 1$.*

Proof. Set $Z = Z(X)$ and assume $|Z| > 1$. Let μ be the meridian of K and let γ be the slope of the surgery to $\mathcal{P}$. Then let $\bar{\mu}$ and $\bar{\gamma}$ be the images of μ and γ in the quotient $\mathcal{Z} = X/Z$.

By Lemma A.6, $\mathcal{Z}$ is an orbifold (with nonempty singular set). Since $\mathcal{P}$ is a homology sphere, $\mathcal{Z}$ is the exterior of a knot $\bar{K}$ in an orbi-lens space $\mathcal{Z}(\bar{\mu})$ that is disjoint from the singular set by Lemma A.3. Because $\mathcal{Z}$ is a suborbifold of $\mathcal{Z}(\bar{\mu})$, the singular set $\Sigma(\mathcal{Z})$ may be identified with the singular set $\Sigma(\mathcal{Z}(\bar{\mu}))$. Thus $\Sigma(\mathcal{Z})$ has at most two components. Since $\mathcal{P} = X(\gamma)$, the quotient orbifold $\mathcal{P}/Z = \mathcal{Z}(\bar{\gamma})$ has a singular set $\Sigma(\mathcal{Z}(\bar{\gamma}))$ consisting of one, two, or three embedded circles according to whether the singular set $\Sigma(\mathcal{Z})$ has one or two components and whether or not the slope $\bar{\gamma}$ is a primitive curve.

Case 1: *$\bar{\gamma}$ is not primitive.* Then $\mathcal{Z}$ is the exterior of one of the components of $\Sigma(\mathcal{Z}(\bar{\gamma}))$. Hence by Lemmas A.4 and A.5, $\mathcal{Z}$ must be Seifert fibered. Thus X must be Seifert fibered contrary to K being a hyperbolic knot.

Case 2: *$\bar{\gamma}$ is primitive.* Let $\bar{K}_\gamma$ be the core of the filling solid torus in $\mathcal{Z}(\bar{\gamma})$ so that the exterior of $\bar{K}_\gamma$ is $\mathcal{Z}$. Since $\bar{\gamma}$ is primitive, the singular set $\Sigma = \Sigma(\mathcal{Z}(\bar{\gamma}))$ of $\mathcal{Z}(\bar{\gamma})$ is contained in the knot exterior $\mathcal{Z}$ and thus has either one or two components. Since $\mathcal{Z}$ is a hyperbolic orbifold in which Σ is a geodesic link of one or two components, the exterior of Σ is a hyperbolic manifold $X' = \mathcal{Z}_\Sigma = \mathcal{Z} - \mathcal{N}(\Sigma)$ [Kojima 1988; Sakai 1991].

Let us write $X'(\alpha, \beta)$ (or $X'(\alpha, \beta_1, \beta_2)$) to denote the Dehn filling of X' along the α and β (or α and β_1, β_2) slopes in the components of $\partial X'$ corresponding to $\partial\mathcal{N}(\bar{K}_\gamma) = \partial\mathcal{Z}$ and $\partial\mathcal{N}(\Sigma)$ respectively. We will use $-$ to indicate that the boundary component is left unfilled.

Case 2*a*: *Σ has just one component.* By Lemma A.4, $X'(\bar{\gamma}, -)$ is a Seifert fibered space over the disk with two exceptional fibers. Thus $X'(\bar{\gamma}, -)$ has a one-parameter family of slopes σ_i giving lens space fillings $X'(\bar{\gamma}, \sigma_i)$. Since $X'(\bar{\mu}, -)$ is a solid torus (as it is the complement of the core of an orbi-solid torus in an orbifold lens space with connected singular set), $X'(\bar{\mu}, \beta_i)$ is also a lens space for each β_i. Hence, by Thurston's hyperbolic dehn filling theorem [1979, Theorem 5.8.2], we can choose β_i so that $X'(-, \beta_i)$ is hyperbolic. However, $\Delta(\bar{\mu}, \bar{\gamma}) = |Z|\Delta(\mu, \gamma) \geq 2$, which contradicts the cyclic surgery theorem [Culler et al. 1987].

Case 2*b*: *Σ has two components.* This case is nearly identical to the previous.

By Lemma A.5, $X'(\bar{\gamma}, -, -)$ is a Seifert fibered space over the annulus with one exceptional fiber. Then there are infinitely many pairs of slopes $\{(\sigma_n, \sigma'_{n,m})\}_{(n,m)\in\mathbb{Z}^2}$ such that $X'(\bar{\gamma}, \sigma_n, \sigma'_{n,m})$ is a lens space. Since $X'(\bar{\mu}, -, -)$ is a thickened torus (being the complement of the two cores of the orbi-solid tori in the orbi-lens space), $X(\bar{\mu}, \sigma_n, \sigma'_{n,m})$ is a lens space for all these pairs of slopes.

Using Thurston's hyperbolic Dehn filling theorem twice, there are infinitely many of these pairs of slopes such that $X(-, \sigma_n, \sigma'_{n,m})$ is hyperbolic. Again, since $\Delta(\bar{\mu}, \bar{\gamma}) = |Z|\Delta(\mu, \gamma) \geq 2$, we obtain a contradiction to the cyclic surgery theorem [Culler et al. 1987] for these pairs. □

It may be tempting to try to use $|Z| \cdot \Delta(\mu, \gamma)$ in the previous lemma place try and improve the bound in [Boyer and Zhang 1996] that $\Delta(\mu, \gamma) \leq 2$. However, this bound is known to be sharp. Also, when $Z(M)$ is trivial, we lose the ability to "match up" the cores of the solid tori coming from the Heegaard splitting of the lens space with the exceptional fibers in $\mathcal{P}$, and so the arguments in this appendix will not apply.

Proof of Theorem A.1. By Proposition A.2,

$$\mathrm{Sym}(X) = \mathrm{Sym}^+(X).$$

From Lemma A.3(2), $Z(X)$ has index at most 2 in $\mathrm{Sym}(X)$. Then since $Z(X) = 1$ by Lemma A.7, we see that $\mathrm{Sym}(X)$ is either trivial or $\mathbb{Z}/2\mathbb{Z}$. According to Lemma A.3(3), the nontrivial element in this latter case must be a strong involution. □

Acknowledgments.

The main article benefited from Baker's discussions with Joshua Greene and Tye Lidman. Parts of this work and writing were done during his visits to the University of Texas and Boston College in the winter of 2015; he thanks them for their hospitality.

Baker was partially supported by a grant from the Simons Foundation (#209184). Hoffman was supported by the Australian Research Council Discovery Grant DP130103694 and by a grant from the Simons Foundation (#524123). The appendix benefited from the thoughtful suggestions and corrections of the referee.

References

[Baker 2014a] K. L. Baker, "Counting genus one fibered knots in lens spaces", *Michigan Math. J.* **63**:3 (2014), 553–569. MR Zbl

[Baker 2014b] K. L. Baker, "The Poincaré homology sphere and almost-simple knots in lens spaces", *Proc. Amer. Math. Soc.* **142**:3 (2014), 1071–1074. MR Zbl

[Baker et al. 2008] K. L. Baker, J. E. Grigsby, and M. Hedden, "Grid diagrams for lens spaces and combinatorial knot Floer homology", *Int. Math. Res. Not.* **2008**:10 (2008), art. id. rnm024, 39 pp. MR Zbl

[Baker et al. 2016] K. L. Baker, D. Buck, and A. G. Lecuona, "Some knots in $S^1 \times S^2$ with lens space surgeries", *Comm. Anal. Geom.* **24**:3 (2016), 431–470. MR Zbl

[Baker et al. 2019] K. L. Baker, N. R. Hoffman, and J. E. Licata, "Jointly primitive knots and surgeries between lens spaces", preprint, 2019. arXiv

[Berge 1990] J. Berge, "Some knots with surgeries yielding lens spaces", preprint, 1990. arXiv

[Birman and Menasco 1993] J. S. Birman and W. W. Menasco, "Studying links via closed braids, III: Classifying links which are closed 3-braids", *Pacific J. Math.* **161**:1 (1993), 25–113. MR Zbl

[Boileau and Otal 1991] M. Boileau and J.-P. Otal, "Scindements de Heegaard et groupe des homéotopies des petites variétés de Seifert", *Invent. Math.* **106**:1 (1991), 85–107. MR Zbl

[Boileau et al. 2003] M. Boileau, S. Maillot, and J. Porti, *Three-dimensional orbifolds and their geometric structures*, Panoramas et Synthèses **15**, Société Mathématique de France, Paris, 2003. MR Zbl

[Boileau et al. 2012] M. Boileau, S. Boyer, R. Cebanu, and G. S. Walsh, "Knot commensurability and the Berge conjecture", *Geom. Topol.* **16**:2 (2012), 625–664. MR Zbl

[Boyer and Zhang 1996] S. Boyer and X. Zhang, "Finite Dehn surgery on knots", *J. Amer. Math. Soc.* **9**:4 (1996), 1005–1050. MR Zbl

[Brody 1960] E. J. Brody, "The topological classification of the lens spaces", *Ann. of Math.* (2) **71** (1960), 163–184. MR Zbl

[Cooper et al. 2000] D. Cooper, C. D. Hodgson, and S. P. Kerckhoff, *Three-dimensional orbifolds and cone-manifolds*, MSJ Memoirs **5**, Math. Soc. Japan, Tokyo, 2000. MR Zbl

[Culler et al. 1987] M. Culler, C. M. Gordon, J. Luecke, and P. B. Shalen, "Dehn surgery on knots", *Ann. of Math.* (2) **125**:2 (1987), 237–300. MR Zbl

[Dunbar 1988] W. D. Dunbar, "Geometric orbifolds", *Rev. Mat. Univ. Complut. Madrid* **1**:1-3 (1988), 67–99. MR Zbl

[Eftekhary 2009] E. Eftekhary, "Seifert fibered homology spheres with trivial Heegaard Floer homology", preprint, 2009. arXiv

[Eftekhary 2011] E. Eftekhary, "Knots which admit a surgery with simple knot Floer homology groups", *Algebr. Geom. Topol.* **11**:3 (2011), 1243–1256. MR Zbl

[Eftekhary 2018] E. Eftekhary, "Bordered Floer homology and existence of incompressible tori in homology spheres", *Compos. Math.* **154**:6 (2018), 1222–1268. MR Zbl

[Eudave-Muñoz et al. 2015] M. Eudave-Muñoz, E. Jasso, K. Miyazaki, and K. Motegi, "Seifert fibered surgeries on strongly invertible knots without primitive/Seifert positions", *Topology Appl.* **196**:B (2015), 729–753. MR Zbl

[Fintushel and Stern 1980] R. Fintushel and R. J. Stern, "Constructing lens spaces by surgery on knots", *Math. Z.* **175**:1 (1980), 33–51. MR Zbl

[Giroux 2002] E. Giroux, "Géométrie de contact: de la dimension trois vers les dimensions supérieures", pp. 405–414 in *Proceedings of the International Congress of Mathematicians* (Beijing, 2002), vol. II, edited by T. Li, Higher Ed. Press, Beijing, 2002. MR Zbl

[Goda and Teragaito 2000] H. Goda and M. Teragaito, "Dehn surgeries on knots which yield lens spaces and genera of knots", *Math. Proc. Cambridge Philos. Soc.* **129**:3 (2000), 501–515. MR Zbl

[Gordon 1991] C. M. Gordon, "Dehn surgery on knots", pp. 631–642 in *Proceedings of the International Congress of Mathematicians* (Kyoto, 1990), vol. I, Math. Soc. Japan, Tokyo, 1991. MR Zbl

[Greene 2013] J. E. Greene, "The lens space realization problem", *Ann. of Math.* (2) **177**:2 (2013), 449–511. MR Zbl

[Hedden 2009] M. Hedden, "On knot Floer homology and cabling, II", *Int. Math. Res. Not.* **2009**:12 (2009), 2248–2274. MR Zbl

[Hedden 2011] M. Hedden, "On Floer homology and the Berge conjecture on knots admitting lens space surgeries", *Trans. Amer. Math. Soc.* **363**:2 (2011), 949–968. MR Zbl

[Hodgson and Rubinstein 1985] C. Hodgson and J. H. Rubinstein, "Involutions and isotopies of lens spaces", pp. 60–96 in *Knot theory and manifolds* (Vancouver, 1983), edited by D. Rolfsen, Lecture Notes in Math. **1144**, Springer, 1985. MR Zbl

[Kobayashi 1984] T. Kobayashi, "Structures of the Haken manifolds with Heegaard splittings of genus two", *Osaka J. Math.* **21**:2 (1984), 437–455. MR Zbl

[Kojima 1988] S. Kojima, "Isometry transformations of hyperbolic 3-manifolds", *Topology Appl.* **29**:3 (1988), 297–307. MR Zbl

[Kronheimer et al. 2007] P. Kronheimer, T. Mrowka, P. Ozsváth, and Z. Szabó, "Monopoles and lens space surgeries", *Ann. of Math.* (2) **165**:2 (2007), 457–546. MR Zbl

[Meyerhoff 1986] R. Meyerhoff, "Sphere-packing and volume in hyperbolic 3-space", *Comment. Math. Helv.* **61**:2 (1986), 271–278. MR Zbl

[Montesinos 1975] J. M. Montesinos, "Surgery on links and double branched covers of S^3", pp. 227–259 in *Knots, groups, and* 3*-manifolds*, edited by L. P. Neuwirth, Ann. of Math. Studies **84**, 1975. MR Zbl

[Moser 1971] L. Moser, "Elementary surgery along a torus knot", *Pacific J. Math.* **38** (1971), 737–745. MR Zbl

[Neumann and Reid 1992] W. D. Neumann and A. W. Reid, "Arithmetic of hyperbolic manifolds", pp. 273–310 in *Topology '90* (Columbus, OH, 1990), edited by B. Apanasov et al., Ohio State Univ. Math. Res. Inst. Publ. **1**, de Gruyter, Berlin, 1992. MR Zbl

[Ni 2007] Y. Ni, "Knot Floer homology detects fibred knots", *Invent. Math.* **170**:3 (2007), 577–608. MR Zbl

[Ni 2009] Y. Ni, "Link Floer homology detects the Thurston norm", *Geom. Topol.* **13**:5 (2009), 2991–3019. MR Zbl

[Ozsváth and Szabó 2005] P. Ozsváth and Z. Szabó, "On knot Floer homology and lens space surgeries", *Topology* **44**:6 (2005), 1281–1300. MR Zbl

[Ozsváth and Szabó 2011] P. S. Ozsváth and Z. Szabó, "Knot Floer homology and rational surgeries", *Algebr. Geom. Topol.* **11**:1 (2011), 1–68. MR Zbl

[Rasmussen 2007] J. Rasmussen, "Lens space surgeries and L-space homology spheres", preprint, 2007. arXiv

[Reid 1991] A. W. Reid, "Arithmeticity of knot complements", *J. London Math. Soc.* (2) **43**:1 (1991), 171–184. MR Zbl

[Sakai 1991] T. Sakai, "Geodesic knots in a hyperbolic 3-manifold", *Kobe J. Math.* **8**:1 (1991), 81–87. MR Zbl

[Tange 2009a] M. Tange, "Lens spaces given from L-space homology 3-spheres", *Experiment. Math.* **18**:3 (2009), 285–301. MR Zbl

[Tange 2009b] M. Tange, "Ozsváth Szabó's correction term and lens surgery", *Math. Proc. Cambridge Philos. Soc.* **146**:1 (2009), 119–134. MR Zbl

[Tange 2011] M. Tange, "On the non-existence of L-space surgery structure", *Osaka J. Math.* **48**:2 (2011), 541–547. MR Zbl

[Thurston 1979] W. P. Thurston, "The geometry and topology of three-manifolds", lecture notes, Princeton University, 1979, Available at http://library.msri.org/books/gt3m/.

[Thurston 1997] W. P. Thurston, *Three-dimensional geometry and topology*, vol. 1, edited by S. Levy, Princeton Mathematical Series **35**, Princeton University Press, 1997. MR Zbl

[Wang and Zhou 1992] S. C. Wang and Q. Zhou, "Symmetry of knots and cyclic surgery", *Trans. Amer. Math. Soc.* **330**:2 (1992), 665–676. MR Zbl

Received May 16, 2017. Revised June 27, 2019.

KENNETH L. BAKER
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MIAMI
CORAL GABLES, FL
UNITED STATES

k.baker@math.miami.edu

NEIL R. HOFFMAN
DEPARTMENT OF MATHEMATICS
OKLAHOMA STATE UNIVERSITY
STILLWATER, OK
UNITED STATES

neil.r.hoffman@okstate.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.29

FUSION SYSTEMS OF BLOCKS OF FINITE GROUPS OVER ARBITRARY FIELDS

ROBERT BOLTJE, ÇİSİL KARAGÜZEL AND DENİZ YILMAZ

To any block idempotent b of a group algebra kG of a finite group G over a field k of characteristic $p > 0$, Puig associated a fusion system and proved that it is saturated if the k-algebra $kC_G(P)e$ is split, where (P, e) is a maximal kGb-Brauer pair. We investigate in the nonsplit case how far the fusion system is from being saturated by describing it in an explicit way as being generated by the fusion system of a related block idempotent over a larger field together with a single automorphism of the defect group.

1. Introduction

Let k be a field of characteristic p, let G be a finite group and let b be a block idempotent of kG. Puig defined a fusion system $\mathcal{F}_{(P,e)}(kGb)$ associated to kGb after choosing a maximal kGb-Brauer pair (P, e). Up to category isomorphism, this fusion system does not depend on the choice of (P, e). Puig also proved that $\mathcal{F}_{(P,e)}(kGb)$ is saturated if the k-algebra $kC_G(P)e$ is split. It is known that in the nonsplit case it can happen that the fusion system associated to kGb is not saturated. In fact, the Sylow axiom can fail, while the extension axiom always holds. In the Main Theorem of this paper (Theorem 5.2) we establish a precise connection between the fusion systems of related blocks in a Galois extension L/K of fields of characteristic p with Galois group Γ. More precisely, let b be a block idempotent of LG and $\tilde{b}$ the unique block idempotent of KG with $b\tilde{b} = b$. Moreover, let (P, e) be a maximal LGb-Brauer pair and let $\tilde{e}$ be the unique block idempotent of $KC_G(P)$ with $e\tilde{e} = e$. Then $(P, \tilde{e})$ is a maximal $KG\tilde{b}$-Brauer pair and one has an inclusion of the fusion systems

$$\mathcal{F} := \mathcal{F}_{(P,e)}(LGb) \subseteq \mathcal{F}_{(P,\tilde{e})}(KG\tilde{b}) =: \widetilde{\mathcal{F}}.$$

Theorem 5.2 states that there exists an element $\sigma \in \mathrm{Aut}_{\widetilde{\mathcal{F}}}(P)$ such that $\widetilde{\mathcal{F}} = \langle \mathcal{F}, \sigma \rangle$. As consequences of the nature of σ we obtain that $\widetilde{\mathcal{F}}$ is saturated if and only if $\mathcal{F}$ is saturated and p does not divide the index $[\Gamma_b : \Gamma_e] = [K(e) : K(b)]$ of the stabilizers of b and e under the Galois action, or equivalently the degree of the field extensions after adjoining the coefficients of e and b to K. In the case that L is chosen such

MSC2010: 20C20.
Keywords: blocks of finite groups, fusion systems.

that $LC_G(P)e$ is split, this gives a particularly handy criterion for a fusion system of a block $KG\tilde{b}$ in the nonsplit case to be saturated; see Theorem 6.3. The main result allows an alternative easy proof for the known fact that the extension axiom holds also in the nonsplit case; see Theorem 6.2. Finally, the Main Theorem implies that a weak form of Alperin's fusion theorem holds also for arbitrary block fusion systems; see Theorem 6.5.

Notation 1.1. We will use the following standard notations without further notice:

For a group G and $x \in G$, we denote by $c_x : G \to G$ the conjugation map $g \mapsto xgx^{-1}$. If k is a commutative ring, its k-linear extension to the group algebra kG is again denoted by $c_x : kG \to kG$. We frequently will use left-exponential notation ${}^x(-) := c_x$ for these maps. The maps c_x, $x \in G$, define an action of G on kG via k-algebra homomorphisms.

For $H \leqslant G$, we denote by $[G/H]$ a set of representatives of the cosets G/H.

If a group G acts on a set X, we usually denote the stabilizer of an element $x \in X$ by G_x. Moreover, for $H \leqslant G$, we denote by X^H the set of H-fixed points of X.

2. Brauer pairs

Throughout this section, G denotes a finite group, k denotes a field of characteristic $p > 0$, and b denotes a block idempotent of kG, i.e., a primitive idempotent of $Z(kG)$. We recall the definition and properties of Brauer pairs for kG following the treatment in [Aschbacher et al. 2011, IV.2]. We note that the blanket assumption in [Aschbacher et al. 2011, IV.2] that k is algebraically closed is not used in the proofs of any of the statements that we cite from there. Alternatively, see also [Linckelmann 2018, Sections 5.9 and 6.3].

Recall that, for a p-subgroup P of G, the *Brauer homomorphism* with respect to P is the k-linear projection map $\mathrm{Br}_P : (kG)^P \to kC_G(P)$, $\sum_{g\in G} \alpha_g g \mapsto \sum_{g\in C_G(P)} \alpha_g g$. This is a surjective k-algebra homomorphism which respects G-conjugation: $c_x \circ \mathrm{Br}_P = \mathrm{Br}_{{}^xP} \circ c_x : (kG)^P \to kC_G({}^xP)$ for $x \in G$. Thus, $\mathrm{Br}_P(b)$ is an idempotent of $Z(kC_G(P)) = (kC_G(P))^{C_G(P)}$. Recall further that a *kG-Brauer pair* is a pair (P, e) consisting of a p-subgroup P of G and a block idempotent e of $kC_G(P)$. If e occurs in the unique decomposition of $\mathrm{Br}_P(b)$ into a sum of primitive idempotents of $Z(kC_G(P))$ (that is, if $\mathrm{Br}_P(b)e = e$), then we call (P, e) a *(kG, b)-Brauer pair*. We denote by $\mathcal{BP}(kG)$ the set of kG-Brauer pairs and by $\mathcal{BP}(kG, b)$ the set of (kG, b)-Brauer pairs. Clearly, $\mathcal{BP}(kG)$ is the disjoint union of the subsets $\mathcal{BP}(kG, b)$, where b runs through the block idempotents of kG. The set $\mathcal{BP}(kG)$ is a G-set under the conjugation action given by ${}^x(P, e) := ({}^xP, {}^xe)$, and the subset $\mathcal{BP}(kG, b)$ is G-stable. Finally, we say that an idempotent i of $(kG)^P$ is *associated* to a kG-Brauer pair (P, e) if

$$e\mathrm{Br}_P(i) = \mathrm{Br}_P(i) \neq 0.$$

Note that if i is primitive in $(kG)^P$ then $e\mathrm{Br}_P(i) \neq 0$ implies that $\mathrm{Br}_P(i) \neq 0$ and that $\mathrm{Br}_P(i)$ is primitive in $kC_G(P)$. One writes $(Q, f) \leqslant (P, e)$ if $Q \leqslant P$ and if any primitive idempotent i of $(kG)^P$ which is associated to (P, e) is also associated to (Q, f); see [Aschbacher et al. 2011, Definition 2.9]. This relation has the following properties.

Theorem 2.1 [Aschbacher et al. 2011, Theorems IV.2.10 and IV.2.16]. (a) *Let $(P, e) \in \mathcal{BP}(kG)$ and let $Q \leqslant P$. Then there exists a unique block idempotent f of $kC_G(Q)$ such that $(Q, f) \leqslant (P, e)$.*

(b) *Let $(Q, f) \leqslant (P, e)$ be in $\mathcal{BP}(kG)$ with $Q \trianglelefteq P$. Then f is the unique block idempotent of $kC_G(Q)$ which is P-stable and satisfies $Br_P(f)e = e$.*

(c) *The relation $\leqslant$ on $\mathcal{BP}(kG)$ is a partial order which is respected by the conjugation action of G.*

Clearly $(\{1\}, b) \in \mathcal{BP}(kG, b)$ and part (b) of the above theorem implies that if $(P, e) \in \mathcal{BP}(kG, b)$ then $(\{1\}, b) \leqslant (P, e)$. Parts (a) and (c) further imply that if $(Q, f) \leqslant (P, e)$ holds for elements in $\mathcal{BP}(kG)$ then $(Q, f) \in \mathcal{BP}(kG, b)$ if and only if $(P, e) \in \mathcal{BP}(kG, b)$.

For Brauer pairs $(Q, f), (P, e) \in \mathcal{BP}(kG)$ one writes $(Q, f) \trianglelefteq (P, e)$ if $Q \trianglelefteq P$, f is P-stable and $\mathrm{Br}_P(f)e = e$, cf. [Aschbacher et al. 2011, Definition IV.2.13]. The following result is well-known to specialists. We state it for convenient future reference and give a proof for the convenience of the reader.

Theorem 2.2. *For $(Q, f), (P, e) \in \mathcal{BP}(kG)$ with $Q \leqslant P$ the following statements are equivalent*:

(i) *One has $(Q, f) \leqslant (P, e)$.*

(ii) *There exist primitive idempotents i of $(kG)^P$ and j of $(kG)^Q$ such that $ij = j = ji$, $\mathrm{Br}_P(i)e \neq 0$ and $\mathrm{Br}_Q(j)f \neq 0$.*

(iii) *There exist Brauer pairs $(Q_i, d_i) \in \mathcal{BP}(kG)$, $i = 0, \dots, n$, such that*

$$(Q, f) = (Q_0, d_0) \trianglelefteq (Q_1, d_1) \trianglelefteq \cdots \trianglelefteq (Q_n, d_n) = (P, e).$$

(iv) *For all primitive idempotents i of $(kG)^P$ with $\mathrm{Br}_P(i)e \neq 0$ one has $\mathrm{Br}_Q(i)f \neq 0$.*

(v) *There exists a primitive idempotent i of $(kG)^P$ such that $\mathrm{Br}_P(i)e \neq 0$ and $\mathrm{Br}_Q(i)f = \mathrm{Br}_Q(i) \neq 0$.*

(vi) *There exists a primitive idempotent i of $(kG)^P$ such that $\mathrm{Br}_P(i)e \neq 0$ and $\mathrm{Br}_Q(i)f \neq 0$.*

Proof. The equivalences (i)$\Longleftrightarrow$(ii)$\Longleftrightarrow$(iii) follow from [Aschbacher et al. 2011, Proposition IV.2.14]. Further, the implications (i)$\Rightarrow$(iv) and (v)$\Rightarrow$(vi) are trivial and the implication (i)$\Rightarrow$(v) follows from the fact that the image of a primitive idempotent under a surjective k-algebra homomorphism is either 0 or a primitive idempotent.

Next we show that (iv) implies (i). Let i be a primitive idempotent of $(kG)^P$ such that $\mathrm{Br}_P(i)e = \mathrm{Br}_P(i) \neq 0$. By (iv), $\mathrm{Br}_Q(i)f \neq 0$. By Theorem 2.1(a) there exists a block idempotent f' of $kC_G(Q)$ such that $(Q, f') \leqslant (P, e)$. Thus, $\mathrm{Br}_Q(i)f' = \mathrm{Br}_Q(i)$ which implies that $0 \neq \mathrm{Br}_Q(i)f = \mathrm{Br}_Q(i)f'f$ and further that $f = f'$ and thus $(Q, f) \leqslant (P, e)$.

Finally, we show that (vi) implies (i). Let i be as in (vi). By Theorem 2.1(a) there exists a block idempotent f' of $kC_G(Q)$ such that $(Q, f') \leqslant (P, e)$. This implies $\mathrm{Br}_Q(i)f' = \mathrm{Br}_Q(i) \neq 0$ and $0 \neq \mathrm{Br}_Q(i)f = \mathrm{Br}_Q(i)f'f$. Thus $f = f'$ and $(Q, f) \leqslant (P, e)$. □

Recall that if $I \leqslant H \leqslant G$ then we have a well-defined trace map

$$\mathrm{Tr}_I^H : (kG)^I \to (kG)^H, \quad a \mapsto \sum_{x \in [H/I]} {}^x a.$$

A subgroup P of G, minimal with the property that $b \in \mathrm{Tr}_P^G((kG)^P)$, is called a *defect group* of the block idempotent b and of the block algebra kGb. The defect groups of kGb form a single G-conjugacy class of p-subgroups of G. Maximal elements in $\mathcal{BP}(kG, b)$ enjoy properties that resemble the Sylow theorem for finite groups.

Theorem 2.3 [Aschbacher et al. 2011, Theorem IV.2.20]. (a) *The maximal elements in $\mathcal{BP}(kG, b)$ with respect to $\leqslant$ form a single G-orbit.*

(b) *For $(P, e) \in \mathcal{BP}(kG, b)$ the following are equivalent*:

(i) *(P, e) is a maximal element in $\mathcal{BP}(kG, b)$.*

(ii) *P is a defect group of kGb.*

(iii) *P is maximal among all p-subgroups of G with the property* $\mathrm{Br}_P(b) \neq 0$.

3. Fusion systems of block algebras

Throughout this section, p is a prime. We first recall the basic notions and properties of fusion systems, a structure introduced by Puig. Our terminology follows [Aschbacher et al. 2011, Chapter I].

For subgroups Q and R of a finite group G we denote by $\mathrm{Hom}_G(Q, R)$ the set of all group homomorphisms $\varphi : Q \to R$ with the property that there exists $x \in G$ with $\varphi(u) = c_x(u)$ for all $u \in Q$. Moreover, we set $\mathrm{Aut}_G(Q) := \mathrm{Hom}_G(Q, Q)$.

Definition 3.1 [Aschbacher et al. 2011, Definition I.2.1]. Let P be a finite p-group. A subcategory $\mathcal{F}$ of the category of finite groups whose objects are the subgroups of P is called a *fusion system* over P if for any two subgroups Q and R of P, the set $\mathrm{Hom}_{\mathcal{F}}(Q, R)$ has the following properties:

(i) $\mathrm{Hom}_P(Q, R) \subseteq \mathrm{Hom}_{\mathcal{F}}(Q, R)$ and every element of $\mathrm{Hom}_{\mathcal{F}}(Q, R)$ is injective.

(ii) For each $\varphi \in \mathrm{Hom}_{\mathcal{F}}(Q, R)$, the group isomorphism $Q \to \varphi(Q)$, $u \mapsto \varphi(u)$, and its inverse are also morphisms in $\mathcal{F}$.

For instance, if G is a finite group and P is a p-subgroup of G, we obtain a fusion system $\mathcal{F}_P(G)$ over P by setting $\mathrm{Hom}_{\mathcal{F}_P(G)}(Q, R) := \mathrm{Hom}_G(Q, R)$, for all subgroups Q and R of P. Note that the intersection of two fusion systems over P is again a fusion system and that a fusion system over P is determined by the isomorphisms it contains. Thus the smallest fusion system over a finite p-group P is the fusion system $\mathcal{F}_P(P)$.

Definition 3.2 [Aschbacher et al. 2011, Definition I.2.4]. Let $\mathcal{F}$ be a fusion system over a finite p-group P. A subgroup Q of P is called *fully $\mathcal{F}$-centralized* if $|C_P(Q)| \geqslant |C_P(Q')|$ for any subgroup Q' of P which is $\mathcal{F}$-isomorphic to Q. Similarly, Q is called *fully $\mathcal{F}$-normalized* if $|N_P(Q)| \geqslant |N_P(Q')|$ for any subgroup Q' of P which is $\mathcal{F}$-isomorphic to Q.

Definition 3.3 [Aschbacher et al. 2011, Definition I.2.2]. Let $\mathcal{F}$ be a fusion system on a finite p-group P and let $\varphi : Q \to R$ be an isomorphism in $\mathcal{F}$. One denotes by N_φ the set of all elements $y \in N_P(Q)$ for which there exists $z \in N_P(R)$ with the property

$$\varphi \circ c_y = c_z \circ \varphi : Q \to R.$$

Note that $QC_P(Q) \leqslant N_\varphi \leqslant N_P(Q)$ and that N_φ does not depend on $\mathcal{F}$, but only on φ and P.

If $\mathcal{F}$ is a fusion system over a finite p-group P and $Q \leqslant P$ then we set $\mathrm{Aut}_{\mathcal{F}}(Q) := \mathrm{Hom}_{\mathcal{F}}(Q, Q)$, a subgroup of the automorphism group of Q. The following definition of saturation goes back to Stancu and is an equivalent reformulation of the original definition; see [Aschbacher et al. 2011, Proposition I.9.3].

Definition 3.4. A fusion system $\mathcal{F}$ over a p-group P is called *saturated* if the following two conditions hold.

(i) *Sylow axiom:* The group $\mathrm{Aut}_P(P)$ is a Sylow p-subgroup of $\mathrm{Aut}_{\mathcal{F}}(P)$.

(ii) *Extension axiom:* For every $Q \leqslant P$ and every $\varphi \in \mathrm{Hom}_{\mathcal{F}}(Q, P)$ such that $\varphi(Q)$ is fully $\mathcal{F}$-normalized there exists a morphism $\psi \in \mathrm{Hom}_{\mathcal{F}}(N_\varphi, P)$ whose restriction to Q equals φ.

For instance, if P is a Sylow p-subgroup of a finite group G then the fusion system $\mathcal{F}_P(G)$ is saturated; see [Aschbacher et al. 2011, Theorem I.2.3].

Definition 3.5 [Aschbacher et al. 2011, Definitions IV.3.15 and IV.2.21]. Let G be a finite group, let k be a field of characteristic p, let b be a block idempotent of kG, and let (P, e) be a maximal (kG, b)-Brauer pair. We define a category $\mathcal{F}_{(P,e)}(kGb)$ as follows. First, for every $Q \leqslant P$ denote by e_Q the unique block idempotent of $kC_G(Q)$ with $(Q, e_Q) \leqslant (P, e)$. The objects of $\mathcal{F}_{(P,e)}(kGb)$ are the subgroups of P and for subgroups Q and R of P let $\mathrm{Hom}_{\mathcal{F}_{(P,e)}(kGb)}(Q, R)$ denote the set of group homomorphisms $\varphi : Q \to R$ such that there exists $x \in G$ with $\varphi(u) = c_x(u)$

for all $u \in Q$ and ${}^x(Q, e_Q) \leqslant (R, e_R)$. Composition in $\mathcal{F}_{(P,e)}(kGb)$ is the usual composition of functions.

Remark 3.6. Let kG, b, and (P, e) be as in Definition 3.5.

(a) It is clear from the definition that $\mathcal{F}_{(P,e)}(kGb)$ is a fusion system over P.

(b) If kGb is the principal block of kG, then by Brauer's third main theorem, $\mathcal{F}_{(P,e)}(kGb)$ is equal to $\mathcal{F}_P(G)$ and P is a Sylow p-subgroup of G. Thus, $\mathcal{F}_{(P,e)}(kGb)$ is saturated in this case.

(c) Example 3.8 below shows that in general the Sylow axiom does not hold for $\mathcal{F}_{(P,e)}(kGb)$. But we will show in Theorem 6.2 that the extension axiom holds for $\mathcal{F}_{(P,e)}(kGb)$.

The following theorem was first proved by Puig. It follows from Theorem IV.3.2 and Proposition IV.3.14 in [Aschbacher et al. 2011]. See also [Linckelmann 2018, Theorem 8.5.2] and note that there the terminology is different: Fusion systems in [Linckelmann 2018] are defined to be saturated fusion systems in our terminology.

Theorem 3.7. *Let kG, b, and (P, e) be as in Definition 3.5 and suppose that the k-algebra $kC_G(P)e$ is split, i.e., for every simple $kC_G(P)e$-module V one has a k-algebra isomorphism $\mathrm{End}_{kC_G(P)e}(V) \cong k$. Then the fusion system $\mathcal{F}_{(P,e)}(kGb)$ is saturated.*

We are grateful to Radha Kessar who suggested the following example to us.

Example 3.8. Let $p = 2$, $k = \mathbb{F}_2$, the field with 2 elements, and $G := D_{24} = (C_3 \times C_4) \rtimes C_2$, the dihedral group with 24 elements, with C_2 acting by inversion on $C_3 \times C_4$. Let g denote a generator of C_3. Then $b := g + g^2$ is a block idempotent of $\mathbb{F}_2G$ and $(P, e) := (C_4, b)$ is a maximal $(\mathbb{F}_2G, b)$-Brauer pair. We have $\mathrm{Aut}_P(P) = \{1\}$, since P is abelian and an easy computation shows that $\mathrm{Aut}_{\mathcal{F}_{(P,e)}(\mathbb{F}_2Gb)}(P) \cong C_2$. Thus, the Sylow axiom does not hold for $\mathcal{F}_{(P,e)}(\mathbb{F}_2Gb)$ and therefore the fusion system $\mathcal{F}_{(P,e)}(\mathbb{F}_2Gb)$ is not saturated.

4. Extension of scalars

Throughout this section L/K denotes a finite Galois extension of fields of characteristic $p > 0$ and Γ denotes its Galois group. Moreover, G denotes a finite group.

Γ acts via K-algebra automorphisms on the group algebra LG and also on $Z(LG)$ by applying $\gamma \in \Gamma$ to the coefficients of an element in LG. Thus, Γ permutes the block idempotents of LG and fixes the block idempotents b of KG. Since $\mathrm{Br}_P : (LG)^P \to LC_G(P)$ commutes with the Γ-action, Theorem 2.3 implies that any Γ-conjugate of b has the same defect groups as b. We denote by Γ_b the stabilizer of b in Γ and set

$$\tilde{b} := \sum_{\gamma \in [\Gamma/\Gamma_b]} {}^{\gamma}b.$$

Clearly, $\tilde{b}$ is an idempotent in $(Z(LG))^\Gamma = Z(KG)$. More precisely one has the following:

Proposition 4.1. (a) *Let b be a block idempotent of LG. Then $\tilde{b} := \sum_{\gamma\in[\Gamma/\Gamma_b]} {}^\gamma b$ is a block idempotent of KG.*

(b) *The map $b \mapsto \tilde{b}$ induces a bijection between the set of Γ-orbits of block idempotents of LG and the set of block idempotents of KG.*

(c) *If b is a block idempotent of LG and $\tilde{b}$ is the block idempotent of KG associated to it as in* (*a*) *then b and $\tilde{b}$ have the same defect groups.*

Proof. (a) By definition, $\tilde{b}$ is the sum of the distinct Γ-conjugates of b, thus an idempotent of $Z(KG)$. To see that $\tilde{b}$ is primitive in $Z(KG)$, assume that $\tilde{b} = c_1 + c_2$ for nonzero orthogonal idempotents $c_1, c_2 \in Z(KG)$ and let I_1 and I_2 denote the set of primitive idempotents of $Z(LG)$ that occur in a primitive decomposition of c_1 and c_2 in $Z(LG)$, respectively. Then I_1 and I_2 are disjoint and Γ-stable. On the other hand $I_1 \cup I_2$ is the single Γ-orbit of b. This is a contradiction.

(b) This is immediate from (a).

(c) Let P be a defect group of $\tilde{b}$. By Theorem 2.3, $\mathrm{Br}_P(\tilde{b}) \neq 0$ in $KC_G(P) \subseteq LC_G(P)$. Thus $0 \neq \mathrm{Br}_P(\tilde{b}) = \sum_{\gamma\in[\Gamma/\Gamma_b]} \mathrm{Br}_P({}^\gamma b)$ implies that some Γ-conjugate of b, and therefore also b, has a defect group Q containing P. Thus, $0 \neq \mathrm{Br}_Q(b) = \mathrm{Br}_Q(b\tilde{b}) = \mathrm{Br}_Q(b)\mathrm{Br}_Q(\tilde{b})$, which implies that $\mathrm{Br}_Q(\tilde{b}) \neq 0$ and therefore $|Q| \leqslant |P|$. This implies $P = Q$. □

Note that Γ acts on $\mathcal{BP}(LG)$ via

$$ {}^\gamma(P, e) = (P, {}^\gamma e), \tag{1}$$

for $\gamma \in \Gamma$ and $(P, e) \in \mathcal{BP}(LG)$. Note that this action commutes with the G-action on $\mathcal{BP}(LG)$ so that we obtain an action of $\Gamma \times G$ on $\mathcal{BP}(LG)$. Moreover, since Br_P commutes with the action of Γ and since the G-action on LG commutes with the Γ-action on LG, $\Gamma \times G$ acts via poset isomorphisms on $\mathcal{BP}(LG)$. Thus, if b is a block idempotent of LG and $\gamma \in \Gamma$, the G-posets $\mathcal{BP}(LGb)$ and $\mathcal{BP}(LG\,{}^\gamma b)$ are isomorphic via (1) and $\Gamma_b \times G$ acts via poset automorphisms on $\mathcal{BP}(LGb)$.

In the next proposition we write $\leqslant_K$ and $\leqslant_L$ for the poset structures of $\mathcal{BP}(KG)$ and $\mathcal{BP}(LG)$, respectively. They are related as follows.

Proposition 4.2. *For $(Q, f), (P, e) \in \mathcal{BP}(LG)$ with $Q \leqslant P$, the following are equivalent*:

(i) *One has $(Q, \tilde{f}) \leqslant_K (P, \tilde{e})$ in $\mathcal{BP}(KG)$.*

(ii) *There exists $\gamma \in \Gamma$ such that $(Q, f) \leqslant_L {}^\gamma(P, e)$ in $\mathcal{BP}(LG)$.*

Proof. Assume first that (i) holds and let i be a primitive idempotent of $(KG)^P$ such that $\mathrm{Br}_P(i)\tilde{e} = \mathrm{Br}_P(i) \neq 0$. Then, by definition also $\mathrm{Br}_Q(i)\tilde{f} = \mathrm{Br}_Q(i) \neq 0$.

Let J be a primitive decomposition of i in $(LG)^P$. Since $\mathrm{Br}_P(i)\tilde{e} \neq 0$, there exists $j \in J$ such that $\mathrm{Br}_P(j)\tilde{e} \neq 0$. Thus, there exists $\gamma \in \Gamma$ such that $\mathrm{Br}_P(j)\,{}^{\gamma}e \neq 0$. Since $\mathrm{Br}_P(j)$ is primitive in $LC_G(P)$, we have $\mathrm{Br}_P(j)\,{}^{\gamma}e = \mathrm{Br}_P(j)$. Let f' be the block idempotent of $LC_G(Q)$ such that $(Q, f') \leqslant_L (P, {}^{\gamma}e) = {}^{\gamma}(P, e)$. Then, by Theorem 2.2 also $\mathrm{Br}_Q(j) f' = \mathrm{Br}_Q(j) \neq 0$. Thus $\mathrm{Br}_Q(j) f' \tilde{f} = \mathrm{Br}_Q(j)\mathrm{Br}_Q(i) f' \tilde{f} = \mathrm{Br}_Q(j) f' \mathrm{Br}_Q(i) \tilde{f} = \mathrm{Br}_Q(j)\mathrm{Br}_Q(i) = \mathrm{Br}_Q(j) \neq 0$ which implies that $f' \tilde{f} \neq 0$. This implies $f' = {}^{\delta}f$ for some $\delta \in \Gamma$. Thus ${}^{\delta}(Q, f) \leqslant_L {}^{\gamma}(P, e)$ and (ii) holds after applying δ^{-1}.

Next assume that $\gamma \in \Gamma$ with $(Q, f) \leqslant_L {}^{\gamma}(P, e)$. By Theorem 2.1(a) there exists a block idempotent f_1 of $LC_G(Q)$ such that $(Q, \tilde{f}_1) \leqslant_K (P, \tilde{e})$. Since we already proved that (i) implies (ii), there exists $\delta \in \Gamma$ such that $(Q, f_1) \leqslant_L {}^{\delta}(P, e)$. Thus we have $(Q, {}^{\gamma^{-1}}f) \leqslant_L (P, e)$ and also $(Q, {}^{\delta^{-1}}f) \leqslant_L (P, e)$. The uniqueness part of Theorem 2.2(a) now implies that f and f_1 are Γ-conjugate. Thus $\tilde{f} = \tilde{f}_1$ and $(Q, \tilde{f}) \leqslant_K (P, \tilde{e})$. □

The following corollaries are now immediate from Proposition 4.2.

Corollary 4.3. *The map*

$$\mathcal{BP}(LG) \to \mathcal{BP}(KG), \quad (P, e) \mapsto (P, \tilde{e}),$$

is a surjective morphism of G-posets, which restricts to a surjective morphism of G-posets $\mathcal{BP}(LGb) \to \mathcal{BP}(KG\tilde{b})$ for every block idempotent b of LG.

Corollary 4.4. *Let b be a block idempotent of LG and let $(P, e) \in \mathcal{BP}(LGb)$ be a maximal LGb-Brauer pair. Then $(P, \tilde{e}) \in \mathcal{BP}(KG\tilde{b})$ is a maximal $(KG\tilde{b})$-Brauer pair and one obtains an inclusion of fusion systems*

$$\mathcal{F}_{(P,e)}(LGb) \to \mathcal{F}_{(P,\tilde{e})}(KG\tilde{b})$$

which is the identity on objects and on morphisms.

5. The Main Theorem

We keep p, G, L/K, and Γ as introduced at the beginning of Section 4. Moreover we fix a block idempotent b of LG and denote by Γ_b the stabilizer of b in Γ. We fix a maximal LGb-Brauer pair $(P, e) \in \mathcal{BP}(LGb)$. For every $Q \leqslant P$, let e_Q denote the unique block idempotent of $LC_G(Q)$ such that $(Q, e_Q) \leqslant (P, e)$ in $\mathcal{BP}(LG)$. By Proposition 4.2, one has $(Q, \widetilde{e_Q}) \leqslant (P, \tilde{e})$ so that $\widetilde{e_Q} = \tilde{e}_Q$. This allows us to use the notation $\tilde{e}_Q$ for both purposes. Recall that $\Gamma \times G$ acts on $\mathcal{BP}(LG)$ and $\Gamma_b \times G$ acts on $\mathcal{BP}(LGb)$ via poset isomorphisms. Note that for any $(Q, f) \in \mathcal{BP}(LGb)$ one has $\Gamma_{(Q,f)} = \Gamma_f$. For the stabilizer in G of a KG-Brauer pair or LG-Brauer pair (Q, f) we will write $N_G(Q, f)$.

Let $p_1 : G \times \Gamma \to G$ and $p_2 : G \times \Gamma \to \Gamma$ denote the projection maps. For any subgroup X of $G \times \Gamma$, we set $k_1(X) := \{g \in G \mid (g, 1) \in X\}$ and $k_2(X) := \{\gamma \in \Gamma \mid (1, \gamma) \in X\}$. As explained in [Bouc 2010, p. 24], one has

$$(2) \quad k_1(X) \trianglelefteq p_1(X) \leqslant G, \; k_2(X) \trianglelefteq p_2(X) \leqslant \Gamma, \text{ and } p_1(X)/k_1(X) \cong p_2(X)/k_2(X)$$

via $gk_1(X) \leftrightarrow \gamma k_2(X)$ if and only if $(g, \gamma) \in X$.

We denote by $K(b)$ and $K(e)$ the subfields of L obtained by adjoining the coefficients of the block idempotents $b \in LG$ and $e \in LC_G(P)$. Thus, $K(b)$ is the fixed field of Γ_b in L and $K(e)$ is the fixed field of Γ_e in L.

Proposition 5.1. *Let b be a block idempotent of LG.*

(a) For any $(R, e_R) \leqslant (Q, e_Q)$ in $\mathcal{BP}(LGb)$ one has $\Gamma_e = \Gamma_{(P,e)} \leqslant \Gamma_{(Q,e_Q)} \leqslant \Gamma_{(R,e_R)} \leqslant \Gamma_{(\{1\},b)} = \Gamma_b$. In particular, $K(b) \subseteq K(e)$.

(b) Let $X := \mathrm{stab}_{G\times\Gamma}(P, e)$ be the stabilizer of the maximal LGb-Brauer pair (P, e). One has

$$k_1(X) = N_G(P, e), \quad p_1(X) = N_G(P, \tilde{e}), \quad k_2(X) = \Gamma_e, \quad \text{and} \quad p_2(X) = \Gamma_b.$$

(c) One has $N_G(P, e) \trianglelefteq N_G(P, \tilde{e})$ and $N_G(P, \tilde{e})/N_G(P, e) \cong \Gamma_b/\Gamma_e$. Moreover, $K(e)/K(b)$ is a Galois extension with cyclic Galois group isomorphic to $N_G(P, \tilde{e})/N_G(P, e)$.

Proof. (a) It suffices to show that $\Gamma_{(Q,e_Q)} \leqslant \Gamma_{(R,e_R)}$. Let $\gamma \in \Gamma_{(Q,e_Q)}$. Then ${}^{\gamma}(R, e_R) \leqslant_L {}^{\gamma}(Q, e_Q) = (Q, {}^{\gamma}e_Q) = (Q, e_Q)$. The uniqueness part of Theorem 2.1(a) implies that ${}^{\gamma}e_R = e_R$. Thus, $\gamma \in \Gamma_{(R,e_R)}$.

(b) The first equation is clear from the definition of $k_1(X)$. For the proof of the second equation, let $g \in p_1(X)$. Then there exists $\gamma \in \Gamma$ with $(P, e) = {}^{(g,\gamma)}(P, e) = ({}^{g}P, {}^{g\gamma}e)$. From ${}^{g\gamma}e = e$ it follows that ${}^{g}\tilde{e} = \tilde{e}$. Thus ${}^{g}(P, \tilde{e}) = (P, \tilde{e})$ and $g \in N_G(P, \tilde{e})$. Conversely, if $g \in N_G(P, \tilde{e})$ then ${}^{g}\tilde{e} = \tilde{e}$ which implies that there exists $\gamma \in \Gamma$ with ${}^{g}e = {}^{\gamma}e$. Thus, ${}^{(g,\gamma^{-1})}(P, e) = (P, e)$ and $g \in p_1(X)$. The third equation follows immediately from the definition of $k_2(X)$. For the proof of the fourth equation let $\gamma \in p_2(X)$. Then there exists $g \in G$ with ${}^{(g,\gamma)}(P, e) = (P, e)$. Since $(\{1\}, b) \leqslant (P, e)$, this implies ${}^{(g,\gamma)}(\{1\}, b) \leqslant {}^{(g,\gamma)}(P, e) = (P, e)$. The uniqueness part in Theorem 2.1(a) implies that ${}^{(g,\gamma)}(\{1\}, b) = (1, b)$ and that $\gamma \in \Gamma_b$. Conversely, assume that $\gamma \in \Gamma_b$. Then $(\{1\}, b) \leqslant (P, e)$ implies $(\{1\}, b) = {}^{(1,\gamma)}(\{1\}, b) \leqslant {}^{(1,\gamma)}(P, e) = (P, {}^{\gamma}e)$. This implies that both (P, e) and ${}^{\gamma}(P, e)$ are maximal LGb-Brauer pairs. By Theorem 2.3(a), there exists $g \in G$ such that ${}^{g}(P, {}^{\gamma}e) = (P, e)$. Thus $(g, \gamma) \in X$ and $\gamma \in p_2(X)$.

(c) The assertions of the first sentence follow from part (b) and (2). For the second statement it suffices to show that Γ_b/Γ_e is cyclic. Note that the coefficients of $e \in LC_G(P)$ generate a finite field extension of the prime field $\mathbb{F}_p$ in L, which we denote by $\mathbb{F}_p(e)$. Since $\Gamma_e \trianglelefteq \Gamma_b$, we have a Galois extension $K(e)/K(b)$ with Galois group $\Delta \cong \Gamma_b/\Gamma_e$. Now, restriction from $K(e)$ to $\mathbb{F}_p(e)$ is an injective group homomorphism from Δ to the cyclic Galois group $\mathrm{Gal}(\mathbb{F}_p(e)/\mathbb{F}_p)$. In fact, if $\delta \in \Delta$ restricts to the identity on $\mathbb{F}_p(e)$, then it is the identity on $\mathbb{F}_p(e)$ and on K, thus on $K(e)$. This completes the proof of Part (c). □

Next we give a more precise picture of the inclusion of fusion systems from Corollary 4.4. In the following theorem the term $\langle\mathcal{F}, \sigma\rangle$ denotes the fusion system *generated* by $\mathcal{F}$ and σ, i.e., the intersection of all fusion systems over P that contain $\mathcal{F}$ and σ.

Theorem 5.2. *Let L/K be a finite Galois extension of fields of characteristic $p > 0$ with Galois group Γ, let b be a block idempotent of LG, and let (P, e) be a maximal LGb-Brauer pair. Set $\mathcal{F} := \mathcal{F}_{(P,e)}(LGb)$ and $\widetilde{\mathcal{F}} := \mathcal{F}_{(P,\tilde{e})}(KG\tilde{b})$. Let $g_0 \in N_G(P, e)$ be such that $g_0 N_G(P, \tilde{e})$ generates $N_G(P, e)/N_G(P, \tilde{e})$ (see Proposition 5.1(c)) and set $\sigma := c_{g_0} \in \mathrm{Aut}(P)$. Then $\widetilde{\mathcal{F}} = \langle\mathcal{F}, \sigma\rangle$.*

More precisely, $\sigma \in \mathrm{Aut}_{\widetilde{\mathcal{F}}}(P)$ and, for any subgroups Q and R of P and any $\varphi \in \mathrm{Hom}_{\widetilde{\mathcal{F}}}(Q, R)$, there exist $i \in \mathbb{Z}$, $\psi \in \mathrm{Hom}_{\mathcal{F}}(Q, \sigma^{-i}(R))$ and $\psi' \in \mathrm{Hom}_{\mathcal{F}}(\sigma^i(Q), R)$ with $\varphi = \sigma^i|_{\sigma^{-i}(R)} \circ \psi = \psi' \circ \sigma^i|_Q$.

Proof. Since $g_0 \in N_G(P, \tilde{e})$, we have $\sigma = c_{g_0} \in \mathrm{Aut}_{\widetilde{\mathcal{F}}}(P)$. It follows that $\langle\mathcal{F}, \sigma\rangle \subseteq \widetilde{\mathcal{F}}$. In order to prove the reverse inclusion, let Q and R be subgroups of P and let $\varphi \in \mathrm{Hom}_{\widetilde{\mathcal{F}}}(Q, R)$. Then there exists $g \in G$ such that $\varphi = c_g : Q \to R$ and ${}^g(Q, \tilde{e}_Q) \leqslant_K (R, \tilde{e}_R)$. By Proposition 4.2 there exists $\gamma \in \Gamma$ such that ${}^g(Q, e_Q) \leqslant_L (R, {}^\gamma e_R)$. Since $(\{1\}, b) = {}^g(\{1\}, b) \leqslant_L {}^g(Q, e_Q) \leqslant_L (R, {}^\gamma e_R)$ and also $(\{1\}, {}^\gamma b) \leqslant_L (R, {}^\gamma e_R)$, Theorem 2.1(a) implies $(\{1\}, b) = (\{1\}, {}^\gamma b)$ so that $\gamma \in \Gamma_b$. Thus, both (P, e) and $(P, {}^\gamma e)$ are maximal LGb-Brauer pairs. Theorem 2.3(a) implies that there exists $h \in G$ such that ${}^h(P, e) = (P, {}^\gamma e)$ and we obtain $(P, e) = {}^{h^{-1}}(P, {}^\gamma e) \geqslant_L {}^{h^{-1}}(R, {}^\gamma e_R) = ({}^{h^{-1}}R, {}^{h^{-1}\gamma}e_R)$. Again, Theorem 2.1(a) implies that ${}^{h^{-1}\gamma}e_R = e_{h^{-1}Rh}$ and thus ${}^{h^{-1}g}(Q, e_Q) \leqslant_L {}^{h^{-1}}(R, {}^\gamma e_R) = ({}^{h^{-1}}R, e_{h^{-1}Rh})$. This in turn implies that the homomorphism $\alpha := c_{h^{-1}g} : Q \to {}^{h^{-1}}R$ belongs to $\mathrm{Hom}_{\mathcal{F}}(Q, {}^{h^{-1}}R)$ and that the homomorphism $\varphi = c_g : Q \to R$ factors as

$$\varphi = c_h \circ \alpha : Q \to {}^{h^{-1}}R \to R. \tag{3}$$

Since ${}^h(P, e) = (P, {}^\gamma e)$, we obtain $h \in N_G(P, \tilde{e})$ and can write $h = g_0^i x$ for some $i \in \mathbb{Z}$ and $x \in N_G(P, e)$. This implies that the map $c_h : P \to P$ factors as $c_h = \sigma^i \circ \beta : P \to P$, where $\sigma^i = c_{g_0^i} : P \to P$ and $\beta := c_x \in \mathrm{Aut}_{\mathcal{F}}(P)$, since $x \in N_G(P, e)$. Restriction to ${}^{h^{-1}}R$ yields the factorization

$$c_h|_{h^{-1}Rh} = \sigma^i|_{\beta(h^{-1}Rh)} \circ \beta|_{h^{-1}Rh} : {}^{h^{-1}}R \to \beta({}^{h^{-1}}R) \to R$$

with $\beta({}^{h^{-1}}R) = \sigma^{-i}(R)$ and $\beta|_{h^{-1}Rh} \in \mathrm{Hom}_{\mathcal{F}}({}^{h^{-1}}R, R)$. Setting $\psi := \beta|_{h^{-1}Rh} \circ \alpha : Q \to \sigma^{-1}(R)$ and using (3) we obtain the desired factorization of φ. This also implies the inclusion $\widetilde{\mathcal{F}} \subseteq \langle\mathcal{F}, \sigma\rangle$.

In order to find ψ' with the desired property we use the elements g, h, x, and i from the first part of the proof and note that

$$(P,e) = {}^{\gamma^{-1}h}(P,e) \geqslant_L {}^{\gamma^{-1}}({}^hQ, {}^he_Q) = ({}^hQ, {}^{\gamma^{-1}h}e_Q),$$

which implies that ${}^{\gamma^{-1}h}e_Q = e_{hQh^{-1}}$. Thus,

$${}^{gh^{-1}}({}^hQ, e_{hQh^{-1}}) = {}^{gh^{-1}}({}^hQ, {}^{\gamma^{-1}h}e_Q) = ({}^gQ, {}^{g\gamma^{-1}}e_Q) \leqslant_L (R, e_R),$$

which implies that $\alpha' := c_{gh^{-1}} : {}^hQ \to R$ belongs to $\mathrm{Hom}_{\mathcal{F}}({}^hQ, R)$. Thus, φ can be factored as

$$\varphi = c_g = c_{gh^{-1}} \circ c_h = \alpha' \circ c_h : Q \to {}^hQ \to R. \tag{4}$$

We can rewrite $h = g_0^i x = x' g_0^i$ for some $x' \in N_G(P,e)$ and obtain an element $\beta' \in \mathrm{Aut}_{\mathcal{F}}(P)$ together with a factorization $c_h = \beta' \circ \sigma^i : P \to P$. Restricting this equation to Q yields a factorization

$$c_h = \beta'|_{\sigma^i(Q)} \circ \sigma^i|_Q : Q \to \sigma^i(Q) \to {}^hQ.$$

Setting $\psi' := \alpha' \circ \beta'|_{\sigma^i(Q)} \in \mathrm{Hom}_{\mathcal{F}}(\sigma^i(Q), R)$, the factorization in (4) can now be expressed as $\varphi = \psi' \circ \sigma^i|_Q$ as claimed. □

6. Consequences of the Main Theorem

In this section we prove several consequences of Theorem 5.2.

Recall that if $\mathcal{F}$ is a fusion system over a p-group P, a subgroup Q of P is called $\mathcal{F}$-centric if $C_P(R) = Z(R)$ for all subgroups R of P which are $\mathcal{F}$-isomorphic to Q.

Proposition 6.1. *Let L/K, b, (P,e) and $\mathcal{F} \subseteq \tilde{\mathcal{F}}$ be as in Theorem 5.2.*

(a) *A subgroup Q of P is fully $\mathcal{F}$-centralized if and only if it is fully $\tilde{\mathcal{F}}$-centralized.*

(b) *A subgroup Q of P is fully $\mathcal{F}$-normalized if and only if it is fully $\tilde{\mathcal{F}}$-normalized.*

(c) *A subgroup Q of P is $\mathcal{F}$-centric if and only it is $\tilde{\mathcal{F}}$-centric.*

Proof. The "if" parts follow immediately from the fact that the $\mathcal{F}$-isomorphism class of Q is a subset of the $\tilde{\mathcal{F}}$-isomorphism class of Q. For the forward implications note that by Theorem 5.2 two subgroups Q and Q' of P are $\tilde{\mathcal{F}}$-isomorphic if and only if there exists a subgroup Q'' of P such that Q is $\mathcal{F}$-isomorphic to Q'' and $Q' = \sigma^i(Q'')$ for some $i \in \mathbb{Z}$. Moreover, $\sigma^i(C_P(Q'')) = C_P(\sigma^i(Q''))$, $\sigma^i(N_P(Q'')) = N_P(\sigma^i(Q''))$, and $\sigma^i(Z(Q'')) = Z(\sigma^i(Q''))$, since σ^i is an automorphism of P. The result is now immediate. □

The following theorem is known to experts. See for instance the part of the proof of [Linckelmann 2018, Theorem 8.5.2] dealing with the extension axiom and note that it does not use any assumptions on the field of coefficients k. Below is a proof with a different approach, using Theorem 5.2.

Theorem 6.2. *Let k be a field of characteristic $p > 0$ and let c be a block idempotent of kG. Then the extension axiom holds for the fusion system of kGc, for any choice of maximal Brauer pair.*

Proof. Let (P, f) be a maximal kGc-Brauer pair. We apply Theorem 5.2 with $K = k$, a splitting field L of $KC_G(P)f$ such that L/K is a finite Galois extension with Galois group Γ, and to a block idempotent b of LG with $cb \neq 0$. Then $c = \tilde{b}$. Moreover, there exists a maximal LGb-Brauer pair (P, e) such that $ef = e$ and therefore $f = \tilde{e}$. We aim to show that the fusion system $\tilde{\mathcal{F}} = \mathcal{F}_{(P,\tilde{e})}(KG\tilde{b})$ satisfies the extension axiom. Note that by Theorem 3.7, the extension axiom holds for $\mathcal{F} = \mathcal{F}_{(P,e)}(LGb)$, since L is a splitting field of $LC_G(P)e$. Let $\varphi \in \mathrm{Hom}_{\tilde{\mathcal{F}}}(Q, P)$ be such that $\varphi(Q)$ is fully $\tilde{\mathcal{F}}$-normalized. By Theorem 5.2 we can factorize $\varphi = \sigma^i \circ \psi$ for some $\psi \in \mathrm{Hom}_{\mathcal{F}}(Q, P)$. With $\varphi(Q)$ also $\psi(Q) = \sigma^{-i}(\varphi(Q))$ is fully $\tilde{\mathcal{F}}$-normalized, since they are $\tilde{\mathcal{F}}$-isomorphic and $N_P(\psi(Q)) = \sigma^{-i}(N_P(\varphi(Q)))$. By Proposition 6.1(b), $\psi(Q)$ is fully $\mathcal{F}$-normalized. Since $\mathcal{F}$ satisfies the extension axiom, there exists $\hat{\psi} \in \mathrm{Hom}_{\mathcal{F}}(N_\psi, P)$ such that $\hat{\psi}|_Q = \psi$. It follows that $\hat{\varphi} := \sigma^i \circ \hat{\psi} \in \mathrm{Hom}_{\tilde{\mathcal{F}}}(N_\psi, P)$ extends φ. To finish the proof it suffices to show that $N_\varphi \subseteq N_\psi$. So let $x \in N_\varphi$. Then $x \in N_P(Q)$ and there exists $y \in N_P(\varphi(Q))$ with $\varphi \circ c_x = c_y \circ \varphi : Q \xrightarrow{\sim} \varphi(Q)$. But this implies

$$\psi \circ c_x = \sigma^{-i} \circ \varphi \circ c_x = \sigma^{-i} \circ c_y \circ \varphi = c_{\sigma^{-i}(y)} \circ \sigma^{-i} \circ \varphi = c_{\sigma^{-i}(y)} \circ \psi,$$

with $\sigma^{-i}(y) \in \sigma^{-i}(N_P(\varphi(Q))) = N_P(\sigma^{-i}(\varphi(Q))) = N_P(\psi(Q))$. Thus, $N_\varphi \subseteq N_\psi$ and the proof is complete. □

Theorem 6.3. *Let L/K, b, (P, e) and $\mathcal{F} \subseteq \tilde{\mathcal{F}}$ be as in Theorem 5.2. The fusion system $\tilde{\mathcal{F}}$ is saturated if and only if the fusion system $\mathcal{F}$ is saturated and p does not divide $[N_G(P, \tilde{e}) : N_G(P, e)] = [\Gamma_b : \Gamma_e] = [K(e) : K(b)]$. In particular, if moreover L is a splitting field for $LC_G(P)e$, then $\tilde{\mathcal{F}}$ is saturated if and only if p does not divide $[N_G(P, \tilde{e}) : N_G(P, e)] = [\Gamma_b : \Gamma_e] = [K(e) : K(b)]$.*

Proof. Note that the map $N_G(P, e) \to \mathrm{Aut}_{\mathcal{F}}(P)$, $g \mapsto c_g$, induces an isomorphism $N_G(P, e)/C_G(P) \to \mathrm{Aut}_{\mathcal{F}}(P)$ which maps $PC_G(P)/C_G(P)$ to $\mathrm{Aut}_P(P)$. Thus, the Sylow axiom holds for $\mathcal{F}$ if and only if $p \nmid [N_G(P, e) : PC_G(P)]$. Similarly, the Sylow axiom holds for $\tilde{\mathcal{F}}$ if and only if $p \nmid [N_G(P, \tilde{e}) : PC_G(P)]$. By Theorem 6.2 it suffices to show that the Sylow axiom holds for $\tilde{\mathcal{F}}$ if and only it holds for $\mathcal{F}$ and $p \nmid [\Gamma_b : \Gamma_e]$. But, by Proposition 5.1(c), one has $[\Gamma_b : \Gamma_e] = [N_G(P, \tilde{e}) : N_G(P, e)] = [K(e) : K(b)]$ which implies the result. □

Next we will show that a weak form of Alperin's fusion theorem holds for arbitrary block fusion systems.

Definition 6.4. Let $\mathcal{F}$ be a fusion system over a p-group P. We say that *Alperin's weak fusion theorem holds for $\mathcal{F}$* if $\mathcal{F} = \langle \mathrm{Aut}_{\mathcal{F}}(Q) \mid Q \in \mathcal{C} \rangle$, where $\mathcal{C}$ is the set of subgroups of P which are $\mathcal{F}$-centric and fully $\mathcal{F}$-normalized.

Theorem 6.5. *Let k be a field of characteristic p and let c be a block idempotent of kG. Then Alperin's weak fusion theorem holds for the fusion system of kGc, for any choice of maximal kGc-Brauer pair.*

Proof. Set $K := k$ and choose L, b, (P, e) as in the proof of Theorem 6.2 with $c = \tilde{b}$ and apply Theorem 5.2 to this situation with $\mathcal{F} := \mathcal{F}_{(P,e)}(LGb)$ and $\widetilde{\mathcal{F}} := \mathcal{F}_{(P,\tilde{e})}(KG\tilde{b})$. We need to show that Alperin's weak fusion theorem holds for $\widetilde{\mathcal{F}}$. Since $\mathcal{F}$ is saturated, Alperin's weak fusion theorem holds for $\mathcal{F}$, see for instance [Linckelmann 2018, Theorem 8.2.8]. Thus, $\mathcal{F} = \langle \mathrm{Aut}_{\mathcal{F}}(Q) \mid Q \in \mathcal{C} \rangle$, where $\mathcal{C}$ denotes the set of subgroups of P which are $\mathcal{F}$-centric and fully $\mathcal{F}$-normalized. Moreover, by Proposition 6.1, $\mathcal{C}$ is equal to the set $\widetilde{\mathcal{C}}$ of subgroups of P which are $\widetilde{\mathcal{F}}$-centric and fully $\widetilde{\mathcal{F}}$-normalized. Thus, by Theorem 5.2, we have

$$\widetilde{\mathcal{F}} = \langle \mathcal{F}, \sigma \rangle = \langle \{\mathrm{Aut}_{\mathcal{F}}(Q) \mid Q \in \mathcal{C}\} \cup \{\sigma\} \rangle \subseteq \langle \mathrm{Aut}_{\widetilde{\mathcal{F}}}(Q) \mid Q \in \mathcal{C} \rangle \subseteq \widetilde{\mathcal{F}}.$$

But this implies $\widetilde{\mathcal{F}} = \langle \mathrm{Aut}_{\widetilde{\mathcal{F}}}(Q) \mid Q \in \mathcal{C} \rangle = \langle \mathrm{Aut}_{\widetilde{\mathcal{F}}}(Q) \mid Q \in \widetilde{\mathcal{C}} \rangle$, which means that Alperin's weak fusion theorem holds for $\widetilde{\mathcal{F}}$. □

References

[Aschbacher et al. 2011] M. Aschbacher, R. Kessar, and B. Oliver, *Fusion systems in algebra and topology*, London Mathematical Society Lecture Note Series **391**, Cambridge University Press, 2011. MR Zbl

[Bouc 2010] S. Bouc, *Biset functors for finite groups*, Lecture Notes in Mathematics **1990**, Springer, Berlin, 2010. MR Zbl

[Linckelmann 2018] M. Linckelmann, *The block theory of finite group algebras*, vol. II, London Mathematical Society Student Texts **92**, Cambridge University Press, 2018. MR Zbl

Received July 24, 2019.

ROBERT BOLTJE
boltje@ucsc.edu

ÇİSİL KARAGÜZEL
ckaraguz@ucsc.edu

DENİZ YILMAZ
deyilmaz@ucsc.edu

(*all*)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
SANTA CRUZ, CA
UNITED STATES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.43

TORSION POINTS AND GALOIS REPRESENTATIONS ON CM ELLIPTIC CURVES

ABBEY BOURDON AND PETE L. CLARK

We prove several results on torsion points and Galois representations for complex multiplication (CM) elliptic curves over a number field containing the CM field. One result computes the degree in which such an elliptic curve has a rational point of order N, refining results of Silverberg (*Compositio Math.* 68:3 (1988), 241–249; *Contemp. Math.* 133 (1992)). Another result bounds the size of the torsion subgroup of an elliptic curve with CM by a nonmaximal order in terms of the torsion subgroup of an elliptic curve with CM by the maximal order. Our techniques also yield a complete classification of both the possible torsion subgroups and the rational cyclic isogenies of a K-CM elliptic curve E defined over $K(j(E))$.

1. Introduction

Let F be a field of characteristic 0, and let $E_{/F}$ be an elliptic curve. We say E has *complex multiplication* (CM) if the endomorphism algebra

$$\mathrm{End}^0 E = \mathrm{End}(E_{/\overline{F}}) \otimes_{\mathbb{Z}} \mathbb{Q}$$

is strictly larger than $\mathbb{Q}$, in which case it is necessarily an imaginary quadratic field K and $\mathcal{O} := \mathrm{End}(E_{/\overline{F}})$ is a $\mathbb{Z}$-order in K.

The general theory of complex multiplication has a long and rich history, with important contributions made by Kronecker, Weber, Fricke, Hasse, Deuring, and

MSC2010: 11G05, 11G15.
Keywords: torsion points, elliptic curves, Galois representations, complex multiplication.

Shimura. For a summary of these foundational results, see [Silverman 1994, Chapter 2]. More recent contributions to the study of torsion points and Galois representations on CM elliptic curves defined over number fields have been made by Olson [1974], Silverberg [1988; 1992], Parish [1989], Aoki [1995; 2006], Ross [1994], Kwon [1999], Prasad and Yogananda [2001], Stevenhagen [2001], Breuer [2010], Lombardo [2017], Lozano-Robledo [2018b], Gaudron and Rémond [2018] and the present authors and our collaborators [Clark et al. 2013; 2014; Bourdon et al. 2017a; 2017b; Clark and Pollack 2015; Bourdon and Pollack 2017]. In this paper, we consider the case of a CM elliptic curve defined over a number field that contains the CM field. The case in which the ground field is a number field not assumed to contain the CM field is pursued in [Bourdon and Clark 2019]. There is related work of Á. Lozano-Robledo [2018a] done concurrently with the present work, which determines all possible images of the ℓ-adic Galois representations of a CM elliptic curve E over $\mathbb{Q}(j(E))$ up to conjugacy.

Throughout this introduction we maintain the following notation: K is an imaginary quadratic field, $\mathcal{O}$ is an order in K, $\mathfrak{f}$ is the conductor of $\mathcal{O}$, $K(\mathfrak{f})$ is the $\mathfrak{f}$-ring class field of K (i.e., $K(\mathfrak{f}) = K(j(E))$ for any $\mathcal{O}$-CM elliptic curve E), F is a number field containing K and N is a positive integer.

1A. ***The Torsion Degree Theorem.*** Let $\mathcal{O}$ be an order in the imaginary quadratic field K, and let $N \in \mathbb{Z}^+$. The following result was first proven by Silverberg [1988; 1992] and then subsequently by Prasad and Yogananda [2001].

Theorem 1.1 (Silverberg). *Let $F \supset K$ be a number field, and suppose that there is an $\mathcal{O}$-CM elliptic curve $E_{/F}$ with an F-rational point of order N. Then*

$$\varphi(N) \le \#\mathcal{O}^\times \cdot [F : K].$$

Theorem 1.1 is a crucial result in the study of torsion subgroups of CM elliptic curves over general number fields. For instance, it was the main tool in the complete enumeration of torsion subgroups of CM elliptic curves defined over number fields of small degree [Clark et al. 2013; 2014].

The hypotheses of Theorem 1.1 force $F \supset K(\mathfrak{f}) = K(j(E))$. Thus it is natural to define $T(\mathcal{O}, N)$ to be the least degree $[F : K(\mathfrak{f})]$ of a number field $F \supset K$ over which some $\mathcal{O}$-CM elliptic curve admits an F-rational point of order N. We show in Theorem 6.2 that

$$\varphi(N) \mid \#\mathcal{O}^\times \cdot T(\mathcal{O}, N), \tag{1}$$

i.e., Theorem 1.1 holds as a divisibility.

Our first main result computes $T(\mathcal{O}, N)$ in all cases and gives the analogous divisibility refinement.

Theorem 1.2. *Let $\mathcal{O}$ be an order in the imaginary quadratic field K, and let N be a positive integer. There is an integer $T(\mathcal{O}, N)$, explicitly computed in Section 7, such that*

(i) *if $F \supset K$ is a number field and $E_{/F}$ is an $\mathcal{O}$-CM elliptic curve with an F-rational point of order N, then $T(\mathcal{O}, N) \mid [F : K(\mathfrak{f})]$, and*

(ii) *there is a number field $F \supset K$ and an $\mathcal{O}$-CM elliptic curve $E_{/F}$ such that $[F : K(\mathfrak{f})] = T(\mathcal{O}, N)$ and $E(F)$ has a point of order N.*

Equivalently, Theorem 1.2 determines the least degree of a closed $\mathcal{O}$-CM point on $X_1(N)_{/K}$ and shows that this degree divides the degree of all closed $\mathcal{O}$-CM points.

1B. *The Isogeny Torsion Theorem.* A key feature of the present work is that we work with *all* imaginary quadratic orders $\mathcal{O}$, not just the maximal order $\mathcal{O}_K$. Working with nonmaximal orders entails certain technical complications. For instance, if $F \supset K$ is a number field, then $E(F)[\text{tors}]$ is a finite $\mathcal{O}$-submodule of $E(\mathbb{C})$. As we will see in Section 2B, every finite $\mathcal{O}$-submodule of $E(\mathbb{C})$ is cyclic if and only if the order $\mathcal{O}$ is maximal.

The phenomenon of "ascending isogenies" can sometimes be used to study $\mathcal{O}$-CM elliptic curves in terms of $\mathcal{O}_K$-CM elliptic curves, and this happens twice in the present paper. Specifically, let E be an $\mathcal{O}$-CM elliptic curve defined over a number field F, and let $\mathfrak{f}'$ be a positive integer that divides $\mathfrak{f}$. Then by [Bourdon and Pollack 2017, Proposition 2.2], there is an elliptic curve $(E_{\mathfrak{f}'})_{/F}$ such that $\mathcal{O}(\mathfrak{f}') := \operatorname{End} E_{\mathfrak{f}'}$ is the order of conductor $\mathfrak{f}'$ in K and an F-rational isogeny $\iota_{\mathfrak{f}'} : E \to E_{\mathfrak{f}'}$ that is cyclic of degree $\mathfrak{f}/\mathfrak{f}'$. There is an embedding $F \hookrightarrow \mathbb{C}$ such that the base change of $\iota_{\mathfrak{f}'}$ to $\mathbb{C}$ is the natural map $\mathbb{C}/\mathcal{O} \to \mathbb{C}/\mathcal{O}(\mathfrak{f}')$ of complex elliptic curves. The map $\iota_{\mathfrak{f}'}$ is universal for maps from an $\mathcal{O}$-CM elliptic curve to an $\mathcal{O}(\mathfrak{f}')$-CM elliptic curve [Bourdon and Clark 2019, §2.6] and is thus unique, up to isomorphism on the target. Here is the first result making use of this canonical isogeny.

Theorem 1.3 (Isogeny Torsion Theorem). *Let $\mathcal{O}$ be an order in an imaginary quadratic field K, of conductor $\mathfrak{f}$, and let $\mathfrak{f}'$ be a positive integer dividing $\mathfrak{f}$. Let $F \supset K$ be a number field, and let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve. Let $\iota_{\mathfrak{f}'} : E \to E_{\mathfrak{f}'}$ be the F-rational isogeny to an elliptic curve $E_{\mathfrak{f}'}$ with CM by the order in K of conductor $\mathfrak{f}'$, as described above. Then we have*

$$\#E(F)[\text{tors}] \mid \#E_{\mathfrak{f}'}(F)[\text{tors}].$$

In particular, taking $\mathfrak{f}' = 1$, we see that $\#E(F)[\text{tors}]$ is bounded by $\#E_1(F)[\text{tors}]$, where $(E_1)_{/F}$ is an $\mathcal{O}_K$-CM elliptic curve. We give examples where the exponent of $E_{\mathfrak{f}'}(F)[\text{tors}]$ is strictly smaller than that of $E(F)[\text{tors}]$, showing in general we cannot view $E(F)[\text{tors}]$ as a subgroup of $E_{\mathfrak{f}'}(F)[\text{tors}]$, and we prove that $\#E_{\mathfrak{f}'}(F)[\text{tors}]/\#E(F)[\text{tors}]$ can be arbitrarily large (see Propositions 6.8 and 6.9). Moreover, the statement is false if we do not require $F \supset K$. Theorem 1.3 has applications to determining fields of moduli of partial level N structures (Sections 6B and 6C).

1C. ***The reduced Galois representation.*** There is a well-known interplay between points on modular curves over number fields and Galois representations of elliptic curves. The proofs of Theorems 1.2 and 1.3 make use of Galois representations, and in the former case we build on a nearly complete description of the image of the mod N Galois representation on an $\mathcal{O}$-CM elliptic $E_{/K(\mathfrak{f})}$.

For an elliptic curve E defined over a number field F and a positive integer N, the $\mathbb{Z}$-linear action of $\mathfrak{g}_F := \operatorname{Aut}(\overline{F}/F)$ on $E[N]$ gives rise to the mod N Galois representation:

$$\rho_N : \mathfrak{g}_F \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}).$$

When E does not have CM, a celebrated result of Serre [1972] asserts that as N varies over all positive integers, the index $[\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho_N(\mathfrak{g}_F)]$ remains bounded. This is certainly not the case when E has CM: as usual, here we consider the case in which F is a number field containing the CM field K. Then for $N \in \mathbb{Z}^+$, Galois acts by $\mathcal{O}$-linear endomorphisms of $E[N]$, which is a free $\mathcal{O}/N\mathcal{O}$-module of rank one. Thus the mod N Galois representation takes the form

$$\rho_N : \mathfrak{g}_F \to (\mathcal{O}/N\mathcal{O})^\times \hookrightarrow \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}).$$

In the CM case, the analogue of Serre's result is the boundedness of the index of ρ_N in $(\mathcal{O}/N\mathcal{O})^\times$ as N varies. In fact more is true: a slight variant of ρ_N is surjective for all $\mathcal{O}$ and N. To motivate this, observe that fixing $\mathcal{O}$ is the same as fixing $j(E)$ (up to Galois conjugacy), but fixing $j(E)$ does not determine the $K(\mathfrak{f})$-rational model of E and thus not ρ_N. One gets from one model to another via a twist by $d \in K(\mathfrak{f})^\times/K(\mathfrak{f})^{\times \#\mathcal{O}^\times}$. If E and E^d are elliptic curves over $K(\mathfrak{f})$ and $\rho_E, \rho_{E^d} : \mathfrak{g}_{K(\mathfrak{f})} \to (\mathcal{O}/N\mathcal{O})^\times$ are their mod N Galois representations, then $\rho_{E^d} = \rho_E \otimes \chi_d$, where $\chi_d : \mathfrak{g}_{K(\mathfrak{f})} \to \mathcal{O}^\times$ is the character corresponding to d. Thus we define the *reduced mod N Cartan subgroup*

$$\overline{C_N(\mathcal{O})} = C_N(\mathcal{O})/q_N(\mathcal{O}^\times)$$

to be the quotient of $C_N(\mathcal{O}) = (\mathcal{O}/N\mathcal{O})^\times$ by the image of $\mathcal{O}^\times$ under the natural map $q_N : \mathcal{O} \to \mathcal{O}/N\mathcal{O}$ and the *reduced mod N Galois representation* to be the composite homomorphism

$$\overline{\rho_N} : \mathfrak{g}_F \xrightarrow{\rho_N} C_N(\mathcal{O}) \to \overline{C_N(\mathcal{O})}.$$

The key feature of $\overline{\rho_N}$ is that it is independent of the $K(\mathfrak{f})$-rational model.

For an elliptic curve E defined over a field F of characteristic 0, there is an F-rational isomorphism $\iota : E/\operatorname{Aut}(E) \xrightarrow{\sim} \mathbb{P}^1$, and a *Weber function* on E is any function $\mathfrak{h} : E \to \mathbb{P}^1$ obtained by composing the quotient map with such an isomorphism ι. Then the field extension cut out by the reduced Galois representation is the field obtained by adjoining to $K(\mathfrak{f})$ the values of the Weber function on the N-torsion points of E:

$$\overline{\mathbb{Q}}^{\ker \overline{\rho_N}} = K(\mathfrak{f})(\mathfrak{h}(E[N])).$$

When $\mathcal{O} = \mathcal{O}_K$ the first main theorem of complex multiplication tells us that for any ideal I of $\mathcal{O}_K$ we have that $K(\mathfrak{f})(\mathfrak{h}(E[I]))$ is $K^{(I)}$, the I-ray class field of K. It turns out that

$$[K^I : K^{(1)}] = \#C_N(\mathcal{O})$$

and thus $\overline{\rho_N}$ is surjective. The case of an arbitrary order is much less classical but still known: Stevenhagen [2001] used Shimura's reciprocity law to show that for all $N \in \mathbb{Z}^+$, the Weber function field $K(\mathfrak{f})(\mathfrak{h}(E[N]))$ is $K(\mathfrak{f})^{N\mathcal{O}_K}$, the *$N$-ray class field of $\mathcal{O}$* — this is the finite abelian extension of K corresponding to the image of the subgroup $\mathbb{C}^\times \times \{x \in \widehat{\mathcal{O}}^\times \mid x \equiv 1 \pmod{N}\}$ in the norm one idèle class group of K. Moreover, it follows from the adelic description [Stevenhagen 2001, p. 8] that

$$\operatorname{Aut}(K(\mathfrak{f})^{N\mathcal{O}_K}/K(\mathfrak{f})) = \overline{C_N(\mathcal{O})}.$$

Thus we have the following result.

Theorem 1.4 (Stevenhagen). *Let $\mathcal{O}$ be an order in the imaginary quadratic field K, and let $N \in \mathbb{Z}^+$. Then the reduced mod N Galois representation*

$$\overline{\rho_N} : \mathfrak{g}_{K(\mathfrak{f})} \to \overline{C_N(\mathcal{O})}$$

is surjective and $\overline{\mathbb{Q}}^{\ker \overline{\rho_N}} = K(\mathfrak{f})^{N\mathcal{O}_K}$, the N-ray class field of $\mathcal{O}$.

We will give a new proof of Theorem 1.4, as follows. Let $E_{/K(\mathfrak{f})}$ be an $\mathcal{O}$-CM elliptic curve. Using the canonical isogeny $\iota_1 : E \to E_1$ to an $\mathcal{O}_K$-CM elliptic curve, we show that the torsion field $K(\mathfrak{f})(E[N])$ contains the ray class field $K^{N\mathcal{O}_K}$: Theorem 2.11(b). Using an observation of Parish we show that $K(\mathfrak{f})(E[N])$ contains the ring class field $K(N\mathfrak{f})$: Theorem 4.1. By Theorem 2.10(c), we have

$$K(\mathfrak{f})(\mathfrak{h}(E[N])) \supset K^{N\mathcal{O}_K} K(N\mathfrak{f}) = K(\mathfrak{f})^{N\mathcal{O}_K},$$

where the last equality can be shown using class field theory (Section 5A). Since

$$[K(\mathfrak{f})(\mathfrak{h}(E[N])) : K(\mathfrak{f})] \le \#\overline{C_N(\mathcal{O})} = [K(\mathfrak{f})^{N\mathcal{O}_K} : K(\mathfrak{f})],$$

we get

$$\overline{\mathbb{Q}}^{\ker \overline{\rho_N}} = K(\mathfrak{f})(\mathfrak{h}(E[N])) = K(\mathfrak{f})^{N\mathcal{O}_K}.$$

Theorem 1.4 has the following useful consequences:

Corollary 1.5. *For all number fields $F \supset K$ and all $\mathcal{O}$-CM elliptic curves $E_{/F}$ we have*

$$[C_N(\mathcal{O}) : \rho_N(\mathfrak{g}_F)] \mid \#\mathcal{O}^\times [F : K(\mathfrak{f})] \le 6[F : K(\mathfrak{f})].$$

Remark 1.6. Corollary 1.5 strengthens a result of D. Lombardo [2017, Theorem 1.5], who showed that $[C_N(\mathcal{O}) : \rho_N(\mathfrak{g}_F)] \le 6[F : K]$.

Corollary 1.7. *Let $N \in \mathbb{Z}^+$. There is a number field $F \supset K$ and an $\mathcal{O}$-CM elliptic curve $E_{/F}$ such that $E[N] = E[N](F)$ and $[F : K(j(E))] = \#\overline{C_N(\mathcal{O})}$.*

Corollary 1.8. *For all $N \in \mathbb{Z}^+$, there is an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ such that $\rho_{E,N}(\mathfrak{g}_{K(\mathfrak{f})}) = C_N(\mathcal{O})$.*

Using Theorems 1.2 and 1.4 we also obtain a complete classification of the set of $N \in \mathbb{Z}^+$ such that an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ admits a $K(\mathfrak{f})$-rational cyclic N-isogeny, and of the possible torsion subgroups of an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$. See Sections 6F and 6G.

2. Preliminaries

2A. *Foundations.* We begin by setting some terminology for orders in imaginary quadratic fields. Let K be an imaginary quadratic field, with ring of integers $\mathcal{O}_K$, and let $w_K := \#\mathcal{O}_K^\times$ be the number of roots of unity in K. Let $\mathcal{O}$ be a $\mathbb{Z}$-order in K. Let $\mathfrak{f} = [\mathcal{O}_K : \mathcal{O}]$ be the conductor of $\mathcal{O}$, and let Δ be the discriminant of $\mathcal{O}$. Then

$$\mathcal{O} = \mathbb{Z} + \mathfrak{f}\mathcal{O}_K, \quad \Delta = \mathfrak{f}^2 \Delta_K.$$

For fixed K and $\mathfrak{f} \in \mathbb{Z}^+$ there is a unique order $\mathcal{O}(\mathfrak{f})$ in K of conductor $\mathfrak{f}$. Thus an imaginary quadratic order is determined by its discriminant Δ, a negative integer which is 0 or 1 modulo 4. Conversely, for any negative integer Δ which is 0 or 1 modulo 4, we put

$$\tau_\Delta = \frac{\Delta + \sqrt{\Delta}}{2},$$

and then $\mathbb{Z}[\tau_\Delta]$ is an order in $K = \mathbb{Q}(\sqrt{\Delta})$ of discriminant Δ.

Throughout this paper we will use the following terminological convention: by "an order $\mathcal{O}$" we always mean a $\mathbb{Z}$-order $\mathcal{O}$ in an imaginary quadratic field, which is determined as the fraction field of $\mathcal{O}$ and denoted by K. We may specify an order $\mathcal{O}$ by giving its discriminant, which also determines K. If K is already given, then we specify an order $\mathcal{O}$ in K by giving the conductor $\mathfrak{f}$.

For any $\mathcal{O}$-CM elliptic curve E we have $K(j(E)) = K(\mathfrak{f})$, the ring class field of K of conductor $\mathfrak{f}$ ([Cox 1989, Theorem 11.1]) Thus $[K(j(E)) : K]$ is determined via the following formula.

Theorem 2.1. *For $N \in \mathbb{Z}^+$, let $K(N)$ denote the N-**ring class field** of K. Then $K(1) = K^{(1)}$ is the Hilbert class field of K, and for all $N \geq 2$ we have*

$$[K(N) : K^{(1)}] = \frac{2}{w_K} N \prod_{p \mid N} \left(1 - \left(\frac{\Delta_K}{p}\right)\frac{1}{p}\right).$$

Proof. See, e.g., [Cox 1989, Corollary 7.24]. □

For a number field F, $N \in \mathbb{Z}^+$ and $E_{/F}$ an elliptic curve, we denote by ρ_N the homomorphism

$$\mathfrak{g}_F \to \operatorname{Aut} E[N] \cong \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}),$$

the *modulo N Galois representation*. If $E_{/F}$ has CM by the order $\mathcal{O}$ in K, then $E[N] \cong_{\mathcal{O}} \mathcal{O}/N\mathcal{O}$ — i.e., we have an isomorphism of $\mathcal{O}$-modules (see [Parish 1989, Lemma 1], generalized in Lemma 2.4 below) — and provided $F \supset K$ we have

$$\rho_N : \mathfrak{g}_F \hookrightarrow \operatorname{Aut}_{\mathcal{O}} E[N] \cong \mathrm{GL}_1(\mathcal{O}/N\mathcal{O}) = (\mathcal{O}/N\mathcal{O})^\times.$$

In other words, the image of the mod N Galois representation lands in the *mod N Cartan subgroup*

$$C_N(\mathcal{O}) = (\mathcal{O}/N\mathcal{O})^\times.$$

Lemma 2.2. *Let $\mathcal{O}$ be an order of discriminant Δ, and let $N = p_1^{a_r} \cdots p_r^{a_r} \in \mathbb{Z}^+$.*

(a) *We have $C_N(\mathcal{O}) = \prod_{i=1}^r C_{p_i^{a_i}}(\mathcal{O})$ (canonical isomorphism).*

(b) *We have $\#C_N(\mathcal{O}) = N^2 \prod_{p \mid N} \left(1 - \left(\frac{\Delta}{p}\right)\frac{1}{p}\right)\left(1 - \frac{1}{p}\right)$.*

Proof. (a) It suffices to tensor the Chinese remainder theorem isomorphism $\mathbb{Z}/N\mathbb{Z} = \prod_{i=1}^r \mathbb{Z}/p_i^{a_i}\mathbb{Z}$ with the $\mathbb{Z}$-module $\mathcal{O}$ and pass to the unit groups.

(b) By [Clark et al. 2013], for any prime number p we have

$$\#C_p(\mathcal{O}) = p^2 \left(1 - \left(\frac{\Delta}{p}\right)\frac{1}{p}\right)\left(1 - \frac{1}{p}\right).$$

The natural map $C_{p^a}(\mathcal{O}) \to C_p(\mathcal{O})$ is surjective with kernel of size p^{2a-2} [Clark and Pollack 2015, p. 3]. Together with part (a) this shows that if $N = p_1^{a_1} \cdots p_r^{a_r}$ then

$$\#C_N(\mathcal{O}) = \prod_{i=1}^r p_i^{2a_i - 2}(p_i - 1)\left(p_i - \left(\frac{\Delta}{p_i}\right)\right) = N^2 \prod_{p \mid N} \left(1 - \left(\frac{\Delta}{p}\right)\frac{1}{p}\right)\left(1 - \frac{1}{p}\right). \quad \square$$

2B. ***Torsion kernels.*** Let $E_{/\mathbb{C}}$ be an $\mathcal{O}$-CM elliptic curve. For a nonzero ideal I of $\mathcal{O}$, we define the *I-torsion kernel*

$$E[I] := \{P \in E \mid \text{for all } \alpha \in I,\ \alpha P = 0\}.$$

There is an invertible ideal $\Lambda \subset \mathcal{O}$ such that

$$E \cong \mathbb{C}/\Lambda.$$

If we put

$$(\Lambda : I) := \{x \in \mathbb{C} \mid xI \subset \Lambda\} = \{x \in K \mid xI \subset \Lambda\}$$

then we have (immediately) that

$$E[I] = \{x \in \mathbb{C}/\Lambda \mid xI \subset \Lambda\} = (\Lambda : I)/\Lambda.$$

Let $|I| := \#\mathcal{O}/I$.

Lemma 2.3. *Let $I, J \subset \mathcal{O}$ be nonzero ideals and $E_{/\mathbb{C}}$ be an $\mathcal{O}$-CM elliptic curve.*

(a) *If $I \subset J$, then $E[J] \subset E[I]$.*

(b) *We have $E[I] \subset E[|I|]$. In particular,*

$$\#E[I] \leq |I|^2.$$

Proof. (a) This is immediate from the definition.

(b) By Lagrange's theorem, every element of $\mathcal{O}/I$ is killed by $|I|$, so $|I| \in |I|\mathcal{O} \subset I$. Apply part (a). □

Lemma 2.4. *If I is an invertible $\mathcal{O}$-ideal, then*

$$E[I] = I^{-1}\Lambda/\Lambda \cong_{\mathcal{O}} \mathcal{O}/I.$$

In particular, $\#E[I] = |I| = \#\mathcal{O}/I$.

Proof. An ideal I is invertible if and only if there is an $\mathcal{O}$-submodule I^{-1} of K such that $II^{-1} = \mathcal{O}$. If so, then for $x \in K$ we have

$$xI \subset \Lambda \iff xII^{-1} = x\mathcal{O} \subset I^{-1}\Lambda \iff x \in I^{-1}\Lambda,$$

giving $E[I] = I^{-1}\Lambda/\Lambda$. Because Λ is a locally free $\mathcal{O}$-module, for all $\mathfrak{p} \in \operatorname{Spec} \mathcal{O}$ we have $\Lambda_{\mathfrak{p}} \cong \mathcal{O}_{\mathfrak{p}}$ and thus $(I^{-1}\Lambda/\Lambda)_{\mathfrak{p}} \cong (I^{-1}/\mathcal{O})_{\mathfrak{p}} \cong (\mathcal{O}/I)_{\mathfrak{p}}$. Thus $I^{-1}\Lambda/\Lambda$ is locally free of rank one as an $\mathcal{O}/I$-module. But the ring $\mathcal{O}/I$ is semilocal, and hence has trivial Picard group: any locally free rank one $\mathcal{O}/I$-module is isomorphic to $\mathcal{O}/I$ [Clark 2015, Corollary 13.38]. □

Lemma 2.5. *Let R be a Dedekind domain, and let M be a cyclic torsion R-module, and let $N \subset M$ be an R-submodule. Then:*

(a) *N is also a cyclic R-module.*

(b) *We have $N \cong R/\operatorname{ann} N$.*

Proof. Let $I = \operatorname{ann} M$. Since M is a finitely generated torsion module over a domain, we have $I \neq 0$ and $M \cong R/I$. Thus $N \cong I'/I$ for some ideal $I' \supset I$. The ring R/I

is principal Artinian [Clark 2015, Theorem 20.11], so the ideal I'/I of R/I is principal. Thus N is a cyclic, torsion R-module, so $N \cong R/\operatorname{ann} N$. □

Theorem 2.6. *Let $E_{/\mathbb{C}}$ be an $\mathcal{O}_K$-CM elliptic curve, and let $M \subset E(\mathbb{C})$ be a finite $\mathcal{O}_K$-submodule. Then $M = E[\operatorname{ann} M] \cong_{\mathcal{O}_K} \mathcal{O}_K/\operatorname{ann} M$ and thus $\#M = |\operatorname{ann} M|$.*

Proof. That $M \subset E[\operatorname{ann} M]$ is a tautology. Because $\mathcal{O}_K$ is a Dedekind domain, every nonzero $\mathcal{O}_K$-ideal is invertible, so Lemma 2.4 gives $\#E[\operatorname{ann} M] = |\operatorname{ann} M|$. On the other hand, let $\mathfrak{t} = \#M$. Then $M \subset E[\mathfrak{t}] \cong_{\mathcal{O}_K} \mathcal{O}_K/\mathfrak{t}\mathcal{O}_K$, a finite cyclic $\mathcal{O}_K$-module. By Lemma 2.5 we have $M \cong \mathcal{O}_K/\operatorname{ann} M$ so $\#M = |\operatorname{ann} M|$. Thus $M = E[\operatorname{ann} M]$, so Lemma 2.4 gives $M \cong \mathcal{O}_K/\operatorname{ann} M$ and $\#M = |\operatorname{ann} M|$. □

Example 2.7. Theorem 2.6 fails for all nonmaximal orders. Indeed, let $\mathcal{O}$ be a nonmaximal order in an imaginary quadratic field K. There is nonzero prime ideal $\mathfrak{p}$ of $\mathcal{O}$ such that the local ring $\mathcal{O}_{\mathfrak{p}}$ is not a DVR. If $\mathfrak{p} \cap \mathbb{Z} = (\ell)$, then $\mathcal{O}/\mathfrak{p} \cong \mathbb{Z}/\ell\mathbb{Z}$. Since every ideal of $\mathcal{O}$ can be generated by two elements, we have $\dim_{\mathcal{O}/\mathfrak{p}} \mathfrak{p}/\mathfrak{p}^2 = 2$. Thus $\#\mathcal{O}/\mathfrak{p}^2 = \ell^3$ and $(\ell^3) \subset \mathfrak{p}^2$. It follows that in the quotient ring $\mathcal{O}/\ell^3\mathcal{O}$, the maximal ideal $\mathfrak{p} + \ell^3\mathcal{O}$ is not principal. Let $E_{/\mathbb{C}}$ be an $\mathcal{O}$-CM elliptic curve, so $E[\ell^3] \cong_{\mathcal{O}} \mathcal{O}/\ell^3\mathcal{O}$. So the $\mathcal{O}$-submodule $M = \mathfrak{p}E[\ell^3]$ of $E[\ell^3]$ is not cyclic and thus not isomorphic to $\mathcal{O}/\operatorname{ann} M$.

For a nonzero ideal I of $\mathcal{O}_K$, let K^I denote the *I-ray class field* of K.

Theorem 2.8 (first main theorem of complex multiplication). *Let $E_{/\mathbb{C}}$ be an $\mathcal{O}_K$-CM elliptic curve, and let I be a nonzero ideal of $\mathcal{O}_K$. Let $\mathfrak{h} : E \to \mathbb{P}^1$ be a Weber function. Then*:

$$K^{(1)}(\mathfrak{h}(E[I])) = K^I.$$

Proof. See, e.g., [Silverman 1994, Theorem II.5.6]. □

Combining Theorems 2.6 and 2.8, we get the class-field theoretic containment corresponding to any finite $\mathcal{O}_K$-submodule of $E(\bar{F})$, for any $\mathcal{O}_K$-CM elliptic curve E defined over a number field $F \supset K$.

For convenience, we record here the formulas for $[K^I : K^{(1)}]$. Here, φ denotes Euler's totient function and $\varphi_K(I)$ the natural generalization for a nonzero ideal I of $\mathcal{O}_K$. That is,

$$\varphi_K(I) = \#(\mathcal{O}_K/I)^\times = |I| \prod_{\mathfrak{p} \mid I} \left(1 - \frac{1}{|\mathfrak{p}|}\right).$$

Lemma 2.9. *Let I be a nonzero ideal of K. We put $U_I(K) = \{x \in U(K) \mid x - 1 \in I\}$ and $U(K) = \mathcal{O}_K^\times$.*

(a) *We have*

$$[K^I : K^{(1)}] = \frac{\varphi_K(I)}{[U(K) : U_I(K)]}.$$

(b) *If* $K \neq \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$, *then*

$$[K^I : K^{(1)}] = \begin{cases} \varphi_K(I), & I \mid (2), \\ \frac{\varphi_K(I)}{2}, & I \nmid (2). \end{cases}$$

(c) *If* $K = \mathbb{Q}(\sqrt{-1})$, *then*

$$[K^I : K^{(1)}] = \begin{cases} \varphi_K(I), & I \mid (1+i), \\ \frac{\varphi_K(I)}{2}, & I \nmid (1+i) \text{ and } I \mid (2), \\ \frac{\varphi_K(I)}{4}, & I \nmid (2). \end{cases}$$

(d) *If* $K = \mathbb{Q}(\sqrt{-3})$, *then*

$$[K^I : K^{(1)}] = \begin{cases} 1, & I = (1), \\ \frac{\varphi_K(I)}{2}, & I \neq (1) \text{ and } I \mid (\zeta_3 - 1), \\ \frac{\varphi_K(I)}{3}, & I = (2), \\ \frac{\varphi_K(I)}{6}, & \text{otherwise}. \end{cases}$$

Proof. Parts (b)–(d) can be deduced from (a), which appears as [Cohen 2000, Corollary 3.2.4]. □

2C. ***On Weber functions.*** Let F be a field of characteristic 0. For an elliptic curve $E_{/F}$, let $\mathfrak{h} : E \to \mathbb{P}^1$ be a Weber function for E (cf. Section 1C).

Theorem 2.10 (Weber function principle). *Let* $N \in \mathbb{Z}^{\geq 2}$, *let* $\mathcal{O}$ *be the order of conductor* $\mathfrak{f}$ *in* K, *and let* $F = K(\mathfrak{f})$. *For an* $\mathcal{O}$*-CM elliptic curve* $E_{/F}$, *fix an embedding* $F \hookrightarrow \mathbb{C}$ *such that* $j(E) = j(\mathbb{C}/\mathcal{O})$. *Define*

$$W(N, \mathcal{O}) = K(\mathfrak{f})(\mathfrak{h}(E[N])).$$

(a) $W(N, \mathcal{O})$ *is a subfield of* $F(E[N])$ *and* $[F(E[N]) : W(N, \mathcal{O})]$ *divides*

$$\begin{cases} \#\mathcal{O}^\times, & N \geq 3, \\ \frac{\#\mathcal{O}^\times}{2}, & N = 2. \end{cases}$$

(b) *There is an elliptic curve* $E_{/F}$ *such that*

$$[F(E[N]) : W(N, \mathcal{O})] = \begin{cases} \#\mathcal{O}^\times, & N \geq 3, \\ \frac{\#\mathcal{O}^\times}{2}, & N = 2. \end{cases}$$

(c) *As we range over all elliptic curves* $E_{/F}$ *with* $j(E) = j(\mathbb{C}/\mathcal{O})$, *we have*

$$\bigcap_E F(E[N]) = W(N, \mathcal{O}).$$

Proof. (a) Let

$$w = \begin{cases} \#\mathcal{O}^\times, & N \geq 3, \\ \frac{\#\mathcal{O}^\times}{2}, & N = 2. \end{cases}$$

Let μ_w be the image of $\mathcal{O}^\times \to C_N(\mathcal{O})$, a cyclic group of order w. The field $F(E[N])/F$ is Galois with Galois group $\rho_N(\mathfrak{g}_F) \subset C_N(\mathcal{O})$. Since $\mathfrak{h}(P) = \mathfrak{h}(Q)$ for points P, Q on E if and only if there is $\xi \in \mathcal{O}^\times$ such that $\xi(P) = Q$ (e.g., [Lang 1987, Theorem I.7]), it follows that

$$W(N, \mathcal{O}) = F(E[N])^{\rho_N(\mathfrak{g}_F)\cap\mu_w}.$$

Thus

$$[F(E[N]) : W(N, \mathcal{O})] \mid w.$$

(b) and (c) If $E_{/F}$, $E'_{/F}$ with $j(E) = j(E')$, then $K(\mathfrak{f})(\mathfrak{h}(E[N])) = K(\mathfrak{f})(\mathfrak{h}(E'[N]))$ by the model independence of the Weber function. So $W(N, \mathcal{O}) \subset \bigcap_E F(E[N])$. To see that equality holds, let $E_{/F}$ have $j(E) = j(\mathbb{C}/\mathcal{O})$. Let $\mathfrak{p}$ be a prime of $\mathcal{O}_F$ that is unramified in $F' = F(E[N])$. By weak approximation, there is $\pi \in \mathfrak{p} \setminus \mathfrak{p}^2$. Put $L = F(\pi^{1/w})$, and let $\chi : \mathfrak{g}_F \to \mu_w$ be a character with splitting field $\overline{F}^{\ker\chi} = L$. (Explicitly, we may take $\chi(\sigma) = \sigma(\pi^{1/w})/\pi^{1/w}$.) Then L/F is totally ramified over $\mathfrak{p}$, so F' and L are linearly disjoint over F. It follows that

$$\rho_{N,E^\chi}(\mathfrak{g}_{F'}) = (\rho_{N,E_{/F'}} \otimes \chi)(\mathfrak{g}_{F'}) = \chi(\mathfrak{g}_{F'}) = \mu_w.$$

Thus

$$w = [F(E^\chi[N]) : F(E[N]) \cap F(E^\chi[N])] \mid [F(E^\chi[N]) : W(N, \mathcal{O})] \mid w,$$

so $F(E^\chi[N])$ has degree w over $W(N, \mathcal{O}) = F(E[N]) \cap F(E^\chi[N])$. □

2D. ***A containment of Weber function fields.***

Theorem 2.11. *Let K be an imaginary quadratic field, and let $\mathcal{O} \subset \mathcal{O}'$ be orders in K.*

(a) *For all $N \in \mathbb{Z}^+$ we have a containment of Weber function fields $W(N, \mathcal{O}) \supset W(N, \mathcal{O}')$.*

(b) *If E is a K-CM elliptic curve defined over a number field $F \supset K$, then we have*

$$F(\mathfrak{h}(E[N])) \supset K^{N\mathcal{O}_K}. \tag{2}$$

Proof. (a) Let $\mathfrak{f}$ (resp. $\mathfrak{f}'$) be the conductor of $\mathcal{O}$ (resp. of $\mathcal{O}'$), so $\mathfrak{f}' \mid \mathfrak{f}$. Let $E_{/K(\mathfrak{f})}$ be an $\mathcal{O}$-CM elliptic curve. Let $\iota = \iota_{\mathfrak{f},\mathfrak{f}'} : E \to E'$ be the canonical cyclic $\mathfrak{f}/\mathfrak{f}'$-isogeny to an $\mathcal{O}'$-CM elliptic curve $(E')_{/K(\mathfrak{f})}$. There is an embedding $K(\mathfrak{f}) \hookrightarrow \mathbb{C}$ such that $E(\mathbb{C}) \cong E/\mathcal{O}$ and $(E')(\mathbb{C}) \cong E/\mathcal{O}'$ and $\iota : E/\mathcal{O} \to E/\mathcal{O}'$ is the natural map. Then ι maps $\frac{1}{N} + \mathcal{O}$ to $\frac{1}{N} + \mathcal{O}'$, which generates $E'[N]$ as an $\mathcal{O}'$-module. Thus

$$K(\mathfrak{f})(E[N]) \supset K(\mathfrak{f})(E'[N]) \supset K(\mathfrak{f}')(E'[N]) \supset W(N, \mathcal{O}'),$$

and Theorem 2.10(c) gives

$$W(N,\mathcal{O}) \supset W(N,\mathcal{O}').$$

(b) This follows from part (a) and Theorem 2.8. □

Though Theorem 2.11 is a consequence of Theorem 1.4, we prove it here as an ingredient in our new proof of Theorem 1.4. Part (b) had been proved earlier in [Bourdon et al. 2017b, Theorem 3.16], but the present argument seems more transparent.

3. Proof of the Isogeny Torsion Theorem

For a quadratic field K and $d \in \mathbb{Z}^+$ we will write $\mathcal{O}(d)$ for the unique order in K of conductor d. We recall the setup: let $F \supset K$ be a number field, let $E_{/F}$ be an elliptic curve with endomorphism ring an order $\mathcal{O}$ of conductor $\mathfrak{f}$ in K, and let $\mathfrak{f}'$ be a positive integer that divides $\mathfrak{f}$. Then there is a canonical F-rational cyclic $\mathfrak{f}/\mathfrak{f}'$-isogeny $\iota_{/\mathfrak{f}'} : E \to E_{\mathfrak{f}'}$ such that $\operatorname{End} E_{\mathfrak{f}'} = \mathcal{O}(\mathfrak{f}')$.

There is a field embedding $F \hookrightarrow \mathbb{C}$ such that $E(\mathbb{C}) \cong \mathbb{C}/\mathcal{O}(\mathfrak{f})$, $E_{\mathfrak{f}'}(\mathbb{C}) \cong \mathbb{C}/\mathcal{O}(\mathfrak{f}')$ and $(\iota_{\mathfrak{f}'})_{/\mathbb{C}}$ is the quotient map $\mathbb{C}/\mathcal{O}(\mathfrak{f}) \to \mathbb{C}/\mathcal{O}(\mathfrak{f}')$. Put $\tau_K = \frac{1}{2}(\Delta_K + \sqrt{\Delta_K})$, so $\mathcal{O}(\mathfrak{f}) = \mathbb{Z}[\mathfrak{f}\tau_K]$ and $\mathcal{O}(\mathfrak{f}') = \mathbb{Z}[\mathfrak{f}'\tau_K]$. For $N \in \mathbb{Z}^+$, let

$$e_{1,\mathfrak{f}}(N) := \frac{1}{N} + \mathcal{O}(\mathfrak{f}), \qquad e_{2,\mathfrak{f}}(N) := \frac{\mathfrak{f}\tau_K}{N} + \mathcal{O}(\mathfrak{f}),$$
$$e_{1,\mathfrak{f}'}(N) := \frac{1}{N} + \mathcal{O}(\mathfrak{f}'), \qquad e_{2,\mathfrak{f}'}(N) := \frac{\mathfrak{f}'\tau_K}{N} + \mathcal{O}(\mathfrak{f}').$$

Then $\ker(E[N] \xrightarrow{\iota_{\mathfrak{f}'}} E_{\mathfrak{f}'}[N])$ is cyclic of order $\gcd(N, \mathfrak{f}/\mathfrak{f}')$, generated by

$$\frac{N}{\gcd\big(N, \frac{\mathfrak{f}}{\mathfrak{f}'}\big)} e_{2,\mathfrak{f}}(N),$$

and $\iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(N)) = e_{1,\mathfrak{f}'}(N)$.

For finite commutative groups T_1 and T_2, we have $\#T_1 \mid \#T_2$ if and only if $\#T_1[\ell^\infty] \mid \#T_2[\ell^\infty]$ for all prime numbers ℓ. So it suffices to show that for all prime numbers ℓ, we have $\#E(F)[\ell^\infty] \mid \#E_{\mathfrak{f}'}(F)[\ell^\infty]$. Write $\mathfrak{f} = \ell^{c_1}\overline{\mathfrak{f}}$ with $\gcd(\overline{\mathfrak{f}}, \ell) = 1$ and $\mathfrak{f}' = \ell^{c_2}\overline{\mathfrak{f}'}$ with $\gcd(\overline{\mathfrak{f}'}, \ell) = 1$. Then we have

$$\#E(F)[\ell^\infty] = \#E_{\ell^{c_1}}(F)[\ell^\infty], \quad \#E_{\mathfrak{f}'}(F)[\ell^\infty] = \#E_{\ell^{c_2}}(F)[\ell^\infty],$$

so we may assume that $\mathfrak{f} = \ell^{c_1}$ and $\mathfrak{f}' = \ell^{c_2}$. Indeed, it is enough to treat the case $c_2 = c_1 - 1$, since repeated application of this case yields the general case. So suppose $\mathfrak{f} = \ell^c$ for some $c \in \mathbb{Z}^+$ and $\mathfrak{f}' = \ell^{c-1}$. By, for example, the Mordell–Weil theorem, there are integers $0 \le a \le b$ such that

$$E(F)[\ell^\infty] \cong \mathbb{Z}/\ell^a\mathbb{Z} \oplus \mathbb{Z}/\ell^b\mathbb{Z}.$$

Let $Q := \frac{1}{\ell^a} + \mathcal{O}(\mathfrak{f}) \in E(F)$. Since $\iota_{\mathfrak{f}'}$ is F-rational, $Q' := \iota_{\mathfrak{f}'}(Q) = \frac{1}{\ell^a} + \mathcal{O}(\mathfrak{f}')$ lies in $E_{\mathfrak{f}'}(F)$ and generates $E_{\mathfrak{f}'}[\ell^a]$ as an $\mathcal{O}(\mathfrak{f}')$-module, so $E_{\mathfrak{f}'}[\ell^a] = E_{\mathfrak{f}'}(F)[\ell^a]$. If $a = b$, it follows that $\#E(F)[\ell^\infty] \mid \#E_{\mathfrak{f}'}(F)[\ell^\infty]$, so we may assume $b > a$. Since $\ker(E[\ell^\infty] \xrightarrow{\iota_{\mathfrak{f}'}} E_{\mathfrak{f}'}[\ell^\infty])$ has order ℓ, we have $\mathbb{Z}/\ell^a\mathbb{Z} \oplus \mathbb{Z}/\ell^{b-1}\mathbb{Z} \hookrightarrow E_{\mathfrak{f}'}(F)[\ell^\infty]$. Thus it suffices to show that $E_{\mathfrak{f}'}(F)$ has either a point of order ℓ^b or has full ℓ^{a+1}-torsion.

Let $P = E(F)$ be a point of order ℓ^b, and write $P = \alpha e_{1,\mathfrak{f}}(\ell^b) + \beta e_{2,\mathfrak{f}}(\ell^b)$ with $\alpha, \beta \in \mathbb{Z}/\ell^b\mathbb{Z}$. If $\ell \nmid \alpha$ then since $\mathfrak{f} = \ell\mathfrak{f}'$ we have that $\iota_{\mathfrak{f}'}(P) = \alpha e_{1,\mathfrak{f}'}(\ell^b) + \ell\beta e_{2,\mathfrak{f}'}(\ell^b)$ has order ℓ^b and we are done, so we may assume that $\ell \mid \alpha$, in which case $\ell \nmid \beta$. With respect to the basis $e_{1,\mathfrak{f}}(\ell^b)$, $e_{2,\mathfrak{f}}(\ell^b)$ of $E[\ell^b]$, the image of the mod ℓ^b Galois representation on E consists of matrices of the form

$$\begin{bmatrix} a & b\ell^{2c}\frac{\Delta_K - \Delta_K^2}{4} \\ b & a + b\ell^c\Delta_K \end{bmatrix} \text{ with } a, b \in \mathbb{Z}/\ell^b\mathbb{Z}. \tag{3}$$

Since $E(F)$ has full ℓ^a-torsion, we have $a \equiv 1 \pmod{\ell^a}$ and $b \equiv 0 \pmod{\ell^a}$. Thus

$$\rho_{\ell^{a+1}}(\mathfrak{g}_F) \subset \left\{ \begin{bmatrix} 1 + \ell^a A & 0 \\ \ell^a B & 1 + \ell^a A \end{bmatrix} \;\middle|\; A, B \in \mathbb{Z}/\ell^{a+1}\mathbb{Z} \right\}.$$

Since $\ell^{b-a-1}P = \alpha e_{1,\mathfrak{f}}(\ell^{a+1}) + \beta e_{2,\mathfrak{f}}(\ell^{a+1})$ is F-rational, all such matrices in the image of Galois satisfy

$$\begin{bmatrix} 1 + \ell^a A & 0 \\ \ell^a B & 1 + \ell^a A \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

and thus $\ell^a\alpha B + \beta + \ell^a A\beta \equiv \beta \pmod{\ell^{a+1}}$. Since $\ell \mid \alpha$, we get

$$\ell^a A\beta \equiv -\ell^a \alpha B \equiv 0 \pmod{\ell^{a+1}},$$

and thus $\ell \mid A$ and $\rho_{\ell^{a+1}}(\mathfrak{g}_F)$ consists of matrices of the form

$$\begin{bmatrix} 1 & 0 \\ \ell^a B & 1 \end{bmatrix}$$

for $B \in \mathbb{Z}/\ell^{a+1}\mathbb{Z}$. It follows that for all $\sigma \in \mathfrak{g}_F$, there is $B \in \mathbb{Z}/\ell^{a+1}\mathbb{Z}$ such that

$$\begin{aligned} \sigma(\iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(\ell^{a+1}))) &= \iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(\ell^{a+1}) + B\ell^a \iota_{\mathfrak{f}'}(e_{2,\mathfrak{f}}(\ell^{a+1})) \\ &= \iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(\ell^{a+1})) + B\ell^a(\ell e_{2,\mathfrak{f}'}(\ell^{a+1})) = \iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(\ell^{a+1})). \end{aligned}$$

Thus $e_{1,\mathfrak{f}'}(\ell^{a+1}) = \iota_{\mathfrak{f}'}(e_{1,\mathfrak{f}}(\ell^{a+1})) \in E_{\mathfrak{f}'}(F)$. Since the $\mathcal{O}(\mathfrak{f}')$-submodule generated by $e_{1,\mathfrak{f}'}(\ell^{a+1})$ is $E_{\mathfrak{f}'}[\ell^{a+1}]$, we get $\mathbb{Z}/\ell^{a+1}\mathbb{Z} \oplus \mathbb{Z}/\ell^{a+1}\mathbb{Z} \hookrightarrow E_{\mathfrak{f}'}(F)$, completing the proof of Theorem 1.3.

4. The projective torsion point field

Let F be a field. For a positive integer N not divisible by the characteristic of F and $E_{/F}$ an elliptic curve, we define the *projective modulo N Galois representation* as the composite map

$$\mathbb{P}\rho_N : \mathfrak{g}_F \xrightarrow{\rho_N} \operatorname{Aut} E[N] \cong \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) \to \mathrm{PGL}_2(\mathbb{Z}/N\mathbb{Z}) := \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/(\mathbb{Z}/N\mathbb{Z})^\times.$$

The *projective torsion field* is

$$F(\mathbb{P}E[N]) = \overline{F}^{\ker \mathbb{P}\rho_N}.$$

Thus $F(\mathbb{P}E[N])$ is the unique minimal field extension of F on which the image of ρ_N consists of scalar matrices. It follows that $F(E[N])/F(\mathbb{P}E[N])$ is a Galois extension with automorphism group a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$.

Observe that the projective Galois representation and thus the projective torsion field are unchanged by *quadratic* twists. If $E_{/F}$ has CM by an order of discriminant $\Delta = \mathfrak{f}^2\Delta_K \neq -3, -4$ and $F \supset K$, then the projective N-torsion field is a well-defined abelian extension of $K(\mathfrak{f})$. When $\Delta = -4$ (resp. $\Delta = -3$) we have quartic twists (resp. sextic twists) which can change the projective Galois representation and the projective torsion field.

Theorem 4.1. *Let $\mathcal{O}$ be an order of discriminant $\Delta = \mathfrak{f}^2\Delta_K$. Let E be an $\mathcal{O}$-CM elliptic curve defined over $F = K(\mathfrak{f})$. Let $N \geq 2$.*

(a) ([Parish 1989]) *We have $F(\mathbb{P}E[N]) \supset K(N\mathfrak{f})$. Thus we may put*

$$d(E, N) = [F(\mathbb{P}E[N]) : K(N\mathfrak{f})].$$

(b) ([Parish 1989]) *If $\Delta \notin \{-3, -4\}$, then $d(E, N) = 1$, i.e., $F(\mathbb{P}E[N]) = K(N\mathfrak{f})$.*

(c) *If $\Delta = -4$, then $d(E, N) \mid 2$.*

(d) *If $\Delta = -3$, then $d(E, N) \mid 3$.*

Proof. For $N \in \mathbb{Z}^+$, let $\mathcal{O}(N)$ be the order of conductor N in K. Thus $\mathcal{O} = \mathcal{O}(\mathfrak{f})$.

Step 1: We show that $F(\mathbb{P}E[N]) \supset K(N\mathfrak{f})$ in all cases.

There is a field embedding $F \hookrightarrow \mathbb{C}$ such that $E(\mathbb{C}) \cong \mathbb{C}/\mathcal{O}$. The $\mathbb{C}$-linear map $z \mapsto Nz$ carries $\mathcal{O}(\mathfrak{f})$ into $\mathcal{O}(N\mathfrak{f})$ and induces a cyclic N-isogeny $\mathbb{C}/\mathcal{O}(\mathfrak{f}) \to \mathbb{C}/\mathcal{O}(N\mathfrak{f})$. Let C be the kernel of this isogeny, viewed as a finite étale subgroup scheme of $E_{/\mathbb{C}}$. Then C has a (unique) minimal field of definition $F(C) \subset F(E[N])$, hence of finite degree over F. The field $F(\mathbb{P}E[N])$ is precisely the compositum of the minimal fields of definition of all order N cyclic subgroup schemes $C \subset E_{/\mathbb{C}}$, so $F(C) \subset F(\mathbb{P}E[N])$. Since C is $F(\mathbb{P}E[N])$-rational, the elliptic curve E/C has a model over this field, and thus

$$F(\mathbb{P}E[N]) \supset K(j(E/C)) = K(N\mathfrak{f}).$$

Step 2: In view of Step 1, we have $F(\mathbb{P}E[N]) \supset K(N\mathfrak{f}) \supset K(\mathfrak{f}) = K(j(E))$, so we have $F(\mathbb{P}E[N]) = K(N\mathfrak{f})$ if and only if $[F(\mathbb{P}E[N]) : K(\mathfrak{f})] \le [K(N\mathfrak{f}) : K(\mathfrak{f})]$. We have

$$[F(\mathbb{P}E[N]) : K(\mathfrak{f})] = \#\mathbb{P}\rho_N(\mathfrak{g}_F) \le \#(\mathcal{O}/N\mathcal{O})^\times/(\mathbb{Z}/N\mathbb{Z})^\times = N\prod_{p\mid N}\left(1-\left(\frac{\Delta}{p}\right)\frac{1}{p}\right).$$

- Suppose $\mathfrak{f} > 1$. Using Theorem 2.1 to compute $[K(N\mathfrak{f}) : K^{(1)}]$ and $[K(\mathfrak{f}) : K^{(1)}]$ gives

$$[K(N\mathfrak{f}) : K(\mathfrak{f})] = \frac{[K(N\mathfrak{f}) : K^{(1)}]}{[K(\mathfrak{f}) : K^{(1)}]} = N\prod_{p\mid N,\ p\nmid\mathfrak{f}}\left(1-\left(\frac{\Delta}{p}\right)\frac{1}{p}\right) = N\prod_{p\mid N}\left(1-\left(\frac{\Delta}{p}\right)\frac{1}{p}\right),$$

  because $1-\left(\frac{\Delta}{p}\right)\frac{1}{p} = 1$ for all $p \mid \mathfrak{f}$. Thus $d(E,N) = 1$ in this case.
- Suppose $\mathfrak{f} = 1$, so $\Delta = \Delta_K$. Then

$$[K(N\mathfrak{f}) : K(\mathfrak{f})] = [K(N) : K^{(1)}] = \frac{2}{w_K}N\prod_{p\mid N}\left(1-\left(\frac{\Delta}{p}\right)\frac{1}{p}\right).$$

If $\Delta \notin \{-3,-4\}$ then $\frac{2}{w_K} = 1$, and again we get $d(E,N) = 1$. If $\Delta = -4$ then $\frac{2}{w_K} = \frac{1}{2}$, so the calculation shows $d(E,N) \in \{1,2\}$, and if $\Delta = -3$ then $\frac{2}{w_K} = \frac{1}{3}$, so the calculation shows $d(E,N) \in \{1,3\}$. □

Remark 4.2. (a) Theorem 4.1(a) and (b) are due to Parish [1989, Proposition 3]. However, he alludes to a calculation of the above sort rather than explicitly carrying it out. Since Theorem 4.1 will play an important role in the proof of Theorem 1.4, we have given a complete proof.

(b) Parish [1989, Proposition 3] assumes $K \ne \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$. Later [1989, p. 263], he claims
  - if $\Delta = -4$ then $F(\mathbb{P}E[N]) = K(N)$ for all $N \ge 3$, and
  - if $\Delta = -3$ then $F(\mathbb{P}E[N]) = K(N)$ for all $N \ge 4$.

  As we will see shortly in Example 4.4, both claims are false.

The following result is an analogue of [Bourdon et al. 2017b, Theorem 5.6] for higher twists.

Proposition 4.3 (higher twisting at the bottom). *For $M \in \mathbb{Z}^+$, we denote the mod M cyclotomic character by χ_M.*

(a) *Let $K = \mathbb{Q}(\sqrt{-1})$ and let $\ell \equiv 5 \pmod 8$ be a prime number. There is a character $\Psi : \mathfrak{g}_K \to (\mathbb{Z}/\ell\mathbb{Z})^\times$ of order $\frac{\ell-1}{4}$ and an $\mathcal{O}_K$-CM elliptic curve $E_{/K}$ such that the mod ℓ Galois representation is*

$$\sigma \mapsto \rho_\ell(\sigma) = \begin{bmatrix}\Psi(\sigma) & 0\\ 0 & \Psi^{-1}(\sigma)\chi_\ell(\sigma)\end{bmatrix}.$$

(b) *Let* $K = \mathbb{Q}(\sqrt{-3})$ *and let* $\ell \equiv 7, 31 \pmod{36}$ *be a prime number. There is a character* $\Psi : \mathfrak{g}_K \to (\mathbb{Z}/\ell\mathbb{Z})^\times$ *of order* $\frac{\ell-1}{6}$ *and an* $\mathcal{O}_K$*-CM elliptic curve* $E_{/K}$ *such that the mod* ℓ *Galois representation is*

$$\sigma \mapsto \rho_\ell(\sigma) = \begin{bmatrix} \Psi(\sigma) & 0 \\ 0 & \Psi^{-1}(\sigma)\chi_\ell(\sigma) \end{bmatrix}.$$

Proof. (a) Let $(E_1)_{/K}$ be an $\mathcal{O}_K$-CM elliptic curve. Because $\ell \equiv 1 \pmod 4$, the Cartan subgroup $C_\ell(\mathcal{O}_K)$ is split. It follows that there are precisely two $C_\ell(\mathcal{O}_K)$-stable one-dimensional $\mathbb{Z}/\ell\mathbb{Z}$-subspaces of $E_1[\ell]$, so we may take basis vectors e_1 and e_2 for $E_1[\ell]$ lying in these two subspaces. For such a basis, the mod ℓ Galois representation has the form

$$\sigma \mapsto \rho_\ell(\sigma) = \begin{bmatrix} \Psi_1(\sigma) & 0 \\ 0 & \Psi_1^{-1}(\sigma)\chi_\ell(\sigma) \end{bmatrix}$$

for a character $\Psi_1 : \mathfrak{g}_K \to (\mathbb{Z}/\ell\mathbb{Z})^\times$. Under this isomorphism, the matrix representation of $i \in \mathcal{O}_K$ is a diagonal matrix

$$\begin{bmatrix} z & 0 \\ 0 & z^{-1} \end{bmatrix},$$

where z is a primitive fourth root of unity in $\mathbb{Z}/\ell\mathbb{Z}$. A general $\mathcal{O}_K$-CM elliptic curve over K is of the form E_1^ψ for a character $\psi : \mathfrak{g}_K \to \mu_4 \subset (\mathbb{Z}/\ell\mathbb{Z})^\times$. Let $Q_4(\ell) = (\mathbb{Z}/\ell\mathbb{Z})^\times/(\mathbb{Z}/\ell\mathbb{Z})^{\times 4}$. Then the image of z in $Q_4(\ell)$ has order 4: if not, there is $w \in (\mathbb{Z}/\ell\mathbb{Z})^\times$ such that $z = w^2$, and then w has order 8 in $(\mathbb{Z}/\ell\mathbb{Z})^\times$, contradicting the assumption that $\ell \equiv 5 \pmod 8$. Thus the natural map $\mu_4 \to Q_4(\ell)$ given by $i \mapsto z \pmod{(\mathbb{Z}/\ell\mathbb{Z})^{\times 4}}$ is an isomorphism; we denote the inverse isomorphism $Q_4(\ell) \to \mu_4$ by ι. Now take

$$\psi : \mathfrak{g}_K \xrightarrow{\Psi_1^{-1}} (\mathbb{Z}/\ell\mathbb{Z})^\times \xrightarrow{q} Q_4(\ell) \xrightarrow{\iota} \mu_4;$$

here q is the quotient map. Let $\Psi_2 = \psi\Psi_1$. Then the twist E_1^ψ has mod ℓ Galois representation

$$\sigma \mapsto \rho_\ell(\sigma) = \begin{bmatrix} \Psi_2(\sigma) & 0 \\ 0 & \Psi_2^{-1}(\sigma)\chi_\ell(\sigma) \end{bmatrix}.$$

The composite $\Psi_2 : \mathfrak{g}_K \to (\mathbb{Z}/\ell\mathbb{Z})^\times \to Q_4(\ell)$ is trivial, so $\Psi_2(\mathfrak{g}_K)$ has order $c \mid \frac{\ell-1}{4}$. Thus

$$\#\rho_{\ell,E_1^\psi}(\mathfrak{g}_K) \mid c(\ell-1) \mid \frac{(\ell-1)^2}{4} = [K^{(\ell)} : K^{(1)}] = [K^{(\ell)} : K],$$

where the last equality holds since K has class number 1. Because $K(E_1^\psi[\ell]) \supset K^{(\ell)}$, we have $\#\rho_{\ell,E_1^\psi}(\mathfrak{g}_K) = \frac{1}{4}(\ell-1)^2$ and $c = \frac{1}{4}(\ell-1)$.

(b) Since $\ell \equiv 1 \pmod 3$, we have a primitive sixth root of unity z in $\mathbb{Z}/\ell\mathbb{Z}$. Since $\ell \equiv 7, 31 \pmod{36}$, we have $4, 9 \nmid \ell - 1$, so z has order 6 in $Q_6(\ell) = (\mathbb{Z}/\ell\mathbb{Z})^\times/(\mathbb{Z}/\ell\mathbb{Z})^{\times 6}$. Also $\frac{1}{6}(\ell-1)^2 = [K^{(\ell)} : K^{(1)}]$. The argument of part (a) carries over. □

Example 4.4. (a) Let $K = \mathbb{Q}(\sqrt{-1})$, and let $\ell \equiv 5 \pmod 8$. Let $E_{/K}$ be an $\mathcal{O}_K$-CM elliptic curve with mod ℓ Galois representation as in Proposition 4.3(a). Then for a number field $L \supset K$, $\rho_\ell|_{\mathfrak{g}_L}$ has scalar image if and only if $\chi_\ell \Psi^{-2}|_{\mathfrak{g}_L}$ is trivial. Since $\chi_\ell : \mathfrak{g}_K \to (\mathbb{Z}/\ell\mathbb{Z})^\times$ has order $\ell - 1$—that is, for all $1 \le k < \ell - 1$, $\chi_\ell^k \ne 1$—and Ψ^{-2} has order dividing $\frac{\ell-1}{4}$, the character $\chi_\ell \Psi^{-2}$ has order $\ell - 1$. Thus $[K(\mathbb{P}E[\ell]) : K] = \ell - 1$, whereas $[K(\ell) : K] = \frac{\ell-1}{2}$. So $d(E, \ell) = 2$.

(b) Let $K = \mathbb{Q}(\sqrt{-3})$, and let $\ell \equiv 7, 31 \pmod{36}$. Let $E_{/K}$ be an $\mathcal{O}$-CM elliptic curve with mod ℓ Galois representation as in Proposition 4.3(b). As in part (a), we have $[K(\mathbb{P}E[\ell]) : K] = \ell - 1$ and $[K(\ell) : K] = \frac{\ell-1}{3}$. So $d(E, \ell) = 3$.

Proposition 4.5. *Let $\mathcal{O}$ be an order of discriminant $\Delta = \mathfrak{f}^2 \Delta_K$, and let $N \in \mathbb{Z}^+$. Then there is an $\mathcal{O}$-CM elliptic curve $E_{/K(N\mathfrak{f})}$ such that the mod N Galois representation consists of scalar matrices.*

Proof. When $\Delta \notin \{-3, -4\}$, this is immediate from Theorem 4.1(b): in that case, the elliptic curve has a model defined over $K(\mathfrak{f})$. Thus we may assume that $\Delta \in \{-3, -4\}$, so $\mathfrak{f} = 1$. Let $\zeta \in \mathcal{O}_K^\times$ be a primitive w_K-th root of unity. Let $\mathcal{O}(N)$ be the order in K of conductor N, let $\tilde{E}_{/K(N)}$ be an $\mathcal{O}(N)$-CM elliptic curve, let $\iota : \tilde{E} \to E$ be the canonical $K(N)$-rational isogeny to an $\mathcal{O}_K$-CM elliptic curve E, let $\iota^\vee : E \to \tilde{E}$ be the dual isogeny, and let C be the kernel of $\iota^\vee$. Identifying $E[N]$ with $N^{-1}\mathcal{O}_K/\mathcal{O}_K \subset \mathbb{C}/\mathcal{O}_K$, $\iota^\vee : \mathbb{C}/\mathcal{O}_K \to \mathbb{C}/\mathcal{O}(N)$ is the map $z + \mathcal{O}_K \mapsto Nz + \mathcal{O}(N)$, so C is the $\mathbb{Z}$-submodule of $\mathbb{C}/\mathcal{O}_K$ generated by $P_1 = \frac{1}{N} + \mathcal{O}_K$. Because C is stable under the action of $\mathfrak{g}_{K(N)}$, this action is given by an isogeny character, say,

$$\sigma(P_1) = \Psi(\sigma) P_1.$$

Let $P_2 = \zeta P_1$. Then $\{P_1, P_2\}$ is a $\mathbb{Z}/N\mathbb{Z}$-basis for $E[N]$. Moreover, for $\sigma \in \mathfrak{g}_{K(N)}$,

$$\sigma P_2 = \sigma \zeta P_1 = \zeta \sigma P_1 = \zeta \Psi(\sigma) P_1 = \Psi(\sigma) \zeta P_1 = \Psi(\sigma) P_2.$$

It follows that $\sigma \in \mathfrak{g}_{K(N)}$ acts on $E[N]$ via the scalar matrix $\Psi(\sigma)$. □

5. Proof of Theorem 1.4 and its corollaries

5A. ***An equality of class fields.*** Let $\mathcal{O}$ and $\mathcal{O}'$ be orders in an imaginary quadratic field K of conductors $\mathfrak{f}$ and $N\mathfrak{f}$, respectively. Here we prove $K(\mathfrak{f})^{N\mathcal{O}_K} = K^{N\mathcal{O}_K} K(N\mathfrak{f})$. We may assume that $N \ge 2$. Class field theory (see, for example, [Stevenhagen 2001, (3.2)]) gives a canonical isomorphism,

$$\Psi : \mathrm{Aut}(K^{\mathrm{ab}}/K(\mathfrak{f})) \xrightarrow{\sim} \widehat{\mathcal{O}}^\times/\mathcal{O}^\times. \tag{4}$$

Thus it suffices to prove an equality of open subgroups of $\widehat{\mathcal{O}_K}^{\times}/\mathcal{O}_K^{\times}$. We abbreviate

$$\mathcal{O}_p := \mathcal{O} \otimes \mathbb{Z}_p.$$

Put

$$A := \{x \in \widehat{\mathcal{O}}^{\times} \mid x \equiv 1 \pmod{N}\} = \prod_{p \nmid N} \mathcal{O}_p^{\times} \times \prod_{p \mid N} (1 + N\mathcal{O}_p),\ \tilde{A} := A\mathcal{O}_K^{\times},$$

$$B := \widehat{\mathcal{O}'}^{\times} = \prod_{p} (\mathcal{O}')_p^{\times},\ \tilde{B} := B\mathcal{O}_K^{\times},$$

$$C := \{x \in \widehat{\mathcal{O}_K}^{\times} \mid x \equiv 1 \pmod{N}\} = \prod_{p \nmid N} (\mathcal{O}_K)_p^{\times} \times \prod_{p \mid N} (1 + N(\mathcal{O}_K)_p),\ \tilde{C} := C\mathcal{O}_K^{\times}.$$

Under class field theory, the field $K(\mathfrak{f})^{N\mathcal{O}_K}$ corresponds to $\tilde{A}$ (cf. [Stevenhagen 2001, p. 9]), the field $K(N\mathfrak{f})$ corresponds to $\tilde{B}$ and the field $K^{N\mathcal{O}_K}$ corresponds to $\tilde{C}$, so showing that $K(\mathfrak{f})^{N\mathcal{O}_K} = K^{N\mathcal{O}_K} K(N\mathfrak{f})$ is equivalent to showing that

$$\tilde{A} = \tilde{B} \cap \tilde{C}.$$

Step 1: We show that $A = B \cap C$. Writing $A_p,\ B_p$ and C_p for the components of p of each of these groups, it is enough to show that

$$A_p = B_p \cap C_p \text{ for all primes } p.$$

Case 1: Suppose $p \nmid N$. Then

$$A_p = \mathcal{O}_p^{\times},$$
$$B_p = (\mathcal{O}')_p^{\times} = A_p,$$
$$C_p = (\mathcal{O}_K)_p^{\times},$$

so $C_p \supset A_p = B_p$ and thus $B_p \cap C_p = A_p$.

Case 2: Suppose $p \mid N$. Write $\mathcal{O}_K = \mathbb{Z}1 + \mathbb{Z}\tau_K$, so $\mathcal{O} = \mathbb{Z}1 + \mathbb{Z}\mathfrak{f}\tau_K$. We have

$$A_p = 1 + N\mathcal{O}_p = 1 + N\mathbb{Z}_p 1 + N\mathfrak{f}\mathbb{Z}_p\tau_K,$$
$$B_p = (1 + \mathbb{Z}_p 1 + N\mathfrak{f}\mathbb{Z}_p\tau_K)^{\times},$$
$$C_p = 1 + N(\mathcal{O}_K)_p = 1 + N\mathbb{Z}_p 1 + N\mathbb{Z}_p\tau_K,$$

so indeed we have $B_p \cap C_p = A_p$.

It follows that $\tilde{B} \cap \tilde{C} = B\mathcal{O}_K^{\times} \cap C\mathcal{O}_K^{\times} \supset A\mathcal{O}_K^{\times} = \tilde{A}$, so it remains to show that $\tilde{B} \cap \tilde{C} \subset \tilde{A}$.

Step 2: Suppose $\Delta_K < -4$, so $\mathcal{O}_K^{\times} = \{\pm 1\}$. Then $\tilde{B} = B$, so if $z \in \tilde{B} \cap \tilde{C}$, then there is $\epsilon \in \{\pm 1\}$ such that $z \in B,\ -z \in B$ and $\epsilon z \in C$, so $\epsilon z \in B \cap C = A$ and thus $z \in \tilde{A}$.

Step 3: Suppose $K = \mathbb{Q}(\sqrt{-1})$ and let ζ be a primitive fourth root of unity, so $\mathcal{O}_K = \mathbb{Z}1 + \mathbb{Z}\zeta$ and $\mathcal{O}_K^\times = \{1, \zeta, \zeta^2, \zeta^3\}$. Suppose $z \in \tilde{B} \cap \tilde{C}$. Then there are $i, j \in \{0, 1, 2, 3\}$, $b \in B$ and $c \in C$ such that

$$z = \zeta^i b = \zeta^j c.$$

We have $z \in \tilde{A}$ if and only if $\zeta^{-j} z \in \tilde{A}$, so we may assume that $j = 0$. If i is even we may argue as in Step 2, so assume that $i \in \{1, 3\}$, and thus we have either $\zeta b = c$ or $\zeta c = b$. But we claim that there are no such elements b and c, which will complete the argument in this case. Indeed, choose a prime p dividing N, and let b_p and c_p be the components at p. There is a reduction map

$$(\mathcal{O}_K)_p \to \mathcal{O}_K \otimes \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/p\mathbb{Z}1 + \mathbb{Z}/p\mathbb{Z}\zeta.$$

Under this map, every element of $B_p \cup C_p$ lands in $\mathbb{Z}/p\mathbb{Z}1$, so $b_p, c_p \in \mathbb{Z}/p\mathbb{Z}1$ while $\zeta b_p, \zeta c_p \in \mathbb{Z}/p\mathbb{Z}\zeta$. Thus we cannot have $\zeta b_p = c_p$ or $\zeta c_p = b_p$.

If $K = \mathbb{Q}(\sqrt{-3})$, then we let ζ be a primitive sixth root of unity, so $\mathcal{O}_K = \mathbb{Z}1 + \mathbb{Z}\zeta$ and $\mathcal{O}_K^\times = \{1, \zeta, \zeta^2, \zeta^3, \zeta^4, \zeta^5\}$, and the argument is very similar: we cannot have $\pm\zeta b_p = c_p$ or $\pm b_p = \zeta c_p$.

5B. ***Proof of Theorem 1.4.*** By Theorems 2.11 and 4.1(a) and Section 5A, we have

$$K(\mathfrak{f})(\mathfrak{h}(E[N])) \supset K^{N\mathcal{O}_K} K(N\mathfrak{f}) = K(\mathfrak{f})^{N\mathcal{O}_K}.$$

For any $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$, the splitting field $\overline{K(\mathfrak{f})}^{\ker \overline{\rho_N}}$ of the reduced mod N Galois representation $\overline{\rho_N}$ on E (cf. Section 1C) is $K(\mathfrak{f})(\mathfrak{h}(E[N]))$, so

$$[K(\mathfrak{f})(\mathfrak{h}(E[N])) : K(\mathfrak{f})] \leq \#\overline{C_N(\mathcal{O})}.$$

As described in the introduction, it is immediate from (4) and the definition of $K(\mathfrak{f})^{N\mathcal{O}_K}$ that

$$\operatorname{Aut}(H_{N,\mathcal{O}}/K(\mathfrak{f})) = \overline{C_N(\mathcal{O})},$$

and thus it follows that

$$K(\mathfrak{f})(\mathfrak{h}(E[N])) = K^{N\mathcal{O}_K} K(N\mathfrak{f}).$$

5C. ***Proof of Corollaries 1.5, 1.7 and 1.8.***

Proof of Corollary 1.5. Theorem 1.4 implies that for any number field $F \supset K(j(E))$ and $N \in \mathbb{Z}^+$, we have

$$[\overline{C_N(\mathcal{O})} : \overline{\rho_N}(\mathfrak{g}_F)] \mid [F : K(j(E))]$$

and thus

$$[C_N(\mathcal{O}) : \rho_N(\mathfrak{g}_F)] \mid \#\mathcal{O}^\times [F : K(j(E))] \leq 6[F : K(j(E))]. \qquad \square$$

Proof of Corollary 1.7. We may assume that $N \geq 2$. Let $w = \#q_N(\mathcal{O}^\times)$, so

$$w = \begin{cases} \#\mathcal{O}^\times, & N \geq 3, \\ \frac{\#\mathcal{O}^\times}{2}, & N = 2. \end{cases}$$

Once again we denote by μ_w the image of $\mathcal{O}^\times$ in $C_N(\mathcal{O})$, a cyclic group of order w. Let $E_{/K(\mathfrak{f})}$ be any $\mathcal{O}$-CM elliptic curve. We may view $G = \operatorname{Aut}(K(\mathfrak{f})(E[N])/K(\mathfrak{f}))$ as a subgroup of $C_N(\mathcal{O})$. Let $H = G \cap \mu_w$ and $L = (K(\mathfrak{f})(E[N]))^H$, so a suitable twist $(E')_{/L}$ of $E_{/L}$ has trivial mod N Galois representation. As shown in the proof of Theorem 2.10, we have $L = K(\mathfrak{f})(\mathfrak{h}(E[N]))$, so by Theorem 1.4 we have $[L : K(\mathfrak{f})] = \#\overline{C_N(\mathcal{O})}$. □

Proof of Corollary 1.8. We may assume that $N \geq 2$. Let $q_N : \mathcal{O}^\times \to C_N(\mathcal{O})$ be the natural homomorphism. By Theorem 2.10(b), there is an elliptic curve $E_{/K(\mathfrak{f})}$ such that

$$[K(\mathfrak{f})(E[N]) : K(\mathfrak{f})(\mathfrak{h}(E[N]))] = \#q_N(\mathcal{O}^\times) = \begin{cases} \#\mathcal{O}^\times, & N \geq 3, \\ \frac{\#\mathcal{O}^\times}{2}, & N = 2. \end{cases}$$

By Theorem 1.4, $[K(\mathfrak{f})(\mathfrak{h}(E[N])) : K(\mathfrak{f})] = \#\overline{C_N(\mathcal{O})}$. Thus $\rho_{E,N}(\mathfrak{g}_{K(\mathfrak{f})}) = C_N(\mathcal{O})$. □

6. Applications

6A. *Divisibility in Silverberg's theorem.*

Lemma 6.1. *Let J, M be subgroups of a group G. If M is normal and $J \cap M = \{1\}$, then $\#J \mid [G : M]$.*

Proof. The composite homomorphism $J \hookrightarrow G \to G/M$ is an injection. □

The following result extends [Bourdon et al. 2017a, Corollary 2.5] from maximal orders to all imaginary quadratic orders, thereby confirming the expectation expressed in [Bourdon et al. 2017a, Remarks following the proof of Corollary 2.5].

Theorem 6.2. *Let $\mathcal{O}$ be an order in an imaginary quadratic field K, and let E be an $\mathcal{O}$-CM elliptic curve defined over a number field $F \supset K$. If $E(F)$ has a point of order $N \in \mathbb{Z}^+$, then*

$$\varphi(N) \mid \frac{\#\mathcal{O}^\times}{2} \frac{[F : \mathbb{Q}]}{\#\operatorname{Pic}\mathcal{O}}.$$

Proof. Let $\mathcal{I}_N = [C_N(\mathcal{O}) : \rho_N(\mathfrak{g}_F)]$ be the index of the mod N Galois representation in the Cartan subgroup. By Corollary 1.5 we have

$$\mathcal{I}_N \mid \#\mathcal{O}^\times [F : K(j(E))] = \frac{\#\mathcal{O}^\times}{2} \frac{[F : \mathbb{Q}]}{\#\operatorname{Pic}\mathcal{O}}.$$

Since there is a rational point of order N, the subgroup $\rho_N(\mathfrak{g}_F)$ contains no scalar matrices other than the identity. Applying Lemma 6.1 with $G = C_N(\mathcal{O})$, $M = \rho_N(\mathfrak{g}_F)$ and J the subgroup of scalar matrices, we get $\varphi(N) \mid \mathcal{I}_N$, and we are done. □

6B. ***A theorem of Franz.*** Let $\mathcal{O}$ be an order in K, of conductor $\mathfrak{f}$, and let $E_{/K(\mathfrak{f})}$ be an $\mathcal{O}$-CM elliptic curve. Choose a field embedding $K(\mathfrak{f}) \hookrightarrow \mathbb{C}$ such that $j(E) = j(\mathbb{C}/\mathcal{O})$ and an isomorphism $E(\mathbb{C}) \xrightarrow{\sim} \mathbb{C}/\mathcal{O}$. This induces an isomorphism $E(\overline{K(\mathfrak{f})})[\text{tors}] \xrightarrow{\sim} \mathbb{C}/\mathcal{O}[\text{tors}]$, which we use to view (the image in $\mathbb{C}/\mathcal{O}$ of) $\tau_K = \frac{1}{2}(\Delta_K + \sqrt{\Delta_K})$ as a point of $E(\overline{K(\mathfrak{f})})[\text{tors}]$ of order $\mathfrak{f}$.

Theorem 6.3 [Franz 1935]. *With notation as above, we have*

$$K(\mathfrak{f})(\mathfrak{h}(\tau_K)) = K^{(\mathfrak{f})}.$$

Proof. As in the proof of Theorem 1.3, over $\mathbb{C}$ we may view the canonical isogeny as $\iota : \mathbb{C}/\mathcal{O} \to \mathbb{C}/\mathcal{O}_K$. We take $e_1 = \frac{1}{\mathfrak{f}} + \mathcal{O}$ and $e_2 = \tau_K + \mathcal{O}$ as a basis for $E[\mathfrak{f}]$. Then e_2 generates $\ker(\iota)$, a $K(\mathfrak{f})$-rational cyclic subgroup of order $\mathfrak{f}$, and with respect to $\{e_1, e_2\}$ the image of the mod $\mathfrak{f}$ Galois representation associated to $E_{/K(\mathfrak{f})}$ consists of matrices of the form

$$\begin{bmatrix} a & b\mathfrak{f}^2\frac{\Delta_K - \Delta_K^2}{4} \\ b & a + b\mathfrak{f}\Delta_K \end{bmatrix} \quad \text{with } a, b \in \mathbb{Z}/\mathfrak{f}\mathbb{Z}.$$

Viewing entries mod $\mathfrak{f}$, we see there is a character $\Psi : \mathfrak{g}_{K(\mathfrak{f})} \to (\mathbb{Z}/\mathfrak{f}\mathbb{Z})^\times$ such that

$$\rho_{E,\mathfrak{f}}(\sigma) = \begin{bmatrix} \Psi(\sigma) & 0 \\ * & \Psi(\sigma) \end{bmatrix}.$$

If $\mathfrak{f} \le 2$, then $K(\mathfrak{f})(\mathfrak{h}(\tau_K)) = K(\mathfrak{f}) = K^{(\mathfrak{f})}$ and the result holds. Thus we may assume $\mathfrak{f} \ge 3$. Let $L := K(\mathfrak{f})(\mathfrak{h}(e_2))$. Since $\mathfrak{f} \ge 3$, we have $j(E) \ne 0, 1728$, so $[L(e_2) : L]$ divides 2 and the restriction $\Psi|_{\mathfrak{g}_L} : \mathfrak{g}_L \to \{\pm 1\}$ defines a quadratic character χ. On the twist E^χ of $E_{/L}$ the point e_2 becomes L-rational. As in the proof of Theorem 5.5 of [Bourdon et al. 2017b], let $\Psi^\pm : \mathfrak{g}_{K(\mathfrak{f})} \to (\mathbb{Z}/\mathfrak{f}\mathbb{Z}^\times)/\{\pm 1\}$ denote the composition of Ψ with the natural map $(\mathbb{Z}/\mathfrak{f}\mathbb{Z})^\times \to (\mathbb{Z}/\mathfrak{f}\mathbb{Z})^\times/\pm 1$. Then $L \subset (\overline{K(\mathfrak{f})})^{\ker \Psi^\pm}$, so $[L : K(\mathfrak{f})] \mid \frac{\varphi(\mathfrak{f})}{2}$. If $\iota : E^\chi \to E'$ is the canonical isogeny, then the proof of Theorem 1.3 shows that $\iota(e_1)$ is an element of $E'(L)$ which generates $E'[\mathfrak{f}]$ as an $\mathcal{O}_K$-module. Thus E' has full $\mathfrak{f}$-torsion over L, so by Theorem 2.8, $K^{(\mathfrak{f})} \subset L$. So

$$[L : K(\mathfrak{f})] \ge [K^{(\mathfrak{f})} : K(\mathfrak{f})] = \frac{\varphi(\mathfrak{f})}{2} \ge [L : K(\mathfrak{f})],$$

and thus $K(\mathfrak{f})(\mathfrak{h}(e_2)) = L = K^{(\mathfrak{f})}$. □

6C. ***The field of moduli of a point of prime order.*** Let $K \ne \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$ be an imaginary quadratic field, and let $\mathcal{O} \subset K$ be the order of conductor $\mathfrak{f}$. Here we use Theorem 1.3 to determine the smallest field $F \supset K$ for which there exists an $\mathcal{O}$-CM elliptic curve $E_{/F}$ with an F-rational point of order $\ell > 2$.

Lemma 6.4. *Let K be an imaginary quadratic field, let $\mathfrak{f} \in \mathbb{Z}^+$, and let $\ell > 2$ be prime. Then $K^{(\ell)} \cap K(\ell\mathfrak{f}) = K(\ell)$.*

Proof. Let $\Delta = \mathfrak{f}^2 \Delta_K$. The statement is immediate if $\mathfrak{f} = 1$, so suppose $\mathfrak{f} > 1$. By Theorem 2.1,

$$[K(\ell\mathfrak{f}) : K(\mathfrak{f})] = \ell - \left(\frac{\Delta}{\ell}\right).$$

Since $[K^{(\ell)} K(\ell\mathfrak{f}) : K(\mathfrak{f})] = \# C_\ell(\mathcal{O})/2$ by Theorem 1.4, in both cases we have

$$[K^{(\ell)} K(\ell\mathfrak{f}) : K(\ell\mathfrak{f})] = \frac{\# C_\ell(\mathcal{O})}{2[K(\ell\mathfrak{f}) : K(\mathfrak{f})]} = \tfrac{1}{2}(\ell - 1).$$

Thus $[K^{(\ell)} : K^{(\ell)} \cap K(\ell\mathfrak{f})] = [K^{(\ell)} K(\ell\mathfrak{f}) : K(\ell\mathfrak{f})] = \frac{1}{2}(\ell - 1)$. As we have $K(\ell) \subset K^{(\ell)} \cap K(\ell\mathfrak{f})$ and $[K^{(\ell)} : K(\ell)] = \frac{1}{2}(\ell - 1)$, the result follows. □

Theorem 6.5. *Let $K \neq \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$ be an imaginary quadratic field, and let $\mathcal{O}$ be the order of conductor $\mathfrak{f}$ in K. Let $F \supset K$.*

(a) *Let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve such that $E(F)$ contains a point of prime order $\ell > 2$. Then there is a prime $\mathfrak{p}$ of $\mathcal{O}_K$ lying over ℓ such that $K(\mathfrak{f})K^{\mathfrak{p}} \subset F$.*

(b) *If $\left(\frac{\Delta}{\ell}\right) \neq -1$, then there is a prime $\mathfrak{p}$ of $\mathcal{O}_K$ lying over ℓ and an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})K^{\mathfrak{p}}}$ such that $E(K(\mathfrak{f})K^{\mathfrak{p}})$ has a point of order ℓ.*

If $\left(\frac{\Delta}{\ell}\right) = -1$, then an $\mathcal{O}$-CM elliptic curve $E_{/F}$ with an F-rational point of order ℓ must have full ℓ-torsion (see [Bourdon et al. 2017b, Theorem 4.8] or Lemma 6.12). In this case, $K(\ell\mathfrak{f})K^{(\ell)} \subset F$ by Theorem 1.4. The existence of an elliptic curve $E_{/K(\ell\mathfrak{f})K^{(\ell)}}$ with full ℓ-torsion is guaranteed by Corollary 1.7.

Proof. (a) Let $F \supset K$ and $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve with an F-rational point of order ℓ. By Theorem 1.3, there is an $\mathcal{O}_K$-CM elliptic curve $E'_{/F}$ with an F-rational point P of order ℓ. If M is the $\mathcal{O}_K$-submodule of $E'(F)$ generated by P, then $M = E'[\operatorname{ann} M]$ and $\# M = |\operatorname{ann} M|$ by Theorem 2.6. Since $\ell \mid \# M$, we must have $\mathfrak{p} \mid \operatorname{ann} M$ for some prime $\mathfrak{p}$ of $\mathcal{O}_K$ above ℓ. By Theorem 2.8 we have

$$K(\mathfrak{f})K^{\mathfrak{p}} \subset K(\mathfrak{f})K^{\operatorname{ann} M} = K(j(E))K^{(1)}(\mathfrak{h}(E'[\operatorname{ann} M])) \subset F.$$

(b) If $\left(\frac{\Delta}{\ell}\right) \neq -1$, then an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ possesses a $K(\mathfrak{f})$-rational cyclic subgroup of order ℓ. (See, e.g., [Clark et al. 2013, p. 13]. This is also a special case of Theorem 6.18.) By [Bourdon et al. 2017b, Theorem 5.5], there is an extension $L/K(\mathfrak{f})$ of degree $(\ell - 1)/2$ and a quadratic twist $(E_1)_{/L}$ such that $E_1(L)$ has a point of order ℓ. By part (a), there is a prime $\mathfrak{p}$ of $\mathcal{O}_K$ lying over ℓ such that $K(\mathfrak{f})K^{\mathfrak{p}} \subset L$, so it will suffice to show that $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] \geq \frac{\ell-1}{2}$.

If $\ell \nmid \mathfrak{f}$, then primes above ℓ are unramified in $K(\mathfrak{f})/K^{(1)}$. Thus $K(\mathfrak{f})$, $K^{\mathfrak{p}}$ are linearly disjoint over $K^{(1)}$, and we have $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] = [K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell - 1)$

since $\left(\frac{\Delta_K}{\ell}\right)=\left(\frac{\Delta}{\ell}\right)\neq -1$. If $\ell \mid \mathfrak{f}$, then applying Lemma 6.4 with $\frac{\mathfrak{f}}{\ell}$ in place of $\mathfrak{f}$, we have

$$K^{\mathfrak{p}} \cap K(\mathfrak{f}) \subset K^{(\ell)} \cap K(\mathfrak{f}) = K(\ell).$$

Thus $K^{\mathfrak{p}} \cap K(\mathfrak{f}) = K^{\mathfrak{p}} \cap K(\ell)$, so

$$[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] = [K^{\mathfrak{p}} : K^{\mathfrak{p}} \cap K(\mathfrak{f})] = [K^{\mathfrak{p}} : K^{\mathfrak{p}} \cap K(\ell)] = [K(\ell)K^{\mathfrak{p}} : K(\ell)]$$

and it is enough to show that $[K(\ell)K^{\mathfrak{p}} : K(\ell)] \geq \frac{\ell-1}{2}$.

• $\left(\frac{\Delta_K}{\ell}\right) = 1$: We will prove that $K^{\mathfrak{p}} \cap K(\ell) = K^{(1)}$ using CM elliptic curves. Let $(E_0)_{/K^{(1)}}$ be an $\mathcal{O}_K$-CM elliptic curve. Then $E_0[\mathfrak{p}]$ is stable under the action of $\mathfrak{g}_{K^{(1)}}$ and generated by a point P of order ℓ. By [Bourdon et al. 2017b, Theorem 5.5], there is an extension $L/K^{(1)}$ of degree $(\ell-1)/2$ and a quadratic twist $(E_1)_{/L}$ such that P becomes L-rational. By Theorem 2.8 we have $K^{\mathfrak{p}} \subset L$, and $K^{\mathfrak{p}} = L$ since $[K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell-1)$. Over $K(\ell)K^{\mathfrak{p}}$, the curve E_1 has a rational point of order ℓ, and the mod ℓ Galois representation is scalar by Theorem 4.1. Thus E_1 has full ℓ-torsion over $K(\ell)K^{\mathfrak{p}}$, and $K^{(\ell)} \subset K(\ell)K^{\mathfrak{p}}$. This implies $\frac{1}{2}(\ell-1) \mid [K(\ell)K^{\mathfrak{p}} : K(\ell)] = [K^{\mathfrak{p}} : K^{\mathfrak{p}} \cap K(\ell)]$. Since $[K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell-1)$, we have $K^{\mathfrak{p}} \cap K(\ell) = K^{(1)}$, and $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] = [K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell-1)$.

• $\left(\frac{\Delta_K}{\ell}\right) = -1$: In this case, $K^{\mathfrak{p}} = K^{(\ell)}$, so $K^{\mathfrak{p}} \cap K(\ell) = K(\ell)$. This implies $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] = [K^{\mathfrak{p}} : K(\ell)] = \frac{1}{2}(\ell-1)$.

• $\left(\frac{\Delta_K}{\ell}\right) = 0$: Since $[K(\ell) : K^{(1)}] = \ell$ and $[K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell-1)$, we have $K^{\mathfrak{p}} \cap K(\ell) = K^{(1)}$. Thus $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] = [K^{\mathfrak{p}} : K^{(1)}] = \frac{1}{2}(\ell-1)$. ☐

Remark 6.6. Assume the setup of Theorem 6.5 but take $K = \mathbb{Q}(\sqrt{-1})$ or $K = \mathbb{Q}(\sqrt{-3})$. Then the assertion of Theorem 6.5(b) is false. Indeed, if $\ell \geq 5$ and $\left(\frac{\Delta}{\ell}\right) \neq -1$, we have $[K(\mathfrak{f})K^{\mathfrak{p}} : K(\mathfrak{f})] \mid \frac{1}{w_K}(\ell-1)$. (See Lemma 2.9.) Suppose $F \supset K$, and let $E_{/F}$ be an elliptic curve with CM by the order in K of conductor $\mathfrak{f}$. If $E(F)$ contains a rational point of order ℓ, then Theorem 6.2 implies $\frac{1}{2}(\ell-1) \mid [F : K(\mathfrak{f})]$. Thus F must properly contain $K(\mathfrak{f})K^{\mathfrak{p}}$.

6D. ***Sharpness in the Isogeny Torsion Theorem.*** The following result was established during the proof of Theorem 1.3.

Lemma 6.7. *Let E be an $\mathcal{O}$-CM elliptic curve defined over a number field F containing the CM field K, and for a positive integer $\mathfrak{f}'$ dividing the conductor $\mathfrak{f}$ of $\mathcal{O}$, let $\iota : E \to E'$ be the canonical F-rational isogeny to an elliptic curve E' with CM by the order in K of conductor $\mathfrak{f}'$. Write*

$$E(F)[\text{tors}] = \mathbb{Z}/s\mathbb{Z} \times \mathbb{Z}/e\mathbb{Z},\ \ E'(F)[\text{tors}] = \mathbb{Z}/s'\mathbb{Z} \times \mathbb{Z}/e'\mathbb{Z},$$

where $s \mid e$ and $s' \mid e'$. Then $s \mid s'$.

Ross [1994, §4], claims that if E is a CM elliptic curve defined over a number field F containing the CM field, then the exponent of the finite group $E(F)[\mathrm{tors}]$ is an invariant of the F-rational isogeny class. In the setting of Lemma 6.7, this would give $e = e'$, and combining this with the conclusion of Lemma 6.7 we would get an injective group homomorphism $E(F)[\mathrm{tors}] \hookrightarrow E(F')[\mathrm{tors}]$. This conclusion is stronger than that of Theorem 1.3. However Ross's claim is false: in the setup of Lemma 6.7, one can have $e' < e$ (in which case there is no injective group homomorphism $E(F)[\mathrm{tors}] \hookrightarrow E'(F)[\mathrm{tors}]$), as the following result shows.

Proposition 6.8. *Let $\ell > 3$ be a prime number, let $K = \mathbb{Q}(\sqrt{-\ell})$, let $n \in \mathbb{Z}^{\geq 3}$, let $\mathcal{O}$ be the order in K of conductor $\mathfrak{f} = \ell^{\lfloor \frac{n}{2} \rfloor}$, and let $F = K(\mathfrak{f})$. For any $\mathcal{O}$-CM elliptic curve $E_{/F}$, there is an extension L/F of degree $\varphi(\ell^n)$ such that $E(L)$ has a point of order ℓ^n, and no $\mathcal{O}_K$-CM elliptic curve has an L-rational point of order ℓ^k for $k > \frac{1}{2}\left(n + 1 + \lfloor \frac{n}{2} \rfloor\right)$ (hence no L-rational point of order ℓ^n).*

Proof. Let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve. As in (3) we may choose a basis $\{e_1, e_2\}$ for $E[\ell^n]$ so that the image of the mod ℓ^n Galois representation consists of matrices

$$\begin{bmatrix} a & b\mathfrak{f}^2 \frac{\Delta_K - \Delta_K^2}{4} \\ b & a + b\mathfrak{f}\Delta_K \end{bmatrix} \quad \text{with } a, b \in \mathbb{Z}/\ell^n\mathbb{Z}.$$

Since ℓ ramifies in K and $\mathfrak{f} = \ell^{\lfloor \frac{n}{2} \rfloor}$, we have $\mathrm{ord}_\ell(b\mathfrak{f}^2(\Delta_K - \Delta_K^2)/4) = 1 + 2\lfloor \frac{n}{2} \rfloor \geq n$, so the matrices have the form

$$\begin{bmatrix} a & 0 \\ b & a + b\mathfrak{f}\Delta_K \end{bmatrix} \quad \text{with } a, b \in \mathbb{Z}/\ell^n\mathbb{Z}.$$

The action of $\mathfrak{g}_F$ on $\langle e_2 \rangle$ gives a character $\Phi : \mathfrak{g}_F \to (\mathbb{Z}/\ell^n\mathbb{Z})^\times$. Take $M = (\overline{F})^{\ker \Phi}$. Then $[M : F] \mid \varphi(\ell^n)$ and $\Phi|_{\mathfrak{g}_M}$ is trivial. Thus there exists an extension L/F with $[L : F] = \varphi(\ell^n)$ such that $E(L)$ contains e_2.

Let $E'_{/L}$ be an $\mathcal{O}_K$-CM elliptic curve, and suppose $E'(L)$ contains a point P of order ℓ^k. Let $\mathfrak{p}$ be the prime ideal of $\mathcal{O}_K$ such that $\ell\mathcal{O}_K = \mathfrak{p}^2$. We claim that the $\mathcal{O}_K$-submodule $M = \langle P \rangle_{\mathcal{O}_K}$ of $E'(L)$ generated by P contains $E[\mathfrak{p}^{2k-1}]$ and so, by Theorem 2.8, that $K^{\mathfrak{p}^{2k-1}} \subset L$. Indeed, by Theorem 2.6, we have $M = E[I]$ for some ideal I of $\mathcal{O}_K$ such that $(\mathcal{O}_K/I, +)$ has ℓ-power order and exponent ℓ^k. Since ℓ ramifies in $\mathcal{O}_K$, this forces I to be of the form $\mathfrak{p}^a$ for some $a \in \mathbb{Z}^+$, and the smallest a such that $(\mathcal{O}_K/\mathfrak{p}^a, +)$ has exponent ℓ^k is $a = 2k - 1$, establishing the claim. Thus

$$\mathrm{ord}_\ell([K^{\mathfrak{p}^{2k-1}} : K^{(1)}]) = 2k - 2 \leq \mathrm{ord}_\ell([L : K^{(1)}]) = \left\lfloor \frac{n}{2} \right\rfloor + n - 1,$$

so $k \leq \frac{1}{2}(n + 1 + \lfloor \frac{n}{2} \rfloor)$. □

In the setting of Theorem 1.3, one wonders whether $\#E(F)[\mathrm{tors}] = \#E'(F)[\mathrm{tors}]$. In fact $\#E'(F)[\mathrm{tors}]/\#E(F)[\mathrm{tors}]$ can be arbitrarily large:

Proposition 6.9. *Let ℓ be an odd prime, let $K \neq \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-3})$ be an imaginary quadratic field, let $\mathcal{O}$ be the order in K of conductor ℓ, and let $F = K(\ell)$. For any $\mathcal{O}$-CM elliptic curve $E_{/F}$ there is an extension L/F such that if $\iota : E \to E'$ is the canonical isogeny to an $\mathcal{O}_K$-CM elliptic curve E', then*

$$\ell \mid \frac{\#E'(L)[\text{tors}]}{\#E(L)[\text{tors}]}.$$

Proof. Let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve. As above, there is a basis $\{e_1, e_2\}$ for $E[\ell]$ such that

$$\rho_\ell(\mathfrak{g}_F) \subset \left\{ \begin{bmatrix} a & 0 \\ b & a \end{bmatrix} \,\middle|\, a, b \in \mathbb{Z}/\ell\mathbb{Z} \right\}$$

and there is an extension L/F with $[L : F] = \ell - 1$ such that $E(L)$ contains e_2. In fact, $E(L)[\ell^\infty] \cong \mathbb{Z}/\ell\mathbb{Z}$. Indeed, E does not have full ℓ-torsion over L since Theorem 1.4 would imply $K^{(\ell)}K(\ell^2) \subset L$ and $\frac{1}{2}\ell(\ell-1) = [K^{(\ell)}K(\ell^2) : K(\ell)]$. In addition, $E(L)$ has no point of order ℓ^2 by Theorem 6.2.

Let $\iota : E \to E'$ be the canonical L-rational isogeny from $E_{/L}$ to $E'_{/L}$, where E' has $\mathcal{O}_K$-CM. Since $e_2 \in E(L)$, the second paragraph of Section 3 shows that $\iota(e_1) \in E'(L)$, and $\iota(e_1)$ generates $E'[\ell]$ as an $\mathcal{O}_K$-module. In other words, $\mathbb{Z}/\ell\mathbb{Z} \times \mathbb{Z}/\ell\mathbb{Z} \hookrightarrow E'(L)[\text{tors}]$. It follows that $\ell \mid (\#E'(L)[\text{tors}]/\#E(L)[\text{tors}])$. ☐

Finally, Theorem 1.3 requires $K \subset F$. This hypothesis cannot be omitted:

Proposition 6.10. *Let $\ell > 3$ be a prime with $\ell \equiv 3 \pmod 4$ and let $n \in \mathbb{Z}^{\geq 3}$. Let $K = \mathbb{Q}(\sqrt{-\ell})$, and let $\mathcal{O}$ be the order in K of conductor $\mathfrak{f} = \ell^{\lfloor \frac{n}{2} \rfloor}$. Let $F = \mathbb{Q}(j(\mathbb{C}/\mathcal{O}))$. There is an elliptic curve $E_{/F}$ and an extension L/F of degree $\frac{\varphi(\ell^n)}{2}$ such that*

(i) $L \not\supset K$,

(ii) *$E(L)$ has a point of order ℓ^n, and*

(iii) *for every $\mathcal{O}_K$-CM elliptic curve $E'_{/L}$ we have $\ell^n \nmid \#E'(L)[\text{tors}]$.*

Proof. Let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve. By [Kwon 1999, Corollary 4.2], E has an F-rational subgroup which is cyclic of order ℓ^n. It follows from [Bourdon et al. 2017b, Theorem 5.6] that there is a twist E_1 of $E_{/F}$ and an extension L/F of degree $\varphi(\ell^n)/2$ such that $E_1(L)$ has a point of order ℓ^n. Note $[L : \mathbb{Q}] = h_K \ell^{\lfloor \frac{n}{2} \rfloor} \frac{\varphi(\ell^n)}{2}$ is odd (see [Cox 1989, Proposition 3.11]), so $K \not\subset L$.

Let $E'_{/L}$ be an $\mathcal{O}_K$-CM elliptic curve. Since $[L : \mathbb{Q}]$ is odd, $E'(L)[\ell^\infty]$ must be cyclic, as full ℓ^k-torsion would imply $\mathbb{Q}(\zeta_{\ell^k}) \subset L$ by the Weil pairing. As in the last paragraph of the proof of Proposition 6.8, $E'(LK)$ contains no point of order ℓ^n. Hence $E'(L)$ contains no point of order ℓ^n, and $\ell^n \nmid \#E'(L)[\text{tors}]$. ☐

6E. ***Minimal and maximal Cartan orbits.*** Let $\mathcal{O}$ be an order, let $N \in \mathbb{Z}^+$, and let $P \in \mathcal{O}/N\mathcal{O}$ be a point of order N. Since $C_N(\mathcal{O})$ contains all scalar matrices, if $P \in \mathcal{O}/N\mathcal{O}$ has order N, then the orbit of $C_N(\mathcal{O})$ on P has size at least $\varphi(N)$. On the other hand, the orbit of $C_N(\mathcal{O})$ on P is certainly no larger than the number of order N points of $\mathcal{O}/N\mathcal{O}$.

In this section we will find all pairs $(\mathcal{O}, N)$ for which there exists a Cartan orbit of this smallest possible size and also all pairs for which there exists a Cartan orbit of this largest possible size.

We introduce the shorthand $H(\mathcal{O}, N)$ to mean: *there is a point P of order N in $\mathcal{O}/N\mathcal{O}$ such that the $C_N(\mathcal{O})$-orbit of P has size $\varphi(N)$.*

Lemma 6.11. *Let $\mathcal{O}$ be an order, and let $N = \ell_1^{a_1} \cdots \ell_r^{a_r} \in \mathbb{Z}^+$. Then $H(\mathcal{O}, N)$ holds if and only if $H(\mathcal{O}, \ell_i^{a_i})$ holds for all $1 \leq i \leq r$.*

Proof. This is an easy consequence of the Chinese remainder theorem. □

Lemma 6.12. *Let $\mathcal{O}$ be the order of discriminant Δ, ℓ a prime number and $a \in \mathbb{Z}^+$.*

(a) *If $\left(\frac{\Delta}{\ell}\right) = 1$, there is an $\mathcal{O}$-submodule of $\mathcal{O}/\ell^a\mathcal{O}$ with underlying $\mathbb{Z}$-module $\mathbb{Z}/\ell^a\mathbb{Z}$.*

(b) *If $\left(\frac{\Delta}{\ell}\right) = -1$, then $C_{\ell^a}(\mathcal{O})$ acts simply transitively on the order ℓ^a elements of $\mathcal{O}/\ell^a\mathcal{O}$.*

Proof. (a) If $\left(\frac{\Delta}{\ell}\right) = 1$, then $\mathcal{O}/\ell\mathcal{O} = \mathcal{O}_K/\ell\mathcal{O}_K \cong \mathbb{Z}/\ell\mathbb{Z} \times \mathbb{Z}/\ell\mathbb{Z}$, so $\mathcal{O} \otimes \mathbb{Z}_\ell$ is isomorphic as a ring to $\mathbb{Z}_\ell \times \mathbb{Z}_\ell$ (see, e.g., [Eisenbud 1995, Corollary 7.5]) and thus $\mathcal{O}/\ell^a\mathcal{O}$ is isomorphic as a ring to $\mathbb{Z}/\ell^a\mathbb{Z} \times \mathbb{Z}/\ell^a\mathbb{Z}$.

(b) If $\left(\frac{\Delta}{\ell}\right) = -1$, then $\mathcal{O} \otimes \mathbb{Z}_\ell = \mathcal{O}_K \otimes \mathbb{Z}_\ell$ is a complete DVR with uniformizer ℓ, so the ring $\mathcal{O}/\ell^a\mathcal{O}$ is finite, local and principal with maximal ideal $\langle \ell \rangle$. An element of $\mathcal{O}/\ell^a\mathcal{O}$ has order ℓ^a if and only if it lies in the unit group $C_{\ell^a}(\mathcal{O})$. □

Lemma 6.13. *Let $\mathcal{O}$ be the order of discriminant Δ, and let $N \in \mathbb{Z}^+$. The following are equivalent:*

(i) *If $2 \mid N$, then $\left(\frac{\Delta}{2}\right) \neq 1$.*

(ii) *The $\mathbb{Z}/N\mathbb{Z}$-subalgebra of $\mathcal{O}/N\mathcal{O}$ generated by $C_N(\mathcal{O})$ is $\mathcal{O}/N\mathcal{O}$.*

Proof. Using the Chinese remainder theorem we reduce to the case of $N = \ell^a$ a power of a prime number ℓ. Let B be the $\mathbb{Z}/\ell^a\mathbb{Z}$-subalgebra generated by $C_{\ell^a}(\mathcal{O})$, so $\#B = \ell^b$ for some $b \leq 2a$.

(i)$\Longrightarrow$(ii) Since $0 \in B \setminus C_{\ell^a}(\mathcal{O})$, we have

$$\begin{aligned}\#B &\geq \#C_{\ell^a}(\mathcal{O}) + 1 \\ &= \ell^{2a}\left(1 - \frac{1}{\ell}\right)\left(1 - \left(\frac{\Delta}{\ell}\right)\frac{1}{\ell}\right) + 1 \geq \begin{cases} \frac{4}{9}\ell^{2a} + 1 > \ell^{2a-1} & \text{if } \ell \geq 3, \\ \frac{1}{2}\ell^{2a} + 1 > \ell^{2a-1} & \text{if } \ell = 2 \text{ and } \left(\frac{\Delta}{2}\right) \neq 1. \end{cases}\end{aligned}$$

Thus $b = 2a$ and $B = \mathcal{O}/\ell^a\mathcal{O}$.

¬(i)⟹ ¬(ii) If $\ell = 2$ and $\left(\frac{\Delta}{2}\right) = 1$, then

$$\mathcal{O}/2^a\mathcal{O} \cong \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \;\middle|\; \alpha, \beta \in \mathbb{Z}/2^a\mathbb{Z} \right\}$$

and $C_{2^a}(\mathcal{O})$ consists of the set of such matrices with $\alpha, \beta \in (\mathbb{Z}/2^a\mathbb{Z})^\times$. Thus $C_{2^a}(\mathcal{O})$ is contained in the subalgebra

$$\mathcal{B} = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \;\middle|\; \alpha, \beta \in \mathbb{Z}/2^a\mathbb{Z} \text{ and } \alpha \equiv \beta \pmod 2 \right\}$$

of order 2^{2a-1}, so $B \subset \mathcal{B} \subsetneq \mathcal{O}/2^a\mathcal{O}$.[1] □

Lemma 6.14. *For an order $\mathcal{O}$ and $N \in \mathbb{Z}^+$, the following are equivalent*:

(i) *There is an ideal I of $\mathcal{O}$ with $\mathcal{O}/I \cong \mathbb{Z}/N\mathbb{Z}$.*

(ii) *There is an $\mathcal{O}$-submodule of $\mathcal{O}/N\mathcal{O}$ with underlying commutative group $\mathbb{Z}/N\mathbb{Z}$.*

(iii) *$H(\mathcal{O}, N)$ holds.*

Proof. (i)⟺(ii) Step 1: Let Λ be a free, rank 2 $\mathbb{Z}$-module, and let Λ' be a $\mathbb{Z}$-submodule of Λ containing $N\Lambda$. By the structure theory of modules over a PID, there is a $\mathbb{Z}$-basis e_1, e_2 for Λ and positive integers $a \mid b$ such that ae_1, be_2 is a $\mathbb{Z}$-basis for Λ'. Thus

$$\Lambda/\Lambda' \cong \mathbb{Z}/a\mathbb{Z} \oplus \mathbb{Z}/b\mathbb{Z}, \qquad \Lambda'/N\Lambda \cong \mathbb{Z}/(N/b)\mathbb{Z} \oplus \mathbb{Z}/(N/a)\mathbb{Z}.$$

It follows that $\Lambda/\Lambda' \cong \mathbb{Z}/N\mathbb{Z} \iff \Lambda'/N\Lambda \cong \mathbb{Z}/N\mathbb{Z}$.

Step 2: If I is an ideal of $\mathcal{O}$ with $\mathcal{O}/I \cong \mathbb{Z}/N\mathbb{Z}$, then $I \supset N\mathcal{O}$, so $I/N\mathcal{O} \cong \mathbb{Z}/N\mathbb{Z}$ by Step 1. Let M be an $\mathcal{O}$-submodule of $\mathcal{O}/N\mathcal{O}$ with underlying $\mathbb{Z}$-module $\mathbb{Z}/N\mathbb{Z}$. Then $M = I/N\mathcal{O}$ for an ideal I of $\mathcal{O}$, and by Step 1 we have $\mathcal{O}/I \cong \mathbb{Z}/N\mathbb{Z}$.

(ii)⟹(iii) Let $P \in \mathcal{O}/N\mathcal{O}$ have order N such that the subgroup generated by P is an $\mathcal{O}$-submodule. For all $g \in C_N(\mathcal{O})$, $gP = a_g P$ for $a_g \in (\mathbb{Z}/N\mathbb{Z})^\times$. Conversely, since $C_N(\mathcal{O})$ contains all scalar matrices, the orbit of $C_N(\mathcal{O})$ on P has size $\varphi(N)$.

(iii)⟹(ii) Case 1: Suppose $2 \nmid N$ or $\left(\frac{\Delta}{2}\right) \neq 1$. Let $P \in \mathcal{O}/N\mathcal{O}$ be a point of order N with $C_N(\mathcal{O})$-orbit of size $\varphi(N)$. There is a $\mathbb{Z}/N\mathbb{Z}$-basis e_1, e_2 of $\mathcal{O}/N\mathcal{O}$ with $e_1 = P$, and our hypothesis gives that with respect to this basis $C_N(\mathcal{O})$ lies in the subalgebra

$$\left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \;\middle|\; a, b, d \in \mathbb{Z}/N\mathbb{Z} \right\}$$

of upper triangular matrices. By Lemma 6.13, $\mathcal{O}/N\mathcal{O}$ also lies in the subalgebra of upper triangular matrices, and thus $\langle P \rangle$ is an $\mathcal{O}$-stable submodule with underlying $\mathbb{Z}$-module $\mathbb{Z}/N\mathbb{Z}$.

[1] Since $\#B \geq \#C_{2^a}(\mathcal{O}) + 1 = 2^{2a-2} + 1 > 2^{2a-2}$, in fact we have $B = \mathcal{B}$.

Case 2: Suppose $2 \mid N$ and $\left(\frac{\Delta}{2}\right) = 1$, and write $N = 2^a N'$ with $2 \nmid N'$. By Lemma 6.12 and the equivalence of (i) and (ii), there is an ideal I_1 in $\mathcal{O}$ with $\mathcal{O}/I_1 \cong \mathbb{Z}/2^a\mathbb{Z}$, and by Case 1 there is an ideal I_2 in $\mathcal{O}$ with $\mathcal{O}/I_2 \cong \mathbb{Z}/N'\mathbb{Z}$. By the Chinese remainder theorem, we have $\mathcal{O}/I_1 I_2 \cong \mathbb{Z}/N\mathbb{Z}$. Since (i) $\iff$ (ii), this suffices. □

Theorem 6.15. *Let $\mathcal{O}$ be an order of discriminant Δ, and let $N \in \mathbb{Z}^+$. The following are equivalent:*

(i) *$H(\mathcal{O}, N)$ holds.*

(ii) *Δ is a square in $\mathbb{Z}/4N\mathbb{Z}$.*

Proof. Using Lemma 6.11, we reduce to the case in which $N = \ell^a$ is a power of a prime number ℓ.

Case 1 (ℓ is odd): Since $\gcd(4, \ell^a) = 1$, we may put $D = \frac{\Delta}{4} \in \mathbb{Z}/\ell^a\mathbb{Z}$. Then Δ is a square in $\mathbb{Z}/4\ell^a\mathbb{Z}$ if and only if D is a square in $\mathbb{Z}/\ell^a\mathbb{Z}$, and

$$\mathcal{O}/\ell^a\mathcal{O} \cong (\mathbb{Z}/\ell^a\mathbb{Z})[t]/(t^2 - D). \tag{5}$$

If there is $s \in \mathbb{Z}/\ell^a\mathbb{Z}$ such that $D = s^2$, then

$$\mathcal{O}/\ell^a\mathcal{O} \cong (\mathbb{Z}/\ell^a\mathbb{Z})[t]/((t+s)(t-s)),$$

so if I is the ideal $\langle t+s, \ell^a \rangle$ of $\mathcal{O}$, then $\mathcal{O}/I \cong \mathbb{Z}/\ell^a\mathbb{Z}$. By Lemma 6.14, $H(\mathcal{O}, \ell^a)$ holds. Conversely, suppose $H(\mathcal{O}, \ell^a)$ holds, so by Lemma 6.14 there is an ideal I of $\mathcal{O}$ with $\mathcal{O}/I \cong \mathbb{Z}/\ell^a\mathbb{Z}$. Since $\ell^a \in I$, we may regard I as an ideal of $\mathcal{O}/\ell^a\mathcal{O}$ such that $(\mathcal{O}/\ell^a\mathcal{O})/I \cong \mathbb{Z}/\ell^a\mathbb{Z}$. In other words, we have a $\mathbb{Z}/\ell^a\mathbb{Z}$-algebra homomorphism

$$f : \mathbb{Z}/\ell^a\mathbb{Z}[t]/(t^2 - D) \to \mathbb{Z}/\ell^a\mathbb{Z}.$$

Then $f(t)^2 = D \in \mathbb{Z}/\ell^a\mathbb{Z}$, so D is a square in $\mathbb{Z}/\ell^a\mathbb{Z}$.

Case 2 ($\ell = 2$, Δ is odd): Here, $\left(\frac{\Delta}{\ell}\right) = \pm 1$.

- If $\left(\frac{\Delta}{\ell}\right) = 1$, then $\Delta \equiv 1 \pmod 8$; by Hensel's lemma, Δ is a square in $\mathbb{Z}/\ell^a\mathbb{Z}$. On the other hand, by Lemmas 6.12(a) and 6.14, $H(\mathcal{O}, \ell^a)$ holds.
- If $\left(\frac{\Delta}{\ell}\right) = -1$, then $\Delta \equiv 5 \pmod 8$, so Δ is not a square modulo 8 and thus not a square modulo $4 \cdot 2^a$. On the other hand, by Lemma 6.12(b), $H(\mathcal{O}, \ell^a)$ does not hold.

Case 3 ($\ell = 2$, Δ is even): Again we may put $D = \frac{\Delta}{4} \in \mathbb{Z}/\ell^a\mathbb{Z}$, and again (5) holds. The argument of Case 1 shows that $H(\mathcal{O}, \ell^a)$ holds if and only if D is a square modulo $\mathbb{Z}/\ell^a\mathbb{Z}$ if and only if Δ is a square modulo $\mathbb{Z}/4\ell^a\mathbb{Z}$. □

Proposition 6.16. *Let $\mathcal{O}$ be an order, and let $N \in \mathbb{Z}^+$. The following are equivalent:*

(i) *$C_N(\mathcal{O})$ acts simply transitively on order N elements of $\mathcal{O}/N\mathcal{O}$.*

(ii) *$C_N(\mathcal{O})$ acts transitively on order N elements of $\mathcal{O}/N\mathcal{O}$.*

(iii) *For all primes $\ell \mid N$ we have $\left(\frac{\Delta}{\ell}\right) = -1$.*

Proof. As usual, we may assume $N = \ell^a$ is a prime power. Certainly (i) $\implies$ (ii).

(ii) $\implies$ (iii): We have

$$\#C_{\ell^a}(\mathcal{O}) = \ell^{2a-2}(\ell-1)\left(\ell - \left(\frac{\Delta}{\ell}\right)\right),$$

whereas the number of elements of order ℓ^a in $\mathcal{O}/\ell^a\mathcal{O}$ is

$$N(\mathcal{O}, \ell^a) := \#\mathcal{O}/\ell^a\mathcal{O} - \#\ell\mathcal{O}/\ell^a\mathcal{O} = \ell^{2a-2}(\ell-1)(\ell+1).$$

Transitivity of the action implies $\#C_{\ell^a}(\mathcal{O}) \geq N(\mathcal{O}, \ell^a)$, which holds if and only if $\left(\frac{\Delta}{\ell}\right) = -1$.

(iii) $\implies$ (i): Since $\left(\frac{\Delta}{\ell}\right) \neq 0$, we have $\mathcal{O}/\ell^a\mathcal{O} \cong \mathcal{O}_K/\ell^a\mathcal{O}_K$, and thus also $C_{\ell^a}(\mathcal{O}) = (\mathcal{O}/\ell^a\mathcal{O})^\times \cong C_{\ell^a}(\mathcal{O}_K)$. Thus $\mathcal{O}/\ell^a\mathcal{O}$ is a finite local principal ring with maximal ideal $\mathfrak{m} = \langle \ell \rangle$ and unit group $C_{\ell^a}(\mathcal{O}) = \mathcal{O}/\ell^a\mathcal{O} \setminus \mathfrak{m}$. The set of order ℓ^a elements of $\mathcal{O}/\ell^a\mathcal{O}$ is $\mathcal{O}/\ell^a\mathcal{O} \setminus \mathfrak{m} = C_{\ell^a}(\mathcal{O})$, so the action of the unit group $C_{\ell^a}(\mathcal{O})$ on this set is the action of $C_{\ell^a}(\mathcal{O})$ on itself, which is simply transitive. □

Corollary 6.17. *Let $\mathcal{O}$ an order of conductor $\mathfrak{f}$. Let $N = \prod_{i=1}^r \ell_i^{a_i} \in \mathbb{Z}^+$ be such that $\left(\frac{\Delta}{\ell_i}\right) = -1$ for all i. Let F be a number field, and let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve such that $E(F)$ has a point of order N. Then*

$$\#\overline{C_N(\mathcal{O})} \mid [FK : K(\mathfrak{f})]. \tag{6}$$

Further for all $\mathcal{O}$ and N satisfying the above conditions, equality can occur in (6).

Proof. Replace F by FK; then $F \supset K(\mathfrak{f})$. By Proposition 6.16, $C_N(\mathcal{O})$ acts transitively on order N elements of $\mathcal{O}/N\mathcal{O}$, so the $\mathcal{O}$-submodule generated by any one of them is $\mathcal{O}/N\mathcal{O}$. Thus the existence of one F-rational point of order N implies that ρ_N is trivial. Applying Theorem 1.4 gives (6). That equality can occur follows from Corollary 1.7. □

6F. *Torsion over $K(j)$: Part I.* Let $\mathcal{O}$ be an order of discriminant $\Delta = \mathfrak{f}^2\Delta_K$. We will give a complete classification of the possible torsion subgroups of $\mathcal{O}$-CM elliptic curves $E_{/K(\mathfrak{f})}$. In this section we will treat the cases $\Delta \neq -3, -4$. For the remaining cases we will make use of Theorem 7.2, so we will come back to those cases in Section 7E.

If $E(K(\mathfrak{f}))$ has a point of order N, then since $[C_N(\mathcal{O}) : \rho_N(\mathfrak{g}_{K(\mathfrak{f})})] \mid \#\mathcal{O}^\times$, there must be some $P \in \mathcal{O}/N\mathcal{O}$ of order N with a $C_N(\mathcal{O})$-orbit of order dividing $\#\mathcal{O}^\times$.

- By Theorem 6.2, if $E(K(\mathfrak{f}))$ has a point of order N, then $\varphi(N) \mid 2$, so

$$N \in \{1, 2, 3, 4, 6\}.$$

- Lemma 2.2(b) implies that for all $N \geq 3$, we have $\#C_N(\mathcal{O}) \geq 4$ (equality holds if $N = 3$ and $\Delta \equiv 1 \pmod 3$). By Theorem 1.4 we cannot have $E[N] = E[N](K(\mathfrak{f}))$.

Thus $E(K(\mathfrak{f}))[\text{tors}]$ is isomorphic to one of the groups in the following list:

$$\{e\},\ \mathbb{Z}/2\mathbb{Z},\ \mathbb{Z}/3\mathbb{Z},\ \mathbb{Z}/4\mathbb{Z},\ \mathbb{Z}/6\mathbb{Z},\ \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z},\ \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/4\mathbb{Z},\ \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}.$$

We will show that all of these groups occur.

Points of order 2. By Theorem 1.4, $E(K(\mathfrak{f}))[2]$ has order 4 if 2 splits in $\mathcal{O}$, order 2 if 2 ramifies in $\mathcal{O}$ and order 1 if 2 is inert in $\mathcal{O}$. Thus:

$$E(K(\mathfrak{f}))[2] \cong \begin{cases} \{e\}, & \Delta \equiv 5 \pmod 8, \\ \mathbb{Z}/2\mathbb{Z}, & \Delta \equiv 0 \pmod 4, \\ \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}, & \Delta \equiv 1 \pmod 8. \end{cases}$$

Points of order 3, 4, or 6. Let $E_{/K(\mathfrak{f})}$ be any $\mathcal{O}$-CM elliptic curve. We claim that for $N \in \{3, 4, 6\}$, there is a quadratic twist E^D of E such that $E^D(K(\mathfrak{f}))$ has a point of order N if and only if $H(\mathcal{O}, N)$ holds. Indeed, as above, since the index of the mod N Galois representation in $C_N(\mathcal{O})$ divides 2, if some $E^D(K(\mathfrak{f}))$ has a point of order N, then $\mathcal{O}/N\mathcal{O}$ has a point of order N with a $C_N(\mathcal{O})$-orbit of size 2. Since $\varphi(N) = 2$, there is a Cartan orbit of size 2 if and only if $H(\mathcal{O}, N)$ holds. Conversely, if $H(\mathcal{O}, N)$ holds then there is a point of order N with a $C_N(\mathcal{O})$-orbit of size 2, hence on some quadratic twist E^D we have an F-rational point of order N. Applying Theorem 6.15, we get:

- Some $\mathcal{O}$-CM $E_{/K(\mathfrak{f})}$ has a point of order 3 if and only if $\Delta \equiv 0, 1 \pmod 3$.
- Some $\mathcal{O}$-CM $E_{/K(\mathfrak{f})}$ has a point of order 4 if and only if $\Delta \equiv 0, 1, 4, 9 \pmod{16}$.
- Some $\mathcal{O}$-CM $E_{/K(\mathfrak{f})}$ has a point of order 6 if and only if $\Delta \equiv 0, 1, 2, 9, 12, 16 \pmod{24}$.

Because the only full N-torsion we can have is full 2-torsion, and 2-torsion is invariant under quadratic twists, we immediately deduce the complete answer in all cases.

- If $\Delta \equiv 0 \pmod{48}$, then there are twists E_1, E_2, E_3 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/4\mathbb{Z}, \quad E_3(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/6\mathbb{Z}.$$

- If $\Delta \equiv 1, 9, 25, 33 \pmod{48}$ then there are twists E_1, E_2 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}.$$

- If $\Delta \equiv 4, 16, 36 \pmod{48}$, then there are twists E_1, E_2, E_3 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f})) \cong \mathbb{Z}/4\mathbb{Z}, \quad E_3(K(\mathfrak{f})) \cong \mathbb{Z}/6\mathbb{Z}.$$

- If $\Delta \equiv 5, 29 \pmod{48}$, then $E(K(\mathfrak{f}))[\text{tors}] = \{e\}$.
- If $\Delta \equiv 8, 44 \pmod{48}$, then $E(K(\mathfrak{f}))[\text{tors}] = \mathbb{Z}/2\mathbb{Z}$.
- If $\Delta \equiv 12, 24, 28, 40 \pmod{48}$, then there are twists E_1, E_2 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f})) \cong \mathbb{Z}/6\mathbb{Z}.$$

- If $\Delta \equiv 13, 21, 37, 45 \pmod{48}$, then there are twists E_1, E_2 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] = \{e\}, \quad E_2(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/3\mathbb{Z}.$$

- If $\Delta \equiv 17, 41 \pmod{48}$, then there are twists E_1, E_2 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}.$$

- If $\Delta \equiv 20, 32 \pmod{48}$, then there are twists E_1, E_2 of E with

$$E_1(K(\mathfrak{f}))[\text{tors}] \cong \mathbb{Z}/2\mathbb{Z}, \quad E_2(K(\mathfrak{f})) \cong \mathbb{Z}/4\mathbb{Z}.$$

6G. *Isogenies over $K(j)$: Part I.*

Theorem 6.18. *Let $\mathcal{O}$ be an order of discriminant $\Delta = \mathfrak{f}^2\Delta_K$, and let $N \in \mathbb{Z}^+$.*

- *If $\Delta \neq -3, -4$, then there is an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ with a $K(\mathfrak{f})$-rational cyclic N-isogeny if and only if Δ is a square in $\mathbb{Z}/4N\mathbb{Z}$.*

- *If $\Delta = -4$, then there is an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ with a $K(\mathfrak{f})$-rational cyclic N-isogeny if and only if N is of the form $2^\epsilon \ell_1^{a_1} \cdots \ell_r^{a_r}$ for primes $\ell_i \equiv 1 \pmod 4$ and $\epsilon, a_1, \ldots, a_r \in \mathbb{N}$ with $\epsilon \leq 2$.*

- *If $\Delta = -3$, then there is an $\mathcal{O}$-CM elliptic curve $E_{/K(\mathfrak{f})}$ with a $K(\mathfrak{f})$-rational cyclic N-isogeny if and only if N is of the form $2^\epsilon 3^a \ell_1^{a_1} \cdots \ell_r^{a_r}$ for primes $\ell_i \equiv 1 \pmod 3$, $\epsilon, a, a_1, \ldots, a_r \in \mathbb{N}$ with $(\epsilon, a) \in \{(0,0),\ (0,1),\ (0,2),\ (1,0),\ (1,1)\}$.*

Proof. Step 1: Let $E_{/K(\mathfrak{f})}$ be an $\mathcal{O}$-CM elliptic curve. If Δ is a square in $\mathbb{Z}/4N\mathbb{Z}$, then by Theorem 6.15 there is a point P of order N in $\mathcal{O}/N\mathcal{O}$ such that $C = \langle P \rangle$ is invariant under $C_N(\mathcal{O})$, so C is $\mathfrak{g}_{K(\mathfrak{f})}$-stable and $E \to E/C$ is a cyclic N-isogeny. If $\Delta \notin \{-4, -3\}$, then the projective Galois representation $\mathbb{P}\rho_N : \mathfrak{g}_{K(\mathfrak{f})} \to C_N(\mathcal{O})/(\mathbb{Z}/N\mathbb{Z})^\times$ is a quotient of the reduced Galois representation, hence surjective. So $K(\mathfrak{f})$-rational cyclic N-isogenies correspond to $C_N(\mathcal{O})$-orbits on $\mathcal{O}/N\mathcal{O}$ of size $\varphi(N)$, which by Theorem 6.15 exist if and only if Δ is a square in $\mathbb{Z}/4N\mathbb{Z}$.

Step 2: If $\Delta \in \{-4, -3\}$, then as above the condition that Δ is a square modulo $4N$ is sufficient for the existence of a $K(\mathfrak{f})$-rational cyclic N-isogeny, but it is no longer clear that it is necessary, and in both cases it turns out not to be. The complete analysis will make use of Theorem 7.2, so we defer the end of the proof until Section 7F. □

7. The Torsion Degree Theorem

7A. ***Statement and preliminary reduction.*** Throughout this section $\mathcal{O}$ denotes an order of conductor $\mathfrak{f}$ and discriminant $\Delta = \mathfrak{f}^2 \Delta_K$.

For $N \in \mathbb{Z}^{\geq 2}$, let $\widetilde{T}(\mathcal{O}, N)$ be the least size of an orbit of $C_N(\mathcal{O})$ on an order N point of $\mathcal{O}/N\mathcal{O}$.

Lemma 7.1. *We have* $\widetilde{T}(\mathcal{O}, 2) = \begin{cases} 1 & \text{if } \left(\frac{\Delta}{2}\right) \neq -1, \\ 3 & \text{if } \left(\frac{\Delta}{2}\right) = -1. \end{cases}$

Proof. By Theorem 6.15, we have $\left(\frac{\Delta}{2}\right) \neq -1$ if and only if there is a $C_2(\mathcal{O})$-orbit of size $\varphi(2) = 1$ on $\mathcal{O}/2\mathcal{O}$ if and only if $\widetilde{T}(\mathcal{O}, 2) = 1$. In the remaining case, $\left(\frac{\Delta}{2}\right) = -1$, we have $\#C_2(\mathcal{O}) = 3$ and no orbit of size 1, hence $\widetilde{T}(\mathcal{O}, 2) = 3$. □

Theorem 7.2 (Torsion Degree Theorem). *Let $\mathcal{O}$ be an order of conductor $\mathfrak{f}$, and let $N \in \mathbb{Z}^{\geq 3}$.*

(a) *There is $T(\mathcal{O}, N) \in \mathbb{Z}^+$ such that*
- *if $F \supset K(\mathfrak{f})$ is a number field and $E_{/F}$ is an $\mathcal{O}$-CM elliptic curve with an F-rational point of order N, then $T(\mathcal{O}, N) \mid [F : K(\mathfrak{f})]$, and*
- *there is a number field $F \supset K(\mathfrak{f})$ with $[F : K(\mathfrak{f})] = T(\mathcal{O}, N)$ and an $\mathcal{O}$-CM elliptic curve $E_{/F}$ with an F-rational point of order N.*

(b) *If $(\Delta, N) = (-3, 3)$, then $T(\mathcal{O}, N) = 1$.*

(c) *Suppose $(\Delta, N) \neq (-3, 3)$. Let $N = \ell_1^{a_1} \cdots \ell_r^{a_r}$ be the prime power decomposition of N. Then*
$$T(\mathcal{O}, N) = \frac{\prod_{i=1}^r \widetilde{T}(\mathcal{O}, \ell_i^{a_i})}{\#\mathcal{O}^\times}.$$

(d) *If $\ell^a = 2$, then $\widetilde{T}(\mathcal{O}, \ell^a) = 2$ is computed in Lemma 7.1. If $\ell^a > 2$, then $\widetilde{T}(\mathcal{O}, \ell^a)$ is as follows, where $k = \operatorname{ord}_\ell(\mathfrak{f})$:*

(1) *If $\ell \nmid \mathfrak{f}$, then* $\widetilde{T}(\mathcal{O}, \ell^a) = \begin{cases} \ell^{a-1}(\ell - 1) & \text{if } \left(\frac{\Delta}{\ell}\right) = 1, \\ \ell^{2a-2}(\ell - 1) & \text{if } \left(\frac{\Delta}{\ell}\right) = 0, \\ \ell^{2a-2}(\ell^2 - 1) & \text{if } \left(\frac{\Delta}{\ell}\right) = -1. \end{cases}$

(2) *If $\ell \mid \mathfrak{f}$, then* $\widetilde{T}(\mathcal{O}, \ell^a) = \begin{cases} \ell^{a-1}(\ell - 1) & \text{if } \left(\frac{\Delta_K}{\ell}\right) = 1, \\ \ell^{a-1}(\ell - 1) & \text{if } \left(\frac{\Delta_K}{\ell}\right) = -1 \text{ and } a \leq 2k, \\ \ell^{2a-2k-1}(\ell - 1) & \text{if } \left(\frac{\Delta_K}{\ell}\right) = -1 \text{ and } a > 2k, \\ \ell^{a-1}(\ell - 1) & \text{if } \left(\frac{\Delta_K}{\ell}\right) = 0 \text{ and } a \leq 2k+1, \\ \ell^{2a-2k-2}(\ell - 1) & \text{if } \left(\frac{\Delta_K}{\ell}\right) = 0 \text{ and } a > 2k+1. \end{cases}$

Remark 7.3. The case $N = 2$ is excluded because of the somewhat anomalous behavior of 2-torsion. But it is easy to see that Theorem 7.2(a) remains true when

$N = 2$, and, moreover:

- If $\Delta \in \{-4, -3\}$ then $T(\mathcal{O}, 2) = 1$.
- Otherwise, $T(\mathcal{O}, 2) = \begin{cases} 1 & \text{if } \left(\frac{\Delta}{2}\right) \neq -1, \\ 3 & \text{if } \left(\frac{\Delta}{2}\right) = -1. \end{cases}$

Let $F \supset K(\mathfrak{f})$ be a number field, and let $E_{/F}$ be an $\mathcal{O}$-CM elliptic curve. As usual, we choose an embedding $F \hookrightarrow \mathbb{C}$ such that $j(E) = j(\mathbb{C}/\mathcal{O})$. Let $P \in E[\text{tors}]$ have order N. We call the field

$$K(\mathfrak{f})(\mathfrak{h}(P))$$

the *field of moduli* of P. It is independent of the chosen model of $E_{/F}$, and there exists an elliptic curve $E'_{/K(\mathfrak{f})(\mathfrak{h}(P))}$ with an isomorphism $\psi : E \to E'$ such that $\psi(P)$ is $K(\mathfrak{f})(\mathfrak{h}(P))$-rational. Further, the pair (E, P) induces a closed point $\mathcal{P}$ on the modular curve $X_1(N)_{/K}$, and $K(\mathfrak{f})(\mathfrak{h}(P))$ is the residue field $K(\mathcal{P})$. Theorem 7.2 concerns the degree $[K(\mathfrak{f})(\mathfrak{h}(P)) : K(\mathfrak{f})]$. Our setup shows that it is no loss of generality to assume $F = K(\mathfrak{f})$.

Let $q_N : \mathcal{O} \to \mathcal{O}/N\mathcal{O}$ be the natural map, and let $q_N^\times : \mathcal{O}^\times \to C_N(\mathcal{O})$ be the induced map on unit groups. As in the introduction, we define the *reduced mod N Cartan subgroup*:

$$\overline{C_N(\mathcal{O})} = C_N(\mathcal{O})/q_N(\mathcal{O}^\times).$$

Let $\overline{E[N]}$ be the set of $\mathcal{O}^\times$-orbits on $E[N]$. Then the action of $C_N(\mathcal{O})$ on $E[N]$ induces an action of $\overline{C_N(\mathcal{O})}$ on $\overline{E[N]}$. The field of moduli $K(\mathfrak{f})(\mathfrak{h}(P))$ depends only on the image $\overline{P}$ of P in $\overline{E[N]}$. By Theorem 1.4, the composite homomorphism

$$\mathfrak{g}_F \xrightarrow{\rho_{E,N}} C_N(\mathcal{O}) \to \overline{C_N(\mathcal{O})}$$

is surjective (and model-independent). Let $H_{\overline{P}} = \{g \in \overline{C_N(\mathcal{O})} \mid g\overline{P} = \overline{P}\}$. It follows that

$$\operatorname{Aut}(K(\mathfrak{f})(\mathfrak{h}(P))/K(\mathfrak{f})) \cong \overline{C_N(\mathcal{O})}/H_{\overline{P}}.$$

Thus $[K(\mathfrak{f})(\mathfrak{h}(P)) : K(\mathfrak{f})]$ is the size of the orbit of the reduced Cartan subgroup $\overline{C_N(\mathcal{O})}$ on $\overline{P}$. (As we will see, in almost every case this is the size of the orbit of $C_N(\mathcal{O})$ on P divided by $\#\mathcal{O}^\times$.) This reduces the proof of Theorem 7.2 to a purely algebraic problem.

7B. ***Generalities.*** For an order N point $P \in \mathcal{O}/N\mathcal{O}$, let $M_P = \{xP \mid x \in \mathcal{O}\}$ be the cyclic $\mathcal{O}$-submodule of $\mathcal{O}/N\mathcal{O}$ generated by P. If we put $I_P = \{x \in \mathcal{O} \mid xP = 0\}$, then we have

$$M_P \cong_{\mathcal{O}} \mathcal{O}/I_P.$$

The isomorphism is canonical and determined by mapping $P \in M_P$ to $1 + I_P \in \mathcal{O}/I_P$.

Lemma 7.4. (*a*) *With notation as above, let*

$$S(I_P) = \{g \in C_N(\mathcal{O}) \mid g \equiv 1 \pmod{I_P}\}.$$

Then with respect to the $C_N(\mathcal{O})$-action, $S(I_P)$ is the stabilizer of P, so as a $C_N(\mathcal{O})$-set the orbit of $C_N(\mathcal{O})$ on P is isomorphic to $C_N(\mathcal{O})/S(I_P)$.

(*b*) *Further, there is a canonical isomorphism of groups $C_N(\mathcal{O})/S(I_P) \xrightarrow{\sim} (\mathcal{O}/I_P)^\times$.*

Proof. (a) For $g \in C_N(\mathcal{O})$, we have $gP = P \iff (g-1)P = 0 \iff (g-1) \in I_P$, giving the first assertion. The orbit stabilizer theorem gives the second assertion.

(b) The ring homomorphism $f : \mathcal{O}/N\mathcal{O} \to \mathcal{O}/I$ induces a homomorphism on unit groups $f^\times : C_N(\mathcal{O}) \to (\mathcal{O}/I_P)^\times$, with kernel $S(I_P)$. Since $\mathcal{O}/N\mathcal{O}$ has finitely many maximal ideals, $f^\times$ is surjective [Clark 2015, Theorem 4.32]. □

Lemma 7.5. *There is a positive integer $M \mid N$ such that*

$$\mathcal{O}/I_P \cong_{\mathbb{Z}} \mathbb{Z}/N\mathbb{Z} \oplus \mathbb{Z}/M\mathbb{Z}.$$

Proof. As a $\mathbb{Z}$-module, $\mathcal{O}/I_P$ is a quotient of $\mathcal{O}/N\mathcal{O} \cong_{\mathbb{Z}} \mathbb{Z}/N\mathbb{Z} \oplus \mathbb{Z}/N\mathbb{Z}$, so

$$\mathcal{O}/I_P \cong_{\mathbb{Z}} \mathbb{Z}/N'\mathbb{Z} \oplus \mathbb{Z}/M\mathbb{Z}$$

with $M \mid N' \mid N$. Since P has order N in $(\mathcal{O}/I_P, +)$, we have $N' = N$. □

The following result computes the size of the reduced Cartan orbit on an order N point of $\mathcal{O}/N\mathcal{O}$ in terms of the size of the Cartan orbit. We recall that we have assumed $N \geq 3$.

Lemma 7.6. (a) *Suppose $(\Delta, N) \neq (-3, 3)$, and let $P \in \mathcal{O}/N\mathcal{O}$ have order N. Then the orbit of $C_N(\mathcal{O})$ on P has size $\#\mathcal{O}^\times$ times the size of the orbit of $\overline{C_N(\mathcal{O})}$ on $\overline{P}$.*

(b) *Suppose $(\Delta, N) = (-3, 3)$. Then the order 3 points of $\mathcal{O}/3\mathcal{O}$ lie in two orbits under $C_3(\mathcal{O})$: one of size 2 and one of size 6. The corresponding reduced Cartan orbits each have size 1.*

Proof. (a) The Cartan orbit has size $\#(\mathcal{O}/I_P)^\times$, and the reduced Cartan orbit is smaller by a factor of the cardinality of the image of $\mathcal{O}^\times \to (\mathcal{O}/I_P)^\times$.

- Suppose $\Delta \notin \{-4, -3\}$. Then $\mathcal{O}^\times = \{\pm 1\}$, and since $N \geq 3$, we have $-1 \not\equiv 1 \pmod{I_P}$.
- Suppose $\Delta = -4$. Since $I_P \not\supset (2)$, by Lemma 2.9 the group $U_{I_P}(K)$ is trivial, and thus the map $\mathcal{O}^\times \to (\mathcal{O}/I_P)^\times$ is injective.
- Suppose $\Delta = -3$. By assumption, $N \geq 4$, so $I_P \nmid (\zeta_3 - 1)$ and the map $\mathcal{O}^\times \to (\mathcal{O}/I_P)^\times$ is injective.

(b) The assertion about Cartan orbits is a case of [Clark et al. 2013, Lemma 19]. (Another proof will be given in the next section.) The fact that both reduced Cartan orbits have size 1 follows from the already established fact that there is an $\mathcal{O}$-CM $E_{/\mathbb{Q}(\sqrt{-3})}$ with full 3-torsion. □

In view of Lemma 7.6, to prove Theorem 7.2 it suffices to compute the least size of an orbit of $C_N(\mathcal{O})$ on an order N point of $\mathcal{O}/N\mathcal{O}$ and show that this divides the size of every such orbit. The following result further reduce us to the case of N a prime power.

Proposition 7.7. *Let $N \geq 2$ have prime power decomposition $N = \ell_1^{a_1} \cdots \ell_r^{a_r}$. Let $P \in \mathcal{O}/N\mathcal{O}$ have order N, and let $I_P = \operatorname{ann} P$. For $1 \leq i \leq r$, let $P_i = N/\ell_i^{a_i} P$, and let $I_{P_i} = \operatorname{ann} P_i$. Then:*

(a) *The ideals $I_{P_1}, \ldots, I_{P_r}$ are pairwise comaximal, so $I_{P_i} + I_{P_j} = \mathcal{O}$ for all $i \neq j$.*

(b) *We have $I_P = I_{P_1} \cdots I_{P_r}$.*

(c) *We have a canonical isomorphism of rings*

$$\mathcal{O}/I_P \xrightarrow{\sim} \prod_{i=1}^{r} \mathcal{O}/I_{P_i}$$

which induces a canonical isomorphism of unit groups

$$(\mathcal{O}/I_P)^\times \xrightarrow{\sim} \prod_{i=1}^{r} (\mathcal{O}/I_{P_i})^\times.$$

(d) *The Cartan orbit of P is isomorphic, as a $C_N(\mathcal{O})$-set, to the direct product of the $C_{\ell_i^{a_i}}(\mathcal{O})$-orbits of the P_i's.*

Proof. (a) For $1 \leq i \leq r$, we have $(\mathcal{O}/I_{P_i}, +) \cong \mathbb{Z}/\ell_i^{a_i}\mathbb{Z} \oplus \mathbb{Z}/\ell_i^{b_i}\mathbb{Z}$ with $0 \leq b_i \leq a_i$; in particular it is an ℓ_i-group. Thus for $i \neq j$, $(\mathcal{O}/(I_i + I_j), +)$ is a homomorphic image of an ℓ_i-group and an ℓ_j-group, so it is trivial.

(b) By the Chinese remainder theorem, we have $I_{P_1} \cdots I_{P_n} = \bigcap_{i=1}^{n} I_{P_i}$. Since P_i is a multiple of P, we have $I_P \subset I_{P_i}$ for all i, and thus $I_P \subset \bigcap_{i=1}^{r} I_{P_i}$. Conversely, choose $y_1, \ldots, y_r \in \mathbb{Z}$ such that $\sum_{i=1}^{r} y_i N/\ell_i^{a_i} = 1$. If $x \in \bigcap_{i=1}^{r} I_{P_i}$ then $xN/\ell_i^{a_i} P = 0$ for all i, hence

$$0 = \sum_{i=1}^{r} y_i \frac{N}{\ell_i^{a_i}} x P = xP,$$

so $x \in I_P$. Thus $I_P = \bigcap_{i=1}^{n} I_{P_i} = I_{P_1} \cdots I_{P_n}$.

(c) The Chinese remainder theorem gives the first isomorphism; the second follows by passing to unit groups.

(d) Apply Lemma 7.4 and part (c). □

7C. *The case $\ell \nmid \mathfrak{f}$.*

Theorem 7.8. *Let $E_{/K(\mathfrak{f})}$ be an $\mathcal{O}$-CM elliptic curve. Let $\ell^a > 2$ be a prime power such that $\ell \nmid \mathfrak{f}$. We will give the sizes and multiplicities of all orbits of $C_{\ell^a}(\mathcal{O})$ on order ℓ^a points of $\mathcal{O}/\ell^a\mathcal{O}$.*

(a) *If $\left(\frac{\Delta}{\ell}\right) = 1$, there are $2a+1$ orbits: two orbits of size $\ell^{a-1}(\ell-1)$, for all $1 \le i \le a-1$, two orbits of size $\ell^{a+i-2}(\ell-1)^2$, and one orbit of size $\ell^{2a-2}(\ell-1)^2$.*

(b) *If $\left(\frac{\Delta}{\ell}\right) = 0$, there are two orbits: an orbit of size $\ell^{2a-2}(\ell-1)$ and an orbit of size $\ell^{2a-1}(\ell-1)$.*

(c) *If $\left(\frac{\Delta}{\ell}\right) = -1$, there is one orbit, of size $\ell^{2a-2}(\ell^2-1)$.*

Proof. Step 1: Suppose $\mathcal{O} = \mathcal{O}_K$. Then every $\mathcal{O}$-submodule of $E[N]$ is of the form $E[I]$ for an ideal $I \supset N\mathcal{O}$, and $E[I] \cong_{\mathcal{O}} \mathcal{O}/I$: thus every submodule is of the form $M_P = \langle P \rangle_{\mathcal{O}}$ and is determined by its annihilator ideal I_P. Conversely, if $I \supset N\mathcal{O}$ is an ideal, then Lemmas 2.3 and 2.4 give that $E[I]$ is an $\mathcal{O}$-submodule of $E[N]$ with annihilator ideal I.

Split case: $\left(\frac{\Delta}{\ell}\right) = 1$. Here, $\ell\mathcal{O} = \mathfrak{p}_1\mathfrak{p}_2$ for distinct prime ideals $\mathfrak{p}_1, \mathfrak{p}_2$ of norm ℓ. The ideals containing $\ell^a\mathcal{O}$ are precisely $\mathfrak{p}_1^c\mathfrak{p}_2^d$ with $\max(c,d) \le a$. We have ring isomorphisms

$$\mathcal{O}/\mathfrak{p}_1^c\mathfrak{p}_2^d \cong \mathcal{O}/\mathfrak{p}_1^c \times \mathcal{O}/\mathfrak{p}_2^d \cong \mathbb{Z}/\ell^c\mathbb{Z} \times \mathbb{Z}/\ell^d\mathbb{Z},$$

and hence unit group isomorphisms

$$(\mathcal{O}/\mathfrak{p}_1^c\mathfrak{p}_2^d)^\times \cong (\mathcal{O}/\mathfrak{p}_1^c)^\times \times (\mathcal{O}/\mathfrak{p}_2^c)^\times \cong (\mathbb{Z}/\ell^c\mathbb{Z})^\times \times (\mathbb{Z}/\ell^d\mathbb{Z})^\times,$$

so

$$\#(\mathcal{O}/\mathfrak{p}_1^c\mathfrak{p}_2^d)^\times = \varphi(\ell^c)\varphi(\ell^d).$$

To get points of order ℓ^a we impose the condition $\max(c,d) = a$. Thus $\mathcal{O}$-modules generated by the points of order ℓ^a are

$$E[\mathfrak{p}_1^a],\ E[\mathfrak{p}_1^a\mathfrak{p}_2], \ldots, E[\mathfrak{p}_1^a\mathfrak{p}_2^a] = E[\ell^a],\ E[\mathfrak{p}_1^{a-1}\mathfrak{p}_2^a], \ldots, E[\mathfrak{p}_1\mathfrak{p}_2^a],\ E[\mathfrak{p}_2^a].$$

So there are $2a+1$ Cartan orbits, one of size $\varphi(\ell^a)\varphi(\ell^a)$ and, for all $0 \le i \le a-1$, two of size $\varphi(\ell^a)\varphi(\ell^i)$. The smallest orbit size is $\ell^{a-1}(\ell-1)$, and all the other orbit sizes are multiples of it.

Ramified case: $\left(\frac{\Delta}{\ell}\right) = 0$. Then $\ell\mathcal{O} = \mathfrak{p}^2$ for a prime ideal $\mathfrak{p}$ of norm ℓ. For any $b \in \mathbb{Z}^+$, the ring $\mathcal{O}/\mathfrak{p}^b$ is local of order ℓ^b with residue field $\mathbb{Z}/\ell\mathbb{Z}$, so the maximal ideal has size ℓ^{b-1} and thus

$$\#(\mathcal{O}/\mathfrak{p}^b)^\times = \ell^b - \ell^{b-1} = \ell^{b-1}(\ell-1).$$

Since $\mathfrak{p}^2 = (\ell)$, the least $c \in \mathbb{N}$ such that $\ell^c \in \mathfrak{p}^b$ is $c = \lceil \frac{b}{2} \rceil$. It follows that

$$(\mathcal{O}/\mathfrak{p}^b, +) \cong_{\mathbb{Z}} \mathbb{Z}/\ell^{\lceil \frac{b}{2} \rceil}\mathbb{Z} \oplus \mathbb{Z}/\ell^{\lfloor \frac{b}{2} \rfloor}\mathbb{Z}.$$

So the annihilator ideals of points of order ℓ^a in $\mathcal{O}/\ell^a\mathcal{O}$ are precisely $\mathfrak{p}^{2a-1}$ and $\mathfrak{p}^{2a}$. We get two Cartan orbits, one of size $\#(\mathcal{O}/\mathfrak{p}^{2a-1})^\times = \ell^{2a-2}(\ell-1)$ and one of size $\#(\mathcal{O}/\mathfrak{p}^{2a})^\times = \ell^{2a-1}(\ell-1)$. The smallest orbit size is $\ell^{2a-2}(\ell-1)$, and the other orbit size is a multiple of it.

Inert case: $\left(\frac{\Delta}{\ell}\right) = -1$**.** Here, $\ell\mathcal{O}$ is a prime ideal, so the ideals containing $\ell^a\mathcal{O}$ are precisely $\ell^i\mathcal{O}$ for $i \le a$. Clearly $\mathcal{O}/\ell^i\mathcal{O}$ has exponent ℓ^a if and only if $i = a$, so the $\mathcal{O}$-module generated by any point of order ℓ^a is $E[\ell^a]$. There is a single Cartan orbit, of size $\#(\mathcal{O}/\ell^a\mathcal{O})^\times = \varphi_K(\ell^a) = \ell^{2a-2}(\ell^2-1)$.

Step 2: Now let $\mathcal{O}$ be an order with $\ell \nmid \mathfrak{f}$. The natural maps $\mathcal{O}/\ell^a\mathcal{O} \to \mathcal{O}_K/\ell^a\mathcal{O}_K$ and $C_{\ell^a}(\mathcal{O}) \to C_{\ell^a}(\mathcal{O}_K)$ are isomorphisms, so the sizes and multiplicities of orbits carry over from $\mathcal{O}_K$ to $\mathcal{O}$. □

7D. ***The case $\ell \mid \mathfrak{f}$.*** Suppose $\ell \mid \mathfrak{f}$. The ring $\mathcal{O}/\ell\mathcal{O}$ is isomorphic to $\mathbb{Z}/\ell\mathbb{Z}[\epsilon]/(\epsilon^2)$ — as one sees, e.g., using the explicit representation of (3) — and is thus a local Artinian ring with maximal ideal $\mathfrak{p}$, say, and residue field $\mathbb{Z}/\ell\mathbb{Z}$. Because $[\mathfrak{p} : \ell O] = \ell$, the only proper nonzero $\mathcal{O}$-submodule of $\mathcal{O}/\ell\mathcal{O}$ is $\mathfrak{p}/\ell$. Thus there are two Cartan orbits on the order ℓ elements of $\mathcal{O}/\ell\mathcal{O}$: one of order $\ell-1$ and one of order $\ell^2-\ell = \#(\mathcal{O}/\ell\mathcal{O})^\times$.

For all $a \in \mathbb{Z}^+$, the ring $\mathcal{O}/\ell^a\mathcal{O}$ is local — for a maximal ideal $\mathfrak{m}$ of $\mathcal{O}$, we have $\ell^a \in \mathfrak{m} \iff \ell \in \mathfrak{m}$ — with residue field $\mathbb{Z}/\ell\mathbb{Z}$. In turn it follows that for any order ℓ^a point $P \in \mathcal{O}/\ell^a\mathcal{O}$ and $I_P = \{x \in \mathcal{O} \mid xP = 0\}$, the ring $\mathcal{O}/I_P$ is local with residue field $\mathbb{Z}/\ell\mathbb{Z}$. By Lemma 7.5, we may write

$$M_P = \mathcal{O}/I_P \cong_{\mathbb{Z}} \mathbb{Z}/\ell^a\mathbb{Z} \oplus \mathbb{Z}/\ell^b\mathbb{Z} \tag{7}$$

for some $0 \le b \le a$, and then

$$\#(\mathcal{O}/I_P)^\times = \#\mathcal{O}/I_P - \frac{\#\mathcal{O}/I_P}{\ell} = \ell^{a+b-1}(\ell-1).$$

So the size of a Cartan orbit on an order ℓ^a element of $\mathcal{O}/\ell^a\mathcal{O}$ is of the form $(\ell-1)\ell^c$ for some $a-1 \le c \le 2a-1$. So in this case it is a priori clear that the minimal size of a Cartan orbit divides the size of all the Cartan orbits. We want to understand how Cartan orbits grow when we lift a point of order ℓ^a to a point of order ℓ^{a+1}. First observe that $x \mapsto \ell x$ gives an $\mathcal{O}$-module isomorphism

$$\mathcal{O}/\ell^a\mathcal{O} \xrightarrow{\sim} \ell\mathcal{O}/\ell^{a+1}\mathcal{O},$$

so we can view $\mathcal{O}/\ell^a\mathcal{O}$ as an $\mathcal{O}$-submodule of $\mathcal{O}/\ell^{a+1}\mathcal{O}$. With P as in (7), let $Q \in \mathcal{O}/\ell^{a+1}\mathcal{O}$ be such that $\ell Q = P$. Put

$$M_Q = \{xQ \mid x \in \mathcal{O}\} \quad \text{and} \quad I_Q = \{x \in \mathcal{O} \mid xQ = 0\},$$

and write

$$M_Q = \mathcal{O}/I_Q \cong_{\mathbb{Z}} \mathbb{Z}/\ell^{a+1}\mathbb{Z} \oplus \mathbb{Z}/\ell^{b'}\mathbb{Z} \tag{8}$$

for $0 \le b' \le a+1$. Because $\ell Q = P$, we have $\ell M_Q = M_P$. Thus we find: if $b = 0$, then $b' \in \{0, 1\}$, whereas if $b \geq 1$ then necessarily $b' = b+1$. So: if the $C_{\ell^a}(\mathcal{O})$-orbit on P has the smallest possible size $\varphi(\ell^a)$, then the $C_{\ell^{a+1}}(\mathcal{O})$-orbit on Q either has size $\varphi(\ell^{a+1})$ or size $\varphi(\ell^{a+2})$ (as we will see shortly, both possibilities can occur), whereas if the $C_{\ell^a}(\mathcal{O})$-orbit on P has size $\varphi(\ell^{a+b}) > \varphi(\ell^a)$, then the $C_{\ell^{a+1}}(\mathcal{O})$-orbit on Q has size $\varphi(\ell^{a+b+2})$: i.e., upon lifting from P to Q the size grows by a factor of ℓ^2.

Since $H(\mathcal{O}, \ell^{a+1})$ implies $H(\mathcal{O}, \ell^a)$, for each fixed ℓ and $\mathcal{O}$ there are two possibilities.

Type I: $H(\mathcal{O}, \ell^a)$ holds for all $a \in \mathbb{Z}^+$.

In Type I, for all $a \in \mathbb{Z}^+$, the least size of a $C_{\ell^a}(\mathcal{O})$-orbit is $\varphi(\ell^a)$.

Type II: There is some $A \in \mathbb{Z}^+$ such that $H(\mathcal{O}, \ell^a)$ holds if and only if $a \leq A$.

In Type II, for $1 \leq a \leq A$, the least size of a $C_{\ell^a}(\mathcal{O})$-orbit is $\varphi(\ell^a)$, but for all $a \geq A$, whenever we lift a point of order ℓ^a to a point of order ℓ^{a+1} the size of the Cartan orbit grows by a factor of ℓ^2, so for all $a > A$ the least size of a $C_{\ell^a}(\mathcal{O})$-orbit is $\ell^{a-A}\varphi(\ell^a)$.

We now determine the smallest size of a $C_{\ell^a}(\mathcal{O})$-orbit on an order ℓ^a point of $\mathcal{O}/\ell^a\mathcal{O}$ by using Theorem 6.15 to determine the type and compute the value of A in Type II.

Case 1: Suppose $\left(\frac{\Delta_K}{\ell}\right) = 1$. Then $H(\mathcal{O}_K, \ell^a)$ holds for all $a \in \mathbb{Z}^+$, so Δ_K is a square modulo $4\ell^a$, hence $\Delta = \mathfrak{f}^2\Delta_K$ is also a square modulo $4\ell^a$, so $H(\mathcal{O}, \ell^a)$ holds, and we are in Type I.

Case 2: Suppose $\left(\frac{\Delta_K}{\ell}\right) = -1$, and put $k = \operatorname{ord}_\ell(\mathfrak{f})$.

- Let $\ell > 2$. If $a \leq 2k$, then $\ell^a \mid \Delta$, so Δ is a square mod ℓ^a and hence also mod $4\ell^a$: thus $H(\mathcal{O}, \ell^a)$ holds. However, if $a = 2k+1$ then we claim $H(\mathcal{O}, \ell^a)$ does not hold. Indeed, suppose there is $s \in \mathbb{Z}$ such that $\Delta = \mathfrak{f}^2\Delta_K \equiv s^2 \pmod{\ell^a}$. Then $\ell^k \mid s$; taking $S = s/\ell^k$ we have $\mathfrak{f}^2/\ell^{2k}\Delta_K \equiv S^2 \pmod{\ell^{a-2k}}$, which implies that Δ_K is a square modulo ℓ, which gives a contradiction. So we are in Type II with $A = 2k$.
- Let $\ell = 2$, and write $\mathfrak{f} = 2^k F$. Suppose $a \leq 2k$. Since $4 \mid \Delta_K - 1$, we have
$$2^{a+2} \mid (2^k F)^2(\Delta_K - 1) = \Delta - (2^k F)^2,$$
so $H(\mathcal{O}, 2^a)$ holds. Suppose $a \geq 2k+1$. If Δ is a square modulo 2^{a+2}, then we find that $\Delta_K \equiv 1 \pmod 8$, so $\left(\frac{\Delta_K}{2}\right) = 1$; this is a contradiction. So we are in Type II with $A = 2k$.

Case 3: Suppose $\left(\frac{\Delta_K}{\ell}\right) = 0$, and put $k = \operatorname{ord}_\ell(\mathfrak{f})$.

- Let $\ell > 2$. If $a \leq 2k+1$, then $\ell^a \mid \Delta$, so Δ is a square mod ℓ^a and hence also mod $4\ell^a$: thus $H(\mathcal{O}, \ell^a)$ holds. However, if $a = 2k+2$ then we claim

$H(\mathcal{O}, \ell^a)$ does not hold. Indeed, $\mathrm{ord}_\ell(\Delta) = 2k+1 < a$, so if $\Delta \equiv s^2 \pmod{\ell^a}$, then $\mathrm{ord}_\ell(s^2) = 2k+2$: contradiction. So we are in Type II with $A = 2k+1$.

- Let $\ell = 2$, and write $\mathfrak{f} = 2^k F$. Suppose $a \le 2k+1$. Since $4 \mid \Delta_K$, there is $s \in \mathbb{Z}$ such that $8 \mid \Delta_K - s^2$, so

$$2^{a+2} \mid 2^{2k+3} \mid (2^k F)^2(\Delta_K - s^2) = \Delta - (2^k F s)^2,$$

so $H(\mathcal{O}, 2^a)$ holds. Suppose $a \ge 2k+2$. If Δ is a square modulo 2^{a+2}, then Δ_K is a square modulo 2^{a+2-2k}, hence modulo 16: contradiction. So we are in Type II with $A = 2k+1$.

7E. *Torsion over $K(j)$: Part II.* We return to complete the classification of torsion on $\mathcal{O}$-CM elliptic curves $E_{/K(\mathfrak{f})}$ begun in Section 6F.

Suppose $\Delta = -4$, so $j = 1728$ and $K(\mathfrak{f}) = K = \mathbb{Q}(\sqrt{-1})$.

- By Theorem 6.2, if $E(K)$ has a point of order N, then $\varphi(N) \mid 4$, so

$$N \in \{1, 2, 3, 4, 5, 6, 8, 10\}.$$

- Using Theorem 7.2 we get

$$T(\mathcal{O}, 1) = T(\mathcal{O}, 2) = T(\mathcal{O}, 4) = T(\mathcal{O}, 5) = T(\mathcal{O}, 10) = 1,$$
$$T(\mathcal{O}, 3) = T(\mathcal{O}, 6) = 2, \;\; T(\mathcal{O}, 8) = 4.$$

- We have $C_2(\mathcal{O}) = \mu_4/\{\pm 1\}$. Thus $\# C_2(\mathcal{O}) = 2$ so every $\mathcal{O}$-CM elliptic curve $E_{/K}$ has a K-rational point of order 2, and some $\mathcal{O}$-CM elliptic curve $E_{/K}$ has $E[2] = E[2](K)$.
- Because $\widetilde{T}(\mathcal{O}, 5) = 4$, if an $\mathcal{O}$-CM elliptic curve $E_{/K}$ has a K-rational point of order 5, the index of the mod 5 Galois representation in $C_5(\mathcal{O})$ is divisible by 4. Because $\# C_2(\mathcal{O}) = 2$, if an $\mathcal{O}$-CM elliptic curve $E_{/K}$ has full 2-torsion then the index of the mod 2 Galois representation in $C_2(\mathcal{O})$ is divisible by 2. Thus if an $\mathcal{O}$-CM elliptic curve $E_{/K}$ had $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z} \hookrightarrow E(K)[\text{tors}]$, the index of the mod 10 Galois representation in $C_{10}(\mathcal{O})$ would be divisible by 8, contradicting Corollary 1.5.
- If $N \ge 3$ then $\# C_N(\mathcal{O}) > \# \mathcal{O}^\times$, so no $\mathcal{O}$-CM elliptic curve $E_{/K}$ has $E[N] = E[N](K)$.
- If there were a CM elliptic curve $E_{/K}$ with $E(K)[\text{tors}] \cong \mathbb{Z}/4\mathbb{Z}$, then there would be an ideal I of $\mathcal{O}$ such that $\mathcal{O}/I$ is isomorphic as a $\mathbb{Z}$-module to $\mathbb{Z}/4\mathbb{Z}$. But there is no such an ideal, a special case of the analysis done in the proof of Theorem 7.8.

Thus the groups which can occur as $E(K)[\text{tors}]$ are precisely

$$\mathbb{Z}/2\mathbb{Z}, \quad \mathbb{Z}/10\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}.$$

Example 7.9. For $K = \mathbb{Q}(\sqrt{-1})$, every $\mathcal{O}_K$-CM elliptic curve $E_{/K}$ is isomorphic over K to

$$E_A : y^2 = x^3 + Ax$$

for some $A \in K^\times$. We exhibit such elliptic curves with all possible torsion subgroups:

A	$E(\mathbb{Q}(\sqrt{-1}))[\text{tors}]$
2	$\mathbb{Z}/2\mathbb{Z}$
$64 - 128\sqrt{-1}$	$\mathbb{Z}/10\mathbb{Z}$
1	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$
4	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$

For the groups $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, we have $A \in \mathbb{Q}^\times$ and thus E_A arises from an elliptic curve defined over $\mathbb{Q}$ via base extension. González-Jiménez [2019, Theorem 1] has shown that for no $A \in \mathbb{Q}^\times$ do we have $E_A(K)[\text{tors}] \cong \mathbb{Z}/10\mathbb{Z}$. Najman [2010; 2011] has classified the torsion subgroups of all elliptic curves (CM or otherwise) defined over K.

Suppose $\Delta = -3$, so $j = 0$ and $K(\mathfrak{f}) = K = \mathbb{Q}(\sqrt{-3})$.

- By Theorem 6.2, if $E(K(\mathfrak{f}))$ has a point of order N, then $\varphi(N) \mid 6$, so
$$N \in \{1, 2, 3, 4, 6, 7, 9, 14, 18\}.$$
- Using Theorem 7.2 we get
$$T(\mathcal{O}, 1) = T(\mathcal{O}, 2) = T(\mathcal{O}, 3) = T(\mathcal{O}, 6) = T(\mathcal{O}, 7) = 1,$$
$$T(\mathcal{O}, 4) = 2,$$
$$T(\mathcal{O}, 9) = T(\mathcal{O}, 14) = 3,$$
$$T(\mathcal{O}, 18) = 9.$$
- We have $C_2(\mathcal{O}) = \mu_6/\{\pm 1\}$. Therefore as we range over all $\mathcal{O}$-CM elliptic curves $E_{/K}$, the group $E(K)[2]$ can be trivial (using Corollary 1.8) or have size 4, but it cannot have size 2.
- We have $C_3(\mathcal{O}) = \mu_6$. Thus there is an $\mathcal{O}$-CM elliptic curve $E_{/K}$ with $E[3] = E[3](K)$.
- If $N \geq 4$ then $\#C_N(\mathcal{O}) > \#\mathcal{O}^\times$, so no $\mathcal{O}$-CM elliptic curve $E_{/K}$ has $E[N] = E[N](K)$.

Thus the groups which can occur as $E(K)[\text{tors}]$ are precisely

$$\{e\}, \quad \mathbb{Z}/3\mathbb{Z}, \quad \mathbb{Z}/7\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}, \quad \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}.$$

Example 7.10. For $K = \mathbb{Q}(\sqrt{-3})$, every $\mathcal{O}_K$-CM elliptic curve $E_{/K}$ is isomorphic over K to

$$E_A : y^2 = x^3 + B$$

for some $B \in K^\times$. We exhibit such elliptic curves with all possible torsion subgroups:

B	$E(\mathbb{Q}(\sqrt{-1}))[\mathrm{tors}]$
2	$\{e\}$
4	$\mathbb{Z}/3\mathbb{Z}$
$6\sqrt{-3}-54$	$\mathbb{Z}/7\mathbb{Z}$
-1	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$
1	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$
16	$\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$

For the groups $\{e\}$, $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$, we have $B \in \mathbb{Q}^\times$ and thus E_B arises from an elliptic curve defined over $\mathbb{Q}$ via base extension. Again González-Jiménez [2019, Theorem 1] has shown that for no $B \in \mathbb{Q}^\times$ do we have $E_B(K)[\mathrm{tors}] \cong \mathbb{Z}/7\mathbb{Z}$. And again Najman [2010; 2011] has classified the torsion subgroups of all elliptic curves (CM or otherwise) defined over K.

Remark 7.11. (a) The calculations in Section 6F where $\Delta \neq -3, -4$ give a more detailed and explicit version of one of the main results of [Parish 1989]. Parish offers addenda on the cases where $\Delta = -3$ and -4, but without proof, and the possibilities $E(K(\mathfrak{f}))[\mathrm{tors}] \cong \mathbb{Z}/10\mathbb{Z}$ when $\Delta = -4$ and $E(K(\mathfrak{f}))[\mathrm{tors}] \cong \mathbb{Z}/7\mathbb{Z}$ and $E(K(\mathfrak{f}))[\mathrm{tors}] \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ when $\Delta = -3$ are not mentioned.

(b) If $\Delta = -3$ or -4, a classification of the possibilities for $E(K(\mathfrak{f}))[\mathrm{tors}]$ *apart* from the "Olson groups" $\{e\}$, $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/6\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ was made in [Bourdon et al. 2017b, Theorem 1.4] using computer calculations on degrees of preimages of $j = 0$ and $j = 1728$ on modular curves [Bourdon et al. 2017b, Table 2]. This result was used to find E_A with $E_A(\mathbb{Q}(\sqrt{-1})[\mathrm{tors}] \cong \mathbb{Z}/10\mathbb{Z}$ in Example 7.9 and E_B with $E_B(\mathbb{Q}(\sqrt{-3})[\mathrm{tors}] \cong \mathbb{Z}/7\mathbb{Z}$ in Example 7.10.

7F. *Isogenies over $K(j)$: Part II.* We return to complete the classification of $K(j)$-rational cyclic isogenies for elliptic curves with CM by the orders of discriminants $\Delta = -4$ and $\Delta = -3$. Recall that these cases have additional complexity coming from the fact that μ_K acts nontrivially on the projectivized torsion group $\mathbb{P}E[N]$. In this case, there is an $\mathcal{O}$-CM elliptic curve $(E_0)_{/K}$ for which the projective mod N Galois representation

$$\mathbb{P}\rho_N : \mathfrak{g}_K \to C_N(\mathcal{O})/(\mathbb{Z}/N\mathbb{Z})^\times$$

is surjective. As we vary over the K-models of E_0, the representation $\mathbb{P}\rho_N$ twists by a character

$$\mathbb{P}\chi : \mathfrak{g}_K \to \mu_K/\{\pm 1\}.$$

Thus the index of $\mathbb{P}\rho_N(\mathfrak{g}_K)$ in $C_N(\mathcal{O})/(\mathbb{Z}/N\mathbb{Z})^\times$ divides 2 when $w_K = 4$ and divides 3 when $w_K = 6$.

We will rule out the existence of K-rational cyclic N-isogenies for various values of N using the following "$\widetilde{T}$-argument": suppose that $\widetilde{T}(\mathcal{O}, N) > \varphi(N)\frac{1}{2}w_K$. Then every $C_N(\mathcal{O})$-orbit on a point of order N in $\mathcal{O}/N\mathcal{O}$ has size a multiple of $\widetilde{T}(\mathcal{O}, N)$, so every $C_N(\mathcal{O})/(\mathbb{Z}/N\mathbb{Z})^\times$-orbit on $\mathbb{P}E[N]$ has size a multiple of $\widetilde{T}(\mathcal{O}, N)/\varphi(N)$, which by our hypothesis is greater than $w_K/2$. So after passing to a field extension L of degree $w_K/2$ to trivialize $\mathbb{P}\chi$, we find that $\mathfrak{g}_L$ acts without fixed points on $\mathbb{P}E[N]$, and there is no L-rational cyclic N-isogeny and thus no K-rational cyclic N-isogeny.

Let $\mathcal{O}$ be the order of discriminant $\Delta = -4$, so $K(j) = K = \mathbb{Q}(\sqrt{-1})$ and $w_K = 4$.

- If $\ell \equiv 1 \pmod 4$, then for all $a \in \mathbb{Z}^+$ we have that -4 is a square in $\mathbb{Z}/4\ell^a\mathbb{Z}$ so there is a K-rational cyclic ℓ^a-isogeny. In fact we get that *every* $\mathcal{O}$-CM elliptic curve $E_{/K}$ has a K-rational cyclic ℓ^a-isogeny.
- If $\ell \equiv 3 \pmod 4$, since
$$\frac{\widetilde{T}(\mathcal{O}, \ell)}{\varphi(\ell)\frac{w_K}{2}} = \frac{\ell^2 - 1}{2(\ell - 1)} = \frac{\ell + 1}{2} > 1,$$
by the $\widetilde{T}$-argument there is no K-rational ℓ-isogeny.
- If $\ell = 2$, then since $T(\mathcal{O}, 4) = 1$, we can have a K-rational point of order 4 (as already seen in Section 7E), hence a cyclic K-rational 4-isogeny. Since
$$\frac{\widetilde{T}(\mathcal{O}, 8)}{\varphi(8)\frac{w_K}{2}} = \frac{16}{4 \cdot 2} > 1,$$
by the $\widetilde{T}$-argument there is no cyclic K-rational 8-isogeny.

Any elliptic curve over a number field admitting a rational cyclic N-isogeny also admits a rational cyclic M-isogeny for all $M \mid N$. Moreover, if an elliptic curve $E_{/F}$ admits F-rational cyclic $N_1, \ldots, N_r$ isogenies for pairwise coprime $N_1, \ldots, N_r$, then the subgroup generated by the kernels of these isogenies is F-rational and cyclic of order $N_1 \cdots N_r$ so E admits an F-rational cyclic $N_1 \cdots N_r$-isogeny. The assertion of Theorem 6.18(b) now follows.

Let $\mathcal{O}$ be the order of discriminant $\Delta = -3$, so $K(j) = K = \mathbb{Q}(\sqrt{-3})$ and $w_K = 6$.

- If $\ell \equiv 1 \pmod 3$, then much as in the $\Delta = -4$ case above we get that every $\mathcal{O}$-CM elliptic curve $E_{/K}$ has a K-rational cyclic ℓ^a-isogeny for all $a \in \mathbb{Z}^+$.
- If $\ell \equiv 2 \pmod 3$ and $\ell > 2$, then since
$$\frac{\widetilde{T}(\mathcal{O}, \ell)}{\varphi(\ell)\frac{w_K}{2}} = \frac{\ell^2 - 1}{3(\ell - 1)} = \frac{\ell + 1}{3} > 1,$$
by the $\widetilde{T}$-argument there is no cyclic K-rational ℓ-isogeny.

- If $\ell = 2$, then since $T(\mathcal{O}, 2) = 1$ there is an $\mathcal{O}$-CM elliptic curve $E_{/K}$ with a K-rational 2-isogeny.
- Since
$$\frac{\widetilde{T}(\mathcal{O}, 4)}{\varphi(4)\frac{w_K}{2}} = \frac{12}{2 \cdot 3} > 1,$$
by the $\widetilde{T}$-argument there is no cyclic K-rational 4-isogeny.
- We claim that there is an $\mathcal{O}$-CM elliptic curve $E_{/K}$ with a K-rational cyclic 9-isogeny. Let $\mathfrak{p}$ be the unique prime ideal of $\mathcal{O}$ lying over 3, and let P be a generator of the cyclic $\mathcal{O}$-module $E[\mathfrak{p}^3] \subset E[9]$, so P has order 9. By Lemma 7.4, the $C_9(\mathcal{O})$-orbit on P can be identified with the unit group $(\mathcal{O}/\mathfrak{p}^3)^\times$, of order 18. The $\mathcal{O}$-module generated by P is also isomorphic to $(\zeta_3 - 1)\mathcal{O}/9\mathcal{O}$, and using this representation it is easy to compute that the group $(\mathcal{O}/\mathfrak{p}^3)^\times$ is generated by the images of the scalar matrices $(\mathbb{Z}/9\mathbb{Z})^\times$ and the cube roots of unity. Thus Galois acts on the image of P in $\mathbb{P}E[9]$ via a character $\mathbb{P}\chi$. After twisting by the inverse of this character, the image of P in $\mathbb{P}E[9]$ becomes fixed by Galois and we get a K-rational cyclic 9-isogeny.
- Since
$$\frac{\widetilde{T}(\mathcal{O}, 18)}{\varphi(18)\frac{w_K}{2}} = \frac{54}{3 \cdot 6} > 1,$$
by the $\widetilde{T}$-argument there is no K-rational cyclic 18-isogeny.
- Since
$$\frac{\widetilde{T}(\mathcal{O}, 27)}{\varphi(27)\frac{w_K}{2}} = \frac{162}{3 \cdot 18} > 1,$$
by the $\widetilde{T}$-argument there is no K-rational cyclic 27-isogeny.
- From Section 7E (or Theorem 7.2) we know there is an $\mathcal{O}$-CM elliptic curve $E_{/K}$ with a rational point of order 6, hence certainly a cyclic K-rational 6-isogeny.

Using the same considerations as in the $\Delta = -4$ case above we get the assertion of Theorem 6.18(c).

Example 7.12. There are 13 imaginary quadratic discriminants Δ such that the corresponding order $\mathcal{O}(\Delta)$ has class number 1. For each such Δ we list in Table 1 the set of $N > 1$ for which there is an $\mathcal{O}(\Delta)$-CM elliptic curve E defined over $K = \mathbb{Q}(\sqrt{\Delta})$ that admits a K-rational cyclic N-isogeny — otherwise put, for which there is an $\mathcal{O}(\Delta)$-CM point on $X_0(N)(\mathbb{Q}(\sqrt{\Delta}))$.

Δ	N		
-3	$2^a 3^b \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1 \pmod 3$,	$(a, b) \in \{(0,0), (1,0), (2,0), (0,1), (1,1)\}$
-4	$2^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1 \pmod 4$,	$a \le 2$
-7	$7^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/7 = 1$,	$a \le 1$
-8	$2^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1, 3 \pmod 8$,	$a \le 1$
-11	$11^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/11 = 1$,	$a \le 1$
-12	$2^a 3^b \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1 \pmod 3$,	$a \le 2, b \le 1$
-16	$2^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1 \pmod 4$,	$a \le 3$
-19	$19^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/19 = 1$,	$a \le 1$
-27	$2^a 3^b \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i \equiv 1 \pmod 3$,	$a \le 2, b \le 3$
-28	$7^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/7 = 1$,	$a \le 1$
-43	$43^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/43 = 1$,	$a \le 1$
-67	$67^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/67 = 1$,	$a \le 1$
-163	$163^a \ell_1^{a_1} \cdots \ell_r^{a_r}$	with $\ell_i/163 = 1$,	$a \le 1$

Table 1. Values of $N > 1$ with an $\mathcal{O}(\Delta)$-CM point on $X_0(N)(\mathbb{Q}(\sqrt{\Delta}))$.

Acknowledgments

We are deeply indebted to R. Bröker for making us aware of [Stevenhagen 2001]. We also thank the anonymous referee for helpful comments and suggestions.

References

[Aoki 1995] N. Aoki, "Torsion points on abelian varieties with complex multiplication", pp. 1–22 in *Algebraic cycles and related topics* (Kitasakado, 1994), edited by F. Hazama, World Sci. Publ., River Edge, NJ, 1995. MR Zbl

[Aoki 2006] N. Aoki, "Torsion points on CM abelian varieties", *Comment. Math. Univ. St. Pauli* **55**:2 (2006), 207–229. MR Zbl

[Bourdon and Clark 2019] A. Bourdon and P. Clark, "Torsion points and isogenies on CM elliptic curves", preprint, 2019. arXiv

[Bourdon and Pollack 2017] A. Bourdon and P. Pollack, "Torsion subgroups of CM elliptic curves over odd degree number fields", *Int. Math. Res. Not.* **2017**:16 (2017), 4923–4961. MR Zbl

[Bourdon et al. 2017a] A. Bourdon, P. L. Clark, and P. Pollack, "Anatomy of torsion in the CM case", *Math. Z.* **285**:3-4 (2017), 795–820. MR Zbl

[Bourdon et al. 2017b] A. Bourdon, P. L. Clark, and J. Stankewicz, "Torsion points on CM elliptic curves over real number fields", *Trans. Amer. Math. Soc.* **369**:12 (2017), 8457–8496. MR Zbl

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

[Clark 2015] P. L. Clark, "Commutative algebra", preprint, 2015, available at http://math.uga.edu/~pete/integral2015.pdf. To appear in *J. London Math. Soc.*

[Clark and Pollack 2015] P. L. Clark and P. Pollack, "The truth about torsion in the CM case", *C. R. Math. Acad. Sci. Paris* **353**:8 (2015), 683–688. MR Zbl

[Clark et al. 2013] P. L. Clark, B. Cook, and J. Stankewicz, "Torsion points on elliptic curves with complex multiplication (with an appendix by Alex Rice)", *Int. J. Number Theory* **9**:2 (2013), 447–479. MR Zbl

[Clark et al. 2014] P. L. Clark, P. Corn, A. Rice, and J. Stankewicz, "Computation on elliptic curves with complex multiplication", *LMS J. Comput. Math.* **17**:1 (2014), 509–535. MR Zbl

[Cohen 2000] H. Cohen, *Advanced topics in computational number theory*, Graduate Texts in Mathematics **193**, Springer, 2000. MR Zbl

[Cox 1989] D. A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory and complex multiplication*, Wiley, New York, 1989. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR Zbl

[Franz 1935] W. Franz, "Die Teilwerte der Weberschen Tau-Funktion", *J. Reine Angew. Math.* **173** (1935), 60–64. MR Zbl

[Gaudron and Rémond 2018] E. Gaudron and G. Rémond, "Torsion des variétés abéliennes CM", *Proc. Amer. Math. Soc.* **146**:7 (2018), 2741–2747. MR Zbl

[González-Jiménez 2019] E. González-Jiménez, "Explicit characterization of the torsion growth of rational elliptic curves with complex multiplication over quadratic fields", preprint, 2019. arXiv

[Kwon 1999] S. Kwon, "Degree of isogenies of elliptic curves with complex multiplication", *J. Korean Math. Soc.* **36**:5 (1999), 945–958. MR Zbl

[Lang 1987] S. Lang, *Elliptic functions*, 2nd ed., Graduate Texts in Mathematics **112**, Springer, 1987. MR Zbl

[Lombardo 2017] D. Lombardo, "Galois representations attached to abelian varieties of CM type", *Bull. Soc. Math. France* **145**:3 (2017), 469–501. MR Zbl

[Lozano-Robledo 2018a] A. Lozano-Robledo, "Galois representations attached to elliptic curves with complex multiplication", preprint, 2018. arXiv

[Lozano-Robledo 2018b] A. Lozano-Robledo, "Uniform boundedness in terms of ramification", *Res. Number Theory* **4**:1 (2018), art. no. 6, 39 pp. MR Zbl

[Najman 2010] F. Najman, "Complete classification of torsion of elliptic curves over quadratic cyclotomic fields", *J. Number Theory* **130**:9 (2010), 1964–1968. MR Zbl

[Najman 2011] F. Najman, "Torsion of elliptic curves over quadratic cyclotomic fields", *Math. J. Okayama Univ.* **53** (2011), 75–82. MR Zbl

[Olson 1974] L. D. Olson, "Points of finite order on elliptic curves with complex multiplication", *Manuscripta Math.* **14** (1974), 195–205. MR Zbl

[Parish 1989] J. L. Parish, "Rational torsion in complex-multiplication elliptic curves", *J. Number Theory* **33**:2 (1989), 257–265. MR Zbl

[Prasad and Yogananda 2001] D. Prasad and C. S. Yogananda, "Bounding the torsion in CM elliptic curves", *C. R. Math. Acad. Sci. Soc. R. Can.* **23**:1 (2001), 1–5. MR Zbl

[Ross 1994] R. Ross, "Minimal torsion in isogeny classes of elliptic curves", *Trans. Amer. Math. Soc.* **344**:1 (1994), 203–215. MR Zbl

[Serre 1972] J.-P. Serre, "Propriétés galoisiennes des points d'ordre fini des courbes elliptiques", *Invent. Math.* **15**:4 (1972), 259–331. MR Zbl

[Silverberg 1988] A. Silverberg, "Torsion points on abelian varieties of CM-type", *Compositio Math.* **68**:3 (1988), 241–249. MR Zbl

[Silverberg 1992] A. Silverberg, "Points of finite order on abelian varieties", pp. 175–193 in *p-adic methods in number theory and algebraic geometry*, edited by A. Adolphson et al., Contemp. Math. **133**, Amer. Math. Soc., Providence, RI, 1992. MR Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR Zbl

[Stevenhagen 2001] P. Stevenhagen, "Hilbert's 12th problem, complex multiplication and Shimura reciprocity", pp. 161–176 in *Class field theory—its centenary and prospect* (Tokyo, 1998), edited by K. Miyake, Adv. Stud. Pure Math. **30**, Math. Soc. Japan, Tokyo, 2001. MR Zbl

Received March 1, 2019. Revised September 15, 2019.

ABBEY BOURDON
DEPARTMENT OF MATHEMATICS AND STATISTICS
WAKE FOREST UNIVERSITY
WINSTON-SALEM, NC
UNITED STATES

bourdoam@wfu.edu

PETE L. CLARK
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
ATHENS, GA
UNITED STATES

plclark@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.89

STABILITY OF THE POSITIVE MASS THEOREM FOR AXISYMMETRIC MANIFOLDS

EDWARD T. BRYDEN

Away from the central axis, we prove the stability of the positive mass theorem in the $W^{1,p}$ sense for asymptotically flat axisymmetric manifolds with nonnegative scalar curvature satisfying some additional technical assumptions. We also derive estimates for the volumes of regions, the areas of axisymmetric surfaces, and the distances between points within the manifolds.

1. Introduction

Based on the formulation of general relativity, our physical intuition leads us to expect a close relationship between the ADM mass of an asymptotically flat Riemannian manifold and its geometry. Recall that the ADM mass of an asymptotically flat Riemannian manifold is defined to be

$$m = \lim_{R\to\infty} \frac{1}{16\pi} \int_{S_R} (g_{ij,j} - g_{jj,i})\nu^i. \tag{1-1}$$

In their celebrated positive mass theorem, Schoen and Yau [1979] proved that if an asymptotically flat manifold has nonnegative scalar curvature, then the ADM mass is nonnegative. They also proved the following rigidity theorem:

$$m = 0 \Longrightarrow M \text{ is isometric to Euclidean space.} \tag{1-2}$$

MSC2010: 53C24, 53C80, 83C99.
Keywords: positive mass theorem, stability, small mass, axisymmetric.

It is natural to ask whether stability also holds; if M has small ADM mass, is M close to Euclidean space? Lee and Sormani [2014] have shown that M need not be smoothly, nor even C^0, close to Euclidean space even in the spherically symmetric setting; there could be increasingly deep thin gravity wells at the center. They conjectured that M is close to Euclidean space in the Sormani–Wenger intrinsic flat (SWIF) sense [Huang et al. 2017; Lee and Sormani 2014]. Proving it will require a method for picking appropriate subregions geometrically and a way to show that these regions converge in the SWIF metric to a subset of Euclidean space.

Lee and Sormani [2014] studied stability in the rotationally symmetric setting. They showed that tubular neighborhoods of fixed radius D about coordinate spheres of fixed area A converge to the Euclidean tubular neighborhood of radius D about a sphere of area A. Earlier, Lee [2009] had proven convergence to Euclidean space outside a compact set in the conformally flat setting. Assuming strong conditions on sectional curvature, Corvino [2005] has proven that an asymptotically flat manifold with nonnegative scalar curvature and small ADM mass must be diffeomorphic to $\mathbb{R}^3$. Finster, Bray and Kath have papers bounding the L^2 norm of the curvature [Bray and Finster 2002; Finster and Kath 2002]. After the Lee–Sormani paper, LeFloch and Sormani [2015] proved that metric tensors converge in the H^1_{loc} sense in the rotationally symmetric setting. Huang, Lee, and Sormani [Huang et al. 2017] proved SWIF convergence in the graph setting and Sormani and Stavrov Allen [2019] proved it in the geometrostatic setting. Allen [2018] proved L^2 convergence in regions where the inverse mean curvature flow is smooth.

Here, we will study the question of stability in the presence of axisymmetry. The class of axisymmetric metrics is both flexible enough to model a range of physically interesting phenomena and restricted enough that we have powerful tools at hand that are not available in the most general setting. Recall that the coordinate expression for an axisymmetric metric in cylindrical coordinates is

$$g = e^{2\alpha-2u}(d\rho^2+dz^2)+\rho^2 e^{-2u}(d\phi + B\,d\rho + A\,dz)^2, \tag{1-3}$$

where all the functions involved depend only on ρ and z. The killing field associated with the axisymmetry of g is $\frac{\partial}{\partial\phi}$.

Since we will be studying large families of asymptotically flat metrics, it is natural to require that the family satisfy some type of uniform falloff condition.

Definition 1.1. Let $\mathcal{M}$ be a family of axisymmetric metrics. Suppose we can parametrize $\mathcal{M}$ by the functions α, u, A, and B in cylindrical coordinates (1-3). If there exist constants C and R_0 such that if g is a metric in $\mathcal{M}$, then for all $\sqrt{\rho^2+z^2} = r \geq R_0$ we have

$$|\partial^I u| \leq \frac{C}{r^{1+|I|}}, \quad |\partial^I \alpha| \leq \frac{C}{r^{1+|I|}}, \quad |\partial^I A| \leq \frac{C}{r^{1+|I|}}, \quad |\partial^I B| \leq \frac{C}{r^{1+|I|}}, \tag{1-4)–(1-7}$$

then we shall call $\mathcal{M}$ uniformly asymptotically flat outside of radius R_0

Chruściel [2008] shows that if (M, g) is a simply connected axisymmetric manifold which is asymptotically flat, then there are cylindrical coordinates (ρ, z, ϕ) in which g takes the form (1-3). In fact, Chruściel's construction works for simply connected axisymmetric manifolds with multiple asymptotically flat ends. In this case, the additional "points at infinity" will be points removed from the z-axis at which the coordinate function u will blow up. If the metric g is assumed to be without conical singularities and smooth, then on the z-axis we will have

$$\alpha = \frac{\partial \alpha}{\partial \rho} = \frac{\partial u}{\partial \rho} = 0, \tag{1-8}$$

away from the points of infinity removed from the z-axis. Chruściel's construction also works for axisymmetric manifolds with boundary, where the boundary has the same killing field as does the rest of the manifold. One must perform a fill-in so that the resulting manifold will have empty boundary [Chruściel 2008], then construction proceeds as in the boundaryless case. However, generally this fill-in will be unphysical. Thus, it is desirable to remove from consideration all points that were filled in out of technical necessity. To accomplish this, one may observe that the form of (1-3) is unchanged by a conformal transformation of the coordinates ρ and z. This allows us to construct cylindrical coordinates for which the boundary of the manifold lies on the axis of symmetry. However, the blow up of the functions α and u at the boundary is much more severe than at points representing other asymptotically flat ends. The effect this has on the analysis of these manifolds is discussed more in Section 2 and Appendix A.

Suppose that g has the standard asymptotically flat falloff rate:

$$|\partial^I (g - \delta_{\mathbb{R}^3})| \leq \frac{C}{r^{1+|I|}}, \tag{1-9}$$

where $\delta_{\mathbb{R}^3}$ is the Euclidean metric. In general, the asymptotic falloff of the functions α, u, A, and B will not be as strong as the those given in Definition 1.1. However, we may make an additional assumption on the growth of the killing field of g in the asymptotic limit which will imply that the functions α, u, A, and B do have the same falloff as in Definition 1.1. Although, the author does not know if making such an assumption uniform over a family of axisymmetric metrics will yield a uniformly asymptotically flat family of axisymmetric metrics. This indicates that there are many families of metrics satisfying the requirements of Definition 1.1, although we do not have a geometric method for picking them out.

In Chruściel's construction of cylindrical coordinates, the coordinate functions ρ and z are both solutions to a PDE determined by the metric g. Specifically, if we let η denote the killing field generating the axisymmetry of g and let q denote the metric on the orbit space induced by g, then both ρ and z solve

$$\Delta_q \omega = \Delta_g \omega - \frac{1}{2|\eta|_g^2} \langle \nabla \omega, \nabla |\eta|_g^2 \rangle_g = 0. \tag{1-10}$$

In fact, ρ and z are uniquely determined up to conformal maps in the plane. In [Gibbons and Holzegel 2006, Section 2], it is noted that if we insist on mapping the axis of symmetry to itself and preserving asymptotic flatness, then ρ is completely fixed. In addition, we can see that z is unique up to translation. This uniqueness justifies our choice to parametrize families of axisymmetric metrics as we did in Definition 1.1

A major obstacle to proving the stability of the positive mass theorem, perhaps the principal one, is that the ADM mass cannot control regions within outermost minimizing surfaces. Classic examples depicting why the Penrose inequality depends on the area of an outermost minimizing surface demonstrate this phenomenon. One way to overcome this difficulty, which was applied in the work of Bray and Finster [2002], Finster and Kath [2002], Huang, Lee, and Sormani [2017], and Allen [2018], is to impose conditions which constrain the location, or prevent the existence, of an outermost minimal surface. We shall follow this approach in making the following definition.

Definition 1.2. Let $\mathcal{M}$ be a family of axisymmetric metrics and let η denote the killing field generating their axisymmetry. Suppose that for each metric $g \in \mathcal{M}$ we have the following inequality

$$\frac{|\eta|_g}{|\nabla \rho|_g}(\rho_0, z) \geq \rho_0. \tag{1-11}$$

Then we shall call $\mathcal{M}$ a family of area enlarging metrics at ρ_0. If the inequality holds for each ρ_0, then we shall simply call the family area enlarging.

Uniqueness of solutions to (1-10) implies that the above is a condition imposed on the family $\mathcal{M}$ and has significance beyond a coordinate condition. However, it is useful to express the above in terms of cylindrical coordinates. In coordinates the condition reads

$$(\alpha - 2u)(\rho_0, z) \geq 0. \tag{1-12}$$

In the appendices we show that the Schwarzschild solution is area enlarging.

Suppose that $\mathcal{M}$ satisfies condition (1-11) for all ρ_0. Let $\delta_{\mathbb{R}^3}$ denote the background Euclidean metric given in the cylindrical coordinates (ρ, z, ϕ). Then in Proposition 5.1 we show that

$$\text{Area}_g(\Sigma) \geq \text{Area}_{\delta_{\mathbb{R}^3}}(\Sigma) \tag{1-13}$$

for axisymmetric surfaces Σ. Together with the Penrose inequality, the above area inequality works to constrain the location of outermost minimal surfaces. In Corollary 5.2 we show that if Σ is an axisymmetric outermost minimal surface which is also a sphere, then

$$\Sigma \subset \rho^{-1}\big([0, 2\sqrt{2}m)\big), \tag{1-14}$$

where m is the ADM mass of the metric under consideration.

As in prior work on stability, we must judiciously decide which regions we will study. In view of the above discussion, the regions

$$\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma) = \left\{\rho_0 + \sigma \le \rho \le \rho_1,\ |z| \le \frac{\rho_1}{2}\right\} \times [0, 2\pi), \tag{1-15}$$

for some fixed ρ_0 and $\sigma \ge 0$, are natural choices. If σ is identically zero, then we shall write $\widetilde{\Omega}_{\rho_0}^{\rho_1}$. Since we mainly work in the orbit space, we shall often only consider the image of $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$ under the projection map, which is simply the rectangle

$$\Omega_{\rho_0}^{\rho_1}(\sigma) = \left\{\rho_0 + \sigma \le \rho \le \rho_1,\ |z| \le \frac{\rho_1}{2}\right\}. \tag{1-16}$$

If σ is taken to be zero, then we shall write $\Omega_{\rho_0}^{\rho_1}$.

Instead of the area enlarging assumption (1-11), we will at first work with another requirement.

Definition 1.3. Let $\mathcal{M}$ be a family of axisymmetric metrics. Suppose that for each metric $g \in \mathcal{M}$ we have the inequality

$$\frac{\partial}{\partial \rho}\left(\frac{1}{\rho}\frac{|\eta|_g}{|\nabla \rho|_g}\right) \le 0 \tag{1-17}$$

on the set $\{\rho = \rho_0\}$. Then we shall call the family radially monotone at ρ_0. If $\mathcal{M}$ is radially monotone at each ρ_0, then we will simply call $\mathcal{M}$ radially monotone.

This too is a geometric condition on a family of axisymmetric metrics. In Proposition B.3 we show that if g is an axisymmetric metric, and ρ is the solution to (1-10), then g is radially monotone if and only if the level sets of the function ρ form a sub-inverse-mean-curvature flow.

The radial monotonicity condition has a useful expression in cylindrical coordinates:

$$\frac{\partial(\alpha - 2u)}{\partial \rho} \le 0. \tag{1-18}$$

In this form, a similar inequality to the above can be found in Section 3.2 of [Chruściel and Nguyen 2011].

One could wonder if there is any relationship between the area enlarging condition and the radial monotonicity condition. Pointwise, there is no such relationship.

However, if radial monotonicity holds everywhere, then the area enlarging condition must also hold everywhere, see Proposition B.4. Thus, radial monotonicity everywhere also constrains the location of minimal surfaces, as in (1-14).

In Appendix B we will show that the Kerr–Newman and axisymmetric geometrostatic metrics satisfy radial monotonicity and the area enlarging condition, respectively. In fact, the Kerr–Newman metrics satisfy radially monotonicity strictly, so that small perturbations of the Kerr–Newman metrics are also radially monotone. The same is true for small perturbations of axisymmetric geometrostatic metrics with regards to the area enlarging condition. However, there is an important difference between the geometric static case and the Kerr–Newman metrics: although there is a minimal surface in the geometric static case, the initial data is extended past this surface "into the black hole," while the explicit form of the Kerr–Newman metric that we use is given only outside of the minimal surface, and the minimal surface is located on the axis of symmetry. As discussed later, this changes the mass formula, though it does not change how we use the mass formula. Until Appendix A, we will assume all of our manifolds have empty boundary, but may have multiple asymptotically flat ends.

We now state the stability of the positive mass theorem in the $W^{1,p}$ sense.

Theorem 1.4. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose that $\mathcal{M}$ is radially monotone at ρ_0 and that for each metric in $\mathcal{M}$, we have*

$$A = B = 0. \tag{1-19}$$

For every $\rho_1 > \max\{\rho_0, R_0\}$, $\epsilon > 0$, $\sigma > 0$, and $1 \leq p < 2$ there exists a $\delta > 0$ such that if the ADM mass of $g \in \mathcal{M}$ is less than δ, then

$$\|g - \delta_{\mathbb{R}^3}\|_{W^{1,p}(\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma))} < \epsilon, \tag{1-20}$$

and

$$\|q - \delta_{\mathbb{R}^2}\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1}(\sigma))} < \epsilon, \tag{1-21}$$

where $\delta_{\mathbb{R}^3}$ denotes the Euclidean metric in cylindrical coordinates, $\delta_{\mathbb{R}^2}$ denotes the Euclidean metric in the (ρ, z) plane, and q denotes the orbit metric of g in the (ρ, z) plane. $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$ denotes the cylinder given in (1-15) and $\Omega_{\rho_0}^{\rho_1}(\sigma)$ denotes its orbit space.

Remark 1.5. Although we are restricting our attention to metrics with no boundary, we are still allowing the possibility of multiple asymptotically flat ends. Thus, there may be closed embedded minimal surfaces in our metric. This shows, once again, how we are using the radial monotonicity condition to handle the presence of these minimal surfaces.

The assumption that the functions A and B vanish is very likely unnecessary, however it does simplify the analysis considerably. That the exponent p is required to be less than two is natural to the problem at hand. Suppose we were able to prove an analogous result for $p > 2$. Then, we would be able to apply the Sobolev embedding theorem to conclude that the convergence was actually C_0 convergence. However, as mentioned before, see [Lee and Sormani 2014], there are counterexamples to C_0 stability.

It is not yet known if $W^{1,p}$ convergence implies SWIF convergence. However, in the course of proving $W^{1,p}$ stability, we obtain similar estimates to those Huang, Lee, and Sormani [Huang et al. 2017] use to prove the stability of the positive mass theorem in the SWIF metric for graphical manifolds. Let $\mathcal{M}$ be a family of three dimensional asymptotically flat graphical manifolds in $\mathbb{R}^4$ and let C_{r_0} denote the infinite cylinder with base a ball of radius r_0 about the origin in $\mathbb{R}^3 \subset \mathbb{R}^4$. Huang, Lee, and Sormani studied the regions $\Omega_{r_0} \subset M \in \mathcal{M}$ defined by

$$\Omega_{r_0} := M \cap C_{r_0}, \tag{1-22}$$

for some appropriately large r_0. Additionally, they assume a uniform diameter bound on the Ω_{r_0}. They then show that as the ADM mass approaches zero, the regions Ω_{r_0} converge in the SWIF metric to a three dimensional Euclidean ball in $\mathbb{R}^4$,

$$B(0, r_0) \times \{0\}. \tag{1-23}$$

Their proof follows from three assertions. First, they showed that the volumes of the Ω_{r_0} converge to the volume of $B(0, r_0)$. Second, they showed that the area of $\partial\Omega_{r_0}$ approaches the area of $\partial B(0, r_0)$. Finally, they showed that $\partial\Omega_{r_0} \cap \partial C_{r_0}$ Lipschitz converges to $\partial B(0, r_0) \times \{0\}$.

We are able to establish volume convergence for the cylinders $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$ defined as in (1-15).

Theorem 1.6. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is radially monotone at ρ_0. For any constants $\epsilon > 0$, $\sigma > 0$, and $\rho_1 > \max\{\rho_0, R_0\}$, there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{1-24}$$

then

$$|\Omega| + \epsilon \geq \mathrm{vol}_g(\Omega) \geq |\Omega| - \epsilon \tag{1-25}$$

for any region Ω *such that*

$$\Omega \subset \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma). \tag{1-26}$$

We are also able to establish control over areas inside our designated regions.

Theorem 1.7. *Let* $\mathcal{M}$ *be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius* R_0*. Suppose also that* $\mathcal{M}$ *is radially monotone at* ρ_0*. For any fixed axisymmetric surface* Σ*, constant* $\epsilon > 0$*, and constant* $\rho_1 > \max\{\rho_0, R_0\}$*, there exists a* $\delta > 0$ *such that if* $m(g) < \delta$*, then*

$$|\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)| + \epsilon \geq \mathrm{Area}_g(\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)) \geq |\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)| - \epsilon. \tag{1-27}$$

We obtain an estimate on distances between certain points in $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$ which can be used to give an upper bound on the diameter of $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$.

Theorem 1.8. *Let* $\mathcal{M}$ *be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius* R_0*. Suppose* $\mathcal{M}$ *is also radially monotone at* ρ_0*. Additionally, assume that* $A = B = 0$ *in the coordinate representations of the metrics under consideration. Suppose we are given* $\epsilon > 0$*,* $\sigma > 0$*, and* $\rho_1 > \max\{\rho_0, R_0\}$*. There exists a constant* $\delta > 0$ *such that if* $m(g) \leq \delta$ *and* x *and* y *are any points such that the Euclidean line segment connecting them lies in* $\Omega_{\rho_0}^{\rho_1}(\sigma) \times \{\phi_0\}$ *for any* ϕ_0*, then*

$$d_g(x, y) \leq d(x, y) + \epsilon. \tag{1-28}$$

For more general pairs of points x and y in $\widetilde{\Omega}_{\rho_0}^{\rho_1}$ we have a pointwise estimate on their distance to each other.

Theorem 1.9. *Let* $\mathcal{M}$ *be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius* R_0*. Suppose also that* $\mathcal{M}$ *is radially monotone at* ρ_0*. Additionally, assume that* $A = B = 0$ *in the coordinate representations of the metrics under consideration. Suppose we are given* $\epsilon > 0$ *and* $\sigma > 0$ *and points* x *and* y *such that the Euclidean line segment connecting them lies in* $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$*. There exists a constant* $\delta > 0$ *such that if* $m(g) \leq \delta$*, then*

$$d_g(x, y) \leq d(x, y) + \epsilon. \tag{1-29}$$

Finally, we are able to establish uniform convergence at large distances from the origin.

Theorem 1.10. *Let* $\mathcal{M}$ *be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of* *radius R_0. Suppose that $\mathcal{M}$ is radially monotone and that for all $g \in \mathcal{M}$ we have*

$$A = B = 0. \tag{1-30}$$

Let $R_1 > R_0$ and let $A(R_0, R_1)$ denote the coordinate spherical annulus centered at the origin. For any given $0 < \beta < 1$ and $\epsilon > 0$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and

$$m(g) < \delta, \tag{1-31}$$

then

$$\|g - \delta_{\mathbb{R}^3}\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon. \tag{1-32}$$

These theorems are proven in Section 5 after we prove a series of lemmas estimating various terms in the coordinate system. All of the above theorems hold if we assume the area enlarging condition (1-11) instead of radial monotonicity (1-17). The only change is that in addition to assuming (1-11), we must assume that our family of manifolds satisfies a stronger uniform asymptotic falloff than the one given in Definition 1.1.

Definition 1.11. Let $\mathcal{M}$ be an uniformly asymptotically flat family of metrics. Suppose that in addition to the uniform asymptotic falloff (Definition 1.1), we have some uniform $\tau > 0$ such that

$$|\alpha| \leq \frac{C}{r^{1+\tau}}. \tag{1-33}$$

Then we shall call $\mathcal{M}$ strongly uniformly asymptotically flat.

In the future we would like to prove the Lee–Sormani stability conjecture that regions outside outermost minimizing surfaces converge in the SWIF sense to regions in Euclidean space. Our volume, area, and distance controls should be useful towards such a proof. Here we used an extra condition (1-11) to constrain, a priori, the location of outer most minimal surfaces. Another approach would be to actually locate outermost minimal surfaces without any assumption. This was done easily in [Lee and Sormani 2014] thanks to spherical symmetry and was a huge challenge in the work of Sormani and Stavrov Allen [2019]. Locating the outermost minimal surfaces in an axisymmetric manifold is of independent interest and would be worthy of a paper on its own.

2. Background information

The ADM mass is calculated by taking a limit of integrals over the boundaries of increasingly large coordinate balls. Thus, it is unclear how the ADM mass should control the geometry inside of these balls. In fact, arbitrary local perturbations of a metric would not change its ADM mass. However, if we restrict our attention to metrics with nonnegative scalar curvature, then we are no longer entirely free in our choice of local perturbation. This restores our hope that the ADM mass can control geometry.

In an attempt to relate ADM mass and the interior geometry, it is natural to make use of the divergence theorem,

$$m(g) = \lim_{R\to\infty} \frac{1}{16\pi}\int_{\partial B_R} (g_{ij,j} - g_{jj,i})\nu^i = \lim_{R\to\infty} \frac{1}{16\pi}\int_{B_R} \operatorname{div}(g_{ij,j} - g_{jj,i}), \tag{2-1}$$

to get an integral over the interior. For now, we are ignoring the question of which metric we should use to take the divergence. Intuitively, we think of scalar curvature as a local energy density. As such, we would like to relate the divergence term to the scalar curvature. Ideally, the nonnegativity of the scalar curvature should give control over the integral of the divergence term. This approach can be successfully carried out in the case of axisymmetric metrics. Furthermore, Witten [1981] used a more sophisticated version of this idea to prove the positive mass theorem for manifolds with spinors.

In cylindrical coordinates for axisymmetric metrics we have the following formula for the scalar curvature [Brill 1959]:

$$R_g = 4e^{2(u-\alpha)}\left[\Delta_{\mathbb{R}^3}\left(u - \tfrac{1}{2}\alpha\right) - \tfrac{1}{2}|\nabla u|_\delta^2 + \frac{1}{2\rho}\frac{\partial\alpha}{\partial\rho} - \frac{\rho^2 e^{-2\alpha}}{8}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial\rho}\right)^2\right]. \tag{2-2}$$

Here we can see that the scalar curvature is indeed closely related to a divergence, namely $\Delta_{\mathbb{R}^3}\left(u - \frac{\alpha}{2}\right)$. This observation leads to a very useful formula for the mass of an axisymmetric metric, including those with multiple asymptotically flat ends [Brill 1959; Chruściel 2008]:

$$m(g) = \frac{1}{16\pi}\int_{\mathbb{R}^3}\left[e^{-2(u-\alpha)}\left[R_g + \frac{\rho^2 e^{-4\alpha+2u}}{2}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial\rho}\right)^2\right] + 2|\nabla u|_\delta^2\right]\rho\, d\rho\, dz\, d\phi. \tag{2-3}$$

If there are multiple asymptotically flat ends, which will be points on the z-axis, then the function u will blow up at these points. In fact, we see that u is roughly the logarithm of the distance to these points in the Euclidean background metric. Since we are integrating over $\mathbb{R}^3$, one may use polar integration to be convinced that (2-3) is finite. For details of the case in which there are multiple asymptotically flat ends, see [Chruściel 2008, Theorems 2.9 and 3.3].

Since all other terms are explicitly nonnegative, if we assume that $R \geq 0$, then the ADM mass immediately gives control over the gradient of u. In an asymptotically flat metric, u must be arbitrarily small on large coordinate spheres. It is therefore reasonable to suppose that we can use the fundamental theorem of calculus to control u everywhere in the manifold. In order to make this precise, we will use the following representation formula to express u in terms of its gradient and its value on large coordinate spheres.

Suppose Ω is a compact region on which the divergence theorem holds and let Γ be the fundamental solution for the Laplacian. Assume further that u is a function which is differentiable on $\mathrm{CL}(\Omega)$. Then we have

$$u(x) = -\int_{\partial\Omega} u(y)\langle \nabla\Gamma(x,y), n\rangle \, dy + \int_{\Omega} \langle \nabla u(y), \nabla\Gamma(x,y)\rangle \, dy. \tag{2-4}$$

In order to see this, we follow the calculations appearing as 2.15 in [Gilbarg and Trudinger 1998], except we use the divergence theorem on the vector field Z defined by

$$Z = u(y)\nabla\Gamma(x,y). \tag{2-5}$$

Since we should not expect to have any physically relevant information inside of a minimal surface, it is reasonable to exclude from consideration all parts of a manifold lying within the outermost minimal surface. As such, it is desirable to include manifolds with minimal surface boundary in our analysis. In fact, we will choose coordinates for which the boundary of the manifold is taken to lie on the axis of symmetry: the boundary will consist of disjoint rods lying on the z-axis. The function u will still blow up logarithmically, but now as the logarithm of the distance to a rod on the axis. Integrating using cylinders should convince one that (2-3) should no longer be finite. In modifying the mass formula to suit manifolds with boundary, we pick up boundary terms which complicate our analysis [Chruściel 2008; Khuri et al. 2019]. In the case of a connected boundary, see [Khuri et al. 2019, Equations (2.10)–(2.12)], the mass formula becomes

$$m(g) = \frac{1}{16\pi}\int_{\mathbb{R}^3} 2|\nabla\bar{u}|^2 + e^{2(u-\alpha)}R_g \, dx + \tfrac{1}{4}\int_{-m_0}^{m_0} \bar{\alpha}(0,z) - 2\bar{u}(0,z)\, dz + m_0, \tag{2-6}$$

where $\bar{u}$ and $\bar{\alpha}$ are appropriate regularizations of u and α, respectively. This formula has a lot in common with the boundaryless case (2-3), however, to the best of the author's knowledge, it has not been demonstrated that

$$\tfrac{1}{4}\int_{-m_0}^{m_0} \bar{\alpha}(0,z) - 2\bar{u}(0,z)\, dz + m_0 \geq 0, \tag{2-7}$$

nor even a lower bound established in general. Thus, it is no longer clear that mass controls the right hand side of (2-6). If (2-7) holds, then Sobolev stability, and all of the related theorems, are still valid in the nonempty connected boundary case, as will be detailed in Appendix A. For now, we will assume that we are in the case of an empty boundary.

The ease with which we can obtain estimates for u is encouraging, however there is one more hurdle. If we want to use mass to control the metric (1-3), then we must be able to turn our estimates for u into estimates for e^u. Luckily, we may use the well known Moser–Trudinger inequality [Gilbarg and Trudinger 1998] to accomplish this.

In view of the coordinate expression for an axisymmetric metric (1-3), we know that if we can control $e^{\alpha-2u}$ as well as e^u, then we have achieved good control over the metric. Although it is less clear, it is possible to use the mass formula (2-3) and the scalar curvature equation (2-2) to show that the ADM mass controls the $W^{1,p}$ norm of $\alpha - 2u$. The process is similar to what we do to estimate u. However, we use Green's representation formula, instead of (2-4), to express $\alpha - 2u$ as a boundary term plus an integral of its derivatives. We recall Green's representation formula now.

Let Ω be a compact region on which the divergence theorem holds and let Γ be the fundamental solution of the Laplacian. Suppose that ω is a twice differentiable function on $\mathrm{CL}(\Omega)$. Then we have the following representation of ω:

$$\omega(x) = \int_{\partial\Omega}\left[\omega(y)\frac{\partial\Gamma(x,y)}{\partial\nu} - \Gamma(x,y)\frac{\partial\omega(y)}{\partial\nu}\right]dy + \int_{\Omega}\Gamma(x,y)\Delta\omega(y)\,dy. \tag{2-8}$$

This result appears in [Gilbarg and Trudinger 1998] as Equation 2.16.

With $W^{1,p}$ estimates for $\alpha - 2u$ in hand, we might hope to use the Moser–Trudinger inequality to get estimates for $e^{\alpha-2u}$. Unfortunately, the Moser–Trudinger inequality doesn't apply in this case. Luckily, because of axisymmetry, we are essentially working in two dimensions. This gives us extra control that does not exist in higher dimensions. In this setting we are able to prove a result similar to the Moser–Trudinger inequality, which allows us to turn $W^{1,p}$ estimates for $\alpha - 2u$ into $W^{1,p}$ estimates for $e^{\alpha-2u}$.

In using (2-4) and (2-8) to control the $W^{1,p}$ norms of u and $\alpha - 2u$, we rely on estimates of the Riesz potential. Recall that the Riesz potential of a function f over a region Ω, denoted $(V_\mu f)(x)$, is defined as

$$(V_\mu f)(x) = \int_\Omega |x-y|^{n(\mu-1)} f(y)\,dy, \tag{2-9}$$

for $\mu \in (0,1]$. Let $0 \le \delta = \delta(p,q) = q^{-1} - p^{-1} < \mu$ and let ω_n denote the volume of the unit n dimensional ball. The following inequality appears as Lemma 7.12 in [Gilbarg and Trudinger 1998]:

$$\|(V_\mu f)\|_p \le \left(\frac{1-\delta}{\mu-\delta}\right)^{1-\delta} \omega_n^{1-\mu} |\Omega|^{\mu-\delta} \|f\|_q, \tag{2-10}$$

where Ω is some open region in $\mathbb{R}^n$ with compact closure and f is in $L^q(\Omega)$.

3. Sobolev estimates for u and e^u

In this section we will see in greater detail the steps needed to estimate the $W^{1,p}$ norm of e^u using the mass formula (2-3). Our end goal is to produce estimates over the regions $\Omega^{\rho_1}_{\rho_0}(\sigma)$, see (1-16). In fact, we are always able to take σ to be zero. To simplify notation, such rectangles will be denoted by $\Omega^{\rho_1}_{\rho_0}$.

To start, the ADM mass only explicitly bounds the $L^2(\mathbb{R}^3)$ norm of ∇u. The following lemma demonstrates that this is enough to get $W^{1,2}(B_{r_0})$ control over u for a ball of fixed radius r_0 about the origin in $\mathbb{R}^3$.

Lemma 3.1. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0, and let B_{r_0} be the ball of radius r_0 about the origin. For any $\epsilon > 0$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{3-1}$$

then

$$\|u\|_{W^{1,2}(B_{r_0})} < \epsilon. \tag{3-2}$$

Proof. We note once again that control over $\|\nabla u\|_{L^2(B_{r_0})}$ is an immediate consequence of the mass formula and the nonnegative scalar curvature assumption. In the calculations that follow we will denote the volume of a three dimensional unit ball by ω_3. First, we look at some very large coordinate ball $B(0, r_1)$ with $r_1 > \max\{r_0, R_0\}$. If we let Γ be the fundamental solution for the Laplacian, then using (2-4) we may express u as

$$u(x) = -\int_{\partial B(0,r_1)} u(y)\langle \nabla\Gamma(x, y), n\rangle\, dy + \int_{B(0,r_1)} \langle \nabla u(y), \nabla\Gamma(x, y)\rangle\, dy \tag{3-3}$$

Taking the absolute value of both sides and using the triangle inequality on the right-hand side shows us that

$$|u(x)| \le \int_{\partial B(0,r_1)} \frac{|u(y)|}{3\omega_3|x-y|^2}\, dy + \int_{B(0,r_1)} \frac{|\nabla u(y)|}{3\omega_3|x-y|^2}\, dy. \tag{3-4}$$

We now integrate $|u|^2$ over $B(0, r_0)$ and use the well known inequality

$$(a+b)^2 \le 2(a^2+b^2) \quad \text{for } a, b \in \mathbb{R} \tag{3-5}$$

to obtain

$$\begin{aligned}\int_{B(0,r_0)} &|u(x)|^2\, dx \\ &\le 2\int_{B(0,r_0)} \left(\int_{\partial B(0,r_1)} \frac{|u(y)|}{3\omega_3|x-y|^2}\, dy\right)^2 + \left(\int_{B(0,r_1)} \frac{|\nabla u(y)|}{3\omega_3|x-y|^2}\, dy\right)^2 dx.\end{aligned} \tag{3-6}$$

To bound the second integral on the right hand side we make use of the mass formula (2-3) and the Riesz potential estimate (2-10) with $\mu = \frac{1}{3}$ and $q = p = 2$ to get

$$\int_{B(0,r_1)} \left(\int_{B(0,r_1)} \frac{|\nabla u(y)|}{3\omega_3 |x-y|^2}\, dy \right)^2 dx \le 8\pi r_1^2 m. \tag{3-7}$$

Using uniform asymptotic flatness (Definition 1.1), we estimate the first integral on the right as follows:

$$\begin{aligned} \int_{B(0,r_0)} \left(\int_{\partial B(0,r_1)} \frac{|u(y)|}{3\omega_3 |x-y|^2}\, dy \right)^2 &\le \frac{1}{9\omega_3^2} \int_{B(0,r_0)} \left(\int_{\partial B(0,r_1)} \frac{C}{|x-y|^2} \frac{1}{r_1}\, dy \right)^2 \\ &\le \frac{\omega_3 r_0^3 C^2 r_1^4}{(r_1 - r_0)^4 r_1^2}. \end{aligned} \tag{3-8}$$

Substituting the above two inequalities into (3-6), we obtain

$$\int_{B(0,r_0)} |u(x)|^2\, dx \le 2 \left[\frac{C^2 \omega_3 r_0^3 r_1^4}{(r_1 - r_0)^4 r_1^2} + 8\pi r_1^2 m \right] \tag{3-9}$$

If we let r_1 grow arbitrarily large, then the first term on the right will become arbitrarily small. We may counter any growth in the second term on the right by choosing the mass to be small enough. □

The next step is to estimate e^u. In order to do that we will apply the Moser–Trudinger inequality to u. Let us now recall the exact statement of the Moser–Trudinger inequality. Let $\Omega \subset \mathbb{R}^n$ and $\omega \in W_0^{1,n}(\Omega)$. Then there exists constants c_1 and c_2 depending only on n, such that

$$\int_\Omega \exp\left(\left(\frac{|\omega|}{c_1 \|\nabla \omega\|_n} \right)^{n/(n-1)} \right) \le c_2 |\Omega|. \tag{3-10}$$

This inequality appears as Theorem 7.15 in [Gilbarg and Trudinger 1998]. Lemma 3.1 gives $W^{1,2}$ control over u, so if we want to apply the Moser–Trudinger inequality, we will have to work over two dimensional domains. Luckily, we have the following almost trivial corollary to Lemma 3.1.

Corollary 3.2. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Let $\Omega_{\rho_0}^{\rho_1}$ denote the region*

$$\left\{ \rho_0 \le \rho \le \rho_1,\ |z| \le \frac{\rho_1}{2} \right\}. \tag{3-11}$$

For every $\epsilon > 0$, $\rho_0 > 0$ *and* $\rho_1 > \rho_0$ *there exists a* $\delta > 0$ *such that if the ADM mass of* $g \in \mathcal{M}$ *is less than* δ, *then*

$$\|u\|_{W^{1,2}(\Omega_{\rho_0}^{\rho_1})} < \epsilon. \tag{3-12}$$

Proof. Consider the region $\widetilde{\Omega}_{\rho_0}^{\rho_1} = \Omega_{\rho_0}^{\rho_1} \times [0, 2\pi)$. Choose r_0 large enough that

$$\widetilde{\Omega}_{\rho_0}^{\rho_1} \subset B_{r_0}. \tag{3-13}$$

In $\Omega_{\rho_0}^{\rho_1}$ we know that $\rho_0 \leq \rho$. Thus, we may observe that

$$\begin{aligned}\int_{\Omega_{\rho_0}^{\rho^1}} u^2 + |\nabla u|^2 \, d\rho \, dz &\leq \frac{1}{2\pi\rho_0} \int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} [u^2 + |\nabla u|^2] \rho \, d\rho \, dz \, d\phi \\ &\leq \frac{1}{2\pi\rho_0} \|u\|^2_{W^{1,2}(B_{r_0})}.\end{aligned} \tag{3-14}$$

Now we may apply Lemma 3.1. □

We're now in a position to estimate the $W^{1,p}$ norm of e^u. For the L^p norm of e^u the proof is an almost direct application of the Moser–Trudinger inequality. To estimate the L^p norm of $\nabla e^u = e^u \nabla u$, we use Hölder's inequality to analyze each term separately For the e^u term we will once again apply the Moser–Trudinger inequality. To estimate ∇u we will rely on Corollary 3.2.

Lemma 3.3. *Let* $\mathcal{M}$ *be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically outside of radius* R_0. *Let* $\Omega_{\rho_0}^{\rho_1}$ *denote the region* $\{(\rho, z) \mid \rho_0 \leq \rho \leq \rho_1, \ |z| \leq \frac{\rho_1}{2}\}$. *For every* $\rho_1 > \rho_0 > 0$, $\epsilon > 0$ *and* $p < 2$ *there exists a* $\delta > 0$ *such that if the ADM mass of* $g \in \mathcal{M}$ *is less than* δ, *then*

$$\|e^{|u|} - 1\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1})} < \epsilon. \tag{3-15}$$

Proof. Since g is smooth, u is bounded and has bounded derivatives in $\Omega_{\rho_0}^{\rho_1}$, though we have not made any assumption on what these bounds might be. Thus, $e^{|u|}$ is Lipschitz, and so

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla(e^{|u|} - 1)|^p = \int_{\Omega_{\rho_0}^{\rho_1}} |\nabla e^{|u|}|^p = \int_{\Omega_{\rho_0}^{\rho_1}} e^{p|u|} |\nabla u|^p. \tag{3-16}$$

Now, we let $r = \frac{2}{p}$ and r' be the conjugate exponent to r. After applying Hölder's inequality with r, we get

$$\int_{\Omega_{\rho_0}^{\rho_1}} e^{p|u|} |\nabla u|^p \leq \left(\int_{\Omega_{\rho_0}^{\rho_1}} e^{r' p|u|} \right)^{1/r'} \left(\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla u|^2 \right)^{p/2}. \tag{3-17}$$

Let $D(0, r_0)$ denote the two dimensional disk centered about the origin with radius r_0. Choose r_0 so that $\Omega_{\rho_0}^{\rho_1} \subset D(0, r_0)$. We may extend u to a function $\bar{u}$ in

$W^{1,2}_0(D(0,r_0))$; see Theorem 4.7 in [Evans and Gariepy 2015]. We may choose the extension $\bar{u}$ such that

$$\|\bar{u}\|_{W^{1,2}_0(D(0,r_0))} \le K\|u\|_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})}, \tag{3-18}$$

where the constant K is independent of the function u. A quick application of the Cauchy–Schwarz inequality gives us the estimate

$$r'p|\bar{u}| \le \tfrac{1}{4}(r'pc_1||\nabla\bar{u}||_2)^2 + \left(\frac{|\bar{u}|}{c_1||\nabla\bar{u}||_2}\right)^2, \tag{3-19}$$

where c_1 is the constant appearing in (3-10). We may now use the Moser–Trudinger inequality (3-10) to see that

$$\begin{aligned}\left(\int_{\Omega^{\rho_1}_{\rho_0}} e^{r'p|u|}\right)^{1/r'} &\le \left(\int_{D(0,r_0)} e^{r'p|\bar{u}|}\right)^{1/r'} \\ &\le \exp\left(\tfrac{1}{4}r'(pc_1||\nabla\bar{u}||_2)^2\right)(c_2|D(0,r_0)|)^{1/r'}.\end{aligned} \tag{3-20}$$

When written entirely in terms of u, the above inequality becomes

$$\left(\int_{\Omega^{\rho_1}_{\rho_0}} e^{r'p|u|}\right)^{1/r'} \le \exp\left[\frac{r'}{4}\left(Kpc_1||u||_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})}\right)^2\right](c_2|D(0,r_0)|)^{1/r'}. \tag{3-21}$$

Combining this with Corollary 3.2 gives

$$\int_{\Omega^{\rho_1}_{\rho_0}} |\nabla e^{|u|}|^p \le \exp\left[\frac{r'}{4}\left(Kpc_1||u||_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})}\right)^2\right](c_2|D(0,r_0)|)^{1/r'}\left(\frac{4m}{\rho_0}\right)^{p/2} \tag{3-22}$$

Now that we have successfully estimated $\nabla(e^{|u|}-1)$, we turn to estimating $e^{|u|}-1$. We use the expansion of $e^{|u|}$ to get that

$$\int_{\Omega^{\rho_1}_{\rho_0}} |e^{|u|}-1|^p = \int_{\Omega^{\rho_1}_{\rho_0}} \left(\sum_1^\infty \frac{|u|^k}{k!}\right)^p \tag{3-23}$$

Factoring out $|u|$ and over estimating the rest shows that the right hand side is bounded above by

$$\int_{\Omega^{\rho_1}_{\rho_0}} |u|^p e^{p|u|} \tag{3-24}$$

Now, we let $r = \frac{2}{p}$ and apply Hölder's inequality to get

$$\int_{\Omega^{\rho_1}_{\rho_0}} |u|^p e^{p|u|} \le \left(\int_{\Omega^{\rho_1}_{\rho_0}} |u|^2\right)^{p/2}\left(\int_{\Omega^{\rho_1}_{\rho_0}} e^{r'p|u|}\right)^{1/r'} \tag{3-25}$$

Finally, we may once again apply Corollary 3.2 and (3-21) to obtain the result. □

4. Sobolev estimates for $\alpha - 2u$ and $e^{\alpha-2u}$

We must now concentrate on estimating $\alpha - 2u$ and $e^{\alpha-2u}$. We will try to imitate as closely as possible the steps which let us successfully estimate u and e^u. First, we obtain $W^{1,p}$ estimates for $\alpha - 2u$ from the mass formula (2-3). Unfortunately, even at this early stage, the process is notably harder than it was for u.

In our attempt to estimate the $W^{1,2}$ norm of u we used a representation formula to express u in terms of its values on a large sphere and its gradient in a large ball. Then we used the asymptotic falloff and the mass formula to control these quantities, respectively. This was a relatively simple process because $\|\nabla u\|$ is a term in the mass formula. However, the gradient of $\alpha - 2u$ does not appear directly in the mass formula. Rather, it is the Laplacian of $\alpha - 2u$ which appears in the mass formula by way of the scalar curvature equation. We will see the precise nature of this relationship in the following lemmas. For now, the important point is that instead of using (2-4) to express $\alpha - 2u$, we should use Green's representation (2-8). It is widely known that one may replace the fundamental solution Γ in (2-8) with a function $G(x, y)$, the Green's function of the domain, which vanishes on the boundary of the domain. This choice simplifies Green's representation formula significantly. Unfortunately, the explicit formula for $G(x, y)$ can be complicated depending on the domain. Thus, although our representation formula has been simplified, it is difficult to estimate $G(x, y)$. Luckily, we are working over very simple domains, namely the rectangles $\Omega^{\rho_1}_{\rho_0}$. Therefore, a compromise is possible. We may simplify the representation formula for any one side of the rectangle. Specifically, we may choose a "Green's" function which vanishes, or whose normal derivative vanishes, on one side of the rectangle. Since we have the least amount of a priori knowledge about the metric near the axis of symmetry, we will choose to simplify our representation formula on the side nearest the axis of symmetry.

For the rectangle $\Omega^{\rho_1}_{\rho_0}$, let $\bar{x}$ denote the reflection of the point x about the vertical line $\{\rho = \rho_0\}$. We can define the following two functions

$$H_N(x, y) = \frac{1}{2\pi}\log(|x - y|) + \frac{1}{2\pi}\log(|\bar{x} - y|) \tag{4-1}$$

and

$$H_D(x, y) = \frac{1}{2\pi}\log(|x - y|) - \frac{1}{2\pi}\log(|\bar{x} - y|). \tag{4-2}$$

A quick check shows that we may replace Γ by either H_N or H_D in (2-8). Furthermore, a calculation shows that

$$\frac{\partial H_N(x, y)}{\partial \nu}\Big|_{\partial\Omega^{\rho_1}_{\rho_0}\cap\{\rho=\rho_0\}} = 0 \tag{4-3}$$

and

$$H_D(x,y)|_{\partial\Omega_{\rho_0}^{\rho_1}\cap\{\rho=\rho_0\}} = 0. \tag{4-4}$$

Since we will be integrating against the functions H_N and H_D in what follows, and since H_N and H_D are sums of functions of the form $\log(|x-y|)$, it will be useful in what follows to have an L^p estimate for $\log(|x-y|)$ over bounded regions.

Lemma 4.1. *Let Ω be a bounded region in $\mathbb{R}^2$ and let*

$$r_0 = \max\{\mathrm{diam}(\Omega), 1\}. \tag{4-5}$$

Then for $y \in \mathrm{cl}(\Omega)$ we have

$$\int_\Omega |\log(|x-y|)|^k\,dx \le \frac{\pi k!}{2^k} + 2\pi(r_0-1)r_0\log(r_0)^k \tag{4-6}$$

for positive integers k.

Proof. We observe that

$$\begin{aligned}\int_\Omega |\log(|x-y|)|^k\,dx &\le \int_{B(y,r_0)} |\log(|x-y|)|^k\,dx \\ &= \int_0^1 (-1)^k 2\pi r\log(r)^k\,dr + \int_1^{r_0} 2\pi r\log(r)^k\,dr\end{aligned} \tag{4-7}$$

The second term on the right has the simple estimate

$$2\pi(r_0-1)r_0\log(r_0)^k. \tag{4-8}$$

To estimate the first term, one must carry out the integration. By induction, we have the following result.

$$\int_0^1 (-1)^k 2\pi r\log(r)^k dr = \frac{\pi k!}{2^k}. \tag{4-9}$$

□

With all of this in mind, we begin the process of estimating the $W^{1,p}$ norm of $\alpha - 2u$.

Proposition 4.2. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose in addition that $\mathcal{M}$ is radially monotone at ρ_0. For every $\rho_1 > \rho_0$, $\epsilon > 0$ and $1 \le p < 2$ there exists a $\delta > 0$ such that if the ADM mass of $g \in \mathcal{M}$ is less than δ, then*

$$\|\alpha - 2u\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1})} < \epsilon \tag{4-10}$$

Applying Green's representation formula to $\alpha - 2u$ over the domain $\Omega_{\rho_0}^{\rho_1}$ gives us

$$(\alpha - 2u)(x) = \int_{\partial\Omega_{\rho_0}^{\rho_1}} \left[(\alpha - 2u) \frac{\partial H_N(x,y)}{\partial \nu} - H_N(x,y) \frac{\partial(\alpha - 2u)}{\partial \nu} \right] dy + \int_{\Omega_{\rho_0}^{\rho_1}} H_N(x,y) \Delta(\alpha - 2u)\, dy. \tag{4-11}$$

The above representation breaks our problem into two pieces. First we must estimate $\Delta(\alpha - 2u)$ over $\Omega_{\rho_0}^{\rho_1}$ and then we must estimate $\alpha - 2u$ on the boundary of $\Omega_{\rho_0}^{\rho_1}$. The necessary estimates are the content of the following two lemmas.

Lemma 4.3. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. If g is a metric in $\mathcal{M}$ and*

$$m(g) \le m, \tag{4-12}$$

then

$$\|\Delta(2u - \alpha)\|_{L^1(\Omega_{\rho_0}^{\rho_1})} \le \frac{4m}{\rho_0} + \frac{4\sqrt{\rho_1 m}}{\rho_0} \tag{4-13}$$

for any $\rho_1 > \rho_0 > 0$.

Proof. We must relate $\Delta(\alpha - 2u)$ to the mass formula. First, we recall that the scalar curvature equation is

$$R_g = 4e^{2(u-\alpha)} \left[\Delta_{\mathbb{R}^3}\left(u - \tfrac{1}{2}\alpha\right) - \tfrac{1}{2}|\nabla u|_\delta^2 + \frac{1}{2\rho}\frac{\partial \alpha}{\partial \rho} - \frac{\rho^2 e^{-2\alpha}}{8}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial \rho}\right)^2 \right], \tag{4-14}$$

where we have written $\Delta_{\mathbb{R}^3}$ to emphasize the fact that it is the three dimensional Laplacian which appears, and not the two dimensional Laplacian Δ. However, if we remember that all of the functions involved don't depend on ϕ, then we can see that

$$\Delta_{\mathbb{R}^3}\left(u - \frac{\alpha}{2}\right) = \Delta\left(u - \frac{\alpha}{2}\right) + \frac{1}{2\rho}\frac{\partial(2u - \alpha)}{\partial \rho}. \tag{4-15}$$

By plugging the above into the scalar curvature equation, we get

$$R_g = 4e^{2(u-\alpha)} \left[\Delta\left(u - \tfrac{1}{2}\alpha\right) - \tfrac{1}{2}|\nabla u|_\delta^2 + \frac{1}{\rho}\frac{\partial u}{\partial \rho} - \frac{\rho^2 e^{-2\alpha}}{8}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial \rho}\right)^2 \right]. \tag{4-16}$$

We now solve the scalar curvature equation for $\Delta(\alpha - 2u)$ and integrate in order to arrive at

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\Delta(\alpha-2u)|\, d\rho\, dz \leq \int_{\Omega_{\rho_0}^{\rho_1}} \frac{e^{2(\alpha-u)}}{2} R_g + |\nabla u|_\delta^2 + \frac{2}{\rho}\left|\frac{\partial u}{\partial \rho}\right| + \frac{\rho^2 e^{-2\alpha}}{4}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial \rho}\right)^2 d\rho\, dz. \tag{4-17}$$

Now, since we are integrating over a region in which $\rho \geq \rho_0$, we have from the mass formula (2-3) that

$$\int_{\Omega_{\rho_0}^{\rho_1}} \frac{e^{2(\alpha-u)}}{2} R_g + |\nabla u|_\delta^2 + \frac{\rho^2 e^{-2\alpha}}{4}\left(\frac{\partial B}{\partial z} - \frac{\partial A}{\partial \rho}\right)^2 d\rho\, dz \leq \frac{4m}{\rho_0}. \tag{4-18}$$

To estimate the final term on the right hand side of (4-17) requires only a little more work. Namely, if we apply Hölder's inequality to

$$\int_{\Omega_{\rho_0}^{\rho_1}} \frac{2}{\rho}\left|\frac{\partial u}{\partial \rho}\right| d\rho\, dz \tag{4-19}$$

and make the simple estimate $\left|\frac{\partial u}{\partial \rho}\right| \leq |\nabla u|_\delta$, then we obtain

$$\int_{\Omega_{\rho_0}^{\rho_1}} \frac{2}{\rho}\left|\frac{\partial u}{\partial \rho}\right| d\rho\, dz \leq \left(\int_{\Omega_{\rho_0}^{\rho_1}} \frac{4}{\rho^2}\right)^{1/2}\left(\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla u|_\delta^2\, d\rho\, dz\right)^{1/2}. \tag{4-20}$$

Using the mass formula once more, we see that

$$\left(\int_{\Omega_{\rho_0}^{\rho_1}} \frac{4}{\rho^2}\right)^{1/2}\left(\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla u|_\delta^2\, d\rho\, dz\right)^{1/2} \leq \frac{4\sqrt{\rho_1 m}}{\rho_0}. \tag{4-21}$$

Putting each of these estimates together gives the desired result. □

We now want to estimate boundary terms on $\partial\Omega_{\rho_0}^{\rho_1}$. Due to the asymptotic falloff conditions (Definition 1.1), it is relatively straight forward to estimate terms on $(\partial\Omega_{\rho_0}^{\rho_1}) - \{\rho = \rho_0\}$. It is more difficult to estimate terms on $(\partial\Omega_{\rho_0}^{\rho^1}) \cap \{\rho = \rho_0\}$.

Lemma 4.4. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Assume that $\mathcal{M}$ is also radially monotone at ρ_0. For $\rho_1 > \max\{\rho_0, R_0\}$, if $g \in \mathcal{M}$ and*

$$m(g) \leq m, \tag{4-22}$$

then

$$\int_{(\partial\Omega_{\rho_0}^{\rho_1})\cap\{\rho=\rho_0\}} \left|\frac{\partial(\alpha-2u)}{\partial \nu}\right| \leq \frac{4m}{\rho_0} + \frac{4\sqrt{\rho_1 m}}{\rho_0} + \frac{6\pi C}{\rho_1}, \tag{4-23}$$

where the constant C is the one appearing in Definition 1.1.

Proof. It is an easy observation that

$$\frac{\partial}{\partial \nu}\Big|_{\partial\Omega_{\rho_0}^{\rho_1}\cap\{\rho=\rho_0\}} = -\frac{\partial}{\partial \rho}. \tag{4-24}$$

If we write the radial monotonicity condition entirely in terms of coordinate functions, then we may see that for $g \in \mathcal{M}$

$$\frac{\partial(\alpha-2u)}{\partial\rho}(\rho_0, z) \leq 0. \tag{4-25}$$

Thus, we observe that

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}\cap\{\rho=\rho_0\}}\left|\frac{\partial(\alpha-2u)}{\partial\nu}\right| = -\int_{-\rho_1/2}^{\rho_1/2}\frac{\partial(\alpha-2u)}{\partial\rho}(\rho, z)\,dz. \tag{4-26}$$

A quick application of Stokes' Theorem over the region

$$\left\{\rho_0 \leq \rho,\ |z| \leq \frac{\rho_1}{2}\right\} \tag{4-27}$$

gives

$$\begin{aligned}\int_{-\rho_1/2}^{\rho_1/2}\frac{\partial(\alpha-2u)}{\partial\rho}&\\ &= -\int_{\{\rho_0\leq\rho,\ |z|\leq\rho_1/2\}}\Delta(\alpha-2u)\,d\rho\,dz + \int_{\{\rho\geq\rho_0,\ |z|=\rho_1/2\}}\frac{\partial(\alpha-2u)}{\partial z}.\end{aligned} \tag{4-28}$$

We may estimate the second integral on the right by plugging in the asymptotic estimates (Definition 1.1). The result is the following inequality

$$\left|\int_{\{\rho\geq\rho_0,\ |z|=\rho_1/2\}}\frac{\partial(\alpha-2u)}{\partial z}\right| \leq \int_{\{\rho\geq\rho_0,\ |z|=\rho_1/2\}}\frac{3C}{|(\rho, z)|^2}\,d\rho. \tag{4-29}$$

We may see by a straightforward integration that

$$\left|\int_{\{\rho\geq\rho_0,\ |z|=\rho_1/2\}}\frac{\partial(\alpha-2u)}{\partial z}\right| \leq \frac{6\pi C}{\rho_1}. \tag{4-30}$$

The last piece of the puzzle is the term

$$\left|\int_{\{\rho_0\leq\rho,\ |z|\leq\rho_1/2\}}\Delta(\alpha-2u)\,d\rho\,dz\right| \leq \int_{\{\rho_0\leq\rho,\ |z|\leq\rho_1/2\}}|\Delta(\alpha-2u)|\,d\rho\,dz. \tag{4-31}$$

We now use the proof of Lemma 4.3 to bound this term. Putting everything together, we get

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}\cap\{\rho=\rho_0\}}\left|\frac{\partial(\alpha-2u)}{\partial\nu}\right| \leq \frac{4m}{\rho_0} + \frac{4\sqrt{\rho_1 m}}{\rho_0} + \frac{6\pi C}{\rho_1}. \tag{4-32}$$

□

We have the necessary estimates to obtain $W^{1,p}$ control over $\alpha - 2u$.

Proof of Proposition 4.2. Consider $\Omega_{\rho_0}^{\tilde{\rho}_1}$ for some $\tilde{\rho}_1 \geq R_0$. We also choose $\tilde{\rho}_1$ to be much larger than ρ_1. As before, we let

$$H_N(x, y) = \frac{1}{2\pi}\log(|x-y|) + \frac{1}{2\pi}\log(|\bar{x}-y|), \tag{4-33}$$

where $\bar{x}$ is the reflection of x about the line $\{\rho = \rho_0\}$. Recall that Green's representation gives us the following formula for $\alpha - 2u$:

$$\begin{aligned}(\alpha-2u)(x) = \int_{\partial\Omega_{\rho_0}^{\tilde{\rho}_1}} (\alpha-2u)(y)\frac{\partial H_N}{\partial\nu}(x, y) - H_N(x, y)\frac{\partial(\alpha-2u)}{\partial\nu}(y)\, dy \\ + \int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} H_N(x, y)\Delta(\alpha-2u)(y)\, dy.\end{aligned} \tag{4-34}$$

We will imitate the estimates that we made for u in Corollary 3.2. Namely, we see that

$$\int_{\Omega_{\rho_0}^{\rho_1}} |(\alpha-2u)(x)|^p\, dx \tag{4-35}$$

is bounded above by

$$\begin{aligned}C(p)\int_{\Omega_{\rho_0}^{\rho_1}}\left(\int_{\partial\Omega_{\rho_0}^{\tilde{\rho}_1}}\left|H_N\frac{\partial(\alpha-2u)}{\partial\nu}\right| + \left|(\alpha-2u)\frac{\partial H_N}{\partial\nu}\right| dy\right)^p \\ + \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |H_N\Delta(\alpha-2u)|\, dy\right)^p dx,\end{aligned} \tag{4-36}$$

for some constant $C(p)$ depending only on p. We estimate each of the three terms above in turn. For the first two terms, we will break $\partial\Omega_{\rho_0}^{\tilde{\rho}_1}$ into

$$\partial\Omega_{\rho_0}^{\tilde{\rho}_1} - \{\rho = \rho_0\} \tag{4-37}$$

and

$$(\partial\Omega_{\rho_0}^{\tilde{\rho}_1}) \cap \{\rho = \rho_0\}. \tag{4-38}$$

Let's start with (4-37). For this piece of the boundary we can use the uniform asymptotically flat condition to obtain the required estimates. First, notice that for x in $\Omega_{\rho_0}^{\rho_1}$ and y in (4-37) we have

$$|H_N(x, y)| \leq \frac{\log\big(2\,\mathrm{diam}(\Omega_{\rho_0}^{\tilde{\rho}_1})\big)}{\pi} \leq \frac{\log(2\sqrt{2}\tilde{\rho}_1)}{\pi}, \tag{4-39}$$

since $\tilde{\rho}_1$ is much larger than ρ_0. From the asymptotic falloff given in Definition 1.1, we see that for y in (4-37)

$$\left|\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right| \leq \frac{3C}{\tilde{\rho}_1^2}. \tag{4-40}$$

Thus, we may see that

$$\begin{aligned}(4\text{-}41)\quad &\int_{\Omega^{\rho_1}_{\rho_0}}\left(\int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}-\{\rho=\rho_0\}}\left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right|dy\right)^p dx\\ &\leq \int_{\Omega^{\rho_1}_{\rho_0}}\left(\frac{9\log(2\sqrt{2}\tilde{\rho}_1)C}{\pi\tilde{\rho}_1}\right)^p dx\\ &\leq \rho_1^2\left(\frac{3\log(2\sqrt{2}\tilde{\rho}_1)C}{\tilde{\rho}_1}\right)^p.\end{aligned}$$

The other term has a similar estimate:

$$(4\text{-}42)\quad \int_{\Omega^{\rho_1}_{\rho_0}}\left(\int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}-\{\rho=\rho_0\}}\left|(\alpha-2u)\frac{\partial H_N}{\partial\nu}\right|dy\right)^p dx \leq \rho_1^2\left(\frac{6C}{\tilde{\rho}_1-\rho_1}\right)^p.$$

Using the two estimates above, we see that

$$\begin{aligned}(4\text{-}43)\quad &\int_{\Omega^{\rho_1}_{\rho_0}}\left(\int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}-\{\rho=\rho_0\}}\left|H_N\frac{\partial(\alpha-2u)}{\partial\nu}\right|+\left|(\alpha-2u)\frac{\partial H_N}{\partial\nu}\right|dy\right)^p\\ &\leq C(p)\left(\rho_1^2\left(\frac{3\log(2\sqrt{2}\tilde{\rho}_1)C}{\tilde{\rho}_1}\right)^p+\rho_1^2\left(\frac{6C}{\tilde{\rho}_1-\rho_1}\right)^p\right).\end{aligned}$$

We can now move to the inner piece of the boundary, (4-38). We will further divide $\partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0\}$ into

$$(4\text{-}44)\quad \partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0,\ |z|\leq\rho_1\}$$

and

$$(4\text{-}45)\quad \partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0,\ |z|\geq\rho_1\}.$$

We now estimate

$$(4\text{-}46)\quad \left(\int_{\Omega^{\rho_1}_{\rho_0}}\left(\int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0,\ |z|\leq\rho_1\}}\left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}\right|dy\right)^p dx\right)^{1/p}.$$

Here we apply Minkowski's inequality for integrals [Folland 1999] to bound the above by

$$(4\text{-}47)\quad \int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0,\ |z|\leq\rho_1\}}\left(\int_{\Omega^{\rho_1}_{\rho_0}}\left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}\right|^p dx\right)^{1/p}dy.$$

We may rewrite this expression as

$$(4\text{-}48)\quad \int_{\partial\Omega^{\tilde{\rho}_1}_{\rho_0}\cap\{\rho=\rho_0,\ |z|\leq\rho_1\}}\left|\frac{\partial(\alpha-2u)}{\partial\nu}\right|\left(\int_{\Omega^{\rho_1}_{\rho_0}}|H_N(x,y)|^p\,dx\right)^{1/p}dy.$$

In view of Lemma 4.4, we must estimate the L^P norm of $H_N(x, y)$ as a function of x over $\Omega_{\rho_0}^{\rho_1}$ for each y in

$$\partial\Omega_{\rho_0}^{\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \leq \rho_1\}. \tag{4-49}$$

We see that the points x and y are both contained in $\Omega_{\rho_0}^{2\rho_1}$, Which has diameter $2\sqrt{2}\rho_1$. Let

$$F(x) = \bar{x}. \tag{4-50}$$

Since F is an isometry, if we apply the change of variable formula to F and note that $y = \bar{y}$ for y in $\{\rho = \rho_0\}$, then we may see that for any q, we have

$$\int_{\Omega_{\rho_0}^{2\rho_1}} |\log(|\bar{x} - y|)|^q\, dx = \int_{F(\Omega_{\rho_0}^{2\rho_1})} |\log(|x - y|)|^q\, dx. \tag{4-51}$$

Thus, we may use (4-6) to see that

$$\int_{\Omega_{\rho_0}^{\rho_1}} |H_N(x, y)|\, dx \leq \int_{\Omega_{\rho_0}^{2\rho_1}} |H_N(x, y)| \leq \tfrac{1}{2} + 16\rho_1^2 \log(2\sqrt{2}\rho_1), \tag{4-52}$$

and

$$\begin{aligned}\left(\int_{\Omega_{\rho_0}^{\rho_1}} |H_N(x, y)|^2\, dx\right)^{1/2} &\leq \left(\int_{\Omega_{\rho_0}^{2\rho_1}} |H_N(x, y)|^2\, dx\right)^{1/2} \\ &\leq \frac{1}{2\pi}\sqrt{2\pi + 64\pi\rho_1^2 \log(2\sqrt{2}\rho_1)^2}.\end{aligned} \tag{4-53}$$

We do a simple interpolation between the above two estimates to get

$$\begin{aligned}\left(\int_{\Omega_{\rho_0}^{\rho_1}} |H_N(x, y)|^p\, dx\right)^{1/p} &\leq \left(\tfrac{1}{2}+16\rho_1^2 \log(2\sqrt{2}\rho_1)\right)^{(2-p)/p} \\ &\quad \times\left(\frac{1}{2\pi}\sqrt{2\pi+64\pi\rho_1^2 \log(2\sqrt{2}\rho_1)^2}\right)^{(2p-2)/p},\end{aligned} \tag{4-54}$$

We now combine the above with Lemma 4.4 to bound (4-48) by

$$\begin{aligned}&\left[\tfrac{1}{2} + 16\rho_1^2 \log(2\sqrt{2}\rho_1)\right]^{(2-p)/p}\left[\frac{1}{2\pi}\sqrt{2\pi + 64\pi\rho_1^2 \log(2\sqrt{2}\rho_1)^2}\right]^{(2p-2)/p} \\ &\qquad \times\left(\frac{4m}{\rho_0} + \frac{4\sqrt{\tilde{\rho}_1 m}}{\rho_0} + \frac{6\pi C}{\tilde{\rho}_1}\right)\end{aligned} \tag{4-55}$$

The term

$$\left(\int_{\Omega_{\rho_0}^{\rho_1}}\left(\int_{\partial\Omega_{\rho_0}^{\tilde{\rho}_1}\cap\{\rho=\rho_0,\ |z|\geq\rho_1\}}\left|H_N(x, y)\frac{\partial(\alpha - 2u)}{\partial\nu}\right| dy\right)^p dx\right)^{1/p} \tag{4-56}$$

is much easier to estimate. In fact, for x in $\Omega_{\rho_0}^{\rho_1}$ and y in $\partial\Omega_{\rho_0}^{\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \geq \rho_1\}$, we have

$$|H_N(x, y)| \leq \frac{1}{\pi} \max\left\{\left|\log\left(\frac{\rho_1}{2}\right)\right|, |\log(2\sqrt{2}\tilde{\rho}_1)|\right\}. \tag{4-57}$$

Once again, combining the above with Lemma 4.4 bounds (4-56) by

$$\frac{(\rho_1)^{2/p}}{\pi} \max\left\{\left|\log(\frac{\rho_1}{2})\right|, |\log(2\sqrt{2}\tilde{\rho}_1)|\right\}\left(\frac{4m}{\rho_0} + \frac{4\sqrt{\tilde{\rho}_1 m}}{\rho_0} + \frac{6\pi C}{\tilde{\rho}_1}\right). \tag{4-58}$$

The final piece of the puzzle is the estimate of

$$\int_{\Omega_{\rho_0}^{\rho_1}} \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} \left|H_N(x, y)\Delta(\alpha - 2u)(y)\right| dy\right)^p dx. \tag{4-59}$$

Here we may use Minkowski's inequality for integrals once more to see that the above is bounded by

$$\left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |\Delta(\alpha - 2u)(y)| \left(\int_{\Omega_{\rho_0}^{\rho_1}} |H_N(x, y)|^p\, dx\right)^{1/p} dy\right)^p. \tag{4-60}$$

Thus, we may bound (4-59) by

$$\left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |\Delta(\alpha - 2u)(y)| \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1} \cup F(\Omega_{\rho_0}^{\tilde{\rho}_1})} |H_N(x, y)|^p\, dx\right)^{1/p} dy\right)^p. \tag{4-61}$$

Again, using the change of variable formula and (4-6), we bound (4-59) by

$$\left(\left[\tfrac{1}{2} + 16\tilde{\rho}_1^2 \log(2\sqrt{2}\tilde{\rho}_1)\right]^{(2-p)/p} \left[\frac{1}{2\pi}\sqrt{2\pi + 64\pi\tilde{\rho}_1^2 \log(2\sqrt{2}\tilde{\rho}_1)^2}\right]^{(2p-2)/p} \times \frac{4m + 4\sqrt{\tilde{\rho}_1 m}}{\rho_0}\right)^p. \tag{4-62}$$

Putting everything above together shows that

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\alpha - 2u|^p \leq C(p)^2 (4\text{-}43) + C(p)^3\left((4\text{-}55)^p + (4\text{-}58)^p\right) + C(p)(4\text{-}62). \tag{4-63}$$

Thus, for any $\epsilon > 0$ and $\rho_1 > \rho_0$ we can pick an appropriate $\tilde{\rho}_1$ and ADM mass m so that

$$\|\alpha - 2u\|_{L^p(\Omega_{\rho_0}^{\rho_1})} < \frac{\epsilon}{2}. \tag{4-64}$$

We can get similar estimates for $\nabla(\alpha - 2u)$ by differentiating the representation formula:

$$\nabla(\alpha-2u)(x) = \int_{\partial\Omega_{\rho_0}^{\tilde{\rho}_1}} (\alpha-2u)\nabla_x \frac{\partial H_N}{\partial \nu} - \nabla_x H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu} + \int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} (\nabla_x H_N)\Delta(\alpha-2u). \tag{4-65}$$

We see that

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla(\alpha-2u)|^p \leq C(p)\int_{\Omega_{\rho_0}^{\rho_1}} \left(\int_{\partial\Omega_{\rho_0}^{\tilde{\rho}_1}} \left|\frac{\partial(\alpha-2u)}{\partial\nu}\nabla_x H_N\right| + \left|(\alpha-2u)\nabla_x\frac{\partial H_N}{\partial\nu}\right|\right)^p + \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |\nabla_x H_N|\,|\Delta(\alpha-2u)|\right)^p. \tag{4-66}$$

As before, we will break $\partial\Omega_{\rho_0}^{\tilde{\rho}_1}$ into (4-37) and (4-38). We start with (4-37). A quick calculation shows that

$$|\nabla_x H_N| \leq \frac{1}{2\pi}\left(\frac{1}{|x-y|} + \frac{1}{|\bar{x}-y|}\right) \tag{4-67}$$

and

$$\left|\nabla_x \frac{\partial H_N}{\partial\nu}\right| \leq \frac{3}{2\pi}\left(\frac{1}{|x-y|^2} + \frac{1}{|\bar{x}-y|^2}\right). \tag{4-68}$$

Estimating the integral over (4-37) now proceeds as before.

As a first step in estimating the integral over (4-38), we note that

$$\nabla_x \frac{\partial H_N}{\partial\nu}\bigg|_{\{\rho=\rho_0\}} = 0. \tag{4-69}$$

Next, we again break (4-38) into (4-44) and (4-45). For both pieces we proceed much as we did before. On (4-44) it is crucial that $p < 2$, since it is only then that the integral

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla_x H_N|^p \tag{4-70}$$

is bounded for all y in (4-44). For (4-45), the necessary changes in the argument are straightforward.

Finally, to estimate

$$\int_{\Omega_{\rho_0}^{\rho_1}} \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |(\nabla_x H_N)\Delta(\alpha-2u)|\right)^p \leq \int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} \left(\int_{\Omega_{\rho_0}^{\tilde{\rho}_1}} |(\nabla_x H_N)\Delta(\alpha-2u)|\right)^p \tag{4-71}$$

we may use the Riesz potential estimates (2-10) with the appropriate choice of constants. Thus, for $\tilde{\rho}_1$ chosen large enough and m chosen small enough, we may conclude that

$$\|\alpha - 2u\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1})} < \epsilon. \tag{4-72}$$

□

In the course of proving Proposition 4.2 we actually proved a little more. For future convenience, we record this result as the following corollary.

Corollary 4.5. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is radially monotone at ρ_0. For any $\rho_1 > \rho_0$, $\epsilon > 0$, and $1 \le p < 2$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{4-73}$$

then

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\alpha - 2u|^p \le \frac{\epsilon}{\rho_0^p} \tag{4-74}$$

and

$$\int_{\Omega_{\rho_0}^{\rho_1}} |\nabla(\alpha - 2u)|^p \le \frac{\epsilon}{\rho_0^p}. \tag{4-75}$$

Having successfully estimated the $W^{1,p}$ norm of $\alpha - 2u$, we must now turn to estimating the $W^{1,p}$ norm of $e^{\alpha-2u}$. As was noted earlier, control over the $W^{1,p}$ norm of $\alpha - 2u$ for $1 \le p < 2$ falls short of what we need to apply the Moser–Trudinger inequality to $\alpha - 2u$. It is thus not immediately clear how to turn estimates for $\alpha - 2u$ into estimates for $e^{\alpha-2u}$. Luckily, the special nature of the fundamental solution to the Laplacian in two dimensions allows us to prove a Moser–Trudinger like inequality which we can use on $\alpha - 2u$.

Lemma 4.6. *Let Ω be a bounded domain in the plane on which the divergence theorem holds and let Γ be the fundamental solution for the Laplacian. Suppose we have $\psi \in C^2(\Omega) \cap C^1(\bar{\Omega})$ and $\Delta\psi \in L^1(\Omega)$. Let Ω_σ denote $= \{x \in \Omega : d(x, \partial\Omega) \ge \sigma\}$ and let $r_0 = \max\{1, \operatorname{diam}(\Omega)\}$. Then we have the estimate*:

$$\int_{\Omega_\sigma} e^{|\psi|} \le \left(|\Omega_\sigma| + \frac{\pi \|\Delta\psi\|_1}{4\pi - \|\Delta\psi\|_1} + 2\pi(r_0 - 1) r_0 [r_0^{\|\Delta\psi\|_1/2\pi} - 1] \right) \times \sup_{x \in \Omega_\sigma} \exp\left(\int_{\partial\Omega} \left| \psi(y) \frac{\partial\Gamma}{\partial\nu}(x, y) \right| + \left| \Gamma(x, y) \frac{\partial\psi}{\partial\nu}(y) \right| dy \right) \tag{4-76}$$

Proof. From Green's representation we have

$$\psi(x) = \int_{\partial\Omega} \psi(y) \frac{\partial\Gamma}{\partial\nu}(x, y) - \Gamma(x, y) \frac{\partial\psi}{\partial\nu}(y)\, dy + \int_{\Omega} \Gamma(x, y) \Delta\psi(y)\, dy \tag{4-77}$$

Using the representation formula to rewrite $\int_{\Omega_\epsilon} e^{|\psi|}$, we obtain

$$\int_{\Omega_\sigma} e^{|\psi(x)|}\,dx \le \int_{\Omega_\sigma} \exp\biggl[\int_{\partial\Omega}\biggl|\psi(y)\frac{\partial\Gamma}{\partial\nu}(x,y) - \Gamma(x,y)\frac{\partial\psi}{\partial\nu}(y)\biggr|\,dy\biggr] \times \exp\biggl[\int_{\Omega}\bigl|\Gamma(x,y)\Delta\psi(y)\bigr|\,dy\biggr]\,dx \tag{4-78}$$

We bound the first term on the right pointwise by its supremum over Ω_σ. Then we may take it outside of the integrand.

$$\int_{\Omega_\sigma} e^{|\psi(x)|}\,dx \le \sup_{x\in\Omega_\sigma} \exp\biggl[\int_{\partial\Omega}\biggl|\psi(y)\frac{\partial\Gamma}{\partial\nu}(x,y)\biggr| + \biggl|\Gamma(x,y)\frac{\partial\psi}{\partial\nu}(y)\biggr|\,dy\biggr] \times \int_{\Omega_\sigma} \exp\biggl[\int_{\Omega}\bigl|\Gamma(x,y)\Delta\psi(y)\bigr|\,dy\biggr]\,dx \tag{4-79}$$

We may now concentrate on estimating

$$\int_{\Omega_\sigma} \exp\biggl[\int_{\Omega}\bigl|\Gamma(x,y)\Delta\psi(y)\bigr|\,dy\biggr] \tag{4-80}$$

The strategy is to expand the above integral using the Taylor series for the exponential function and then bound each term appearing in the expansion:

$$\int_{\Omega_\sigma}\bigl(e^{\int_\Omega|\Gamma(x,y)\Delta(\alpha-2u)(y)|\,dy}\bigr)\,dx = \sum_{k=0}^{\infty}\int_{\Omega_\sigma}\frac{\bigl(\int_\Omega|\Gamma(x,y)\Delta\psi(y)|\,dy\bigr)^k}{k!}\,dx. \tag{4-81}$$

First, recall that the fundamental solution of the Laplacian in two dimensions is given by

$$\frac{1}{2\pi}\log|x-y| \tag{4-82}$$

Second, after observing that $\Omega_\sigma \subset \Omega$, and pulling constants out, we get the inequality

$$\int_{\Omega_\sigma}\frac{\bigl|\int_\Omega \Gamma(x-y)\Delta\psi(y)\,dy\bigr|^k}{k!} \le \frac{1}{k!(2\pi)^k}\int_\Omega\biggl(\int_\Omega|\Delta\psi(y)|\,|\log(|x-y|)|\,dy\biggr)^k dx \tag{4-83}$$

We apply Jensen's inequality to the integral on the right to obtain

$$\frac{1}{(2\pi)^k k!}\int_\Omega\biggl(\int_\Omega|\log(|x-y|)|\,|\Delta\psi(y)|\,dy\biggr)^k dx \le \frac{\|\Delta\psi\|_1^{k-1}}{(2\pi)^k k!}\int_\Omega\int_\Omega|\log(|x-y|)|^k|\Delta\psi(y)|\,dy\,dx \tag{4-84}$$

We now use Tonelli's theorem to switch the order of integration to get

$$\int_\Omega\int_\Omega|\log(|x-y|)|^k\,|\Delta\psi(y)|\,dy\,dx = \int_\Omega|\Delta\psi(y)|\int_\Omega|\log(|x-y|)|^k\,dx\,dy \tag{4-85}$$

Putting (4-6), (4-85), and (4-84) together gives

$$\frac{1}{k!}\int_\Omega\left|\int_\Omega \frac{1}{2\pi}\log(|x-y|)\Delta\psi(y)\,dy\right|^k dx \le \frac{\|\Delta\psi\|_1^k}{(2\pi)^k k!}\left(\frac{\pi k!}{2^k}+2\pi(r_0-1)r_0\log(r_0)^k\right) \tag{4-86}$$

After a quick application of the monotone convergence theorem to the summation over k from $k=1$ to infinity of (4-83) we get

$$\int_{\Omega_\sigma} e^{\left|\int_\Omega \Gamma(x,y)\Delta\psi(y)dy\right|}\,dx \le |\Omega_\sigma| + \frac{\pi\|\Delta\psi\|_1}{4\pi-\|\Delta\psi\|_1} + (r_0-1)r_0\left[\exp\left(\frac{\log(r_0)\|\Delta\psi\|_1}{2\pi}\right)-1\right]. \qquad \square \tag{4-87}$$

We have the following corollary, which is the actual inequality we will use.

Corollary 4.7. *Suppose* $\psi\in C^2(\Omega_{\rho_0}^{\rho_1})\cap C^1(\mathrm{cl}(\Omega_{\rho_0}^{\rho_1}))$ *and let* $r_0=\max\{1,\,\mathrm{diam}(\Omega_{\rho_0}^{\rho_1})\}$. *Then*

$$\int_{(\Omega_{\rho_0}^{\rho_1})_\sigma} e^{|\psi|} \tag{4-88}$$

is bounded above by

$$e^{C(\sigma,\rho_1)\|\Delta\psi\|_1}\left(|(\Omega_{\rho_0}^{\rho_1})_\sigma| + \frac{\pi\|\Delta\psi\|_1}{4\pi-\|\Delta\psi\|_1} + r_0^2[r_0^{\|\Delta\psi\|_1/2\pi}-1]\right) \times \sup_{x\in(\Omega_{\rho_0}^{\rho_1})_\sigma} \exp\left(\int_{\partial\Omega}\left|\psi(y)\frac{\partial H_N}{\partial\nu}\right| + \left|H_N\frac{\partial\psi}{\partial\nu}(y)\right| dy\right), \tag{4-89}$$

where $C(\sigma,\rho_1)=\frac{1}{2\pi}\max\{|\log(\sigma)|,\,|\log(2\sqrt{2}\rho_1)|\}$.

Proof. If we replace Γ by H_N in (4-77), then the right hand side of (4-79) becomes

$$\sup_{x\in(\Omega_{\rho_0}^{\rho_1})_\sigma} \exp\left[\int_{\partial\Omega_{\rho_0}^{\rho_1}}\left|\psi(y)\frac{\partial H_N}{\partial\nu}\right| + \left|H_N\frac{\partial\psi}{\partial\nu}(y)\right| dy + \int_{\Omega_{\rho_0}^{\rho_1}}|\Gamma(\bar{x},y)\Delta\psi|\right] \times \int_{(\Omega_{\rho_0}^{\rho_1})_\sigma}\exp\left[\int_\Omega|\Gamma(x,y)\Delta\psi(y)|\,dy\right]. \tag{4-90}$$

We see that

$$\sup_{x\in(\Omega_{\rho_0}^{\rho_1})_\sigma}\int_{\Omega_{\rho_0}^{\rho_1}}|\Gamma(\bar{x},y)\Delta\psi| \le C(\sigma,\rho_1)\|\Delta\psi\|_{L^1(\Omega_{\rho_0}^{\rho_1})}. \tag{4-91}$$

The corollary now follows from Lemma 4.6. $\square$

In order to apply Corollary 4.7 to $\alpha-2u$, we need an L^1 bound on $\Delta(\alpha-2u)$ and an uniform bound on the boundary. In Lemma 4.3 we established the necessary L^1

bound. Now, we will demonstrate the needed uniform control on the boundary. The following result is very similar to Lemma 4.4, however, due to technical necessities, the statement and proof are slightly different.

Lemma 4.8. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that M is radially monotone at ρ_0. Let $\Omega_{\rho_0}^{\rho_1}$ denote the region*

$$\left\{(\rho, z) \mid \rho_0 \le \rho \le \rho_1,\ |z| \le \frac{\rho_1}{2}\right\}, \tag{4-92}$$

and $(\Omega_{\rho_0}^{\rho_1})_\sigma$ denote $\{x \in \Omega_{\rho_0}^{\rho_1} \mid d(x, \partial\Omega_{\rho_0}^{\rho_1}) > \sigma\}$. Let $\rho_1 \ge R_0$. If $g \in \mathcal{M}$ and the ADM mass of g is less than m, then

$$\sup_{x\in(\Omega_{\rho_0}^{\rho_1})_\sigma} \exp\left(\int_{\partial\Omega_{\rho_0}^{\rho_1}} \left|H_N(x, y)\frac{\partial(2u-\alpha)}{\partial\nu}(y)\right| + \left|(2u-\alpha)(y)\frac{\partial H_N}{\partial\nu}(x, y)\right| dy\right) \le \exp[C(m, \sigma, \rho_1, \rho_0)] \tag{4-93}$$

where

$$C(m, \sigma, \rho_1, \rho_0) = \max\left\{|\log 2\sqrt{2}\rho_1|, |\log\sigma|\right\}\left(\frac{4m+4\sqrt{\rho_1 m}}{\pi\rho_0} + \frac{9C}{\rho_1}\right) + \frac{3C}{\sigma}. \tag{4-94}$$

Proof. As we observed earlier, for three sides of the rectangle $\Omega_{\rho_0}^{\rho_1}$, the necessary estimates to control the left-hand side of (4-93) follow from the uniformly asymptotically flat condition. Let's make this more precise. First, consider those pieces of the rectangle parallel to the ρ-axis.

From the definition of uniform asymptotic flatness, we know that

$$\left|\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right| = \left|\frac{\partial(\alpha-2u)}{\partial z}(y)\right| \le \frac{3C}{|y|^2} \tag{4-95}$$

Analogously, we have

$$|\alpha - 2u| \le \frac{3C}{|y|}. \tag{4-96}$$

In fact, the same is true on the final edge, so the above estimates are true on all of $\partial\Omega_{\rho_0}^{\rho_1} - \{\rho = \rho_0\}$.

Armed with these estimates, let's take a look at the integral

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}} \left|H_N(x, y)\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right| + \left|(\alpha-2u)(y)\frac{\partial H_N}{\partial\nu}(x, y)\right| dy \tag{4-97}$$

Since the point x is at a distance of at least σ away from the boundary, we know that

$$\frac{\partial H_N}{\partial\nu} \le \frac{1}{\sigma\pi} \tag{4-98}$$

and

$$|H_N(x,y)| \le \frac{1}{\pi}\max\Bigl\{|\log(2\sqrt{2}\rho_1)|, |\log(\sigma)|\Bigr\} \tag{4-99}$$

To start, we can bound

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}} \left|(\alpha-2u)(y)\frac{\partial H_N}{\partial\nu}(x,y)\right| dy \tag{4-100}$$

from above by

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}} \frac{3C}{\sigma\pi|y|}\,dy \le \frac{3C}{\sigma}, \tag{4-101}$$

since $|y|\ge\rho_1$ for y in $\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}$. We now make a similar estimate for

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}} \left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right| dy. \tag{4-102}$$

As we did before, we may bound this quantity from above by

$$\int_{\partial\Omega_{\rho_0}^{\rho_1}-\{\rho=\rho_0\}} \frac{3C}{\pi|y|^2}\max\Bigl\{|\log 2\sqrt{2}\rho_1|, |\log\sigma|\Bigr\}\,dy \le \frac{3C}{\rho_1}\max\Bigl\{|\log 2\sqrt{2}\rho_1|, |\log\sigma|\Bigr\}. \tag{4-103}$$

We need to estimate

$$\int_{(\partial\Omega_{\rho_0}^{\rho_1})\cap\{\rho=\rho_0\}} \left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}\right| \tag{4-104}$$

for $x\in(\Omega_{\rho_0}^{\rho_1})_\sigma$. Using (4-99) and Lemma 4.4 we get

$$\int_{(\partial\Omega_{\rho_0}^{\rho_1})\cap\{\rho=\rho_0\}} \left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}\right| \le \frac{1}{\pi}\max\Bigl\{|\log(2\sqrt{2}\rho_1)|, |\log(\sigma)|\Bigr\}\left(\frac{4m}{\rho_0}+\frac{4\sqrt{\rho_1 m}}{\rho_0}+\frac{6\pi C}{\rho_1}\right). \tag{4-105}$$

Putting the estimates together gives

$$\sup_{x\in(\Omega_{\rho_0}^{\rho_1})_\sigma} \exp\left(\int_{\partial\Omega_{\rho_0}^{\rho_1}} \left|H_N(x,y)\frac{\partial(\alpha-2u)}{\partial\nu}(y)\right| + \left|(\alpha-2u)(y)\frac{\partial H_N}{\partial\nu}(x,y)\right| dy\right) \le C(m,\sigma,\rho_1,\rho_0). \quad\square \tag{4-106}$$

With all of the above estimates in hand, controlling the $W^{1,p}$ norm of $e^{\alpha-2u}$ is relatively straightforward. The technical requirements of Corollary 4.7 force us to consider regions $\Omega_{\rho_0}^{\rho_1}(\sigma)$ for positive σ, see (1-16).

Lemma 4.9. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Let $\Omega_{\rho_0}^{\rho_1}$ denote the region $\{(\rho, z) \mid \rho_0 \le \rho \le \rho_1,\ |z| \le \frac{\rho_1}{2}\}$. Suppose that $\mathcal{M}$ is also radially monotone at ρ_0. For every $\rho_1 > \max\{\rho_0, R_0\}$, $\epsilon > 0$, $\sigma > 0$, and $1 \le p < 2$ there exists a $\delta > 0$ such that if the ADM mass of $g \in \mathcal{M}$ is less than δ, then*

$$\|e^{|\alpha-2u|} - 1\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1}(\sigma))} < \epsilon. \tag{4-107}$$

Proof. By assumption, $\alpha - 2u$ is bounded and has bounded derivatives, although we make no assumption on what these bounds might be. Thus, we have that $e^{|\alpha-2u|}$ is Lipschitz. As in Lemma 3.3, we get

$$\int_{\Omega_{\rho_0}^{\rho_1}(\sigma)} |\nabla e^{|\alpha-2u|} - 1|^p \le \int_{\Omega_{\rho_0}^{\rho_1}(\sigma)} |\nabla(\alpha - 2u)|^p e^{p|\alpha-2u|}. \tag{4-108}$$

Let $r > 1$ be such that $rp < 2$. Applying Hölder's inequality to the above gives

$$\left(\int_{\Omega_{\rho_0}^{\rho_1}(\sigma)} |\nabla(\alpha - 2u)|^{rp}\right)^{1/r} \left(\int_{\Omega_{\rho_0}^{\rho_1}(\sigma)} e^{r'p|\alpha-2u|}\right)^{1/r'}, \tag{4-109}$$

where r' is the conjugate exponent to r. In order to control the left hand side we appeal to Proposition 4.2. In order to bound the right hand side we first note that

$$\Omega_{\rho_0}^{\rho_1}(\sigma) \subset (\Omega_{\rho_0}^{\rho_1+\sigma})_\sigma. \tag{4-110}$$

Thus

$$\int_{\Omega_{\rho_0}^{\rho_1}(\sigma)} e^{r'p|\alpha-2u|} \le \int_{(\Omega_{\rho_0}^{\rho_1+\sigma})_\sigma} e^{r'p|\alpha-2u|}. \tag{4-111}$$

We may apply Corollary 4.7 to the function $r'p(\alpha - 2u)$ and modify Lemma 4.8 as necessary in order to see that

$$\int_{(\Omega_{\rho_0}^{\rho_1+\sigma})_\sigma} e^{r'p|\alpha-2u|} \tag{4-112}$$

is uniformly bounded for all m small enough. Thus, combining the two estimates above shows that

$$\|\nabla e^{|\alpha-2u|}\|_{L^p(\Omega_{\rho_0}^{\rho_1}(\sigma))} < \frac{\epsilon}{2} \tag{4-113}$$

for sufficiently small m. Similarly, for m small enough, we can show that

$$\|e^{|\alpha-2u|}\|_{L^p(\Omega_{\rho_0}^{\rho_1}(\sigma))} < \frac{\epsilon}{2}. \tag{4-114}$$

□

5. Proofs of the theorems

In this section we will apply the lemmas to prove the theorems stated in the introduction. Most of the above lemmas analyzed functions over the rectangles $\Omega_{\rho_0}^{\rho_1}$. Now we move our focus to the cylindrical annuli

$$\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma) = \Omega_{\rho_0}^{\rho_1}(\sigma) \times [0, 2\pi), \tag{5-1}$$

see (1-15). Except for the final theorem, this change of focus doesn't involve any new difficulties.

Proof of Theorem 1.4: We first restate the theorem.

Theorem 1.4. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose that $\mathcal{M}$ is radially monotone at ρ_0 and that for each metric in $\mathcal{M}$, we have*

$$A = B = 0. \tag{1-19}$$

For every $\rho_1 > \max\{\rho_0, R_0\}$, $\epsilon > 0$, $\sigma > 0$, and $1 \le p < 2$ there exists a $\delta > 0$ such that if the ADM mass of $g \in \mathcal{M}$ is less than δ, then

$$\|g - \delta_{\mathbb{R}^3}\|_{W^{1,p}(\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma))} < \epsilon, \tag{1-20}$$

and

$$\|q - \delta_{\mathbb{R}^2}\|_{W^{1,p}(\Omega_{\rho_0}^{\rho_1}(\sigma))} < \epsilon, \tag{1-21}$$

where $\delta_{\mathbb{R}^3}$ denotes the Euclidean metric in cylindrical coordinates, $\delta_{\mathbb{R}^2}$ denotes the Euclidean metric in the (ρ, z) plane, and q denotes the orbit metric of g in the (ρ, z) plane. $\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)$ denotes the cylinder given in (1-15) and $\Omega_{\rho_0}^{\rho_1}(\sigma)$ denotes its orbit space.

Proof. Since we have assumed that $A = B = 0$, in order to show that g is $W^{1,p}$ close to $\delta_{\mathbb{R}^3}$ for small ADM mass, we need only show that

$$\|\rho^2 e^{-2u} - \rho^2\|_{W^{1,p}(\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma))} < \epsilon \tag{5-2}$$

and

$$\|e^{2\alpha-2u} - 1\|_{W^{1,p}(\widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma))} < \epsilon \tag{5-3}$$

if the ADM mass is sufficiently small. For (5-2) this follows quickly from Lemma 3.3. Demonstrating (5-3) is only a little more difficult.

As before, we see that

$$\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |e^{2(\alpha-u)} - 1|^p \le \int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2\alpha - 2u|^p e^{2p(\alpha-u)}. \tag{5-4}$$

After applying Hölder's inequality to the above with some $r > 1$ such that $rp < 2$ we obtain

$$\left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2(\alpha - u)|^{rp}\right)^{1/r} \left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{2pr'(\alpha-u)}\right)^{1/r'}. \tag{5-5}$$

In order to estimate the above, we first observe that

$$2(\alpha - u) = 2u + 2(\alpha - 2u). \tag{5-6}$$

We can now estimate the left hand term using the triangle inequality, Corollary 3.2, and Proposition 4.2 for the exponent $rp < 2$. For the right hand side we have

$$\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{2pr'(\alpha-u)} = \int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{2pr'u} e^{2pr'(\alpha-2u)}. \tag{5-7}$$

After applying Hölder's inequality, we may use Lemma 4.9 and Lemma 3.3 applied to $2pr'u$ and $2pr'(\alpha - 2u)$, respectively, to bound the L^p norm of $e^{2\alpha-2u}$. In fact, in the same way, for any fixed q we can bound the L^q norm of $e^{2\alpha-2u}$ for all m small enough, depending on ρ_1, ρ_0, and q. For what follows, we pick q large enough, depending on p. If we take the gradient of $e^{2\alpha-2u}$ we get

$$(e^{2\alpha-2u})\nabla(2\alpha - 2u) = e^{2\alpha-2u}(\nabla 2u + 2\nabla(\alpha - 2u)). \tag{5-8}$$

We again use Hölder's inequality, Lemma 3.3, Proposition 4.2 and Lemma 4.9 to control the L^p norm of $\nabla e^{2\alpha-2u}$. □

Proof of Theorem 1.6: Let us first restate the theorem:

Theorem 1.6. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is radially monotone at ρ_0. For any constants $\epsilon > 0$, $\sigma > 0$, and $\rho_1 > \max\{\rho_0, R_0\}$, there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{1-24}$$

then

$$|\Omega| + \epsilon \geq \mathrm{vol}_g(\Omega) \geq |\Omega| - \epsilon \tag{1-25}$$

for any region Ω such that

$$\Omega \subset \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma). \tag{1-26}$$

Proof. A quick calculation shows that the volume form of g in cylindrical coordinates is

$$\rho e^{2\alpha-3u}\, d\rho\, dz\, d\phi. \tag{5-9}$$

Thus, we have that

$$|\mathrm{vol}_g(\Omega) - |\Omega|| = \left| \int_\Omega (e^{2\alpha - 3u} - 1)\rho \, d\rho \, dz \, d\phi \right| \leq \int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |e^{2\alpha - 3u} - 1| \rho \, d\rho \, dz \, d\phi. \tag{5-10}$$

As we have done before, we can see that

$$\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |e^{2\alpha - 3u} - 1| \rho \, d\rho \, dz \, d\phi \leq \int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2\alpha - 3u| e^{|2\alpha - 3u|} \rho \, d\rho \, dz \, d\phi. \tag{5-11}$$

We may now apply Hölder's inequality to the above in order to see that

$$\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2\alpha - 3u| e^{|2\alpha - 3u|} \leq \left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2\alpha - 3u|^p \right)^{1/p} \left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{p'|2\alpha - 3u|} \right)^{1/p'}, \tag{5-12}$$

where p and p' are conjugate exponents and $1 \leq p < 2$. We may use the triangle inequality to make the estimate

$$\left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} |2\alpha - 3u|^p \right)^{1/p} \leq \|u\|_{W^{1,p}} + 2\|\alpha - 2u\|_{W^{1,p}}. \tag{5-13}$$

We may combine Corollary 3.2 and Proposition 4.2 to control the above. For the exponential term, we use the estimate

$$e^{p'|2\alpha - 3u|} \leq e^{p'|u|} e^{2p'|\alpha - 2u|} \tag{5-14}$$

and Hölder's inequality once more to see that

$$\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{p'|2\alpha - 3u|} \leq (e^{2p'|u|})^{1/2} \left(\int_{\widetilde{\Omega}_{\rho_0}^{\rho_1}} e^{4p'|\alpha - 2u|} \right)^{1/2}. \tag{5-15}$$

We now wish to apply Lemma 3.3 and 4.9 to the above to see that it is uniformly bounded for m small enough, depending on ρ_1, ρ_0 and p. Combining the two estimates finishes the proof. □

Proof of Theorem 1.7: Let us first restate the theorem.

Theorem 1.7. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is radially monotone at ρ_0. For any fixed axisymmetric surface Σ, constant $\epsilon > 0$, and constant $\rho_1 > \max\{\rho_0, R_0\}$, there exists a $\delta > 0$ such that if $m(g) < \delta$, then*

$$|\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)| + \epsilon \geq \mathrm{Area}_g(\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)) \geq |\Sigma \cap \widetilde{\Omega}_{\rho_0}^{\rho_1}(\sigma)| - \epsilon. \tag{1-27}$$

Proof. Let s be a fixed curve in the (ρ, z) plane representing an axisymmetric surface, which we will call Σ. A calculation shows that the area form associated with Σ is

$$\rho \circ s(t) e^{(\alpha-2u)\circ s} |\dot{s}|_\delta \, dt \, d\phi. \tag{5-16}$$

Note that the Euclidean area form for Σ is

$$\rho \circ s(t) |\dot{s}|_\delta \, dt \, d\phi. \tag{5-17}$$

From Lemma 4.9 we deduce that for any $\epsilon > 0$

$$\|\rho e^{\alpha-2u} - \rho\|_{W^{1,1}(\Omega^{\rho_1}_{\rho_0}(\sigma))} < \epsilon, \tag{5-18}$$

if the ADM mass is small enough. Now, the curve segment $s \cap \Omega^{\rho_1}_{\rho_0}(\sigma)$ is part of the boundary of some region in $\Omega^{\rho_1}_{\rho_0}(\sigma)$. Thus, we may use the trace inequality [Evans and Gariepy 2015] to conclude that

$$\|\rho e^{\alpha-2u} - \rho\|_{L^1(s\cap\Omega^{\rho_1}_{\rho_0})} < \epsilon. \tag{5-19}$$

This proves the theorem. □

If the family $\mathcal{M}$ is area enlarging everywhere, then we also have a stronger lower bound on the area of axisymmetric surfaces than the one given above.

Proposition 5.1. *Let g be an axisymmetric metric. Let (ρ, z, ϕ) be the cylindrical coordinates for g, let $\delta_{\mathbb{R}^3}$ be the flat metric in cylindrical coordinates, and let Σ be a C^1 axisymmetric surface. If g is area enlarging, then we have*

$$\mathrm{Area}_g(\Sigma) \geq \mathrm{Area}_{\delta_{\mathbb{R}^3}}(\Sigma) \tag{5-20}$$

Proof. Let Σ be a C^1 axisymmetric surface. Let $s(t)$ be the C^1 curve in the (ρ, z) plane which, when revolved around the ρ-axis, gives Σ. We get the following map

$$(t, \phi) \to (s(t), \phi) \tag{5-21}$$

from $I \times [0, 2\pi)$ to Σ. Let A_g denote the area form of the surface with respect to the metric induced by g, and let $A_{\delta_{\mathbb{R}^3}}$ denote the area form induced by the background Euclidean metric. Then using (5-16) and (5-17) we see that

$$A_g = e^{\alpha-2u} A_{\delta^3_{\mathbb{R}}}. \tag{5-22}$$

In coordinates, the area enlarging condition is equivalent to the nonnegativity of $\alpha - 2u$. Thus, we know that $e^{\alpha-2u}$ is greater than 1. The result now follows. □

We may combine the well known Penrose Inequality with the above proposition to constrain the location of outer most minimal surfaces.

Corollary 5.2. *Let $\mathcal{M}$ be a family of uniformly asymptotically flat metrics with nonnegative scalar curvature. Suppose $\mathcal{M}$ is either radially monotone or area enlarging. Let g be a metric in $\mathcal{M}$ and Σ be the outermost minimal surface. If Σ is axisymmetric and topologically a sphere, and*

$$m(g) \le m, \tag{5-23}$$

then

$$\Sigma \subset \rho^{-1}\big([0, 2\sqrt{2}m)\big). \tag{5-24}$$

Proof. Let

$$\rho_0 = \max\{\rho : (\rho, z) \in \Sigma\}, \tag{5-25}$$

let x_0 be a point in Σ point at which ρ attains the maximum ρ_0, and let $[x_0]$ denote its orbit under the killing field. From the Penrose inequality, we know that

$$m \ge \sqrt{\frac{\text{Area}_g(\Sigma)}{16\pi}}. \tag{5-26}$$

Since Σ is axisymmetric and topologically a sphere, it must be represented in the (ρ, z) plane by a curve γ which intersects the axis of symmetry twice. In particular, γ must emanate from the axis, then touch the point $[x_0]$ and then make its way back to the axis. Let D_{x_0} denote the disk represented by a line connecting the axis to the point $[x_0]$. Since this disk has minimal Euclidean area among axisymmetric surfaces with boundary $[x_0]$, we may conclude that

$$\text{Area}_{\delta_{\mathbb{R}^3}}(\Sigma) > 2\,\text{Area}_{\delta_{\mathbb{R}^3}}(D_{x_0}) = 2\pi\rho_0^2. \tag{5-27}$$

Thus, combining the Penrose inequality with the above and the area enlarging inequality (5-20) gives

$$m > \frac{\rho_0}{2\sqrt{2}}. \tag{5-28}$$ □

If the metric g in the above has positive scalar curvature, then it is a well known result that the outermost minimal surface must be a sphere. The author does not know if in an axisymmetric metric an outermost minimal surface must also be axisymmetric, though it does seem plausible.

Proof of Theorems 1.8 and 1.9.

Theorem 1.8. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose $\mathcal{M}$ is also radially monotone at ρ_0. Additionally, assume that $A = B = 0$ in the coordinate representations of the metrics under consideration. Suppose we are given $\epsilon > 0$, $\sigma > 0$, and $\rho_1 > \max\{\rho_0, R_0\}$. There exists a constant*

$\delta > 0$ *such that if* $m(g) \leq \delta$ *and* x *and* y *are any points such that the Euclidean line segment connecting them lies in* $\Omega^{\rho_1}_{\rho_0}(\sigma) \times \{\phi_0\}$ *for any* ϕ_0, *then*

$$d_g(x, y) \leq d(x, y) + \epsilon. \tag{1-28}$$

Proof. We use the extension theorem for Sobolev functions, appearing as Theorem 4.7 in [Evans and Gariepy 2015]. Following the notation of [Evans and Gariepy 2015], if we let $U = \Omega^{\rho_1}_{\rho_0}(\Sigma)$, $V = 2\Omega^{\rho_1}_{\rho_0}(\Sigma)$, and $p = 1$, then we may see that there is a constant K, depending on $\Omega^{\rho_1}_{\rho_0}(\sigma)$, and extensions of the functions $e^{\alpha-u} - 1$, also denoted $e^{\alpha-u} - 1$, such that

$$\|e^{\alpha-u} - 1\|_{W^{1,1}(\mathbb{R}^2)} \leq K\|e^{\alpha-u} - 1\|_{W^{1,1}(\Omega^{\rho_1}_{\rho_0}(\sigma))}. \tag{5-29}$$

In order to obtain an upper estimate for $d_g(x, y)$, it suffices to estimate the length of one curve connecting the points x and y. Let γ_{xy} denote the Euclidean line in $\Omega^{\rho_1}_{\rho_0}(\sigma) \times \{\phi_0\}$ connecting x to y parametrized by Euclidean arc length In orbit space

$$\gamma_{xy}(t) = (\gamma^{\rho}_{xy}(t), \gamma^{z}_{xy}(t)). \tag{5-30}$$

Every such curve lies on the boundary of a square of side length the diameter of $\Omega^{\rho_1}_{\rho_0}(\sigma)$. All such squares are rotations or translations of each other. Thus, there exists a single constant C such that if Ω is a square with side length the diameter of $\Omega^{\rho_1}_{\rho_0}(\sigma)$, then the trace inequality holds with constant C:

$$\|\omega\|_{L^1(\partial\Omega)} \leq C\|\omega\|_{W^{1,1}(\Omega)}. \tag{5-31}$$

Let $l_g(\gamma)$ be the length of γ as measured in the metric g. Then we have

$$l_g(\gamma) = \int_0^{d(x,y)} e^{(\alpha-u)\circ\gamma(t)}\, dt. \tag{5-32}$$

We now use the trace inequality [Evans and Gariepy 2015] to see that

$$\begin{aligned} |d(x, y) - l_g(\gamma)| &\leq \int_0^{d(x,y)} |e^{(\alpha-u)\circ\gamma(t)} - 1|\, dt \\ &\leq \int_{\partial\Omega} |e^{\alpha-u} - 1| \leq C\|e^{\alpha-u} - 1\|_{W^{1,1}(\Omega)}, \end{aligned} \tag{5-33}$$

where γ lies on the boundary of Ω. Furthermore, we have

$$\|e^{\alpha-u} - 1\|_{W^{1,1}(\Omega)} \leq \|e^{\alpha-u} - 1\|_{W^{1,1}(\mathbb{R}^2)} \leq K\|e^{\alpha-u} - 1\|_{W^{1,1}(\Omega^{\rho_1}_{\rho_0}(\sigma))}. \tag{5-34}$$

We may now use Theorem 1.4 to conclude that

$$|d(x, y) - l_g(\gamma)| < \epsilon \tag{5-35}$$

for small enough ADM mass. □

Very similarly, we can prove a pointwise upper bound on $d_g(x, y)$ for more general x and y in $\widetilde{\Omega}^{\rho_1}_{\rho_0}$.

Theorem 1.9. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is radially monotone at ρ_0. Additionally, assume that $A = B = 0$ in the coordinate representations of the metrics under consideration. Suppose we are given $\epsilon > 0$ and $\sigma > 0$ and points x and y such that the Euclidean line segment connecting them lies in $\widetilde{\Omega}^{\rho_1}_{\rho_0}(\sigma)$. There exists a constant $\delta > 0$ such that if $m(g) \leq \delta$, then*

(1-29) $$d_g(x, y) \leq d(x, y) + \epsilon.$$

Proof. As before, let γ be the Euclidean line connecting x to y. Then we have that

(5-36) $$|l_g(\gamma_{xy}) - 1| \leq \int_0^{d(x,y)} \left|\sqrt{e^{2(\alpha-u)\circ\gamma}((\gamma'_\rho)^2 + (\gamma'_z)^2) + \gamma_\rho^2 e^{-2u\circ\gamma}(\gamma'_\phi)^2} - 1\right| dt.$$

Let

(5-37) $$Z = e^{\alpha-u}\left(\gamma'_\rho \frac{\partial}{\partial\rho} + \gamma'_z \frac{\partial}{\partial z}\right) + e^{-u}\gamma'_\phi \frac{\partial}{\partial\phi}.$$

Using the reverse triangle inequality, we observe that

(5-38) $$||Z| - 1| = ||Z| - |\gamma'|| \leq |Z - \gamma'|,$$

where we are working with the Euclidean metric in cylindrical coordinates. Thus, we may estimate the above integral by

(5-39) $$\int_0^{d(x,y)} \sqrt{(e^{(\alpha-u)\circ\gamma} - 1)^2((\gamma'_\rho)^2 + (\gamma'_z)^2) + (e^{-u\circ\gamma} - 1)^2\gamma_\rho^2(\gamma'_\phi)^2}\, dt.$$

Using the triangle inequality and the bounds

(5-40) $$(\tilde{\gamma}'_\rho)^2 + (\tilde{\gamma}'_z)^2 \leq 1,$$

and

(5-41) $$|\gamma_\rho \gamma'_\phi| \leq 1,$$

we see that the above is bounded in turn by

(5-42) $$\int_0^{d(x,y)} |e^{(\alpha-u)\circ\gamma} - 1|\, dt + \int_0^{d(x,y)} |e^{-u\circ\gamma} - 1|\, dt.$$

Let $\tilde{\gamma}$ be the projection of γ to the (ρ, z) plane. $\tilde{\gamma}$ lies in the boundary of a region Ω. Since u and α don't depend on ϕ, we see that $u \circ \gamma = u \circ \tilde{\gamma}$ and $\alpha \circ \gamma = \alpha \circ \tilde{\gamma}$. We can now use the trace theorem, and then apply Theorem 1.4 as we did before to show that for ADM mass small enough, we have

$$\int_0^{d(x,y)} |e^{(\alpha-u)\circ\tilde{\gamma}} - 1|\, dt + \int_0^{d(x,y)} |e^{-u\circ\tilde{\gamma}} - 1|\, dt < \epsilon. \tag{5-43}$$ □

Proof of Theorem 1.10. We restate the theorem.

Theorem 1.10. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose that $\mathcal{M}$ is radially monotone and that for all $g \in \mathcal{M}$ we have*

$$A = B = 0. \tag{1-30}$$

Let $R_1 > R_0$ and let $A(R_0, R_1)$ denote the coordinate spherical annulus centered at the origin. For any given $0 < \beta < 1$ and $\epsilon > 0$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and

$$m(g) < \delta, \tag{1-31}$$

then

$$\|g - \delta_{\mathbb{R}^3}\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon. \tag{1-32}$$

Proof. Since we have assumed that $A = B = 0$, the proof will be established if we can show that

$$\|e^{2\alpha-2u} - 1\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon \tag{5-44}$$

and

$$\|e^{-2u} - 1\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon \tag{5-45}$$

for small enough ADM mass. The above inequalities will follow if we can show that

$$\|\alpha - u\|_{C^{0,\beta}(A(R_0,R_1))} < \tilde{\epsilon} \tag{5-46}$$

and

$$\|u\|_{C^{0,\beta}(A(R_),R_1)} < \tilde{\epsilon} \tag{5-47}$$

for small enough ADM mass, where $\tilde{\epsilon}$ depends on ϵ above. Using the triangle inequality, we see that it is sufficient to bound the $C^{0,\beta}$ norms of u and $\alpha - 2u$. These bounds are the content of Lemma 5.3 and Lemma 5.7 below, respectively. □

Lemma 5.3. *Suppose $\mathcal{M}$ is a collection of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside a ball of radius R_0. Let u be the function appearing in the axisymmetric coordinate* *representation of g. Let R_1 be greater than R_0 and $A(R_0, R_1)$ be the spherical annulus centered at the origin. For $\epsilon > 0$ and $0 < \beta_0 < 1$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{5-48}$$

then

$$\|u\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon. \tag{5-49}$$

Proof. Since we are working in the asymptotically flat regime, we have uniform upper bounds on the $C^1(A(R_0, R_1))$ norms of the metric functions. From Lemma 3.1 we may bound the $W^{1,2}(A(R_0, R_1))$ norm of u. We now interpolate between these two estimates to bound the $W^{1,q}$ norm of u for arbitrarily large q. Specifically, we write

$$\int_{A(R_0,R_1)} u^q = \int_{A(R_0,R_1)} u^2 u^{q-2} \le \|u\|_\infty^{q-2} \int_{A(R_0,R_1)} u^2 \tag{5-50}$$

We may do the same for the derivatives of u. In the end, we get the following bounds:

$$\|u\|_q \le \|u\|_2^{2/q} \|u\|_\infty^{1-2/q} \tag{5-51}$$

and

$$\|\nabla u\|_q \le \|\nabla u\|_2^{2/q} \|\nabla u\|_\infty^{1-2/q}. \tag{5-52}$$

By assumption $\|u\|_\infty + \|\nabla u\|_\infty \le C$. Furthermore, by Lemma 3.1, we know $\|u\|_{W^{1,2}(A(R_0,R_1))} < \tilde{\epsilon}$ for sufficiently small m. Thus, we obtain the estimate

$$\|u\|_{W^{1,q}} \le C^{1-2/q} \tilde{\epsilon}^{2/q}. \tag{5-53}$$

We may now choose q large enough and appeal to the Sobolev embedding theorem to get C^{0,β_0} bounds on u for $\beta_0 < 1$. □

Remark 5.4. It is important to note that we didn't use the hypothesis of radial monotonicity in the above. We only need radial monotonicity to control $\alpha - 2u$.

We will try to produce similar uniform estimates for $\alpha - 2u$. However, as before, the process is harder. Whereas for u we started off with $W^{1,p}_{\text{loc}}(\mathbb{R}^3)$ control, for $\alpha - 2u$ we only have $W^{1,p}_{\text{loc}}(\mathbb{R}^2_+)$ control. Even worse, the estimates we were able to prove become weaker as we approach the axis $\{\rho = 0\}$, see Corollary 4.5. In order to work our way around this conundrum, we must use the extra factor of ρ present in integrating over B_R in $\mathbb{R}^3$ to control the bad behavior seen in Corollary 4.5.

Lemma 5.5. *Let f be a measurable function on $\Omega_0^{\rho_1}$. Suppose for each t we have the estimate*

$$\int_{\Omega_t^{\rho_1}} |f| \leq \frac{\epsilon}{t^q} \tag{5-54}$$

for some $\epsilon > 0$ and $q > 0$. Suppose $\sigma > q$. Then, there exists a constant, denoted $C(\sigma, q)$, depending only on σ and q such that

$$\int_{\Omega_0^{\rho_1}} \rho^\sigma |f| \leq C(\sigma, q)\epsilon. \tag{5-55}$$

Proof. Let $t_n = 2^{-n}\rho_1$ and let $\Omega_{t_n,t_{n-1}}$ be the following rectangle:

$$\Omega_{t_n,t_{n-1}} = \left\{ t_n < \rho \leq t_{n-1},\ |z| \leq \frac{\rho_1}{2} \right\}. \tag{5-56}$$

From the monotone convergence theorem we see that

$$\int_{\Omega_{\rho_0}^{\rho_1}} \rho^\sigma |f| = \int_{\Omega_{0,t_0}} \rho^\sigma |f| = \sum_1^\infty \int_{\Omega_{t_n,t_{n-1}}} \rho^\sigma |f|. \tag{5-57}$$

We now make the estimate

$$\int_{\Omega_{t_n,t_{n-1}}} \rho^\sigma |f| \leq t_{n-1}^\sigma \frac{\epsilon}{t_n^q} = 2^\sigma \rho_1^{\sigma-q} (2^{\sigma-q})^{-n} \epsilon. \tag{5-58}$$

This gives a convergent series so long as $\sigma > q$. In total, we have the estimate

$$\int_{\Omega_{0,t_0}} \rho^\sigma |f|^p \leq C(\sigma, q)\epsilon. \tag{5-59}$$

□

We now make use of the above lemma to control the $W^{1,1}$ norm of $\alpha - 2u$ over the ball of radius R about the origin in $\mathbb{R}^3$.

Lemma 5.6. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside of radius R_0. Suppose that $\mathcal{M}$ is also a radially monotone family of metrics. For any R and $\epsilon > 0$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{5-60}$$

then

$$\|\alpha - 2u\|_{W^{1,1}(B_R)} < \epsilon. \tag{5-61}$$

Proof. Let D_R be the two dimensional half disk of radius R about the origin. Then

$$\int_{B_R} |\alpha - 2u| = 2\pi \int_{D_R} \rho|\alpha - 2u|. \tag{5-62}$$

For some $\mu > 0$, to be specified later, we rewrite the above quantity as

$$\int_{D_R} \rho^{-\mu} \rho^{1+\mu} |\alpha - 2u|. \tag{5-63}$$

Let $1 < q < 2$ and q' be conjugate exponents. We apply Hölder's inequality to the above to get

$$\left(\int_{D_R} \rho^{-\mu q'} \right)^{1/q'} \left(\int_{D_R} \rho^{(1+\mu)q} |\alpha - 2u|^q \right)^{1/q}. \tag{5-64}$$

Choose μ small enough that

$$\mu q' < 1. \tag{5-65}$$

We may pick large ρ_1 enough that $D_R \subset \Omega_0^{\rho_1}$. From Corollary 4.5 and Lemma 5.5, we see that for some constant $C(\mu, q, R)$,

$$\int_{D_R} \rho |\alpha - 2u| \leq C(\mu, q, R)\epsilon \tag{5-66}$$

if m is chosen small enough. The same argument can be made for

$$\int_{D_R} \rho |\nabla(\alpha - 2u)|. \tag{5-67}$$

□

We now make an estimate on the uniform norm of $\alpha - 2u$ similar to Lemma 5.3.

Lemma 5.7. *Suppose $\mathcal{M}$ is a collection of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is uniformly asymptotically flat outside a ball of radius R_0. Let R_1 be greater than R_0 and $A(R_0, R_1)$ be the spherical annulus centered at the origin. For $\epsilon > 0$ and $0 < \beta < 1$ there exists a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{5-68}$$

then

$$\|\alpha - 2u\|_{C^{0,\beta}(A(R_0,R_1))} < \epsilon. \tag{5-69}$$

Proof. We imitate the proof of Lemma 5.3. As before, we write

$$\int_{A(R_0,R_1)} |\alpha - 2u|^q \leq \|\alpha - 2u\|_\infty^{q-1} \int_{A(R_0,R_1)} |\alpha - 2u|. \tag{5-70}$$

We also have

$$\int_{A(R_0,R_1)} |\nabla(\alpha - 2u)|^q \leq \|\nabla(\alpha - 2u)\|_\infty^{q-1} \int_{A(R_0,R_1)} |\nabla(\alpha - 2u)|. \tag{5-71}$$

By the asymptotic flatness assumption, we know that

$$\|\alpha - 2u\|_\infty + \|(\alpha - 2u)\|_\infty \le C \tag{5-72}$$

For some C depending only on the uniform falloff in Definition 1.1. Thus, for any exponent q we can use Lemma 5.6 to control the Sobolev norm $\|\alpha - 2u\|_{W^{1,q}(A(R_0,R_1))}$ by the ADM mass. Using the Sobolev embedding theorem, we see that

$$\|\alpha - 2u\|_{C^{0,\beta}} \le C\|\alpha - 2u\|^{1/q}_{W^{1,1}(A(R_0,R_1))}, \tag{5-73}$$

where $\beta = 1 - \frac{3}{q}$, the constant C depends only on the uniform falloff in Definition 1.1, the region $A(R_0, R_1)$, and q. Now we can use Lemma 5.6 to control the uniform norm $\alpha - 2u$ on $A(R_0, R_1)$. □

6. Area enlarging case

We now show that all the theorems stated hold when we assume our family of uniformly asymptotically flat metrics is area enlarging and strongly uniformly asymptotically flat, instead of radially monotone. The only steps required are to prove a lemma analogous to Lemma 4.4 and a proposition analogous to Proposition 4.2. The main difference between the radially monotone case and the area enlarging one is in the choice of function for Green's representation formula. Instead of working with $H_N(x, y)$, we will use $H_D(x, y)$ (4-2). We also focus on slightly different rectangles,

$$\Omega^L_{\rho_0\rho_1} := \left\{(\rho, z) : \rho_0 \le \rho \le \rho_1,\ |z| \le \frac{L}{2}\right\}. \tag{6-1}$$

We now prove the first key lemma for the area enlarging and strongly uniformly asymptotically flat case.

Lemma 6.1. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is strongly uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is area enlarging at ρ_0. For any $\rho_1 > \rho_0$, $L > 0$, and $\epsilon > 0$ there exists a $\delta > 0$ such that if*

$$m(g) < \delta, \tag{6-2}$$

then

$$\int_{\partial\Omega^L_{\rho_0\rho_1}\cap\{\rho=\rho_0\}} |\alpha - 2u| < \epsilon. \tag{6-3}$$

Proof. Observe that if $\tilde{L} > L$, then

$$\int_{\partial\Omega^{\tilde{L}}_{\rho_0\rho_1}\cap\{\rho=\rho_0\}} |\alpha - 2u| \ge \int_{\partial\Omega^L_{\rho_0\rho_1}\cap\{\rho=\rho_0\}} |\alpha - 2u|. \tag{6-4}$$

In order to take advantage of asymptotically flat conditions given in Definition 1.1 it we will often consider $\tilde{L}$ sufficiently larger than $\max\{L, R_0\}$. We will then use the above inequality to relate any estimates we obtain back to our original situation. Similarly, we will look at $\tilde{\rho}_1 > \max\{\rho_1, R_0\}$.

If we write the area enlarging condition (1-11) in terms of the coordinate functions, then we see that

$$(\alpha - 2u)(\rho_0, z) \geq 0. \tag{6-5}$$

From this, it quickly follows that

$$\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}\cap\{\rho=\rho_0\}} |\alpha - 2u| = \int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}\cap\{\rho=\rho_0\}} \alpha - 2u. \tag{6-6}$$

In order to estimate the above, we once again take advantage of the fundamental theorem of calculus to write

$$\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}\cap\{\rho=\rho_0\}} (\alpha - 2u)\, dz = \int_{-\tilde{L}/2}^{\tilde{L}/2} \int_{\rho_0}^{\tilde{\rho}_1} -\frac{\partial(\alpha - 2u)}{\partial\rho}\, d\rho\, dz + \int_{-\tilde{L}/2}^{\tilde{L}/2} (\alpha - 2u)(\tilde{\rho}_1, z)\, dz. \tag{6-7}$$

We may switch the order of integration for the integral on the right to get

$$\int_{\rho_0}^{\tilde{\rho}_1} \int_{-\tilde{L}/2}^{\tilde{L}/2} -\frac{\partial(\alpha - 2u)}{\partial\rho}\, dz\, d\rho. \tag{6-8}$$

As before (4-24), from Stokes' theorem we get

$$\int_{-\tilde{L}/2}^{\tilde{L}/2} -\frac{\partial(\alpha - 2u)}{\partial\rho}(\rho, z)\, dz = \int_{\{\rho\leq s,\, |z|\leq\tilde{L}/2\}} \Delta(\alpha - 2u)(s, z) - \int_{\{\rho\leq s,\, |z|=\tilde{L}/2\}} \frac{\partial(\alpha - 2u)}{\partial\nu}. \tag{6-9}$$

Taking the absolute value of the above and plugging it into (6-7) gives us the estimate

$$\int_{-\tilde{L}/2}^{\tilde{L}/2} |\alpha - 2u| \leq \int_{\rho_0}^{\tilde{\rho}_1} \left(\int_{\{\rho\leq s,\, |z|=\tilde{L}/2\}} |\Delta(\alpha - 2u)| + \int_{\{\rho\leq s,\, |z|=\tilde{L}/2\}} \left| \frac{\partial(\alpha - 2u)}{\partial z} \right| ds \right) d\rho + \int_{-\tilde{L}/2}^{\tilde{L}/2} |\alpha - 2u|(\tilde{\rho}_1, z)\, dz. \tag{6-10}$$

We now proceed to estimate the right hand side term by term.

We start with the term

$$\int_{\rho_0}^{\tilde{\rho}_1}\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}\left|\frac{\partial(\alpha-2u)}{\partial z}\right|\,ds\,d\rho. \tag{6-11}$$

Using the asymptotic flatness condition, we estimate

$$\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}\left|\frac{\partial(\alpha-2u)}{\partial z}\right|\,ds \le \int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}\frac{3C}{|(s,z)|^2}\,ds. \tag{6-12}$$

Once more, a simple integration bounds the above by

$$\frac{6\pi C}{\tilde{L}}. \tag{6-13}$$

Thus, we see that

$$\int_{\rho_0}^{\tilde{\rho}_1}\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}\left|\frac{\partial(\alpha-2u)}{\partial z}\right|\,ds\,d\rho \le \frac{6\pi C\tilde{\rho}_1}{\tilde{L}}. \tag{6-14}$$

We may bound

$$\int_{\rho_0}^{\tilde{\rho}_1}\left(\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}|\Delta(\alpha-2u)|\right)d\rho \tag{6-15}$$

by modifying Lemma 4.3 slightly to get

$$\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}|\Delta(\alpha-2u)| \le \frac{4m+4\sqrt{\tilde{L}m}}{\rho} \tag{6-16}$$

and then integrating. We see that

$$\int_{\rho_0}^{\tilde{\rho}_1}\left(\int_{\{\rho\le s,\,|z|=\tilde{L}/2\}}|\Delta(\alpha-2u)|\right)d\rho \le (4m+4\sqrt{\tilde{L}m})\log\left(\frac{\tilde{\rho}_1}{\rho_0}\right). \tag{6-17}$$

Finally, we must bound

$$\int_{-\tilde{L}/2}^{\tilde{L}/2}|\alpha-2u|(\tilde{\rho}_1,z)\,dz. \tag{6-18}$$

Oddly enough, this turns out to be the most delicate estimate, and the point where we need our extra assumption on the asymptotic falloff of the function α. From Lemma 5.3, we know that the $C^{0,\beta}$ norm of u is controlled by m. Recalling (5-49), we see that there is a constant $\tilde{\epsilon}(\tilde{\rho}_1,m)$ such that

$$\int_{-\tilde{L}/2}^{\tilde{L}/2}|u(\tilde{\rho}_1,z)|\,dz \le \tilde{L}\tilde{\epsilon}(m,\tilde{\rho}_1). \tag{6-19}$$

Again, looking at Lemma 5.3, we see that for fixed $\tilde{\rho}_1$

$$\lim_{m\to 0} \tilde{\epsilon}(\tilde{\rho}_1, m) = 0. \tag{6-20}$$

From the extra assumption on the asymptotic falloff of α, we see that

$$\int_{-\tilde{L}/2}^{\tilde{L}/2} |\alpha(\tilde{\rho}_1, z)|\, dz \le \int_{-\tilde{L}/2}^{\tilde{L}/2} \frac{C}{|(\tilde{\rho}_1, z)|^{1+\tau}}\, dz \le C(\tau)(\tilde{\rho}_1)^{-\tau}, \tag{6-21}$$

where $C(\tau)$ is a constant depending only on τ. We may put all of this together to see that

$$\begin{aligned}\int_{-\tilde{L}/2}^{\tilde{L}/2} |\alpha - 2u|\, dz \\ \le (4m + 4\sqrt{\tilde{L}m}) \log\left(\frac{\tilde{\rho}_1}{\rho_0}\right) + \frac{6\pi C \tilde{\rho}_1}{\tilde{L}} + \tilde{L}\tilde{\epsilon}(\tilde{\rho}_1, m) + C(\tau)(\tilde{\rho}_1)^{-\tau}.\end{aligned} \tag{6-22}$$

By choosing $\tilde{\rho}_1$ and $\tilde{L}$ to be as large as necessary and choosing m to be as small as necessary, we see that the above quantity can be made as small as we desire. □

The following corollary to Lemma 6.1 is analogous to Lemma 4.8.

Corollary 6.2. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is strongly uniformly asymptotically flat outside of radius R_0. Suppose also that M is area enlarging at ρ_0. Let*

$$\Omega^L_{\rho_0\rho_1} := \left\{(\rho, z) \,\middle|\, \rho_0 \le \rho \le \rho_1,\, |z| \le \frac{L}{2}\right\}, \quad \textit{and} \quad (\Omega^L_{\rho_0\rho_1})_\sigma := \left\{x \in \Omega^L_{\rho_0\rho_1} \,|\, d(x, \partial\Omega^{\rho_1}_{\rho_0}) > \sigma\right\}.$$

Then for $m > 0$, $\sigma > 0$, $L > R_0$, and $\rho_1 > R_0$ there is a constant $C(\tau, m, \sigma, L, \rho_1, \rho_0)$ such that if $g \in \mathcal{M}$ and the ADM mass of g is less than m, then

$$\begin{aligned}\sup_{x\in(\Omega^L_{\rho_0\rho_1})_\sigma} \exp\left(\int_{\partial\Omega^L_{\rho_0\rho_1}} \left|H_D(x, y)\frac{\partial(\alpha - 2u)}{\partial\nu}(y)\right| + \left|(\alpha - 2u)(y)\frac{\partial H_D}{\partial\nu}(x, y)\right| dy\right) \\ \le \exp\big[C(\tau, m, \sigma, L, \rho_1, \rho_0)\big],\end{aligned} \tag{6-23}$$

where τ is the constant appearing in (1-33) and $C(, \tau, m, \sigma, L, \rho_1, \rho_0)$ is a constant depending on τ, m, σ, L, ρ_1, and ρ_0.

Proof. Much of the proof remains the same as it was in the radially monotone case. The only difference is that we need to estimate

$$\int_{\partial\Omega^L_{\rho_0\rho_1}\cap\{\rho=\rho_0\}} |\alpha - 2u|, \tag{6-24}$$

instead of

$$\int_{\partial\Omega^L_{\rho_0\rho_1}\cap\{\rho=\rho_0\}} \left|\frac{\partial(\alpha - 2u)}{\partial\nu}\right|. \tag{6-25}$$

This we did in Lemma 6.1. □

We now estimate the $W^{1,p}$ norm of $\alpha - 2u$. Using the function H_D instead of H_N complicates our estimate of $\|\nabla(\alpha - 2u)\|_{L^p(\Omega^L_{\rho_0\rho_1})}$. We resort to shrinking our region a bit.

Lemma 6.3. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is strongly uniformly asymptotically flat outside of radius R_0. Suppose also that $\mathcal{M}$ is area enlarging at ρ_0. For any $\rho_1 > \rho_0$, L, $1 \le p < 2$, $\sigma > 0$, and $\epsilon > 0$ there is a $\delta > 0$ such that if $g \in \mathcal{M}$ and*

$$m(g) < \delta, \tag{6-26}$$

then

$$\|\alpha - 2u\|_{W^{1,p}((\Omega^L_{\rho_0\rho_1})_\sigma)} < \epsilon. \tag{6-27}$$

Here

$$(\Omega^L_{\rho_0\rho_1})_\sigma := \big\{x \in \Omega^L_{\rho_0\rho_1} : d(x, \partial\Omega^L_{\rho_0\rho_1}) \ge \sigma\big\}. \tag{6-28}$$

Proof. We may estimate the L^p norm of $\alpha - 2u$ much as we did in Proposition 4.2. We once again consider $\tilde{L} > L$ and $\tilde{\rho}_1 > \rho_0$. As before,

$$\begin{aligned}\int_{(\Omega^L_{\rho_0\rho_1})_\sigma} &|\alpha - 2u|^p \\ &\le C(p) \int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \bigg(\int_{\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}} \bigg|(\alpha - 2u)\frac{\partial H_D}{\partial \nu}\bigg| + \bigg|H_D \frac{\partial(\alpha - 2u)}{\partial \nu}\bigg|\bigg)^p \\ &\qquad + \bigg(\int_{\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}} |H_D \Delta(\alpha - 2u)|\bigg)^p \, dx.\end{aligned} \tag{6-29}$$

On $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} - \{\rho = \rho_0\}$ we have the following bound on the boundary terms

$$\frac{24C\tilde{\rho}_1}{\pi\tilde{L}|\tilde{L} - L|} + \frac{3C\tilde{L}}{\pi\tilde{\rho}_1|\tilde{\rho}_1 - \rho_1|} + \frac{24C\tilde{\rho}_1 \log\big(2\sqrt{\tilde{L}^2 + \tilde{\rho}_1^2}\big)}{\pi\tilde{L}^2} + \frac{3C\tilde{L}\log\big(\sqrt{\tilde{L}^2 + \tilde{\rho}_1^2}\big)}{\pi\tilde{\rho}_1^2}. \tag{6-30}$$

Using the proof of Lemma 6.1 for terms on $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0\}$, we have the estimate

$$\frac{1}{\pi\sigma}\bigg(\big(4m + 4\sqrt{\tilde{L}m}\big)\log\bigg(\frac{\tilde{\rho}_1}{\rho_0}\bigg) + \frac{6\pi C\tilde{\rho}_1}{\tilde{L}} + \tilde{L}\tilde{\epsilon}(\tilde{\rho}_1, m) + C(\tau)(\tilde{\rho}_1)^{-\tau}\bigg). \tag{6-31}$$

If we let $\tilde{\rho}_1 = \tilde{L}^{2/3}$, then we may see that we may pick $\tilde{L}$ large enough and m small enough to ensure

$$\|\alpha - 2u\|_{L^p((\Omega^L_{\rho_0\rho_q})_\sigma)} < \frac{\epsilon}{2}. \tag{6-32}$$

If we differentiate Green's representation formula with H_D we get

$$\begin{aligned}\nabla(\alpha-2u)(x) = \int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}} (\alpha-2u)\nabla_x\left(\frac{\partial H_D}{\partial \nu}\right) - \nabla_x(H_D(x,y))\frac{\partial(\alpha-2u)}{\partial\nu}\,dy \\ + \int_{\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}} \nabla_x(H_D(x,y))\Delta(\alpha-2u)\,dy.\end{aligned} \tag{6-33}$$

On $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0\}$ the above expression is particularly difficult to work with. The issue is that we cannot integrate

$$\left|\nabla_x\left(\frac{\partial H_D}{\partial\nu}\right)\right| \sim \frac{1}{|x-y|^2} \tag{6-34}$$

for x near the boundary, and so we cannot complete the estimate of $\|\alpha - 2u\|_{W^{1,p}}$ in the same way we proved Proposition 4.2.

As we have done before, we take the absolute value of both sides and raise the result to the power p and then integrate to see that

$$\int_{(\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1})_\sigma} |\nabla(\alpha-2u)|^p \tag{6-35}$$

is bounded above by

$$\begin{aligned}C(p)\int_{(\Omega^L_{\rho_0\rho_1})_\sigma}\left(\left(\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}}\left|\frac{\partial(\alpha-2u)}{\partial\nu}\nabla_x H_D\right| + \left|(\alpha-2u)\nabla_x\frac{\partial H_D}{\partial\nu}\right|dy\right)^p\right. \\ \left. + \left(\int_{\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}}\left|\Delta(\alpha-2u)\nabla_x H_D\right|dy\right)^p\right)dx.\end{aligned} \tag{6-36}$$

We once again split the first term into the following two pieces:

$$\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} - \{\rho = \rho_0\} \tag{6-37}$$

and

$$\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0\}. \tag{6-38}$$

Both pieces are relatively easy to estimate. For the first piece the estimates are similar to the above.

As was noted earlier, the gradient of $\nabla_x \frac{\partial H_D}{\partial \nu}$ isn't integrable over $\Omega^L_{\rho_0\rho_1}$ for y in $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0\}$. However, $\nabla_x \frac{\partial H_D}{\partial \nu}$ is much better behaved away from $\partial\Omega^L_{\rho_0\rho_1}$. We now attempt to estimate

$$\int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \left(\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1}} \left| (\alpha - 2u)\nabla_x \frac{\partial H_D}{\partial \nu} \right| dy \right)^p dx. \tag{6-39}$$

As we did before, we split $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0\}$ into

$$\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \le L\} \tag{6-40}$$

and

$$\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| > L\}. \tag{6-41}$$

We start with the piece (6-40). We may use Minkowski's integral inequality [Folland 1999] to see that

$$\left(\int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \left(\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho=\rho_0,\ |z|\le L\}} \left| (\alpha - 2u)\nabla \frac{\partial H_D}{\partial \nu} \right| dy \right)^p dx \right)^{1/p} \tag{6-42}$$

is bounded above by

$$\int_{\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho=\rho_0,\ |z|\le L\}} |\alpha - 2u| \left(\int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \left| \nabla \frac{\partial H_D}{\partial \nu} \right|^p dx \right)^{1/p} dy. \tag{6-43}$$

We now estimate

$$\int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \left| \nabla \frac{\partial H_D}{\partial \nu} \right|^p dx \tag{6-44}$$

for y in $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \le L\}$. Both $\partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \le L\}$ and $(\Omega^L_{\rho_0\rho_1})_\sigma$ are contained in $\Omega^{2L}_{\rho_0\rho_1}$. Thus, if we let r_0 be the diameter of $\Omega^{2L}_{\rho_0\rho_1}$, then for all $y \in \partial\Omega^{\tilde{L}}_{\rho_0\tilde{\rho}_1} \cap \{\rho = \rho_0,\ |z| \le L\}$ we have

$$\begin{aligned} \int_{(\Omega^L_{\rho_0\rho_1})_\sigma} \left| \nabla_x \frac{\partial H_D}{\partial \nu} \right|^p &\le \int_{B(y,r_0)\setminus B(y,\sigma)} \frac{3^p}{\pi^p |x-y|^{2p}}\, dx \\ &= 3^p \pi^{1-p} 2 \int_\sigma^{r_0} C r^{-2p+1}\, dr = C(p, L, \rho_1, \sigma). \end{aligned} \tag{6-45}$$

Thus, we may see that

$$\left(\int_{(\Omega^L_{\rho_0\rho_1})_\sigma}\left(\int_{\partial\Omega^{\tilde L}_{\rho_0\tilde\rho_1}\cap\{\rho=\rho_0,\,|z|\le L\}}\left|(\alpha-2u)\nabla\frac{\partial H_D}{\partial\nu}\right|dy\right)^p dx\right)^{1/p} \le C(p,L,\rho_1,\sigma)^{1/p}\int_{\partial\Omega^{\tilde L}_{\rho_0\tilde\rho_1}\cap\{\rho=\rho_0,\,|z|\le L\}}|\alpha-2u|. \tag{6-46}$$

Over (6-41) we have

$$\left|\nabla\frac{H_D}{\partial\nu}\right|\le\frac{12}{\pi L^2}. \tag{6-47}$$

Thus, we have

$$\left(\int_{(\Omega^L_{\rho_0\rho_1})_\sigma}\left(\int_{\partial\Omega^{\tilde L}_{\rho_0\tilde\rho_1}\cap\{\rho=\rho_0,\,|z|>L\}}\left|(\alpha-2u)\nabla\frac{\partial H_D}{\partial\nu}\right|dy\right)^p dx\right)^{1/p} \le\left(\frac{12\rho_1}{L}\right)^{1/p}\int_{\partial\Omega^{\tilde L}_{\rho_0\tilde\rho_1}\cap\{\rho=\rho_0,\,|z|>L\}}|(\alpha-2u)|\,dy. \tag{6-48}$$

For the last term in (6-36) we may use the Riesz potential estimate as we have done before. Putting everything together gives us the result. □

In fact, the steps required in the above proof give us a corollary analogous to Corollary 4.5.

Corollary 6.4. *Let $\mathcal{M}$ be a family of axisymmetric metrics with nonnegative scalar curvature and empty boundary which is strongly uniformly asymptotically flat outside of radius R_0. Suppose $\mathcal{M}$ is area enlarging as well. For any $L,\ \rho_1,\ 1\le p<2$, and $\epsilon>0$ there exist a $\delta>0$ such that if $g\in\mathcal{M}$ and*

$$m(g)<\delta, \tag{6-49}$$

then

$$\int_{\Omega^L_{\rho_0,\rho_1}}|\alpha-2u|^p<\frac{\epsilon|\log\rho_0|^p}{\rho_0^p} \tag{6-50}$$

and

$$\int_{\Omega^L_{\rho_0,\rho_1}}|\nabla(\alpha-2u)|^p\le\frac{\epsilon|\log\rho_0|^p}{\rho_0^p}. \tag{6-51}$$

Proof. The proofs of (6-50) and (6-51) are similar. We only prove (6-51). Observe that

$$\Omega^L_{2\rho_0\rho_1}\subset(\Omega^{L+\sigma}_{\rho_0(\rho_1+\sigma)})_\sigma. \tag{6-52}$$

In particular, we see from the estimates in the above theorem that

$$\int_{\Omega^L_{2\rho_0\rho_1}} |\nabla(\alpha-2u)|^p \le \int_{(\Omega^{L+\sigma}_{\rho_0(\rho_1+\sigma)})_\sigma} |\nabla(\alpha-2u)|^p \tag{6-53}$$

is bounded above by

$$C(p,L,\rho_1,\sigma)\biggl[\bigl(4m+4\sqrt{\tilde{L}m}\bigr)\log\biggl(\frac{\tilde{\rho}_1}{\rho_0}\biggr)+D(m,\tilde{L},\tilde{\rho}_1,\tau)\biggr]^p \\ +E(p,\tilde{L},\tilde{\rho}_1)\biggl(\frac{4m+4\sqrt{\tilde{L}m}}{\rho_0}\biggr)^p+F(p,\tilde{L},\tilde{\rho}_1), \tag{6-54}$$

where $C(p,L,\rho_1,\sigma)$ is a combination of the constants found in (6-46) and (6-48), $D(m,\tilde{L},\tilde{\rho}_1,\tau)$ is the remainder of (6-22), $E(p,\tilde{L},\tilde{\rho}_1)$ comes from the Riesz potential estimate, and $F(p,\tilde{L},\tilde{\rho}_1)$ is the bound on the remaining boundary terms estimated in (6-36). A simple calculation shows that for $1<p<2$

$$C(L,\rho_1,\sigma)\le C(p)\sigma^{-p}, \tag{6-55}$$

since $2-2p>-p$. For $p=1$, we have

$$C(L,\rho_1,\sigma)\le C(L,\rho_1)\log(\sigma). \tag{6-56}$$

If we plug the above into (6-54) with $\sigma=\rho_0$, then we may see that choosing $\tilde{L}$ and $\tilde{\rho}_1$ large enough, and choosing mass to be small enough gives the result. □

We may now prove a theorem analogous to Proposition 4.2.

Lemma 6.5. *Let $\mathcal{M}$ be an uniformly asymptotically flat family of metrics with nonnegative scalar curvature and empty boundary. Suppose that $\mathcal{M}$ is area enlarging. Let $\Omega^L_{\rho_0\rho_1}$ denote the rectangle given by $\{(\rho,z)\mid \rho_0\le\rho\le\rho_1,\ |z|\le\frac{L}{2}\}$ and let $(\Omega^L_{\rho_0\rho_1})_\sigma$ denote $\{x\in\Omega^L_{\rho_0\rho_1}\mid d(x,\partial\Omega^L_{\rho_0\rho_1})>\sigma\}$. For any $1\le p<2$, $\sigma>0$, $\rho_0>0$, and $\epsilon>0$ there exists a $\delta>0$ such that if g is in our collection of uniformly asymptotically flat metrics, the ADM mass of g is less than δ, and, in the axisymmetric coordinate representation of g then*

$$\|e^{|\alpha-2u|}-1\|_{W^{1,p}((\Omega^L_{\rho_0\rho_1})_\sigma)}<\epsilon. \tag{6-57}$$

Proof. The proof follows the same line as in the radially monotone case, except we use Lemma 6.3 instead of Proposition 4.2. It can be shown that Corollary 4.7 can be adapted to the function H_D. Thus, we also use Corollary 6.2 instead of Lemma 4.8. □

Now that we have analogues of all the estimates we made in the radially monotone case, the proofs of Theorems 1.4, 1.6, 1.7, 1.8 and 1.10 follow almost exactly as they did in the radially monotone case. The only theorem whose modification to the area-enlarging case requires a little care is Theorem 1.10. Since Corollary 6.4 has a slightly different hypothesis than Corollary 4.5, we must show that the conclusion of Lemma 5.5 holds with a slightly weaker hypothesis.

Lemma 6.6. *Let f be a measurable function on $\Omega^L_{0\rho_1}$. Suppose for each t we have the estimate*

$$\int_{\Omega_t^{\rho_1}} |f| \leq \frac{\epsilon |\log(t)|^{\tilde{q}}}{t^q} \tag{6-58}$$

for some $\epsilon > 0$, $q > 0$, and $\tilde{q}$. Suppose $\sigma > q$. Then, there exists a constant, denoted $C(\sigma, q, \tilde{q})$, depending only on σ, q, and $\tilde{q}$ such that

$$\int_{\Omega_0^{\rho_1}} \rho^\sigma |f| \leq C(\sigma, q, \tilde{q})\epsilon. \tag{6-59}$$

Proof. As before, let $t_n = 2^{-n}\rho_1$ and let $\Omega_{t_n,t_{n-1}}$ be the following rectangle.

$$\Omega_{t_n,t_{n-1}} = \left\{ t_n \leq \rho \leq t_{n-1},\ |z| \leq \frac{L}{2} \right\} \tag{6-60}$$

From the monotone convergence theorem we see that

$$\int_{\Omega_{0,t_0}} \rho^\sigma |f| = \sum_1^\infty \int_{\Omega_{t_n,t_{n-1}}} \rho^\sigma |f|. \tag{6-61}$$

We now make the estimate

$$\int_{\Omega_{t_n,t_{n-1}}} \rho^\sigma |f| \leq t_{n-1}^\sigma \frac{\epsilon |\log(t_n)|^{\tilde{q}}}{t_n^q} = 2^q \rho_1^{\sigma-q} (2^{\sigma-q})^{-n} |\log(2^{-n}\rho_1)|^{\tilde{q}} \epsilon. \tag{6-62}$$

This gives a convergent series so long as $\sigma > q$, where we have used that $\sigma - q = \lambda > 0$ and

$$\lim_{n\to\infty} \rho_1 2^{-n} |\log(\rho_1 2^{-n})|^{2\tilde{q}/\lambda} = 0. \tag{6-63}$$

In total, we have the estimate

$$\int_{\Omega_{0,t_0}} \rho^\sigma |f| \leq C(\sigma, q, \tilde{q})\epsilon. \tag{6-64}$$ □

Now we can show that Lemma 5.6 holds in the area-enlarging case and so Theorem 1.10 also holds in the area-enlarging case.

Appendix A: The case of nonempty boundaries

Recall that it is physically desirable to explicitly include manifolds with minimal surface boundary, since we shouldn't expect to have any physical knowledge of the metric inside of a minimal surface. It is possible to deduce the following mass formula for axisymmetric manifolds with connected boundary [Khuri et al. 2019]:

$$m(g)=\frac{1}{16\pi}\int_{\mathbb{R}^3} 2|\nabla\bar{u}|^2+e^{2(u-\alpha)}R_g\,dx+\tfrac{1}{4}\int_{-m_0}^{m_0}\bar{\alpha}(0,z)-2\bar{u}(0,z)\,dz+m_0, \tag{A-1}$$

where $\bar{\alpha}$ and $\bar{u}$ are regularizations of the coordinate functions α and u, respectively, and m_0 is a positive constant determined uniquely by the metric g. Explicitly, the functions $\bar{\alpha}$ and $\bar{u}$ are given by

$$\bar{u}=u-u_0, \tag{A-2}$$

$$\bar{\alpha}=\alpha-\alpha_0. \tag{A-3}$$

where α_0 and u_0 are the coordinate functions associated to the Schwarzschild metric of mass m_0 in Weyl coordinates, coordinates in which the minimal surface is given by a rod of length $2m_0$:

$$\alpha_0=\tfrac{1}{2}\log\left[\frac{\left(\sqrt{\rho^2+(z-m_0)^2}+\sqrt{\rho^2+(z+m_0)^2}\right)^2-4m_0^2}{4\sqrt{\rho^2+(z-m_0)^2}\sqrt{\rho^2+(z+m_0)^2}}\right], \tag{A-4}$$

$$u_0=\tfrac{1}{2}\log\left[\frac{\sqrt{\rho^2+(z-m_0)^2}+\sqrt{\rho^2+(z+m_0)^2}-2m_0}{\sqrt{\rho^2+(z-m_0)^2}+\sqrt{\rho^2+(z+m_0)^2}+2m_0}\right]. \tag{A-5}$$

Chruściel and Nguyen [2011] have shown that the constant m_0 is bounded by

$$m(g)\geq\frac{\pi}{4}m_0, \tag{A-6}$$

given the hypothesis of the positive mass theorem. We have the following theorem:

Theorem A.1. *Let $\mathcal{M}$ be a family of axisymmetric uniformly asymptotically flat metrics with nonnegative scalar curvature. Suppose that $\mathcal{M}$ is either area enlarging, with the corresponding stronger asymptotic falloff, or radially monotone. Additionally, we allow any (M,g) in $\mathcal{M}$ to have a connected minimal surface boundary. In this case, we use the cylindrical coordinates for which the minimal surface is a rod on the axis of symmetry of length $2m_0$ centered about the origin, and we assume (M,g) satisfies the following inequality on its minimal surface boundary*:

$$\tfrac{1}{4}\int_{-m_0}^{m_0}\bar{\alpha}-2\bar{u}(0,z)\,dz+m_0\geq 0, \tag{A-7}$$

where $\bar{\alpha}$ and $\bar{u}$ are as above. Then, for any $\epsilon>0$ there exists a $\delta>0$ such that if (M,g) is in $\mathcal{M}$ and $m(g)<\delta$, then

$$\|g-\delta_{\mathbb{R}^3}\|_{W^{1,p}(\widetilde{\Omega}^{\rho_1}_{\rho_0}(\sigma))}<\epsilon \tag{A-8}$$

$$\|q-\delta_{\mathbb{R}^2}\|_{W^{1,p}(\Omega^{\rho_1}_{\rho_0}(\sigma))}<\epsilon. \tag{A-9}$$

Remark A.2. It is important to note that in the case of a nonempty boundary, we have had to specify the cylindrical coordinates for which the boundary is on the axis. Then, the radial monotonicity and area-enlarging inequality are stated with respect to these coordinates. These conditions are geometric, since in choosing the boundary to be on the axis, we have removed any freedom in the choice of a conformal transformation.

Proof. In Weyl coordinates, with the boundary of the manifold represented as a rod on the axis, we see that for any fixed parameter $\rho_0 > 0$, we have that the functions α_0 and u_0, and their gradients, converge uniformly to zero on $\rho^{-1}[\rho_0, \infty)$ as $m_0 \to 0$. It thus follows that on any compact set away from the axis, say Ω, we have

$$\|\alpha_0\|_{W^{1,2}(\Omega)} \to 0, \tag{A-10}$$

$$\|u_0\|_{W^{1,2}(\Omega)} \to 0, \tag{A-11}$$

as $m_0 \to 0$. Finally, we recall that $m \geq \frac{\pi}{4} m_0$ [Chruściel and Nguyen 2011]. We now have all the ingredients necessary to extend the proofs of this paper to the case of manifolds with boundary. Note that an analogue of Corollary 3.2 holds for $\bar{u}$ by the mass formula (A-1) and (A-7). Thus, we may use the Cauchy–Schwartz inequality to show that

$$\|u\|^2_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})} \leq 2\big(\|u_0\|^2_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})} + \|\bar{u}\|^2_{W^{1,2}(\Omega^{\rho_1}_{\rho_0})}\big) \tag{A-12}$$

is bounded by the mass. At this point, the rest of the proof is the same as in the case of empty boundary. □

As we see in the next section, the nonextreme Kerr–Newman metrics satisfy all of the conditions in the above theorem strictly. Thus, small perturbations will also satisfy these conditions.

Appendix B: Examples

Kerr–Newman. In this section, we show that the Kerr–Newman family of metrics satisfy the radial monotone condition and the area enlarging condition, and (A-7). This is done by a direct calculation. We take the familiar Brill–Lindquist coordinates and transform them into cylindrical coordinates. Unfortunately, the simple expression of the Kerr–Newman metric in Brill–Lindquist coordinates becomes rather complicated when it is written in cylindrical coordinates. The procedure itself is uncomplicated, since there is an explicit map between these two coordinates. The change of coordinates depends on the charge, angular momentum, and mass of the Kerr–Newman metric. Once the map has been constructed, we use the expression for the metric in Brill–Lindquist to write down the expression for the metric in cylindrical coordinates.

We now describe in detail the coordinate change from Brill–Lindquist coordinates to cylindrical coordinates and write down the exact formula for the metric functions u and α. It is convenient to introduce a third coordinate system between Brill–Lindquist and cylindrical. We shall use the prolate-spheroidal coordinates. We will first consider the map from prolate-spheroidal coordinates to Brill–Lindquist coordinates, and then pull back the metric. Let a denote the angular momentum parameter, let e denote the charge parameter, and let m denote the mass parameter, then, in Brill–Lindquist coordinates, the Kerr metric takes the form

$$g = \frac{\sigma}{\gamma}\,dr^2 + \sigma\,d\theta^2 + \frac{\sin^2(\theta)}{\sigma}\big[(r^2+a^2)^2 - a^2\sin^2(\theta)\gamma(r)\big]\,d\phi^2 \tag{B-1}$$

for

$$\gamma(r) = r^2 - 2mr + a^2 + e^2 \tag{B-2}$$

and

$$\sigma(r,\theta) = r^2 + a^2\cos^2(\theta). \tag{B-3}$$

The map from prolate spheroidal coordinates (x, y, ϕ) to Brill–Lindquist coordinates (r, θ, ϕ) is given by

$$r = x\sqrt{m^2-(a^2+e^2)} + m \tag{B-4}$$

$$\theta = \cos^{-1}(y) \tag{B-5}$$

It turns out that the parameter m_0 appearing in Appendix A is given by

$$m_0 = \sqrt{m^2-(a^2+e^2)}. \tag{B-6}$$

The map from cylindrical coordinates to prolate spheroidal is, unfortunately, less simple.

$$x = \frac{\sqrt{\rho^2+(z+m_0)^2}+\sqrt{\rho^2+(z-m_0)^2}}{2m_0} \tag{B-7}$$

$$y = \frac{\sqrt{\rho^2+(z+m_0)^2}-\sqrt{\rho^2+(z-m_0)^2}}{2m_0} \tag{B-8}$$

One may observe that the minimal surface in the Kerr–Newman metric is a rod on the ρ axis.

We now pull back the Kerr–Newman metric twice to obtain the formulas for the functions u and α in cylindrical coordinates. The end results of this process are the following formulas:

$$u(\rho,z) = -\tfrac{1}{2}\log\left[\frac{(1-y^2)\big([(m_0x+m)^2+a^2]^2 - a^2m_0^2[1-y^2][x^2-1]\big)}{\rho^2([m_0x+m]^2+a^2y^2)}\right], \tag{B-9}$$

$$\alpha(\rho,z)=\tfrac{1}{2}\log\left[\frac{(m_0x+m)^2+a^2y^2}{m_0^2(x^2-y^2)}\right]+u(\rho,z). \tag{B-10}$$

When written entirely in terms of (ρ, z), these two equations are very cumbersome. Luckily, for the purpose of verifying the radial monotonicity condition and the area enlarging condition, writing everything in terms of (ρ, z) turns out to be unnecessary.

Proposition B.1. *Nonextreme Kerr–Newman metrics are radially monotone in the coordinates for which the minimal surface is a rod on the axis.*

A straight forward calculation shows that

$$\frac{\partial}{\partial\rho}=\frac{\rho}{(\rho^2+(z+m_0)^2)^{1/2}(\rho^2+(z-m_0)^2)^{1/2}}\left(x\frac{\partial}{\partial x}-y\frac{\partial}{\partial y}\right). \tag{B-11}$$

Thus, we see that

$$\begin{aligned}\frac{\partial(\alpha-2u)}{\partial\rho}&=f(\rho,z)\left(x\frac{\partial}{\partial x}-y\frac{\partial}{\partial y}\right)\\&\quad\times\tfrac{1}{2}\log\left(\frac{[(m_0x+m)^2+a^2]^2-a^2m_0^2[1-y^2][x^2-1]}{m_0^4(x^2-1)(x^2-y^2)}\right),\end{aligned} \tag{B-12}$$

where $f(\rho, z)$ is the nonnegative function appearing in front of the derivatives in (B-11). Since $f(\rho, z)$ is nonnegative, we may restrict our analysis to the second term on the right. Taking the derivatives and collecting terms leaves us with

$$\begin{aligned}&\frac{4m_0x(m_0x+m)[(m_0x+m)^2+a^2]-2a^2m_0^2x^2(1-y^2)}{[(m_0x+m)^2+a^2]^2-a^2m_0^2(1-y^2)(x^2-1)}\\&\qquad-\frac{2x^2((x^2-1)+(x^2-y^2))}{(x^2-1)(x^2-y^2)}\\&\qquad-\left[\frac{2a^2m_0^2(x^2-1)y^2}{[(m_0x+m)^2+a^2]^2-a^2m_0^2(1-y^2)(x^2-1)}+\frac{2y^2}{x^2-y^2}\right].\end{aligned} \tag{B-13}$$

The third term in brackets is nonnegative, so we must analyze the interplay of the first two terms.

We expand

$$\frac{2x^2((x^2-1)+(x^2-y^2))}{(x^2-1)(x^2-y^2)} \tag{B-14}$$

to

$$\frac{2x^2}{x^2-1}+\frac{2x^2}{x^2-y^2}. \tag{B-15}$$

From the range of values that x and y can take, we may deduce that the denominators of both fractions are smaller than x^2. Thus, we have

$$\frac{2x^2}{x^2-1}+\frac{2x^2}{x^2-y^2}>4. \tag{B-16}$$

We now observe that

$$[(m_0x+m)^2+a^2]^2-a^2m_0^2(1-y^2)(x^2-1)\geq(m_0x+m)^4+a^2(m_0x+m)^2. \tag{B-17}$$

As a consequence, we have that

$$\frac{4m_0x(m_0x+m)[(m_0x+m)^2+a^2]-2a^2m_0^2x^2(1-y^2)}{[(m_0x+m)^2+a^2]^2-a^2m_0^2(1-y^2)(x^2-1)}\leq 4. \tag{B-18}$$

Putting everything together shows that

$$\frac{\partial(\alpha-2u)}{\partial\rho}<0. \tag{B-19}$$

Luckily, showing that Kerr–Newman metrics satisfy (A-7) follows quickly from the above expressions for u and α. In fact, one may check that $\bar{\alpha}-2\bar{u}$ is nonnegative on the rod giving the minimal surface.

Proposition B.2. *Let g be a nonextreme Kerr–Newman metric, and let $\bar{\alpha}$ and $\bar{u}$ be as described above. Then, we have that*

$$(\bar{\alpha}-2\bar{u})(0,z)\geq 0, \tag{B-20}$$

for $|z|\leq m_0$. The inequality is strict, unless g is a Schwarzschild metric.

Proof. Once again, the proof consists of a calculation. Using the above expressions for α and u coming from a Kerr–Newman metric, we see that

$$\bar{\alpha}-2\bar{u}=\tfrac{1}{2}\log\left[\frac{([m_0x+m]^2+a^2)^2-a^2m_0^2(1-y^2)(x^2-1)}{m_0^4(x+1)^4}\right]. \tag{B-21}$$

In prolate spheroidal coordinates, the minimal surface rod is given by

$$\{(x,y,\phi):x=1\}.$$

Thus, the above simplifies to

$$\tfrac{1}{2}\log\left[\frac{([m_0+m]^2+a^2)^2}{16m_0^4}\right]. \tag{B-22}$$

Since $m\geq m_0$ and $a\geq 0$, it follows that the above is nonnegative, and only zero in the case that the metric g is Schwarzschild. □

It is interesting to explore some of the geometric meaning behind the condition of radial monotonicity. In coordinates, radial monotonicity implies that

$$\frac{\partial(\alpha - 2u)}{\partial\rho} \leq 0. \tag{B-23}$$

Recall from the proof of Proposition 5.1 that the coordinate function $\alpha - 2u$ controls the area of axisymmetric surfaces. Thus, it is reasonable to suppose that the radial monotonicity condition is an assumption on the mean curvature of the level sets of the function ρ, which is the solution to (1-10). It turns out that this is the case, although in a slightly round about way.

Proposition B.3. *Suppose that g is an asymptotically flat axisymmetric metric and ρ is the solution to (1-10) for g. The metric g is radially monotone if and only if the level sets of ρ form a family of surfaces evolving by a sub-inverse-mean-curvature flow.*

Proof. Let η denote the killing field generating the axisymmetry of (M, g). We start by observing that we may lift any function ω on M/S^1 to a function on M, which we also denote ω. When considered as a function on M we have

$$g(\nabla\omega, \eta) = 0, \tag{B-24}$$

since we lifted ω by transporting it along the flow lines of η. Let q denote the orbit metric of M/S^1. Recall that

$$q(X, Y) = g(\overline{X}, \overline{Y}) - \frac{g(\overline{X}, \eta)g(\overline{Y}, \eta)}{|\eta|_g^2}, \tag{B-25}$$

where X and Y are the images of $\overline{X}$ and $\overline{Y}$ under the projection map, respectively. From the above, we may conclude that for any two functions ω and h on M/S^1 we have

$$q(\nabla\omega, \nabla h) = g(\nabla\omega, \nabla h). \tag{B-26}$$

We have abused notation slightly in using ∇ to denote both the gradient in $(M/S^1, q)$ and in (M, g).

It is a standard computation to see that the mean curvature of the level sets of ρ is given by

$$H = \operatorname{div}_g\left(\frac{\nabla\rho}{|\nabla\rho|_g}\right). \tag{B-27}$$

We expand out the right hand side to get

$$\operatorname{div}_g\left(\frac{\nabla\rho}{|\nabla\rho|_g}\right) = \frac{1}{|\nabla\rho|_g}\left(\Delta_g\rho - \frac{g(\nabla\rho, \nabla|\nabla\rho|)}{|\nabla\rho|}\right) \tag{B-28}$$

We now use Equation (1-10) for ρ to rewrite the above as

$$\text{(B-29)}\quad \frac{1}{|\nabla\rho|}\left(\frac{g(\nabla\rho,\nabla|\eta|)}{|\eta|}-\frac{g(\nabla\rho,\nabla|\nabla\rho|)}{|\nabla\rho|}\right)=\frac{1}{|\nabla\rho|}g\left(\nabla\rho,\nabla\log\frac{|\eta|}{|\nabla\rho|}\right).$$

From axisymmetry, $|\nabla\rho|$ and $|\eta|$ are functions on M/S^1. In particular

$$\text{(B-30)}\quad g\left(\nabla\rho,\nabla\log\frac{|\eta|}{|\nabla\rho|}\right)=q\left(\nabla\rho,\nabla\log\frac{|\eta|}{|\nabla\rho|}\right).$$

Recalling the radial monotonicity condition (1-17) and noting that log is a monotone increasing function, we see that

$$\text{(B-31)}\quad q\left(\nabla\rho,\nabla\log\left(\frac{|\eta|}{\rho|\nabla\rho|}\right)\right)\leq 0,$$

since in the orbit space M/S^1 we have

$$\text{(B-32)}\quad \frac{\partial}{\partial\rho}=\left|\frac{\partial}{\partial\rho}\right|_q^2\nabla\rho.$$

We may plug (B-29) and (B-27) into (B-31) to see that

$$\text{(B-33)}\quad 0\geq q\left(\nabla\rho,\nabla\log\left(\frac{|\eta|}{|\nabla\rho|}\right)\right)-q(\nabla\rho,\nabla\log\rho)=|\nabla\rho|H-|\nabla\rho||\nabla\log\rho|.$$

Dividing both sides by $|\nabla\rho|$ and rearranging terms gives

$$\text{(B-34)}\quad |\nabla\log\rho|\geq H.$$

The above equation is precisely the statement that the level sets of ρ give a sub-inverse-mean-curvature flow. □

It is relatively easy to see that if a metric is radially monotone everywhere, then it must also be area enlarging everywhere. In particular, the following proposition implies that Kerr–Newman metrics are area enlarging.

Proposition B.4. *Let g be an asymptotically flat metric which is everywhere radially monotone. Then g is everywhere area enlarging.*

Proof. Since g is assumed to be globally radially monotone, we have

$$\text{(B-35)}\quad \frac{\partial(\alpha-2u)}{\partial\rho}\leq 0.$$

As g is asymptotically flat, we know that

$$\text{(B-36)}\quad \lim_{\rho\to\infty}(\alpha-2u)(\rho,z)=0$$

for all z. Thus, using the fundamental theorem, we may see that

$$0 \leq -\int_{\rho_0}^{\infty} \frac{\partial(\alpha - 2u)}{\partial \rho}(\rho, z)\, d\rho = (\alpha - 2u)(\rho_0, z). \tag{B-37}$$

This is precisely the coordinate expression of the area enlarging condition. □

We now find several examples of metrics which are area enlarging and strongly asymptotically flat.

Axisymmetric geometrostatic. Here we show that the axisymmetric geometrostatic metrics are area-enlarging and strongly asymptotically flat. Recall that the general form of a geometrostatic metric is

$$(M, g) = \left(\mathbb{R}^3 \backslash \{x_i\}_1^n, (\chi\psi)^2 \delta_{\mathbb{R}^3}\right), \tag{B-38}$$

where for positive numbers $\{a_i\}_1^n$ and $\{b_i\}_1^n$ we have

$$\chi(x) = 1 + \sum_{i=1}^{n} \frac{a_i}{|x - x_i|} \tag{B-39}$$

and

$$\psi(x) = 1 + \sum_{i=1}^{n} \frac{b_i}{|x - x_i|}. \tag{B-40}$$

If the points $\{x_i\}$ lie on a common line, then the resulting metric will be axisymmetric. The axis of symmetry will be the line on which the x_i lie. After a rotation, we may suppose that the axis of symmetry is the z-axis. We may now see that the usual Euclidean cylindrical coordinates are also cylindrical coordinates for (M, g). In particular

$$g = (\chi\psi)^2 (d\rho^2 + dz^2 + \rho^2 d\phi^2). \tag{B-41}$$

A quick calculation shows that the coordinate function α vanishes and

$$u = -\log(\chi\psi). \tag{B-42}$$

Since both χ and ψ are strictly larger than one, we see that u is negative. Since $\alpha = 0$, it is clear that

$$\alpha - 2u \geq 0. \tag{B-43}$$

This is precisely the coordinate expression of the area-enlarging condition. That (M, g) is also strongly asymptotically flat follows trivially from the fact that $\alpha = 0$.

Conformal metrics. Here we show that asymptotically flat axisymmetric metrics with nonnegative scalar curvature which are conformal to Euclidean space and have an axisymmetric, minimal, and connected boundary, or an empty one, satisfy the area enlarging condition and the strongly asymptotically flat condition.

Suppose (M, g) is as above. Then there is some constant m_1 [Chruściel and Nguyen 2011] and function u such that

$$(M, g) = \big(\mathbb{R}^3 \backslash B_{m_1}(0), e^{-2u}\delta_{\mathbb{R}^3}\big). \tag{B-44}$$

Written in cylindrical coordinates

$$g = e^{-2u}(d\rho^2 + dz^2) + \rho^2 e^{-2u}\, d\phi^2 \tag{B-45}$$

Since ∂B_{m_1} is a minimal surface, from the formula for mean curvature we see that [Chruściel and Nguyen 2011]

$$\frac{\partial u}{\partial \nu}\bigg|_{\partial B_{m_1}} = \frac{1}{m_1}. \tag{B-46}$$

Since we have assumed that the scalar curvature is nonnegative, we may use the scalar curvature formula (2-2) together with the Hopf lemma and the maximum principle to conclude that

$$\sup_{B_{r_0} \backslash B_{m_1}} u = \sup_{\partial B_{r_0}} u. \tag{B-47}$$

Since we know from the fact that g is asymptotically flat that u vanishes at infinity, we may conclude that

$$u \le 0, \tag{B-48}$$

and consequently (M, g) satisfies the area enlarging condition (1-11). In fact, if we apply the strong maximum principle, we may see that

$$u < 0, \tag{B-49}$$

unless we are dealing with flat space. Since α vanishes identically, we see that (M, g) is also strongly asymptotically flat.

Acknowledgments

I would like to thank my advisor, Marcus Khuri, for his guidance and for proposing this problem. I'm grateful to Ye Sle Cha, Piotr Chruściel, and Luc Nguyen for their interest in the final result. I'm grateful to Christina Sormani for the support she gives young mathematicians and for the workshops she organizes (funded by DMS 1309360). I'm grateful to Brian Allen, Lisa Hernandez, Sajjad Lakzian, and Dan

Lee for the discussions we had at those workshops. I'm also grateful to Lan-Hsuan Huang for a very helpful conversation we had.

References

[Allen 2018] B. Allen, "IMCF and the stability of the PMT and RPI under L^2 convergence", *Ann. Henri Poincaré* **19**:4 (2018), 1283–1306. MR Zbl

[Bray and Finster 2002] H. Bray and F. Finster, "Curvature estimates and the positive mass theorem", *Comm. Anal. Geom.* **10**:2 (2002), 291–306. MR Zbl

[Brill 1959] D. R. Brill, "On the positive definite mass of the Bondi–Weber–Wheeler time-symmetric gravitational waves", *Ann. Physics* **7** (1959), 466–483. MR

[Chruściel 2008] P. T. Chruściel, "Mass and angular-momentum inequalities for axi-symmetric initial data sets, I: Positivity of mass", *Ann. Physics* **323**:10 (2008), 2566–2590. MR Zbl

[Chruściel and Nguyen 2011] P. T. Chruściel and L. Nguyen, "A lower bound for the mass of axisymmetric connected black hole data sets", *Classical Quantum Gravity* **28**:12 (2011), art. id. 125001, 19 pp. MR Zbl

[Corvino 2005] J. Corvino, "A note on asymptotically flat metrics on $\mathbb{R}^3$ which are scalar-flat and admit minimal spheres", *Proc. Amer. Math. Soc.* **133**:12 (2005), 3669–3678. MR Zbl

[Evans and Gariepy 2015] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Revised ed., CRC Press, Boca Raton, FL, 2015. MR Zbl

[Finster and Kath 2002] F. Finster and I. Kath, "Curvature estimates in asymptotically flat manifolds of positive scalar curvature", *Comm. Anal. Geom.* **10**:5 (2002), 1017–1031. MR Zbl

[Folland 1999] G. B. Folland, *Real analysis: modern techniques and their applications*, 2nd ed., Wiley, New York, 1999. MR Zbl

[Gibbons and Holzegel 2006] G. W. Gibbons and G. Holzegel, "The positive mass and isoperimetric inequalities for axisymmetric black holes in four and five dimensions", *Classical Quantum Gravity* **23**:22 (2006), 6459–6478. MR Zbl

[Gilbarg and Trudinger 1998] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der mathematischen Wissenschaften **224**, Springer, 1998. MR Zbl

[Huang et al. 2017] L.-H. Huang, D. A. Lee, and C. Sormani, "Intrinsic flat stability of the positive mass theorem for graphical hypersurfaces of Euclidean space", *J. Reine Angew. Math.* **727** (2017), 269–299. MR Zbl

[Khuri et al. 2019] M. Khuri, B. Sokolowsky, and G. Weinstein, "A Penrose-type inequality with angular momentum and charge for axisymmetric initial data", *Gen. Relativity Gravitation* **51**:9 (2019), art. id. 118, 23 pp. MR Zbl

[Lee 2009] D. A. Lee, "On the near-equality case of the positive mass theorem", *Duke Math. J.* **148**:1 (2009), 63–80. MR Zbl

[Lee and Sormani 2014] D. A. Lee and C. Sormani, "Stability of the positive mass theorem for rotationally symmetric Riemannian manifolds", *J. Reine Angew. Math.* **686** (2014), 187–220. MR Zbl

[LeFloch and Sormani 2015] P. G. LeFloch and C. Sormani, "The nonlinear stability of rotationally symmetric spaces with low regularity", *J. Funct. Anal.* **268**:7 (2015), 2005–2065. MR Zbl

[Schoen and Yau 1979] R. Schoen and S. T. Yau, "On the proof of the positive mass conjecture in general relativity", *Comm. Math. Phys.* **65**:1 (1979), 45–76. MR Zbl

[Sormani and Stavrov Allen 2019] C. Sormani and I. Stavrov Allen, "Geometrostatic manifolds of small ADM mass", *Comm. Pure Appl. Math.* **72**:6 (2019), 1243–1287. MR Zbl

[Witten 1981] E. Witten, "A new proof of the positive energy theorem", *Comm. Math. Phys.* **80**:3 (1981), 381–402. MR Zbl

Received June 20, 2018. Revised June 28, 2019.

EDWARD T. BRYDEN
DEPARTMENT OF MATHEMATICS
STONY BROOK UNIVERSITY
STONY BROOK, NY
UNITED STATES
Current address:
TUEBINGEN UNIVERSITY
TUEBINGEN
GERMANY
ebryden@math.sunysb.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.153

INDEX ESTIMATES FOR FREE BOUNDARY CONSTANT MEAN CURVATURE SURFACES

MARCOS P. CAVALCANTE AND DARLAN F. DE OLIVEIRA

We consider compact constant mean curvature surfaces with boundary immersed in a mean convex region of the Euclidean space or in the unit sphere. We prove that the weak Morse index is bounded from below by a linear function of the genus and number of boundary components.

1. Introduction

Let W be a Riemannian manifold with nonempty boundary such that its boundary ∂W is a union of smooth hypersurfaces. Let $M \subset W$ be a compact constant mean curvature hypersurface such that M intersects the regular part of ∂W along its boundary in a right angle. It is well known that such hypersurfaces are critical points of the area functional for variations of M that preserve the enclosed volume and keep the boundary freely on ∂W. We recall that the variations allowed of M are variations $\phi : (-\varepsilon, \varepsilon) \times M \to W$ whose immersions $\phi_t : M \to W$ satisfy $\phi_t(\text{int } M) \subset \text{int } W$ and $\phi_t(\partial M) \subset \text{int } \partial W$ for all $t \in (-\varepsilon, \varepsilon)$; see [Ros and Vergasta 1995, Section 1]. These hypersurfaces arise in many geometrical and physical problems and are referred as *free boundary CMC hypersurfaces* (FBCMC hypersurfaces, for short). They have been studied since the 19th century and still form a very active topic in differential geometry. We refer the reader to the books of Finn [1986] and López [2013] for a nice introduction to this subject.

An important problem about FBCMC hypersurfaces is to classify those ones that are *stable*, that is, whose second variation of the area is nonnegative for volume preserving variations. For instance, in the case that W is a geodesic ball in a space form, a well-known conjecture asserts that the totally geodesic ball and the spherical caps are the only solutions. It was confirmed by Ros and Vergasta [1995] and Nunes [2017] in dimension two, and more recently by Wang and Xia [2019] in any dimension and also for capillary hypersurfaces. Other results on stable FBCMC hypersurfaces can be found for instance in [Ainouz and Souam 2016; Athanassenas

Cavalcante was partially supported by CNPq-Brazil (309543/2015-0), CAPES-Brazil (897/18) and FAPEAL-Brazil (14/2016).
MSC2010: 35P15, 49Q10, 53A10.
Keywords: constant mean curvature surfaces, Morse index, free boundary surfaces.

1987; Barbosa 2018; Choe and Koiso 2016; Li and Xiong 2017; 2018; López 2014; Ros 2008; Ros and Souam 1997].

When M is a stable FBCMC surface immersed in a mean convex region $W \subset \mathbb{R}^3$, Ros [2008, Theorem 9] showed that there are just a few possibilities for the genus and the number of boundary components of M (see Corollary 1.2 below).

When M is not stable, there is a Schrödinger operator L associated to the second variation of the area which has nonzero *weak Morse index* (see Section 2 for precise definitions). Geometrically, the weak Morse index of M is the number of directions whose volume preserving variations decrease area. It will be denoted by $\mathrm{Ind}_{\mathrm{w}}(M)$.

The case of free boundary minimal surfaces is of special interest and the works of Fraser and Schoen [2011; 2013; 2016] have motivated much research in this case. For free boundary minimal hypersurfaces all variations keeping the boundary freely in the boundary are allowed, not only volume preserving variations. In this setting Sargent [2017] proved that if M is a free boundary minimal surface immersed in a convex region $W \subset \mathbb{R}^3$ then the index is bounded from below by $(2g+k-1)/3$, where g is the genus of M and k is the number of boundary components. In [Ambrozio et al. 2018], Ambrozio, Carlotto and Sharp proved that it is valid in weakly mean convex domains in $\mathbb{R}^3$. In fact, their results apply more generally to domains in Riemannian manifolds with boundary satisfying suitable curvature conditions. In higher dimensions, they also obtained lower bounds for the index in terms of the dimension of the first relative homology group with real coefficients. The technique presented in these results uses the coordinates of harmonic forms as test functions and is inspired by previous works on eigenvalue estimates and index estimates for minimal hypersurfaces without boundary; see [Ros 2006; Savo 2010; Ambrozio et al. 2018].

Following these lines, in this paper, we obtain lower bounds for the weak Morse index of FBCMC surfaces in weakly mean convex regions of the Euclidean space or the unit sphere. More precisely, our results are the following.

Theorem 1.1. *Let W be a region of $\mathbb{R}^3$ such that its boundary is a union of smooth weakly mean convex surfaces, and let M^2 be a compact, orientable, FBCMC surface immersed in the mean convex side of W and whose boundary intersects the regular part of ∂W. If M has genus g and k boundary components, then*

$$\mathrm{Ind}_{\mathrm{w}}(M) \geq \frac{2g+k-4}{6}.$$

As an immediate consequence we obtain the result of Ros cited above:

Corollary 1.2 [Ros 2008, Theorem 9]. *Under the conditions of Theorem 1.1, if M is stable, then the only possibilities for g and k are*

(1) $g = 0$ *and* $k \leq 4$;

(2) $g = 1$ *and* $k = 1$ *or* 2.

In the case of FBCMC surfaces immersed in weakly mean convex domains of $\mathbb{S}^3$, the result reads as follows:

Theorem 1.3. *Let W be a region in the unit sphere $\mathbb{S}^3$ such that its boundary is a union of smooth weakly mean convex surfaces, and let M^2 be a compact, orientable, FBCMC surface immersed in the mean convex side of W and whose boundary intersects the regular part of ∂W. If M has genus g and k boundary components,*

$$\mathrm{Ind}_{\mathrm{w}}(M) \geq \frac{2g+k-5}{8}.$$

Corollary 1.4. *Under the conditions of Theorem 1.3, if M is stable, then the only possibilities for g and k are*

(1) $g=0$ *and* $k\leq 5$;

(2) $g=1$ *and* $k\leq 3$;

(3) $g=2$ *and* $k=1$.

This paper is organized as follows. In Section 2, we present some definitions and basic results to be used in the proofs. Section 3 is devoted to computing the Jacobi operator of the test functions given by the coordinates of harmonic forms. In Section 4, we present the proof of Theorem 1.1. The proof of Theorem 1.3 is analogous and it is sketched in Section 5.

2. Preliminaries

Let us denote by W a connected domain of the Euclidean space $\mathbb{R}^3$ which is not necessarily compact. For simplicity, let us assume that W has smooth boundary and fix a unit normal vector field ν along each component of ∂W. We recall that the second fundamental form and the mean curvature of ∂W with respect to ν are defined respectively by

$$II^{\partial W}(X,Y) = \langle -D_X\nu, Y\rangle, \qquad \text{for } X, Y \in T\partial W,$$

and

$$H^{\partial W} = \tfrac{1}{2}\,\mathrm{tr}\, II^{\partial W},$$

where D is the Levi-Civita connection in the Euclidean space. Since the boundary ∂W is orientable, at a point on ∂W we have two choices for ν, one pointing inward from W and the other one pointing outward from W. From now on, we fix the vector field ν pointing outward from W. In this case, we say that W is *weakly convex* if $II^{\partial W}$ is nonpositive defined. If $H^{\partial W}\leq 0$, W is said to be *weakly mean convex*.

Let $x : M \to W$ be a compact oriented surfaced with boundary which is properly immersed, that is, $x(M)\cap W = x(\partial M)$. Fixing a unit normal vector field N along x, we denote by A the shape operator associated to the second fundamental form II^M of M with respect to N, namely

$$AX = -D_X N, \qquad \text{for } X \in TM,$$
$$II^M(X, Y) = \langle AX, Y \rangle, \qquad \text{for } X, Y \in TM.$$

We say that M is free boundary if $x(\partial M)$ meets ∂W orthogonally.

From now on, let us assume that W is a mean convex domain of $\mathbb{R}^3$ and M is a *free boundary constant mean curvature surface* properly immersed in W. Such surfaces are critical points of the area functional for normal variations whose variational vector field is given by $X = uN$, where $u \in \mathcal{F}$ and

$$\mathcal{F} = \left\{ u : M \to \mathbb{R} : u \text{ is smooth up to the boundary and } \int_M u \, dM = 0 \right\}.$$

The second variation of area functional is given by the quadratic form $Q : C^\infty(M) \times C^\infty(M) \to \mathbb{R}$ (see [Ros and Vergasta 1995; Ros 2008]),

$$Q(u, u) = \int_M (u \Delta u - \|A\|^2 u^2) \, dM + \int_{\partial M} (u \eta(u) + II^{\partial W}(N, N) u^2) \, ds.$$

Here η is the outward unit conormal vector field on ∂M, that is, the unique unit vector field on ∂M that is tangent to M, normal to ∂M and pointing outward from M. Note that, under our notations, the free boundary condition means that $\eta = \nu$ along ∂M. We point out that in this paper we are using the geometric definition of the Laplacian operator, that is, $\Delta u = \operatorname{div} \nabla u$, where $\operatorname{div} X = -\operatorname{tr} \nabla X$.

The *weak Morse index* of M, denoted by $\operatorname{Ind}_{\mathrm{w}}(M)$, is defined as the maximal dimension of a subspace of $\mathcal{F}$ on which $Q|_{\mathcal{F}}$ is negative definite. Geometrically, the index indicates the number of directions whose variations decrease area. In particular, we say that M is *stable* if the weak Morse index is zero.

We say that $u \in \mathcal{F}$ is an eigenfunction of $Q|_{\mathcal{F}}$ associated to the eigenvalue $\lambda \in \mathbb{R}$ if and only if $Q|_{\mathcal{F}}(u, v) = \lambda \int_M uv \, dM$ for all $v \in \mathcal{F}$. This is equivalent to saying that u solves the following eigenvalue problem:

(2-1)
$$\begin{cases} Lu = \lambda u & \text{in } M, \\ \dfrac{\partial u}{\partial \eta} = -II^{\partial W}(N, N) u & \text{on } \partial M, \end{cases}$$

where $L : \mathcal{F} \to \mathcal{F}$ is given by

(2-2)
$$Lu = Ju - \frac{1}{\operatorname{vol}(M)} \int_M Ju \, dM$$

and $J = \Delta - \|A\|^2$ is the Jacobi operator. We conclude that the weak Morse index coincides with the number of negative eigenvalues of the boundary problem (2-1).

Also, it is well known that such eigenvalues are given in a nondecreasing sequence $\lambda_1^L \le \lambda_2^L \le \cdots \le \lambda_k^L \le \cdots \nearrow \infty$ associated to a $L^2(M)$-orthonormal basis,

$$\{\phi_1, \phi_2, \ldots, \phi_k, \ldots\} \quad \text{of} \quad L^2(M) \cap \mathcal{F}$$

of solutions of the eigenvalue problem (2-1), satisfying the min-max characterization

$$\lambda_k^L = \min_{u \in \mathcal{J}^{k-1} \setminus \{0\}} \frac{Q|_{\mathcal{F}}(u, u)}{\int_M u^2 dM},$$

where $\mathcal{J}^{k-1} = \langle \phi_1, \ldots, \phi_{k-1} \rangle^{\perp}$.

In order to give lower bounds for the weak Morse index of M in terms of its topological invariants we will construct admissible eigenfunctions in $\mathcal{F}$ using harmonic vector fields, or equivalently harmonic 1-forms. Let us denote by $i : \partial M \to M$ the inclusion map. We also set $\mathcal{H}_T^1(M)$ the space of closed and coclosed 1-forms that are tangential at ∂M, that is,

$$\mathcal{H}_T^1(M) := \{w \in \Omega^1(M), dw = 0, \delta w = 0 \text{ and } i_\eta w = 0 \text{ along } \partial M\}.$$

Here d is the exterior derivative operator and δ is the interior derivative operator defined by $\delta = -\star d \star$, where $\star : \Omega^1(M) \to \Omega^1(M)$ is the Hodge star operator.

It is important to note that $\mathcal{H}_T^1(M)$ coincides with the space of harmonic 1-forms satisfying the *absolute boundary conditions*, that is

$$\mathcal{H}_T^1(M) = \{w \in \Omega^1(M), \Delta w = 0, i_\eta w = 0 \text{ and } i_\eta dw = 0 \text{ along } \partial M\}.$$

This space is closed related to the topology of the underline manifold. In fact we have the following result; see [Ambrozio et al. 2018] or [Sargent 2017].

Lemma 2.1. *Let M^2 be a compact, orientable surface with nonempty boundary ∂M. If M has genus g and $k \ge 1$ boundary components, then*

$$\dim \mathcal{H}_T^1(M) = 2g + k - 1.$$

3. Test functions and harmonic vector fields

Denoting by $\mathcal{E} = \{\bar{E}_1, \bar{E}_2, \bar{E}_3\}$ the canonical basis in $\mathbb{R}^3$ we will consider $E_i := \bar{E}_i - \langle \bar{E}_i, N \rangle N$, the orthogonal projection of $\bar{E}_i$ on TM. We also consider the smooth support functions $g_i : M \to \mathbb{R}$, $g_i := \langle \bar{E}_i, N \rangle$, for $1 \le i \le 3$.

Given a 1-form ω on M we denote by ξ its dual vector field, that is, $\xi^\flat = \omega$. Abusing notation slightly, we denote by $\star\xi$ the vector field dual of $\star\omega$. In the following, we will use the coordinates of ξ and $\star\xi$ as test functions. Namely, for each $1 \le i \le 3$, we define $w_i, \bar{w}_i : M \to \mathbb{R}$ as

$$w_i := \omega(E_i) = \langle E_i, \xi \rangle \quad \text{and} \quad \bar{w}_i := \star\omega(E_i) = \langle E_i, \star\xi \rangle.$$

To compute the Jacobi operator of w_i and $\bar{w}_i$ we need the following lemma of local nature proved in [Cavalcante and de Oliveira 2020]; see also [Ros 2007].

Lemma 3.1. *Let M^2 be an orientable CMC surface in $\mathbb{R}^3$. Then, using the above notation we have*

$$\Delta w_i = (\|A\|^2 - 4H^2)w_i + 2H\langle AE_i, \xi\rangle - 2g_i\langle A, \nabla\xi\rangle + \langle E_i, \Delta\xi\rangle,$$
$$\Delta \overline{w}_i = (\|A\|^2 - 4H^2)\overline{w}_i + 2H\langle AE_i, \star\xi\rangle - 2g_i\langle A, \nabla\star\xi\rangle + \langle E_i, \Delta\star\xi\rangle,$$

for $1 \le i \le 3$.

Now we note that when the vector field ξ is harmonic and tangential along ∂M its coordinates are admissible functions to compute the weak Morse index of CMC surfaces. More precisely we have:

Lemma 3.2. *If $\xi \in TM$ is a harmonic vector field which is tangential in ∂M, then $w_i \in \mathcal{F}$, that is,*

$$\int_M w_i \, dM = 0,$$

for $1 \le i \le 3$.

Proof. Note that $E_i = \nabla x_i$, $1 \le i \le 3$, where $x = (x_1, x_2, x_3) : M \to W$ is the immersion map. Then we have

$$\int_M w_i \, dM = \int_M \langle \nabla x_i, \xi\rangle \, dM$$
$$= \int_M x_i \operatorname{div}\xi \, dM + \int_{\partial M} x_i \langle \xi, \eta\rangle \, ds = 0.$$

In fact, $\operatorname{div}\xi = 0$ since ξ is harmonic, and $\langle \xi, \eta\rangle = 0$ since ξ tangential to ∂M. □

Remark 3.3. In general the functions $\overline{w}_i$, $1 \le i \le 3$, do not have mean value zero. However, we will see in Section 4 that if $\dim \mathcal{H}^1_T(M)$ is large enough then we can choose ξ such that

$$\int_M \overline{w}_i \, dM = 0, \qquad \text{for } 1 \le i \le 3.$$

We conclude this section by computing the boundary term of the quadratic form Q on w_i and $\overline{w}_i$,.

Lemma 3.4. *If $\xi \in TM$ is a vector field such that its dual 1-form satisfies the absolute boundary condition, then*

$$\sum_i \int_{\partial M} (w_i \eta(w_i) + II^{\partial W}(N, N) w_i^2) \, ds = 2\int_{\partial M} H^{\partial W} \|\xi\|^2 \, ds, \tag{3-1}$$

$$\sum_i \int_{\partial M} (\overline{w}_i \eta(\overline{w}_i) + II^{\partial W}(N, N) \overline{w}_i^2) \, ds = 2\int_{\partial M} H^{\partial W} \|\xi\|^2 \, ds. \tag{3-2}$$

Proof. We first note that for any vector field $X \in TM$ we have

$$\begin{aligned}\langle \nabla_\eta E_i, X\rangle &= \eta\langle \bar{E}_i, X\rangle - \langle \bar{E}_i, \nabla_\eta X\rangle \\ &= \langle \bar{E}_i, D_\eta X - \nabla_\eta X\rangle \\ &= \langle \bar{E}_i, N\rangle\langle X, A\eta\rangle.\end{aligned}$$

Let ω be the dual 1-form of the vector field ξ. Since $i_\eta d\omega = 0$ we have

$$0 = d\omega(\eta, \xi) = \langle \nabla_\eta \xi, \xi\rangle - \langle \nabla_\xi \xi, \eta\rangle.$$

Thus,

$$\begin{aligned}\sum_i \int_{\partial M} w_i \eta(w_i)\, ds &= \sum_i \int_{\partial M} w_i(\langle \nabla_\eta \xi, E_i\rangle + \langle \bar{E}_i, N\rangle\langle \xi, A\eta\rangle)\, ds \\ &= \int_{\partial M} \langle \nabla_\eta \xi, \xi\rangle\, ds = \int_{\partial M} \langle \nabla_\xi \xi, \eta\rangle\, ds \\ &= -\int_{\partial M} \langle \nabla_\xi \eta, \xi\rangle\, ds = \int_{\partial M} II^{\partial W}(\xi, \xi)\, ds.\end{aligned}$$

Since ξ and N form an orthogonal basis of the tangent space of ∂W along ∂M we conclude the proof by noting that

$$II^{\partial W}(\xi, \xi) + II^{\partial W}(N, N)\|\xi\|^2 = 2H^{\partial W}\|\xi\|^2.$$

The proof of assertion (3-2) follows the same steps as above, noting additionally that the Levi-Civita connection ∇ commutes with the Hodge star operator $\star$. □

4. Proof of Theorem 1.1

Proof. The proofs follow the same spirit as our proofs in [Cavalcante and de Oliveira 2020] but take into account the boundary term. Let $\xi_1, \xi_2, \dots, \xi_m$, be the first m eigenfunctions of the Hodge Laplacian Δ on M, which satisfy the absolute boundary condition [Gilkey et al. 1999, Theorem 1.5.4]. Set $\mathcal{L}_m^\Delta = \operatorname{span}\{\xi_1, \dots, \xi_m\}$ the vector space generated by these functions. By Lemma 2.1, we know that $\dim \mathcal{H}_T^1(M) = 2g + k - 1$. Let us assume that $m \geq 2g - k - 1$, and so $\mathcal{H}_T^1(M)$ is a subspace of $\mathcal{L}_m^\Delta$.

Next, we choose an orthonormal basis of $L^2(M)$ given by eigenfunctions of the operator L defined in (2-2), say $\{\phi_1, \phi_2, \dots, \phi_k, \dots\}$. We denote by $\mathcal{J}^n := \langle \phi_1, \dots, \phi_n\rangle^\perp$ the linear subspace of $\mathcal{F}$ orthogonal to the first n eigenfunctions of L.

Initially, we look for harmonic forms $\xi \in \mathcal{L}_m^\Delta$ such that the functions w_i, $\bar{w}_i \in \mathcal{J}^{\alpha-1}$, for some $\alpha \in \mathbb{N}$ and $i \in \{1, 2, 3\}$. It is equivalent to find a solution to the following system with $6(\alpha - 1)$ homogenous linear equations in the variable ξ:

$$\int_M w_i \phi_k\, dM = \int_M \bar{w}_i \phi_k\, dM = 0, \tag{4-1}$$

where $1 \leq i \leq 3$ and $1 \leq k \leq \alpha - 1$. In particular, if $m(\alpha) := \dim\ \mathcal{L}_m^\Delta > 6(\alpha - 1)$, then the system (4-1) has at least one nontrivial solution $\xi \in \mathcal{L}_m^{\Delta}$ such that $w_i, \overline{w}_i \in \mathcal{J}^{\alpha-1}$ for all $1 \le i \le 3$. By min-max characterization we have

$$\lambda_\alpha^J \int_M w_i^2 \, dM \le Q(w_i, w_i) \quad \text{and} \quad \lambda_\alpha^J \int_M \overline{w}_i^2 \, dM \le Q(\overline{w}_i, \overline{w}_i).$$

Now, using Lemma 3.1 we get

$$\begin{aligned} \lambda_\alpha^J \int_M w_i^2 \, dM \le -4H^2 \int_M w_i^2 \, dM + 2H \int_M \langle E_i, A\xi \rangle w_i \, dM \\ + \int_M \langle E_i, \Delta\xi \rangle w_i \, dM - 2 \int_M g_i \langle A, \nabla\xi \rangle w_i \, dM \\ + \int_{\partial M} (w_i \eta(w_i) + II^{\partial W(N,N)} w_i^2) \, ds. \end{aligned}$$

Summing up $i = 1, 2, 3$ and using Lemma 3.4 we obtain

$$\begin{aligned} \lambda_\alpha^J \int_M \|\xi\|^2 \, dM \le -4H^2 \int_M \|\xi\|^2 \, dM + 2H \int_M \langle A\xi, \xi \rangle \, dM \\ + \int_M \langle \Delta\xi, \xi \rangle \, dM + 2 \int_{\partial M} H^{\partial W} \|\xi\|^2 \, dM. \end{aligned}$$

Applying the same arguments to the test functions $\overline{w}_i$ we get

$$\begin{aligned} \lambda_\alpha^J \int_M \|\xi\|^2 \, dM \le -4H^2 \int_M \|\xi\|^2 \, dM + 2H \int_M \langle A\star\xi, \star\xi \rangle \, dM \\ + \int_M \langle \Delta\star\xi, \star\xi \rangle \, dM + 2 \int_{\partial M} H^{\partial W} \|\xi\|^2 \, ds. \end{aligned}$$

Then, summing these last two inequalities and noting that $\langle A\xi, \xi \rangle + \langle A\star\xi, \star\xi \rangle = 2H\|\xi\|^2$, we have

$$\begin{aligned} \text{(4-2)} \quad \lambda_\alpha^J \int_M \|\xi\|^2 dM \le +2 \int_{\partial M} H^{\partial W} \|\xi\|^2 dM - 2H^2 \int_M \|\xi\|^2 dM \\ + \tfrac{1}{2} \int_M (\langle \Delta\xi, \xi \rangle + \langle \Delta\star\xi, \star\xi \rangle) dM. \end{aligned}$$

Finally, if $\xi \in \mathcal{L}_m^{\Delta}$ we get $\xi = \sum_i \alpha_i \xi_i$ and therefore

$$\text{(4-3)} \qquad \int_M \langle \Delta\star\xi, \star\xi \rangle dM = \int_M \langle \Delta\xi, \xi \rangle dM = \lambda_{m(\alpha)} \int_M \|\xi\|^2 dM.$$

Substituting (4-3) into (4-2) and using the fact that $H^{\partial W} \le 0$ we obtain

$$\lambda_\alpha^J \le -2H^2 + \lambda_{m(\alpha)}^{\Delta},$$

where $m(\alpha) > 6(\alpha - 1)$. This concludes the first part of Theorem 1.1.

In order to get the lower bound for the weak Morse index of M we take $\mathcal{J}^{\alpha-1} := \langle \phi_1, \dots, \phi_{\alpha-1} \rangle$, where $\phi_1, \dots, \phi_{\alpha-1}$ are the first eigenfunctions of the eigenvalue equation (2-1). From Lemma 3.2 we know that if $\xi \in \mathcal{H}^1_T(M)$, then the test functions w_1, w_2 and w_3, belong to $\mathcal{F}$.

We look for vector fields $\xi \in \mathcal{H}^1_T(M)$ such that for $1 \le i \le 3$, the test functions $w_i, \bar{w}_i$ are in $\mathcal{J}^{\alpha-1}$, for some $\alpha \in \mathbb{N}$, and $\bar{w}_i$ is in $\mathcal{F}$. In this case, we have the following system with $6\alpha - 3$ homogeneous linear equations in the variable ξ:

$$\int_M \bar{w}_i = \int_M w_i \phi_k = \int_M \bar{w}_i \phi_k = 0, \tag{4-4}$$

where $1 \le i \le 3$ and $1 \le k \le \alpha - 1$.

If $\dim \mathcal{H}^1_T(M) = 2g + k - 1 > 6\alpha - 3$, then the system (4-4) has at least one nontrivial solution $\xi \in \mathcal{H}^1_T(M)$. Following the same steps as above we get

$$\lambda^L_\alpha \int_M \|\xi\|^2 \le -2H^2 \int_M \|\xi\|^2.$$

This implies that $\lambda^L_\alpha < 0$ and then $\mathrm{Ind}_{\mathrm{w}}(M) \ge \alpha$. Since α can be chosen as the largest integer such that $2g + k - 1 > 6\alpha - 3$ we get

$$\mathrm{Ind}_{\mathrm{w}}(M) \ge \frac{2g + k - 4}{6}.$$

□

5. Proof of Theorem 1.3

Proof. Composing the immersion $x : M \to W \subset \mathbb{S}^3$ with the canonical immersion of the unit sphere into the Euclidean space, we may consider $x : M \to \mathbb{R}^4$. Let $\mathcal{E} = \{\bar{E}_1, \bar{E}_2, \bar{E}_3, \bar{E}_4\}$ be the canonical basis in $\mathbb{R}^4$ and $E_i := \bar{E}_i - \langle \bar{E}_i, N \rangle N - \langle \bar{E}_i, x \rangle x$, the orthogonal projections of $\bar{E}_i$ on TM. Choosing $\nu = -x$ as an orientation of $\mathbb{S}^3$ we have

$$D_Y X - \nabla_Y X = \langle AX, Y \rangle N + \langle X, Y \rangle \nu, \qquad X, Y \in TM,$$

and also

$$\begin{aligned} \langle \nabla_X E_i, Y \rangle = X \langle \bar{E}_i, Y \rangle - \langle \bar{E}_i, \nabla_X Y \rangle &= \langle \bar{E}_i, D_Y X - \nabla_Y X \rangle \\ &= \langle AX, Y \rangle \langle \bar{E}_i, N \rangle + \langle X, Y \rangle \langle \bar{E}_i, \nu \rangle. \end{aligned}$$

Using that $\langle \xi, N \rangle = \langle \star\xi, N \rangle = \langle \xi, \nu \rangle = \langle \star\xi, \nu \rangle = 0$ we get

$$\begin{aligned} \sum_i \int_{\partial M} w_i \eta(w_i)\, ds &= \sum_i \int_{\partial M} w_i (\langle \nabla_\eta \xi, E_i \rangle) ds \\ &\quad + \sum_i \int_{\partial M} w_i (\langle \bar{E}_i, N \rangle \langle \xi, A\eta \rangle + \langle \eta, \xi \rangle \langle \bar{E}_i, \nu \rangle)\, ds \\ &= \int_{\partial M} \langle \nabla_\eta \xi, \xi \rangle\, ds + \int_{\partial M} \langle \xi, N \rangle \langle \xi, A\eta \rangle + \int_{\partial M} \langle \eta, \xi \rangle \langle \xi, \nu \rangle \\ &= \int_{\partial M} \langle \nabla_\eta \xi, \xi \rangle\, ds, \end{aligned}$$

and analogously

$$\sum_i \int_{\partial M} \bar{w}_i \eta(\bar{w}_i)\, ds = \int_{\partial M} \langle \nabla_\eta \star\xi, \star\xi\rangle\, ds = \int_{\partial M} \langle \nabla_\eta \xi, \xi\rangle\, ds.$$

So, Lemma 3.4 holds in the spherical case. The Laplacian of the test functions w_i and $\bar{w}_i$ are given by (see [Cavalcante and de Oliveira 2020])

$$\Delta w_i = (\|A\|^2 - 4H^2) w_i + 2H\langle AE_i, \xi\rangle - 2g_i\langle A, \nabla\xi\rangle + \langle E_i, \Delta\xi\rangle - 2\langle x, \bar{E}_i\rangle \operatorname{div}\xi,$$
$$\Delta \bar{w}_i = (\|A\|^2 - 4H^2) \bar{w}_i + 2H\langle AE_i, \bar{\xi}\rangle - 2g_i\langle A, \nabla\bar{\xi}\rangle + \langle E_i, \Delta\bar{\xi}\rangle - 2\langle x, \bar{E}_i\rangle \operatorname{div}\bar{\xi}.$$

Under these considerations, and taking into account that the Jacobi operator for immersions into $\mathbb{S}^3$ is given by $J = \Delta - (\|A\|^2 + 2)$ the proof follows as in the proof of Theorem 1.1. □

Acknowledgments

The authors would like to thank Lucas Ambrozio and Ivaldo Nunes for useful comments, and Fernando Codá Marques for illuminating conversations about this work.

References

[Ainouz and Souam 2016] A. Ainouz and R. Souam, "Stable capillary hypersurfaces in a half-space or a slab", *Indiana Univ. Math. J.* **65**:3 (2016), 813–831. MR Zbl

[Ambrozio et al. 2018] L. Ambrozio, A. Carlotto, and B. Sharp, "Index estimates for free boundary minimal hypersurfaces", *Math. Ann.* **370**:3-4 (2018), 1063–1078. MR Zbl

[Athanassenas 1987] M. Athanassenas, "A variational problem for constant mean curvature surfaces with free boundary", *J. Reine Angew. Math.* **377** (1987), 97–107. MR Zbl

[Barbosa 2018] E. Barbosa, "On CMC free-boundary stable hypersurfaces in a Euclidean ball", *Math. Ann.* **372**:1-2 (2018), 179–187. MR Zbl

[Cavalcante and de Oliveira 2020] M. P. Cavalcante and D. F. de Oliveira, "Lower bounds for the index of compact constant mean curvature surfaces in $\mathbb{R}^3$ and $\mathbb{S}^3$", *Rev. Mat. Iberoam.* **36**:1 (2020), 195–206. MR

[Choe and Koiso 2016] J. Choe and M. Koiso, "Stable capillary hypersurfaces in a wedge", *Pacific J. Math.* **280**:1 (2016), 1–15. MR Zbl

[Finn 1986] R. Finn, *Equilibrium capillary surfaces*, Grundlehren der Mathematischen Wissenschaften **284**, Springer, 1986. MR Zbl

[Fraser and Schoen 2011] A. Fraser and R. Schoen, "The first Steklov eigenvalue, conformal geometry, and minimal surfaces", *Adv. Math.* **226**:5 (2011), 4011–4030. MR Zbl

[Fraser and Schoen 2013] A. Fraser and R. Schoen, "Minimal surfaces and eigenvalue problems", pp. 105–121 in *Geometric analysis, mathematical relativity, and nonlinear partial differential equations*, edited by M. Ghomi et al., Contemp. Math. **599**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Fraser and Schoen 2016] A. Fraser and R. Schoen, "Sharp eigenvalue bounds and minimal surfaces in the ball", *Invent. Math.* **203**:3 (2016), 823–890. MR Zbl

[Gilkey et al. 1999] P. B. Gilkey, J. V. Leahy, and J. Park, *Spectral geometry, Riemannian submersions, and the Gromov–Lawson conjecture*, Chapman & Hall, Boca Raton, FL, 1999. MR Zbl

[Li and Xiong 2017] H. Li and C. Xiong, "Stability of capillary hypersurfaces with planar boundaries", *J. Geom. Anal.* **27**:1 (2017), 79–94. MR Zbl

[Li and Xiong 2018] H. Li and C. Xiong, "Stability of capillary hypersurfaces in a Euclidean ball", *Pacific J. Math.* **297**:1 (2018), 131–146. MR Zbl

[López 2013] R. López, *Constant mean curvature surfaces with boundary*, Springer, 2013. MR Zbl

[López 2014] R. López, "Capillary surfaces with free boundary in a wedge", *Adv. Math.* **262** (2014), 476–483. MR Zbl

[Nunes 2017] I. Nunes, "On stable constant mean curvature surfaces with free boundary", *Math. Z.* **287**:1-2 (2017), 473–479. MR Zbl

[Ros 2006] A. Ros, "One-sided complete stable minimal surfaces", *J. Differential Geom.* **74**:1 (2006), 69–92. MR Zbl

[Ros 2007] A. Ros, "Stable periodic constant mean curvature surfaces and mesoscopic phase separation", *Interfaces Free Bound.* **9**:3 (2007), 355–365. MR Zbl

[Ros 2008] A. Ros, "Stability of minimal and constant mean curvature surfaces with free boundary", *Mat. Contemp.* **35** (2008), 221–240. MR Zbl

[Ros and Souam 1997] A. Ros and R. Souam, "On stability of capillary surfaces in a ball", *Pacific J. Math.* **178**:2 (1997), 345–361. MR Zbl

[Ros and Vergasta 1995] A. Ros and E. Vergasta, "Stability for hypersurfaces of constant mean curvature with free boundary", *Geom. Dedicata* **56**:1 (1995), 19–33. MR Zbl

[Sargent 2017] P. Sargent, "Index bounds for free boundary minimal surfaces of convex bodies", *Proc. Amer. Math. Soc.* **145**:6 (2017), 2467–2480. MR Zbl

[Savo 2010] A. Savo, "Index bounds for minimal hypersurfaces of the sphere", *Indiana Univ. Math. J.* **59**:3 (2010), 823–837. MR Zbl

[Wang and Xia 2019] G. Wang and C. Xia, "Uniqueness of stable capillary hypersurfaces in a ball", *Math. Ann.* **374**:3-4 (2019), 1845–1882. MR Zbl

Received March 25, 2018. Revised March 25, 2019.

MARCOS P. CAVALCANTE
UNIVERSIDADE FEDERAL DE ALAGOAS
MACEIÓ, AL
BRAZIL
marcos@pos.mat.ufal.br

DARLAN F. DE OLIVEIRA
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
FEIRA DE SANTANTA, BA
BRAZIL
darlanfdeoliveira@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.165

A CRITERION FOR MODULES OVER GORENSTEIN LOCAL RINGS TO HAVE RATIONAL POINCARÉ SERIES

ANJAN GUPTA

We prove that modules over an Artinian Gorenstein local ring R have rational Poincaré series sharing a common denominator if $R/\mathrm{socle}(R)$ is a Golod ring. If R is a Gorenstein local ring with square of the maximal ideal being generated by at most two elements, we show that modules over R have rational Poincaré series sharing a common denominator. By a result of Şega, it follows that R satisfies the Auslander–Reiten conjecture. We provide a different proof of a result of Rossi and Şega (*Adv. Math.* 259 (2014), 421–447) concerning rationality of Poincaré series of modules over compressed Gorenstein local rings. We also give a new proof of the fact that modules over Gorenstein local rings of codepth at most 3 have rational Poincaré series sharing a common denominator, which is originally due to Avramov, Kustin and Miller (*J. Algebra* 118:1 (1988), 162–204).

1. Introduction

Let R be a commutative Noetherian local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. Let M be a finitely generated module over R. The Poincaré series of M over R is a formal power series in $\mathbb{Z}[|t|]$ defined as

$$P_M^R(t) = \sum_{i\geq 0} \beta_i^R(M)t^i \in \mathbb{Z}[|t|],$$

where $\beta_i^R(M) = \dim_k \mathrm{Tor}_i^R(M, k)$ denotes the i-th Betti number of M. We say that a formal power series $P(t) \in \mathbb{Z}[|t|]$ is a rational function if there exists a polynomial $g(t) \in \mathbb{Z}[t]$ such that $g(t)P(t)$ is a polynomial in $\mathbb{Z}[t]$. An example due to Anick [1982] shows that the Poincaré series $P_k^R(t)$ is not a rational function in general. Bøgvad [1983] observed that $P_k^R(t)$ may not be a rational function even if R is a Gorenstein ring.

Following Roos [2005, Definition 2.1], we say that a ring R is *good* if there exists a polynomial $d_R(t) \in \mathbb{Z}[t]$ such that $d_R(t)P_M^R(t) \in \mathbb{Z}[t]$ for every finitely

MSC2010: 13D02, 13D40, 13H10.
Keywords: Poincaré series, Golod rings, DG algebras, fibre products, connected sums.

generated R-module M and *bad* otherwise. Roos proved that bad rings exist [2005, Theorem 2.4]. Nevertheless there are an abundance of good rings, e.g., regular local rings, local complete intersections; see [Gulliksen 1974, Corollary 4.2]. We refer to [Avramov et al. 1988; 1994] for more examples of good rings and a detailed account of applications of rationality of Poincaré series.

We use $\mu(-)$ to denote the minimal number of generators. The embedding dimension of R $(= \dim_k \mathfrak{m}/\mathfrak{m}^2)$ is denoted by $\operatorname{edim}(R)$. Let $\hat{R}$ denote the $\mathfrak{m}$-adic completion of R. By Cohen's structure theorem, there is a regular local ring Q with maximal ideal $\mathfrak{n}$ and a surjective ring homomorphism $\eta : Q \to \hat{R}$ such that $\ker \eta = I \subset \mathfrak{n}^2$. The map η is called a minimal Cohen presentation of R. The Loewy length of R is defined as $\operatorname{ll}(R) = \max\{i : \mathfrak{m}^i \neq 0\}$ if R is Artinian and infinity otherwise.

We recall a few more examples of good rings collected from existing literature. Precise references are given with each of the examples.

Examples. Let R be a Gorenstein local ring such that $\operatorname{edim}(R) = n \geq 2$, $\operatorname{ll}(R) = s$ and $\mu(I) = r$. If R satisfies one of the conditions (1)–(3) below, then $P_k^R(t) = (1+t)^n/d_R(t)$ for some polynomial $d_R(t) \in \mathbb{Z}[t]$ and $d_R(t)P_M^R(t) \in \mathbb{Z}[t]$ for every finitely generated R-module M.

(1) R is a compressed Artinian Gorenstein ring (see Definition 4.3) and $s \geq 2$, $s \neq 3$. If $\eta : Q \to R$ is a minimal Cohen presentation of R, then the polynomial $d_R(t)$ is given by $1 - t(P_R^Q(t) - 1) + t^{n+1}(1+t)$; see [Rossi and Şega 2014, Theorem 5.1].

(2) R is an Artinian Gorenstein ring and $\mu(\mathfrak{m}^2) = 1$. The polynomial $d_R(t)$ is given by $1 - nt + t^2$; see [Sally 1980, Theorem 2] and [Croll et al. 2018, Theorem 5.4]

(3) R is not a complete intersection and $\operatorname{codepth}(R) = \operatorname{edim}(R) - \operatorname{depth}(R) \leq 3$. The polynomial $d_R(t)$ is equal to $1 - rt^2 - rt^3 + t^5$; see [Wiebe 1969, Satz 9] and [Avramov et al. 1988, Theorem 6.4].

The main objective of the present article is to give a criterion for Gorenstein local rings to be good, which provides a common method to prove the good property in each of the above examples. As a new application we show that if R is an Artinian Gorenstein ring and $\mu(\mathfrak{m}^2) = 2$, then R is a good ring.

We recall a few definitions. Let $\phi : R \to S$ be a surjective homomorphism of local rings and k be the common residue field of R and S. From the standard change of rings spectral sequence of Tor, Serre proved the following term-wise inequality of power series:

$$P_k^S(t) \prec \frac{P_k^R(t)}{1 - t(P_S^R(t) - 1)}.$$

The homomorphism ϕ is called a Golod homomorphism if the above inequality is an equality. The most widespread method to show that a ring R is good is to use a result of Levin (Theorem 2.2) which states that a ring R is good if there is a surjective Golod homomorphism from a complete intersection onto R.

Let $\eta : Q \twoheadrightarrow \hat{R}$ be a minimal Cohen presentation of R. We say that R is a Golod ring if η is a Golod homomorphism. Let $\operatorname{edim}(R) = n$ and K^R denote the Koszul complex of R on a minimal set of generators of maximal ideal $\mathfrak{m}$. It follows that R is a Golod ring whenever one has

$$P_k^R(t) = \frac{(1+t)^n}{1 - \sum_{i=1}^n \dim_k \mathrm{H}_i(K^R) t^{i+1}}.$$

We refer to [Avramov 1998, §3] for more details on Golod rings and Golod homomorphisms. The main result of the present article is the following:

Theorem I. *Let R be an Artinian Gorenstein local ring of embedding dimension $n \geq 2$ such that $R/\operatorname{socle}(R)$ is a Golod ring. Let $\eta : Q \to R$ be a minimal Cohen presentation, $\mathfrak{n}$ denote the maximal ideal of Q and $I = \ker(\eta) \subset \mathfrak{n}^2$. Then the following hold.*

(1) *For any $f \in I \setminus \mathfrak{n}I$, the induced map $Q/(f) \twoheadrightarrow R$ is a Golod homomorphism.*

(2) *Let $d_R(t) = 1 - t(P_R^Q(t) - 1) + t^{n+1}(1+t)$. Then for any R-module M we have $d_R(t) P_M^R(t) \in \mathbb{Z}[t]$.*

It is worth noting that if R is an Artinian Gorenstein ring, $\operatorname{edim}(R) \geq 2$ and $R/\operatorname{socle}(R)$ is a Golod ring, then with the notation used in the above theorem,

$$P_k^R(t) = \frac{(1+t)^n}{d_R(t)}$$

by a result of Rossi and Şega [2014, Proposition 6.2]. Therefore, statement (2) is an immediate consequence of statement (1) and the result of Levin.

The following is proved as an application of Theorem I.

Theorem II. *Let R be an Artinian Gorenstein local ring with maximal ideal $\mathfrak{m}$ and residue field k. Let M be a finitely generated R-module. Assume that $\operatorname{edim}(R) = n$ and $\mu(\mathfrak{m}^2) \leq 2$. Then the following hold.*

(1) *If $n = 1$, then $P_k^R(t) = \frac{1}{1-t}$ and $(1-t)P_M^R(t) \in \mathbb{Z}[t]$.*

(2) *If $n \geq 2$, then $P_k^R(t) = \frac{1}{1-nt+t^2}$ and $(1+t)^n(1-nt+t^2)P_M^R(t) \in \mathbb{Z}[t]$.*

(3) *If $\operatorname{Ext}^i(M, M) = 0$ for all $i \geq 1$, then M is a free R-module.*

Statement (3) follows from statements (1) and (2) by an argument of Şega [2003]. It implies that R satisfies the Auslander–Reiten conjecture [1975].

The rings considered in Theorem II are called *stretched* when $\mu(\mathfrak{m}^2) = 1$ and *almost stretched* when $\mu(\mathfrak{m}^2) = 2$ (see Definition 3.11). Stretched Cohen–Macaulay local rings were introduced by Sally [1980]. She proved that $P_k^R(t)$ is rational for such a ring R [1980, Theorem 2]. Later Elias and Valla introduced almost stretched Cohen–Macaulay local rings. They proved that if R is an almost stretched Gorenstein local ring and the residue field k of R has characteristic zero, then $P_k^R(t)$ is rational [Elias and Valla 2009, Theorem 1.1]. In a recent article [Croll et al. 2018, Corollary 5.6], stretched Cohen–Macaulay local rings are shown to be good. Using Theorem II, we prove that stretched Cohen Macaulay and almost stretched Gorenstein rings are good without any assumption on residue fields. We also prove that such rings satisfy the Auslander–Reiten conjecture.

Now we briefly describe the organisation of the article. In Section 2, we prove Theorem I. The proof extensively uses a characterisation theorem for Golod algebras (see Theorem 2.1) and chain derivations on acyclic closures whose construction dates back to the work of Gulliksen. The connected sum of Gorenstein local rings was introduced in [Ananthnarayan et al. 2012] (see Definition 3.2). In Section 3, we provide a criterion for connected sum decompositions of Gorenstein local rings. We show that if R is an Artinian Gorenstein local ring with maximal ideal $\mathfrak{m}$ and $\mu(\mathfrak{m}^2) \leq 2$, then R decomposes as a connected sum unless $\mu(\mathfrak{m}) \leq 2$ (Corollary 3.5). We use this decomposition to show that quotients of such a ring R by nonzero powers of maximal ideal $\mathfrak{m}$ are Golod rings (Lemma 3.8). This fact is crucially used in the proof of Theorem II. Finally, Section 4 contains new proofs of Examples (1) and (3) using Theorem I. We identify a certain quotient C of the Koszul algebra K^R of an Artinian Gorenstein ring R such that R is a surjective image of a complete intersection under a Golod homomorphism whenever C is a Golod DG algebra (Section 2B). We show that for the ring considered in Examples (1) and (3), this quotient is a Golod algebra. We make it a point to advertise here that our versions are slightly stronger than the earlier ones in both examples since we constructed Golod homomorphisms from hypersurfaces given by any choice of generator belonging to a minimal generating set of the defining ideal.

We conclude with a remark that our approach only constructs Golod homomorphisms from hypersurface rings. We hope that the present approach can be generalised further to find criteria for existence of Golod homomorphisms from complete intersections of higher codimension.

All rings in this article are Noetherian local rings with $1 \neq 0$. All modules are nonzero and finitely generated. Throughout this article, the expression "local ring $(R, \mathfrak{m}, k)$" refers to a commutative Noetherian local ring R with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. When information on the residue field is not necessary, we denote a local ring R with maximal ideal $\mathfrak{m}$ simply by $(R, \mathfrak{m})$.

2. The main result

Let $(R, \mathfrak{m}, k)$ be a local ring. A DG algebra (A, ∂) over the ring R consists of a nonnegatively graded strictly skew-commutative R-algebra $A = \oplus_{i \geq 0} A_i$ such that

$A_0 = R/I$ for some ideal I of R and an R-linear differential map ∂ of degree -1 satisfying the Leibniz rule. A DG-algebra homomorphism $\phi : A \to B$ is a chain map of complexes which induces a ring homomorphism $\phi^\# : A^\# \to B^\#$ between underlying skew-commutative rings $A^\#$ and $B^\#$ after forgetting the differential maps on A and B. The DG-algebra B is called a semifree extension of A if $B^\#$ is a free module over $A^\#$.

The DG algebra (A, ∂) is augmented if it is equipped with a surjective DG algebra homomorphism $\epsilon : A \twoheadrightarrow k$. If $\tilde{\epsilon} : \mathrm{H}(A) \to k$ is the induced map on homology, we set $\mathrm{I}A = \ker \epsilon$, $\mathrm{IH}(A) = \ker \tilde{\epsilon}$ and $\mathrm{IZ}(A) = \mathrm{I}A \cap \mathrm{Z}(A)$. We say that the DG algebra (A, ∂) is minimal if $\partial(A) \subset \mathfrak{m}A$. A minimal DG algebra A is augmented naturally with the surjective map $\epsilon = q \circ pr$ where $pr : A \twoheadrightarrow A_0$ is the projection and $q : A_0 \twoheadrightarrow k$ is the natural quotient map. We refer to [Avramov 1998] for more information on DG algebras and related terminologies.

2A. ***Tate resolutions.*** Tate described a method to construct a DG algebra resolution of the residue field k over R. The method involves an iterated process of adjoining exterior variables to kill cycles of even degrees and divided powers variables to kill cycles of odd degrees starting from R. In literature, this construction is known as Tate resolution. Later Gulliksen proved that if the number of variables added at each step of killing cycles of a certain degree is the minimum possible, the resulting Tate resolution becomes a minimal free resolution of the residue field k. In this case, the DG algebra is called the acyclic closure of k over R which is unique up to isomorphism of DGΓ algebras. We refer the reader to [Avramov 1998, §6] and [Gulliksen and Levin 1969, Chapter 1] for more details.

In this article, by Tate resolution we mean a surjective R-linear quasi-isomorphism $\epsilon : R\langle X\rangle \twoheadrightarrow k$ where $R\langle X\rangle$ is the acyclic closure of k. The adjoined set of variables $X = \{X_i : i \geq 1\}$ is ordered such that $1 \leq \deg(X_i) \leq \deg(X_j)$ for $i < j$. Note that by construction the acyclic closure $R\langle X\rangle$ is a semifree extension of R.

An R-linear derivation of degree n on the acyclic closure $R\langle X\rangle$ is an R-linear map $\eta : R\langle X\rangle \to R\langle X\rangle$ of degree n satisfying the following properties:

(1) $\eta(R) = 0$ (R-linearity).

(2) η satisfies the Leibniz rule; that is, $\eta(uv) = \eta(u)v + (-1)^{n \deg(u)} x \eta(v)$ for $u, v \in R\langle X\rangle$.

(3) $\eta(X_i^{(i)}) = \eta(X_i) X_i^{(i-1)}$, with $X_i^{(i)}$ being the i-th divided power of a variable X_i of even positive degree.

The derivation η is called a chain derivation if it commutes with the differential ∂ of $R\langle X\rangle$ in the graded sense, i.e., $\eta \circ \partial = (-1)^n \partial \circ \eta$.

Gulliksen and Levin [1969, Theorem 1.6.2] constructed a sequence of R-linear chain derivations η_j on the acyclic closure $R\langle X\rangle$ such that $\eta_j(X_j) = 1$ and $\eta_j(X_i) = 0$ for $i < j$.

2B. ***Golod algebras.*** An augmented DG algebra A over R with augmentation map $\epsilon : A \twoheadrightarrow k$ is called a Golod algebra if A admits a trivial Massey operation, i.e., there are a graded k-basis $\mathfrak{b}_R = \{h_\lambda\}_{\lambda\in\Lambda}$ of $\mathrm{IH}(A)$, a function $\mu : \bigsqcup_{i=1}^{\infty} \mathfrak{b}_R^i \to A$ such that $\mu(h_\lambda) \in \mathrm{IZ}(A)$ with $\mathrm{cls}(\mu(h_\lambda)) = h_\lambda$, and setting $\bar{a} = (-1)^{i+1}a$ for $a \in A_i$ one has

$$\partial\mu(h_{\lambda_1}, \ldots, h_{\lambda_p}) = \sum_{j=1}^{p-1} \overline{\mu(h_{\lambda_1}, \ldots, h_{\lambda_j})}\mu(h_{\lambda_{j+1}}, \ldots, h_{\lambda_p}).$$

The following is proved in [Levin 1976, Theorem 1.5] and also follows from [Levin 1985, Theorem 1.1].

Theorem 2.1. *Let $f : (R, \mathfrak{m}) \twoheadrightarrow (S, \mathfrak{n})$ be a surjective homomorphism of local rings with common residue field k and $\epsilon : R\langle X\rangle \twoheadrightarrow k$ be a DG algebra resolution of k over R. Set $A = R\langle X\rangle \otimes_R S$. Consider A augmented with the augmentation $\epsilon \otimes_R S$. Then the following are equivalent*:

(1) *The DG algebra A is a Golod algebra.*

(2) *The map f is a Golod homomorphism.*

(3) *The induced maps* $\mathrm{Tor}^R(k, k) \to \mathrm{Tor}^S(k, k)$ *and* $\mathrm{Tor}^R(\mathfrak{n}, k) \to \mathrm{Tor}^S(\mathfrak{n}, k)$ *are injective.*

We recall the following result of Levin recorded in [Avramov et al. 1988, Proposition 5.18].

Theorem 2.2. *Let $(R, \mathfrak{m}, k)$ be a local ring and $\phi : P \twoheadrightarrow R$ be a surjective Golod homomorphism from a local complete intersection P of embedding dimension n onto R. Then there exists a polynomial $d_R(t) \in \mathbb{Z}[t]$ such that for any finitely generated R-module M, we have $d_R(t)P_M^R(t) \in \mathbb{Z}[t]$. Further, $d_R(t)P_k^R(t) = (1+t)^n$.*

We are now equipped to prove the main result.

2C. *Proof of Theorem I.*

Proof. Let the maximal ideal of R be $\mathfrak{m}$ and $k = R/\mathfrak{m}$ denote the residue field of R. We know that $\mathrm{H}_1(K^R) \cong I/\mathfrak{n}I$. Therefore, the minimal generators of I are in one-to-one correspondence with the generators of $\mathrm{H}_1(K^R)$. Let the maximal ideal $\mathfrak{n}$ of Q be minimally generated by $y_1, \ldots, y_n$. The Koszul complex of Q is $K^Q = Q\langle X_i : \partial(X_i) = y_i, 1 \le i \le n\rangle$. Let $f = \sum_{i=1}^n a_i y_i$ and $P = Q/(f)$. Note that $K^P = K^Q \otimes_Q P$ and $K^R = K^Q \otimes_Q R$ are Koszul complexes of P and R, respectively. Set $z = \sum_{i=1}^n a_i X_i \in K_1^Q$. Then its residue class $\bar{z}$ is a cycle in $\mathrm{Z}_1(K^P)$. Let $V = K^P\langle T : \partial(T) = \bar{z}\rangle$ be the extension of K^P by adjoining a divided powers variable T of degree 2 to kill the cycle $\bar{z}$. By [Avramov 1998, Theorem 7.3.3], the natural augmentation $V \twoheadrightarrow k$ is the Tate resolution of k over P.

Set $U = V \otimes_R R = K^R\langle T : \partial(T) = \bar{z}\rangle$. Since $f \in I \setminus \mathfrak{n}I$, we have $\bar{z} \in \mathrm{Z}_1(K^R) \setminus B_1(K^R)$. Therefore, we can adjoin variables to U to obtain the acyclic closure $\mathfrak{X}$ of the residue field k over R. The augmentation map $\epsilon : \mathfrak{X} \twoheadrightarrow k$ is the Tate resolution of k over R.

By Theorem 2.1, to show that $P \twoheadrightarrow R$ is a Golod homomorphism, we need to prove that the induced maps $\operatorname{Tor}^P(k,k) \to \operatorname{Tor}^R(k,k)$ and $\operatorname{Tor}^P(\mathfrak{m},k) \to \operatorname{Tor}^R(\mathfrak{m},k)$ are injective. Both V and $\mathfrak{X}$ are minimal algebras. Therefore, the first map is $U \otimes_R k \to \mathfrak{X} \otimes_R k$ which is obviously injective since $\mathfrak{X}$ is a semifree extension of U. The second map is $i_* : \mathrm{H}(\mathfrak{m}U) \to \mathrm{H}(\mathfrak{m}\mathfrak{X})$ which is induced by the inclusion $\mathfrak{i} : \mathfrak{m}U \to \mathfrak{m}\mathfrak{X}$. We prove that $\mathfrak{i}_*$ is an injective map.

We have an R-linear chain derivation $v : \mathfrak{X} \to \mathfrak{X}$ of degree -2 such that $v(T) = 1$. Set $\bar{R} = R/\operatorname{socle}(R)$. Note that $\operatorname{socle}(R) \subset \mathfrak{m}^2$, so $K^{\bar{R}} = \bar{R} \otimes_R K^R$ is the Koszul complex of $\bar{R}$. Now $K^{\bar{R}}$ can be extended to the acyclic closure $\mathfrak{Y}$ over $\bar{R}$. Let $\mathfrak{j} : K^{\bar{R}} \to \mathfrak{Y}$ denote the inclusion. The augmentation $\epsilon_{\mathfrak{Y}} : \mathfrak{Y} \twoheadrightarrow k$ is an algebra homomorphism over K^R. The acyclic closure $\mathfrak{X}$ is semifree over K^R. Therefore, the augmentation $\epsilon_{\mathfrak{X}} : \mathfrak{X} \twoheadrightarrow k$ lifts to a DG algebra homomorphism $\beta : \mathfrak{X} \to \mathfrak{Y}$ over K^R [Avramov 1998, Proposition 2.1.9]. Let $\alpha : K^R \to K^{\bar{R}}$ denote the quotient map. By abuse of notation we denote restriction of a map by the same symbol. We have the following commutative diagram:

$$\begin{array}{ccc} \mathfrak{m}K^R & \xrightarrow{\mathfrak{i}} & \mathfrak{m}\mathfrak{X} \\ \Big\downarrow{\scriptstyle\alpha} & & \Big\downarrow{\scriptstyle\beta} \\ \mathfrak{m}K^{\bar{R}} & \xrightarrow{\mathfrak{j}} & \mathfrak{m}\mathfrak{Y} \end{array}$$

A cycle y in $\mathfrak{m}U$ can be written as $y = \sum_{k=0}^{m} a_k T^{(m-k)}$, $a_i \in \mathfrak{m}K^R$. Suppose $\mathfrak{i}(y)$ is in the boundary of $\mathfrak{m}\mathfrak{X}$. We prove by induction on m that y is in the boundary of $\mathfrak{m}U$.

First assume that $m = 0$. Then $y \in \mathfrak{m}K^R$. Since $\mathfrak{i}(y)$ is in the boundary of $\mathfrak{m}\mathfrak{X}$, $\mathfrak{j} \circ \alpha(y)$ is in the boundary of $\mathfrak{m}\mathfrak{Y}$ by the commutative diagram. Now $\bar{R}$ is a Golod ring, so $\mathfrak{j}$ induces an injective map $\mathfrak{j}_* : \mathrm{H}(\mathfrak{m}K^{\bar{R}}) \to \mathrm{H}(\mathfrak{m}\mathfrak{Y})$. Therefore, $\alpha(y)$ is in the boundary of $\mathfrak{m}K^{\bar{R}}$. This implies that $y = sy_1 + \partial(y_2)$ where $y_1 \in K^R$, $y_2 \in \mathfrak{m}K^R$ and $\operatorname{socle}(R) = (s)$.

We know from [Levin and Avramov 1978, Lemma 1.2] that $\operatorname{socle}(R)K_i^R \subset (0 : \mathfrak{m}^2)\,\mathrm{B}_i(K^R)$ for $1 \le i \le n-1$. If $\deg(y) = \deg(y_1) < n$, then $sy_1 \in (0 : \mathfrak{m}^2)\,\mathrm{B}(K^R)$ and consequently $y \in \mathfrak{m}\,\mathrm{B}(K^R) \subset \mathfrak{m}\,\mathrm{B}(U)$. On the other hand if $\deg(y) = n$, then $y_2 = 0$ and $y = sy_1 = asX_1 \cdots X_n$, $a \in R$. Since $\mathrm{H}(K^R)$ is a Poincaré algebra [Avramov and Golod 1971], there is a $z' \in \mathrm{Z}_{n-1}(K^R)$ such that $\bar{z}z' = sX_1 \cdots X_n$. We conclude $y = a\bar{z}z' = \partial(aTz') \in \mathfrak{m}\,\mathrm{B}(U)$. Therefore, the induction step for $m = 0$ follows.

Now we assume that $m > 0$. Note that $a_0 = v^m(\mathfrak{i}(y))$. Since $\mathfrak{i}(y) \in \mathfrak{m}\,\mathrm{B}(\mathfrak{X})$ and the chain derivation v commutes with the differential of $\mathfrak{X}$, we have $a_0 \in \mathfrak{m}K^R \cap \mathrm{B}(\mathfrak{m}\mathfrak{X})$. We consider two cases.

Suppose $\deg(a_0) = n$. Then $a_0 \in Z_n(\mathfrak{m}K^R) = \operatorname{socle}(R)K_n^R$. Therefore, $a_0 = asX_1 \cdots X_n,\ a \in R$. One observes $a_0T^{(m)} = a\bar{z}z'T^{(m)} = \partial(az'T^{(m+1)}) \in \mathfrak{m}\,B(U)$. Therefore, $\mathfrak{i}\big(\sum_{k=0}^{m-1} a_k T^{(m-k)}\big) = \mathfrak{i}(y) - a_0T^{(m)}$ is in the boundary of $\mathfrak{m}\mathfrak{X}$. Consequently, $\sum_{k=0}^{m-1} a_k T^{(m-k)}$ is in the boundary of $\mathfrak{m}U$ by the induction hypothesis. We conclude that y is in the boundary of $\mathfrak{m}U$.

Suppose $\deg(a_0) < n$. Then by the argument in the induction step $m = 0$, one has $a_0 = \partial(y_3)$ for $y_3 \in \mathfrak{m}K^R$. We can write

$$y=\partial(y_3)T^{(m)}+\sum_{k=0}^{m-1}a_kT^{(m-k)}=\partial(y_3T^{(m)})+\Big[(-1)^{\deg(y_3)+1}y_3\bar{z}T^{(m-1)}+\sum_{k=0}^{m-1}a_kT^{(m-k)}\Big].$$

The first summand is in $\mathfrak{m}\,B(U)$. This implies that the second summand is in the boundary of $\mathfrak{m}\mathfrak{X}$ and therefore also in the boundary of $\mathfrak{m}U$ by induction hypothesis. We conclude that y is in the boundary of $\mathfrak{m}U$. This completes the induction step. Hence $\mathfrak{i}_*$ is an injective map and statement (1) follows.

The ring $\bar{R}$ is Golod. The Poincaré series of R is computed in [Rossi and Şega 2014, Proposition 6.2] as

$$P_k^R(t) = \frac{(1+t)^n}{1 - t(P_R^Q(t) - 1) + t^{n+1}(1+t)}$$

so statement (2) follows from Theorem 2.2. □

3. Stretched and almost stretched rings

Our aim in this section is to prove that stretched and almost stretched Gorenstein rings are good. The key step is to show that rings of these types decompose as connected sums. If the residue field is infinite, then Lemma 3.1 follows from [Eakin and Sathaye 1976, Theorem 1]. The proof of the lemma was suggested by the anonymous referee.

Lemma 3.1. *Let $(R, \mathfrak{m}, k)$ be a local ring such that $\mu(\mathfrak{m}^2) \le 2$. Then there exists an $x \in \mathfrak{m} \setminus \mathfrak{m}^2$ such that $\mathfrak{m}^2 = x\mathfrak{m}$. Furthermore, if $\operatorname{ll}(R) \ge 3$, then $x^2 \notin \mathfrak{m}^3$.*

Proof. If $\mathfrak{m}^2 = x\mathfrak{m} + \mathfrak{m}^3$ and $x \in \mathfrak{m} \setminus \mathfrak{m}^2$, then by Nakayama's lemma we have $m^2 = x\mathfrak{m}$. Therefore, to prove the first assertion, it is enough to assume that $\mathfrak{m}^3 = 0$, i.e., $\mathfrak{m}^2$ is a k-vector space. If $\mathfrak{m}^2 = 0$, then $\mathfrak{m}^2 = x\mathfrak{m} = 0$ for all $x \in \mathfrak{m} \setminus \mathfrak{m}^2$. If $\mu(\mathfrak{m}^2) = 1$, then for any $x \in \mathfrak{m} \setminus \mathfrak{m}^2$ such that $x\mathfrak{m} \neq 0$, we have $\mathfrak{m}^2 = x\mathfrak{m}$. Therefore, we only need to consider the case when $\mu(\mathfrak{m}^2) = 2$, i.e., $\mathfrak{m}^2$ is a vector space of dimension two.

Let $x_1, \dots, x_n$ be a minimal generating set of $\mathfrak{m}$. Let r be such that $x_i\mathfrak{m} \neq 0$ for all i with $1 \le i \le r$ but $x_i\mathfrak{m} = 0$ for $i > r$. Assume, by way of contradiction, that $\mathfrak{m}^2 \neq x\mathfrak{m}$ for all $x \in \mathfrak{m} \setminus \mathfrak{m}^2$. Thus, if $i \le r$ then $x_i\mathfrak{m}$ is a one-dimensional vector space. We may also assume $\mathfrak{m}^2 = x_1\mathfrak{m} + x_2\mathfrak{m}$. Clearly $x_1\mathfrak{m} \neq x_2\mathfrak{m}$ since otherwise $\mathfrak{m}^2 = x_1\mathfrak{m}$, a contradiction to our assumption.

If $i \leq r,\ j \leq r$ and $x_i x_j \neq 0$, then one observes that $x_i\mathfrak{m} = x_j\mathfrak{m}$. This is true because $x_i\mathfrak{m}$ and $x_j\mathfrak{m}$ are both one-dimensional vector spaces and they share the nonzero element $x_i x_j$. Since $x_1\mathfrak{m} \neq x_2\mathfrak{m}$, we must therefore have $x_1x_2 = 0$. Now $x_1\mathfrak{m} \neq 0$ and $x_2\mathfrak{m} \neq 0$, so there exist i, j with $i \leq r,\ j \leq r$ such that $x_1x_i \neq 0$ and $x_2x_j \neq 0$. This implies $x_1\mathfrak{m} = x_i\mathfrak{m}$ and $x_2\mathfrak{m} = x_j\mathfrak{m}$. In particular, $x_i\mathfrak{m} \neq x_j\mathfrak{m}$ and hence $x_ix_j = 0$. We have $(x_2 + x_i)x_1 = x_1x_i \neq 0$ and $(x_2 + x_i)x_j = x_2x_j \neq 0$, hence $(x_2 + x_i)\mathfrak{m} = x_1\mathfrak{m}$ and $(x_2 + x_i)\mathfrak{m} = x_j\mathfrak{m} = x_2\mathfrak{m}$. This yields $x_1\mathfrak{m} = x_2\mathfrak{m}$, a contradiction. Therefore, the first part of the lemma follows.

If $x^2 \in \mathfrak{m}^3$, then $\mathfrak{m}^3 = x^2\mathfrak{m} \subset \mathfrak{m}^4$. By Nakayama's lemma, $\mathfrak{m}^3 = 0$ which implies $\mathrm{ll}(R) \leq 2$. Therefore, $x^2 \notin \mathfrak{m}^3$ if $\mathrm{ll}(R) \geq 3$. □

We recall definitions of fibre products and connected sums [Ananthnarayan et al. 2012].

Definition 3.2. Let $(R, \mathfrak{m}_R, k)$ and $(S, \mathfrak{m}_S, k)$ be local rings with a common residue field k. Let $\pi_R : R \to k$ and $\pi_S : S \to k$ be natural quotient maps from R and S onto k, respectively. The fibre product of R and S is defined as the ring $R \times_k S = \{(r, s) \in R \times S : \pi_R(r) = \pi_S(s)\}$. The ring $R \times_k S$ is local with maximal ideal $\mathfrak{m}_R \oplus \mathfrak{m}_S$.

Now assume that both R and S are Artinian Gorenstein local rings with one-dimensional socles $\mathrm{socle}(R) = \langle \delta_R \rangle$ and $\mathrm{socle}(S) = \langle \delta_S \rangle$. Then the connected sum of R and S is defined as

$$R\#S = \frac{R \times_k S}{\langle (\delta_R, -\delta_S) \rangle}.$$

We say a Gorenstein local ring Q is decomposable as a connected sum if there are rings R and S such that $Q = R\#S,\ l(R) < l(Q)$ and $l(S) < l(Q)$. Here $l(-)$ denotes the length function.

Define a left module structure on the polynomial ring $T = k[Y_1, \ldots, Y_n]$ over the ring $S = k[X_1, \ldots, X_n],\ X_i = Y_i^{-1}$, by defining the action of X_i on a monomial $M \in T$ as the usual multiplication if $X_iM \in T$ and zero otherwise. Macaulay's inverse system establishes a one-to-one correspondence between local Artinian Gorenstein algebras

$$R = \frac{k[X_1, \ldots, X_n]}{I}$$

such that $I \subset (X_1, \ldots, X_n)$ and polynomials F in the ring T up to a unit multiple. The correspondence is given by $I = \mathrm{ann}\, F$; see [Eisenbud 1995, Theorem 21.6]. If Gorenstein local rings R and S correspond to $F \in k[Y_1, \ldots, Y_m]$ and $G \in k[Y_{m+1}, \ldots, Y_n]$, respectively, then the connected sum $R\#S$ corresponds to $F + G \in k[Y_1, \ldots, Y_n]$; see [Ananthnarayan 2009, Remark 4.24].

The following result follows from [Ananthnarayan et al. 2019, Proposition 4.1] and can be proved easily for Artinian k-algebras using Macaulay's inverse system.

Theorem 3.3. *Let $(Q, \mathfrak{n}, k)$ be a regular local ring and $I \subset \mathfrak{n}^2$ be an ideal such that $R = Q/I$ is an Artinian Gorenstein local ring. Let $\mathfrak{n}$ be minimally generated by $x_1, \dots, x_m, y_1, \dots, y_n$ such that $(x_1, \dots, x_m)(y_1, \dots, y_n) \subset I$. Let $\max\{i : (x_1, \dots, x_m)^i \not\subset I\} = s$ and $\max\{i : (y_1, \dots, y_n)^i \not\subset I\} = t$. Then there are ideals I_1 and I_2 in Q containing $(x_1, \dots, x_m)$ and $(y_1, \dots, y_n)$, respectively, such that the following hold.*

(1) *The rings $S = Q/I_1$ and $T = Q/I_2$ are Gorenstein rings. Further, $\operatorname{edim}(S) = n$, $\operatorname{edim}(T) = m$, $\operatorname{ll}(S) = t$, and $\operatorname{ll}(T) = s$.*

(2) *$R = S\#_k T$.*

The next theorem is the key to decomposing an Artinian Gorenstein local ring $(R, \mathfrak{m}, k)$ with $\mu(\mathfrak{m}^2) \le 2$ as a connected sum. When $\operatorname{edim}(R) = 2$ and $\operatorname{char}(k) = 0$, the theorem follows from [Elias and Valla 2008, Theorem 4.1].

Theorem 3.4. *Let $(R, \mathfrak{m}, k)$ be a local Artinian Gorenstein ring. Let $\operatorname{edim}(R) = n$, $\operatorname{ll}(R) \ge 3$ and $\dim_k \mathfrak{m}^2/\mathfrak{m}^3 = m < n$. Assume that $\mathfrak{m}$ admits a generator x_1 such that $\mathfrak{m}^2 = x_1\mathfrak{m}$. Then there exists a minimal generating set $\{x_1, x_2, \dots, x_n\}$ of $\mathfrak{m}$ extending x_1 such that the following hold.*

(1) $\mathfrak{m}^2 = (x_1^2, x_1x_2, \dots, x_1x_m)$.

(2) $(x_1, \dots, x_m)(x_{m+1}, \dots, x_n) = 0$.

(3) $(x_{m+1}, \dots, x_n)^2 = \operatorname{socle}(R)$.

The ring R decomposes as a connected sum $R = S\#T$ such that $\operatorname{edim}(S) = m$, $\operatorname{edim}(T) = n - m$, $\operatorname{ll}(S) = \operatorname{ll}(R)$ and $\operatorname{ll}(T) = 2$.

Proof. Since $\operatorname{ll}(R) \ge 3$, we have $x_1^2 \notin \mathfrak{m}^3$. Therefore, we can choose a minimal generating set $\{x_1, x_2, \dots, x_n\}$ of $\mathfrak{m}$ such that $\mathfrak{m}^2 = (x_1^2, x_1x_2, \dots, x_1x_m)$. Statement (1) follows.

We have $x_1x_j = \alpha_{1j}x_1^2 + \alpha_{2j}x_1x_2 + \cdots + \alpha_{mj}x_1x_m$, $\alpha_{ij} \in R$ for $1 \le i \le m$ and $m+1 \le j \le n$. This gives $x_1(x_j - \alpha_{1j}x_1 - \alpha_{2j}x_2 - \cdots - \alpha_{mj}x_m) = 0$. Replacing $x_j - \alpha_{1j}x_1 - \alpha_{2j}x_2 - \cdots - \alpha_{mj}x_m$ by x_j, we assume that $x_1x_j = 0$ for $m+1 \le j \le n$.

If $m = 1$, property (2) is satisfied. We assume that $\mu(\mathfrak{m}^2) = m \ge 2$. Since $\mathfrak{m}^2$ is minimally generated by $\{x_1^2, x_1x_2, \dots, x_1x_m\}$ and $x_1(x_{m+1}, \dots, x_n) = 0$, we have

$$(0 : x_1) \subset \mathfrak{m}(x_1, \dots, x_m) + (x_{m+1}, \dots, x_n).$$

This implies that the residue classes of elements $x_{m+1}, \dots, x_n$ form a k-basis of $((0 :_R x_1) + \mathfrak{m}^2)/\mathfrak{m}^2$. Therefore, $\dim_k(((0 :_R x_1) + \mathfrak{m}^2)/\mathfrak{m}^2) = n - m$.

Claim 1: $\mathfrak{m}[(0 :_R x_1) \cap \mathfrak{m}^2] = \operatorname{socle}(R)$.

Proof. Note $\mathfrak{m}^2(0:_R x_1) = x_1\mathfrak{m}(0:_R x_1) = 0$. This implies $\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2] \subset \operatorname{socle}(R)$. Therefore, it is enough to prove that $\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2] \neq 0$. We have $\mathfrak{m}^{i+1} = x_1\mathfrak{m}^i$ for $i \geq 1$. This means that $\mu(\mathfrak{m}^{i+1}) \leq \mu(\mathfrak{m}^i)$, $i \geq 1$. Let $t = \max\{i : \mu(\mathfrak{m}^i) = m\}$. Then $t \geq 2$. Since $m \geq 2$, we have $\operatorname{ll}(R) \geq t+1$. The map $\mathfrak{m}^t/\mathfrak{m}^{t+1} \xrightarrow{\cdot x_1} \mathfrak{m}^{t+1}/\mathfrak{m}^{t+2}$ is not injective since $\dim_k(\mathfrak{m}^t/\mathfrak{m}^{t+1}) > \dim_k(\mathfrak{m}^{t+1}/\mathfrak{m}^{t+2})$. Therefore, we find $y \in \mathfrak{m}^t \setminus \mathfrak{m}^{t+1}$ such that $yx_1 \in \mathfrak{m}^{t+2}$. Note that $\mathfrak{m}^{t+2} = x_1\mathfrak{m}^{t+1}$. It follows that $yx_1 = x_1 m$ for some $m \in \mathfrak{m}^{t+1}$. Consequently $x_1(y-m) = 0$. Clearly, $y-m \in [(0:_R x_1)\cap \mathfrak{m}^2]$ and $y-m \notin \mathfrak{m}^{t+1}$. Since $\operatorname{socle}(R) \subset \mathfrak{m}^{t+1}$, we have $y-m \notin \operatorname{socle}(R)$. Therefore, $[(0:_R x_1)\cap \mathfrak{m}^2] \not\subset \operatorname{socle}(R)$. We conclude that $\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2] \neq 0$ and the claim is proved. □

Claim 2: $\dim_k\big([(0:_R x_1)\cap \mathfrak{m}^2]/(\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2])\big) = m-1$.

Proof. We know that $(0:_R x_1) = \operatorname{Hom}_R(R/(x_1), R)$. We have $\mathfrak{m}^2 = x_1\mathfrak{m} \subset (x_1)$. By Matlis duality $l(0:_R x_1) = l(R/x_1R) = l(R/\mathfrak{m}^2) - l((x_1R+\mathfrak{m}^2)/\mathfrak{m}^2) = 1+n-1 = n$. Now we have

$$l[(0:_R x_1)\cap \mathfrak{m}^2] = l(0:_R x_1) + l(\mathfrak{m}^2) - l[(0:_R x_1)+\mathfrak{m}^2]$$
$$= l(0:_R x_1) - l\left[\frac{(0:_R x_1)+\mathfrak{m}^2}{\mathfrak{m}^2}\right] = n-(n-m) = m.$$

Therefore,

$$\dim_k \frac{[(0:_R x_1)\cap \mathfrak{m}^2]}{\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]} = l\left(\frac{[(0:_R x_1)\cap \mathfrak{m}^2]}{\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]}\right) = l[(0:_R x_1)\cap \mathfrak{m}^2]-1 = m-1. \quad \square$$

Claim 3: The pairing

$$\frac{(x_2,\ldots,x_m)}{\mathfrak{m}(x_2,\ldots,x_m)} \times \frac{[(0:_R x_1)\cap \mathfrak{m}^2]}{\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]} \to \operatorname{socle}(R)$$

given by $(\bar{x},\bar{y}) \to xy$ is well defined and nondegenerate.

Proof. We have $\mathfrak{m}(x_2,\ldots,x_m)(0:_R x_1) = 0$ since $\mathfrak{m}^2 = x_1\mathfrak{m}$. This implies that $(x_2,\ldots,x_m)[(0:_R x_1)\cap \mathfrak{m}^2] \subset \operatorname{socle}(R)$. Therefore, the above pairing exists. Note that $\mathfrak{m}^2(x_{m+1},\ldots,x_n) = 0$. As a result, if $y \in [(0:_R x_1)\cap \mathfrak{m}^2]$ and $y(x_2,\ldots,x_m) = 0$, we have $y \in \operatorname{socle}(R) = \mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]$. This implies that the map

$$\frac{[(0:_R x_1)\cap \mathfrak{m}^2]}{\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]} \to \operatorname{Hom}\left(\frac{(x_2,\ldots,x_m)}{\mathfrak{m}(x_2,\ldots,x_m)}, \operatorname{socle}(R)\right)$$

induced by the above pairing is injective. We have

$$\dim_k \frac{(x_2,\ldots,x_m)}{\mathfrak{m}(x_2,\ldots,x_m)} = \dim_k \frac{[(0:_R x_1)\cap \mathfrak{m}^2]}{\mathfrak{m}[(0:_R x_1)\cap \mathfrak{m}^2]} = m-1.$$

Therefore, the above map is an isomorphism and consequently the pairing is nondegenerate. □

Note that $\mathfrak{m}(x_2, \dots, x_m)(x_{m+1}, \dots, x_n) = 0$ so

$$(x_2, \dots, x_m)(x_{m+1}, \dots, x_n) \subset \operatorname{socle}(R).$$

The multiplication by x_j defines a map $(x_2, \dots, x_m)/(\mathfrak{m}(x_2, \dots, x_m)) \to \operatorname{socle}(R)$ for $m+1 \le j \le n$. Since the pairing in Claim 3 is nondegenerate, we have $\hat{x}_j \in [(0 :_R x_1) \cap \mathfrak{m}^2]$ such that $x_j x_i = \hat{x}_j x_i$ for $2 \le i \le m$ and $m+1 \le j \le n$. We also have $x_1 x_j = x_1 \hat{x}_j = 0$ for $m+1 \le j \le n$. It follows that $(x_j - \hat{x}_j)(x_1, \dots, x_m) = 0$, $m+1 \le j \le n$. Therefore, replacing $(x_j - \hat{x}_j)$ by x_j we have $(x_1, \dots, x_m)(x_{m+1}, \dots, x_n) = 0$ and property (2) is satisfied.

Let $\operatorname{socle}(R) = \langle \delta \rangle$ and $K = (x_{m+1}, \dots, x_n)$. We have $\mathfrak{m}^2 K = x_1 K \mathfrak{m} = 0$, so $K^2 \subset \mathfrak{m}K \subset \operatorname{socle}(R)$. Note that $K^2 \ne 0$ for otherwise each of the x_j, $m+1 \le j \le n$, is in $\operatorname{socle}(R) \subset \mathfrak{m}^2$, a contradiction. Therefore, $K^2 = \mathfrak{m}K = \operatorname{socle}(R)$ and the property (3) is satisfied.

The last statement follows from Theorem 3.3. □

The following is a consequence of Lemma 3.1 and Theorem 3.4.

Corollary 3.5. *Let $(R, \mathfrak{m})$ be an Artinian Gorenstein local ring such that* $\operatorname{ll}(R) \ge 3$ *and* $\mu(\mathfrak{m}^2) \le \min\{2, \operatorname{edim}(R) - 1\}$. *Then $R = S\#T$ where $(S, \mathfrak{p})$ and $(T, \mathfrak{q})$ are Gorenstein local rings*, $\operatorname{edim}(S) = \mu(\mathfrak{m}^2)$, $\operatorname{ll}(S) = \operatorname{ll}(R)$ *and* $\operatorname{ll}(T) = 2$.

Lemma 3.6. *Let $(R, \mathfrak{m})$ be an Artinian Gorenstein local ring such that* $\operatorname{edim}(R) \ge 2$ *and* $\operatorname{ll}(R) = s$. *Then the quotient ring $R/\mathfrak{m}^i$ is not a Gorenstein ring for $2 \le i \le s$.*

Proof. If possible assume that $R/\mathfrak{m}^i$ is a Gorenstein ring for some i satisfying $2 \le i \le s$. Then the injective hull of k over $R/\mathfrak{m}^i$ is $E_{R/\mathfrak{m}^i}(k) \cong R/\mathfrak{m}^i$. We know that $E_{R/\mathfrak{m}^i}(k) \cong \operatorname{Hom}_R(R/\mathfrak{m}^i, R) = (0 :_R \mathfrak{m}^i)$. Consequently $(0 :_R \mathfrak{m}^i) = (x)$, a principal ideal for some $x \in R$. Now $\mathfrak{m}^{s-i+1} \subset (0 :_R \mathfrak{m}^i)$ and $\mathfrak{m}^{s-i+1} \not\subset \mathfrak{m}(0 :_R \mathfrak{m}^i)$ for otherwise $\mathfrak{m}^{s-i+1} \subset \mathfrak{m}(0 :_R \mathfrak{m}^i) \subset (0 :_R \mathfrak{m}^{i-1})$ which implies $\mathfrak{m}^s = \mathfrak{m}^{s-i+1}\mathfrak{m}^{i-1} = 0$, a contradiction. Since $(0 :_R \mathfrak{m}^i)$ is principal, we have $(0 :_R \mathfrak{m}^i) = \mathfrak{m}^{s-i+1} = (x)$.

Apply Macaulay's theorem characterising Hilbert function [Bruns and Herzog 1993, Theorem 4.2.10] to the associated graded ring $\operatorname{gr}_{\mathfrak{m}}(R)$. We obtain $\mu(\mathfrak{m}^{n+1}) \le \mu(\mathfrak{m}^n)^{\langle n \rangle}$ for all $n \ge 1$. We already have $\mu(\mathfrak{m}^{s-i+1}) = 1$. This implies that $\mathfrak{m}^j$ is a principal ideal for j satisfying the inequality $s - i + 1 \le j \le s$ and so $l(0 :_R \mathfrak{m}^i) = l(\mathfrak{m}^{s-i+1}) = i$. By Matlis duality, $l(R/\mathfrak{m}^i) = l \operatorname{Hom}(R/\mathfrak{m}^i, R) = l(0 :_R \mathfrak{m}^i) = i$. We have $\sum_{j=0}^{i-1} [l(\mathfrak{m}^j/\mathfrak{m}^{j+1}) - 1] = l(R/\mathfrak{m}^i) - i = 0$ and each summand is nonnegative. This shows that $\operatorname{edim}(R) = l(\mathfrak{m}/\mathfrak{m}^2) = 1$, a contradiction. □

The following theorem is proved in [Dress and Krämer 1975, Satz 2].

Theorem 3.7. *Let $(S, \mathfrak{m}_S, k)$ and $(T, \mathfrak{m}_T, k)$ be two local rings and $R = S \times_k T$. Then,*

$$\frac{1}{P_k^R(t)} = \frac{1}{P_k^S(t)} + \frac{1}{P_k^T(t)} - 1.$$

If M is an S-module, then

$$\frac{1}{P_M^R(t)} = \frac{P_k^S(t)}{P_M^S(t)}\left(\frac{1}{P_k^S(t)} + \frac{1}{P_k^T(t)} - 1\right) = \frac{P_k^S(t)}{P_M^S(t)\,P_k^R(t)}.$$

The following is a consequence of Lemma 3.1 and Theorem 3.4.

Lemma 3.8. *Let $(R, \mathfrak{m}, k)$ be an Artinian Gorenstein local ring such that $\mu(\mathfrak{m}) = n$, $\operatorname{ll}(R) \geq 2$ and $\mu(\mathfrak{m}^2) \leq 2$. Let i be an integer satisfying $2 \leq i \leq \operatorname{ll}(R)$. Then $R/\mathfrak{m}^i$ is a Golod ring and $P_k^{R/\mathfrak{m}^i}(t) = \frac{1}{1-nt}$.*

Proof. Fix i satisfying $2 \leq i \leq \operatorname{ll}(R)$ and set $\bar{R} = R/\mathfrak{m}^i$. The result is clear when $n = 1$ or $\operatorname{ll}(R) = 2$. First we assume that $n = 2$ and $\operatorname{ll}(R) > 2$. A result of Scheja [1964] states that a codepth 2 local ring is either a Gorenstein (equivalently complete intersection) or a Golod ring. The ring $\bar{R}$ cannot be a Gorenstein ring by Lemma 3.6 so $\bar{R}$ is a Golod ring. Since R is a complete intersection, the defining ideal of $\bar{R}$ is minimally generated by three elements. It follows that $\kappa^{\bar{R}}(t) = \sum_{i \geq 0} \dim_k \mathrm{H}_i(K^{\bar{R}})t^i = 1 + 3t + 2t^2$ and

$$(1 - t(\kappa^{\bar{R}}(t) - 1)) = 1 - t(1 + 3t + 2t^2 - 1) = (1+t)^2(1-2t).$$

We have

$$P_k^{\bar{R}}(t) = \frac{(1+t)^2}{1 - t(\kappa^{\bar{R}}(t) - 1)} = \frac{1}{1-2t}.$$

Now we assume that $n > 2$ and $\operatorname{ll}(R) > 2$. By Corollary 3.5, it follows that $R = S \# T$ where $(S, \mathfrak{p})$ and $(T, \mathfrak{q})$ are Gorenstein local rings, $\operatorname{edim}(S) = \mu(\mathfrak{m}^2) \leq 2$, $\operatorname{ll}(S) = \operatorname{ll}(R) \geq 3$ and $\operatorname{ll}(T) = 2$. One has $\bar{R} = \bar{S} \times_k \bar{T}$ where $\bar{S} = S/\mathfrak{p}^i$ and $\bar{T} = T/\mathfrak{q}^2$. Both $\bar{S}$ and $\bar{T}$ are Golod rings. The ring $\bar{R}$ is a Golod ring because a fibre product of Golod rings is Golod; see [Lescot 1983, Theorem 4.1].

$P_k^{\bar{S}}(t) = 1/(1 - \operatorname{edim}(\bar{S})t)$ by the case $n = 2$ and $P_k^{\bar{T}}(t) = 1/(1 - \operatorname{edim}(\bar{T})t)$ ([Avramov 1998, Example 4.2.2]). The formula for the Poincaré series follows from Theorem 3.7. □

The following is a well known fact.

Lemma 3.9. *Let $(R, \mathfrak{m}, k)$ be a local ring, x be a nonzero divisor of R and $S = R/(x)$. If there exists a polynomial $d(t) \in \mathrm{Z}[t]$ such that $d(t)P_M^S(t) \in \mathbb{Z}[t]$ for all S-modules M, then $d(t)P_M^R(t) \in \mathbb{Z}[t]$ for all R-modules M. Now assume further that $x \in \mathfrak{m} \setminus \mathfrak{m}^2$, then $P_k^S(t) = P_k^R(t)/(1+t)$. The ring S is Golod if and only if R is so.*

Proof. Let M be an R-module and N be the first syzygy of M. We have the exact sequence

$$0 \to N \to R^{\mu(M)} \to M \to 0.$$

This implies that $P^R_M(t) = \mu(M) + t P^R_N(t)$. Therefore, it is enough to show that $d(t)P^R_N(t) \in \mathbb{Z}[t]$. Let $F_* \twoheadrightarrow N$ be the minimal free resolution of N over R. Note that x is also a nonzero divisor of N. This implies that $S \otimes_R F_* \twoheadrightarrow S \otimes_R N$ is a minimal free resolution of $S \otimes_R N = \bar{N}$ as an S module. As a result, we have $P^R_N(t) = P^S_{\bar{N}}(t)$. Therefore, the first part of the lemma follows from the hypothesis.

The assertions regarding Poincaré series and Golod property follow from Propositions 3.3.5 (1) and 5.2.4 in [Avramov 1998], respectively. □

The following is due to Şega [2003, Proposition 1.5].

Proposition 3.10. *Let R be a local ring such that there is a $d_R(t) \in \mathbb{Z}[t]$ satisfying $d_R(t)P^R_M(t) \in \mathbb{Z}[t]$ for each finitely generated R-module M. Let $d_R(t) = p(t)q(t)r(t)$ where $p(t)$ is 1 or irreducible, $q(t)$ has nonnegative coefficients, $r(t)$ is 1 or irreducible and has no positive real root among its complex roots of minimal absolute value. Then the following hold for each pair of R-modules M, N.*

(1) *If $\operatorname{Tor}^R_i(M, N) = 0$ for $i \gg 0$, either M or N has finite projective dimension.*
(2) *If $\operatorname{Ext}^i_R(M, N) = 0$ for $i \gg 0$, either M has a finite projective dimension or N has a finite injective dimension.*

We are ready to prove Theorem II.

3A. *Proof of Theorem II.*

Proof. The case $n = 1$ is easy. We skip the details.

Now we assume that $n \geq 2$. The quotient map $R \twoheadrightarrow \frac{R}{\operatorname{socle}(R)}$ is a Golod homomorphism and

$$P_k^{R/\operatorname{socle}(R)}(t) = \frac{P^R_k(t)}{1 - t^2 P^R_k(t)}$$

[Levin and Avramov 1978, Theorem 2]. Therefore, we have

$$P^R_k(t) = \frac{P_k^{R/\operatorname{socle}(R)}(t)}{1 + t^2 P_k^{R/\operatorname{socle}(R)}(t)} = \frac{1}{1/(P_k^{R/\operatorname{socle}(R)}(t)) + t^2} = \frac{1}{1 - nt + t^2}.$$

The last equality follows from Lemma 3.8. The same lemma states that $R/\operatorname{socle}(R)$ is a Golod ring. By Theorem I, the ring R is a surjective image of a complete intersection under a Golod homomorphism. The second part of statement (2) is a consequence of Theorem 2.2.

Now we prove (3). If the projective dimension $\operatorname{pd}_R(M)$ of M is finite, then $\operatorname{Ext}^{\operatorname{pd}_R(M)}_R(M, M) \neq 0$. Therefore, it is enough to show that $\operatorname{pd}_R(M) < \infty$. When $n \leq 2$, R is a complete intersection. The statement follows from [Avramov and Buchweitz 2000, Theorem 4.2]. When $n \geq 3$, the polynomial $(1 - nt + t^2)$ is irreducible.

The statement follows from Proposition 3.10 (take $p(t) = (1 - nt + t^2)$, $q(t) = (1+t)^n$ and $r(t) = 1$). Here one uses the fact that a module over a Gorenstein ring has a finite projective dimension if and only if it has a finite injective dimension. □

Stretched and almost stretched rings were introduced by Sally [1980] and Elias and Valla [2008], respectively.

Definition 3.11. An Artinian local ring R with maximal ideal $\mathfrak{m}$ is called stretched if $\mathfrak{m}^2$ is a principal ideal and almost stretched if $\mathfrak{m}^2$ is minimally generated by two elements.

Let R be a Cohen–Macaulay local ring of dimension d with maximal ideal $\mathfrak{m}$. Then R is called stretched (almost stretched) if there exists a minimal reduction $\underline{x} = x_1, \dots, x_d$ of $\mathfrak{m}$ such that $R/(\underline{x})$ is a stretched (almost stretched) Artinian ring. Here by minimal reduction we mean that $\underline{x}$ satisfies $\mathfrak{m}^{r+1} = (x_1, \dots, x_d)\mathfrak{m}^r$ for some nonnegative integer r.

Stretched Cohen–Macaulay local rings were shown to be good in [Croll et al. 2018, Corollary 5.6]. We outline a different method. In statement (2) of the following corollary, we find a more efficient common denominator of Poincaré series of modules over such rings.

Corollary 3.12. *Let $(R, \mathfrak{m}, k)$ be a d-dimensional stretched Cohen–Macaulay local ring and M be an R-module. Let $n = \dim_k \mathfrak{m}/\mathfrak{m}^2$ and $r = \dim_k \operatorname{Ext}^d_R(k, R)$ denote the type of R. Then the following hold.*

(1) *If $r = n - d$, then R is a Golod ring, $P^R_k(t) = (1+t)^d/(1-(n-d)t)$ and $(1-(n-d)t)P^R_M(t) \in \mathbb{Z}[t]$.*
(2) *If $r \neq n-d$, then $(1+t)^{n-d-r+1}(1-(n-d)t+t^2)P^R_M(t) \in \mathbb{Z}[t]$ and $P^R_k(t) = (1+t)^d/(1-(n-d)t+t^2)$.*
(3) *If $\operatorname{Ext}^i(M, M \oplus R) = 0$ for all $i \geq 1$, then M is a free R-module.*

Proof. To prove statements (1) and (2), it is enough to assume that R is a stretched Artinian ring, i.e., $d = 0$ (see Lemma 3.9). We have $\operatorname{edim}(R) = n$ and $\dim_k \operatorname{socle}(R) = r$. Let $\mathfrak{m} = (x_1, \dots, x_n)$ and $\operatorname{ll}(R) = s$. The ideals $\mathfrak{m}^i$, $i \geq 2$, are principal ideals. If $\operatorname{socle}(R) \subset \mathfrak{m}^2$, then $\operatorname{socle}(R) = \mathfrak{m}^s$ is a principal ideal, so R is a Gorenstein ring. Both statements (1) and (2) follow from Theorem II.

Otherwise assume that $x_1 \in \operatorname{socle}(R) \setminus \mathfrak{m}^2$. One observes that $(x_1) \cap (x_2, \dots, x_n) = 0$ and $(x_1) + (x_2, \dots, x_n) = \mathfrak{m}$. For any two ideals I, J in a ring R, we know that $R = R/I \times_{R/I+J} R/J$. Therefore, it follows that $R = R/(x_1) \times_k R/(x_2, \dots, x_n)$.

The maximal ideal of the ring $R/(x_2, \dots, x_n)$ is generated by the residue class of x_1, so its square is zero since $x_1 \in \operatorname{socle}(R)$. On the other hand, the ring $R/(x_1)$ is a stretched Artinian ring. If the socle of $R/(x_1)$ is contained in the square of its maximal ideal, it is a Gorenstein ring. Otherwise we decompose $R/(x_1)$ again as before.

After a finite number of steps, we have $R = S \times_k T$ where $(S, \mathfrak{m}_S)$ is a stretched Artinian Gorenstein ring and $(T, \mathfrak{m}_T)$ is a local ring with $\mathfrak{m}_T^2 = 0$. Clearly $r = 1+\operatorname{edim}(T)$. This implies that $\operatorname{edim}(T) = r-1$ and $\operatorname{edim}(S) = \operatorname{edim}(R) - \operatorname{edim}(T) = n - r + 1$. By Theorem II, we have

$$P_k^S(t) = \begin{cases} \frac{1}{1-t(n-r+1)+t^2} & \text{when } n \geq r+1, \\ \frac{1}{1-t} & \text{when } n = r. \end{cases}$$

On the other hand T is a Golod ring and $P_k^T(t) = \frac{1}{1-(r-1)t}$. The rational expression of $P_k^R(t)$ follows by the following computation:

$$\begin{aligned} \frac{1}{P_k^R(t)} &= \frac{1}{P_k^S(t)} + \frac{1}{P_k^T(t)} - 1 \\ &= \begin{cases} (1-t) + 1 - (r-1)t - 1 & \text{when } n = r, \\ (1 - t(n-r+1) + t^2) + 1 - (r-1)t - 1 & \text{when } n \geq r+1, \end{cases} \\ &= \begin{cases} 1 - nt & \text{when } n = r, \\ 1 - nt + t^2 & \text{when } n \geq r+1. \end{cases} \end{aligned}$$

If $n = r$, then $\operatorname{edim}(S) = 1$. This implies that S is a Golod ring. Therefore, R is also a Golod ring since a fibre product of Golod rings is Golod [Lescot 1983, Theorem 4.1].

Now we find a polynomial $d_R(t) \in \mathbb{Z}[t]$ such that $d_R(t) P_M^R(t) \in \mathbb{Z}[t]$ for any R-module M. The second syzygy of M is a direct sum of two modules, one is over S and another over T; see [Dress and Krämer 1975, Remark 3]. Therefore, it suffices to assume that $M = M_1 \oplus M_2$ where M_1 and M_2 are modules over S and T, respectively. By Theorem 3.7 we have

$$\frac{P_M^R(t)}{P_k^R(t)} = \frac{1}{P_k^R(t)} (P_{M_1}^R(t) + P_{M_2}^R(t)) = \frac{P_{M_1}^S(t)}{P_k^S(t)} + \frac{P_{M_2}^T(t)}{P_k^T(t)}.$$

We observe that

$$\frac{P_{M_2}^T(t)}{P_k^T(t)} = \frac{1}{P_k^T(t)} (1 + t P_{\operatorname{Syz}_1^T(M_2)}^T(t)) = 1 - (r-1)t + t \frac{P_{\operatorname{Syz}_1^T(M_2)}^T(t)}{P_k^T(t)}.$$

Since the square of the maximal ideal of T is zero, the first syzygy $\operatorname{Syz}_1^T(M_2)$ is a k-vector space. Therefore, $P_{M_2}^T(t)/P_k^T(t)$ is a polynomial in $\mathbb{Z}[t]$.

If $n = r$, we have $\operatorname{edim}(S) = 1$. By (1) of Theorem II, $P_{M_1}^S(t)/P_k^S(t)$ is a polynomial. Hence we conclude that $(1-nt) P_M^R(t) = P_M^R(t)/P_k^R(t) \in \mathbb{Z}[t]$ if $n = r$.

If $n \geq r+1$, we have $\operatorname{edim}(S) = n - r + 1$. By (2) of Theorem II, we have

$$(1+t)^{(n-r+1)} \frac{P_{M_1}^S(t)}{P_k^S(t)} \in \mathbb{Z}[t].$$

Hence we conclude that

$$(1+t)^{(n-r+1)}(1-nt+t^2)P_M^R(t) = (1+t)^{(n-r+1)}\frac{P_M^R(t)}{P_k^R(t)} \in \mathbb{Z}[t]$$

if $n \geq r+1$.

Therefore, both statements (1) and (2) follow. To prove statement (3), it suffices to show that $\mathrm{pd}_R(M) < \infty$. If $n-d=2$, then R is either a complete intersection or a Golod ring; see [Scheja 1964]. In both cases, rings are known to satisfy statement (3); see, for instance, [Avramov and Buchweitz 2000, Theorem 4.2] when R is a complete intersection and [Jorgensen and Şega 2004, Proposition 1.4] when R is a Golod ring. If $n-d>2$, then we see at once that $\mathrm{pd}_R(M) < \infty$ from Proposition 3.10. Here one observes that if the injective dimension of $M \oplus R$ is finite then R is Gorenstein and both projective and injective dimensions of M are finite. □

The following result follows from Lemma 3.9 and Theorem II.

Corollary 3.13. *Let* $(R, \mathfrak{m}, k)$ *be an almost stretched Gorenstein local ring of dimension* d *and embedding dimension* n. *Let* M *be a finitely generated* R*-module. Then the following hold.*

(1) *If* $n-d=1$, *then* $P_k^R(t) = (1+t)^d/(1-t)$ *and* $(1-t)P_M^R(t) \in \mathbb{Z}[t]$.

(2) *If* $n-d \geq 2$, *then* $(1+t)^{n-d}(1-(n-d)t+t^2)P_M^R(t) \in \mathbb{Z}[t]$ *and* $P_k^R(t) = (1+t)^d/(1-(n-d)t+t^2)$.

(3) *If* $\mathrm{Ext}^i(M, M) = 0$ *for all* $i \geq 1$, *then* M *is a free* R*-module.*

4. Revisiting known results

In this section, we provide proofs of Examples (1) and (3) in the introduction. As the section title suggests, these examples were found by other authors. Our proofs are different and obtained using Theorem I. Further, our versions are slightly stronger; see Remark 4.7. We recall the following from [Levin and Avramov 1978, Theorem 1].

Theorem 4.1. *Let* $(R, \mathfrak{m})$ *be an Artinian Gorenstein local ring of embedding dimension* n *and* K^R *be the Koszul complex on a minimal set of generators of the maximal ideal* $\mathfrak{m}$. *Set* $\mathrm{socle}(R) = (s)$, $\bar{R} = R/sR$ *and* $K^{\bar{R}} = \bar{R} \otimes_R K^R$, *the Koszul complex of* $\bar{R}$. *Define a DG algebra structure on*

$$K^R \oplus \frac{K^R}{\mathfrak{m}K^R}[-1]$$

with multiplication $(k, \bar{l})(k', \bar{l'}) = (kk', \overline{lk'} + (-1)^{\deg(k)}\overline{kl'})$ *and differential* $\partial(k, \bar{l}) = (\partial(k) + sl, 0)$. *Then the chain map*

$$K^R \oplus \frac{K^R}{\mathfrak{m}K^R}[-1] \to K^{\bar{R}}, \qquad (k, \bar{l}) \mapsto \bar{k},$$

is a quasi-isomorphism. If

$$\overline{H} = \frac{\mathrm{H}(K^R)}{\mathrm{H}_n(K^R)} \quad \text{and} \quad \overline{K} = \frac{K^R \otimes_R k}{K^R_n \otimes_R k}[-1],$$

then $\mathrm{H}(K^{\overline{R}}) = \overline{H} \ltimes \overline{K}$.

Lemma 4.2. *Let $(R, \mathfrak{m})$ be an Artinian Gorenstein local ring of embedding dimension $n \geq 2$. Let K^R denote the Koszul complex on a minimal set of generators of $\mathfrak{m}$ and C denote the quotient of K^R defined by*

$$C_i = \begin{cases} K_i^R & \text{for } 0 \leq i \leq n-2, \\ K^R_{n-1}/\mathrm{B}_{n-1}(K^R) & \text{for } i = n-1, \\ 0 & \text{for } i = n. \end{cases}$$

Then C has a DG algebra structure. Assume that C is a Golod DG algebra with natural augmentation. Then $R/\operatorname{socle}(R)$ is a Golod ring and R satisfies assertions (1) *and* (2) *of Theorem I.*

Proof. The fact that C is a DG algebra is straightforward because $K^R_n \oplus \mathrm{B}_{n-1}(K^R)$ is a DG ideal of the Koszul algebra K^R. Let $q : K^R \twoheadrightarrow C$ be the quotient map and $\operatorname{socle}(R) = (s)$. Let $G : K^R/(\mathfrak{m}K^R) \to K^R$ and $H : C/\mathfrak{m}C \to C$ denote chain maps induced by multiplications by s on K^R, C respectively. We have $q \circ G = H \circ \bar{q}$ where $\bar{q} : K^R/(\mathfrak{m}K^R) \twoheadrightarrow C/(\mathfrak{m}C)$ is the map induced by q. Let

$$\mathfrak{c}(G) = K^R \oplus \frac{K^R}{\mathfrak{m}K^R}[-1] \quad \text{and} \quad \mathfrak{c}(H) = C \oplus \frac{C}{\mathfrak{m}C}[-1]$$

be the cones of G and H respectively. Define $\alpha : \mathfrak{c}(G) \twoheadrightarrow \mathfrak{c}(H)$ by $\alpha(k, \bar{l}) = (q(k), \bar{q}(\bar{l}))$. Both $\mathfrak{c}(G)$ and $\mathfrak{c}(H)$ have DG algebra structure and α is a surjective DG algebra homomorphism. The kernel of α is the complex

$$0 \to \frac{K^R_n}{\mathfrak{m}K^R_n} \xrightarrow{\cdot s} K^R_n \xrightarrow{\partial} \mathrm{B}_{n-1}(K^R) \to 0$$

which is exact. Thus, α are a quasi-isomorphism of DG algebras. By Theorem 4.1, $\mathfrak{c}(G)$ is quasi-isomorphic to $K^{R/\operatorname{socle}(R)}$. Therefore, to show that $R/\operatorname{socle}(R)$ is a Golod ring it is enough to prove that $\mathfrak{c}(H)$ is a Golod algebra.

Since C is a Golod algebra, there are a k-basis $\mathfrak{b}_C = \{h_\lambda\}_{\lambda \in \Lambda}$ of $\mathrm{H}_{\geq 1}(C)$ and a function (trivial Massey operation) $\mu : \bigsqcup_{i=1}^{\infty} \mathfrak{b}_C^i \to C$ such that $\mu(h_\lambda) \in \mathrm{Z}_{\geq 1}(C)$ with $\operatorname{cls}(\mu(h_\lambda)) = h_\lambda$ for all $\lambda \in \Lambda$ and

$$\partial\mu(h_{\lambda_1}, \ldots, h_{\lambda_p}) = \sum_{j=1}^{p-1} \overline{\mu(h_{\lambda_1}, \ldots, h_{\lambda_j})}\mu(h_{\lambda_{j+1}}, \ldots, h_{\lambda_p}).$$

By (2) of [Avramov 1998, Lemma 4.1.6], $\{x \in C : \partial(x) \in \mathfrak{m}^2 C\} \subset \mathfrak{m}C$. Since $\mu(h_\lambda) \in \mathfrak{m}C$, by induction on p we conclude that $\mu(h_{\lambda_1}, \ldots, h_{\lambda_p}) \in \mathfrak{m}C$.

By [Levin and Avramov 1978, Lemma 1.2], $\operatorname{socle}(R)K_i^R \subset (0:\mathfrak{m}^2)\,\mathrm{B}_i(K^R)$ for $1 \le i \le n-1$ and so we have $sC \subset (0:\mathfrak{m}^2)\,\mathrm{B}(C)$. Let $z \in \mathrm{Z}(C)$ be such that $(z,0) = \partial(y_1, \bar{y}_2)$ for $(y_1, \bar{y}_2) \in \mathfrak{c}(H)$. Then $z = \partial(y_1) + sy_2 \in \mathrm{B}(C)$. Therefore, the inclusion $C \hookrightarrow \mathfrak{c}(H)$ induces an injective map $\mathrm{H}(C) \hookrightarrow \mathrm{H}(\mathfrak{c}(H))$. By abuse of notation we write $(c,0)$ as c. It follows that $\mathfrak{b}_C = \{h_\lambda\}_{\lambda\in\Lambda}$ is a linearly independent set in $\mathrm{H}_{\ge 1}(\mathfrak{c}(H))$ and $\mu(h_{\lambda_1},\dots,h_{\lambda_p}) \in \mathfrak{m}\,\mathfrak{c}(H)$, $p \ge 1$, satisfy the properties above.

We extend $\mathfrak{b}_C$ to a basis $\mathfrak{b}_{\mathfrak{c}(H)} = \{h_\lambda\}_{\lambda\in\Lambda\sqcup\Lambda'}$ of $\mathrm{H}_{\ge 1}(\mathfrak{c}(H))$. Let $h_{\lambda'}$, $\lambda' \in \Lambda'$, be the homology class of $(c_{\lambda'}, \bar{d}_{\lambda'}) \in \mathrm{Z}_{\ge 1}(\mathfrak{c}(H))$. Now $\partial(c_{\lambda'}, \bar{d}_{\lambda'}) = 0$ implies that $\partial(c_{\lambda'}) + sd_{\lambda'} = 0$. We have $sd_{\lambda'} = \partial(e_{\lambda'})$ for some $e_{\lambda'} \in (0:\mathfrak{m}^2)C$. We write $(c_{\lambda'}, \bar{d}_{\lambda'}) = (c_{\lambda'} + e_{\lambda'}, 0) + (-e_{\lambda'}, \bar{d}_{\lambda'})$. Note that $c_{\lambda'} + e_{\lambda'} \in \mathrm{Z}(C)$. Therefore, after subtracting suitable R-linear combinations of h_λ, $\lambda \in \Lambda$, from each $h_{\lambda'}$, $\lambda' \in \Lambda'$, if necessary, we may assume that $h_{\lambda'}$ is a homology class of some cycle in $(0:\mathfrak{m}^2)C \oplus \frac{C}{\mathfrak{m}C}$.

We define $\mu(h_{\lambda'})$, $\lambda' \in \Lambda'$, to be an element in $(0:\mathfrak{m}^2)C \oplus \frac{C}{\mathfrak{m}C}$ whose homology class is $h_{\lambda'}$. We extend μ from $\mathfrak{b}_C^i$ to a Massey operation on $\mathfrak{b}_{\mathfrak{c}(H)}^i$, $i > 1$, such that $\mu : \mathfrak{b}_{\mathfrak{c}(H)}^i \setminus \mathfrak{b}_C^i \to (0:\mathfrak{m}^2)C$, $i > 1$, by induction on i. Note that $\overline{\mu(h_\lambda)}\mu(h_{\lambda'}) \in sC \in (0:\mathfrak{m}^2)\,\mathrm{B}(C)$ for $(\lambda,\lambda') \notin \Lambda \times \Lambda$. We choose $\mu(h_\lambda, h_{\lambda'}) \in (0:\mathfrak{m}^2)C$ such that $\partial(\mu(h_\lambda, h_{\lambda'})) = \overline{\mu(h_\lambda)}\mu(h_{\lambda'})$. Now assume $\mu(h_{\delta_1},\dots,h_{\delta_i}) \in (0:\mathfrak{m}^2)C$, with $(h_{\delta_1},\dots,h_{\delta_i}) \in \mathfrak{b}_{\mathfrak{c}(H)}^i \setminus \mathfrak{b}_C^i$ satisfying the desired relations, are constructed for all $i \le p$. We choose $(h_{\delta_1},\dots,h_{\delta_{p+1}}) \in \mathfrak{b}_{\mathfrak{c}(H)}^{p+1} \setminus \mathfrak{b}_C^{p+1}$ and observe that $\sum_{j=1}^{p} \overline{\mu(h_{\delta_1},\dots,h_{\delta_j})}\mu(h_{\delta_{j+1}},\dots,h_{\delta_{p+1}})$ is an element in sC. Therefore, we can choose $\mu(h_{\delta_1},\dots,h_{\delta_{p+1}}) \in (0:\mathfrak{m}^2)C$ such that $\partial(\mu(h_{\delta_1},\dots,h_{\delta_{p+1}})) = \sum_{j=1}^{p} \overline{\mu(h_{\delta_1},\dots,h_{\delta_j})}\mu(h_{\delta_{j+1}},\dots,h_{\delta_{p+1}})$. Thus by induction μ extends to a trivial Massey operation on $\mathfrak{b}_{\mathfrak{c}(H)}$. Therefore, $\mathfrak{c}(H)$ is a Golod algebra and the result follows. □

We recall the definition of compressed Gorenstein local rings.

Definition 4.3. Let $(R,\mathfrak{m},k)$ be an Artinian Gorenstein local ring of Loewy length s and embedding dimension $n \ge 2$. Set

$$\varepsilon_i = \min\left\{ \binom{n-1+s-i}{n-1}, \binom{n-1+i}{n-1} \right\}$$

for all i with $0 \le i \le s$. Then it is shown in [Rossi and Şega 2014, Proposition 4.2] that $l(R) \le \sum_{i=0}^{n} \varepsilon_i$. The ring R is called a compressed Gorenstein ring if equality holds.

We provide a different proof of the result of Rossi and Şega [2014, Theorem 5.1] in Theorem 4.4.

Theorem 4.4. *Let R be a compressed Gorenstein local ring such that* $\operatorname{edim}(R) = n \ge 2$ *and* $\operatorname{ll}(R) = s$, $s \ge 2$, $s \ne 3$. *Then $R/\operatorname{socle}(R)$ is a Golod ring. Consequently, R satisfies assertions* (1) *and* (2) *of Theorem I.*

Proof. We follow notation as set in the proof of Theorem I and Lemma 4.2. Let $t = \max\{i : I \subset \mathfrak{n}^i\}$. It is proved in [Rossi and Şega 2014, Proposition 4.2] that $t = \lceil \frac{s+1}{2} \rceil$, the least integer not less than $\frac{s+1}{2}$. By [Rossi and Şega 2014, Lemma 1.4], the map $\mathrm{H}_{\geq 1}(R/\mathfrak{m}^t \otimes_Q K^Q) \to \mathrm{H}_{\geq 1}(R/\mathfrak{m}^{t-1} \otimes_Q K^Q)$ induced by the surjection $R/\mathfrak{m}^t \twoheadrightarrow R/\mathfrak{m}^{t-1}$ is a zero map. This implies that $\mathrm{Z}_{\geq 1}(K^R) \subset \mathrm{B}_{\geq 1}(K^R) + \mathfrak{m}^{t-1} K^R_{\geq 1}$ and therefore $\mathrm{Z}_{\geq 1}(C) \subset \mathrm{B}_{\geq 1}(C) + \mathfrak{m}^{t-1} C$. Thus we find a basis $\mathfrak{b} = \{h_\lambda\}_{\lambda \in \Lambda}$ of $\mathrm{H}_{\geq 1}(C)$ represented by cycles in $\mathfrak{m}^{t-1} C$.

Lemma 4.4 in [Rossi and Şega 2014] proves that the map $\psi : \mathrm{H}_{<n}(\mathfrak{m}^{r+1} K^R) \to \mathrm{H}_{<n}(\mathfrak{m}^r K^R)$ induced by the inclusion $\mathfrak{m}^{r+1} \hookrightarrow \mathfrak{m}^r$ is zero for $r = s + 1 - t$. Since, $s \geq 2$, $s \neq 3$, we have $t - 1 \leq r \leq r + 1 \leq 2(t - 1)$. This implies that the map $\mathrm{H}_{<n}(\mathfrak{m}^{2(t-1)} K^R) \to \mathrm{H}_{<n}(\mathfrak{m}^{t-1} K^R)$ is also zero since it factors through ψ. Therefore, we have $\mathrm{Z}_{<n}(\mathfrak{m}^{2t-2} K^R) \subset \mathrm{B}(\mathfrak{m}^{t-1} K^R)$ which implies $\mathrm{Z}(\mathfrak{m}^{2t-2} C) \subset \mathrm{B}(\mathfrak{m}^{t-1} C)$. It is worth pointing out that both the lemmas used here are independent of all other results in [Rossi and Şega 2014].

We construct inductively a trivial Massey operation $\mu : \bigsqcup_{i=1}^{\infty} \mathfrak{b}^i \to \mathfrak{m}^{t-1} C$. Define $\mu(h_\lambda)$ to be a cycle in $\mathfrak{m}^{t-1} C$ such that the homology class of $\mu(h_\lambda)$ is h_λ. Now $\overline{\mu(h_\lambda)} \mu(h_{\lambda'}) \in \mathrm{Z}(\mathfrak{m}^{2t-2} C) \subset \mathrm{B}(\mathfrak{m}^{t-1} C)$, and so we choose $\mu(h_\lambda, h_{\lambda'}) \in \mathfrak{m}^{t-1} C$ such that $\partial(\mu(h_\lambda, h_{\lambda'})) = \overline{\mu(h_\lambda)} \mu(h_{\lambda'})$. The method carries over to the next steps of construction. Thus, C is a Golod DG algebra and the result follows from Lemma 4.2. □

Lemma 4.5. *Let $(R, \mathfrak{m}, k)$ be an Artinian Gorenstein local ring but not a complete intersection. Let $\eta : Q \twoheadrightarrow R$ be a minimal Cohen presentation of R and $I = \ker(\eta)$. Assume that $\mu(I) = r$ and $\mathrm{edim}(R) = n \leq 3$. Then $R/\mathrm{socle}(R)$ is a Golod ring, so R satisfies both assertions (1) and (2) of Theorem I. If $d_R(t) = 1 - rt^2 - rt^3 + t^5$, then for any finitely generated R-module M, we have $d_R(t) P_M^R(t) \in Z[t]$. The Poincaré series of k is given by*

$$P_k^R(t) = \frac{(1+t)^n}{1 - rt^2 - rt^3 + t^5}.$$

Proof. As before, we follow notation as set in the proof of Theorem I and Lemma 4.2. By [Wiebe 1969, Satz 7], we have $\mathrm{H}_1(K^R)^2 = 0$ giving $\mathrm{H}_1(C)^2 = 0$. For a proof written in English, we refer to [Bruns and Herzog 1993, Corollary 3.4.8]. Now C is a DG algebra of length 2. Therefore, any basis of $\mathrm{H}_{\geq 1}(C)$ admits a trivial Massey operation and so C is a Golod algebra. The first part of the result follows from Lemma 4.2.

We compute the denominator. The Koszul complex of R is of length 3. We see $\dim_k \mathrm{H}_0(K^R) = 1$, $\dim_k \mathrm{H}_1(K^R) = \mu(I) = r$, $\dim_k \mathrm{H}_3(K^R) = \dim_k \mathrm{socle}(R) = 1$ and so $\dim_k \mathrm{H}_2(K^R) = r$ since $\sum_{i \geq 0} (-1)^i \dim_k \mathrm{H}_i(K^R)$ must be zero. With the notation used in Theorem I, we have $P_R^Q(t) = \sum_{i=0}^{3} \dim_k \mathrm{H}_i(K^R) = 1 + rt + rt^2 + t^3$. Therefore, we have

$$d_R(t) = 1 - t(P_R^Q(t) - 1) + t^{n+1}(1+t) = 1 - t^2(r + rt + t^2) + t^4(1+t) = 1 - rt^2 - rt^3 + t^5.$$

The formula for $P_k^R(t)$ follows from in [Rossi and Şega 2014, Proposition 6.2]. □

If $(R, \mathfrak{m}, k)$ is a Gorenstein local ring satisfying the hypothesis of the theorem below, then the rational expression of $P_k^R(t)$ was computed by Wiebe [1969, Satz 9]. It was proved in [Avramov et al. 1988, Theorem 6.4] that if $(R, \mathfrak{m})$ is any Artinian ring such that $\operatorname{edim}(R) - \operatorname{depth}(R) \leq 3$, then all R-modules have rational Poincaré series. We prove a weaker version.

Theorem 4.6. *Let $(R, \mathfrak{m}, k)$ be a Gorenstein local ring but not a complete intersection such that $\operatorname{edim}(R) - \operatorname{depth}(R) = n \leq 3$. Let $\eta : Q \twoheadrightarrow \hat{R}$ be a minimal Cohen presentation of R, $\ker \eta = I$ and $\mu(I) = r$. Then for any $f \in I \setminus \mathfrak{n}I$, the induced map $Q/(f) \twoheadrightarrow \hat{R}$ is a Golod homomorphism.*

Let $d_R(t) = 1 - rt^2 - rt^3 + t^5$. Then $P_k^R(t) = (1+t)^{\operatorname{edim}(R)}/d_R(t)$ and for any R-module M we have $d_R(t) P_M^R(t) \in \mathbb{Z}[t]$.

Proof. Let $\dim(R) = \operatorname{depth}(R) = d$ and the maximal ideal $\mathfrak{m}$ be minimally generated by $x_1, \dots, x_e$ such that $x_1, \dots, x_d$ form an R-sequence. Then $S = R/(x_1, \dots, x_d)$ is an Artinian Gorenstein local ring and $\operatorname{edim}(S) = e - d = n \leq 3$. Let $K^R = R\langle X_i : \partial(X_i) = x_i, 1 \leq i \leq n\rangle$ and $K^S = S\langle X_i : \partial(X_i) = x_i, d+1 \leq i \leq n\rangle$ denote the Koszul complexes. The quotient map $q : K^R \twoheadrightarrow K^S$ is a quasi-isomorphism; see [Avramov 1998, Lemma 4.1.6].

To prove that the map $Q/(f) \twoheadrightarrow \hat{R}$ is a Golod homomorphism for any $f \in I \setminus \mathfrak{n}I$, it is equivalent to show that for any cycle $z \in \mathrm{Z}_1(K^R) \setminus \mathrm{B}_1(K^R)$, the semifree extension $K^R\langle T \mid \partial(T) = z\rangle$ is a Golod algebra. Now one observes that q extends to a surjective quasi-isomorphism

$$\tilde{q} : K^R\langle T \mid \partial(T) = z\rangle \twoheadrightarrow K^S\langle T \mid \partial(T) = q(z)\rangle;$$

see, for instance, [Gulliksen and Levin 1969, Proposition 1.3.5]. Lemma 4.5 applies to S. We conclude that the image of $\tilde{q}$ is a Golod algebra. Therefore, $K^R\langle T \mid \partial(T) = z\rangle$ is a Golod algebra.

The statement about Poincaré series is an immediate consequence of Lemmas 3.9 and 4.5. □

Remark 4.7. In both Theorems 4.4 and 4.6, we constructed a Golod homomorphism from a hypersurface ring which is a quotient of an arbitrary generator belonging to a minimal generating set of the defining ideal. Thus both theorems are slightly stronger than their earlier versions.

Acknowledgements

Work on this article started as a result of discussions with Liana M. Şega during the conference in honour of Craig Huenke held in Ann Arbor in July 2016. I would like to thank her for clarifications on her results about the Auslander–Reiten conjecture. I sincerely thank the anonymous referee for suggesting the proof of Lemma 3.1. I owe many thanks to Matteo Varbaro for his inspiration and encouragement throughout. This research is supported by the INdAM-COFUND-2012 Fellowship cofounded by Marie Curie actions, Italy.

References

[Ananthnarayan 2009] H. Ananthnarayan, *Approximating Artinian rings by Gorenstein rings and 3-standardness of the maximal ideal*, Ph.D. thesis, University of Kansas, 2009, available at https://search.proquest.com/docview/304917043.

[Ananthnarayan et al. 2012] H. Ananthnarayan, L. L. Avramov, and W. F. Moore, "Connected sums of Gorenstein local rings", *J. Reine Angew. Math.* **667** (2012), 149–176. MR Zbl

[Ananthnarayan et al. 2019] H. Ananthnarayan, E. Celikbas, J. Laxmi, and Z. Yang, "Decomposing Gorenstein rings as connected sums", *J. Algebra* **527** (2019), 241–263. MR Zbl

[Anick 1982] D. J. Anick, "A counterexample to a conjecture of Serre", *Ann. of Math.* (2) **115**:1 (1982), 1–33. MR Zbl

[Auslander and Reiten 1975] M. Auslander and I. Reiten, "On a generalized version of the Nakayama conjecture", *Proc. Amer. Math. Soc.* **52** (1975), 69–74. MR Zbl

[Avramov 1994] L. L. Avramov, "Local rings over which all modules have rational Poincaré series", *J. Pure Appl. Algebra* **91**:1-3 (1994), 29–48. MR Zbl

[Avramov 1998] L. L. Avramov, "Infinite free resolutions", pp. 1–118 in *Six lectures on commutative algebra* (Bellaterra, 1996), edited by J. M. Giral et al., Progr. Math. **166**, Birkhäuser, Basel, 1998. MR Zbl

[Avramov and Buchweitz 2000] L. L. Avramov and R.-O. Buchweitz, "Support varieties and cohomology over complete intersections", *Invent. Math.* **142**:2 (2000), 285–318. MR Zbl

[Avramov and Golod 1971] L. L. Avramov and E. S. Golod, "The homology of algebra of the Koszul complex of a local Gorenstein ring", *Mat. Zametki* **9** (1971), 53–58. MR Zbl

[Avramov et al. 1988] L. L. Avramov, A. R. Kustin, and M. Miller, "Poincaré series of modules over local rings of small embedding codepth or small linking number", *J. Algebra* **118**:1 (1988), 162–204. MR Zbl

[Bøgvad 1983] R. Bøgvad, "Gorenstein rings with transcendental Poincaré-series", *Math. Scand.* **53**:1 (1983), 5–15. MR Zbl

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1993. MR Zbl

[Croll et al. 2018] A. Croll, R. Dellaca, A. Gupta, J. Hoffmeier, V. Mukundan, L. M. Şega, G. Sosa, P. Thompson, and D. R. Tracy, "Detecting Koszulness and related homological properties from the algebra structure of Koszul homology", *Nagoya Math. J.* (online publication June 2018).

[Dress and Krämer 1975] A. Dress and H. Krämer, "Bettireihen von Faserprodukten lokaler Ringe", *Math. Ann.* **215** (1975), 79–82. MR Zbl

[Eakin and Sathaye 1976] P. Eakin and A. Sathaye, "Prestable ideals", *J. Algebra* **41**:2 (1976), 439–454. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR Zbl

[Elias and Valla 2008] J. Elias and G. Valla, "Structure theorems for certain Gorenstein ideals", *Michigan Math. J.* **57** (2008), 269–292. MR Zbl

[Elias and Valla 2009] J. Elias and G. Valla, "A family of local rings with rational Poincaré series", *Proc. Amer. Math. Soc.* **137**:4 (2009), 1175–1178. MR Zbl

[Gulliksen 1974] T. H. Gulliksen, "A change of ring theorem with applications to Poincaré series and intersection multiplicity", *Math. Scand.* **34** (1974), 167–183. MR Zbl

[Gulliksen and Levin 1969] T. H. Gulliksen and G. Levin, *Homology of local rings*, Queen's Paper in Pure and Applied Mathematics **20**, Queen's University, Kingston, ON, 1969. MR Zbl

[Jorgensen and Şega 2004] D. A. Jorgensen and L. M. Şega, "Nonvanishing cohomology and classes of Gorenstein rings", *Adv. Math.* **188**:2 (2004), 470–490. MR Zbl

[Lescot 1983] J. Lescot, "La série de Bass d'un produit fibré d'anneaux locaux", pp. 218–239 in *Paul Dubreil and Marie-Paule Malliavin algebra seminar, 35th year* (Paris, 1982), edited by M.-P. Malliavin, Lecture Notes in Math. **1029**, Springer, 1983. MR Zbl

[Levin 1976] G. Levin, *Lectures on Golod homomorphisms*, Matematiska institutionen, Stockholms universitet **15**, University of Stockholm, 1976. Zbl

[Levin 1985] G. Levin, "Modules and Golod homomorphisms", *J. Pure Appl. Algebra* **38**:2-3 (1985), 299–304. MR Zbl

[Levin and Avramov 1978] G. L. Levin and L. L. Avramov, "Factoring out the socle of a Gorenstein ring", *J. Algebra* **55**:1 (1978), 74–83. MR Zbl

[Roos 2005] J.-E. Roos, "Good and bad Koszul algebras and their Hochschild homology", *J. Pure Appl. Algebra* **201**:1-3 (2005), 295–327. MR Zbl

[Rossi and Şega 2014] M. E. Rossi and L. M. Şega, "Poincaré series of modules over compressed Gorenstein local rings", *Adv. Math.* **259** (2014), 421–447. MR Zbl

[Sally 1980] J. D. Sally, "The Poincaré series of stretched Cohen–Macaulay rings", *Canadian J. Math.* **32**:5 (1980), 1261–1265. MR Zbl

[Scheja 1964] G. Scheja, "Über die Bettizahlen lokaler Ringe", *Math. Ann.* **155** (1964), 155–172. MR Zbl

[Şega 2003] L. M. Şega, "Vanishing of cohomology over Gorenstein rings of small codimension", *Proc. Amer. Math. Soc.* **131**:8 (2003), 2313–2323. MR Zbl

[Wiebe 1969] H. Wiebe, "Über homologische Invarianten lokaler Ringe", *Math. Ann.* **179** (1969), 257–274. MR Zbl

Received January 30, 2018. Revised September 21, 2019.

ANJAN GUPTA
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH BHOPAL
BHOPAL
INDIA
agmath@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.189

GENERALIZED CARTAN MATRICES ARISING FROM NEW DERIVATION LIE ALGEBRAS OF ISOLATED HYPERSURFACE SINGULARITIES

Naveed Hussain, Stephen S.-T. Yau and Huaiqing Zuo

Dedicated to professor Shing Tung Yau on the occasion of his 70th birthday

Let V be a hypersurface with an isolated singularity at the origin defined by the holomorphic function $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$. The Yau algebra $L(V)$ is defined to be the Lie algebra of derivations of the moduli algebra $A(V) := \mathcal{O}_n/\big(f, \frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\big)$, i.e., $L(V) = \mathrm{Der}(A(V), A(V))$. It is known that $L(V)$ is finite dimensional and its dimension $\lambda(V)$ is called the Yau number. We introduced a new Lie algebra $L^*(V)$ which was defined to be the Lie algebra of derivations of

$$A^*(V) = \mathcal{O}_n \Big/ \left(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}, \mathrm{Det}\left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n} \right),$$

i.e., $L^*(V) = \mathrm{Der}(A^*(V), A^*(V))$. $L^*(V)$ is finite dimensional and $\lambda^*(V)$ is the dimension of $L^*(V)$. In this paper we compute the generalized Cartan matrix $C(V)$ and other various invariants arising from the new Lie algebra $L^*(V)$ for simple elliptic singularities and simple hypersurface singularities. We use the generalized Cartan matrix to characterize the ADE singularities.

1. Introduction

Recall that simple (Kleinian, rational double point) singularities, consist of two series $A_k : \{x^2 + y^2 - z^{k+1} = 0\}$, $k \geq 1$, $D_k : \{x^2 + y^2 z + z^{k-1} = 0\}$, $k \geq 4$ and three exceptional singularities E_6, E_7, E_8 defined by polynomials

$$x^2 + y^3 + z^4, \quad x^2 + y^3 + yz^3, \quad x^2 + y^3 + z^5,$$

respectively. $\widetilde{E}_6$ is a simple elliptic singularity defined by

$$\{(x, y, z) \in \mathbb{C}^3 \mid x^3 + y^3 + z^3 = 0\}.$$

Both Yau and Zuo are supported by NSFC Grants 11961141005 and 11531007. Zuo is supported by NSFC Grant 11771231. Yau is supported by Tsinghua university start-up fund and Tsinghua university education foundation fund (042202008), He is grateful to the National Center for Theoretical Sciences (NCTS) for providing an excellent research environment while part of this research was done.
MSC2010: 14B05, 32S05.
Keywords: isolated singularity, Lie algebra, generalized Cartan matrix.

Its (μ, τ)-constant family is given by

$$V_t = \big\{(x, y, z) \in \mathbb{C}^3 \mid f_t(x, y, z) = x^3 + y^3 + z^3 + txyz = 0\big\},$$

with $t^3 + 27 \neq 0$ (see [Yau 1983]). $\widetilde{E}_7$ is a simple elliptic singularity defined by $\{(x, y, z) \in \mathbb{C}^3 \mid x^4 + y^4 + z^2 = 0\}$. Seeley and Yau [1990] showed that its (μ, τ)-constant family is given by

$$V_t = \big\{(x, y, z) \in \mathbb{C}^3 \mid f_t(x, y, z) = x^4 + y^4 + tx^2y^2 + z^2 = 0\big\},$$

with $t^2 \neq 4$. The simple elliptic singularity $\widetilde{E}_8$ defined by

$$\{(x, y, z) \in \mathbb{C}^3 \mid x^6 + y^3 + z^2 = 0\}.$$

In [Seeley and Yau 1990], the authors had studied the (μ, τ)-constant family of $\widetilde{E}_8$, which is given by

$$V_t = \{(x, y, z) \in \mathbb{C}^3 \mid f_t = x^6 + y^3 + z^2 + tx^4y = 0\},$$

with $4t^3 + 27 \neq 0$.

Finite dimensional Lie algebras are the semidirect product of the semisimple Lie algebras and solvable Lie algebras. Brieskorn gave the connection between simple Lie algebras and simple singularities. The Lie algebra L is called nilpotent if the lower central series of ideals: $L_0 = L,\ L_1 = [L, L],\ L_i = [L, L_{(i-1)}],\ i = 2, 3, \ldots$ terminates. Simple Lie algebras have been well understood, but not the nilpotent Lie algebras.

The problem of classifying nilpotent Lie algebra was studied for the first time by Umlauf [1891], a student of Engle. Umlauf gave the complete list over $\mathbb{C}$ up to dimension 6 and a certain complex family at dimensions 7, 8 and 9. The introduction of the root systems for the nilpotent Lie algebras was given by Bratzlavsky [1974] and Favre [1972; 1973]. The concept of root system constitutes an important step in the classification of nilpotent Lie algebras. By using these root systems, Santharoubane [1983] established a link between the nilpotent Lie algebras and the Kac–Moody Lie algebras (which generalize the semisimple Lie algebras and are of infinite dimension). Yau [1986], introduced many numerical invariants, namely, dimension of the Lie algebra $L(V)$; dimension of the maximal nilpotent subalgebras (i.e., nilradical of nilpotent Lie algebra) $g(V)$ of $L(V)$; dimension of the maximal torus of $g(V)$; generalized Cartan matrix $C(V)$ (see Definition 2.6); type and nilpotency of singularity. Benson and Yau [1987] computed the generalized Cartan matrix $C(V)$ for simple hypersurface singularities by using the Yau algebra. Moreover, Seeley and Yau [1991] computed the generalized Cartan matrix $C(V)$ by using Yau algebras of simple elliptic singularities. Thus it is extremely important to establish connection between singularities and nilpotent Lie algebras.

It is well-known that for any isolated hypersurface singularity $(V, 0) \subset (\mathbb{C}^n, 0)$, where $V = V(f) = \{f = 0\}$, one can consider the finite dimensional moduli algebra

$$A(V) := \mathcal{O}_n \Big/ \left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right).$$

The $\mathcal{O}_n$ is the algebra of convergent power series in n indeterminates and $f \in \mathcal{O}_n$. Mather and Yau [1982] proved that the complex structure of $(V, 0)$ is determined by its moduli algebra. Subsequently, Yau [1986] introduced the Lie algebra $L(V)$ to $(V, 0)$, which is the algebra of derivations of $A(V)$, i.e., $L(V) := \mathrm{Der}(A(V), A(V))$. Yau and his collaborators have been systematically studied the Lie algebras of isolated hypersurface singularities since the 1980s [Yau 1983; 1984; 1986; 1991; Benson and Yau 1987; 1990; Seeley and Yau 1990; 1991, Yau and Zuo 2016a; 2016b; Chen et al. 1995; 2019; ≥ 2020a; ≥ 2020b; Hussain et al. 2018; 2020; 2019a; 2019b; Hussain 2018]. One can construct nilpotent Lie algebras from the Yau algebras, however the Yau algebras can not be used to distinguish the ADE singularities [Elashvili and Khimshiashvili 2006]. Recently, in [Chen et al. ≥ 2020b], a new natural connection between the set of complex analytic isolated hypersurface singularities and the set of finite dimensional nilpotent Lie algebras has been constructed. We introduced a new Lie algebra $L^*(V) := \mathrm{Der}(A^*(V), A^*(V))$, to be the Lie algebra of derivations of the Artinian algebra

$$A^*(V) = \mathcal{O}_n \Big/ \left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}, \mathrm{Det}\left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\ldots,n} \right),$$

and $\lambda^*(V)$ is the dimension of $L^*(V)$. In [Chen et al. ≥ 2020b], we have used it to distinguish ADE singularities and prove Torelli-type theorems for some simple elliptic singularities. This new Lie algebra is a subtle invariant associated to an isolated hypersurface singularity. In this paper we shall study the new Lie algebra $L^*(V)$ for simple hypersurface singularities and simple elliptic singularities. We shall introduce many numerical invariants, namely, dimension of the maximal nilpotent subalgebras (i.e., nilradical of nilpotent Lie algebra) $g(V)$ of $L^*(V)$; dimension of the maximal torus of $g(V)$; generalized Cartan matrix $C(V)$, etc. We shall compute different numerical invariants such as dimension of maximal torus of $g(V)$; type and nilpotency of singularity and generalized Cartan matrix $C(V)$ and so on. We use the generalized Cartan matrix to characterize the ADE singularities with one exceptional case (i.e, A_6 and D_5) and obtain the following result.

Theorem 1.1. *The generalized Cartan matrix characterizes the simple* (*ADE*) *hypersurface singularities except A_6 and D_5 singularities. I.e.*, *if X and Y are two simple hypersurface singularities*, *then $C(X) = C(Y)$ if and only if X and Y are analytically isomorphic.*

2. Preliminaries

2A. ***Isolated hypersurface singularities.*** Let $\mathbb{C}[x_1, x_2, \ldots, x_n]$ be the algebra of complex polynomials in n indeterminates. Denote by $\mathcal{O}_n$ the algebra of germs of holomorphic functions at the origin of $\mathbb{C}^n$. Obviously, $\mathcal{O}_n$ can be naturally identified with the algebra of convergent power series in n indeterminates with complex coefficients. For a polynomial $f \in \mathbb{C}[x_1, x_2, \ldots, x_n]$, denote by $V = V(f)$ the germ at the origin of $\mathbb{C}^n$ of hypersurface $\{f = 0\} \subset \mathbb{C}^n$. We say that V is a germ of isolated hypersurface singularity if the origin is an isolated zero of the gradient of f. The local (function) algebra of V is defined as the (commutative associative) algebra $F(V) \cong \mathcal{O}_n/(f)$, where (f) is the principal ideal generated by the germ of f at the origin. According to Hilbert's Nullstellensatz for an isolated singularity $V = V(f) = \{f = 0\}$ the factor-algebra $A(V) = \mathcal{O}_n/\big(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\big)$ is finite dimensional. This factor-algebra is called the moduli algebra of V and its dimension $\tau(V)$ is called Tjurina number. The well-known Mather–Yau theorem states that

Theorem 2.1 [Mather and Yau 1982]. *The analytic isomorphism type of an isolated hypersurface singularity is determined by the isomorphism class of its moduli algebras i.e.*,

$$(V_1, 0) \cong (V_2, 0) \Leftrightarrow A(V_1) \cong A(V_2).$$

2B. ***Yau algebra.*** Recall that a derivation of commutative associative algebra A is defined as a linear endomorphism D of A satisfying the Leibniz rule: $D(ab) = D(a)b + aD(b)$. Thus for such an algebra A one can consider the Lie algebra of its derivations $\operatorname{Der}(A, A)$ with the bracket defined by the commutator of linear endomorphisms.

Definition 2.2. Let $f(x_1, \ldots, x_n)$ be a complex polynomial and $V = \{f = 0\}$ be a germ of an isolated hypersurface singularity at the origin in $\mathbb{C}^n$. Let $A(V)$ be the moduli algebra and $L(V) := \operatorname{Der}(A(V), A(V))$. Yu [1996] calls $L(V)$ the Yau algebra of V. The dimension of $L(V)$ is called the Yau number by Elashvili and Khimshiashvili [2006] and is denoted by $\lambda(V)$.

2C. ***New derivation Lie algebra.*** We recall the following beautiful theorem due to Dimca.

Theorem 2.3 [Dimca 1984]. *Two zero-dimensional isolated complete intersection singularities X and Y are isomorphic if and only if their singular subspaces* $\operatorname{Sing}(X)$ *and* $\operatorname{Sing}(Y)$ *are isomorphic.*

Remark 2.4. Let $V = V(f)$ be an isolated quasihomogeneous hypersurface singularity. Assume that X defined by $\big(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\big)$ is a zero-dimensional isolated complete intersection singularity. Then $\mathrm{Sing}(X)$ is defined by

$$\left(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}, \mathrm{Det}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\dots,n}\right).$$

Theorem 2.3 implies that in order to study the analytic isomorphism type of zero dimensional isolated complete intersection singularity X, we only need to consider the Artinian local algebra $A^*(V)$ which is the coordinate ring of $\mathrm{Sing}(X)$. Thus $A^*(V)$ is defined as the quotient

$$\mathcal{O}_n \Big/ \left(f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}, \mathrm{Det}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\dots,n}\right).$$

Combing Theorem 2.3 with the Mather–Yau theorem, we know that the $A^*(V)$ is a complete invariant of quasihomogeneous isolated hypersurface singularities (i.e., $A^*(V)$ determines and is determined by the analytic isomorphism type of the singularity). We call $A^*(V)$ the generalized moduli algebra of V. Based on this important observation, we introduced the following new invariants for isolated hypersurface singularities [Chen et al. ≥ 2020b].

Definition 2.5. Let $V = \{f = 0\}$ be a germ of isolated hypersurface singularity at the origin of $\mathbb{C}^n$ defined by $f \in \mathbb{C}[x_1, x_2, \dots, x_n]$. The new Lie algebra arising from the isolated hypersurface singularity V is defined as $L^*(V) := \mathrm{Der}(A^*(V), A^*(V))$ (we simply denote it by $\mathrm{Der}(A^*(V))$). The dimension of this new Lie algebra is denoted by $\lambda^*(V)$.

2D. ***Kac–Moody Lie algebras and isolated hypersurface singularities.*** Let $(V, 0)$ be an isolated hypersurface singularity. Let $g(V)$ be the maximal ideal of $L^*(V)$ consisting of nilpotent elements. It is follows from [Santharoubane 1983] that a generalized Cartan matrix $C(V)$, constructed from $g(V)$, is an invariant of $(V, 0)$ (see [Yau 1986]).

Definition 2.6. An $l \times l$ matrix with entries in $\mathbb{Z}$, $C = (c_{ij})$ is a generalized Cartan matrix if

(a) $$c_{ii} = 2 \quad \text{for all } i = 1, \dots, l,$$

(b) $$c_{ij} \le 0 \quad \text{for all } i, j = 1, \dots, l,\ i \ne j,$$

(c) $$c_{ij} = 0 \text{ if and only if } c_{ji} = 0 \quad \text{for all } i, j = 1, \dots, l,\ i \ne j.$$

To each generalized Cartan matrix $C(V)$, one can associate a Lie algebra $\mathrm{KM}(C)$ (called a Kac–Moody Lie algebra) defined by generators

$$\{f_1, \dots, f_l, h_1, \dots, h_l, e_1, \dots, e_l\}$$

and relations:

$$[h_i, e_j] = c_{ij} e_j, \quad [h_i, f_j] = -c_{ij} f_j, \quad \text{for all } i, j = 1, \dots, l,$$
$$[h_i, h_j] = 0, \quad \text{for all } i, j = 1, \dots, l, \quad [e_i, f_i] = h_i,$$
$$[e_i, f_j] = 0, \quad (\operatorname{ad} e_i)^{-c_{ij}+1} e_j = 0 = (\operatorname{ad} f_i)^{-c_{ij}+1} f_j, \quad \text{for all } i \neq j.$$

Let $H = \mathbb{C}h_i + \cdots + \mathbb{C}h_l$; denote $\xi_+(C)$ (resp. $\xi_-(C)$) the subalgebra of $\mathrm{KM}(C)$ generated by $\{e_1, \dots, e_l\}$ (resp. $(f_1, \dots, f_l)$) one shows that:

$$\mathrm{KM}(C) = \xi_+(C) \oplus H \oplus \xi_-(C).$$

One can also define $\xi_+(C)$ by generators $\{e_1, \dots, e_l\}$, and relations

$$(\operatorname{ad} e_i)^{-c_{ij}+1} e_j = 0 \quad \text{for all } i, j = 1, \dots, l, \ i \neq j.$$

We shall construct the generalized Cartan matrix from an isolated hypersurface singularity $(V, 0)$. Let $g(V)$ be the set of all nilpotent elements in $L^*(V)$; then $g(V)$ is the maximal nilpotent Lie subalgebra of $L^*(V)$ and $\mathrm{Der}(g(V))$ is its derivation algebra.

Definition 2.7. A torus on $g(V)$ is a commutative subalgebra of $\mathrm{Der}(g(V))$ whose elements are semisimple endomorphism. A maximal torus is a torus not contained in any other torus. The dimension of the maximal torus is called the generalized Mostow number (GMN). The GMN is an invariant of the isolated singularity $(V, 0)$.

Theorem 2.8 (Mostow's theorem [Santharoubane 1983]). *If T_1 and T_2 are maximal tori of $g(V)$, then there exists $\varphi \in \operatorname{Aut} g(V)$* (automorphism group of $g(V)$) *such that $\varphi T_1 \varphi^{-1} = T_2$.*

Let T be a maximal torus and consider the root space decomposition of $g(V)$ relative to T [Santharoubane 1983]:

$$g(V) = \sum_{\beta \in R(T)} g(V)^\beta,$$
$$g(V)^\beta = \{x \in g(V) \mid tx = \beta(t)x, \ \forall t \in T\},$$

and

$$R(T) = \{\beta \in T^* \mid g(V)^\beta \neq (0)\} \quad \text{(root system)}\ ,$$
$$R^1(T) = \big\{\beta \in R(T) \mid g(V)^\beta \not\subseteq [g(V), g(V)]\big\},$$
$$l_\beta = \dim\left(\frac{g(V)^\beta}{[g(V), g(V)] \cap g(V)^\beta}\right), \quad \text{for all } \beta \in R^1(T),$$
$$d_\beta = \dim(g(V)^\beta), \quad \beta \in R^1(T).$$

The map $\beta \mapsto d_\beta$, $R^1(T) \to \mathbb{N}^*$ gives the partition

$$R^1(T) = R^1(T)_{p_1} \cup \cdots \cup R^1(T)_{p_q}, \quad p_1 < \cdots < p_q, \quad R^1(T)_{p_i} \neq \varnothing,$$
$$R^1(T)_p = \{\beta \in R^1(T) \mid d_\beta = p\}.$$

Set $s_i = \sharp R^1(T)_{p_i}$ and $s = s_1 + \cdots + s_q$. We let $d_{\beta_i} = d_i$ and $l_{\beta_i} = l_i$.

Let $f : \{1, \dots, l\} \to \{1, \dots, s\}$ be defined by

$$f_i = \begin{cases} 1, & 1 \leqslant i \leqslant l_1, \\ 2, & l_1 < i \leqslant l_1 + l_2, \\ \vdots & \\ s, & l_1 + l_2 + \cdots + l_{s-1} < i \leqslant l. \end{cases}$$

Theorem 2.9 [Santharoubane 1983]. *For $i, j \in \{1, \dots, l\}$ $i \neq j$ let*

$$-c_{ij}(T) = \min\bigl\{-n \in N \mid (\operatorname{ad} v)^{-n+1} w = 0,\ \forall v \in g(V)^{\beta_{f(i)}},\ \forall w \in g(V)^{\beta_{f(j)}}\bigr\},$$

with $(\operatorname{ad} 0)^0 = 0$ *and let* $c_{ii}(T) = 2$ *for* $i = 1, \dots, l$. *Then*

$$C(T) = (c_{ij}(T))_{1 \le i, j \le l}$$

is a Cartan matrix.

3. Main Results

Now we apply the above theory to study the new Lie algebra $L^*(V)$ of simple hypersurface singularities and simple elliptic singularities. We use the following convention: $g^1 = [g, g], \dots, g^{p+1} = [g, g^p]$. We use N to denote the set of positive integers.

Proposition 3.1. *Let $V = \{x, y, z \in \mathbb{C}^3 \mid x^2 + y^2 - z^{k+1} = 0\}$ be the A_k singularity, $k \geq 1$. Then*

$$C(A_k) = \begin{cases} \textit{is not defined}, & k{=}1,2,3,4, \\ \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, & k{=}5, \\ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, & k{=}6, \\ \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}, & k{=}7, \\ \begin{pmatrix} 2 & -(k-5) \\ -(k-5)/2 & 2 \end{pmatrix}, & k \text{ is odd and } k \geq 9, \\ \begin{pmatrix} 2 & -(k-5) \\ -(k-4)/2 & 2 \end{pmatrix}, & k \text{ is even and } k \geq 8. \end{cases}$$

Proof. It is easy to see that $A^*(V) = \langle 1, z, z^2, \dots, z^{k-2} \rangle$ for $k \geq 2$ with the multiplication rule $z^{k-1} = 0$, and $A^*(V) = 0$ for $k = 1$. After a simple calculation we get

$$L^*(V) = \begin{cases} \langle z\partial_z, z^2\partial_z, \dots, z^{k-2}\partial_z\rangle, & k \geq 3, \\ 0, & k = 1, 2, \end{cases}$$

and

$$g(V) = \begin{cases} \langle z^2\partial_z, \dots, z^{k-2}\partial_z\rangle, & k \geq 4, \\ \langle z\partial_z\rangle, & k = 3, \\ 0, & k = 2. \end{cases}$$

It is easy to see the Cartan matrix is not defined when $1 \leq k \leq 4$.

For the A_5 singularity,

$$g(V) = \langle z^2\partial_z, z^3\partial_z\rangle.$$

By setting $x_1 = z^2\partial_z,\ x_2 = z^3\partial_z$, we get the multiplication table $[x_1, x_2] = 0$. The type of A_5 singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of A_5 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 0$. It is easy to see from ([Benson and Yau 1987]) that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\ x_1 &\mapsto x_1, & x_1 &\mapsto 0, \\ x_2 &\mapsto 0, & x_2 &\mapsto x_2. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$. Since $\dim T = 2 =$ the type of A_5, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \end{aligned}$$

(x_1, x_2) is a T-minimal system of generators. Therefore the generalized Cartan matrix is

$$C(A_5) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

For the A_6 singularity, we have the following multiplication table $[x_1, x_2] = x_3$. The type of A_6 singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of the A_6 singularity $= \min\big\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\big\} = 1$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\ x_1 &\mapsto x_1, & x_1 &\mapsto 0, \\ x_2 &\mapsto 0, & x_2 &\mapsto x_2, \\ x_3 &\mapsto x_3, & x_3 &\mapsto x_3. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$. Since $\dim T = 2 =$ the type of A_6, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_1+\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3. \end{aligned}$$

Note that (x_1, x_2) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(A_6) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

For the A_7 singularity, we have the following multiplication table:

$$[x_1, x_2] = x_3, \quad [x_1, x_3] = 2x_4.$$

The type of A_7 singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of the A_7 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 2$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\ x_1 &\mapsto x_1, & x_1 &\mapsto 0, \\ x_2 &\mapsto 0, & x_2 &\mapsto x_2, \\ x_3 &\mapsto x_3, & x_3 &\mapsto x_3, \\ x_4 &\mapsto 2x_4, & x_4 &\mapsto x_4. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$. Since $\dim T = 2 =$ the type of A_7, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_1+\beta_2} \oplus g^{2\beta_1+\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus Cx_3 \oplus \mathbb{C}x_4, \end{aligned}$$

(x_1, x_2) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(A_7) = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}.$$

For the A_k singularity $k \geq 8$,

$$g(V) = \langle z^2\partial_z, z^3\partial_z, \ldots, z^{k-2}\partial_z \rangle.$$

By setting $x_1 = z^2\partial_z, x_2 = z^3\partial_z, \ldots, x_{k-3} = z^{k-2}\partial_z$, we have the following multiplication table:

$$\begin{aligned} &[x_1, x_2] = x_3, \quad [x_1, x_3] = 2x_4, \quad \ldots, \quad [x_1, x_{k-4}] = (k-5)x_{k-3}, \\ &[x_2, x_3] = x_5, \quad [x_2, x_4] = 2x_6, \quad \ldots, \quad [x_2, x_{k-5}] = (k-7)x_{k-3}. \end{aligned}$$

The type of A_k ($k \geq 8$) singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of the A_k ($k \geq 8$) singularity is

$$\min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = k-5.$$

It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t : g(V) &\to g(V), \\ x_1 &\mapsto x_1, \\ x_2 &\mapsto 2x_2, \\ x_3 &\mapsto 3x_3, \\ &\vdots \\ x_{k-3} &\mapsto (k-3)x_{k-3}. \end{aligned}$$

It is follows from [Benson and Yau 1987] that T is the maximal torus of $g(V)$. Let $\beta : T \to \mathbb{C}$ be a linear map with $\beta(t) = 1$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{2\beta_2} \oplus \cdots \oplus g^{(k-3)\beta_1} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \cdots \oplus \mathbb{C}x_{k-3}, \end{aligned}$$

(x_1, x_2) is a T-minimal system of generators. It is noted that $(\operatorname{ad} x_1)^{k-4} x_2 = 0$, but $(\operatorname{ad} x_1)^{k-5} x_2 \neq 0$. Therefore $c_{12} = -(k-5)$. In order to compute the c_{21} we divide it into two cases.

Case 1. k is odd and $k = 2l + 7 \geq 9$, $l \geq 1$, then

$$(\operatorname{ad} x_2)^{l+1} x_1 = -(2\cdot 2-1)(2\cdot 3-1)\cdots(2l-1)x_{2l+3}, \quad (\operatorname{ad} x_2)^{l+2} x_1 = 0,$$
$$\Rightarrow c_{21} = -(l+1) = -\frac{k-5}{2}.$$

Case 2. k is even and $k = 2l + 6 \geq 8$, $l \geq 1$, then

$$(\operatorname{ad} x_2)^{l+1} x_1 = -(2\cdot 2-1)(2\cdot 3-1)\cdots(2l-1)x_{2l+3}, \quad (\operatorname{ad} x_2)^{l+2} x_1 = 0,$$
$$\Rightarrow c_{21} = -(l+1) = -\frac{k-4}{2}.$$

Therefore the generalized Cartan matrix is

$$C(A_k) = \begin{cases} \begin{pmatrix} 2 & -(k-5) \\ -\frac{k-5}{2} & 2 \end{pmatrix}, & k \text{ is odd and } k \geq 9, \\ \begin{pmatrix} 2 & -(k-5) \\ -\frac{k-4}{2} & 2 \end{pmatrix}, & k \text{ is even and } k \geq 8. \end{cases}$$

□

Proposition 3.2. *Let $V = \{x, y, z \in \mathbb{C}^3 \mid x^2 + y^2 z + z^{k-1} = 0\}$ be the D_k singularity, $k \geq 4$. Then*

$$C(D_k) = \begin{cases} \textit{is not defined}, & k=4, \\ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, & k=5, \\ \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, & k=6, \\ \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, & k=7, \\ \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -1 & 2 \end{pmatrix}, & k=8, \\ \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -(k-6) \\ 0 & 0 & -\frac{k-5}{2} & 2 \end{pmatrix}, & k \textit{ is odd and } k \geq 9, \\ \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -(k-6) \\ 0 & 0 & -\frac{k-6}{2} & 2 \end{pmatrix}, & k \textit{ is even and } k \geq 10. \end{cases}$$

Proof. It is easy to see from [Arnold et al. 2004, Theorem 13.1],

$$A^*(V) = \langle z^i,\ 0 \leq i \leq k-3,\ y \rangle.$$

After a simple calculation we get

$$L^*(V) = \langle y\partial_y, z^{k-3}\partial_y, y\partial_z, z^i\partial_z,\ 1 \leq i \leq k-3 \rangle,$$
$$g(V) = \langle z^{k-3}\partial_y, y\partial_z, z^i\partial_z,\ 2 \leq i \leq k-3,\ k \geq 5 \rangle.$$

It is easy to see, the Cartan matrix is not defined when $k = 4$. For the D_5 singularity,

$$g(V) = \langle y\partial_z, z^2\partial_y, z^2\partial_z \rangle.$$

We set $x_1 = y\partial_z,\ x_2 = z^2\partial_y$ and $x_3 = z^2\partial_z$. we have following multiplication table: $[x_1, x_2] = -x_3$. The type of D_5 singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of the A_5 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 1$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\ x_1 &\mapsto x_1, & x_1 &\mapsto 0, \\ x_2 &\mapsto 0, & x_2 &\mapsto x_2, \\ x_3 &\mapsto x_3, & x_3 &\mapsto x_3. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$. Since $\dim T = 2 =$ the type of D_5, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_1+\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3, \end{aligned}$$

(x_1, x_2) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(D_5) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

For the D_6 singularity, we have the following multiplication table: $[x_1, x_2] = -x_4$. The type of A_6 singularity $= \dim g(V)/[g(V), g(V)] = 2$. The nilpotency of A_6 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 1$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), & t_3 : g(V) &\to g(V), \\ x_1 &\mapsto x_1, & x_1 &\mapsto 0, & x_1 &\mapsto 0, \\ x_2 &\mapsto 0, & x_2 &\mapsto x_2, & x_2 &\mapsto 0, \\ x_3 &\mapsto 0, & x_3 &\mapsto 0, & x_3 &\mapsto x_3, \\ x_4 &\mapsto x_4, & x_4 &\mapsto x_4, & x_4 &\mapsto 0. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2 + \mathbb{C}t_3$. Since $\dim T = 3 =$ the type of D_6, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2, 3$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_3} \oplus g^{\beta_1+\beta_3} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4, \end{aligned}$$

(x_1, x_2, x_3) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(D_6) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

For the D_7 singularity, we have the following multiplication table: $[x_1, x_2] = -x_5$, $[x_3, x_4] = x_5$. The type of D_7 singularity $= \dim g(V)/[g(V), g(V)] = 4$. The nilpotency of the D_7 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 1$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{array}{lll} t_1 : g(V) \to g(V), & t_2 : g(V) \to g(V), & t_3 : g(V) \to g(V), \\ \quad x_1 \mapsto x_1, & \quad x_1 \mapsto 0, & \quad x_1 \mapsto 0, \\ \quad x_2 \mapsto 0, & \quad x_2 \mapsto x_2, & \quad x_2 \mapsto 0, \\ \quad x_3 \mapsto 0, & \quad x_3 \mapsto 0, & \quad x_3 \mapsto x_3, \\ \quad x_4 \mapsto x_4, & \quad x_4 \mapsto x_4, & \quad x_4 \mapsto -x_4, \\ \quad x_5 \mapsto x_5, & \quad x_5 \mapsto x_5, & \quad x_5 \mapsto 0. \end{array}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2 + \mathbb{C}t_3$ is a unique maximal torus on $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2, 3$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_3} \oplus g^{\beta_1+\beta_2-\beta_3} \oplus g^{\beta_1+\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4 \oplus \mathbb{C}x_5, \end{aligned}$$

(x_1, x_2, x_3, x_4) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(D_7) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

For the D_8 singularity, we have the following multiplication table: $[x_1, x_2] = -x_6$, $[x_3, x_4] = x_5$, $[x_3, x_5] = 2x_6$. The type of D_8 singularity $= \dim g(V)/[g(V), g(V)] = 4$. The nilpotency of the D_8 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 2$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{array}{lll} t_1 : g(V) \to g(V), & t_2 : g(V) \to g(V), & t_3 : g(V) \to g(V), \\ \quad x_1 \mapsto x_1, & \quad x_1 \mapsto 0, & \quad x_1 \mapsto 0, \\ \quad x_2 \mapsto 0, & \quad x_2 \mapsto x_2, & \quad x_2 \mapsto 0, \\ \quad x_3 \mapsto 0, & \quad x_3 \mapsto 0, & \quad x_3 \mapsto x_3, \\ \quad x_4 \mapsto x_4, & \quad x_4 \mapsto x_4, & \quad x_4 \mapsto -2x_4, \\ \quad x_5 \mapsto x_5, & \quad x_5 \mapsto x_5, & \quad x_5 \mapsto -x_5, \\ \quad x_6 \mapsto x_6, & \quad x_6 \mapsto x_6, & \quad x_6 \mapsto 0. \end{array}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2 + \mathbb{C}t_3$ is a unique maximal torus on $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2, 3$;

$$\begin{aligned} g(V) &= g^{\beta_1} \oplus g^{\beta_2} \oplus g^{\beta_3} \oplus g^{\beta_1+\beta_2-2\beta_3} \oplus g^{\beta_1+\beta_2-\beta_3} \oplus g^{\beta_1+\beta_2} \\ &= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4 \oplus \mathbb{C}x_5 \oplus Cx_6, \end{aligned}$$

(x_1, x_2, x_3, x_4) is a T-minimal system of generators. The generalized Cartan matrix is

$$C(D_8) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

For the D_k singularity $k \geq 9$,

$$g(V) = \langle y\partial_z, z^{k-3}\partial_y, z^2\partial_z, z^3\partial_z, z^4\partial_z, \dots, z^{k-3}\partial_z \rangle.$$

By setting $x_1 = z^2\partial_z,\ x_2 = z^3\partial_z,\ \dots,\ x_{k-2} = z^{k-3}\partial_z$, we have the following multiplication table:

$$\begin{aligned} &[x_1, x_2] = -x_{k-2}, \\ &[x_3, x_4] = x_5,\ \ [x_3, x_5] = 2x_6,\ \ [x_3, x_6] = 3x_7,\ \ \dots,\ \ [x_3, x_{k-3}] = (k-6)x_{k-2}, \\ &[x_4, x_5] = x_7,\ \ [x_4, x_6] = 2x_8,\ \ [x_4, x_7] = 3x_9,\ \ \dots,\ \ [x_4, x_{k-4}] = (k-8)x_{k-2}, \\ &[x_5, x_6] = x_9,\ \ [x_5, x_7] = 2x_{10},\ [x_5, x_8] = 3x_{11},\ \ \dots,\ \ [x_5, x_{k-5}] = (k-10)x_{k-2}. \end{aligned}$$

The type of D_k singularity (for $k \geq 9$) $= \dim g(V)/[g(V), g(V)] = 4$. The nilpotency of the D_k singularity (for $k \geq 9$) $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = k-6$.

Case 1. When k is odd:

It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\ x_1 &\mapsto (k-4)x_1, & x_1 &\mapsto 0 \\ x_2 &\mapsto 0, & x_2 &\mapsto (k-4)x_2, \\ x_3 &\mapsto x_3, & x_3 &\mapsto x_3, \\ x_4 &\mapsto 2x_4, & x_4 &\mapsto 2x_4, \\ x_5 &\mapsto 3x_5, & x_5 &\mapsto 3x_5, \\ &\vdots & &\vdots \\ x_{k-2} &\mapsto (k-4)x_{k-2}, & x_{k-2} &\mapsto (k-4)x_{k-2}. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$ is a unique maximal torus on $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$g(V) = g^{(k-4)\beta_1} \oplus g^{(k-4)\beta_2} \oplus g^{\beta_1+\beta_2} \oplus g^{2(\beta_1+\beta_2)} \oplus g^{3(\beta_1+\beta_2)} \oplus \cdots \oplus g^{(k-4)(\beta_1+\beta_2)}$$
$$= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4 \oplus \mathbb{C}x_5 \oplus \cdots \oplus \mathbb{C}x_{k-2},$$

(x_1, x_2, x_3, x_4) is a T-minimal system of generators. Since $(\operatorname{ad} x_3)^{k-5}x_4 = 0$, and $(\operatorname{ad} x_3)^{k-6}x_4 \neq 0$, so $c_{34} = -(k-6)$. In order to compute the c_{43}, we have $k = 2l + 7 \geq 9$;

$$(\operatorname{ad} x_4)^{l+1}x_3 = -(2\cdot 2-1)(2\cdot 3-1)\cdots(2l-1)x_{2l+5}, \quad (\operatorname{ad} x_4)^{l+2}x_3 = 0,$$
$$\Rightarrow c_{43} = -(l+1) = -\frac{k-5}{2}.$$

Therefore the generalized Cartan matrix is

$$C(D_k) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -(k-6) \\ 0 & 0 & -\frac{k-5}{2} & 2 \end{pmatrix}.$$

Case 2. When k is even:

It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V),\\ x_1 &\mapsto (k-4)^2 x_1, & x_1 &\mapsto 0,\\ x_2 &\mapsto 0, & x_2 &\mapsto (k-4)^2 x_2,\\ x_3 &\mapsto (k-4)x_3, & x_3 &\mapsto (k-4)x_3,\\ x_4 &\mapsto 2(k-4)x_4, & x_4 &\mapsto 2(k-4)x_4,\\ x_5 &\mapsto 3(k-4)x_5, & x_5 &\mapsto 3(k-4)x_5,\\ &\vdots & &\vdots\\ x_{k-2} &\mapsto (k-4)^2 x_{k-2}, & x_{k-2} &\mapsto (k-4)^2 x_{k-2}. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$ is a unique maximal torus on $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$g(V) = g^{(k-4)^2\beta_1} \oplus g^{(k-4)^2\beta_2} \oplus g^{(k-4)(\beta_1+\beta_2)} \oplus g^{2(k-4)(\beta_1+\beta_2)} \oplus g^{3(k-4)(\beta_1+\beta_2)} \oplus \cdots$$
$$\cdots \oplus g^{(k-4)^2(\beta_1+\beta_2)}$$
$$= \mathbb{C}x_1 \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4 \oplus \mathbb{C}x_5 \oplus \cdots \oplus \mathbb{C}x_{k-2},$$

(x_1, x_2, x_3, x_4) is a T-minimal system of generators. It is noted that $(\operatorname{ad} x_3)^{k-5}x_4 = 0$, but $(\operatorname{ad} x_3)^{k-6}x_4 \neq 0$. Therefore $c_{34} = -(k-6)$. In order to compute the c_{43}, we have $k = 2l + 8 \geq 9$;

$$(\operatorname{ad} x_4)^{l+1} x_3 = -(2\cdot 2-1)(2\cdot 3-1)\cdots(2l-1)x_{2l+5}, \quad (\operatorname{ad} x_4)^{l+2} x_3 = 0,$$
$$\Rightarrow c_{43} = -(l+1) = -\frac{k-6}{2}.$$

Therefore the generalized Cartan matrix is

$$C(D_k) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & -(k-6) \\ 0 & 0 & -\frac{k-6}{2} & 2 \end{pmatrix}. \qquad \square$$

Proposition 3.3. *Let* $V = \{x, y, z \in \mathbb{C}^3 \mid x^2 + y^3 + z^4 = 0\}$ *be the* E_6 *singularity. Then*

$$C(E_6) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -2 & 2 & -2 & -2 \\ 0 & -2 & 2 & 0 \\ 0 & -2 & 0 & 2 \end{pmatrix}.$$

Proof. It is easy to see that $A^*(V) = \langle 1, y, z, yz, z^2\rangle$ with the multiplication rules $y^2 = 0 = z^3 = yz^2$. We have the following basis of the new Lie algebra of the E_6 singularity:

$$L^*(V) = \langle y\partial_y, yz\partial_y, z^2\partial_y, y\partial_z, yz\partial_z, z\partial_z, z^2\partial_z\rangle.$$

The nilradical of the new Lie algebra of the E_6 singularity is spanned by

$$g(V) = \langle yz\partial_y, z^2\partial_y, y\partial_z, yz\partial_z, z^2\partial_z\rangle.$$

We set $x_1 = yz\partial_y,\ x_2 = z^2\partial_y,\ \ldots,\ x_5 = z^2\partial_z$. The multiplication table of the nilradical of the Lie algebra is given as

$$[x_1, x_3] = x_4, \quad [x_2, x_3] = -2x_1 + x_5, \quad [x_3, x_5] = 2x_4.$$

The type of E_6 singularity $= \dim g(V)/[g(V), g(V)] = 3$. The nilpotency of the E_6 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 2$. It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{array}{lll} t_1 : g(V) \to g(V), & t_2 : g(V) \to g(V), & t_3 : g(V) \to g(V), \\ x_1 \mapsto x_1, & x_1 \mapsto x_5, & x_1 \mapsto 0, \\ x_2 \mapsto 2x_2, & x_2 \mapsto -4x_2, & x_2 \mapsto -x_2, \\ x_3 \mapsto -x_3, & x_3 \mapsto 3x_3, & x_3 \mapsto x_3, \\ x_4 \mapsto 0, & x_4 \mapsto 0, & x_4 \mapsto x_4, \\ x_5 \mapsto x_5, & x_5 \mapsto 4x_1, & x_5 \mapsto 0. \end{array}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2 + \mathbb{C}t_3$. Since $\dim T = 3 =$ the type of E_6, T is the maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2, 3$;

$$\begin{aligned} g(V) &= g^{\beta_1+2\beta_2} \oplus g^{\beta_1-2\beta_2} \oplus g^{2\beta_1-4\beta_2-\beta_3} \oplus g^{-\beta_1+3\beta_2+\beta_3} \oplus g^{\beta_3} \\ &= \mathbb{C}\Big(\frac{x_1}{2} + x_5\Big) \oplus \mathbb{C}\Big(-\frac{x_1}{2} + x_5\Big) \oplus \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4, \end{aligned}$$

$\big(x_2, x_3, \frac{x_1}{2} + x_5, -\frac{x_1}{2} + x_5\big)$ is a T-minimal system of generators. The generalized Cartan matrix is

$$C(E_6) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -2 & 2 & -2 & -2 \\ 0 & -2 & 2 & 0 \\ 0 & -2 & 0 & 2 \end{pmatrix}. \qquad \square$$

Proposition 3.4. *Let $V = \{x, y, z \in \mathbb{C}^3 \mid x^2 + y^3 + yz^3 = 0\}$ be the E_7 singularity. Then*

$$C(E_7) = \begin{pmatrix} 2 & 0 & 0 & -1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -2 & -2 & -3 & 2 \end{pmatrix}.$$

Proof. It is easy to see that $A^*(V) = \langle 1, y, z, yz, y^2, z^2\rangle$ with the multiplication rules

$$y^2z = 0 = y^4 = z^3, \quad y^3 + 3z^2 = 0, \quad 6z^2 - 3y^3 = 0.$$

We have the following basis of the new Lie algebra of the E_7 singularity:

$$L^*(V) = \langle z\partial_y, yz\partial_y, y^2\partial_y, z^2\partial_y, yz\partial_z, y^2\partial_z, z^2\partial_z, 3y\partial_y + 2z\partial_z\rangle.$$

The nilradical of the new Lie algebra of E_7 singularity is spanned by

$$g(V) = \langle yz\partial_y, y^2\partial_y, z^2\partial_y, yz\partial_z, y^2\partial_z, z^2\partial_z, z\partial_y\rangle.$$

We set $x_1 = yz\partial_y$, $x_2 = y^2\partial_y$, $\ldots$, $x_7 = z\partial_y$. The multiplication table of the nilradical of the new Lie algebra is given as

$$\begin{aligned} &[x_1, x_7] = -x_3, && [x_1, x_5] = 3x_3, && [x_2, x_5] = -6x_6, \\ &[x_2, x_7] = -2x_1, && [x_4, x_7] = x_1 - x_6, && [x_4, x_5] = 3x_6, \\ &[x_5, x_7] = x_2 - 2x_4, && [x_6, x_7] = x_3. \end{aligned}$$

The type of E_7 singularity $= \dim g(V)/[g(V), g(V)] = 3$. The nilpotency of the E_7 singularity $= \min\big\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\big\} = 3$ It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$
\begin{aligned}
t_1 : g(V) &\to g(V), & t_2 : g(V) &\to g(V), \\
x_1 &\mapsto 3x_1, & x_1 &\mapsto 0, \\
x_2 &\mapsto 2x_2, & x_2 &\mapsto -10x_2 - 3x_4, \\
x_3 &\mapsto 4x_3, & x_3 &\mapsto -8x_3, \\
x_4 &\mapsto 2x_4, & x_4 &\mapsto -12x_2 - 10x_4, \\
x_5 &\mapsto x_5, & x_5 &\mapsto -8x_5 + 3x_7, \\
x_6 &\mapsto 3x_6, & x_6 &\mapsto 3x_1 - 12x_6, \\
x_7 &\mapsto x_7, & x_7 &\mapsto 4x_7.
\end{aligned}
$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2$ is a unique maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2$;

$$
\begin{aligned}
g(V) &= g^{3\beta_1} \oplus g^{2\beta_1 - 4\beta_2} \oplus g^{2\beta_1 - 16\beta_2} \oplus g^{4\beta_1 - 8\beta_2} \oplus g^{\beta_1 - 8\beta_2} \oplus g^{3\beta_1 - 12\beta_2} \oplus g^{\beta_1 + 4\beta_2} \\
&= \mathbb{C}(4x_1 + x_6) \oplus \mathbb{C}\Big(-\frac{x_2}{2} + x_4\Big) \oplus \mathbb{C}\Big(\frac{x_2}{2} + x_4\Big) \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_5 \oplus \mathbb{C}x_6 \\
&\qquad \oplus \mathbb{C}\Big(\frac{x_5}{4} + x_7\Big),
\end{aligned}
$$

$\big(x_2/2 + x_4, -x_2/2 + x_4, x_5, x_5/4 + x_7\big)$ is a T-minimal system of generators. The generalized Cartan matrix is

$$
C(E_7) = \begin{pmatrix} 2 & 0 & 0 & -1 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -2 & -2 & -3 & 2 \end{pmatrix}. \qquad \square
$$

Proposition 3.5. *Let* $V = \{x, y, z \in \mathbb{C}^3 \mid x^2 + y^3 + z^5 = 0\}$ *be the* E_8 *singularity. Then*

$$
C(E_8) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -2 & 2 & 0 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}.
$$

Proof. It is noted that $A^*(V) = \langle 1, y, z, yz, yz^2, z^2, z^3 \rangle$ with multiplication rules

$$
y^2 = 0 = z^4 = yz^3.
$$

We have the following basis of the new Lie algebra of the E_8 singularity,

$$
L^*(V) = \langle y\partial_y, yz\partial_y, yz^2\partial_y, z^3\partial_y, y\partial_z, yz\partial_z, yz^2\partial_z, z\partial_z, z^2\partial_z, z^3\partial_z \rangle.
$$

The nilradical of the new Lie algebra of the E_8 singularity is spanned by

$$g(V) = \langle yz\partial_y, yz^2\partial_y, z^3\partial_y, y\partial_z, yz\partial_z, yz^2\partial_z, z^2\partial_z, z^3\partial_z\rangle.$$

We set $x_1 = yz\partial_y$, $x_2 = yz^2\partial_y, \dots, x_8 = z^3\partial_z$. The multiplication table of nilradical of new Lie algebra is given as

$$\begin{aligned} &[x_1, x_4] = x_5, && [x_1, x_5] = x_6, && [x_1, x_7] = -x_2,\\ &[x_2, x_4] = x_6, && [x_3, x_4] = -3x_2 + x_8, && [x_4, x_7] = 2x_5,\\ &[x_4, x_8] = 3x_6, && [x_5, x_7] = x_6. \end{aligned}$$

The type of E_8 singularity $= \dim g(V)/[g(V), g(V)] = 4$. The nilpotency of the E_8 singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 2$ It is easy to see from [Benson and Yau 1987] that the torus T of $g(V)$ is spanned by

$$\begin{aligned} &t_1 : g(V) \to g(V), && t_2 : g(V) \to g(V), && t_3 : g(V) \to g(V),\\ &\quad x_1 \mapsto -2x_7, && \quad x_1 \mapsto x_1, && \quad x_1 \mapsto 0,\\ &\quad x_2 \mapsto -3x_2, && \quad x_2 \mapsto 2x_2, && \quad x_2 \mapsto 0,\\ &\quad x_3 \mapsto 0, && \quad x_3 \mapsto 0, && \quad x_3 \mapsto x_3,\\ &\quad x_4 \mapsto -3x_4, && \quad x_4 \mapsto 2x_4, && \quad x_4 \mapsto -x_4,\\ &\quad x_5 \mapsto -5x_5, && \quad x_5 \mapsto 3x_5, && \quad x_5 \mapsto -x_5,\\ &\quad x_6 \mapsto -6x_6, && \quad x_6 \mapsto 4x_6, && \quad x_6 \mapsto -x_6,\\ &\quad x_7 \mapsto x_1 - 3x_7, && \quad x_7 \mapsto x_7, && \quad x_7 \mapsto 0,\\ &\quad x_8 \mapsto -3x_8, && \quad x_8 \mapsto 2x_8, && \quad x_8 \mapsto 0. \end{aligned}$$

Thus $T = \mathbb{C}t_1 + \mathbb{C}t_2 + \mathbb{C}t_3$ is a maximal torus of $g(V)$. Let $\beta_i : T \to \mathbb{C}$ be a linear map with $\beta_i(t_j) = \delta_{ij}$ for $i, j = 1, 2, 3$;

$$\begin{aligned} g(V) &= g^{-3\beta_1+2\beta_2} \oplus g^{\beta_3} \oplus g^{-3\beta_1+2\beta_2-\beta_3} \oplus g^{-5\beta_1+3\beta_2-\beta_3} \oplus g^{-6\beta_1+4\beta_2-\beta_3}\\ &\qquad \oplus g^{-\beta_1+\beta_2} \oplus g^{-2\beta_1+\beta_2} \oplus g^{-3\beta_1+2\beta_2}\\ &= \mathbb{C}x_2 \oplus \mathbb{C}x_3 \oplus \mathbb{C}x_4 \oplus \mathbb{C}x_5 \oplus \mathbb{C}x_6 \oplus \mathbb{C}(2x_1 + x_7) \oplus \mathbb{C}(x_1 + x_7) \oplus \mathbb{C}x_8, \end{aligned}$$

$(x_3, x_4, 2x_1 + x_7, x_1 + x_7)$ is a T-minimal system of generators. The generalized Cartan matrix is

$$C(E_8) = \begin{pmatrix} 2 & -1 & 0 & 0\\ -2 & 2 & 0 & -1\\ 0 & 0 & 2 & -1\\ 0 & -1 & -1 & 2 \end{pmatrix}. \qquad \square$$

From Propositions 3.1–3.5, we have the following conclusion.

Theorem 1.1. *The generalized Cartan matrix characterizes the simple (ADE) hypersurface singularities except A_6 and D_5 singularities. I.e., if X and Y are two simple hypersurface singularities, then $C(X) = C(Y)$ if and only if X and Y are analytically isomorphic.*

Remark 3.6. In Theorem 1.1, in order for the generalized Cartan matrix to make sense, the simple singularities there should be $A_k, k \geq 5,\ D_k, k \geq 5,\ E_6, E_7, E_8$.

However, Theorem 1.1 is not true for general singularities. In the following Propositions 3.7, 3.8 and 3.9, we shall see that the generalized Cartan matrix is constant for simple elliptic singularities. Therefore it does not characterize the simple elliptic singularities.

Proposition 3.7. *Let $V_t = \{(x, y, z) \in \mathbb{C}^3 \mid f_t(x, y, z) = x^3 + y^3 + z^3 + txyz = 0\}$ with $t^3 + 27 \neq 0$, the $\widetilde{E}_6$ singularity. Then*

$$C(\widetilde{E}_6) = \begin{pmatrix} 2&0&0&0&0&0&0&0&0\\ 0&2&0&0&0&0&0&0&0\\ 0&0&2&0&0&0&0&0&0\\ 0&0&0&2&0&0&0&0&0\\ 0&0&0&0&2&0&0&0&0\\ 0&0&0&0&0&2&0&0&0\\ 0&0&0&0&0&0&2&0&0\\ 0&0&0&0&0&0&0&2&0\\ 0&0&0&0&0&0&0&0&2 \end{pmatrix}.$$

Proof. It is easy to see that new moduli algebra $A^*(V_t) = \langle 1, x, y, z, xy, xz, yz\rangle$ with multiplication rules, $x^2 = -\frac{t}{3}yz,\ y^2 = -\frac{t}{3}xz,\ z^2 = -\frac{t}{3}xy$. A basis for the new Lie algebra $L^*(V_t)$ (for $t^3 \neq 0, -27, 216$) is

$$\begin{aligned} &e_0 = x\partial_x + y\partial_y + z\partial_z, && e_1 = xy\partial_x, && e_2 = xy\partial_y, && e_3 = xy\partial_z, && e_4 = xz\partial_x,\\ &e_5 = xz\partial_y, && e_6 = xz\partial_z, && e_7 = yz\partial_x, && e_8 = yz\partial_y, && e_9 = yz\partial_z. \end{aligned}$$

For $t = 0$, $\{e_0\}$ is replaced by $\{e_{10} = x\partial_x, e_{11} = y\partial_y, e_{12} = z\partial_z\}$. For $t = -3$, $\{e_0\}$ is replaced by $\{e_0, z\partial_x + x\partial_y + y\partial_z, y\partial_x + z\partial_y + x\partial_z\}$. For $t = 6$, $\{e_0\}$ is replaced by $\{e_0, -z\partial_x - x\partial_y - y\partial_z, -y\partial_x - z\partial_y - x\partial_z\}$.

The type and nilpotency of $\widetilde{E}_6$ singularity are zero.

The nilradical $g(V)$ of $L^*(V_t)$ is spanned by $\langle e_1, \dots, e_9\rangle$ (for all t such that $t^3 \neq 0, -27, 216$).

It is easy to see that multiplication table for nilradical of the new Lie algebra is zero.

It is follows from [Seeley and Yau 1991], in case of generic t the derivation of the nilradical of $L^*(V_t)$ has a basis of the form

$$f_1 : a_{11} = 1, \quad f_2 : a_{22} = 1, \quad f_3 : a_{33} = 1, \quad f_4 : a_{44} = 1, \quad f_5 : a_{55} = 1,$$
$$f_6 : a_{66} = 1, \quad f_7 : a_{77} = 1, \quad f_8 : a_{88} = 1, \quad f_9 : a_{99} = 1.$$

We have following decomposition of the nilradical of the new Lie algebra with respect to the maximal torus [Seeley and Yau 1991]:

$$g^{\beta_1} = \langle e_1 \rangle, \quad g^{\beta_2} = \langle e_2 \rangle, \quad g^{\beta_3} = \langle e_3 \rangle, \quad g^{\beta_4} = \langle e_4 \rangle, \quad g^{\beta_5} = \langle e_5 \rangle,$$
$$g^{\beta_6} = \langle e_6 \rangle, \quad g^{\beta_7} = \langle e_7 \rangle, \quad g^{\beta_8} = \langle e_8 \rangle, \quad g^{\beta_9} = \langle e_9 \rangle.$$

Therefore we have following Cartan matrix

$$C(\widetilde{E}_6) = \begin{pmatrix} 2&0&0&0&0&0&0&0&0\\ 0&2&0&0&0&0&0&0&0\\ 0&0&2&0&0&0&0&0&0\\ 0&0&0&2&0&0&0&0&0\\ 0&0&0&0&2&0&0&0&0\\ 0&0&0&0&0&2&0&0&0\\ 0&0&0&0&0&0&2&0&0\\ 0&0&0&0&0&0&0&2&0\\ 0&0&0&0&0&0&0&0&2 \end{pmatrix}. \qquad \square$$

Proposition 3.8. *Let $V_t = \{(x, y, z) \in \mathbb{C}^3 \mid f_t(x, y, z) = x^4 + y^4 + tx^2y^2 + z^2 = 0\}$ with $t^2 \neq 4$, be the $\widetilde{E}_7$ singularity. Then*

$$C(\widetilde{E}_7) = \begin{cases} \begin{pmatrix} 2&-1&-1&-1&-1&-1\\ -1&2&-1&-1&-1&-1\\ -1&-1&2&-1&-1&-1\\ -1&-1&-1&2&-1&-1\\ -1&-1&-1&-1&2&-1\\ -1&-1&-1&-1&-1&2 \end{pmatrix}, & t \neq 0, \\ \begin{pmatrix} 2&-1&-1&0&-1&-1\\ -1&2&-1&-1&-1&-1\\ -1&-1&2&-1&-1&-1\\ 0&-1&-1&2&-1&-1\\ -1&-1&-1&-1&2&0\\ -1&-1&-1&-1&0&2 \end{pmatrix}, & t = 0. \end{cases}$$

Proof. In [Seeley and Yau 1990], the authors showed that (μ, τ)-constant family for $\widetilde{E}_7$ is

$$V_t = \{(x, y, z) \in \mathbb{C}^3 \mid f_t(x, y, z) = x^4 + y^4 + tx^2y^2 + z^2 = 0\},$$

with $t^2 \neq 4$. The moduli algebra

$$A^*(V_t) = A(V_t)/(x^2y^2) = \langle 1, x, y, x^2, xy, y^2, x^2y, xy^2 \rangle,$$

with the multiplication rules

$$x^3 = -\frac{t}{2}xy^2, \quad y^3 = -\frac{t}{2}x^2y, \quad x^3y = 0 = xy^3 = x^2y^2.$$

A basis of the new Lie algebra $L^*(V_t)$ (for $t \neq 0, -6, 6$) is:

deg0: $e_0 = x\partial_x + y\partial_y$,

deg1: $e_1 = x^2\partial_x$, $e_2 = y^2\partial_y$, $e_3 = y^2\partial_x$, $e_4 = x^2\partial_y$, $e_5 = xy\partial_x$, $e_6 = xy\partial_y$,

deg2: $e_7 = x^2y\partial_x$, $e_8 = xy^2\partial_y$, $e_9 = xy^2\partial_x$, $e_{10} = x^2y\partial_y$.

For $t = 0$, $\{e_0\}$ is replaced by $\{x\partial_x, y\partial_y\}$. For $t = 6$, $\{e_0\}$ is replaced by $\{e_0, y\partial_x + x\partial_y\}$. For $t = -6$, $\{e_0\}$ is replaced by $\{e_0, y\partial_x - x\partial_y\}$.

The type of $\widetilde{E}_7$ singularity $= \dim g(V)/[g(V), g(V)] = 4$.

The nilpotency of the $\widetilde{E}_7$ singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 1$

The nilradical $g(V)$ of $L^*(V_t)$ is of dimension 10 and is spanned by $\langle e_1, \dots, e_{10} \rangle$ (for all t such that $t^2 \neq 4$). We have the following multiplication table:

$$\begin{array}{lll}
[e_0, e_i] = \deg(e_i)\, e_i, & [e_2, e_3] = -te_7, & [e_3, e_4] = 2e_8 - 2e_7, \\
[e_1, e_2] = 0, & [e_2, e_4] = -2e_{10}, & [e_3, e_5] = -te_7/2, \\
[e_1, e_3] = -2e_9, & [e_2, e_5] = e_9, & [e_3, e_6] = -2e_9 - te_{10}/2, \\
[e_1, e_4] = -te_8, & [e_2, e_6] = -e_8, & [e_4, e_5] = -te_9/2 - 2e_{10}, \\
[e_1, e_5] = -e_7, & [e_5, e_6] = e_8 - e_7. & [e_4, e_6] = -te_8/2, \\
[e_1, e_6] = e_{10}, & &
\end{array}$$

Other Lie brackets are 0. It follows from [Seeley and Yau 1991], that we can consider derivations which preserve degree to find a maximal torus of derivations on nilradical $g(V)$. Let δ be a such derivation;

$$\delta e_1 = \sum_{j=1}^{6} a_{6j}e_j, \quad \delta e_7 = \sum_{j=7}^{10} a_{10j}e_j.$$

Case 1. In case of generic t the derivation of the nilradical of $L^*(V_t)$ is spanned by

$$f_1 : \frac{a_{11}}{2} = \frac{a_{22}}{2} = \frac{a_{33}}{2} = \frac{a_{44}}{2} = \frac{a_{55}}{2} = \frac{a_{66}}{2} = a_{77} = a_{88} = a_{99} = a_{10,10} = 1,$$

and otherwise $a_{ij} = 0$. In case of $t = 0$ the derivation of the nilradical of $L^*(V_t)$ is spanned by

$$f_1\colon a_{11} = -a_{33} = 2a_{44} = a_{66} = a_{77} = a_{88} = 2a_{10,10} = 1,$$
$$f_2\colon a_{22} = 2a_{33} = -a_{44} = a_{55} = a_{77} = a_{88} = 2a_{99} = 1,$$

and other $a_{ij} = 0$. Thus for generic t we have a torus of dimension 1 spanned by $\delta = \mathrm{ad}_{e_0}$:

$$\delta e_i = \begin{cases} e_i, & i = 1, 2, 3, 4, 5, 6, \\ 2e_i, & i = 7, 8, 9, 10. \end{cases}$$

For generic t, let $\beta(\delta) = 1$. Then we have the following decomposition of the nilpotent Lie algebra with respect to the maximal torus:

$$g^{\beta} = \langle e_1, e_2, e_3, e_4, e_5, e_6 \rangle,$$
$$C_{ij} = \begin{cases} 2, & i = j, \\ -1, & i \neq j. \end{cases}$$

We have the following generalized Cartan matrix:

$$C(\widetilde{E}_7) = \begin{pmatrix} 2 & -1 & -1 & -1 & -1 & -1 \\ -1 & 2 & -1 & -1 & -1 & -1 \\ -1 & -1 & 2 & -1 & -1 & -1 \\ -1 & -1 & -1 & 2 & -1 & -1 \\ -1 & -1 & -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & -1 & -1 & 2 \end{pmatrix}.$$

Case 2. For $t = 0$ we have a torus of dimension 2, spanned by degree derivations $\delta_1 = \mathrm{ad}_{(x\partial_x)}$ and $\delta_2 = \mathrm{ad}_{(y\partial_y)}$:

$$\delta_1 e_i = \begin{cases} e_i, & i = 1, 5, 6, 8, 9, \\ 2e_i, & i = 4, 7, 10, \\ 0, & i = 2, 3, \end{cases} \quad \text{and} \quad \delta_2 e_i = \begin{cases} e_i, & i = 5, 6, 7, 10, \\ 2e_i, & i = 2, 3, 8, 9, \\ 0, & i = 1, 4. \end{cases}$$

For $t = 0$, we have following decomposition of the nilpotent Lie algebra:

$$\beta_1(\delta_1) = 1, \quad \beta_1(\delta_2) = 0 \quad \Rightarrow \quad g^{\beta_1} = \langle e_1 \rangle,$$
$$\beta_2(\delta_1) = 1, \quad \beta_2(\delta_2) = 1 \quad \Rightarrow \quad g^{\beta_2} = \langle e_5, e_6 \rangle,$$
$$\beta_3(\delta_1) = 2, \quad \beta_3(\delta_2) = 0 \quad \Rightarrow \quad g^{\beta_3} = \langle e_4 \rangle,$$
$$\beta_4(\delta_1) = 0, \quad \beta_4(\delta_2) = 2 \quad \Rightarrow \quad g^{\beta_4} = \langle e_2, e_3 \rangle.$$

It is easy to see that

$$
\begin{aligned}
&f(1)=1, \quad f(2)=2, \quad f(3)=2, \quad f(4)=3, \quad f(5)=4, \quad f(6)=4,\\
&(\operatorname{ad} e_1)^2 e_5=(\operatorname{ad} e_1)^2 e_6=0 \quad \Rightarrow C_{12}=C_{13}=C_{21}=C_{31}=-1,\\
&(\operatorname{ad} e_1) e_4=0 \quad \Rightarrow C_{14}=C_{41}=0,\\
&(\operatorname{ad} e_1)^2 e_2=0, \quad (\operatorname{ad} e_1)^2 e_3=[e_1,-2e_9]=0 \quad \Rightarrow C_{15}=C_{51}=C_{16}=C_{61}=-1,\\
&(\operatorname{ad}(\gamma e_5+e_6))^2 e_5=[\gamma e_5+e_6, \gamma e_8-\gamma e_7]=0, \quad (\operatorname{ad} e_5)^2 e_6=[e_5, e_8-e_7]=0\\
&\qquad \Rightarrow C_{23}=C_{32}=-1,\\
&(\operatorname{ad} e_5)^2 e_4=(\operatorname{ad} e_6)^2 e_4=0 \quad \Rightarrow C_{24}=C_{34}=C_{42}=C_{43}=-1,\\
&(\operatorname{ad} e_5)^2 e_2=0, \quad (\operatorname{ad} e_6)^2 e_2=0, \quad (\operatorname{ad} e_5)^2 e_3=0, \quad (\operatorname{ad} e_6)^2 e_3=0\\
&\qquad \Rightarrow C_{25}=C_{26}=C_{35}=C_{36}=C_{52}=C_{53}=C_{62}=C_{63}=-1,\\
&(\operatorname{ad} e_4)^2 e_2=[e_4, 2e_{10}]=0, \quad (\operatorname{ad} e_4)^2 e_3=[e_4, 2e_7-2e_8]=0\\
&\qquad \Rightarrow C_{45}=C_{46}=C_{54}=C_{64}=-1,\\
&(\operatorname{ad}(\gamma e_2+e_3)) e_2=(\operatorname{ad} e_2) e_3=0 \quad \Rightarrow C_{56}=C_{65}=0.
\end{aligned}
$$

We have the following generalized Cartan matrix:

$$
C(\widetilde{E}_7)=\begin{pmatrix} 2 & -1 & -1 & 0 & -1 & -1\\ -1 & 2 & -1 & -1 & -1 & -1\\ -1 & -1 & 2 & -1 & -1 & -1\\ 0 & -1 & -1 & 2 & -1 & -1\\ -1 & -1 & -1 & -1 & 2 & 0\\ -1 & -1 & -1 & -1 & 0 & 2 \end{pmatrix}. \qquad \square
$$

Proposition 3.9. *Let* $V_t=\{(x,y,z)\in\mathbb{C}^3 \mid f_t=x^6+y^3+z^2+tx^4y=0\}$ *with* $4t^3+27\neq 0$, *be the* $\widetilde{E}_8$ *singularity. Then*

$$
C(\widetilde{E}_8)=\begin{cases} \begin{pmatrix} 2 & -3 & -3\\ -3 & 2 & -3\\ -3 & -3 & 2 \end{pmatrix}, & t\neq 0,\\ \begin{pmatrix} 2 & -1 & -1\\ -1 & 2 & -2\\ -1 & -2 & 2 \end{pmatrix}, & t=0. \end{cases}
$$

Proof. In [Seeley and Yau 1990], the authors had studied the (μ,τ)-constant family of $\widetilde{E}_8$, which is given by

$$
V_t=\big\{(x,y,z)\in\mathbb{C}^3 \mid f_t=x^6+y^3+z^2+tx^4y=0\big\},
$$

with $4t^3+27\neq 0$. The new moduli algebra

$$A^*(V_t)=\langle 1,x,x^2,y,x^3,xy,x^4,x^2y,x^3y\rangle,$$

with the multiplication rules $y^2=-\frac{t}{3}x^4,\ x^5=-\frac{2t}{3}x^3y,\ x^4y=0$.

By calculation, a basis of the new Lie algebra $L^*(V_t)$ (for $4t^3+27\neq 0$ and $t\neq 0$), is the following:

$$\begin{aligned}
&\deg 0: && e_0=x\partial_x+2y\partial_y,\\
&\deg 1: && e_1=x^2\partial_x+2xy\partial_y,\quad e_2=3y\partial_x-2tx^3\partial_y,\quad e_3=ty\partial_x-3xy\partial_y,\\
&\deg 2: && e_4=x^3\partial_x,\quad e_5=xy\partial_x,\quad e_6=x^2y\partial_y,\quad e_7=x^4\partial_y,\\
&\deg 3: && e_8=x^4\partial_x,\quad e_9=x^2y\partial_x,\quad e_{10}=x^3y,\partial_y,\\
&\deg 4: && e_{11}=x^3y\partial_x.
\end{aligned}$$

For $t=0$, $\{e_0\}$ is replaced by $\{x\partial_x, y\partial_y\}$. Let $g(V)$ be the nilradical of $L^*(V_t)$, which is of dimension 11, spanned by $\langle e_1,e_2,\dots,e_{11}\rangle$ (for all t such that $4t^3+27\neq 0$). By calculation, the multiplication table of $g(V)$ is given as follows:

$$[e_1,e_2]=0,\quad [e_1,e_3]=-3e_6+\frac{2t^2}{3}e_7,\quad [e_2,e_3]=-2t^2e_4+9e_5+6t^2e_6+9te_7,$$
$$[e_1,e_4]=e_8-2e_{10}\quad [e_2,e_4]=9e_9-4t^2e_{10},\quad [e_1,e_5]=e_9-\frac{4t^2}{9}e_{10},$$
$$[e_2,e_5]=-3te_8+6te_{10},\quad [e_1,e_6]=2e_{10},\quad [e_2,e_6]=-3e_9+\frac{8t^2}{3}e_{10},$$
$$[e_1,e_7]=-\frac{4t}{3}e_{10},\quad [e_2,e_7]=-3e_8+12e_{10},\quad [e_1,e_8]=-\frac{4t}{3}e_{11},$$
$$[e_2,e_8]=12e_{11},\quad [e_1,e_9]=2e_{11},\quad [e_2,e_9]=\frac{8t^2}{3}e_{11},\quad [e_1,e_{10}]=0,$$
$$[e_2,e_{10}]=-3e_{11},\quad [e_3,e_4]=3te_9+3e_{10},\quad [e_3,e_5]=-\frac{t^2}{3}e_8-3e_9+\frac{2t^2}{3}e_{10},$$
$$[e_4,e_5]=-2e_{11},\quad [e_3,e_6]=-te_9+\frac{4t^3}{9}e_{10},\quad [e_4,e_6]=0,$$
$$[e_3,e_7]=-te_8+2te_{10},\quad [e_4,e_7]=0,\quad [e_3,e_8]=4te_{11},\quad [e_5,e_6]=-e_{11},$$
$$[e_3,e_9]=\Big(\frac{4t^3}{9}-3\Big)e_{11},\quad [e_5,e_7]=\frac{2t}{3}e_{11},\quad [e_3,e_{10}]=-te_{11},\quad [e_6,e_7]=0.$$

Other Lie brackets $[e_i,e_j]$ $(i<j)$ are zero. It is follows from [Seeley and Yau 1991], that we can consider derivations which preserve degree to find a maximal torus of derivations on the nilradical $g(V)$. Let δ be a such derivation:

$$\delta e_1=\sum_{j=1}^{3}a_{3j}e_j,\quad \delta e_4=\sum_{j=4}^{7}a_{7j}e_j,\quad \delta e_8=\sum_{j=8}^{10}a_{10j}e_j,\quad \delta e_{11}=a_{11,11}e_{11}.$$

It is follows from the multiplication table that

$$[g(V), g(V)] = \left\langle -3e_6 + \frac{2t^2}{3}e_7, -2t^2e_4 + 9e_5 + 6t^2e_6 + 9te_7, e_8, e_9, e_{10}, e_{11} \right\rangle.$$

Therefore the type of $\widetilde{E}_8$ singularity $= \dim g(V)/[g(V), g(V)] = 5$. The nilpotency of $\widetilde{E}_8$ singularity $= \min\{p \in N \cup \{0\} \mid g(V)^{p+1} = 0\} = 4$

Case 1. It is follows from [Seeley and Yau 1991] in case of generic t the derivation of nilradical of $L^*(V_t)$ is spanned by

$$f_1 \colon \frac{3a_{10,10}}{4} = \frac{a_{11}}{4} = a_{11,11} = \frac{a_{22}}{4} = \frac{a_{33}}{4} = \frac{a_{44}}{2} = \frac{a_{55}}{2} = \frac{a_{66}}{2} = \frac{a_{77}}{2} = \frac{3a_{88}}{4} = \frac{3a_{99}}{4} = 1,$$

and other $a_{ij} = 0$. In case of $t = 0$, the derivation of the nilradical of $L^*(V_t)$ is spanned by

$$f_1 \colon -a_{11,11} = -a_{22} = -a_{55} = a_{77} = -a_{99} = 1,$$
$$f_2 \colon -3a_{10,10} = -a_{11} = 3a_{22} = -a_{33} = -2a_{44} = 2a_{55} = -2a_{66} = -6a_{77} = -3a_{88} = a_{99} = 1,$$

and other $a_{ij} = 0$. Thus for generic t we have a torus of dimension 1 spanned by $\delta = \mathrm{ad}_{e_0}$;

$$\delta e_i = \begin{cases} e_i, & i = 1, 2, 3, \\ 2e_i, & i = 4, 5, 6, 7, \\ 3e_i, & i = 8, 9, 10, \\ 4e_i, & i = 11. \end{cases}$$

Let $\beta(\delta) = 1$, and $g^\beta = \langle e_1, e_2, e_3 \rangle$. Since $(\mathrm{ad}\, e_1)^4 e_2 = 0$, $(\mathrm{ad}\, e_1)^4 e_3 = 0$ and $(\mathrm{ad}\, e_2)^4 e_3 = 0$:

$$C_{ij} = \begin{cases} 2, & i = j, \\ -3, & i \neq j. \end{cases}$$

So we have the following generalized Cartan matrix:

$$C(\widetilde{E}_8) = \begin{pmatrix} 2 & -3 & -3 \\ -3 & 2 & -3 \\ -3 & -3 & 2 \end{pmatrix}.$$

Case 2. For $t = 0$ we have a torus of dimension 2, spanned by degree derivation $\delta_1 = \mathrm{ad}_{(x\partial_x)}$ and $\delta_2 = \mathrm{ad}_{(y\partial_y)}$;

$$\delta_1 e_i = \begin{cases} e_i, & i = 1, 2, 3, 9, \\ 2e_i, & i = 4, 6, 11, \\ 3e_i, & i = 8, 10, \\ 4e_i, & i = 7, \\ 0, & i = 5, \end{cases} \quad \text{and} \quad \delta_2 e_i = \begin{cases} e_i, & i = 2, 5, 7, 9, 11, \\ 0, & i = 1, 3, 4, 6, 8, 10. \end{cases}$$

For $t = 0$, we have the following decomposition of the nilpotent Lie algebra:

$$\beta_1(\delta_1) = 1, \quad \beta_1(\delta_2) = 1 \quad \Rightarrow \quad g^{\beta_1} = \langle e_2, e_9 \rangle,$$
$$\beta_2(\delta_1) = 1, \quad \beta_2(\delta_2) = 0 \quad \Rightarrow \quad g^{\beta_2} = \langle e_1, e_3 \rangle.$$

It is noted that $e_9 \in [g(V), g(V)]$ and we have $f(1) = 1,\ f(2) = f(3) = 2$,

$$(\operatorname{ad} e_2)^2 e_1 = (\operatorname{ad} e_2)^2 e_3 = (\operatorname{ad} e_9)^2 e_1 = (\operatorname{ad} e_9)^2 e_3 = 0$$
$$\Rightarrow C_{12} = C_{21} = C_{31} = C_{13} = -1,$$
$$(\operatorname{ad}(\gamma e_1 + e_3))^3 e_1 = (\operatorname{ad} e_1)^3 e_3 = 0 \qquad \Rightarrow C_{32} = C_{23} = -2.$$

We have the following generalized Cartan matrix:

$$C(\widetilde{E}_8) = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -2 \\ -1 & -2 & 2 \end{pmatrix}. \qquad \square$$

References

[Arnold et al. 2004] V. I. Arnold, A. N. Varchenko, and S. M. Gusein-Zade, Особенности дифференцируемых отображений, 2nd ed., MCNMO, Moscow, 2004.

[Benson and Yau 1987] M. Benson and S. S.-T. Yau, "Lie algebras and their representations arising from isolated singularities: computer method in calculating the Lie algebras and their cohomology", pp. 3–58 in *Complex analytic singularities*, edited by T. Suwa and P. Wagreich, Adv. Stud. Pure Math. **8**, North-Holland, Amsterdam, 1987. MR Zbl

[Benson and Yau 1990] M. Benson and S. S.-T. Yau, "Equivalences between isolated hypersurface singularities", *Math. Ann.* **287**:1 (1990), 107–134. MR Zbl

[Bratzlavsky 1974] F. Bratzlavsky, "Sur les algèbres admettant un tore d'automorphismes donné", *J. Algebra* **30** (1974), 305–316. MR Zbl

[Chen et al. 1995] H. Chen, Y.-J. Xu, and S. S.-T. Yau, "Nonexistence of negative weight derivation of moduli algebras of weighted homogeneous singularities", *J. Algebra* **172**:2 (1995), 243–254. MR Zbl

[Chen et al. 2019] H. Chen, S. S.-T. Yau, and H. Zuo, "Non-existence of negative weight derivations on positively graded Artinian algebras", *Trans. Amer. Math. Soc.* **372**:4 (2019), 2493–2535. MR Zbl

[Chen et al. ≥ 2020a] B. Chen, H. Chen, S. S.-T. Yau, and H. Zuo, "The non-existence of negative weight derivations on positive dimensional isolated singularities: generalized Wahl conjecture", To appear in *J. Differential Geom.*

[Chen et al. ≥ 2020b] B. Chen, N. Hussain, S. S.-T. Yau, and H. Zuo, "Variation of complex structures and variation of Lie algebras II: New Lie algebras arising from singularities", To appear in *J. Differential Geom.*

[Dimca 1984] A. Dimca, "Are the isolated singularities of complete intersections determined by their singular subspaces?", *Math. Ann.* **267**:4 (1984), 461–472. MR Zbl

[Elashvili and Khimshiashvili 2006] A. Elashvili and G. Khimshiashvili, "Lie algebras of simple hypersurface singularities", *J. Lie Theory* **16**:4 (2006), 621–649. MR Zbl

[Favre 1972] G. Favre, *Système de poids sur une algèbre de Lie nilpotente*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 1972.

[Favre 1973] G. Favre, "Système de poids sur une algèbre de Lie nilpotente", *Manuscripta Math.* **9** (1973), 53–90. MR Zbl

[Hussain 2018] N. Hussain, "Survey on derivation Lie algebras of isolated singularities", *Methods Appl. Anal.* **25**:4 (2018), 307–322. MR Zbl

[Hussain et al. 2018] N. Hussain, S. S.-T. Yau, and H. Zuo, "On the derivation Lie algebras of fewnomial singularities", *Bull. Aust. Math. Soc.* **98**:1 (2018), 77–88. MR Zbl

[Hussain et al. 2019a] N. Hussain, S. S.-T. Yau, and H. Zuo, "kth Yau number of isolated hypersurface singularities and an inequality conjecture", *J. Aust. Math. Soc.* (online publication April 2019).

[Hussain et al. 2019b] N. Hussain, S. S.-T. Yau, and H. Zuo, "On two conjectures for the k-th Yau numbers of isolated hypersurface singularities", preprint, 2019.

[Hussain et al. 2020] N. Hussain, S. S.-T. Yau, and H. Zuo, "On the new k-th Yau algebras of isolated hypersurface singularities", *Math. Z.* **294**:1-2 (2020), 331–358. MR Zbl

[Mather and Yau 1982] J. N. Mather and S. S.-T. Yau, "Classification of isolated hypersurface singularities by their moduli algebras", *Invent. Math.* **69**:2 (1982), 243–251. MR Zbl

[Santharoubane 1983] L. J. Santharoubane, "Kac–Moody Lie algebras and the universal element for the category of nilpotent Lie algebras", *Math. Ann.* **263**:3 (1983), 365–370. MR Zbl

[Seeley and Yau 1990] C. Seeley and S. S.-T. Yau, "Variation of complex structures and variation of Lie algebras", *Invent. Math.* **99**:3 (1990), 545–565. MR Zbl

[Seeley and Yau 1991] C. Seeley and S. S.-T. Yau, "Algebraic methods in the study of simple-elliptic singularities", pp. 216–237 in *Algebraic geometry* (Chicago, IL, 1989), edited by S. Bloch et al., Lecture Notes in Math. **1479**, Springer, 1991. MR Zbl

[Umlauf 1891] K. A. Umlauf, *Über die Zusammensetzung der endlichen continuierlichen Transformationsgruppen insbesondre der Gruppen vom Range Null*, Ph.D. thesis, Universität Leipzig, 1891, Available at https://tinyurl.com/UmlaufThesis. Zbl

[Yau 1983] S. S. T. Yau, "Continuous family of finite-dimensional representations of a solvable Lie algebra arising from singularities", *Proc. Nat. Acad. Sci. U.S.A.* **80**:24 (1983), 7694–7696. MR Zbl

[Yau 1984] S. S.-T. Yau, "Criteria for right-left equivalence and right equivalence of holomorphic functions with isolated critical points", pp. 291–297 in *Complex analysis of several variables* (Madison, WI, 1982), edited by Y. T. Siu, Proc. Sympos. Pure Math. **41**, Amer. Math. Soc., Providence, RI, 1984. MR Zbl

[Yau 1986] S. S.-T. Yau, "Solvable Lie algebras and generalized Cartan matrices arising from isolated singularities", *Math. Z.* **191**:4 (1986), 489–506. MR Zbl

[Yau 1991] S. S.-T. Yau, "Solvability of Lie algebras arising from isolated singularities and nonisolatedness of singularities defined by $\mathrm{sl}(2, \mathbf{C})$ invariant polynomials", *Amer. J. Math.* **113**:5 (1991), 773–778. MR Zbl

[Yau and Zuo 2016a] S. S.-T. Yau and H. Zuo, "Derivations of the moduli algebras of weighted homogeneous hypersurface singularities", *J. Algebra* **457** (2016), 18–25. MR Zbl

[Yau and Zuo 2016b] S. S.-T. Yau and H. Zuo, "A sharp upper estimate conjecture for the Yau number of a weighted homogeneous isolated hypersurface singularity", *Pure Appl. Math. Q.* **12**:1 (2016), 165–181. MR Zbl

[Yu 1996] Y. Yu, "On Jacobian ideals invariant by a reducible $\mathrm{sl}(2, \mathbf{C})$ action", *Trans. Amer. Math. Soc.* **348**:7 (1996), 2759–2791. MR Zbl

Received January 20, 2019. Revised July 24, 2019.

NAVEED HUSSAIN
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING
CHINA

and

SCHOOL OF DATA SCIENCES
HUASHANG COLLEGE GUANGDONG UNIVERSITY OF FINANCE AND ECONOMICS
GUANGZHOU GUANGDONG
CHINA

naveed1573@gmail.com

STEPHEN S.-T. YAU
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING
CHINA

yau@uic.edu

HUAIQING ZUO
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING
CHINA

hqzuo@mail.tsinghua.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.219

ON THE COMMUTATIVITY OF COSET PRESSURE

BING LI AND WEN-CHIAO CHENG

We establish the conditional entropy with a coset partition and coset pressure for subadditive potentials via separated sets on a compact metric group. Analogues of the variational principle and thermodynamic formalism are shown in this system. The major finding of this study consists of its presentation of the commutativity proposition for those conjugate invariants.

1. Introduction

A basic issue in the theory of dynamical systems is the study of the complexity of orbits. This has led to the development of many different subjects in mathematics. In 1958, Kolmogorov applied the notion of entropy from information theory to ergodic theory. Since then, the notion of entropy has played an important role in understanding the complexity of various dynamical systems. The two main types of entropy are measure-theoretic (or metric) entropy and topological entropy. The former measures the maximal loss of information of the iteration of finite partitions in a measure preserving transformation. The latter measures the maximal exponential growth rate of orbits for an arbitrary topological dynamical system. These two notions are connected by the so-called variational principle. This relation states that the topological entropy is the supremum of the metric entropies for all invariant probability measures of a given topological system, and has received considerable attention.

As a natural generalization of topological entropy, topological pressure is a quantity which belongs to one of the concepts in the thermodynamic formalism. The thermodynamic formalism itself is a generalization of the concepts of statistical physics to the area of mathematical dynamical systems theory. Ruelle [1973] first introduced the concept of topological pressure of additive potentials for expansive dynamical systems. Walters [1982] then extended this concept to the compact space with the continuous transformation. Moreover, in some cases, the values of entropy functions can be expressed as the topological pressure of certain functions related to dynamical systems. Numerous nonlinear physical problems involve a complicated discrete dynamical system. Topological pressure contains information

---

Cheng is the corresponding author.
MSC2010: 28D20.
Keywords: thermodynamic formalism, coset pressure, conditional entropy, commutativity.

on the dynamics of the system; these dynamics can be extracted by varying the potential energy function. Related studies include [Falconer 1988; Gelfert and Wolf 2008; Huang et al. 2008; Huang and Yi 2007; Molaei 2008; Pollner and Vattay 1996; Spandl 2008]. The framework presented by Bowen [1971] has caused topological pressure to become a fundamental tool for studying the multifractal formalism of dimension theory, especially for examining nonconformal dynamical systems in statistical mechanics; see [Falconer 1988; Pesin 1997].

Furthermore, the notions of topological entropy and topological pressure introduced above are applied to autonomous dynamical systems. Kolyada, Misiurewicz, and Snoha [Kolyada et al. 1999; Kolyada and Snoha 1996] introduced topological entropy for a nonautonomous dynamical system given by a sequence $\{f_n\}_{n=1}^{\infty}$ of continuous self-maps of a compact metric space. More precisely, Kolyada and Snoha [1996] showed that the topological entropy of the composition of two continuous self-maps of a compact metric space does not depend on the order in which functions compose, that is, $h_{\text{top}}(S \circ T) = h_{\text{top}}(T \circ S)$. This fact demonstrates that the dynamics of $S \circ T$ must exhibit some common features with that of $T \circ S$, although $S \circ T$ and $T \circ S$ do not usually coincide. Kong, Cheng, and Li [Kong et al. 2015] generalized topological entropy to the topological pressure of $(X;\, f_{1,\infty})$. They analyzed those basic pressure propositions concerning a nonautonomous dynamical system $(X;\, f_{1,\infty})$ given by a compact metric space X and a sequence $f_{1,\infty} = \{f_n\}_{n=1}^{\infty}$ of continuous self-maps of X. They also showed that, for any continuous maps T and S from a compact metric space into itself, the maps $T \circ S$ and $S \circ T$ have the same topological pressure. Different propositions regarding nonautonomous dynamical systems also can be found in [Cánovas 2011; Kuang et al. 2013].

Essentially, the thermodynamic formalism can be described as a rigorous study of certain mathematical structures inspired in thermodynamics. Balibrea, López and Peña [Balibrea et al. 1999b] presented the commutativity proposition for topological pressure by using thermodynamic formalism. The commutativity proposition for the sequence entropy on the interval also can be obtained in [Balibrea et al. 1999a]. Here we pursue the commutativity for coset pressure and conditional entropy function. The outline of the paper is as follows. Section 2 establishes basic entropy with a coset partition and shows the variational principle. Then we discuss the commutativity of conditional entropy. Section 3 studies coset pressure with a sequence of subadditive potential functions and demonstrates the thermodynamic formalism. Then, along with [Balibrea et al. 1999b] and based on thermodynamic formalism, the main result of the commutativity proposition also holds for the coset pressure.

2. Commutativity of entropy

In dynamical systems and ergodic theory it is understood that a reasonable measure-theoretic or topological entropy should be a measure of the uncertainty of the system and that they should be invariant under measurable or topological change of coordinates, respectively. This section reviews the concept of conditional entropy in the context of a probability space and topological entropy in the context of the compact metric space. We then show the commutativity property of conditional entropy as follows.

Assume that (X, d) is a compact metric space with metric d and $T : X \to X$ is a continuous selfmap. For any $n \in \mathbb{N}$, first we define the Bowen metric on X as

$$d_n(x, y) = \max_{0 \le i \le n-1} d(T^i x, T^i y) \quad \text{for all } x, y \in X.$$

Let K be a compact subset of X. For any $n \in \mathbb{N}$ and $\epsilon > 0$, a subset F of X is said to be an (n, ϵ)-spanning set of K with respect to T if for all $x \in K$, there exists $y \in F$ with $d_n(x, y) \le \epsilon$, i.e.,

$$K \subset \bigcup_{y \in F} \bigcap_{i=0}^{n-1} T^{-i} \bar{B}(T^i y; \epsilon),$$

where $B(T^i y; \epsilon)$ represents the open ball with center $T^i y$ and radius ϵ in the metric d, and $\bar{B}(T^i y; \epsilon)$ is the corresponding closed ball. Let $r(n, \epsilon, K)$ denote the smallest cardinality of (n, ϵ)-spanning set for K with respect to T. A subset E of K is said to be (n, ϵ)-separated with respect to T if $x, y \in E, x \neq y$, implies $d_n(x, y) > \epsilon$. In other words, for $x \in E$, the set $\bigcap_{i=0}^{n-1} T^{-i} \bar{B}(T^i x; \epsilon)$ contains no other points of E except x itself. Let $s(n, \epsilon, K)$ be the largest cardinality of a (n, ϵ)-separated subset of K with respect to T.

In the compact metric space, the topological entropy of a set K introduced by Bowen and defined by the separated or spanning sets can also be given using open covers. Let α be an open cover of X and denote by $\aleph(\alpha|_K)$ the number of sets in a finite subcover of α with the smallest cardinality for K. The entropy of α on K is defined by $H(\alpha|_K) = \log \aleph(\alpha|_K)$, and the topological entropy of T for K is as follows:

$$\begin{aligned} h_{\text{top}}(T|K) &= \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(n, \epsilon, K) \\ &= \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon, K) \\ &= \sup_{\text{open cover } \beta} \limsup_{n \to \infty} \frac{1}{n} \log \aleph\big(\textstyle\bigvee_{i=0}^{n-1} T^{-i} \beta|_K\big), \end{aligned}$$

where

$$\textstyle\bigvee_{i=0}^{n-1} T^{-i} \beta = \{A_{i_0} \cap T^{-1} A_{i_1} \cap \cdots \cap T^{-(n-1)} A_{i_{n-1}} : A_{i_j} \in \beta\}.$$

We assume throughout that $(X, d, \cdot)$ is a compact group with metric d and group product $\cdot$. $T : X \to X$ is a continuous endomorphism, and B is a closed T-invariant subgroup of X, where T-invariant means $T^{-1}(B) = B$. For convenience, some more notations are as follows:

- $[B]$ is the coset partition, i.e.,

$$[B] = \{B \cdot x : x \in X\},$$

  where $B \cdot x = \{y \cdot x : y \in B\}$. It is easy to check that the collection of sets $B \cdot x$ forms a partition of X.

- $\mathcal{M}(X)$ is the set of all Borel probability measures on X. $\mathcal{M}(X, T)$ is the subset of $\mathcal{M}(X)$ with all T-invariant measures. $\mathcal{E}(X, T)$ is the subset of $\mathcal{M}(X, T)$ with all T-invariant ergodic measures.

- For any partition ξ of X,

$$\xi^n = \bigvee_{i=0}^{n-1} T^{-i}\xi = \{A_{i_0} \cap T^{-1}A_{i_1} \cap \cdots \cap T^{-(n-1)}A_{i_{n-1}} : A_{i_j} \in \xi, 0 \le j \le n-1\}.$$

Under the same assumption as above, the topological entropy of T for $B \cdot x$ is defined to be

$$\begin{aligned} h_{\text{top}}(T \mid B \cdot x) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log r(n, \epsilon, B \cdot x) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon, B \cdot x) \\ &= \sup_{\text{open cover } \beta} \limsup_{n\to\infty} \frac{1}{n} \log \aleph\Big(\bigvee_{i=0}^{n-1} T^{-i}\beta|_{B\cdot x}\Big). \end{aligned}$$

The conditional entropy is usually defined as follows. Let $(X, \mathcal{B}, \mu)$ be a probability space, and let $\mathcal{A}$ and $\mathcal{C}$ be two partitions of $(X, \mathcal{B}, \mu)$ with

$$\mathcal{A} = \{A_1, \dots, A_k\} \quad \text{and} \quad \mathcal{C} = \{C_1, \dots, C_p\}.$$

The entropy of $\mathcal{A}$ given $\mathcal{C}$ is the number

$$H_\mu(\mathcal{A} \mid \mathcal{C}) = -\sum_{j=1}^{p} \mu(C_j) \sum_{i=1}^{k} \frac{\mu(A_i \cap C_j)}{\mu(C_j)} \log \frac{\mu(A_i \cap C_j)}{\mu(C_j)}$$

omitting the j-terms when $\mu(C_j) = 0$. Here, the summation of each quantity is 1, i.e., $\sum_{j=1}^{p} \mu(C_j) = 1$.

Next, let $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ be a measure preserving transformation of probability space $(X, \mathcal{B}, \mu)$ (i.e., if $A \in \mathcal{B}$, then $T^{-1}A \in \mathcal{B}$ and $\mu(T^{-1}A) = \mu(A)$). For any finite partition α of X, we consider the refinement α^n of α. Again, B is a closed T-invariant subgroup. Then $B \cdot x$ is closed and $[B] = T^{-1}[B]$. Therefore each element in the partition $[B]$ belongs to $\mathcal{B}$. Then we consider the conditional entropy given by the coset partition $[B]$, defined by $H_\mu(\alpha^n \mid [B]) = H_\mu\big(\bigvee_{i=0}^{n-1} T^{-i}\alpha \mid [B]\big)$.

Good references for these entropy invariants are [Brucks and Bruin 2004; Katok and Hasselblatt 1995], which contain many of the relevant earlier references.

Lemma 2.1 [Cheng 2006]. *The sequence* $a_n = H_\mu(\alpha^n \mid [B])$ *is subadditive, that is,*

$$a_{n+m} = H_\mu(\alpha^{n+m} \mid [B]) \le H_\mu(\alpha^n \mid [B]) + H_\mu(\alpha^m \mid [B]) = a_n + a_m.$$

The conditional entropy of α given $[B]$ with respect to T is the value

$$\begin{aligned} h_\mu(T \mid [B], \alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha^n \mid [B]) \\ &= \inf_{n\ge 1} \frac{1}{n} H_\mu(\alpha^n \mid [B]) \end{aligned}$$

and the conditional entropy of T with respect to μ and $[B]$ is defined to be

$$h_\mu(T \mid [B]) = \sup_\alpha h_\mu(T \mid [B], \alpha),$$

where α ranges over all finite partitions of X.

The well-known variational principle for entropy shows the relationship between topological entropy and metric entropy. We only can obtain the variational inequality for the invariant partition. However, under the condition of coset partition and using Misiurewicz's technique and more advanced ergodic analysis, a similar type of local variational principle can be obtained as follows.

Theorem 2.2 [Cheng 2006] (variational principle). *Let* $T : X \to X$ *be a continuous endomorphism of the compact group* $(X, d, \cdot\,)$ *with closed* T*-invariant subgroup* B*. Then we have*

$$\sup_{x\in X} h_{\text{top}}(T \mid B\cdot x) = \sup_{\mu\in\mathcal{M}(X,T)} h_\mu(T \mid [B]).$$

Next, we investigate the invariant measure for the measurable selfmap. Assume that $S : X \to X$ and $T : X \to X$ are functions. The composition of S and T is given by $S\circ T(x) = S(T(x))$. It is obvious that $\mathcal{M}(X, S) \cap \mathcal{M}(X, T) \subseteq \mathcal{M}(X, S\circ T)$. Let $\widetilde{S} : \mathcal{M}(X) \to \mathcal{M}(X)$ be defined as follows. For any $\mu \in \mathcal{M}(X)$, the measure $\widetilde{S}\mu$ is defined by $\widetilde{S}\mu(E) = \mu(S^{-1}(E))$ for all measurable set $E \subset X$. If μ is an invariant measure of S, then $\widetilde{S}\mu = \mu$. If $\mu \in \mathcal{M}(X, S)\cap\mathcal{M}(X, S\circ T)$, then $\widetilde{S}\mu = \mu$ and hence, $\widetilde{T}\mu = \mu$. Thus μ is an invariant measure for S and T.

Lemma 2.3 [Balibrea et al. 1999b]. *Given the same conditions as indicated above, we can see that* $\widetilde{S\circ T} = \widetilde{S}\circ\widetilde{T}$.

Lemma 2.4 [Walters 1982]. *Given the same conditions as indicated above and that* $\phi : X \to R$ *is a continuous real-value function, we can see that*

$$\int \phi \, d(\widetilde{S}\mu) = \int \phi\circ S \, d\mu.$$

Lemma 2.5 [Balibrea et al. 1999b]. *Given the same conditions as indicated above, the maps*

$$\widetilde{T}\,|_{\mathcal{M}(X,S\circ T)} : \mathcal{M}(X, S\circ T) \to \mathcal{M}(X, T\circ S),$$
$$\widetilde{S}\,|_{\mathcal{M}(X,T\circ S)} : \mathcal{M}(X, T\circ S) \to \mathcal{M}(X, S\circ T)$$

are well defined bijections.

Theorem 2.6. *Assume that $S, T : X \to X$ are continuous endomorphisms and let $\mu \in \mathcal{M}(X, S\circ T)$. B is a closed S, T-invariant subgroup of X, with coset partition $[B]$. Then, we have*

$$h_\mu(S\circ T \mid [B]) = h_{\widetilde{T}\mu}(T\circ S \mid [B]).$$

Proof. Assume $\mathcal{A}$ is a finite partition of X, then we have that

$$H_\mu(S^{-1}\mathcal{A} \mid [B]) = H_{\widetilde{S}\mu}(\mathcal{A} \mid [B])$$

and clearly that

$$h_\mu(S\circ T \mid [B], \mathcal{A}) = h_\mu(S\circ T \mid [B], (S\circ T)^{-1}\mathcal{A}).$$

Moreover,

$$\begin{aligned} H_\mu\big(\textstyle\bigvee_{i=0}^{n-1}(S\circ T)^{-i}\mathcal{A} \mid [B]\big) &= H_\mu\big(T^{-1}\textstyle\bigvee_{i=0}^{n-2}(S\circ T)^{-i}(S^{-1}\mathcal{A}) \mid [B]\big) \\ &= H_\mu\big(T^{-1}\textstyle\bigvee_{i=0}^{n-2}(S\circ T)^{-i}(S^{-1}\mathcal{A}) \mid T^{-1}[B]\big) \\ &= H_{\widetilde{T}\mu}\big(\textstyle\bigvee_{i=0}^{n-2}(S\circ T)^{-i}(S^{-1}\mathcal{A}) \mid [B]\big). \end{aligned}$$

We can divide by n and let n go to ∞ to obtain

$$h_\mu(S\circ T \mid [B], \mathcal{A}) = h_{\widetilde{T}\mu}(T\circ S \mid [B], S^{-1}\mathcal{A}).$$

This implies $h_\mu(S\circ T \mid [B], \mathcal{A}) = h_{\widetilde{T}\mu}(T\circ S \mid [B], S^{-1}\mathcal{A}) \le h_{\widetilde{T}\mu}(T\circ S \mid [B])$. Furthermore, since $\mu \in \mathcal{M}(X, S\circ T)$, we have $\widetilde{S}\widetilde{T}\mu = \mu$, and

$$h_{\widetilde{T}\mu}(T\circ S \mid [B], \mathcal{A}) = h_\mu(S\circ T) \mid [B], T^{-1}\mathcal{A}) \le h_\mu(S\circ T \mid [B]).$$

Thus $h_\mu(S\circ T \mid [B]) = h_{\widetilde{T}\mu}(T\circ S \mid [B])$. □

We have $\widetilde{T}\mu = \mu$ if $\mu \in \mathcal{M}(X, T)$. Therefore the commutativity for conditional entropy holds under some situations.

Lemma 2.7. *Assume that $S, T : X \to X$ are continuous endomorphisms and let $\mu \in \mathcal{M}(X, S) \cap \mathcal{M}(X, T)$. If B is a closed S, T-invariant subgroup of X, with coset partition $[B]$, then we have*

$$h_\mu(S\circ T \mid [B]) = h_\mu(T\circ S \mid [B]).$$

3. Commutativity of pressure

Topological pressure is an important generalization of topological entropy. It is well known that the topological pressure with a potential function plays a fundamental role in the study of the Hausdorff dimension of repellers and the hyperbolic set. It roughly measures the orbit structure complexity of the iterated map on the potential function. During this section, along with the discussion of the approximation provided in the study by Cao, Feng and Huang [Cao et al. 2008], we define the subadditivity of a sequence of functions and coset pressure in the following. Then we state the thermodynamic formalism and use this formalism to demonstrate the commutativity of coset pressure.

A sequence $\mathcal{F} = \{\log f_n\}_{n=1}^{\infty}$ of functions on X is called subadditive if each f_n is a positive real-valued continuous function on X with

$$f_{n+m}(x) \leq f_n(x) f_m(T^n x) \quad \text{for all } x \in X, m, n \in \mathbb{N}.$$

Assume $[B]$ is a coset partition of this compact group X and $T : X \to X$ is a continuous endomorphism. Therefore the coset pressure is defined as follows:

$$P_n(T, \mathcal{F}, \epsilon, B \cdot x) = \sup_E \Big\{ \sum_{y \in E} f_n(y) : E \text{ is an } (n, \epsilon)\text{-separated subset of } B \cdot x \Big\}.$$

and

$$P(T, \mathcal{F}, \epsilon, B \cdot x) = \limsup_{n \to \infty} \frac{1}{n} \log P_n(T, \mathcal{F}, \epsilon, B \cdot x).$$

Then

$$P(T, \mathcal{F}, [B], \epsilon) = \sup_{x \in X} P(T, \mathcal{F}, \epsilon, B \cdot x).$$

It is easy to show that $P(T, \mathcal{F}, [B], \epsilon)$ is a decreasing function of ϵ. Therefore the coset pressure of T with respect to $\mathcal{F}$ and $[B]$ is defined as follows:

$$P(T, \mathcal{F}, [B]) = \lim_{\epsilon \to 0} P(T, \mathcal{F}, [B], \epsilon).$$

Furthermore, we consider a sequence of positive real-valued functions $\{f_n\}$ on X. For a T-invariant Borel probability measure μ, denote

$$\mathcal{F}_*(\mu) = \lim_{n \to \infty} \frac{1}{n} \int \log f_n \, d\mu.$$

A standard subadditive argument assures the existence of the above limit. When $\mu \in \mathcal{E}(X, T)$, by the subadditive ergodic theorem, the above limit exists μ-almost everywhere without integrating against μ.

The basic propositions of coset pressure are provided in [Zhao and Cheng 2014]. The following thermodynamic formalism is the main finding of coset pressure, and gives the relation among $P(T, \mathcal{F}, [B])$, $h_\mu(T \mid [B])$ and $\mathcal{F}_*(\mu)$. The process of the proof is influenced by the methods in [Cao et al. 2008].

Theorem 3.1. *Let $\mathcal{F} = \{\log f_n\}_{n=1}^{\infty}$ be a sequence of subadditive potential functions on the compact group X and let $T : X \to X$ be a continuous endomorphism with a closed T-invariant subgroup B. Then*

$$P(T,\mathcal{F},[B])=\begin{cases}-\infty & \text{if } \mathcal{F}_*(\mu)=-\infty \text{ for all } \mu\in M(X,T),\\ \sup\{h_\mu(T\mid[B])+\mathcal{F}_*(\mu):\mu\in M(X,T),\mathcal{F}_*(\mu)\neq-\infty\} & \text{otherwise.}\end{cases}$$

In physics, a continuous function is regarded as a complicated potential function. Thus, setting $\mathcal{F} = \phi$, the usual variational principle is as follows. Furthermore, we can present the commutativity proposition for coset pressure by using thermodynamic formalism.

$$P(T,\phi,[B]) = \sup_{\mu\in\mathcal{M}(X,T)} \{h_\mu(T\mid[B]) + \int \phi\, d\mu\}.$$

Theorem 3.2. *Assume that both $S, T : X \to X$ are continuous endomorphisms and $\phi : X \to R$ is a continuous real-value function. If B is a closed S, and T-invariant subgroup of X, with coset partition $[B]$. Then, we have*

$$P(S\circ T,\phi,[B]) = P(T\circ S,\phi\circ S,[B]).$$

Proof. With the variational principle and Lemmas 2.4, 2.5, we have

$$\begin{aligned}P(T\circ S,\phi\circ S,[B]) &= \sup_{\mu\in\mathcal{M}(X,T\circ S)} \{h_\mu(T\circ S) + \int \phi\circ S\, d\mu\}\\ &= \sup_{\mu\in\mathcal{M}(X,T\circ S)} \{h_{\widetilde{S}\mu}(S\circ T) + \int \phi\, d(\widetilde{S}\mu)\}\\ &= \sup_{\widetilde{S}\mu\in\mathcal{M}(X,S\circ T)} \{h_{\widetilde{S}\mu}(S\circ T) + \int \phi\, d(\widetilde{S}\mu)\}\\ &= P(S\circ T,\phi,[B]).\end{aligned}$$

□

Using Theorem 3.2, we have the following lemma:

Lemma 3.3. *If $\mathcal{M}(X, T\circ S) = \mathcal{M}(X, S)$, then*

$$P(T\circ S,\phi,[B]) = P(S\circ T,\phi,[B]).$$

Next, assume $\mathcal{F} = \{\log f_n\}_{n=1}^{\infty}$ is a sequence of subadditive potential functions. Let $\mathcal{F}\circ S = \{\log f_n\circ S\}_{n=1}^{\infty}$ and $\mathcal{F}_*\circ S(\mu) = \lim_{n\to\infty} \frac{1}{n}\int \log f_n\circ S\, d\mu$. With the same process as was used for the proof, the next proposition is trivial.

Theorem 3.4. *Assume that $S, T : X \to X$ are continuous endomorphisms and that $\mathcal{F} = \{\log f_n\}_{n=1}^{\infty}$ is a sequence of subadditive potential functions. If B is a closed S, T-invariant subgroup of X, with coset partition $[B]$, then we have*

$$P(S\circ T,\mathcal{F},[B]) = P(T\circ S,\mathcal{F}\circ S,[B]).$$

Similarly, using Theorem 3.4, we have the following lemma:

Lemma 3.5. *If* $\mathcal{M}(X, T \circ S) = \mathcal{M}(X, S)$, *then*

$$P(T \circ S, \mathcal{F}, [B]) = P(S \circ T, \mathcal{F}, [B]).$$

Acknowledgements

This work was carried out when Bing Li visited NCTS at the National Taiwan University and Wen-Chiao Cheng visited the University of Paris-Est Cretéil during the summer of 2016. The authors sincerely appreciate the warm hospitality of the hosts. Wen-Chiao Cheng is partially supported by NSC Grant 105-2115-M-034-001. Bing Li is partially supported by NSFC 11671151 and Guangdong Natural Science Foundation 2018B0303110005.

References

[Balibrea et al. 1999a] F. Balibrea, J. S. Cánovas Peña, and V. Jiménez López, "Commutativity and non-commutativity of topological sequence entropy", *Ann. Inst. Fourier* (*Grenoble*) **49**:5 (1999), 1693–1709. MR Zbl

[Balibrea et al. 1999b] F. Balibrea, V. Jiménez López, and J. S. Cánovas Peña, "Some results on entropy and sequence entropy", *Int. J. Bifur. Chaos Appl. Sci. Eng.* **9**:9 (1999), 1731–1742. MR Zbl

[Bowen 1971] R. Bowen, "Entropy for group endomorphisms and homogeneous spaces", *Trans. Amer. Math. Soc.* **153** (1971), 401–414. MR Zbl

[Brucks and Bruin 2004] K. M. Brucks and H. Bruin, *Topics from one-dimensional dynamics*, Lond. Math. Soc. Student Texts **62**, Cambridge Univ. Press, 2004. MR Zbl

[Cánovas 2011] J. S. Cánovas, "Li–Yorke chaos in a class of nonautonomous discrete systems", *J. Difference Equ. Appl.* **17**:4 (2011), 479–486. MR Zbl

[Cao et al. 2008] Y.-L. Cao, D.-J. Feng, and W. Huang, "The thermodynamic formalism for sub-additive potentials", *Discrete Contin. Dyn. Syst.* **20**:3 (2008), 639–657. MR Zbl

[Cheng 2006] W.-C. Cheng, "Conditional entropy with a coset partition", *Far East J. Dyn. Syst.* **8**:2 (2006), 193–208. MR Zbl

[Falconer 1988] K. J. Falconer, "A subadditive thermodynamic formalism for mixing repellers", *J. Phys. A* **21**:14 (1988), L737–L742. MR Zbl

[Gelfert and Wolf 2008] K. Gelfert and C. Wolf, "Topological pressure via saddle points", *Trans. Amer. Math. Soc.* **360**:1 (2008), 545–561. MR Zbl

[Huang and Yi 2007] W. Huang and Y. Yi, "A local variational principle of pressure and its applications to equilibrium states", *Israel J. Math.* **161** (2007), 29–74. MR Zbl

[Huang et al. 2008] X. Huang, X. Wen, and F. Zeng, "Topological pressure of nonautonomous dynamical systems", *Nonlinear Dyn. Syst. Theory* **8**:1 (2008), 43–48. MR Zbl

[Katok and Hasselblatt 1995] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyc. Math. Appl. **54**, Cambridge Univ. Press, 1995. MR Zbl

[Kolyada and Snoha 1996] S. Kolyada and L. Snoha, "Topological entropy of nonautonomous dynamical systems", *Random Comput. Dynam.* **4**:2-3 (1996), 205–233. MR Zbl

[Kolyada et al. 1999] S. Kolyada, M. Misiurewicz, and L. Snoha, "Topological entropy of nonautonomous piecewise monotone dynamical systems on the interval", *Fund. Math.* **160**:2 (1999), 161–181. MR Zbl

[Kong et al. 2015] M. Kong, W.-C. Cheng, and B. Li, "Topological pressure for nonautonomous systems", *Chaos Solitons Fractals* **76** (2015), 82–92. MR Zbl

[Kuang et al. 2013] R. Kuang, W.-C. Cheng, and B. Li, "Fractal entropy of nonautonomous systems", *Pacific J. Math.* **262**:2 (2013), 421–436. MR Zbl

[Molaei 2008] M. R. Molaei, "Dynamically defined topological pressure", *J. Dyn. Syst. Geom. Theor.* **6**:1 (2008), 75–81. MR Zbl

[Pesin 1997] Y. B. Pesin, *Dimension theory in dynamical systems: contemporary views and applications*, Univ. Chicago Press, 1997. MR Zbl

[Pollner and Vattay 1996] P. Pollner and G. Vattay, "New method for computing topological pressure", *Phys. Rev. Lett.* **76**:22 (1996), 4155–4158.

[Ruelle 1973] D. Ruelle, "Statistical mechanics on a compact set with Z^v action satisfying expansiveness and specification", *Trans. Amer. Math. Soc.* **185** (1973), 237–251. MR Zbl

[Spandl 2008] C. Spandl, "Computability of topological pressure for sofic shifts with applications in statistical physics", *J. Univ. Comput. Sci.* **14**:6 (2008), 876–895. MR Zbl

[Walters 1982] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Math. **79**, Springer, 1982. MR Zbl

[Zhao and Cheng 2014] Y. Zhao and W.-C. Cheng, "Coset pressure with sub-additive potentials", *Stoch. Dyn.* **14**:1 (2014), art. id. 1350012. MR Zbl

Received December 11, 2016. Revised September 3, 2019.

BING LI
SCHOOL OF MATHEMATICS
SOUTH CHINA UNIVERSITY OF TECHNOLOGY
GUANGZHOU
CHINA
scbingli@scut.edu.cn

WEN-CHIAO CHENG
DEPARTMENT OF APPLIED MATHEMATICS
CHINESE CULTURE UNIVERSITY
YANGMINGSHAN
TAIPEI
TAIWAN
zwq2@faculty.pccu.edu.tw

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.229

SIGNATURE INVARIANTS RELATED TO THE UNKNOTTING NUMBER

CHARLES LIVINGSTON

New lower bounds on the unknotting number of a knot are constructed from the classical knot signature function. These bounds can be twice as strong as previously known signature bounds. They can also be stronger than known bounds arising from Heegaard Floer and Khovanov homology. Results include new bounds on the Gordian distance between knots and information about four-dimensional knot invariants. By considering a related nonbalanced signature function, bounds on the unknotting number of slice knots are constructed; these are related to the property of double-sliceness.

1. Introduction

The unknotting number of a knot $K \subset S^3$, denoted $u(K)$, is the minimum number of crossing changes that is required to convert K into an unknot. This is among the most intractable knot invariants. For instance, the unknotting numbers of several 10-crossing knots are still unknown. Scharlemann [1985] proved that the connected sum of two unknotting number 1 knots has unknotting number 2, but little beyond this is known concerning the additivity of the unknotting number.

Many knot invariants offer tools for estimating the unknotting number; these include the rank of the homology of branched covers [Kinoshita 1957; Wendt 1937], the Murasugi signature [1965], $\sigma(K)$, the Levine–Tristram signature function [Levine 1969; Tristram 1969], $\sigma_K(\omega)$, defined for $\omega \in S^1 \subset \mathbb{C}$, and the Witt group of the rational numbers, $W(\mathbb{Q})$, studied in [Jabuka 2009]. Heegaard Floer homology and Khovanov homology have provided smooth knot invariants τ, Υ, and s (see [Ozsváth and Szabó 2003; Ozsváth et al. 2017; Rasmussen 2010]) that also offer lower bounds on the unknotting number. (See also [Cochran and Lickorish 1986; 2008; Owens 2010; Owens and Strle 2016].) The precise bound on the unknotting number that has been known to arise from the signature function is easily described. Let $a = \max(\sigma_K(\omega))$ and $b = \min(\sigma_K(\omega))$. Then $u(K) \geq (a-b)/2$. Here we will observe that the knot signature function offers much stronger constraints on the

This work was supported by a grant from the National Science Foundation, NSF-DMS-1505586.
MSC2010: 57M25.
Keywords: knot, unknotting number, signature.

unknotting number; in some cases the new bounds will be seen to be twice as strong as this previously known bound. Examples also demonstrate that the new bounds can exceed those arising from Heegaard Floer and Khovanov homology.

There is a refined version of the unknotting number that incorporates the signs of the crossing changes that unknot K. Let $U(K)$ be the set of integer pairs (p, n) for which K can be unknotted using p crossing changes from positive to negative and n crossing changes from negative to positive. Then $U(K)$ is called the *signed unknotting set of* K. Observe that $u(K) = \min\{p+n \mid (p, n) \in U(K)\}$. Finding constraints on $U(K)$ is especially difficult. The results we present here depend critically on the signs of crossing changes, and thus they are able to extract information about $U(K)$ that cannot be attained with previously known techniques. In turn, these can be used to strengthen the bounds on $u(K)$.

The invariants we develop here also provide lower bounds on the *Gordian distance* between knots K and J, denoted $d_g(K, J)$; this is the minimum number of crossing changes that are required to convert K into J. Clearly $d_g(K, J) \le u(K) + u(J)$; lower bounds are more difficult to find.

The results presented here also have applications to four-dimensional knot invariants. For instance, we provide new lower bounds on the *clasp number* of knots; this invariant is defined to be the minimum number of transverse double points in an immersed disk in B^4 bounded by K; it is also referred to as the *four-ball crossing number* and is related to the notion of *kinkiness* defined by Gompf [1984].

The signature function is built from a *nonbalanced signature function*, $s_K(\omega)$. The two functions agree almost everywhere, but s_K is not a concordance invariant. In a final section we discuss how s_K provides bounds on the unknotting number that can be nontrivial for slice knots, and we present applications of this to double slicing of knots, a concept dating to such work as [Sumners 1971; Terasaka and Hosokawa 1961].

1.1. ***Outline and summary of results.*** In Section 2 we will review the definition of the signature function of a knot, $\sigma_K(\omega)$. This is an integer-valued step function on the set of unit length complex numbers $\omega \in S^1 \subset \mathbb{C}$; discontinuities can occur only at roots of the Alexander polynomial, $\Delta_K(t)$. The definition of σ_K is such that at each discontinuity its value is equal to its two-sided average at that point. There is also a related jump function,

$$J_K(e^{2\pi i t}) = \tfrac{1}{2}\Big(\lim_{\tau\to t^+} \sigma_K(e^{2\pi i\tau}) - \lim_{\tau\to t^-} \sigma_K(e^{2\pi i\tau})\Big).$$

The signature function is defined in terms of a Witt group; in Sections 2 and 3 we study this group and how crossing changes affect the Witt class associated to a knot. Section 3 presents the proof of our key result. In the statement of the theorem and throughout this paper, we denote the unit circle in the complex plane by $\mathbb{S}^1$.

Proposition 1. *Let K_+ be a knot, let δ be an irreducible rational polynomial, and let $\{\alpha_1, \dots, \alpha_k\} \subset \mathbb{S}^1$ with $k > 0$ satisfy $\delta(\alpha_i) = 0$ for all i. If a crossing in a diagram for K_+ is changed from positive to negative to yield a knot K_-, then one of the following two possibilities occurs.*

(1) *For every α_i, $J_{K_-}(\alpha_i) - J_{K_+}(\alpha_i) = 0$ and $\sigma_{K_-}(\alpha_i) - \sigma_{K_+}(\alpha_i) \in \{0, 2\}$.*

(2) *For every α_i, $J_{K_-}(\alpha_i) - J_{K_+}(\alpha_i) \in \{-1, 1\}$ and $\sigma_{K_-}(\alpha_i) - \sigma_{K_+}(\alpha_i) = 1$.*

In Section 4 we prove a corollary to this proposition.

Theorem 2. *Let $K \subset S^3$ be a knot, let $\delta(x)$ be a rational irreducible polynomial, and let $\{\alpha_1, \dots, \alpha_k\} \subset \mathbb{S}^1$ with $k > 0$ satisfy $\delta(\alpha_i) = 0$ for all i. Let $\mathfrak{J}_\delta$ denote the maximum of $\{|J_K(\alpha_i)|\}$ and let $\underline{\mathfrak{S}}_\delta$ and $\overline{\mathfrak{S}}_\delta$ denote the minimum and maximum of $\{\sigma_K(\alpha_i)\}$, respectively. Suppose that $\overline{\mathfrak{S}}_\delta \geq 0$.*

(1) *If $\underline{\mathfrak{S}}_\delta \leq \mathfrak{J}_\delta$, then $u(K) \geq \mathfrak{J}_\delta + (\overline{\mathfrak{S}}_\delta - \underline{\mathfrak{S}}_\delta)/2$.*

(2) *If $\underline{\mathfrak{S}}_\delta \geq \mathfrak{J}_\delta$, then $u(K) \geq (\mathfrak{J}_\delta + \overline{\mathfrak{S}}_\delta)/2$.*

Note. Letting $-K$ denote the mirror image of K, we have that $\sigma_{-K}(\omega) = -\sigma_K(\omega)$. We also have that $u(-K) = u(K)$. Thus, the condition $\overline{\mathfrak{S}} \geq 0$ does not limit the generality of Theorem 2. The set of polynomials that are relevant in applying this theorem are symmetric factors of the Alexander polynomial of K, $\Delta_K(x)$. The strongest obstructions arise by letting $\{\alpha_1, \dots, \alpha_k\}$ be the full set of unit length roots of δ.

Section 4 also presents an analog of Theorem 2 in the case of signed unknotting numbers.

Theorem 3. *Let K, $\mathfrak{J}_\delta$, $\underline{\mathfrak{S}}_\delta$, and $\overline{\mathfrak{S}}_\delta$ be as in the statement of Theorem 2. Suppose that $\overline{\mathfrak{S}}_\delta \geq 0$.*

(1) *If $\underline{\mathfrak{S}}_\delta \leq \mathfrak{J}_\delta$, then unknotting K requires at least $(\mathfrak{J}_\delta + \overline{\mathfrak{S}}_\delta)/2$ negative to positive crossings and $(\mathfrak{J}_\delta - \underline{\mathfrak{S}}_\delta)/2$ positive to negative crossing changes.*

(2) *If $\underline{\mathfrak{S}}_\delta \geq \mathfrak{J}_\delta$, then unknotting K requires at least $(\mathfrak{J}_\delta + \overline{\mathfrak{S}}_\delta)/2$ negative to positive crossing changes.*

In Section 5 we observe that the signed unknotting data obtained from different choices of polynomials can be complementary. Using this, we provide examples for which combining the bounds that arise from different polynomials yields bounds on the (unsigned) unknotting number that are stronger than what can be obtained from either one of the polynomials.

In Section 6 we construct explicit examples to demonstrate that the bounds on the unknotting number provided by Theorem 2 can be twice as strong as previously known signature bounds. We also prove that our new bounds cannot exceed twice the classical bound.

Section 7 discusses the application of these results to bounding the Gordian distance between knots.

In Section 8 we describe a four-dimensional perspective on these results. The obstructions we develop actually bound the number of crossing changes required to convert K into a knot with trivial signature function. Thus, they also bound the number of crossing changes required to convert K into a slice knot (the *slicing number* of K) and the number of crossing changes required to convert K into an algebraically slice knot (the *algebraic slicing number*). Past work on these invariants includes [Livingston 2002; Owens 2010; Owens and Strle 2016].

In the remainder of Section 8 the focus is on the clasp number [Murakami and Yasuhara 2000] of the knot K, which is the minimum number of transverse double points in an immersed disk bounded by K in the four-ball. In the course of the work we also present a new simplified proof of a result in [Livingston 2011] that offers strong bounds on the cobordism distance between knots K and J; this is the minimum genus of a cobordism (W, F) between (S^3, K) and (S^3, J) with $W \cong S^3 \times I$. References include [Baader 2006; 2012; Feller 2016; Feller and Krcatovich 2017; Hirasawa and Uchida 2002; Kawamura 2002; Owens 2010; Owens and Strle 2016].

In Section 9 we briefly discuss the nonbalanced signature function, $s_K(\omega)$, defined as the signature of the matrix denoted W_F in Section 2. (The standard signature function, $\sigma_K(\omega)$, is built as the two-sided average of $\sigma_K(\omega)$. The two functions agree almost everywhere, but $s_K(\omega)$ is not a concordance invariant. As explained in the section, $s_K(\omega)$ provides bounds on the unknotting number of slice knots; from the four-dimensional perspective it is related to double-sliceness of knots.

1.2. *Example.* To conclude this introduction, we provide a simple example illustrating Theorem 2.

Example 4. We prove the knot $5_1 \# 10_{132}$ has unknotting number 3. To simplify our work, we let $K = -5_1 \# -10_{132}$ and prove $u(K) = 3$. Working with the standard diagrams for 5_1 and 10_{132}, such as illustrated in [Cha and Livingston 2019; Rolfsen 1976], one can quickly show that their unknotting numbers are at most 2 and 1, respectively, and thus $u(K) \le 3$. We will prove that $u(K) = 3$ by showing $u(K) \ge 3$.

The signature functions for -5_1 and 10_{132} and the signature function for the difference, K, are illustrated in Figure 1, graphed as functions of t, where $\omega = e^{2\pi i t}$, $0 \le t \le 1/2$. Let δ be the tenth cyclotomic polynomial, ϕ_{10}, having roots $\omega_1 = e^{2\pi i(1/10)}$ and $\omega_2 = e^{2\pi i(3/10)}$ on the upper half circle. As seen in Figure 1, the jumps at ω_1 and ω_2 for K are 0 and 2, respectively. The signatures are 0 and 2 at these points.

In the notation of Theorem 2 we have $\mathfrak{J}_\delta = 2$, $\underline{\mathfrak{S}}_\delta = 0$, and $\overline{\mathfrak{S}}_\delta = 2$, and from that theorem we have $u(K) \ge 2 + (2-0)/2 = 3$, as desired. For this knot, the classical

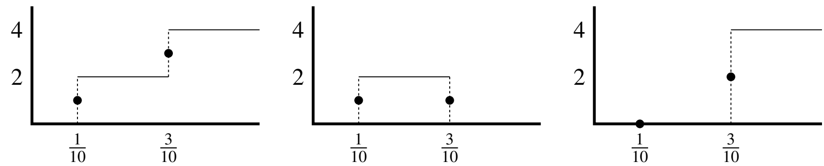

Figure 1. Signature functions for -5_1, 10_{132} and $-5_1 \# -10_{132}$.

lower bound on the unknotting number that arises from the signature function is 2. (The Rasmussen invariant s, the tau invariant τ and the Upsilon invariant, Υ, all provide lower bounds of 1. For the first two, the values have been tabulated [Cha and Livingston 2019]. Because 10_{132} is nonalternating, the computation of Υ_K is more complicated and will not be presented here.)

Applying Theorem 3, we see that to unknot $5_1 \# 10_{132}$ requires at least two crossing changes from positive to negative and one crossing change from negative to positive.

2. Witt class invariants of knots and the signature function

2.1. ***The Witt class of a knot.*** Let $F \subset S^3$ denote a genus g compact oriented surface with connected boundary K. We will also write F to denote the surface along with a choice of homology basis, calling this a *based surface*. Associated to F there is a $2g \times 2g$ Seifert matrix V_F. Given V_F, there is the matrix $W_F \in M_{2g,2g}(\mathbb{Q}(x))$ defined by

$$W_F = (1-x)V_F + (1-x^{-1})V_F^{\mathsf{T}}.$$

Here $\mathbb{Q}(x)$ is the quotient field of $\mathbb{Q}[x, x^{-1}]$. Elements of $\mathbb{Q}[x, x^{-1}]$ are Laurent polynomials; we will refer to elements of $\mathbb{Q}[x, x^{-1}]$ simply as polynomials.

For future reference, we recall that the Alexander polynomial of K is given by $\Delta_K(x) = \det(V_F - xV_F^{\mathsf{T}}) \in \mathbb{Z}[x, x^{-1}]$ and note that $\det(W_F) = (1-x)^{2g}\Delta_K(x^{-1})$. The Alexander polynomial is well-defined up to multiplication by $\pm x^k$ for some k.

The Witt group $W(\mathbb{Q}(x))$ is defined to be the set of equivalence classes of nonsingular hermitian matrices with coefficients in $\mathbb{Q}(x)$, a field with involution $x \to x^{-1}$. Two such matrices, A and B, of ranks m and n, are called Witt equivalent if $m \equiv n \mod 2$ and the form defined by $A \oplus -B$ vanishes on a subspace of dimension $(m+n)/2$. The group structure on $W(\mathbb{Q}(x))$ is induced by direct sums and inversion is given by multiplication by -1. Congruent matrices represent the same element in the Witt group. For details concerning this Witt group, see [Litherland 1984]. We have the following fundamental result of Levine [1969].

Proposition 5. *If F_1 and F_2 are based Seifert surfaces for a knot K, then W_{F_1} is Witt equivalent to W_{F_2}.*

This permits us to define $W_K \in W(\mathbb{Q}(x))$ to be the Witt class represented by W_F for an arbitrary choice of based Seifert surface F for K.

2.2. ***The signature function of a Witt class.*** Suppose that $\mathsf{w} \in W(\mathbb{Q}(x))$ can be represented by two matrices, $A(x)$ and $B(x)$. For almost all $\alpha \in \mathbb{S}^1$, the matrices $A(\alpha)$ and $B(\alpha)$ are defined and nonsingular. For all such α, the signatures of $A(\alpha)$ and $B(\alpha)$, denoted σ_A and σ_B, will be equal. Thus for all real t, there is an equality of limits:

$$\tfrac{1}{2}\Big(\lim_{\tau\to t^+}\sigma_A(e^{2\pi i\tau})+\lim_{\tau\to t^-}\sigma_A(e^{2\pi i\tau})\Big)=\tfrac{1}{2}\Big(\lim_{\tau\to t^+}\sigma_B(e^{2\pi i\tau})+\lim_{\tau\to t^-}\sigma_B(e^{2\pi i\tau})\Big).$$

For $\omega = e^{2\pi i t}$, we denote this limit $\sigma_{\mathsf{w}}(\omega)$ and for $\mathsf{w} = W_K$, we denote it $\sigma_K(\omega)$. This is a step function that is integer-valued except perhaps at its discontinuities, where it equals its two-sided average. Modulo 2, its value (except at the discontinuities) equals the rank of a representative; thus, for a knot, it is even-valued away from the discontinuities and is integer-valued at the discontinuities. As in the introduction, for such a w we write

$$J_{\mathsf{w}}(e^{2\pi i t})=\tfrac{1}{2}\Big(\lim_{\tau\to t^+}\sigma_{\mathsf{w}}(e^{2\pi i\tau})-\lim_{\tau\to t^-}\sigma_{\mathsf{w}}(e^{2\pi i\tau})\Big),$$

and in the case $\mathsf{w} = W_K$ we write $J_K(\omega)$.

Both of the functions σ_K and J_K are invariant under complex conjugation. They are defined on the set of unit complex numbers, which we henceforth write as $\mathbb{S}^1 = \{\omega \in \mathbb{C} \mid |\omega| = 1\}$. A fairly simple exercise shows that for a knot K, the fact that $\det(V_K - V_K^{\mathsf{T}}) = \pm 1$ implies that $\sigma_K(\omega) = 0$ for all ω close to 1. Given the properties of σ_K, when we graph $\sigma_K(e^{2\pi i t})$, we will restrict to $t \in (0, 1/2)$.

2.3. ***The signature and the four-genus.*** We now briefly summarize a well-known result that follows immediately from [Taylor 1979].

Theorem 6. *If K bounds a surface of genus h in B^4, for instance if $g_4(K) = h$, then W_K has a $2h$-dimensional representative.*

Proof. According to [Taylor 1979], if K bounds a Seifert surface of genus g and bounds a surface of genus $h \le g$ in B^4, then with respect to some basis, the upper left $(g-h)\times(g-h)$ block of the Seifert matrix V_F has all entries 0. It then follows that W_F is Witt equivalent to a sum $A \oplus B$, where A is $2(g-h)\times 2(g-h)$ and is Witt trivial, and B is $2h \times 2h$. □

Corollary 7. *For all $t \in (0, 2)$, $g_4(K) \ge \frac{1}{2}|\sigma_K(e^{2\pi i t})|$.*

3. Diagonalization and crossing changes

3.1. ***Diagonalization.*** The field $\mathbb{Q}(t)$ has characteristic 0, and thus the matrix W_F associated to a Seifert surface F is congruent to a diagonal matrix; that is, it can be diagonalized using simultaneous row and column operations. We will write a diagonalization by listing its diagonal elements: $[d_1, d_2, \ldots, d_{2g}]$. By scaling the corresponding basis of the underling vector space, we can clear the denominators of these diagonal elements and divide by factors of the form $f(t)f(t^{-1})$. Thus, we can assume that each d_i is a product of distinct irreducible symmetric polynomials in $\mathbb{Q}[t, t^{-1}]$.

We now have the following.

Theorem 8. *Let* $\mathsf{w} \in W(\mathbb{Q}(t))$, *let* δ *be an irreducible symmetric polynomial, and let* $\boldsymbol{\alpha} = \{\alpha_i\} \subset \mathbb{S}^1$ *denote a subset of the roots of* δ *that lie on the unit circle. Then if* w *is represented by an* $N \times N$ *matrix, we have*

$$N \geq \left(\max_{\alpha_i \in \boldsymbol{\alpha}}\{|J_K(\alpha_i)|\} + \max_{\alpha_i \in \boldsymbol{\alpha}}\{|\sigma_K(\alpha_i)|\}\right).$$

For all i *and* j, $J_K(\alpha_i) = \sigma_K(\alpha_j) \bmod 2$.

Proof. Choose a diagonal representation of w of the form $[f_1\delta, \ldots, f_m\delta, g_1, \ldots, g_n]$, where the f_k and g_k are symmetric polynomials that are relatively prime to δ. The jump at α_i is given as the sum of the jumps, each ± 1, arising from the diagonal elements of the form $f_k\delta$. Thus, $m \geq |J(\alpha_i)|$ for all i. The signature is determined by the signs of the g_k at α_i. Thus, $n \geq |\sigma_K(\alpha_i)|$ for all i. This completes the proof of the inequality.

To prove the last statement, concerning the parities of the jumps and signatures, we observe that modulo 2, $J_K(\alpha_i) = m \bmod 2$ and $\sigma_K(\alpha_j) = n \bmod 2$. In addition, $m + n = N$. Finally, for a knot K, W_F is a $2g \times 2g$ matrix. The proof is completed by noting that Witt equivalence preserves the rank of a representative, modulo 2. □

3.2. ***Crossing changes.*** In considering signed crossing changes, the following result is useful. A proof can be constructed from a careful examination of Seifert's algorithm for constructing Seifert surfaces. One proof is presented in [Kim and Livingston 2005], where the focus is on the effect of crossing changes on the Alexander polynomial.

Theorem 9. *If* K_+ *and* K_- *differ by a crossing change from positive to negative, then they bound Seifert surfaces* F_+ *and* F_- *of the same genus,* g, *with the following property: for appropriate choices of bases for homology, the Seifert forms are identical except for the lower right entry:* $V_{F_-}^{2g,2g} - V_{F_+}^{2g,2g} = 1$.

Lemma 10. *Let* δ *be a symmetric irreducible polynomial. The Witt classes for* K_+ *and* K_- *decompose as*

$$W_{K_\pm} = [f_1\delta, \ldots, f_m\delta, g_1, \ldots, g_n] \oplus \begin{pmatrix} a(x) & \overline{b(x)} \\ b(x) & d(x) + \epsilon_\pm(1-x)(1-x^{-1}) \end{pmatrix},$$

where the f_i *and* g_i *are symmetric polynomials that are relatively prime to* δ, $\epsilon_+ = 0$, *and* $\epsilon_- = 1$. *Furthermore,* $m + n + 2 = 2g$.

Proof. Consider the matrix representation of $W_{K_\pm}$ determined by the Seifert forms given in Theorem 9. The determinant is nonzero: an elementary exercise in linear algebra shows that the upper left $(2g-1)\times(2g-1)$ submatrix has nullity at most 1. Thus, this block can be diagonalized via a change of basis so that the first $(2g-2)$ diagonal entries are nonzero. The resulting $2g\times 2g$ matrix can have nonzero entries in the last column and bottom row, but the diagonal entries can be used to clear these out, with the possible exception of the last two rows and columns. This yields the desired decomposition. □

We can now prove Proposition 1, which we restate.

Proposition 1. *Let K_+ be a knot, let δ be an irreducible rational polynomial, and let $\{\alpha_1,\ldots,\alpha_k\}\subset\mathbb{S}^1$ with $k>0$ satisfy $\delta(\alpha_i)=0$ for all i. If a crossing in a diagram for K_+ is changed from positive to negative to yield a knot K_-, then one of the following two possibilities occurs.*

(1) *For every α_i, $J_{K_-}(\alpha_i)-J_{K_+}(\alpha_i)=0$ and $\sigma_{K_-}(\alpha_i)-\sigma_{K_+}(\alpha_i)\in\{0,2\}$.*

(2) *For every α_i, $J_{K_-}(\alpha_i)-J_{K_+}(\alpha_i)\in\{-1,1\}$ and $\sigma_{K_-}(\alpha_i)-\sigma_{K_+}(\alpha_i)=1$.*

Proof. The difference $W_{K_-}-W_{K_+}$ is represented by the differences of the corresponding 2×2 block matrices given in Lemma 10, so we restrict our attention to these, calling them w_- and w_+. If the entry $a(x)$ is 0, then w_+ and w_- are both Witt trivial, so the difference of jumps is 0, as is the difference of the signature; thus Case (1) is satisfied.

If $a(x)\neq 0$, then the forms can be further diagonalized so that the only place at which they differ is the last diagonal element. This diagonal element will be of the form

$$\frac{p(x)}{q(x)}+\epsilon_\pm(1-x)(1-x^{-1}),$$

for some $p(x)$ and $q(x)$. We write the 1×1 matrices as

$$v_+=\left(\frac{p(x)}{q(x)}\right)\quad\text{and}\quad v_-=\left(\frac{p(x)}{q(x)}+(1-x)(1-x^{-1})\right).$$

We will refer to the entries of these two matrices as $v_+(x)$ and $v_-(x)$. It remains to analyze the jump functions and signature functions associated to these two matrices.

For each value of i, we associate to $v_\pm$ the jump and signature of the form at α_i, denoting these $j^i_\pm$ and $\sigma^i_\pm$. We proceed in a series of steps.

Step 1: Consider v_+. At points α close to but not equal to α_i, the signature is either 1 or -1. If the signature changes sign at α_i, then $|j^i_+|=1$ and $\sigma^i_+=0$. On the other hand, if the signature doesn't change sign, then $|j^i_+|=0$ and $\sigma^i_+=\pm 1$. The same properties hold for v_-.

Step 2: Since $(1-\omega)(1-\omega^{-1}) > 0$ for all $\omega \in \mathbb{S}^1$ with $\omega \neq 1$, we have that $\sigma^i_- - \sigma^i_+ \geq 0$.

Step 3: Given Steps 1 and 2, the only possible nontrivial changes of the pairs $(|j^i_+|, \sigma^i_+) \to (|j^i_-|, \sigma^i_-)$ are:

- Type 1: $(0,-1) \to (0,1)$.
- Type 2: $(0,-1) \to (1,0)$.
- Type 3: $(1,0) \to (0,1)$.

These are consistent with the statement of Proposition 1.

Step 4: The proof of the proposition is completed by showing that if a nontrivial change of Type 2 or Type 3 occurs at some α_i, then the same change occurs at all α_i. After changes of basis, the forms can be written as Witt equivalent forms, for which we use the same names,

$$v_+ = (f_+(x)\delta(x)^{\epsilon_+}), \quad v_+ = (f_-(x)\delta(x)^{\epsilon_-}).$$

Here $f_\pm$ are symmetric polynomials that are relatively prime to δ and $\epsilon_\pm$ are either 0 or 1. There are four cases to consider.

- If $(\epsilon_+, \epsilon_-) = (0,0)$, then there are no nontrivial jumps at any α_i, so no changes of Type 2 or 3 occur.
- If $(\epsilon_+, \epsilon_-) = (1,0)$, then at each α_i there is a jump for v_+ but not for v_-, so for all α_i we see a change of Type 3.
- If $(\epsilon_+, \epsilon_-) = (0,1)$, then at each α_i there is no v_+ jump but for v_- there is a jump, so for all α_i we see a change of Type 2.
- If $(\epsilon_+, \epsilon_-) = (1,1)$, then at each α_i both v_+ and v_- have nonzero jumps, so no change of Type 2 or 3 occurs at any of the α_i.

Together, these steps complete the proof of the proposition. □

Proof. Changes of the first type in Proposition 1 clearly leave $\mathfrak{J}$ unchanged and leave each of $\underline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ unchanged or increased by 2.

Changes of the second type, since they increase every signature by 1, increase the minimum and maximum signature by 1. The condition on the change in $\mathfrak{J}$ is a little more subtle. For instance, if it were possible that one jump is 1 and one jump is 2, then after the change, it might be that the first jump is 2 and the second is 1, and thus the maximum absolute value would not change. However, as stated in Theorem 8, the jumps all have the same parity. Thus, the parity of the jumps switches for such a crossing change, so the maximum absolute value must also change. □

4. Bounds on the unknotting number and signed unknotting number

To begin, we have a corollary of Proposition 1.

Corollary 11. *Let $K \subset S^3$ be a knot and let $\{\alpha_1, \dots, \alpha_k\} \subset \mathbb{S}^1$ be a nonempty subset of the complex roots of an irreducible rational polynomial. Let $\mathfrak{J}$ denote the maximum of $\{|J_K(\alpha_i)|\}$ and let $\underline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ denote the minimum and maximum of $\{\sigma_K(\alpha_i)\}$. A crossing change in K from positive to negative either leaves $\mathfrak{J}$ unchanged and leaves each of $\underline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ unchanged or increased by 2; or it changes $\mathfrak{J}$ by 1 and increases both $\underline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ by 1.*

4.1. *Unsigned unknotting number bounds.* Our main goal in this section is the following theorem, as stated in the introduction.

Theorem 2. *Let $K \subset S^3$ be a knot and let $\{\alpha_1, \dots, \alpha_k\} \subset \mathbb{S}^1$ be a nonempty subset of the complex roots of an irreducible rational polynomial. Let $\mathfrak{J}$ denote the maximum of $\{|J_K(\alpha_i)|\}$ and let $\underline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}$ denote the minimum and maximum of $\{\sigma_K(\alpha_i)\}$. Suppose that $\overline{\mathfrak{S}} \geq 0$. If $\underline{\mathfrak{S}} \leq \mathfrak{J}$ then $u(K) \geq \mathfrak{J} + (\overline{\mathfrak{S}} - \underline{\mathfrak{S}})/2$. If $\underline{\mathfrak{S}} \geq \mathfrak{J}$ then $u(K) \geq (\mathfrak{J} + \overline{\mathfrak{S}})/2$.*

Proof. We consider the set

$$\Lambda = \{(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) \in \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \mid \mathfrak{j} = \underline{\mathfrak{s}} = \overline{\mathfrak{s}} \bmod 2\}.$$

We define two sets of functions from Λ to itself. The first set consists of what we call *F-type functions*. These, which do not change the value of $\mathfrak{j}$, are as follows:

$$\begin{aligned} F_1^-(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}} - 2, \overline{\mathfrak{s}}), & F_2^-(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}} - 2), \\ F_3^-(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}} - 2, \overline{\mathfrak{s}} - 2), & F_1^+(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}} + 2, \overline{\mathfrak{s}}), \\ F_2^+(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}} + 2), & F_3^+(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j}, \underline{\mathfrak{s}} + 2, \overline{\mathfrak{s}} + 2). \end{aligned}$$

Functions of the second type, *G-type functions*, change the value of $\mathfrak{j}$. These are defined as follows:

$$\begin{aligned} G_1^-(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j} - 1, \underline{\mathfrak{s}} - 1, \overline{\mathfrak{s}} - 1), & G_2^-(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j} + 1, \underline{\mathfrak{s}} - 1, \overline{\mathfrak{s}} - 1), \\ G_1^+(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j} - 1, \underline{\mathfrak{s}} + 1, \overline{\mathfrak{s}} + 1), & G_2^+(\mathfrak{j}, \underline{\mathfrak{s}}, \overline{\mathfrak{s}}) &= (\mathfrak{j} + 1, \underline{\mathfrak{s}} + 1, \overline{\mathfrak{s}} + 1). \end{aligned}$$

For a given knot K, a crossing change affects the value of the associated pair $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ by applying one of these functions. The superscripts $+$ and $-$ correspond to whether the crossing changes are positive to negative or negative to positive, respectively. A sequence of crossing changes that results in an unknot yields a sequence of these functions which in composition carry $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ to $(0, 0, 0)$.

We now consider a given element $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}}) \in \Lambda$. For the proof of the theorem, we can assume $\mathfrak{J} \geq 0$ and $\underline{\mathfrak{S}} \leq \overline{\mathfrak{S}}$. We ask for the minimum length of a sequence of these functions that can reduce $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ to $(0, 0, 0)$. A simple observation is that the F-type functions commute with the G-type functions, so we can assume that a minimal length sequence consists of a sequence of G-type functions followed by a sequence of F-type functions. (Here, the order of the sequence is in terms of the order of composition; in function notation, $f \circ g$ denotes g followed by f.)

Since the F-type functions do not change the value of $\mathfrak{j}$, the initial application of the G-type functions reduces $\mathfrak{J}$ to 0. It follows that by commuting elements in the initial sequence of G-type functions, we can assume the sequence begins with $\mathfrak{J}$ terms of type $G_1^{\pm}$ (which together decrease the $\mathfrak{j}$-coordinate to 0) followed by a sequence of G-type functions that alternately increase and decrease the $\mathfrak{j}$-coordinate by 1.

Next observe that a pair of G-functions that raise and then lower the $\mathfrak{j}$-coordinate compose to give a single function, either the identity or one of F_3^- or F_3^+. Thus, in a minimum length sequence, such pairs do not appear, and hence there are precisely $\mathfrak{J}$ of the G-type functions followed by a sequence of F-type functions.

If one considers all possible sequences of G-type functions of length $\mathfrak{J}$ that convert $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ to a triple with $\mathfrak{j}$-coordinate 0, the possible ending values of $(\underline{\mathfrak{s}}, \overline{\mathfrak{s}})$ are $(\underline{\mathfrak{S}}+\alpha, \overline{\mathfrak{S}}+\alpha)$, where $-\mathfrak{J} \le \alpha \le \mathfrak{J}$. Each F-type function reduces the difference $\overline{\mathfrak{S}} - \underline{\mathfrak{S}}$ by at most 2. Thus at least

$$((\overline{\mathfrak{S}}+\alpha) - (\underline{\mathfrak{S}}+\alpha))/2 = (\overline{\mathfrak{S}} - \underline{\mathfrak{S}})/2$$

applications of F-type functions are required to reduce this pair to (0, 0). In fact, if for some α the interval $(\underline{\mathfrak{S}}+\alpha, \overline{\mathfrak{S}}+\alpha)$ contains 0, a sequence of that length will suffice. There will be such an α if $\underline{\mathfrak{S}} \le \mathfrak{J}$. Thus, in this setting the minimum length sequence is $\mathfrak{J} + (\overline{\mathfrak{S}} - \underline{\mathfrak{S}})/2$, as desired.

On the other hand, if $\underline{\mathfrak{S}} > \mathfrak{J}$, then we also have $\overline{\mathfrak{S}} > \mathfrak{J}$. In this case, the sequence of G-type functions has reduced the $\overline{\mathfrak{s}}$-coordinate to no less than $\overline{\mathfrak{S}} - \mathfrak{J}$, so at least another $(\overline{\mathfrak{S}} - \mathfrak{J})/2$ steps are required. Thus, the minimal length of the sequence is at least

$$\mathfrak{J} + (\underline{\mathfrak{S}} - \mathfrak{J})/2 + (\overline{\mathfrak{S}} - \underline{\mathfrak{S}})/2 = (\mathfrak{J} + \overline{\mathfrak{S}})/2. \qquad \square$$

4.2. ***Signed unknotting number bounds.*** In the proof of Theorem 2, at one step we considered the condition that an interval $[\underline{\mathfrak{S}}+\alpha, \overline{\mathfrak{S}}+\alpha]$ contained 0. If the argument is examined closely, in the case that $\underline{\mathfrak{S}} < \mathfrak{J}$ there can be more than one α for which this holds. This will complicate the count of negative and positive shifts that will appear in the sequence of functions that reduce the jumps and signatures to 0.

Theorem 3. *Let K and $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ be as in Theorem 2. Suppose that $\overline{\mathfrak{S}} \ge 0$.*

(1) *If $\underline{\mathfrak{S}} \le \mathfrak{J}$, then unknotting K requires at least $(\mathfrak{J} + \overline{\mathfrak{S}})/2$ negative to positive crossings and $(\mathfrak{J} - \underline{\mathfrak{S}})/2$ positive to negative crossing changes.*

(2) *If $\underline{\mathfrak{S}} \ge \mathfrak{J}$, then unknotting K requires at least $(\mathfrak{J} + \overline{\mathfrak{S}})/2$ negative to positive crossing changes.*

Proof. Suppose that the sequence of functions that reduces $\mathfrak{J}$ to 0 has a terms that lower the $\underline{\mathfrak{s}}$ and $\bar{\mathfrak{s}}$ coordinates. That sequence has $\mathfrak{J}-a$ terms that increase the $\underline{\mathfrak{s}}$ and $\bar{\mathfrak{s}}$ coordinates. The application of these functions carries the pair $(\underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ to $(\underline{\mathfrak{S}}+\mathfrak{J}-2a, \overline{\mathfrak{S}}+\mathfrak{J}-2a)$. Assume this interval contains 0. Then the sequence of F-type functions that carry this pair to $(0,0)$ must have $-(\underline{\mathfrak{S}}+\mathfrak{J}-2a)/2$ terms that increase the smaller coordinate and $(\overline{\mathfrak{S}}+\mathfrak{J}-2a)/2$ terms that decrease the larger coordinate. Summing these counts gives the desired result. □

Example 12. From Example 4 we see $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}})$ is $(2,0,2)$ or $(2,2,4)$ for the knots $-5_1 \# -10_{132}$ and $-5_1, \#10_{132}$, respectively. In both cases $\underline{\mathfrak{S}} \le \mathfrak{J}$. Applying Theorem 3, we see unknotting $-5_1 \# -10_{132}$ requires at least two crossing changes from negative to positive and one crossing change from positive to negative. To unknot $-5_1 \# 10_{132}$ requires at least three crossing changes from negative to positive.

5. Polynomial splittings and signed unknotting numbers

The bounds on the unknotting number developed in the previous sections depend on a choice of polynomial. This section presents an example for which there are two relevant polynomials to consider. Either one provides a lower bound of three for the unknotting number. However, for one of the polynomials, when signs are considered it will be seen that unknotting requires at least two changes from negative to positive and one change from positive to negative. Using the other polynomial, we will see that at least three changes from negative to positive are required. Combining these two results, we see that at least three changes from negative to positive and one change from positive to negative are required, and hence the unknotting number must be at least four.

Example 13. We consider the knot $K = 2(3_1) - 5_1 - 8_2 + 10_{132} - 11_{n6}$. Figure 2 illustrates the graph of its signature function. The scale is such that $\sigma_K(\alpha_1) = 2$. The data we use, including the signature function, can be found in [Cha and Livingston 2004].

The relevant roots of the Alexander polynomial are of the form $e^{2\pi i t}$ with $0 < t < 1/2$.

- $K = 3_1$:
  (1) $\delta_1 = \Delta_K(x) = x^2 - x + 1$.
  (2) roots (γ), $t \approx 0.167$.
- $K = 5_1$ or $K = 10_{132}$:
  (1) $\delta_2 = \Delta_K(x) = x^4 - x^3 + x^2 - x + 1$.
  (2) roots (α_1, α_2), $t = 0.1$, $t = 0.3$.
- $K = 8_2$ or $K = 11_{n6}$:
  (1) $\delta_3 = \Delta_K(x) = x^6 - 3x^5 + 3x^4 - 3x^3 + 3x^2 - 3x + 1$.
  (2) roots (β_1, β_2), $t \approx 0.132$, $t \approx 0.322$.

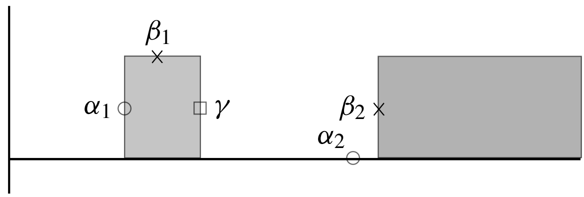

Figure 2. Signature function for $2(3_1) - 5_1 - 8_2 + 10_{132} - 11_{n6}$.

The jump and signature data is as follows:

- $\mathfrak{J}_{\delta_1}(K) = 2,\ \underline{\mathfrak{S}}_{\delta_1}(K) = 2,\ \overline{\mathfrak{S}}_{\delta_1}(K) = 2.$
- $\mathfrak{J}_{\delta_2}(K) = 2,\ \underline{\mathfrak{S}}_{\delta_2}(K) = 0,\ \overline{\mathfrak{S}}_{\delta_2}(K) = 2.$
- $\mathfrak{J}_{\delta_3}(K) = 2,\ \underline{\mathfrak{S}}_{\delta_3}(K) = 2,\ \overline{\mathfrak{S}}_{\delta_3}(K) = 4.$

From the δ_2 invariants and the δ_3 invariants we see that at least three crossing changes are required to unknot K. However, from the δ_2 invariants we see that an unknotting requires at least two negative to positive changes and at least one positive to negative change is required. From δ_3 we see that at least three negative to positive changes are required. Combining these observations, we see that at least three negative to positive changes are required, and at least one positive to negative change is needed. Thus, the unknotting number is at least four.

Note. It is evident and can be proved in a number of ways that the unknotting number of this knot is much greater than four. This example becomes more interesting when considered from the four-dimensional perspective, as discussed in Section 8. It follows from the results there that this knot does not bound an immersed disk in B^4 having fewer than four double points. The best lower bound on this clasp number that can be obtained from previous signature based bounds is 2.

6. Comparison of bounds

Example 13 illustrates a general procedure for finding a lower bound on the unknotting number. For the moment we will call the outcome of that process $u_2(K)$. We will not present a formal definition of this invariant. (The reader is invited to write down the details of the definition; it requires defining the invariant that captures the minimum number of positive and negative crossing changes for each symmetric irreducible δ and then taking the maximums of each of these separately over all symmetric irreducible factors of $\Delta_K(x)$. One must also consider the case $\overline{\mathfrak{S}}_\delta(K) < 0$, which we did not write down.)

In this section we will compare $u_2(K)$ with the classical knot signature bound on $u(K)$; we will temporarily denote the classical bound by $u_1(K)$.

Example 13 presented a knot for which $u_2(K) = 2u_1(K)$. By taking multiples of K we can construct, for each $N > 0$, a knot for which the classical signature bound

	$\mathcal{N} \geq$	$\mathcal{P} \geq$
$\overline{\mathfrak{S}} \geq 0, \underline{\mathfrak{S}} \leq \mathfrak{J}$	$(\mathfrak{J}+\overline{\mathfrak{S}})/2$	$(\mathfrak{J}-\underline{\mathfrak{S}})/2$
$\overline{\mathfrak{S}} \geq 0, \underline{\mathfrak{S}} > \mathfrak{J}$	$(\mathfrak{J}+\overline{\mathfrak{S}})/2$	0
$\overline{\mathfrak{S}} < 0, -\overline{\mathfrak{S}} \leq \mathfrak{J}$	$(\mathfrak{J}+\overline{\mathfrak{S}})/2$	$(\mathfrak{J}-\underline{\mathfrak{S}})/2$
$\overline{\mathfrak{S}} < 0, -\overline{\mathfrak{S}} > \mathfrak{J}$	0	$(\mathfrak{J}-\underline{\mathfrak{S}})/2$

Table 1. Bounds on required signed crossing changes.

on the unknotting number is $u(K) \geq 2N$, but for which our stronger invariants show that $u(K) \geq 4K$. The next result states that this is the best possible.

Theorem 14. *For all knots K, $u_1(K) \leq u_2(K) \leq 2u_1(K)$.*

Proof. Denote the minimum and maximum values of $\sigma_K(\omega)$ with a and A. Since the signature function takes on the value 0 near $\omega = 1$ ($t = 0$), we have $a \leq 0 \leq A$. By definition, $u_1(K) = (A-a)/2$.

For the convenience of the reader, we present the bounds on the signed number of crossing changes in Table 1, covering the four possible cases. For the moment, we let $\mathcal{N}$ denote the minimum number of required changes from negative to positive and let $\mathcal{P}$ denote the minimum number of required crossing changes from positive to negative. The table summarizes the result of Theorem 3, including the cases in which $\overline{\mathfrak{S}} < 0$.

Part 1: $u_2(K) \leq 2u_1(K)$. Our bound on the unknotting number is the sum of an entry from the "$\mathcal{N}$" column in the table, arising from some polynomial δ_1, and an entry from the "$\mathcal{P}$" column arising from a polynomial δ_2, which might equal δ_1. This sum will involve either one or two values of $\mathfrak{J}$, each divided by two. Since each $\mathfrak{J}$ satisfies $0 \leq \mathfrak{J} \leq (A-a)/2$, the sum, after dividing by two, is less than or equal to $(A-a)/2$.

The sum of two entries also involves terms of the form $(\overline{\mathfrak{S}} - \underline{\mathfrak{S}})/2$, where each term might arise from a different δ_i. (There are also cases in which either the $\overline{\mathfrak{S}}$ or $\underline{\mathfrak{S}}$ terms are replaced with 0.) In any case, this sum is also bounded above by $(A-a)/2$.

Adding together these two sums yields a total that is less than or equal to $(A-a)$, which is $2u_1(K)$, as desired.

Part 2: $u_1(K) \leq u_2(K)$. We observe first that $A = |J_K(\omega)| + \sigma_K(\omega)$ for some value of ω. That value of ω is the root of an irreducible polynomial δ. For that δ, we can consider the possibilities that are listed in Table 1 and with care find that the bound on $\mathcal{N}$ must be of the form $(\mathfrak{J} + \overline{\mathfrak{S}})/2$. Since $A \geq 0$, we must have $-\sigma_K(\omega) \leq |J_K(\omega)|$. Thus, the constraint that arises for $\mathcal{N}$ must be at least $A/2$.

In a similar manner, but working with $-K$, we see that the constraint on the $\mathcal{P}$ must be at least $-a/2$. Thus, the total must be at least $(A-a)/2$, as desired. □

7. Gordian distance

The Gordian distance between knots K and J, denoted $d_g(K, J)$, is defined to be the minimum number of crossing changes required to convert K into J. Initial interest in d_g arose in the classical knot theory setting, but as we will describe in Section 8, it is related to four-dimensional properties of knots and in particular is closely tied to a natural metric defined on the knot concordance group. References include [Baader 2006; 2010; Blair et al. 2017; Borodzik et al. 2016; Feller 2014; Hirasawa and Uchida 2002; Miyazawa 2011; Murakami 1985].

Theorem 15. $$d_g(K, J) \geq u_2(K \# -J).$$

Proof. Since u_2 depends only on the signature function, it gives a lower bound on the number of crossing changes required to convert a given knot into a knot with trivial signature function. If K can be converted into J with u crossing changes, then $K \# -J$ can be converted into $J \# -J$ with u crossing changes. The knot $J \# -J$ is a slice knot, and thus has trivial signature. (We will say more about such four-dimensional issues in Section 8.) □

Example 16. As our only application, we consider the connected sum of torus knots

$$K = T(3, 10) \# -T(2, 15) \# -T(5, 6).$$

Its signature function is illustrated in Figure 3. The Alexander polynomial factors as cyclotomic polynomials $\phi_6(x)^2\phi_{10}(x)^2\phi_{15}(x)^2\phi_{30}(x)^3$. In the graphs, the points on the graph above these roots are marked, with the α points corresponding to roots of ϕ_{30}; similarly, β, γ, and η, points correspond to roots of ϕ_{15}, ϕ_{10}, and ϕ_6, respectively.

The classical signature bound on the unknotting number of K is $16/2 = 8$. Considering the polynomial ϕ_{30} we have the set of jumps $\{1, 1, -1, 3\}$ and the

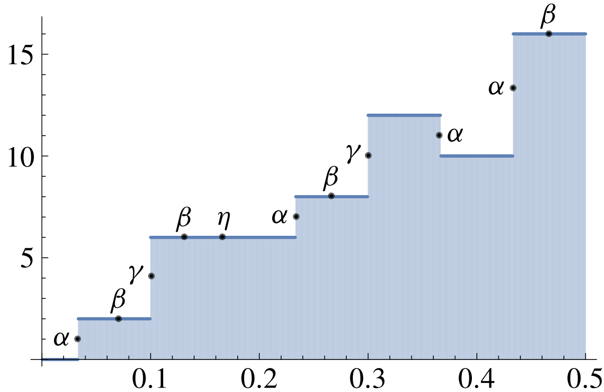

Figure 3. Signature function for $T(3, 10) \# -T(2, 15) \# -T(5, 6)$.

set of signatures is $\{1, 7, 11, 13\}$. Thus $(\mathfrak{J}, \underline{\mathfrak{S}}, \overline{\mathfrak{S}}) = (3, 1, 13)$. Thus, we see that unknotting K would require at least $(3+13)/2 = 8$ crossing changes from negative to positive, and at least $(3-1)/2 = 1$ crossing changes from positive to negative. All together we have a total of at least 9 crossing changes required.

In conclusion, we have the bound on the Gordian distance

$$d_g(T(3, 10), T(2, 15) \# T(5, 6)) \geq 9.$$

We will not compute Khovanov or Heegaard Floer invariants here, but note that the τ, Υ and s-invariant bounds on the unknotting number and Gordian distances are all 8.

8. Four-dimensional perspective

If a knot K can be unknotted with k crossing changes, then it bounds an immersed disk in B^4 with k transverse double points. The minimum number of double points in an immersed disk bounded by K in B^4, taken over all such immersed disks, has been called the *clasp number*, *the four-dimensional clasp number* or the *four-ball crossing number*. For the moment, we denote this invariant $c(K)$. References include [Kawamura 2002; Murakami and Yasuhara 2000; Owens and Strle 2016]. This invariant can be refined by considering the number of positive and negative double points in the immersed disk.

In a similar way, if a sequence of crossing changes converts a knot K into a knot J, then there is an immersion of $S^1 \times I$ into $S^3 \times I$ (a *singular concordance*) with boundary $K \times \{0\} \cup J \times \{1\}$ such that the number of double points equals the number of crossing changes. There is, in a way, a converse. From [Owens and Strle 2016, Proposition 2.1] we have the following.

Theorem 17. *If K and J bound a singular concordance with p positive and n negative double points, then there are knots K' and J', concordant to K and J, respectively, such that K' can be converted into J' with p positive and n negative crossing changes.*

For knots K and J, we can define a distance $d_c(K, J)$ as the minimum number of double points in a singular concordance between the knots. This induces a metric on the concordance group. For U the unknot, we have that the four-dimensional clasp number is equal to $d_c(K, U)$.

Since concordant knots have the same signature functions, we have the following.

Theorem 18. *For knots K and J, $d_c(K, J) \geq u_2(K \# -J)$.*

Example 19. The same computation as in Example 4 shows any singular concordance from 5_1 to -10_{132} must have at least two positive double points and one negative double point. In particular, the four-dimensional clasp number of $5_1 \# 10_{132}$ is 3.

Example 20. From Example 16 we have that $d_c(T(3, 10), T(2, 15) \# T(5, 6)) \geq 9$.

8.1. ***The four-genus.*** The clasp number and four-genus of a knot are related by the bound $g_4(K) \leq c(K)$. This can be enhanced with the following observation: If K bounds an immersed disk with p positive and n negative double points, then $g_4(K) \leq \max(p, n)$. Thus, lower bounds on g_4 provide lower bounds on the clasp number (and unknotting number).

In the case that a knot is slice, $g_4(K) = 0$, the clasp number is also 0. Multiples of the square knot, $T(2, 3) \# -T(2, 3)$, provide examples of slice knots for which the unknotting number can be arbitrarily large. In fact, using the homology of the 2-fold branched cover [Kinoshita 1957; Wendt 1937], one shows that

$$u(N(T(2, 3) \# -T(2, 3))) = 2N.$$

It is unknown whether there exists a knot K with $g_4(K) = 1$, but $c(K) > 2$; in [Livingston 2002; Owens and Strle 2016] it is shown that there are knots with four-genus 1 that cannot be converted into slice knots using one crossing change. Owens [2010] has identified two-bridge knots K_n with $g_4(K_n) = n,\ \sigma(K) = 2n$, and which cannot be converted into a slice knot with n crossing changes from negative to positive and any number of positive to negative crossing changes. In particular, the knot $K_1 = -7_4$ has $g_4(K_1) = 1,\ \sigma(K_1) = 2$, but any sequence of crossing changes that converts K_1 into a slice knot must include at least two crossing changes from negative to positive.

Our main results, since they are providing *lower* bounds on $u(K)$ and $c(K)$, might not offer bounds on $g_4(K)$. However, it is worth noting that the observations made in this paper offer a much simpler proof of this result from [Livingston 2011].

Corollary 21. *Let $K \subset S^3$ be a knot and let $\{\alpha_1, \dots, \alpha_k\} \subset \mathbb{S}^1$ be a nonempty subset of the complex roots of an irreducible rational polynomial δ. Then*

$$g_4(K) \geq \tfrac{1}{2}(\max\{|J_K(\alpha_i)|\} + \max\{|\sigma_K(\alpha_i)|\}).$$

Proof. As stated in Theorem 6, it follows from [Taylor 1979] that if $g_4(K) = g$, then $W_K \in W(\mathbb{Q}(x))$ has a $2g \times 2g$ representative. Consider the diagonal form of such a representative, $W_K = [f_1\delta, \dots, f_m\delta, g_1, \dots, g_n]$. It is clear that for all α_i, $|J_K(\alpha_i)| \leq m$ and $\max|\sigma_K(\alpha_i)| \leq n$. It follows that

$$\max\{|J_K(\alpha_i)|\} + \max\{|\sigma_K(\alpha_i)|\} \leq m + n = 2g. \qquad \square$$

9. The nonbalanced signature function

For a knot K with Seifert matrix V_K, the signature of the matrix $(1-\omega)V_K + (1-\overline{\omega})V_K^{\mathsf{T}}$ yields a well-defined function $s_K(\omega)$ on the unit circle. The proof that $s_K(\omega)$ is a knot invariant is a consequence of the fact that any two Seifert surfaces for a knot are stably equivalent [Murasugi 1965; Trotter 1962]. The signature function we have been considering, σ_K, is defined by taking the two-sided average of s_K. We have focused our attention on the balanced function because it defines a homomorphism on the knot concordance group. There are examples of slice knots for which s_K is nontrivial [Cha and Livingston 2004; Levine 1989], and thus it is not a knot concordance invariant. Here we briefly explore how s_K can be used to extract unknotting information that is not accessible via σ_K. Of course, these results do not generalize to give concordance invariants.

Let $\delta(x)$ be an irreducible Alexander polynomial having roots $\{\alpha_1, \dots, \alpha_n\}$ on the unit circle. Let $\Lambda_\delta = \mathbb{Q}[x, x^{-1}]_{(\delta)}$ denote the ring formed by inverting all irreducible elements in $\mathbb{Q}[x, x^{-1}]$ other than $\delta(x)$. The proof that hermitian forms over $\mathbb{Q}(x)$ can be diagonalized is easily modified to the case of hermitian forms over Λ_δ. One needs to check that the step-by-step diagonalization process can be adjusted so that division by $\delta(x)$ is not required.

Given this, the proof of Lemma 10 generalizes to the setting of Λ_δ, and so the effect of crossing changes is determined by the signature functions for a pair of matrices of the form

$$W_{\pm 1}(x) = \begin{pmatrix} a(x) & \overline{b(x)} \\ b(x) & d(x) + \epsilon_\pm (1-x)(1-x^{-1}) \end{pmatrix}.$$

Here, all polynomials are in Λ_δ, $\epsilon_- = 1$, and $\epsilon_+ = 0$.

We now factor out powers of δ to rewrite this as

$$\begin{pmatrix} a'(x)\delta(x)^i & \overline{b'(x)}\delta(x)^j \\ b'(x)\delta(x)^j & d(x) + \epsilon_\pm (1-x)(1-x^{-1}) \end{pmatrix}.$$

Considering the difference of signatures, $\operatorname{sign}(W_+(\alpha_i)) - \operatorname{sign}(W_-(\alpha_i))$, yields the following cases, all of which are easily analyzed by considering diagonalizations.

- If $i = 0$, the difference of signatures is determined by the difference of values of
$$d'(\alpha_i) + \epsilon_\pm (1-\alpha_i)(1-\alpha_i^{-1})$$
for some $d' \in \Lambda_\delta$.
- If $i \neq 0$ and $j = 0$, then both signatures are 0.
- If $i \neq 0$ and $j \neq 0$, then the difference of signatures is determined by the difference of values of
$$d'(\alpha_i) + \epsilon_\pm (1-\alpha_i)(1-\alpha_i^{-1})$$
for some $d' \in \Lambda_\delta$.

The approach of our previous work now applies, and a simple consequence is the following.

Lemma 22. *If $\delta(x) \in \mathbb{Z}[t, t^{-1}]$ has roots $\{\alpha_1, \dots, \alpha_n\}$ on the unit circle and a crossing change from positive to negative is made to a knot K, then either all the values of $s_K(\alpha_i)$ increase by* 1, *or some increase by* 2 *and others are unchanged.*

Rather than exploit the dependance of this result on the choice of δ, here we will present an easily stated result, expressed in terms of the floor and ceiling function. Recall that $\max\{s_K(\omega)\} \geq 0$ and $\min\{s_K(\omega)\} \leq 0$.

Theorem 23. *For a knot K, let $M = \max\{s_K(\omega)\}$ and $m = \min\{s_K(\omega)\}$. The unknotting number satisfies*

$$u(K) \geq \lceil M/2 \rceil - \lfloor m/2 \rfloor.$$

Example 24. The knot 8_{20} is slice, and hence its signature function σ_K is identically 0. However, its Alexander polynomial is $(t^2 - t + 1)^2$, having a root at the sixth root of unity, ξ_6. A direct computation shows that $s_K(\xi_6) = 1$, and this is the only nonzero value of the upper unit circle. It is shown in [Cha and Livingston 2004] that similar, but less explicit examples abound.

Using such knots as 8_{20}, one can construct a knot K for which there are two numbers on the upper half circle, ω_1 and ω_2, with the property that $s_K(\omega_1) = 3$, $s_K(\omega_2) = -3$, and $s_k(\omega) = 0$ for all other ω on the upper half circle. According to Theorem 23, this knot has unknotting number at least 4. Notice that this is one more than $(\max(s_K(\omega)) - \min(s_K(\omega)))/2$, as might be expected from the classic bound based on $\sigma_K(\omega)$.

9.1. ***Doubly slice knots.*** A knot K is called doubly slice if it is the cross-section of an unknotted two-sphere embedded in S^4. The first proof of the existence of such knots appeared in [Terasaka and Hosokawa 1961]. Invariants of such knots have since been studied in much finer detail; see for instance, [Kearton 1975; Kim 2006; Livingston and Meier 2015; Meier 2015; Stoltzfus 1978; Sumners 1971]. The nonbalanced signature function of a doubly slice knot is identically 0, and this provides a means of proving that some slice knots are not doubly slice.

The slicing number of a knot and the algebraic slicing number of a knot are the number of crossing changes required to convert a knot into a slice, respectively algebraically slice, knot. One could similarly define a *double slicing number* of a knot. Hence, we have:

Theorem 25. *For a knot K, let $M = \max\{s_K(\omega)\}$ and $m = \min\{s_K(\omega)\}$. The number of crossing changes required to convert K into a doubly slice knot is greater than or equal to $\lceil M/2 \rceil - \lfloor m/2 \rfloor$.*

Acknowledgments

Many thanks to Pat Gilmer, whose careful reading and suggestions helped eliminate several gaps and greatly clarified the exposition. Thanks are also due to Jeff Meier and Matthias Nagel for helpful comments, and to the referee for providing a valuable list of suggestions that enhanced the quality of this paper.

References

[Baader 2006] S. Baader, "Note on crossing changes", *Q. J. Math.* **57**:2 (2006), 139–142. MR Zbl

[Baader 2010] S. Baader, "Unknotting sequences for torus knots", *Math. Proc. Cambridge Philos. Soc.* **148**:1 (2010), 111–116. MR Zbl

[Baader 2012] S. Baader, "Scissor equivalence for torus links", *Bull. Lond. Math. Soc.* **44**:5 (2012), 1068–1078. MR Zbl

[Blair et al. 2017] R. Blair, M. Campisi, J. Johnson, S. A. Taylor, and M. Tomova, "Neighbors of knots in the Gordian graph", *Amer. Math. Monthly* **124**:1 (2017), 4–23. MR Zbl

[Borodzik et al. 2016] M. Borodzik, S. Friedl, and M. Powell, "Blanchfield forms and Gordian distance", *J. Math. Soc. Japan* **68**:3 (2016), 1047–1080. MR Zbl

[Cha and Livingston 2004] J. C. Cha and C. Livingston, "Knot signature functions are independent", *Proc. Amer. Math. Soc.* **132**:9 (2004), 2809–2816. MR Zbl

[Cha and Livingston 2019] J. C. Cha and C. Livingston, "KnotInfo: table of knot invariants", electronic reference, 2019, Available at http://www.indiana.edu/∼knotinfo.

[Cochran and Lickorish 1986] T. D. Cochran and W. B. R. Lickorish, "Unknotting information from 4-manifolds", *Trans. Amer. Math. Soc.* **297**:1 (1986), 125–142. MR Zbl

[Feller 2014] P. Feller, "Gordian adjacency for torus knots", *Algebr. Geom. Topol.* **14**:2 (2014), 769–793. MR Zbl

[Feller 2016] P. Feller, "Optimal cobordisms between torus knots", *Comm. Anal. Geom.* **24**:5 (2016), 993–1025. MR Zbl

[Feller and Krcatovich 2017] P. Feller and D. Krcatovich, "On cobordisms between knots, braid index, and the upsilon-invariant", *Math. Ann.* **369**:1-2 (2017), 301–329. MR Zbl

[Gompf 1984] R. E. Gompf, "Infinite families of Casson handles and topological disks", *Topology* **23**:4 (1984), 395–400. MR Zbl

[Hirasawa and Uchida 2002] M. Hirasawa and Y. Uchida, "The Gordian complex of knots", *J. Knot Theory Ramifications* **11**:3 (2002), 363–368. MR Zbl

[Jabuka 2009] S. Jabuka, "The rational Witt class and the unknotting number of a knot", preprint, 2009. arXiv

[Kawamura 2002] T. Kawamura, "On unknotting numbers and four-dimensional clasp numbers of links", *Proc. Amer. Math. Soc.* **130**:1 (2002), 243–252. MR Zbl

[Kearton 1975] C. Kearton, "Simple knots which are doubly-null-cobordant", *Proc. Amer. Math. Soc.* **52** (1975), 471–472. MR Zbl

[Kim 2006] T. Kim, "New obstructions to doubly slicing knots", *Topology* **45**:3 (2006), 543–566. MR Zbl

[Kim and Livingston 2005] S.-G. Kim and C. Livingston, "Knot mutation: 4-genus of knots and algebraic concordance", *Pacific J. Math.* **220**:1 (2005), 87–105. MR Zbl

[Kinoshita 1957] S. i. Kinoshita, "On Wendt's theorem of knots", *Osaka Math. J.* **9** (1957), 61–66. MR Zbl

[Levine 1969] J. Levine, "Invariants of knot cobordism", *Invent. Math.* **8** (1969), 98–110. MR Zbl

[Levine 1989] J. P. Levine, "Metabolic and hyperbolic forms from knot theory", *J. Pure Appl. Algebra* **58**:3 (1989), 251–260. MR Zbl

[Litherland 1984] R. A. Litherland, "Cobordism of satellite knots", pp. 327–362 in *Four-manifold theory* (Durham, NH, 1982), edited by C. Gordon and R. Kirby, Contemp. Math. **35**, Amer. Math. Soc., Providence, RI, 1984. MR Zbl

[Livingston 2002] C. Livingston, "The slicing number of a knot", *Algebr. Geom. Topol.* **2** (2002), 1051–1060. MR Zbl

[Livingston 2011] C. Livingston, "Knot 4-genus and the rank of classes in $W(\mathbb{Q}(t))$", *Pacific J. Math.* **252**:1 (2011), 113–126. MR Zbl

[Livingston and Meier 2015] C. Livingston and J. Meier, "Doubly slice knots with low crossing number", *New York J. Math.* **21** (2015), 1007–1026. MR Zbl

[Meier 2015] J. Meier, "Distinguishing topologically and smoothly doubly slice knots", *J. Topol.* **8**:2 (2015), 315–351. MR Zbl

[Miyazawa 2011] Y. Miyazawa, "Gordian distance and polynomial invariants", *J. Knot Theory Ramifications* **20**:6 (2011), 895–907. MR Zbl

[Murakami 1985] H. Murakami, "Some metrics on classical knots", *Math. Ann.* **270**:1 (1985), 35–45. MR Zbl

[Murakami and Yasuhara 2000] H. Murakami and A. Yasuhara, "Four-genus and four-dimensional clasp number of a knot", *Proc. Amer. Math. Soc.* **128**:12 (2000), 3693–3699. MR Zbl

[Murasugi 1965] K. Murasugi, "On a certain numerical invariant of link types", *Trans. Amer. Math. Soc.* **117** (1965), 387–422. MR Zbl

[Owens 2008] B. Owens, "Unknotting information from Heegaard Floer homology", *Adv. Math.* **217**:5 (2008), 2353–2376. MR Zbl

[Owens 2010] B. Owens, "On slicing invariants of knots", *Trans. Amer. Math. Soc.* **362**:6 (2010), 3095–3106. MR Zbl

[Owens and Strle 2016] B. Owens and S. Strle, "Immersed disks, slicing numbers and concordance unknotting numbers", *Comm. Anal. Geom.* **24**:5 (2016), 1107–1138. MR Zbl

[Ozsváth and Szabó 2003] P. Ozsváth and Z. Szabó, "Knot Floer homology and the four-ball genus", *Geom. Topol.* **7** (2003), 615–639. MR Zbl

[Ozsváth et al. 2017] P. S. Ozsváth, A. I. Stipsicz, and Z. Szabó, "Concordance homomorphisms from knot Floer homology", *Adv. Math.* **315** (2017), 366–426. MR Zbl

[Rasmussen 2010] J. Rasmussen, "Khovanov homology and the slice genus", *Invent. Math.* **182**:2 (2010), 419–447. MR Zbl

[Rolfsen 1976] D. Rolfsen, *Knots and links*, Mathematics Lecture Series **7**, Publish or Perish, Berkeley, CA, 1976. MR Zbl

[Scharlemann 1985] M. G. Scharlemann, "Unknotting number one knots are prime", *Invent. Math.* **82**:1 (1985), 37–55. MR Zbl

[Stoltzfus 1978] N. W. Stoltzfus, "Algebraic computations of the integral concordance and double null concordance group of knots", pp. 274–290 in *Knot theory* (Plans-sur-Bex, Switzerland, 1977), edited by J.-C. Hausmann, Lecture Notes in Math. **685**, Springer, 1978. MR Zbl

[Sumners 1971] D. W. Sumners, "Invertible knot cobordisms", *Comment. Math. Helv.* **46** (1971), 240–256. MR Zbl

[Taylor 1979] L. R. Taylor, "On the genera of knots", pp. 144–154 in *Topology of low-dimensional manifolds* (Chelwood Gate, England, 1977), Lecture Notes in Math. **722**, Springer, 1979. MR Zbl

[Terasaka and Hosokawa 1961] H. Terasaka and F. Hosokawa, "On the unknotted sphere S^2 in E^4", *Osaka Math. J.* **13** (1961), 265–270. MR Zbl

[Tristram 1969] A. G. Tristram, "Some cobordism invariants for links", *Proc. Cambridge Philos. Soc.* **66** (1969), 251–264. MR Zbl

[Trotter 1962] H. F. Trotter, "Homology of group systems with applications to knot theory", *Ann. of Math.* (2) **76** (1962), 464–498. MR Zbl

[Wendt 1937] H. Wendt, "Die gordische Auflösung von Knoten", *Math. Z.* **42**:1 (1937), 680–696. MR Zbl

Received July 10, 2018. Revised July 12, 2019.

CHARLES LIVINGSTON
DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, IN
UNITED STATES
livingst@indiana.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.251

THE GLOBAL WELL-POSEDNESS AND SCATTERING FOR THE 5-DIMENSIONAL DEFOCUSING CONFORMAL INVARIANT NLW WITH RADIAL INITIAL DATA IN A CRITICAL BESOV SPACE

CHANGXING MIAO, JIANWEI YANG AND TENGFEI ZHAO

We obtain the global well-posedness and scattering for the radial solution to the defocusing conformal invariant nonlinear wave equation with initial data in the critical Besov space $\dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$. This is the 5-dimensional analogue of Dodson's result (2019), which was the first on the global well-posedness and scattering of the energy subcritical nonlinear wave equation without the uniform boundedness assumption on the critical Sobolev norms employed as a substitute of the missing conservation law with respect to the scaling invariance of the equation. The proof is based on exploiting the structure of the radial solution, developing the Strichartz-type estimates and incorporation of Dodson's strategy (2019), where we also avoid a logarithm-type loss by employing the inhomogeneous Strichartz estimates.

1. Introduction

We consider the solutions u to

$$\text{(1-1)} \qquad \begin{cases} \partial_{tt}u - \Delta u + \mu |u|^p u = 0, & (t,x) \in \mathbb{R} \times \mathbb{R}^d, \\ (u(0), \partial_t u(0)) = (u_0, u_1), & x \in \mathbb{R}^d, \end{cases}$$

where $\mu = \pm 1$, $d \geq 1$, and $p > 0$. If $\mu = 1$, (1-1) is described as defocusing, otherwise focusing. There is a natural scaling symmetry for (1-1), i.e., if we let $u_\lambda(t,x) = \lambda^{2/p} u(\lambda t, \lambda x)$ for $\lambda > 0$, then u_λ is also a solution to (1-1) with initial data $(\lambda^{2/p} u_0(\lambda x), \lambda^{(2/p)+1} u_1(\lambda x))$ preserving the $\dot{H}^{s_p} \times \dot{H}^{s_p - 1}(\mathbb{R}^d)$ norm of the initial data, where we define the critical regularity as $s_p = \frac{d}{2} - \frac{2}{p}$. At least, the solutions to (1-1) formally conserve the energy

$$\text{(1-2)} \quad E(u(t), \partial_t u(t)) := \frac{1}{2} \int_{\mathbb{R}^5} |\nabla_x u(t)|^2 \, dx + \frac{1}{2} \int_{\mathbb{R}^5} |\partial_t u(t)|^2 \, dx + \frac{\mu}{p+2} \int_{\mathbb{R}^5} |u(t)|^{p+2} \, dx,$$

MSC2010: primary 35B40, 35L05; secondary 35Q40.
Keywords: nonlinear wave equation, Strichartz estimates, scattering, hyperbolic coordinates, Morawetz estimates.

which is also invariant under the scaling if $s_p = 1$. In view of this, we say the Cauchy problem (1-1) is energy critical when $s_p = 1$, subcritical for $s_p < 1$ and supercritical when $s_p > 1$.

Lindblad and Sogge [1995] proved the local theory of the Cauchy problem (1-1) in the minimal regularity spaces. In fact, if $d \geq 2$ and $p \geq (d+3)/(d-1)$, the Cauchy problem (1-1) with initial data in the critical spaces $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^d)$ is locally well-posed. The global theory for the Cauchy problem (1-1) with $\mu = 1$ and $s_p \leq 1$ has been studied extensively. While for the focusing case, even the solution with smooth initial data may blow up at finite time. For more related results see [Sogge 1995].

We will consider global existence and scattering of the solutions to (1-1). In general, a solution u is said to be scattering if it is a global solution and approaches a linear solution as $t \to \pm\infty$. In the cases of $d \geq 2$ and $p \geq (d+3)/(d-1)$, the solution to (1-1) with small initial datum in the critical Sobolev spaces is globally well-posed and scattering; see [Lindblad and Sogge 1995].

For the defocusing energy critical wave equation (1-1), Grillakis [1990] first established the global existence theory for classical solution when $d = 3$. The results for other dimensions are proved in [Grillakis 1992; Shatah and Struwe 1993]. Scattering results for large energy data are proved in [Bahouri and Gérard 1999; Bahouri and Shatah 1998; Nakanishi 1999] by establishing variants of the Morawetz estimates [1968]

$$\iint_{\mathbb{R}^{1+d}} \frac{|u|^{\frac{2d}{d-2}}}{|x|}\,dx\,dt \leq C_d E(u_0, u_1), \tag{1-3}$$

where C_d is a constant depending on d. While in focusing energy critical cases, the Morawetz estimates (1-3) fails. The scattering results do not hold in general, since (1-1) has a ground state

$$W(x) = \left(1 + \frac{|x|^2}{d(d-2)}\right)^{-\frac{d-2}{2}}.$$

In the cases of $3 \leq d \leq 5$, Kenig and Merle [2008] proved the scattering result for solution with initial data such that $E(u_0, u_1) < E(W, 0)$ and $\|u_0\|_{\dot{H}^1(\mathbb{R}^d)} < \|W\|_{\dot{H}^1(\mathbb{R}^d)}$. In their proofs, the main ingredient is the concentration compactness/rigidity theorem method introduced by [Kenig and Merle 2006]. This method is powerful and plays an important role in study of many other nonlinear dispersive equations. We refer to [Killip and Vişan 2013; Koch et al. 2014; Kenig 2015].

For the defocusing subcritical equation (1-1), the global existence has been proved for solution with initial data in the energy space $\dot{H}^1 \times L^2(\mathbb{R}^d)$ by Ginibre and Velo [1985; 1989]. However, there are no scattering results even for solutions with initial datum in $(\dot{H}^1 \cap \dot{H}^{s_p}) \times (L^2 \cap \dot{H}^{s_p-1})(\mathbb{R}^d)$.

Recently, Dodson [2019] proved scattering results for the defocusing cubic wave equation with the initial datum belonging to the space $\dot{B}^2_{1,1} \times \dot{B}^1_{1,1}(\mathbb{R}^3)$, which is a subspace of $\dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^3)$. We remark that this is the first work that gives scattering results for large data in the critical Sobolev space without any a priori bound on the critical norm of the solution. Dodson's strategy consists of three steps:

(1) By establishing some new Strichartz-type estimates, one can show that the solution is in the energy space $\dot{H}^1 \times L^2(\mathbb{R}^3)$ up to some free evolutions. Then this decomposition enables one to prove the global well-posedness of the solution.

(2) To obtain the scattering result, a conformal transformation is applied to show that the energy part of the solution has finite energy in hyperbolic coordinates. Then from the conformal invariance of the equation and a Morawetz-type inequality, one can deduce that $\|u\|_{L^4_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \le C(\|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}, \delta_1)$, where the parameter δ_1 relies on the scaling and spatial profiles of the initial data.

(3) Finally, one can remove the dependence of δ_1 by employing the profile decomposition, which completes the proof.

Let $S(t)(f, g)$ be the solution of Cauchy problem to the free wave equation

$$\begin{cases} \partial_{tt} v - \Delta v = 0, & (t, x) \in \mathbb{R} \times \mathbb{R}^5, \\ (v, \partial_t v)|_{t=0} = (f(x), g(x)), & x \in \mathbb{R}^5. \end{cases} \tag{1-4}$$

For the sake of statement, we introduce the following notation as

$$\dot{S}(t)(f, g) \triangleq \partial_t S(t)(f, g), \quad \text{and} \quad \vec{S}(t)(f, g) \triangleq \big(S(t)(f, g), \dot{S}(t)(f, g)\big).$$

We consider the Cauchy problem of nonlinear wave equation

$$\begin{cases} \partial_{tt} u - \Delta u + |u|u = 0, & (t, x) \in \mathbb{R} \times \mathbb{R}^d, \\ (u(0), \partial_t u(0)) = (u_0, u_1), & x \in \mathbb{R}^d, \end{cases} \tag{1-5}$$

Our main result can be stated as:

Theorem 1.1. *For any radial initial data* $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$, *the solution u to (1-5) is globally well-posed and scattering, i.e., there exists* $(u_0^{\pm}, u_1^{\pm}) \in \dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$ *such that*

$$\lim_{t\to\pm\infty} \|(u(t), \partial_t u(t)) - \vec{S}(t)(u_0^{\pm}, u_1^{\pm})\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \to 0. \tag{1-6}$$

Furthermore, there is a function $A : [0, \infty) \to [0, \infty)$, *such that*

$$\|u\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \le A(\|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}). \tag{1-7}$$

Remark 1.2. (1) This theorem extends the results of [Dodson 2019] to the 5-dimension case. The proof will utilize the strategy given in [Dodson 2019], but it is highly nontrivial.

(2) Unlike the 3-dimensional case, the dispersive estimate (see (2-20)) gives a decay in time of order -2, which may cause a logarithmic failure when one estimates

$$\|u\|_{L_t^{4/3}L_x^4(J\times\mathbb{R}^5)} + \sup_{t\in J}\big(t^{3/4}\|u\|_{L_x^4(\mathbb{R}^5)}\big),$$

where $0 \in J$ is a local time interval. We circumvent this difficulty by using the inhomogeneous Strichartz estimates in [Taggart 2010] and prove the global well-posedness of u.

(3) For the scattering result, by reductions, we need to bound the $L^3_{t,x}$ of w on the light cone $\{|x| \le t + \frac{1}{2}\}$. We will define the hyperbolic energy by rewriting (1-5) as the form

$$\partial_{tt}(r^2 w) - \partial_{rr}(r^2 w) = -2w - r^2|w|w.$$

Observing that the additional term $2w$ and the nonlinear term $r^2|w|w$ enjoy the same sign, we can bound the $L^3_{t,x}$ norm of w by applying a Morawetz-type inequality, if we assume the hyperbolic energy of w is bounded.

(4) To certify the above assumption, we will make full use of (2-19) for radial solution and the sharp Hardy inequality. In contrast to the 3-dimensional case, some terms in (2-19) seem more difficult to dealt with. However, the integration domains of these terms are symmetric about the radius r, which is also consistent with the Huygens principle. This fact allows us apply the Hardy–Littlewood maximal functions to verify the assumption.

Now, we give the outline of the proof. By the Strichartz estimates and a standard fix point argument, for initial data (u_0, u_1), there exists a maximal time interval $I \subset \mathbb{R}$ such that there exists a unique solution u (see Definition 2.9 in Section 2) to (1-5) on $I \times \mathbb{R}^5$. We consider the global well-posedness by developing some Strichartz-type estimates (3-30). Utilizing the standard blowup criterion, we can show the global well-posedness of u.

Next, we claim the following proposition:

Proposition 1.3. *For every radial initial data* $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$, *let* u *be the corresponding solution to* (1-5). *Then there exists a parameter* δ_1 *depending the initial data* (u_0, u_1) *and a function* $A : [0,\infty)^2 \to [0,\infty)$ *such that*

$$\|u\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \le A(\|(u_0,u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}, \delta_1). \tag{1-8}$$

We prove Theorem 1.1 by employing this and establishing Proposition 4.2, where the proof provides an alternate proof of Lemma 6.2 in [Dodson 2019].

Finally, we need to prove Proposition 1.3. From the partition $u = v + w$, it suffices to show the boundedness of $L^3_{t,x}$ norm of w. We prove the hyperbolic energy of w is uniformly bounded. Then, a Morawetz-type inequality yields that the $L^3_{t,x}$ norm of w is bounded in the cone, which finishes the proof.

This paper is organized as follows: Section 2 gives some tools from harmonic analysis and basic properties for the wave equation. In Section 3 we give the decomposition of u and prove its global well-posedness. The existence of the function A in (1-7) is shown in Section 4 based on the Proposition 1.3. Finally, in Section 5, we complete the proof by showing Proposition 1.3.

We end the introduction with some notations used throughout this paper. We use $\mathcal{S}(\mathbb{R}^d)$ to denote the space of Schwartz functions on $\mathbb{R}^d$. For $1 \le p \le \infty$, we define $L^p(\mathbb{R}^d)$ by the spaces of Lebesgue measurable functions with finite $L^p(\mathbb{R}^d)$-norm, which is defined by

$$\|f\|_{L^p(\mathbb{R}^d)} = \Big(\int_{\mathbb{R}^d} |f(x)|^p \,dx\Big)^{1/p}, \quad \text{for } 1 \le p < \infty,$$

and $\|f\|_{L^\infty(\mathbb{R}^d)} = \operatorname{ess\,sup}_{x\in\mathbb{R}^d} |f(x)|$. We let ℓ^p be the spaces of complex number sequences $\{a_n\}_{n\in\mathbb{Z}}$ such that $\{a_n\}_{n\in\mathbb{Z}} \in l^p$ if and only if

$$\|\{a_n\}\|_{\ell^p_n(\mathbb{Z})} \triangleq \Big(\sum_n |a_n|^p\Big)^{1/p} < \infty, \quad \text{for } 1 \le p < \infty$$

and $\|\{a_n\}\|_{\ell^\infty_n(\mathbb{Z})} := \sup_n |a_n| < \infty$. We use $X \lesssim Y$ to mean that there exists a constant $C > 1$ such that $X \le CY$, where the dependence of C on the parameters will be clear from the context. We use $X \sim Y$ to denote $X \lesssim Y$ and $Y \lesssim X$. $A \ll B$ denotes there is a sufficiently large number C such that $A \le C^{-1}B$.

2. Basic tools and some elementary properties for the wave equation

In this section, we recall some tools from harmonic analysis and useful results for the wave equation.

2A. ***Some tools from harmonic analysis.*** Recall the Fourier transform of $f \in \mathcal{S}(\mathbb{R}^d)$ is defined by

$$\widehat{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x)e^{-ix\xi} \,dx,$$

which can be extended to Schwartz distributions naturally. We will make frequent use of the Littlewood–Paley projection operators. Specifically, we let φ be a radial smooth function supported on the ball $|\xi| \le 2$ and equal to 1 on the ball $|\xi| \le 1$. For $j \in \mathbb{Z}$, we define the Littlewood–Paley projection operators by

$$\begin{aligned} \widehat{P_{\le j} f}(\xi) &:= \varphi(\xi/2^j)\widehat{f}(\xi), \\ \widehat{P_{> j} f}(\xi) &:= (1-\varphi(\xi/2^j))\widehat{f}(\xi), \\ \widehat{P_j f}(\xi) &:= (\varphi(\xi/2^j) - \varphi(\xi/2^{j-1}))\widehat{f}(\xi). \end{aligned}$$

The Littlewood–Paley operators commute with derivative operators and are bounded on the general Sobolev spaces. These operators also obey the following standard Bernstein estimates:

Lemma 2.1 (Bernstein estimates). *For* $1 \le r \le q \le \infty$ *and* $s \ge 0$,

$$\begin{aligned} \big\| |\nabla|^{\pm s} P_j f \big\|_{L^r_x(\mathbb{R}^d)} &\sim 2^{\pm js} \| P_j f \|_{L^r_x(\mathbb{R}^d)}, \\ \big\| |\nabla|^{s} P_{\le j} f \big\|_{L^r_x(\mathbb{R}^d)} &\lesssim 2^{js} \| P_{\le j} f \|_{L^r_x(\mathbb{R}^d)}, \\ \| P_{>j} f \|_{L^r_x(\mathbb{R}^d)} &\lesssim 2^{-js} \big\| |\nabla|^{s} P_{>j} f \big\|_{L^r_x(\mathbb{R}^d)}, \\ \| P_{\le j} f \|_{L^q(\mathbb{R}^d)} &\lesssim 2^{(d/r-d/q)j} \| P_{\le j} f \|_{L^r_x(\mathbb{R}^d)}, \end{aligned}$$

where the fractional derivative operator $|\nabla|^\sigma$ *is defined by* $\widehat{|\nabla|^\sigma f}(\xi) = |\xi|^\sigma \widehat{f}(\xi)$, *for* $\sigma \in \mathbb{R}$.

Definition 2.2 (homogeneous Besov spaces). Let s be a real number and let $1 \le p, r \le \infty$. We denote the homogeneous Besov norm by

$$\| f \|_{\dot{B}^s_{p,r}(\mathbb{R}^d)} := \big\| \{ 2^{js} \| P_j f \|_{L^p(\mathbb{R}^d)} \} \big\|_{\ell^r_j(\mathbb{Z})}, \tag{2-1}$$

for $f \in \mathcal{S}(\mathbb{R}^d)$. Then the Besov space $\dot{B}^s_{p,r}(\mathbb{R}^d)$ is the completion of the Schwartz function under this norm.

We shall give the following radial Sobolev-type inequalities, which are analogous to the 3-dimensional cases established in [Dodson 2019]. We denote radial derivative by $\partial_r f(x) = \big(\frac{x}{|x|} \cdot \nabla\big) f(x)$ for any function f defined on $\mathbb{R}^5$.

Lemma 2.3 (radial Sobolev-type inequalities in Besov spaces). *For any radial function* $f \in \mathcal{S}(\mathbb{R}^5)$, *we have*

$$\| |x|^2 f \|_{L^\infty(\mathbb{R}^5)} \lesssim \| f \|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)}. \tag{2-2}$$

Let $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$ *be a radial function*; *then we have*

$$\begin{aligned} \Big\| \frac{1}{|x|^2} \partial_r u_0(x) \Big\|_{L^1_x(\mathbb{R}^5)} + \Big\| \frac{1}{|x|} \partial_{rr} u_0(x) \Big\|_{L^1_x(\mathbb{R}^5)} \\ + \Big\| \frac{1}{|x|^3} u_0(x) \Big\|_{L^1_x(\mathbb{R}^5)} + \big\| |x|^3 \partial_r u_0 \big\|_{L^\infty_x(\mathbb{R}^5)} \lesssim \| u_0 \|_{\dot{B}^3_{1,1}(\mathbb{R}^5)}, \end{aligned} \tag{2-3}$$

$$\Big\| \frac{1}{|x|} \partial_r u_1(x) \Big\|_{L^1_x(\mathbb{R}^5)} + \Big\| \frac{1}{|x|^2} u_1(x) \Big\|_{L^1_x(\mathbb{R}^5)} + \big\| |x|^3 u_1(x) \big\|_{L^\infty_x(\mathbb{R}^5)} \lesssim \| u_1 \|_{\dot{B}^2_{1,1}(\mathbb{R}^5)}. \tag{2-4}$$

Proof. We first consider (2-2). Since f is radial, using polar coordinates, we have

$$P_j f(|x|) = P_j f(x) = \int_{\mathbb{R}^5} \widehat{P_j f}(\xi) e^{ix\xi}\, d\xi = \int_0^\infty \int_{\mathbb{S}^4} \widehat{P_j f}(r) r^4 e^{irx\omega}\, d\sigma(\omega)\, dr. \tag{2-5}$$

Recall the decay estimates of Fourier transform of the surface measure on the sphere

$$\widehat{d\sigma_{\mathbb{S}^4}}(\xi) \le C(1+|\xi|)^{-2},$$

which, with Hölder's inequality, yields

$$|P_j f(|x|)| \lesssim \int_0^\infty |\widehat{P_j f}(r)| r^2 |x|^{-2}\, dr \lesssim |x|^{-2} 2^{\frac{1}{2}j} \|P_j f\|_{L^2}. \tag{2-6}$$

Then the inequality (2-2) follows from (2-6) and the definition of the Besov space.

Next, we consider (2-3) and (2-4). By the density of Schwartz functions in $\dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$, we may assume that $u_0, u_1 \in \mathcal{S}(\mathbb{R}^5)$. We claim it suffices to show

$$\left\| \frac{1}{|x|^2} \partial_r u_0(x) \right\|_{L^1_x(\mathbb{R}^5)} + \left\| \frac{1}{|x|} |\Delta u_0(x)| \right\|_{L^1_x(\mathbb{R}^5)} + \left\| \frac{1}{|x|^3} u_0(x) \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_0\|_{\dot{B}^3_{1,1}(\mathbb{R}^5)}, \tag{2-7}$$

$$\left\| \frac{1}{|x|} \partial_r u_1(x) \right\|_{L^1_x(\mathbb{R}^5)} + \left\| \frac{1}{|x|^2} u_1(x) \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_1\|_{\dot{B}^2_{1,1}(\mathbb{R}^5)}. \tag{2-8}$$

To see this, by using the fact $\Delta f = \partial_{rr} f + \frac{4}{r} \partial_r f$ for radial function $f(x)$ on $\mathbb{R}^5$, we have

$$\left\| \frac{1}{|x|} \partial_{rr} u_0(x) \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \left\| \frac{1}{|x|^2} \partial_r u_0(x) \right\|_{L^1_x(\mathbb{R}^5)} + \left\| \frac{1}{|x|} |\Delta u_0(x)| \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_0\|_{\dot{B}^3_{1,1}(\mathbb{R}^5)} \tag{2-9}$$

From the fundamental theorem of calculus and polar coordinates, for $y \in \mathbb{R}^5 \setminus \{0\}$,

$$|y|^3 |\partial_r u_0(y)| \lesssim \int_{|y|}^\infty \int_{\mathbb{S}^4} r^3 |\partial_{rr} u_0(r)|\, d\sigma(\omega)\, dr \lesssim \left\| \frac{1}{|x|} \partial_{rr} u_0(x) \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_0\|_{\dot{B}^3_{1,1}(\mathbb{R}^5)} \tag{2-10}$$

and

$$|y|^3 |u_1(y)| \lesssim \int_{|y|}^\infty \int_{\mathbb{S}^4} r^3 |\partial_{rr} u_1(r)|\, d\sigma(\omega)\, dr \lesssim \left\| \frac{1}{|x|} \partial_r u_1(x) \right\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_1\|_{\dot{B}^2_{1,1}(\mathbb{R}^5)}. \tag{2-11}$$

Hence, we are reduced to proving (2-7) and (2-8). We just give the estimate for the first term on the left-hand side of (2-7), since others can be handled similarly. For $j \in \mathbb{Z}$, utilizing Bernstein's estimates and polar coordinates, we obtain

$$\begin{aligned}\Big\|\frac{1}{|x|^2}\partial_r P_j u_0(x)\Big\|_{L^1_x(\mathbb{R}^5)} &\lesssim \int_0^\infty \int_{\mathbb{S}^4} r^2 \partial_r P_j u_0(r)\, d\sigma(\omega)\, dr \\ &\lesssim \int_0^{2^{-j}} \frac{1}{r^2}|\partial_r(P_j u_0)|r^4\, dr + \int_{2^{-j}}^\infty \frac{1}{r^2}|\partial_r(P_j u_0)|r^4\, dr \\ &\lesssim 2^{-3j}\|\partial_r(P_j u_0)\|_{L^\infty_r(\mathbb{R}_+)} + 2^{2j}\|\partial_r(P_j u_0)\|_{L^1_x(\mathbb{R}^5)} \\ &\lesssim 2^{3j}\|P_j u_0\|_{L^1_x(\mathbb{R}^5)}.\end{aligned} \tag{2-12}$$

Thus, we have $\|(1/|x|^2)\partial_r u_0(x)\|_{L^1_x(\mathbb{R}^5)} \lesssim \|u_0\|_{\dot{B}^3_{1,1}(\mathbb{R}^5)}$. □

As a direct consequence of Lemma 2.3, we have

$$\begin{aligned}\big\||x|^{1/2}\partial_r u_0(x)\big\|_{L^2_x(\mathbb{R}^5)} + \big\||x|^{-1/2}u_0(x)\big\|_{L^2_x(\mathbb{R}^5)} + \big\||x|^{1/2}u_1(x)\big\|_{L^2_x(\mathbb{R}^5)} \\ \lesssim \|(u_0,u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}.\end{aligned} \tag{2-13}$$

Lemma 2.4. *Suppose $\chi(x) \in C_c^\infty(\mathbb{R}^5)$. Let $R = 2^k$ be a dyadic number for $k \in \mathbb{Z}$ and denote $\chi_R(x) = \chi(\frac{x}{R})$. Then we have*

$$\|\chi_R(x) f\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)} \lesssim \|f\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)}, \tag{2-14}$$

$$\|\chi_R(x) g\|_{\dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} \lesssim \|g\|_{\dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)}, \tag{2-15}$$

where the bound is independent of R. Furthermore, if $\chi(x) = 1$ on $|x| \le 1$, then for $(f,g) \in \dot{B}^{1/2}_{2,1} \times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)$, we have

$$\lim_{R\to\infty} \|(1-\chi_R(x))f\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)} + \|(1-\chi_R(x))g\|_{\dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} = 0. \tag{2-16}$$

Proof. By scaling, to prove the inequalities (2-14) and (2-15), it suffices to prove the cases for $R = 1$, which follows from a similar proof of Lemma 2.2 in [Dodson 2019]. On the other hand, (2-16) follows from (2-14), (2-15), and the fact that $C_c^\infty \times C_c^\infty(\mathbb{R}^5)$ is dense in $\dot{B}^{1/2}_{2,1} \times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)$. □

Finally, we need the following chain rule estimates for later use.

Lemma 2.5 (C^1-fractional chain rule [Christ and Weinstein 1991]). *Suppose $G \in C^1(\mathbb{C})$, $s \in (0,1]$, and $1 < q, q_1, q_2 < \infty$ satisfying $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$. Then*

$$\big\||\nabla|^s G(u)\big\|_{L^q(\mathbb{R}^d)} \lesssim \big\|G'(u)\big\|_{L^{q_1}(\mathbb{R}^d)}\big\||\nabla|^s u\big\|_{L^{q_2}(\mathbb{R}^d)}. \tag{2-17}$$

2B. ***Fundamental properties of the wave equations.*** Throughout the paper, by abuse of notations, we often write $u(t) = u(t,x)$ for simplicity and $u(t,r) = u(t,x)$ when $u(t,\cdot)$ is radially symmetric.

Recall the explicit formula for solution to the linear wave equation in 5 dimensions,

$$\begin{aligned} S(t)(f,g)(x) &= \cos(t|\nabla|)f(x) + \frac{\sin(t|\nabla|)}{|\nabla|}g(x) \\ &= \frac{1}{3\omega_5}\partial_t\Big[\frac{1}{t}\partial_t\Big]\Big(t^3\int_{|y|=1} f(x+ty)\,d\sigma(y)\Big) \\ &\qquad + \frac{1}{3\omega_5}\frac{1}{t}\partial_t\Big(t^3\int_{|y|=1} g(x+ty)\,d\sigma(y)\Big), \end{aligned} \tag{2-18}$$

where ω_5 is the surface area of the unit sphere in $\mathbb{R}^5$. When (f,g) is radially symmetric, for $t > 0$, (2-18) can be rewritten as

$$\begin{aligned} S(t)(f,g)(r) &= \frac{1}{2r^2}\big[(r-t)^2 f(r-t) + (r+t)^2 f(r+t)\big] \\ &\quad - \frac{t}{2r^3}\int_{|r-t|}^{r+t} s f(s)\,ds + \frac{1}{4r^3}\int_{|r-t|}^{r+t} s(s^2+r^2-t^2)g(s)\,ds. \end{aligned} \tag{2-19}$$

See also [Rammaha 1987; Lindblad and Sogge 1996; Colzani et al. 2002] for the radial solutions to general dimensions linear wave equation. From the explicit formula (2-18), we can obtain the following dispersive estimate.

Proposition 2.6 (dispersive estimate).

$$\|S(t)(u_0,u_1)\|_{L^\infty(\mathbb{R}^5)} \lesssim \frac{1}{t^2}\big[\|\nabla^3 u_0\|_{L^1(\mathbb{R}^5)} + \|\nabla^2 u_1\|_{L^1(\mathbb{R}^5)}\big]. \tag{2-20}$$

Proof. We give the proof for completeness. A similar proof for the 3-dimensional case can be found in [Killip and Visan 2011]. By (2-18), the free solution $S(t)(u_0,u_1)$ can be rewritten as

$$\begin{aligned} &\frac{1}{\omega_5}\int_{|y|=1} u_0(x+ty)\,d\sigma(y) + \frac{5t}{3\omega_5}\int_{|y|=1} y(\nabla u_0)(x+ty)\,d\sigma(y) \\ &\quad + \frac{t^2}{3\omega_5}\int_{|y|=1} y(\nabla^2 u_0)(x+ty)y\,d\sigma(y) + \frac{t}{\omega_5}\int_{|y|=1} u_1(x+ty)\,d\sigma(y) \\ &\qquad + \frac{t^2}{3\omega_5}\int_{|y|=1} y(\nabla u_1)(x+ty)\,d\sigma(y), \end{aligned} \tag{2-21}$$

which, with the fundamental theorem of calculus, yields (2-20). For instance, using polar coordinates, we can estimate the first term of (2-21) as

$$\begin{aligned}
\left|\frac{1}{\omega_5}\int_{|y|=1} u_0(x+ty)\,d\sigma(y)\right| \\
&= \left|\frac{1}{\omega_5}\int_t^\infty\int_s^\infty\int_\tau^\infty\int_{|y|=1}\frac{d^3}{d\rho^3}[u_0(x+\rho y)]\,d\sigma(y)\,d\rho\,d\tau\,ds\right| \\
&\lesssim \int_t^\infty\int_s^\infty\int_\tau^\infty\int_{|y|=1}|\nabla^3 u_0|(x+\rho y)\,d\sigma(y)\,d\rho\,d\tau\,ds \\
&\lesssim \int_t^\infty\int_s^\infty\frac{1}{\tau^4}\,d\tau\,ds\|\nabla^3 u_0\|_{L^1(\mathbb{R}^5)} \\
&\lesssim \frac{1}{t^2}\|\nabla^3 u_0\|_{L^1(\mathbb{R}^5)}.
\end{aligned} \tag{2-22}$$

The other terms can be dealt with similarly. □

We recall the Strichartz estimates of the wave equation in $\mathbb{R}^5$. Let $I\subset\mathbb{R}$ be an interval. We denote the spacetime norm $L_t^q W_x^{s,r}(I\times\mathbb{R}^5)$ of a function $u(t,x)$ on $I\times\mathbb{R}^5$ by

$$\|u\|_{L_t^q W_x^{s,r}(I\times\mathbb{R}^5)} := \big\|\|u(t,x)\|_{W_x^{s,r}(\mathbb{R}^5)}\big\|_{L_t^q(I)},$$

for $s\in\mathbb{R}$, $1\le q,\ r\le\infty$. We denote that a pair (q,r) of exponents is admissible, if

$$2\le q\le\infty,\quad 2\le r<\infty,\quad \text{and}\quad \frac{1}{q}+\frac{2}{r}\le 1. \tag{2-23}$$

Moreover, we say (q,r) is wave acceptable, provided

$$1\le q<\infty,\quad 2\le r\le\infty,\quad \frac{1}{q}<4\Big(\frac{1}{2}-\frac{1}{r}\Big), \tag{2-24}$$

or $(q,r)=(\infty,2)$.

Proposition 2.7 (Strichartz estimates [Lindblad and Sogge 1995; Ginibre and Velo 1995; Keel and Tao 1998]). *Let $(u_0,u_1)\in\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)$ and (q,r), $(\tilde q,\tilde r)$ be two admissible pairs. If u is a weak solution to the wave equation $\partial_{tt}u-\Delta u=F(t,x)$ with initial data (u_0,u_1), then we have*

$$\begin{aligned}
\big\||\nabla|^\rho u\big\|_{L_t^q L_x^r(I\times\mathbb{R}^5)} + \sup_{t\in I}\|(u,u_t)\|_{\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)} \\
\lesssim \|(u_0,u_1)\|_{\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)} + \big\||\nabla|^{-\mu}F\big\|_{L_t^{\tilde q'}L^{\tilde r'}(I\times\mathbb{R}^5)},
\end{aligned} \tag{2-25}$$

provided that

$$\rho=\frac{1}{q}+\frac{5}{r}-2\quad \text{and}\quad \mu=\frac{1}{\tilde q}+\frac{5}{\tilde r}-2. \tag{2-26}$$

Proposition 2.8 (inhomogeneous Strichartz estimates [Taggart 2010]). *Suppose that the exponent pairs* (q_1, r_1) *and* $(\tilde{q}_1, \tilde{r}_1)$ *are wave acceptable, and satisfy the scaling condition*

$$\frac{1}{q} + \frac{1}{\tilde{q}} = 2 - 2\Big(\frac{1}{r_1} + \frac{1}{\tilde{r}_1}\Big)$$

and the conditions

$$\frac{1}{q} + \frac{1}{\tilde{q}} < 1, \qquad \frac{1}{2} \le \frac{\tilde{r}_1}{r_1} \le 2.$$

Let $r \ge r_1,\ \tilde{r} \ge \tilde{r}_1,\ \rho \in \mathbb{R}$ *be such that*

$$\rho + 5\Big(\frac{1}{2} - \frac{1}{r}\Big) - \frac{1}{q} = 1 - \Big(\tilde{\rho} + 5\Big(\frac{1}{2} - \frac{1}{\tilde{r}}\Big) - \frac{1}{\tilde{q}}\Big).$$

If $F(t, x)$ *is in* $L_t^{\tilde{q}'}(\mathbb{R};\ \dot{B}^{-\tilde{\rho}}_{\tilde{r}',2}(\mathbb{R}^5))$ *and* u *is a weak solution to the inhomogeneous wave equation*

$$-\partial_t^2 u + \Delta u = F(t, x), \qquad u(0) = u_t(0) = 0, \tag{2-27}$$

then

$$\|u\|_{L_t^q(\mathbb{R};\dot{B}^{\rho}_{r,2}(\mathbb{R}^5))} \lesssim \|F(t, x)\|_{L_t^{\tilde{q}'}(\mathbb{R};\dot{B}^{-\tilde{\rho}}_{\tilde{r}',2}(\mathbb{R}^5))}. \tag{2-28}$$

Next, we recall the well-posedness theory and the perturbation theory of the Cauchy problem (1-5).

Definition 2.9 (solution). Let I be a time interval such that $0 \in I$. We say function $u : I \times \mathbb{R}^5 \to \mathbb{R}$ is a (strong) solution to the Cauchy problem (1-5) in I if it satisfies $(u, u_t)(0) = (u_0, u_1)$,

$$(u, u_t) \in C(I;\ \dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)) \cap L^3_{t,x}(I \times \mathbb{R}^5),$$

and the integral equation

$$u(t) = S(t)(u_0, u_1) - \int_0^t S(t-\tau)(0, |u|u(\tau))\, d\tau \tag{2-29}$$

for all $t \in I$.

Theorem 2.10 (local well-posedness [Lindblad and Sogge 1995; Rodriguez 2017]). *Let* $(u_0, u_1) \in \dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$ *with*

$$\|(u_0, u_1)\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le A.$$

There exists $\delta = \delta(A) > 0$ *such that, if*

$$\|S(t)(u_0, u_1)\|_{L^3_{t,x}([0,T]\times\mathbb{R}^5)} \le \delta, \qquad \text{for some } T > 0, \tag{2-30}$$

then there exists a unique solution u *to* (1-5) *in* $[0, T] \times \mathbb{R}^5$, *such that*

$$\sup_{0\le t\le T} \|(u, u_t)\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} + \|u\|_{L^3_{t,x}([0,T]\times\mathbb{R}^5)} \le C(A). \tag{2-31}$$

In addition, if $A > 0$ is small enough, we can take $T = \infty$.

We define $T_+(u_0, u_1) := \sup I$, where I is the maximal interval of existence of the solution u.

Lemma 2.11 (standard blowup criterion). *Suppose u is the solution to the Cauchy problem (1-5) with initial data $(u_0, u_1) \in \dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$ and $T_+(u_0, u_1) < \infty$. Then we have*

$$\|u\|_{L^3_{t,x}([0,T_+(u_0,u_1))\times\mathbb{R}^5)} = \infty. \tag{2-32}$$

The proof is standard and similar to the energy critical case in [Kenig 2015].

We end this section by recalling the stability lemma for the Cauchy problem (1-5), which plays an important role in the Theorem 4.1.

Theorem 2.12 (perturbation theory [Rodriguez 2017]). *Let $I \subset \mathbb{R}$ be a time interval with $0 \in I$. Let $(u_0, u_1) \in \dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$ and some constants $M, A, A' > 0$ be given. Let $\tilde{u}$ be defined on $I \times \mathbb{R}^5$ and satisfy*

$$\sup_{t\in I} \|(\tilde{u}, \partial_t \tilde{u})\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le A, \tag{2-33}$$

$$\|\tilde{u}\|_{L^3_{t,x}(I\times\mathbb{R}^5)} \le M. \tag{2-34}$$

Assume that $\partial_{tt}\tilde{u} - \Delta\tilde{u} = -|\tilde{u}|\tilde{u} + e$ on $I \times \mathbb{R}^5$,

$$\big\|\big(u_0 - \tilde{u}(0), u_1 - \partial_t\tilde{u}(0)\big)\big\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le A' \tag{2-35}$$

and that

$$\|e\|_{L^{3/2}_{t,x}(I\times\mathbb{R}^5)} + \big\|S(t)\big[(\tilde{u}(0), \partial_t\tilde{u}(0)) - (u_0, u_1)\big]\big\|_{L^3_{t,x}(I\times\mathbb{R}^5)} < \varepsilon. \tag{2-36}$$

Then, there exist $\beta > 0$ and $\varepsilon_0 = \varepsilon_0(M, A, A') > 0$ such that, for $0 < \varepsilon < \varepsilon_0$, there exists a solution u to (1-5) in I such that $(u(0), \partial_t u(0)) = (u_0, u_1)$, with

$$\|u\|_{L^3_{t,x}(I\times\mathbb{R}^5)} \le C(M, A, A'), \tag{2-37}$$

$$\sup_{t\in I} \|(\partial_t\tilde{u}(t), \tilde{u}(t)) - (u, \partial_t u(t))\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le C(M, A, A')(A' + \varepsilon^\beta). \tag{2-38}$$

3. Decomposition of the solution and global well-posedness

In this section, we will prove that for any given initial data $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$, the corresponding solution u to (1-5) is globally well-posed. To prove this, we first show the solution u belongs to some suitable Strichartz-type spaces on a local time interval. Then, we split it into two parts: $u = v + w$. Based on the inhomogeneous Strichartz estimates (2-28), we will derive a decay property for v and prove that w is in the energy space $\dot{H}^1 \times L^2(\mathbb{R}^5)$. We remark that the constants in "$\lesssim$" in this section depend upon $\|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}$.

For the sake of simplicity, we denote $F(u) = |u|u$. Recall that u_λ is also a solution to the Cauchy problem (1-5) with initial data $(\lambda^2 u_0(\lambda x), \lambda^3 u_1(\lambda x))$, where

$$u_\lambda(t,x) = \lambda^2 u(\lambda t, \lambda x). \tag{3-1}$$

Given $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$, for any $\eta > 0$, there exists $j_0 = j_0(u_0, u_1, \eta) < \infty$ such that

$$\sum_{j \geq j_0} 2^{2j} \|P_j u_0\|_{L^1(\mathbb{R}^5)} < \eta. \tag{3-2}$$

Replace u by u_λ for $\lambda = 2^{-j_0}$, then we have

$$\sum_{j\geq 0} 2^{3j} \|P_j u_0\|_{L^1(\mathbb{R}^5)} + \sum_{j \geq 0} 2^{2j} \|P_j u_1\|_{L^1(\mathbb{R}^5)} < \eta. \tag{3-3}$$

Lemma 3.1. *Let $\epsilon_0 > 0$ be a small constant and $\eta \ll \epsilon_0$. If the initial data $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}(\mathbb{R}^5)$ satisfies (3-3) and u is the solution to (1-5) with initial data (u_0, u_1) given by Theorem 2.10, then there exists*

$$\delta = \delta(\epsilon_0, \|(u_0,u_1)\|_{\dot{B}^3_{1,1}\times \dot{B}^2_{1,1}}) > 0$$

such that

$$\|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \leq \sum_{j\in\mathbb{Z}} \|P_j u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \lesssim \epsilon_0, \tag{3-4}$$

$$\|u\|_{L^\infty_t([-\delta,\delta];\dot{B}^{1/2}_{2,1}(\mathbb{R}^5))} \lesssim \sum_{j\in\mathbb{Z}} \|P_j u\|_{L^\infty_t \dot{H}^{1/2}([-\delta,\delta]\times\mathbb{R}^5)} \lesssim 1. \tag{3-5}$$

Proof. By Strichartz's estimates in Proposition 2.7 and (3-3), we obtain

$$\|S(t)P_{\geq 0}(u_0,u_1)\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \leq \tfrac{1}{2}\epsilon_0. \tag{3-6}$$

On the other hand, for every $t \in \mathbb{R}$, by Bernstein, we have

$$\|S(t)P_{\leq 0}(u_0,u_1)\|_{L^3_x(\mathbb{R}^5)} \lesssim 1. \tag{3-7}$$

Hence, taking δ small enough, we have,

$$\|S(t)(u_0,u_1)\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \leq \epsilon_0. \tag{3-8}$$

Then, by the Strichartz estimates, we have

$$\|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \lesssim \|S(t)(u_0,u_1)\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} + \|u\|^2_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)}, \tag{3-9}$$

from which, by a standard continuity argument, we deduce that

$$\|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \lesssim \epsilon_0. \tag{3-10}$$

Let

$$a_k = \|P_k u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} + 2^{\frac{1}{2}k}\|P_k u\|_{L^\infty_t L^2_x([-\delta,\delta]\times\mathbb{R}^5)} + 2^{\frac{1}{4}k}\|P_k u\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)}, \tag{3-11}$$

$$b_k = 2^{\frac{1}{2}k}\|P_k u_0\|_{L^2_x} + 2^{-\frac{1}{2}k}\|P_k u_1\|_{L^2_x}. \tag{3-12}$$

By Young's inequality, it suffices to show there is a recurrence relation

$$a_k \lesssim b_k + \epsilon_0 \sum_j 2^{-\frac{1}{4}|j-k|} a_j. \tag{3-13}$$

To prove this, making use of the Strichartz estimates, we have

$$a_k \lesssim b_k + 2^{\frac{1}{4}k}\|P_k F(u)\|_{L^2_t L^{4/3}_x([0,\delta]\times\mathbb{R}^5)}. \tag{3-14}$$

First, we consider the low frequency part of the second term in the right-hand side of (3-14). By Lemma 2.5 and Hölder, we have

$$\begin{aligned} \|P_k F(u_{\le k})\|_{L^2_t L^{4/3}_x([-\delta,\delta]\times\mathbb{R}^5)} &\lesssim 2^{-\frac{1}{2}k}\|P_k|\nabla_x|^{\frac{1}{2}} F(u_{\le k})\|_{L^2_t L^{4/3}_x([-\delta,\delta]\times\mathbb{R}^5)} \\ &\lesssim 2^{-\frac{1}{2}k}\|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \big\||\nabla|^{1/2}(P_{\le k}u)\big\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)} \\ &\lesssim \|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} \sum_{j\le k} 2^{-\frac{1}{2}(k-j)}\|P_j u\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)}, \end{aligned} \tag{3-15}$$

from which it follows that

$$\begin{aligned} 2^{\frac{k}{4}}\|P_k F(u_{\le k})\|_{L^2_t L^{4/3}_x([-\delta,\delta]\times\mathbb{R}^5)} &\lesssim \epsilon_0 \sum_{j\le k} 2^{-\frac{1}{4}(k-j)} 2^{\frac{1}{4}j}\|P_j u\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)} \\ &\lesssim \epsilon_0 \sum_{j\le k} 2^{-\frac{1}{4}(k-j)} a_j. \end{aligned} \tag{3-16}$$

On the other hand, by Hölder's inequality,

$$\begin{aligned} &2^{\frac{1}{4}k}\|P_k(F(u)-F(u_{\le k}))\|_{L^2_t L^{4/3}_x([-\delta,\delta]\times\mathbb{R}^5)} \\ &\lesssim \|u\|_{L^3_{t,x}([-\delta,\delta]\times\mathbb{R}^5)} 2^{\frac{1}{4}k}\|P_{\ge k+1}u\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)} \\ &\lesssim \epsilon_0 \sum_{j\ge k+1} 2^{-\frac{1}{4}(j-k)} 2^{\frac{1}{4}j}\|P_j u\|_{L^6_t L^{12/5}_x([-\delta,\delta]\times\mathbb{R}^5)} \lesssim \epsilon_0 \sum_{j\ge k+1} 2^{-\frac{1}{4}(j-k)} a_j. \end{aligned} \tag{3-17}$$

Then the recurrence relation (3-13) follows from (3-16) and (3-17). □

Note that by the inequality (3-5), the inequalities (3-16) and (3-17) yield that

$$\sum_{k\in\mathbb{Z}} 2^{\frac{1}{4}k}\|P_k(F(u))\|_{L^2_t L^{4/3}_x([0,\delta]\times\mathbb{R}^5)} \lesssim \epsilon_0. \tag{3-18}$$

As an application of Lemma 3.1 and the radial Sobolev inequality (2-2), we will see that the solution u possesses some space decay property in the region $\{|x| \geq |t| + C\}$ for some large constant $C > 0$. Let $\chi(x)$ be a smooth cutoff function such that $\chi(x) = 1$ for $|x| \leq \frac{1}{2}$ and $\chi(x) = 0$ for $|x| \geq 1$. By Lemma 2.4, there exists a dyadic integer $R = R(u_0, u_1, \epsilon_0)$ such that $\|(1-\chi(\frac{x}{R}))(u_0, u_1)\|_{\dot{B}^{1/2}_{2,1} \times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} \leq \epsilon_0$. Then by scaling, we have

$$\big\|(1-\chi(2x))\big((2R)^2 u_0(2Rx), (2R)^3 u_1(2Rx)\big)\big\|_{\dot{B}^{1/2}_{2,1} \times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} \leq \epsilon_0. \tag{3-19}$$

By abuse of notations, we will still use (u_0, u_1) to represent the initial data $\big((2R)^2 u_0(2Rx), (2R)^3 u_1(2Rx)\big)$. Then we have,

$$\|(1-\chi(2x))(u_0, u_1)\|_{\dot{B}^{1/2}_{2,1} \times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} \leq \epsilon_0. \tag{3-20}$$

In addition, by Lemma 3.1, we have

$$\|u\|_{L^3_{t,x}([-\delta/(2R),\, \delta/(2R)] \times \mathbb{R}^5)} \lesssim \epsilon_0, \tag{3-21}$$

$$\|u\|_{L^\infty_t([-\delta/(2R),\, \delta/(2R)];\, \dot{B}^{1/2}_{2,1}(\mathbb{R}^5))} \lesssim 1. \tag{3-22}$$

Lemma 3.2. *Let $J \subset \mathbb{R}$ be an interval such that u is a solution to (1-5) on J. Then*

$$\|u\|_{L^3_{t,x}(\{(t,x) \in J \times \mathbb{R}^5 :\, |x| \geq |t| + \frac{1}{2}\})} + \sup_{t \in J} \big\| |x|^2 u(t,x) \big\|_{L^\infty_x(\{x \in \mathbb{R}^5 :\, |x| \geq |t| + \frac{1}{2}\})} \lesssim \epsilon_0. \tag{3-23}$$

Proof. Let $U(t,x)$ be the solution to (1-5) with initial data $(1-\chi(2x))(u_0(x), u_1(x))$. Employing Theorem 2.10 and arguing by similar arguments in Lemma 3.1, one can deduce (3-14) when u is replaced by U and $[-\delta, \delta]$ is replaced by $\mathbb{R}$. Thus

$$\|U\|_{L^3_{t,x}(\mathbb{R} \times \mathbb{R}^5)} + \|U\|_{L^\infty_t \dot{B}^{1/2}_{2,1}(\mathbb{R} \times \mathbb{R}^5)} \lesssim \epsilon_0. \tag{3-24}$$

Due to the finite propagation speed property of the wave equation (1-5), we have $u(t,x) = U(t,x)$ when $|x| \geq |t| + \frac{1}{2}$. Then (3-23) follows from (3-24) and the radial Sobolev inequality (2-2). □

Next, we want to show the following local properties, which will play an important role in Section 3B. Unlike the case of three dimensions in [Dodson 2019], we will make use of the inhomogeneous Strichartz estimates (2-28) to conquer the difficulties caused by the higher order decay of time.

Lemma 3.3. *If ϵ_0 is sufficiently small and δ is as given in Lemma 3.1, then, for $3 < r < 4$, we have*

$$\sup_{-\frac{2\delta}{2R} < t < \frac{2\delta}{2R}} t^{(2r-5)/r} \|u\|_{L^r_x(\mathbb{R}^5)} + \|u\|_{L^{5/4}_t L^{25/6}_x([0,\, \delta/(2R)] \times \mathbb{R}^5)} \lesssim 1. \tag{3-25}$$

We remark that for $3 < r < \infty$, the space $L^{r/(2r-5)}_t L^r_x(\mathbb{R} \times \mathbb{R}^5)$ is $\dot{H}^{1/2}$-critical but not admissible.

Proof. First, we consider the estimates for the linear part. Utilizing dispersive estimate (2-20), we have

$$\|S(t)P_j(u_0,u_1)\|_{L^\infty_x(\mathbb{R}^5)} \lesssim \frac{1}{t^2}\big[2^{3j}\|P_ju_0\|_{L^1(\mathbb{R}^5)} + 2^{2j}\|P_ju_1\|_{L^1(\mathbb{R}^5)}\big]. \tag{3-26}$$

By Bernstein, we have

$$\begin{aligned}\|S(t)P_j(u_0,u_1)\|_{L^2_x(\mathbb{R}^5)} &\lesssim \|P_ju_0\|_{L^2_x(\mathbb{R}^5)} + 2^{-j}\|P_ju_1\|_{L^2_x(\mathbb{R}^5)}\\ &\lesssim 2^{\frac{5}{2}j}\|P_ju_0\|_{L^1(\mathbb{R}^5)} + 2^{\frac{3}{2}j}\|P_ju_1\|_{L^1(\mathbb{R}^5)}.\end{aligned} \tag{3-27}$$

Interpolating this inequality with the estimate (3-26) yields that,

$$\|S(t)P_j(u_0,u_1)\|_{L^r_x(\mathbb{R}^5)} \lesssim t^{-2(1-\frac{2}{r})}2^{-\frac{j}{r}}\big[2^{3j}\|P_ju_0\|_{L^1(\mathbb{R}^5)} + 2^{2j}\|P_ju_1\|_{L^1(\mathbb{R}^5)}\big]. \tag{3-28}$$

On the other hand, for $r \geq \frac{5}{2}$, employing Bernstein's estimates, we have

$$\begin{aligned}\|S(t)P_j(u_0,u_1)\|_{L^r_x(\mathbb{R}^5)} &\lesssim 2^{5(\frac{1}{2}-\frac{1}{r})j}\|S(t)P_j(u_0,u_1)\|_{L^2_x(\mathbb{R}^5)}\\ &\lesssim 2^{(2-\frac{5}{r})j}\big[2^{3j}\|P_ju_0\|_{L^1(\mathbb{R}^5)} + 2^{2j}\|P_ju_1\|_{L^1(\mathbb{R}^5)}\big].\end{aligned} \tag{3-29}$$

This estimate and (3-28) yield that, for $r \geq \frac{5}{2}$,

$$\begin{aligned}\sup_{t\in\mathbb{R}} t^{(2r-5)/r}\|S(t)(u_0,u_1)\|_{L^r_x(\mathbb{R}^5)} + \|S(t)(u_0,u_1)\|_{L_t^{r/(2r-5)}L^r_x(\mathbb{R}\times\mathbb{R}^5)}\\ \lesssim \|(u_0,u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}.\end{aligned} \tag{3-30}$$

By the reversal property of the wave equation, it suffices to prove (3-25) for $t \geq 0$. Using the inhomogeneous Strichartz estimates (2-28), Lemmas 2.5, 3.1, and Hölder, we have

$$\begin{aligned}\Big\|\int_0^t S(t-\tau)(0,F(u(\tau)))\,d\tau\Big\|_{L_t^{5/4}L_x^{25/6}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}\\ \lesssim \big\||\nabla|^{1/4}F(u)\big\|_{L_t^{30/29}L_x^{300/197}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}\\ \lesssim \big\||\nabla|^{1/2}u\big\|^{1/2}_{L_t^\infty L_x^2([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}\|u\|^{1/2}_{L^3_{t,x}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}\|u\|_{L_t^{5/4}L_x^{25/6}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}.\end{aligned} \tag{3-31}$$

This estimates together with the estimate (3-30) yields

$$\|u\|_{L_t^{5/4}L_x^{25/6}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)} \lesssim 1, \tag{3-32}$$

provided $0 < \epsilon_0 \ll \min(1, \|(u_0,u_1)\|^{-1}_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)})$.

Let $c \in (0,1)$ to be chosen later. First, employing the dispersive estimate (2-20), Lemma 2.5, and interpolation, for $r \in (3,4)$, we have

$$
\begin{aligned}
(3\text{-}33)\quad &\sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \Big\| \int_{(1-c)t}^{t} S(t-\tau)(0, F(u(\tau)))\,d\tau \Big\|_{L^r(\mathbb{R}^5)} \\
&\lesssim \sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \int_{(1-c)t}^{t} \frac{1}{(t-\tau)^{2-\frac{4}{r}}} \big\| |\nabla|^{2-\frac{6}{r}} F(u(\tau)) \big\|_{L^{r'}(\mathbb{R}^5)}\,d\tau \\
&\lesssim \sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \int_{(1-c)t}^{t} \frac{1}{(t-\tau)^{2-\frac{4}{r}}} \|u(\tau)\|_{L^r(\mathbb{R}^5)}^{\frac{r-1}{2r-5}} \\
&\qquad \times \|u(\tau)\|_{\dot H_x^{1/2}(\mathbb{R}^5)}^{4-\frac{12}{r}} \|u(\tau)\|_{L_x^{5/2}(\mathbb{R}^5)}^{(\frac{12}{r}-2-\frac{r-1}{2r-5})}\,d\tau \\
&\lesssim \Big(\sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \|u(t)\|_{L^r(\mathbb{R}^5)}\Big)^{\frac{r-1}{2r-5}} \|u\|_{L_t^\infty \dot H_x^{1/2}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}^{4-\frac{12}{r}} \\
&\qquad \times \|u\|_{L_t^\infty L_x^{5/2}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}^{(\frac{12}{r}-2-\frac{r-1}{2r-5})} \cdot \int_{(1-c)t}^{t} \frac{1}{(t-\tau)^{2-\frac{4}{r}}} t^{1-\frac{4}{r}}\,d\tau \\
&\lesssim c^{4/r-1} \Big(\sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \|u(t)\|_{L^r(\mathbb{R}^5)}\Big)^{\frac{r-1}{2r-5}}.
\end{aligned}
$$

For the remainder part, we utilize the dispersive estimate (2-20), Lemma 2.5, the Hölder inequality, and interpolation to obtain

$$
\begin{aligned}
(3\text{-}34)\quad & t^{\frac{2r-5}{r}} \Big\| \int_0^{(1-c)t} S(t-\tau)(0, F(u(\tau)))\,d\tau \Big\|_{L^r(\mathbb{R}^5)} \\
&\lesssim t^{\frac{2r-5}{r}} \int_0^{(1-c)t} \frac{1}{(t-\tau)^{2-\frac{4}{r}}} \big\| |\nabla|^{2-\frac{6}{r}} F(u(\tau)) \big\|_{L_x^{r'}(\mathbb{R}^5)}\,d\tau \\
&\lesssim c^{\frac{4}{r}-2} \big\| |\nabla|^{2-\frac{6}{r}} F(u) \big\|_{L_{t,x}^{r'}(\mathbb{R}\times\mathbb{R}^5)} \\
&\lesssim c^{\frac{4}{r}-2} \|u\|_{L_t^{r/(2r-5)} L_x^r(\mathbb{R}\times\mathbb{R}^5)} \|u\|_{L_t^\infty \dot H_x^{1/2}(\mathbb{R}\times\mathbb{R}^5)}^{4-\frac{12}{r}} \|u\|_{L_{t,x}^3(\mathbb{R}\times\mathbb{R}^5)}^{\frac{12}{r}-3} \\
&\lesssim c^{\frac{4}{r}-2} \epsilon_0^{\frac{159}{7r}-\frac{39}{7}},
\end{aligned}
$$

where in the last step we used the fact that

$$
\|u\|_{L_t^{r/(2r-5)} L_x^r([0,\frac{\delta}{2R}]\times\mathbb{R}^5)} \lesssim \|u\|_{L_{t,x}^3([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}^{\frac{75}{7}(\frac{1}{r}-\frac{6}{25})} \|u\|_{L_t^{5/4} L_x^{25/6}([0,\frac{\delta}{2R}]\times\mathbb{R}^5)}^{\frac{25}{7}(1-\frac{3}{r})} \lesssim \epsilon_0^{\frac{75}{7}(\frac{1}{r}-\frac{6}{25})}.
$$

Hence, by (3-30) and (3-32)–(3-34), taking $c > 0$ small enough and using a continuity method, there exists $\epsilon_0 = \epsilon_0(\|(u_0,u_1)\|_{\dot B_{1,1}^3 \times \dot B_{1,1}^2}) > 0$ such that

$$
(3\text{-}35)\qquad \sup_{t\in[0,\frac{\delta}{2R}]} t^{\frac{2r-5}{r}} \|u\|_{L_x^r} \lesssim 1, \qquad \text{for } 3 < r < 4. \qquad \square
$$

We denote $\delta_1 = \frac{\delta}{2R}$ for simplicity. Let $\psi \in C_0^\infty(\mathbb{R}^5)$ be supported in $|x| \le \frac{\delta_1}{10}$ and $\psi(x) = 1$ when $|x| \le \frac{\delta_1}{20}$. Then we can assume that $|(\nabla\psi)(x)| \lesssim \frac{1}{\delta_1}$. For $t \ge \delta_1$, we split $u(t,x) = v(t,x) + w(t,x)$, where

$$v(t) = S(t)(\psi u_0, \psi u_1) - \int_0^{\delta_1/10} S(t-\tau)(0, \psi F(u(\tau)))\,d\tau. \tag{3-36}$$

We will prove that v has a decay property and w has finite energy.

3A. ***The decay part of the solution u.***

Lemma 3.4. *For $t \ge \delta_1$, we have*

$$\|v(t)\|_{L_x^\infty(\mathbb{R}^5)} \lesssim \delta_1^{-1/2} t^{-2}. \tag{3-37}$$

In addition, we have

$$\|v\|_{L_t^\infty(\mathbb{R};\dot{B}_{2,1}^{1/2}(\mathbb{R}^5))} + \|v\|_{L_{t,x}^3(\mathbb{R}\times\mathbb{R}^5)} \lesssim \delta_1^{-1/2}. \tag{3-38}$$

Proof. We first estimate the linear part of v. By the Huygens principle, the radial Sobolev inequality (2-2) and Lemma 2.4, we have, for $t \ge \delta_1$

$$\|S(t)(\psi u_0, \psi u_1)\|_{L_x^\infty(\mathbb{R}^5)} \lesssim \frac{1}{t^2}\|(u_0,u_1)\|_{\dot{B}_{2,1}^{1/2}\times\dot{B}_{2,1}^{-1/2}} \lesssim \frac{1}{t^2}. \tag{3-39}$$

For the second part of v and $t \ge \delta_1$, using the Radial Sobolev inequality (2-2), the Huygens principle and the Strichartz estimates, we obtain

$$\begin{aligned}
&\Big\|\int_0^{\delta_1/10} S(t-\tau)(0,\psi F(u(\tau)))\,d\tau\Big\|_{L_x^\infty(\mathbb{R}^5)} \\
&\quad\lesssim \frac{1}{t^2}\Big\|\int_0^{\delta_1/10} S(t-\tau)(0,\tilde{\chi}F(u(\tau)))\,d\tau\Big\|_{\dot{B}_{2,1}^{1/2}(\mathbb{R}^5)} \\
&\quad\lesssim \frac{1}{t^2}\Big\|\int_0^{\delta_1/10} \frac{\sin(\tau|\nabla|)}{|\nabla|}(\psi F(u(\tau)))\,d\tau\Big\|_{\dot{B}_{2,1}^{1/2}} \\
&\qquad\qquad + \frac{1}{t^2}\Big\|\int_0^{\delta_1/10} \frac{\cos(\tau|\nabla|)}{|\nabla|}(\psi F(u(\tau)))\,d\tau\Big\|_{\dot{B}_{2,1}^{1/2}}
\end{aligned} \tag{3-40}$$

$$\lesssim \frac{1}{t^2}\sum_{j<0} 2^{-\frac{1}{2}j}\big\|P_j\big[\psi F(u(\tau))\big]\big\|_{L_t^1L_x^2([0,\delta_1/10]\times\mathbb{R}^5)} \tag{3-41}$$

$$+ \frac{1}{t^2}\sum_{j\ge 0} 2^{\frac{1}{4}j}\big\|P_j\big[\psi F(u(\tau))\big]\big\|_{L_t^2L_x^{4/3}([0,\delta_1/10]\times\mathbb{R}^5)} \tag{3-42}$$

For the low frequency part (3-41), by Bernstein's estimates and Hölder's inequality, we have

$$t^2 \times (3\text{-}41) \lesssim \sum_{j<0} 2^{\frac{j}{3}}\|F(u)\|_{L_t^1L_x^{3/2}([0,\delta_1/10]\times\mathbb{R}^5)} \lesssim \delta_1^{\frac{1}{3}}\|u\|_{L_{t,x}^3([0,\delta_1/10]\times\mathbb{R}^5)}^2 \lesssim 1. \tag{3-43}$$

For (3-42), it suffices to estimate

$$\sum_{j\geq 0} 2^{\frac{1}{4}j} \big\| [P_j, \psi] F(u(\tau)) \big\|_{L_t^2 L_x^{4/3}([0,\delta_1/10]\times\mathbb{R}^5)} \tag{3-44}$$

$$+ \sum_{j\geq 0} 2^{\frac{1}{4}j} \big\| \psi P_j F(u(\tau)) \big\|_{L_t^2 L_x^{4/3}([0,\delta_1/10]\times\mathbb{R}^5)}. \tag{3-45}$$

For (3-44), by commutator estimates, Young's inequality, the Sobolev embedding and Lemma 3.1, we have

$$\begin{aligned} (3\text{-}44) &\lesssim \sum_{j\geq 0} 2^{-\frac{3}{4}j} \delta_1^{-1} \| Q_j F(u(\tau)) \|_{L_\tau^2 L_x^{4/3}([0,\delta_1/10]\times\mathbb{R}^5)} \\ &\lesssim \delta_1^{-1} \sum_{j\geq 0} 2^{-\frac{3}{4}j} 2^{\frac{1}{4}j} \| F(u(\tau)) \|_{L_\tau^2 L_x^{5/4}([0,\delta_1/10]\times\mathbb{R}^5)} \\ &\lesssim \delta_1^{-\frac{1}{2}} \|u\|^2_{L_t^\infty L_x^{5/2}([0,\delta_1/10]\times\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}, \end{aligned} \tag{3-46}$$

where

$$Q_j f(x) = 2^{6j} \int_{\mathbb{R}^5} |y| |\phi(2^j y)| |f|(x-y)\, dy$$

and in the first inequality we used the mean value theorem. For (3-45), by the estimate (3-18), we have

$$(3\text{-}45) \lesssim \sum_{j\geq 0} 2^{\frac{1}{4}j} \| P_j F(u) \|_{L_\tau^2 L_x^{4/3}} \lesssim \epsilon_0. \tag{3-47}$$

Hence, by (3-39)–(3-47), we have $\|v(t)\|_{L_x^\infty(\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}} t^{-2}$.

Now we consider (3-38). For simplicity, we write

$$\|v\|_{S(\mathbb{R})} := \|v\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} + \|v\|_{L_t^\infty(\mathbb{R};\dot{B}^{1/2}_{2,1}(\mathbb{R}^5))}.$$

For the linear part, by the Strichartz estimates and Lemma 2.4, we have

$$\|S(t)(\psi u_0, \psi u_1)\|_{S(\mathbb{R})} \lesssim \|(u_0, u_1)\|_{\dot{B}^{1/2}_{2,1}\times \dot{B}^{-1/2}_{2,1}(\mathbb{R}^5)} \lesssim 1. \tag{3-48}$$

By the Strichartz estimates and repeating the arguments that deal with (3-40),

$$\begin{aligned} &\Big\| \int_0^{\delta_1/10} S(t-\tau)(0, \psi F(u(\tau)))\, d\tau \Big\|_{S(\mathbb{R})} \\ &\quad \lesssim \Big\| \int_0^{\delta_1/10} \frac{\sin(\tau|\nabla|)}{|\nabla|} (\psi F(u(\tau)))\, d\tau \Big\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)} \\ &\qquad\qquad + \Big\| \int_0^{\delta_1/10} \frac{\cos(\tau|\nabla|)}{|\nabla|} (\psi F(u(\tau)))\, d\tau \Big\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}. \end{aligned}$$

This completes the proof. □

3B. *The energy part of the solution u.*

Lemma 3.5. *We have*

$$\|\vec{w}(\delta_1)\|_{\dot{H}^1\times L^2(\mathbb{R}^5)} \lesssim \delta_1^{-1/2}. \tag{3-49}$$

Proof. By the definition of w, for $t \geq \delta_1$, we have

$$\begin{aligned} w(t) &= S(t)\big((1-\psi)u_0, (1-\psi)u_1\big) \\ &\quad - \int_0^{\frac{\delta_1}{10}} S(t-\tau)(0, (1-\psi)F(u(\tau)))\,d\tau - \int_{\frac{\delta_1}{10}}^t S(t-\tau)(0, F(u(\tau)))d\tau. \end{aligned} \tag{3-50}$$

First, we consider the contribution of the third term of (3-50). Taking $r = \frac{50}{13}$ in (3-25), by interpolation, we have

$$\begin{aligned} &\|u\|^2_{L_t^2 L_x^4([\delta_1/10,\delta_1]\times\mathbb{R}^5)} \\ &\quad = \int_{\delta_1/10}^{\delta_1} \|u\|_{L_x^{50/13}(\mathbb{R}^5)} \|u\|_{L_x^{25/6}(\mathbb{R}^5)}\,dt \\ &\quad \lesssim \delta_1^{-\frac{1}{2}} \sup_{t\in[\frac{\delta_1}{10},\delta_1]} \Big[t^{\frac{7}{10}} \|u(t)\|_{L_x^{50/13}(\mathbb{R}^5)}\Big] \|u\|_{L_t^{5/4} L_x^{25/6}([\frac{\delta_1}{10},\delta_1]\times\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}. \end{aligned} \tag{3-51}$$

From this inequality and Strichartz, we have

$$\begin{aligned} &\Big\|\int_{\delta_1/10}^{\delta_1} S(\delta_1-\tau)(0, F(u(\tau)))d\tau\Big\|_{\dot{H}_x^1(\mathbb{R}^5)} \\ &\quad + \Big\|\partial_t\Big[\int_{\delta_1/10}^{t} S(t-\tau)(0, F(u(\tau)))\,d\tau\Big]_{t=\delta_1}\Big\|_{L_x^2(\mathbb{R}^5)} \\ &\quad\quad \lesssim \|F(u)\|_{L_\tau^1 L_x^2([\frac{\delta_1}{10},\delta_1]\times\mathbb{R}^5)} \lesssim \|u\|^2_{L_\tau^2 L_x^4([\frac{\delta_1}{10},\delta_1]\times\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}. \end{aligned} \tag{3-52}$$

By Strichartz, radial Sobolev inequality (2-2) and Hölder, the second term of (3-50) can be estimated as

$$\begin{aligned} &\Big\|\int_0^{\delta_1/10} \vec{S}(t-\tau)\big(0, (1-\psi)F(u(\tau))\big)\,d\tau\Big|_{t=\delta_1}\Big\|_{\dot{H}^1\times L^2(\mathbb{R}^5)} \\ &\quad \lesssim \|(1-\psi)F(u(\tau))\|_{L_\tau^1 L_x^2([0,\frac{\delta_1}{10}]\times\mathbb{R}^5)} \\ &\quad \lesssim \delta_1^{\frac{1}{2}} \|(1-\psi)u\|^{\frac{1}{2}}_{L_{t,x}^\infty([0,\frac{\delta_1}{10}]\times\mathbb{R}^5)} \|u\|^{\frac{3}{2}}_{L_{t,x}^3([0,\frac{\delta_1}{10}]\times\mathbb{R}^5)} \\ &\quad \lesssim \delta_1^{-\frac{1}{2}} \|u\|^{\frac{1}{2}}_{L_t^\infty \dot{B}_{2,1}^{1/2}} \lesssim \delta_1^{-\frac{1}{2}}. \end{aligned} \tag{3-53}$$

Hence, it remains to estimate

$$\|(1-\psi)(u_0, u_1)\|_{\dot{H}^1\times L^2(\mathbb{R}^5)}. \tag{3-54}$$

For u_0, by radial Sobolev inequality (2-2) and polar coordinates,

$$\|(\nabla\psi)u_0\|_{L^2(\mathbb{R}^5)}^2 \lesssim \delta_1^{-1}\int_{\delta_1/100}^{2\delta_1}\int_{\mathbb{S}^4} u_0^2(r)\,d\sigma(\omega)r^4\,dr \lesssim \|u_0\|_{\dot{B}^{1/2}_{2,1}(\mathbb{R}^5)}^2 \lesssim 1. \tag{3-55}$$

By the inequality (2-13), we have

$$\|(1-\psi)\partial_r u_0\|_{L^2_x(\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}. \tag{3-56}$$

For u_1, by the inequality (2-8) and polar coordinates, one can deduce that

$$\|(1-\psi)u_1\|_{L^2_x(\mathbb{R}^5)}^2 \lesssim \int_{\frac{\delta_1}{10}}^{\infty}\int_{\mathbb{S}^4} |u_1(r)|^2 r^4 d\sigma(\omega)\,dr \lesssim \int_{\frac{\delta_1}{10}}^{\infty} r^{-2}\,dr \lesssim \delta_1^{-1}. \tag{3-57}$$

This inequality together with (3-55) and (3-56) implies that

$$\|(1-\psi)(u_0, u_1)\|_{\dot{H}^1\times L^2(\mathbb{R}^5)} \lesssim \delta_1^{-\frac{1}{2}}. \tag{3-58}$$

This completes the proof. □

3C. *Global well-posedness.* In this subsection, we prove that the solution u is globally well-posed. We emphasize that the constants in " $\lesssim$ " in this subsection depend only upon δ_1 and $\|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}$.

Theorem 3.6. *Let u be the solution to (1-5) with initial data $(u_0, u_1) \in \dot{B}^3_{1,1} \times \dot{B}^2_{1,1}$. Then u is globally well-posed and such that for any compact interval $J \subset \mathbb{R}$,*

$$\|u\|_{L^3_{t,x}(J\times\mathbb{R}^5)} < C\big(J, \|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}}, \delta_1\big). \tag{3-59}$$

Proof. By Lemma 2.11, it suffices to show (3-59). By the time reversibility of the wave equation, we just need consider the part of $t \geq 0$.

For $t \geq \delta_1$, by $u(t) = w(t) + v(t)$ and the formula (3-36) of $v(t)$, we have

$$w_{tt} - \Delta w = -|u|u. \tag{3-60}$$

We define the energy of w as (1-2) by

$$E(w(t)) = \frac{1}{2}\int_{\mathbb{R}^5} |\nabla_t w(t)|^2\,dx + \frac{1}{2}\int_{\mathbb{R}^5} |\nabla_x w(t)|^2\,dx + \frac{1}{3}\int_{\mathbb{R}^5} |w(t)|^3\,dx. \tag{3-61}$$

By the estimates (3-22), (3-38), Lemma 3.5 and interpolation, we have

$$E(w(\delta_1)) \lesssim 1. \tag{3-62}$$

Now, we consider

$$\begin{aligned}(3\text{-}63)\quad \Big|\frac{d}{dt}E(w(t))\Big| &= \Big|\int_{\mathbb{R}^5}(|w|w-|u|u)w_t\,dx\Big| \\ &\lesssim \|v\|_{L_x^6(\mathbb{R}^5)}\,\|w_t\|_{L_x^2(\mathbb{R}^5)}\,\|w\|_{L_x^3(\mathbb{R}^5)}+\|w_t\|_{L_x^2(\mathbb{R}^5)}\,\|v\|_{L_x^6(\mathbb{R}^5)}\,\|v\|_{L_x^3(\mathbb{R}^5)}.\end{aligned}$$

By interpolation and the dispersive estimate (3-37) of v, we have

$$\begin{aligned}(3\text{-}64)\quad \|v\|_{L_x^6(\mathbb{R}^5)}\|w_t\|_{L_x^2(\mathbb{R}^5)}\|w\|_{L^3(\mathbb{R}^5)} &\lesssim \frac{1}{t}\|v\|^{\frac{1}{2}}_{L_x^3(\mathbb{R}^5)}\|w_t\|_{L_x^2(\mathbb{R}^5)}\|w\|_{L_x^3(\mathbb{R}^5)} \\ &\lesssim \frac{1}{t}E(w(t))^{\frac{5}{6}}\|v\|^{\frac{1}{2}}_{L_x^3(\mathbb{R}^5)} \\ &\lesssim \frac{1}{t}\big[E(w)(t)+\|v\|^3_{L_x^3(\mathbb{R}^5)}\big],\end{aligned}$$

$$\begin{aligned}(3\text{-}65)\quad \|w_t\|_{L_x^2(\mathbb{R}^5)}\,\|v\|_{L_x^6(\mathbb{R}^5)}\,\|v\|_{L_x^3(\mathbb{R}^5)} &\lesssim \frac{1}{t}E(w(t))^{\frac{1}{2}}\|v\|^{\frac{3}{2}}_{L_x^3(\mathbb{R}^5)} \\ &\lesssim \frac{1}{t}\big[E(w)(t)+\|v\|^3_{L_x^3(\mathbb{R}^5)}\big].\end{aligned}$$

Substituting (3-64) and (3-65) into (3-63), we obtain

$$(3\text{-}66)\qquad \frac{d}{dt}E(w(t)) \le C\frac{1}{t}(E(w(t))+\|v\|^3_{L_x^3(\mathbb{R}^5)}).$$

Hence,

$$(3\text{-}67)\qquad \frac{d}{dt}[t^{-C}(E(w(t)))] \le t^{-C-1}\|v\|^3_{L_x^3(\mathbb{R}^5)}.$$

This estimate and the inequality (3-38) imply that

$$(3\text{-}68)\qquad E(w(t)) \le C_1(1+|t|)^{C_2}.$$

Thus, for any compact interval $J\subset\mathbb{R}$, we have

$$(3\text{-}69)\qquad \|u\|_{L^3_{t,x}(J\times\mathbb{R}^5)} \le \|v\|_{L^3_{t,x}(J\times\mathbb{R}^5)}+\|w\|_{L^3_{t,x}(J\times\mathbb{R}^5)}<\infty. \qquad \square$$

4. Scattering

In this section we prove Theorem 1.1 by assuming Proposition 1.3, that is, removing the dependence of δ_1 in (1-8). From the arguments in Section 3, we have $\delta_1=\delta/(2R)$, where δ and R depending the scaling and spatial profile of the initial data, respectively. We give the heuristic idea of the proof by analyzing the effect of the parameters δ and R on the critical norm $L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)$.

Note that the critical norm $L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)$ of the solution to the nonlinear equation (1-5) is invariant under the scaling transform. Hence the parameter δ may not be the main difficulty in proving Theorem 1.1. On the other hand, the latter parameter R relies on the spatial profile of the initial data. For example, let R be the parameter corresponding to the initial data (u_0, u_1) with compact support. The linear evolution $\vec{S}(t)$ for $t \in \mathbb{R}$ does not change the critical norm $\dot{H}^{1/2} \times \dot{H}^{-1/2}$ by the Plancherel theorem, but owing to the Huygens principle for the odd-dimension linear wave equations, it does change the spatial support of the initial datum. Thus, for the initial data $\vec{S}(t_0)(u_0, u_1)$, the spatial parameter (may be chosen as $R + t_0$) is likely very huge, when t_0 is large enough. However, the $\dot{B}^3_{1,1} \times \dot{B}^2_{1,1}$ norm may become huge under the evolution of $\vec{S}(t)$. Indeed, if $\|(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)} = 1$, then $\|\vec{S}(t_0)(u_0, u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)} \to \infty$ as $t_0 \to \infty$.[1] Hence, if

$$\big\|\vec{S}(t_0)(u_0, u_1)\big\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)} \lesssim 1,$$

then t_0 remains bounded. Taking account of this fact, one may conquer the difficulties caused by the parameter R.

To finish the proof of Theorem 1.1, we need the following theorem of profile decomposition.

4A. ***Profile decomposition.*** Now, we recall the linear profile decomposition from [Ramos 2012] in the radial case. We refer to [Bahouri and Gérard 1999] for the profile decompositions in the energy critical spaces.

Theorem 4.1 (profile decomposition). *Let $C > 0$ be a fixed number and let $(u_0^n, u_1^n)_n$ be a sequence in $\dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$, with*

$$\|(u_0^n, u_1^n)\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le C. \tag{4-1}$$

Then there exist a subsequence of (u_0^n, u_1^n) (still denoted by (u_0^n, u_1^n)), a sequence

$$(\phi_0^j, \phi_1^j)_{j\in\mathbb{N}} \subset \dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5),$$

a sequence

$$(R^N_{0,n}, R^N_{1,n})_{N\ge 1} \subset \dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)$$

and a sequence of parameters $(t_j^n, \lambda_j^n) \subset \mathbb{R}\times(0,\infty)$ such that for each $N \ge 1$,

$$S(t)(u_0^n, u_1^n) = \sum_{j=1}^{N} (\lambda_j^n)^2 \big[S(\lambda_j^n(t - t_j^n))(\phi_0^j, \phi_1^j)\big](\lambda_j^n x) + S(t)(R^N_{0,n}, R^N_{1,n}) \tag{4-2}$$

with

$$\lim_{N\to\infty} \limsup_{n\to\infty} \|S(t)(R^N_{0,n}, R^N_{1,n})\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} = 0. \tag{4-3}$$

[1] By interpolation and density, it suffices to show that, for $f \in \mathcal{S}(\mathbb{R}^5)$, $\lim_{t\to\infty} \|e^{it|\nabla|} f\|_{\dot{B}^{-2}_{\infty,\infty}(\mathbb{R}^5)} = 0$, which follows from the dispersive estimate (2-20) and Bernstein's estimates.

In addition, the parameters (t_j^n, λ_j^n) *satisfy the orthogonality property: for any* $j \neq k$,

$$\lim_{n\to\infty}\left(\frac{\lambda_j^n}{\lambda_k^n}+\frac{\lambda_k^n}{\lambda_j^n}+(\lambda_k^n)^{\frac12}(\lambda_j^n)^{\frac12}|t_j^n-t_k^n|\right)=\infty. \tag{4-4}$$

Furthermore, for every $N \geq 1$,

$$\begin{aligned}\|(u_0^n, u_1^n)\|^2_{\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)} \\ = \sum_{j=1}^N \|(\phi_0^j,\phi_1^j)\|^2_{\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)} + \|(R_{0,n}^N, R_{1,n}^N)\|^2_{\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)} + o_n(1).\end{aligned} \tag{4-5}$$

4B. *End of the proof of the main theorem.* Now, we apply the strategy in [Dodson 2019] to finish the proof of Theorem 1.1, that is, remove the δ_1 in Proposition 1.3.

We prove Theorem 1.1 by contradiction. We assume u is the solution to (1-5) with the initial data such that $(u_0, u_1) \in \dot B^3_{1,1}\times\dot B^2_{1,1}(\mathbb{R}^5)$. For $M \geq 0$, let

$$f(M) = \sup\{\|u\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} : \|(u_0,u_1)\|_{\dot B^3_{1,1}\times\dot B^2_{1,1}(\mathbb{R}^5)} \leq M\}. \tag{4-6}$$

Then by Proposition 1.3, f is well defined. Also, by Bernstein and Theorem 2.10, $f(M) < \infty$ when M is small enough. From the definition, $f(M)$ is nondecreasing as M increases.

If Theorem 1.1 fails, then there exist $M_0 < \infty$ and a sequence $\{(u_0^n, u_1^n)\}_{n\in\mathbb{N}} \subset \dot B^3_{1,1}\times\dot B^2_{1,1}(\mathbb{R}^5)$, such that

$$\|(u_0^n,u_1^n)\|_{\dot B^3_{1,1}\times\dot B^2_{1,1}(\mathbb{R}^5)} \leq M_0, \tag{4-7}$$

and the solution u^n to (1-5) with $(u^n(0), (\partial_t u^n)(0)) = (u_0^n, u_1^n)$ satisfying

$$\|u^n\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \to \infty, \tag{4-8}$$

as $n \to \infty$. By Theorem 4.1, we have

$$S(t)(u_0^n,u_1^n) = \sum_{j=1}^N (\lambda_j^n)^2\big[S(\lambda_j^n(t-t_j^n))(\phi_0^j,\phi_1^j)\big](\lambda_j^n x) + S(t)(R_{0,n}^N, R_{1,n}^N). \tag{4-9}$$

In the proof of Theorem 4.1, Ramos [2012] actually proved that

$$\vec S(t+t_j^n\lambda_j^n)\left[(\lambda_j^n)^{-2}u_0^n\left(\frac{\cdot}{\lambda_j^n}\right), (\lambda_j^n)^{-3}u_1^n\left(\frac{\cdot}{\lambda_j^n}\right)\right]\Bigg|_{t=0} \rightharpoonup (\phi_0^j,\phi_1^j), \tag{4-10}$$

weakly in $\dot H^{1/2}\times\dot H^{-1/2}(\mathbb{R}^5)$ as $n\to\infty$. From this we can prove the following:

Proposition 4.2. *For fixed* $j \in \mathbb{N}$, *if* $(\phi_0^j, \phi_1^j) \neq 0$, *then* $|t_j^n\lambda_j^n|$ *is bounded as* $n \to \infty$.

Proof. First, if $t_j^n\lambda_j^n$ is unbounded for $n\in\mathbb{N}$, then by taking a subsequence of n (still denoted by n), we assume that $|t_j^n\lambda_j^n|\to\infty$, as $n\to\infty$. In light of the heuristic analysis at the beginning of this section, we have

$$\vec{S}(t_j^n\lambda_j^n)\Big[(\lambda_j^n)^{-2}u_0^n\Big(\frac{\cdot}{\lambda_j^n}\Big),\,(\lambda_j^n)^{-3}u_1^n\Big(\frac{\cdot}{\lambda_j^n}\Big)\Big]\to(0,0). \tag{4-11}$$

in $L_x^3\times W_x^{-1,3}(\mathbb{R}^5)$. In fact, using (4-7) and the estimate (3-30) in Section 3, we have

$$\begin{aligned}\Big\|S(t_j^n\lambda_j^n)\Big[(\lambda_j^n)^{-2}u_0^n\Big(\frac{\cdot}{\lambda_j^n}\Big),\,(\lambda_j^n)^{-3}u_1^n\Big(\frac{\cdot}{\lambda_j^n}\Big)\Big]\Big\|_{L_x^3(\mathbb{R}^5)}&\\ \lesssim|t_j^n\lambda_j^n|^{-\frac13}\Big\|\Big((\lambda_j^n)^{-2}u_0^n\Big(\frac{\cdot}{\lambda_j^n}\Big),\,(\lambda_j^n)^{-3}u_1^n\Big(\frac{\cdot}{\lambda_j^n}\Big)\Big)\Big\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}&\\ \lesssim|t_j^n\lambda_j^n|^{-\frac13}\|(u_0^n,u_1^n)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}\to0,&\end{aligned} \tag{4-12}$$

as $n\to\infty$. Similarly, by the dispersive estimate (2-20), Bernstein and interpolation, we have

$$\begin{aligned}\Big\|\dot{S}(t_j^n\lambda_j^n)\Big[(\lambda_j^n)^{-2}u_0^n\Big(\frac{\cdot}{\lambda_j^n}\Big),\,(\lambda_j^n)^{-3}u_1^n\Big(\frac{\cdot}{\lambda_j^n}\Big)\Big]\Big\|_{\dot{W}_x^{-1,3}(\mathbb{R}^5)}&\\ \lesssim|t_j^n\lambda_j^n|^{-\frac13}\|(u_0^n,u_1^n)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}\to0,&\end{aligned} \tag{4-13}$$

as $n\to\infty$. Hence, from the weak convergence relation (4-10), (4-12) and (4-13) imply that $(\phi_0^j,\phi_1^j)=(0,0)$. □

For simplicity, we assume that every (ϕ_0^j,ϕ_1^j) in (4-9) is nontrivial. By Proposition 4.2, $t_j^n\lambda_j^n$ is bounded for each fixed j, and therefore after taking a suitable subsequence of n (still denoted by n), we can assume $t_j^n\lambda_j^n\to t_j\in\mathbb{R}$ as $n\to\infty$. Hence, if we denote $(\varphi_0^j,\varphi_1^j)=\vec{S}(-t_j)(\phi_0^j,\phi_1^j)$, then

$$\vec{S}(-\lambda_j^nt_j^n)(\phi_0^j,\phi_1^j)-(\varphi_0^j,\varphi_1^j)\to0, \tag{4-14}$$

in $\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)$ as $n\to\infty$. Let

$$\begin{cases}\tilde{R}_{0,n}^N=R_{0,n}^N+\sum_{j=1}^N(\lambda_j^n)^2\big[S(-\lambda_j^nt_j^n)(\phi_0^j,\phi_1^j)\big](\lambda_j^nx)-(\lambda_j^n)^2(\varphi_0^j,\varphi_1^j)(\lambda_j^nx),\\ \tilde{R}_{1,n}^N=R_{1,n}^N+\sum_{j=1}^N(\lambda_j^n)^3\big[\dot{S}(-\lambda_j^nt_j^n)(\phi_0^j,\phi_1^j)\big](\lambda_j^nx)-(\lambda_j^n)^3[(\varphi_0^j,\varphi_1^j)](\lambda_j^nx),\end{cases} \tag{4-15}$$

then

$$\lim_{N\to\infty}\limsup_{n\to\infty}\|S(t)(\tilde{R}_{0,n}^N,\tilde{R}_{1,n}^N)\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)}=0. \tag{4-16}$$

Taking $t=0$ in (4-9), by (4-14) and (4-15), we have

$$(u_0^n, u_1^n) = \sum_{j=1}^{N}\big((\lambda_j^n)^2\big(\varphi_0^j(\lambda_j^n x), (\lambda_j^n)^3\varphi_1^j(\lambda_j^n x)\big)\big) + (\tilde{R}_{0,n}^N, \tilde{R}_{1,n}^N). \tag{4-17}$$

In addition, by the orthogonality (4-4) and Proposition 4.2, we have for each $j \neq k$

$$\lim_{n\to\infty} \frac{\lambda_j^n}{\lambda_k^n} + \frac{\lambda_k^n}{\lambda_j^n} = \infty, \tag{4-18}$$

as $n \to \infty$. Thus, for fixed $j \in \mathbb{N}$, we have

$$\left((\lambda_j^n)^{-2}u_0^n\left(\frac{x}{\lambda_j^n}\right), (\lambda_j^n)^{-3}u_1^n\left(\frac{x}{\lambda_j^n}\right)\right) \rightharpoonup (\varphi_0^j, \varphi_1^j) \tag{4-19}$$

weakly in $\dot{H}^{1/2} \times \dot{H}^{-1/2}(\mathbb{R}^5)$ as $n \to \infty$. By Fatou's lemma, this fact and the inequality (4-7) imply

$$\|(\varphi_0^j, \varphi_1^j)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)} \le M_0. \tag{4-20}$$

On the other hand, (4-5) and (4-14) yield that

$$\sum_{j\ge1}\|(\varphi_0^j, \varphi_1^j)\|^2_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \lesssim \sup_{n\ge1}\|(u_0^n, u_1^n)\|^2_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \lesssim C_0^2. \tag{4-21}$$

Hence, for fixed $\epsilon > 0$, there exists a finite integer N_0 such that

$$\sum_{j\ge N_0+1} \big\|(\varphi_0^j, \varphi_1^j)\big\|^2_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)} \le \epsilon. \tag{4-22}$$

By the local well-posedness theory, if $\epsilon > 0$ is small enough, then the solution v^j to (1-5) with the initial data $(\varphi_0^j, \varphi_1^j)$ is globally well-posed and

$$\|v^j\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \lesssim \|(\varphi_0^j, \varphi_1^j)\|_{\dot{H}^{1/2}\times\dot{H}^{-1/2}(\mathbb{R}^5)}, \quad \text{for every } j \ge N_0+1. \tag{4-23}$$

For $1 \le j \le N_0$, as a consequence of Proposition 1.3, the solution to (1-5) with the initial data $(\varphi_0^j, \varphi_1^j)$ is globally well-posed and such that

$$\|v^j\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \lesssim_{M_0,j} 1. \tag{4-24}$$

By the orthogonality property (4-18), for any $j \neq k$,

$$\lim_{n\to\infty}\iint_{\mathbb{R}\times\mathbb{R}^5} |(\lambda_j^n)^2 v^j(\lambda_j^n t, \lambda_j^n x)|^2\, |(\lambda_k^n)^2 v^k(\lambda_k^n t, \lambda_k^n x)|\, dx\, dt = 0. \tag{4-25}$$

This together with the estimates (4-22)–(4-24) implies

$$\sup_{N\ge N_0+1} \lim_{n\to\infty}\Big\|\sum_{1\le j\le N}(\lambda_j^n)^2 v^j(\lambda_j^n t, \lambda_j^n x)\Big\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \tag{4-26}$$

is bounded. Similarly, as a consequence of the trivial estimate

$$\Big|F\Big(\sum_{j=1}^{N}(\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x)\Big)-\sum_{j=1}^{N}F((\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x))\Big| \lesssim \sum_{1\le j,k\le N,\, j\ne k} |(\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x))|\,|(\lambda_k^n)^2 v^k(\lambda_k^n t,\lambda_k^n x)|$$

and the orthogonality property (4-18), we have

$$\lim_{n\to\infty}\Big\|F\Big(\sum_{j=1}^{N}(\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x)\Big)-\sum_{j=1}^{N}F((\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x))\Big\|_{L^{3/2}_{t,x}(\mathbb{R}\times\mathbb{R}^5)}=0. \tag{4-27}$$

Let u_N^n be an approximate solution to (1-5) defined by

$$u_N^n=\sum_{j=1}^{N}(\lambda_j^n)^2 v^j(\lambda_j^n x)+S(t)(\tilde{R}_{0,n}^N,\tilde{R}_{1,n}^N).$$

Then, recall the property (4-16) for $(\tilde{R}_{0,n}^N,\tilde{R}_{1,n}^N)$ and the fact that (4-26) is uniformly bounded for $N\ge N_0+1$, we obtain

$$\limsup_{N\to\infty}\lim_{n\to\infty}\|u_N^n\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)}\lesssim 1. \tag{4-28}$$

Moreover, combining (4-27), the property (4-16) for $(\tilde{R}_{0,n}^N,\tilde{R}_{1,n}^N)$, and Hölder's inequality, we have

$$\limsup_{N\to\infty}\lim_{n\to\infty}\Big\|F(u_N^n)-\sum_{j=1}^{N}F\big((\lambda_j^n)^2 v^j(\lambda_j^n t,\lambda_j^n x)\big)\Big\|_{L^{3/2}_{t,x}(\mathbb{R}\times\mathbb{R}^5)}=0. \tag{4-29}$$

Utilizing Theorem 2.12, by (4-17), (4-28) and (4-29), we have that for n large enough, the solution u^n to (1-5) with initial data (u_0^n,u_1^n) is global and such that

$$\lim_{n\to\infty}\|u^n\|_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \tag{4-30}$$

is bounded, which contradicts the hypothesis (4-8) of u^n. Thus, we have proved Theorem 1.1.

5. Hyperbolic coordinates and spacetime estimates

In this section, we will finish the proof of Proposition 1.3. We first reduce Proposition 1.3 to estimating the $L^3_{t,x}$ norm of w on the region Ω_2, which will be defined below. Without loss of generality, we assume that $\delta_1<\frac{1}{4}$. As in Theorem 3.6, we also note the constants in "$\lesssim$" in this section may be different in each step and are dependent on δ_1 and $\|(u_0,u_1)\|_{\dot{B}^3_{1,1}\times\dot{B}^2_{1,1}(\mathbb{R}^5)}$.

5A. ***Reduction of the proof of Proposition 1.3.*** Now we consider the $L^3_{t,x}$ norm of u on $\mathbb{R}_+ \times \mathbb{R}^5$. First, we split time-spatial region $\mathbb{R}_+ \times \mathbb{R}^5$ as the union

$$\mathbb{R}_+ \times \mathbb{R}^5 = \Omega_1 \cup \Omega_2 \cup \Omega_3, \tag{5-1}$$

where

$$\Omega_1 = \left\{(t,x) \in \mathbb{R}_+ \times \mathbb{R}^5 : |x| \geq t + \tfrac{1}{2}\right\},$$
$$\Omega_2 = \{(t,x) \in \mathbb{R}_+ \times \mathbb{R}^5 : (t + (1-\delta_1))^2 - |x|^2 \geq 1\}.$$

Since $\delta_1 < \frac{1}{4}$, there exists a large constant $C > 0$, such that

$$\Omega_3 \subset \{(t,x) \in \mathbb{R}_+ \times \mathbb{R}^5 : t + |x| \leq C\}.$$

Recalling the estimate (3-23) in Section 3, we obtain $\|u\|_{L^3_{t,x}(\Omega_1)} \lesssim 1$. For the bounded region Ω_3, Theorem 3.6 yields $\|u\|_{L^3_{t,x}(\Omega_3)} \lesssim 1$. Hence, we just need to consider the $L^3_{t,x}$ norm of u on the region Ω_2. By the estimate (3-38) for v, we are reduced to showing $\|w\|_{L^3_{t,x}(\Omega_2)} \lesssim 1$.

5B. ***Hyperbolic coordinates.*** For the radial solution $u(t,x)$ to (1-5), if we denote $u(t,r) = u(t,x)$ for $r = |x|$, then

$$\partial_{tt}(r^2 u) - \partial_{rr}(r^2 u) = -2u - r^2|u|u. \tag{5-2}$$

Denote $\bar{u}(t,r) = u(t-(1-\delta_1), r)$ and denote $\bar{v}, \bar{w}$ similarly. Let $(t,r) = (e^\tau \cosh s, e^\tau \sinh s)$; then $dr\,dt = e^{2\tau} d\tau\, ds$. We denote the hyperbolic transforms by

$$\tilde{u} = \frac{e^{2\tau}\sinh^2 s}{s^2}\bar{u}(e^\tau \cosh s, e^\tau \sinh s), \tag{5-3}$$

$$\tilde{v} = \frac{e^{2\tau}\sinh^2 s}{s^2}\bar{v}(e^\tau \cosh s, e^\tau \sinh s), \tag{5-4}$$

$$\widetilde{w} = \frac{e^{2\tau}\sinh^2 s}{s^2}\overline{w}(e^\tau \cosh s, e^\tau \sinh s). \tag{5-5}$$

Hence, we have

$$\partial_{\tau\tau}(s^2\tilde{u}) - \partial_{ss}(s^2\tilde{u}) = -\frac{2s^2}{\sinh^2 s}\tilde{u} - \frac{s^4}{\sinh^2 s}|\tilde{u}|\tilde{u}, \tag{5-6}$$

$$\partial_{\tau\tau}(s^2\tilde{v}) - \partial_{ss}(s^2\tilde{v}) = -\frac{2s^2}{\sinh^2 s}\tilde{v}, \tag{5-7}$$

$$\partial_{\tau\tau}(s^2\widetilde{w}) - \partial_{ss}(s^2\widetilde{w}) = -\frac{2s^2}{\sinh^2 s}\widetilde{w} - \frac{s^4}{\sinh^2 s}|\tilde{u}|\tilde{u}. \tag{5-8}$$

Define the hyperbolic energy of $\widetilde{w}$ by

$$E_h(\widetilde{w})(\tau) = \int_0^\infty \left[\frac{1}{2}|(s^2\widetilde{w})_\tau|^2 + \frac{1}{2}|(s^2\widetilde{w})_s|^2 + \frac{|s^2\widetilde{w}|^2}{\sinh^2 s} + \frac{1}{3}\frac{|s^2\widetilde{w}|^3}{\sinh^2 s}\right] ds. \tag{5-9}$$

5C. *The hyperbolic energy for some $\tau_0 \geq 0$.* First, we want to prove $E_h(\widetilde{w})(\tau)$ is bounded for some $\tau_0 \geq 0$. We claim that it suffices to show the boundedness of

$$\int_0^\infty \left[|(s^2\widetilde{w})_\tau(\tau_0, s)|^2 + |(s^2\widetilde{w})_s(\tau_0, s)|^2\right] ds \tag{5-10}$$

for some $\tau_0 > 0$.

To prove this claim, we need the sharp Hardy inequality,

$$\left(\frac{d-2}{2}\right)^2 \int_{\mathbb{R}^d} \frac{|f(x)|^2}{|x|^2}\, dx \leq \int_{\mathbb{R}^d} |\nabla f|^2(x)\, dx. \tag{5-11}$$

By polar coordinates, we rewrite this inequality for radial functions,

$$\left(\frac{d-2}{2}\right)^2 \int_0^\infty |f(r)|^2 r^{d-3}\, dr \;\leq\; \int_0^\infty |\partial_r f(r)|^2(r) r^{d-1}\, dr. \tag{5-12}$$

Then, this inequality and integration by parts imply that

$$\begin{aligned} \int_0^\infty &|(s^2\widetilde{w}(\tau_0))_s|^2\, ds \\ &= \int_0^\infty s^4\widetilde{w}_s^2(\tau_0)\, ds + 4\int_0^\infty s^2\widetilde{w}(\tau_0) s\widetilde{w}_s(\tau_0)\, ds + 4\int_0^\infty s^2\widetilde{w}^2(\tau_0)\, ds \\ &= \int_0^\infty s^4\widetilde{w}_s^2(\tau_0)\, ds - 2\int_0^\infty s^2\widetilde{w}(\tau_0)^2\, ds \geq \frac{1}{9}\int_0^\infty s^4\widetilde{w}_s^2(\tau_0)\, ds. \end{aligned} \tag{5-13}$$

In addition, by Hölder and Sobolev in polar coordinates, we have

$$\begin{aligned} \int_0^\infty \frac{|s^2\widetilde{w}(\tau_0)|^3}{\sinh^2 s}\, ds &= \int_0^\infty \frac{s^2}{\sinh^2 s}|\widetilde{w}(\tau_0)|^3 s^4\, ds \\ &\lesssim \left(\int_0^\infty |\widetilde{w}(\tau_0)|^{\frac{10}{3}} s^4\, ds\right)^{\frac{9}{10}} \lesssim \left(\int_0^\infty |\widetilde{w}_s(\tau_0)|^2 s^4\, ds\right)^{\frac{3}{2}}. \end{aligned} \tag{5-14}$$

By Hardy's inequality, we have

$$\int_0^\infty \frac{|s^2\widetilde{w}(\tau_0)|^2}{\sinh^2 s}\, ds \lesssim \int_0^\infty \frac{1}{s^2}|\widetilde{w}(\tau_0)|^2 s^4\, ds \lesssim \int_0^\infty |\widetilde{w}_s(\tau_0)|^2 s^4\, ds. \tag{5-15}$$

Hence the claim follows.

5C1. ***The hyperbolic energy for*** $s > s_0 > 0$. For $\tau \in [0,1]$ and sufficiently large $s_0 > 0$, we can assume that $e^{\tau - s_0} < \frac{1}{2} - \delta_1$. By the finite speed of propagation, $v(t,r)$ are supported in the region $\{(t,r) \in \mathbb{R}\times\mathbb{R}_+ : r - t \lesssim \delta_1/5\}$. Then, for $\tau \in [0,1]$ and $s \ge s_0$, we have

$$e^\tau \sinh s - [e^\tau \cosh s - (1-\delta_1)] = 1 - \delta_1 - e^{\tau - s} > \frac{1}{2} > \frac{\delta_1}{5},$$

which leads to $v(e^\tau \cosh s - (1-\delta_1), e^\tau \sinh s) = 0$. Hence, for $\tau \in [0,1]$, we have

$$\int_{s_0}^{\infty} \tfrac{1}{2}|(s^2\widetilde{w})_\tau(\tau,s)|^2 + \tfrac{1}{2}|(s^2\widetilde{w})_s(\tau,s)|^2\,ds = \int_{s_0}^{\infty} \tfrac{1}{2}|(s^2\tilde{u})_\tau(\tau,s)|^2 + \tfrac{1}{2}|(s^2\tilde{u})_s(\tau,s)|^2\,ds. \tag{5-16}$$

Since u is a radial solution to (1-5), we have, by (2-19),

$$\begin{aligned} r^2 u(t,r) &= \frac{1}{2}[(r-t)^2 u_0(r-t) + (r+t)^2 u_0(r+t)] \\ &\quad - \frac{1}{2} t r^{-1} \int_{r-t}^{r+t} s u_0(s)\,ds + \frac{1}{4r}\int_{r-t}^{r+t} s(s^2+r^2-t^2)u_1(s)\,ds \\ &\quad + \frac{1}{4r}\int_0^t \int_{r-t+s}^{r+t-s} \rho(\rho^2 + r^2 - (t-s)^2)|u|u(s,\rho)\,d\rho\,ds, \end{aligned} \tag{5-17}$$

for $r \ge t \ge 0$. Hence, by the hyperbolic transform (5-3), we have

$$s^2\tilde{u}(\tau,s)$$

$$= \frac{1}{2}\Big[(1-\delta_1-e^{\tau-s})^2 u_0(1-\delta_1-e^{\tau-s}) + (e^{\tau+s}-(1-\delta_1))^2 u_0(e^{\tau+s}-(1-\delta_1))\Big] \tag{5-18}$$

$$-\frac{1}{2}(e^\tau\cosh s-(1-\delta_1))(e^\tau\sinh s)^{-1}\int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_0(\rho)\,d\rho \tag{5-19}$$

$$+\frac{1}{4}\int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho\frac{\rho^2+(e^{\tau+s}-(1-\delta_1))(1-\delta_1-e^{\tau-s})}{e^\tau\sinh s} u_1(\rho)\,d\rho \tag{5-20}$$

$$+\frac{1}{4}\int_{1-\delta_1}^{e^\tau\cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t} \rho\frac{\rho^2+(e^{\tau+s}-t)(t-e^{\tau-s})}{e^\tau \sinh s}|\bar{u}|\bar{u}(t,\rho)\,d\rho\,dt. \tag{5-21}$$

For (5-18), by a direct calculation, we obtain

$$\begin{aligned}(\partial_\tau+\partial_s)\text{(5-18)} &= 2(e^{\tau+s}-(1-\delta_1))e^{\tau+s}u_0(e^{\tau+s}-(1-\delta_1)) \\ &\quad + (e^{\tau+s}-(1-\delta_1))^2 u_0'(e^{\tau+s}-(1-\delta_1))e^{\tau+s}, \end{aligned} \tag{5-22}$$

$$\begin{aligned}(\partial_\tau-\partial_s)\text{(5-18)} &= 2(1-\delta_1-e^{\tau-s})e^{\tau-s}u_0(1-\delta_1-e^{\tau-s}) \\ &\quad + ((1-\delta_1-e^{\tau-s})^2 u_0'(1-\delta_1-e^{\tau-s})e^{\tau-s}. \end{aligned} \tag{5-23}$$

Using the estimate (2-13) in Section 2 and polar coordinates, we deduce that

$$\int_{s_0}\big|(e^{\tau+s}-(1-\delta_1))e^{\tau+s}u_0(e^{\tau+s}-(1-\delta_1))\big|^2 ds \lesssim \int_0^\infty |u_0(r)|^2 r^3\,dr \lesssim 1, \tag{5-24}$$

$$\int_{s_0}\big|(e^{\tau+s}-(1-\delta_1))^2 e^{\tau+s}u_0'(e^{\tau+s}-(1-\delta_1))\big|^2 ds \lesssim \int_0^\infty |\partial_r u_0(r)|^2 r^5\,dr \lesssim 1. \tag{5-25}$$

By the radial Sobolev inequality (2-2), we have $|u_0(r)|\lesssim r^{-2}$. This estimate and the inequality (2-3) imply

$$\begin{aligned}\int_{s_0}^\infty \big|(1-\delta_1-e^{\tau-s})e^{\tau-s}u_0(1-\delta_1-e^{\tau-s})\big|^2\,ds \\ +\int_{s_0}^\infty \big|((1-\delta_1-e^{\tau-s})^2 u_0'(1-\delta_1-e^{\tau-s})e^{\tau-s}\big|^2\,ds \\ \lesssim \int_{s_0}^\infty e^{-2s}\,ds \lesssim 1.\end{aligned} \tag{5-26}$$

We now take the derivatives of (5-19) with respect to τ and s,

$$\partial_\tau(5\text{-}19) = \frac{1-\delta_1}{2e^\tau \sinh s}\int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_0(\rho)\,d\rho + I_1 + I_2 \tag{5-27}$$

$$\partial_s(5\text{-}19) = \partial_s\left(\frac{e^\tau \cosh s-(1-\delta_1)}{2e^\tau \sinh s}\right)\int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_0(\rho)\,d\rho + I_1 - I_2, \tag{5-28}$$

where

$$I_1 = \frac{e^\tau \cosh s-(1-\delta_1)}{2e^\tau \sinh s}e^{\tau+s}(e^{\tau+s}-(1-\delta_1))u_0(e^{\tau+s}-(1-\delta_1)), \tag{5-29}$$

$$I_2 = \frac{e^\tau \cosh s-(1-\delta_1)}{2e^\tau \sinh s}e^{\tau-s}(1-\delta_1-e^{\tau-s})u_0(1-\delta_1-e^{\tau-s}). \tag{5-30}$$

For the first term in the right-hand side of (5-27), by the inequality (2-3), we have

$$\int_{s_0}^\infty \left|e^{-s}\int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_0(\rho)\,d\rho\right|^2 ds \lesssim \int_{s_0}^\infty e^{-2s}\,ds \lesssim 1. \tag{5-31}$$

By similar estimates, one can find that the contribution of the first term in the right-hand side of (5-28) to (5-16) is finite. For I_1, a change of variables and the inequality (2-13) yield

$$\int_{s_0}^\infty |I_1|^2\,ds \lesssim \int_{s_0}^\infty |e^{2s}u_0(e^{\tau+s}-(1-\delta_1))|^2\,ds \lesssim \int_{\frac{1}{2}e^{s_0}}^\infty \rho^3|u_0(\rho)|^2\,d\rho \lesssim 1. \tag{5-32}$$

For I_2, by (2-2), we obtain

$$\int_{s_0}^{\infty} |I_2|\, ds \lesssim \int_{s_0}^{\infty} |e^{-s}|^2\, ds \lesssim 1. \tag{5-33}$$

Next, we consider the contribution of (5-20) to (5-16). For simplicity, we consider

$$\tfrac{1}{2}(\partial_\tau - \partial_s)\text{(5-20)}$$
$$= e^{\tau-s}(1-\delta_1-e^{\tau-s})^2 u_1(1-\delta_1-e^{\tau-s}) \tag{5-34}$$
$$+ \frac{e^{\tau-s}(e^{\tau+s}-(1-\delta_1))^2}{(e^{\tau+s}-e^{\tau-s})^2} \int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_1(\rho)\, d\rho \tag{5-35}$$
$$+ \frac{e^{\tau-s}}{(e^{\tau+s}-e^{\tau-s})^2} \int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho^3 u_1(\rho)\, d\rho, \tag{5-36}$$

and

$$\tfrac{1}{2}(\partial_\tau + \partial_s)\text{(5-20)}$$
$$= e^{\tau+s}(e^{\tau+s}-(1-\delta_1))^2 u_1(e^{\tau+s}-(1-\delta_1)) \tag{5-37}$$
$$+ \frac{e^{\tau+s}(1-\delta_1-e^{\tau-s})^2}{(e^{\tau+s}-e^{\tau-s})^2} \int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho u_1(\rho)\, d\rho \tag{5-38}$$
$$+ \frac{e^{\tau+s}}{(e^{\tau+s}-e^{\tau-s})^2} \int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho^3 u_1(\rho)\, d\rho. \tag{5-39}$$

Using the estimates (2-4) and (2-13), we can easily estimate the contributions of (5-34)–(5-38) to (5-16). Let $\mathbb{1}_J(y)$ be the characteristic function of an interval $J \subset \mathbb{R}$. For (5-39), by the inequality (2-13) and a change of variables, we see that

$$\begin{aligned}
\int_{s_0}^{\infty} \left| e^{-s} \int_{1-\delta_1-e^{\tau-s}}^{e^{\tau+s}-(1-\delta_1)} \rho^3 u_1(\rho)\, d\rho \right|^2 ds &\lesssim \int_{s_0}^{\infty} \left| e^{-s} \int_0^{2e^s} \rho^3 |u_1(\rho)|\, d\rho \right|^2 ds \\
&\lesssim \int_0^{\infty} \left| \frac{1}{\eta} \int_0^{2\eta} \rho^3 |u_1(\rho)|\, d\rho \right|^2 \frac{1}{\eta}\, d\eta \\
&\lesssim \int_0^{\infty} \left| \frac{1}{\eta} \int_0^{2\eta} \rho^{\frac{5}{2}} |u_1(\rho)|\, d\rho \right|^2 d\eta \\
&\lesssim \int_0^{\infty} \left| \mathcal{M}\big(\mathbb{1}_{[0,\infty)}(\rho) \rho^{\frac{5}{2}} u_1(\rho)\big) \right|^2 (\eta)\, d\eta \\
&\lesssim \int_0^{\infty} |u_1(\rho)|^2 \rho^5\, d\rho \lesssim 1,
\end{aligned} \tag{5-40}$$

where $\mathcal{M}$ is the Hardy–Littlewood maximal function and we used the fact that $\mathcal{M}$ is bounded in L^2.

Next, we consider the contribution of (5-21) to the energy (5-16). Also, for simplicity, we consider

$$
\begin{aligned}
&(\partial_\tau+\partial_s)(5\text{-}21)\\
&= e^{\tau+s}\int_{1-\delta_1}^{e^\tau\cosh s}(e^{\tau+s}-t)^2(|\bar u|\bar u)(t,e^{\tau+s}-t)\,dt && (5\text{-}41)\\
&\quad+\frac{e^{\tau+s}}{(e^{\tau+s}-e^{\tau-s})^2}\int_{1-\delta_1}^{e^\tau\cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t}\rho(t-e^{\tau-s})^2(|\bar u|\bar u)(t,\rho)\,d\rho\,dt && (5\text{-}42)\\
&\quad-\frac{e^{\tau+s}}{(e^{\tau+s}-e^{\tau-s})^2}\int_{1-\delta_1}^{e^\tau\cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t}\rho^3(|\bar u|\bar u)(t,\rho)\,d\rho\,dt, && (5\text{-}43)
\end{aligned}
$$

and

$$
\begin{aligned}
&(\partial_\tau-\partial_s)(5\text{-}21)\\
&= e^{\tau-s}\int_{1-\delta_1}^{e^\tau\cosh s}(t-e^{\tau-s})^2|\bar u|\bar u(t,t-e^{\tau-s})\,dt && (5\text{-}44)\\
&\quad-\frac{e^{\tau-s}}{(e^{\tau+s}-e^{\tau-s})^2}\int_{1-\delta_1}^{e^\tau\cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t}\rho(e^{\tau+s}-t)^2|\bar u|\bar u(t,\rho)\,d\rho\,dt && (5\text{-}45)\\
&\quad+\frac{e^{\tau-s}}{(e^{\tau+s}-e^{\tau-s})^2}\int_{1-\delta_1}^{e^\tau\cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t}\rho^3|\bar u|\bar u(t,\rho)\,d\rho dt. && (5\text{-}46)
\end{aligned}
$$

By the definition of $\bar u$, the inequality (3-23) and Hölder, the contribution of (5-41) can be estimated as

$$
\begin{aligned}
&\int_{s_0}^\infty\Big|e^{\tau+s}\int_{1-\delta_1}^{e^\tau\cosh s}(e^{\tau+s}-t)^2|\bar u|^2(t,e^{\tau+s}-t)\,dt\Big|^2\,ds && (5\text{-}47)\\
&\lesssim\int_{s_0}^\infty\int_{1-\delta_1}^{e^\tau\cosh s}|\bar u|^4(t,e^{\tau+s}-t)e^{7s}\,dt\,ds\\
&\lesssim\int_{\frac14 e^{s_0}}^\infty\int_{1-\delta_1}^{\rho}|\bar u|^4(t,\rho)\rho^6\,dt\,d\rho\lesssim\int_0^\infty\int_{\rho>t+\frac12}|u|^4(t,\rho)\rho^6\,d\rho\,dt\\
&\lesssim\|u\|^3_{L^3(\{|x|>|t|+\frac12\})}\sup_{t\ge0}\big\||x|^2u(t,x)\big\|_{L^\infty_x(\{|x|>t+\frac12\})}\lesssim 1.
\end{aligned}
$$

Now, we consider (5-42) and (5-43). By Hölder, a change of variables, and the inequality (3-23), we have

$$
\begin{aligned}
&\int_{s_0}^\infty\Big|\int_{1-\delta_1}^{e^\tau\cosh s}e^{-s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t}(t^2\rho+\rho^3)|\bar u|^2(t,\rho)\,d\rho\,dt\Big|^2\,ds && (5\text{-}48)\\
&\lesssim\int_{s_0}^\infty\Big|\int_{1-\delta_1}^{e^\tau\cosh s}\mathcal{M}\big(\mathbb{1}_{[t-\frac12+\delta_1,\infty)}(\rho)\rho^3|\bar u|^2(t,\rho)\big)(e^\tau\sinh s)\,dt\Big|^2\,ds
\end{aligned}
$$

$$\lesssim \int_{s_0}^{\infty}\int_{1-\delta_1}^{e^\tau \cosh s} \big[\mathcal{M}\big(\mathbb{1}_{[t-\frac{1}{2}+\delta_1,\infty)}(\rho)\rho^3|\bar{u}|^2(t,\rho)\big)(e^\tau \sinh s)\big]^2 e^\tau \cosh s\, dt\, ds$$

$$\lesssim \int_{1-\delta_1}^{\infty}\int_{e^{s_0}}^{\infty} \big[\mathcal{M}\big(\mathbb{1}_{[t-\frac{1}{2}+\delta_1,\infty)}(\rho)\rho^3|\bar{u}|^2(t,\rho)\big)(r)\big]^2\, dr\, dt$$

$$\lesssim \int_{1-\delta_1}^{\infty}\int_{r>t-\frac{1}{2}+\delta_1} r^6|\bar{u}|^4(t,r)\, dr\, dt$$

$$\lesssim \int_{0}^{\infty}\int_{\rho>t+\frac{1}{2}} |u|^4(t,\rho)\rho^6\, d\rho\, dt$$

$$\lesssim \|u\|^3_{L^3(\{|x|>|t|+\frac{1}{2}\})} \sup_{t\ge 0}\big\||x|^2 u(t,x)\big\|_{L^\infty_x(\{|x|>t+\frac{1}{2}\})}$$

$$\lesssim 1.$$

Thus, the contribution of (5-42) and (5-43) to (5-16) is finite.

For (5-44), by the fact that $e^{\tau-s_0} < \frac{1}{2}-\delta_1$, the definition of $\bar{u}$, and the inequality (3-23), we have

$$\int_{s_0}^{\infty} e^{-2s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s} (t-e^{\tau-s})^2|\bar{u}|^2(t,t-e^{\tau-s})\, dt\Big|^2\, ds \lesssim \int_{s_0}^{\infty} e^{-2s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s} t^{-2}\, dt\Big|^2 \lesssim 1. \tag{5-49}$$

Similarly, for (5-45) and (5-46), by the fact that $e^{\tau-s_0} < \frac{1}{2}-\delta_1$ and the inequality (3-23), we can obtain that

$$\begin{aligned}\int_{s_0}^{\infty} e^{-6s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t} [\rho^3+\rho(e^{\tau+s}-t)^2]|\bar{u}|^2(t,\rho)\, d\rho\, dt\Big|^2 ds &\lesssim \int_{s_0}^{\infty} e^{-2s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t} \rho|\bar{u}|^2(t,\rho)\, d\rho\, dt\Big|^2 ds\\ &\lesssim \int_{s_0}^{\infty} e^{-2s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s}\int_{t-e^{\tau-s}}^{e^{\tau+s}-t} \rho^{-3}\, d\rho\, dt\Big|^2 ds\\ &\lesssim \int_{s_0}^{\infty} e^{-2s}\Big|\int_{1-\delta_1}^{e^\tau \cosh s} t^{-2}\, dt\Big|^2 ds\\ &\lesssim 1.\end{aligned} \tag{5-50}$$

Hence, combining (5-24)–(5-50), we have

$$\int_{s_0}^{\infty} \tfrac{1}{2}|(s^2\widetilde{w})_\tau|^2 + \tfrac{1}{2}|(s^2\widetilde{w})_s|^2\, ds \lesssim 1. \tag{5-51}$$

5C2. ***The hyperbolic energy for*** $0 \le s \le s_0$. By the hyperbolic transform (5-5), we have

$$(s^2\widetilde{w})_\tau(\tau,s) = 2e^{2\tau}\sinh^2 s\,\overline{w}(e^\tau\cosh s, e^\tau\sinh s) + e^{3\tau}\sinh^2 s\cosh s\,\overline{w}_t(e^\tau\cosh s, e^\tau\sinh s) + e^{3\tau}\sinh^3 s\,\overline{w}_r(e^\tau\cosh s, e^\tau\sinh s), \tag{5-52}$$

$$(s^2\widetilde{w})_s(\tau,s) = 2e^{2\tau}\sinh s\cosh s\,\overline{w}(e^\tau\cosh s, e^\tau\sinh s) + e^{3\tau}\sinh^3 s\,\overline{w}_t(e^\tau\cosh s, e^\tau\sinh s) + e^{3\tau}\sinh^2 s\cosh s\,\overline{w}_r(e^\tau\cosh s, e^\tau\sinh s). \tag{5-53}$$

Hence

$$\int_0^1\int_0^{s_0} |(s^2\widetilde{w})_\tau|^2 + |(s^2\widetilde{w})_s|^2\,ds\,d\tau$$

$$\lesssim \int_0^1\int_0^{s_0} e^{4\tau}\sinh^2 s(\sinh^2 s+\cosh^2 s)|\overline{w}|^2(e^\tau\cosh s, e^\tau\sinh s)\,ds\,d\tau \tag{5-54}$$

$$+\int_0^1\int_0^{s_0} e^{6\tau}\sinh^4 s(\sinh^2 s+\cosh^2 s)[\overline{w}_t^2+\overline{w}_r^2](e^\tau\cosh s, e^\tau\sinh s)\,ds\,d\tau. \tag{5-55}$$

Taking $C_0 = e^{1+s_0}$, by a change of variables, the Hardy inequality and the inequality (3-68), we obtain

$$\text{(5-54)} \lesssim \iint_{|x|+|t|\le C_0} \frac{1}{|x|^2}|\overline{w}|^2(t,x)\,dx\,dt \lesssim \sup_{0<t<C_0} \|\nabla_x\overline{w}\|_{L_x^2(\mathbb{R}^5)} \lesssim 1. \tag{5-56}$$

Similarly, for (5-55), by a change of variables, we have

$$\text{(5-55)} \lesssim \iint_{|x|+|t|\le C_0} |\nabla_{t,x}\overline{w}|^2(t,x)\,dx\,dt \lesssim \sup_{0<t<C_0} \|\nabla_{t,x}\overline{w}\|_{L_x^2(\mathbb{R}^5)} \lesssim 1. \tag{5-57}$$

Then, by the mean value theorem, there exists $\tau_0 \in [0, 1]$, such that

$$\int_0^{s_0} |(s^2\widetilde{w})_\tau|^2(\tau_0,s) + |(s^2\widetilde{w})_s|^2(\tau_0,s)\,ds \lesssim 1. \tag{5-58}$$

This estimate along with (5-51) implies

$$\int_0^{\infty} |(s^2\widetilde{w})_\tau|^2(\tau_0,s) + |(s^2\widetilde{w})_s|^2(\tau_0,s)\,ds \lesssim 1. \tag{5-59}$$

5D. ***Uniform boundedness of the hyperbolic energy of*** $\widetilde{w}$. We are now going to show that $E_h(\widetilde{w})(\tau)$ is uniformly bounded for $\tau \in \mathbb{R}_+$.

A simple calculation gives

$$\frac{d}{d\tau}E_h(\widetilde{w}(\tau)) = \int \frac{|s^2\widetilde{w}|s^2\widetilde{w} - |s^2\tilde{u}|s^2\tilde{u}}{\sinh^2 s} s^2\widetilde{w}_\tau\, ds. \tag{5-60}$$

Utilizing the decay property (3-37) of v, we have, for $\tau, s \geq 0$,

$$(e^\tau \cosh s - (1-\delta_1))^2 v(e^\tau \cosh s - (1-\delta_1), e^\tau \sinh s) \lesssim \delta_1^{-\frac{1}{2}}. \tag{5-61}$$

The Huygens principle implies that $v(e^\tau \cosh s - (1-\delta_1), e^\tau \sinh s) = 0$ unless $1 - \frac{6}{5}\delta_1 \leq e^{\tau - s} \leq 1 - \frac{4}{5}\delta_1$. Thus, for $\tau, s \geq 0$, we have

$$\begin{aligned}\frac{s^2|\tilde{v}(\tau,s)|}{\sinh^2 s} &= e^{2\tau}\left|v(e^\tau \cosh s - (1-\delta_1), e^\tau \sinh s)\right| \\ &\lesssim \frac{e^{2\tau}\mathbb{1}_{\{s\geq 0:\ e^{\tau-s}\leq 1-\frac{4}{5}\delta_1\}}}{(e^\tau \cosh s - (1-\delta_1))^2} \lesssim e^{-\tau}.\end{aligned} \tag{5-62}$$

Hence, by Hölder and interpolation, we have

$$\begin{aligned}&\int_0^\infty \frac{1}{\sinh^2 s}|s^2\widetilde{w}_\tau||s^2\tilde{v}|^2\, ds \\ &\lesssim \left(\int_0^\infty |s^2\widetilde{w}_\tau|^2\, ds\right)^{\frac{1}{2}} \left(\int_0^\infty \frac{1}{\sinh^2 s}|s^2\tilde{v}|^3\, ds\right)^{\frac{1}{2}} \left\|\frac{s^2}{\sinh^2 s}\tilde{v}(\tau,s)\right\|_{L_s^\infty}^{\frac{1}{2}} \\ &\lesssim e^{-\tau/2}E_h(\widetilde{w}(\tau))^{\frac{1}{2}}\left(\int_0^\infty \frac{1}{\sinh^2 s}|s^2\tilde{v}|^3\, ds\right)^{\frac{1}{2}},\end{aligned} \tag{5-63}$$

$$\begin{aligned}&\int_0^\infty \frac{1}{\sinh^2 s}|s^2\widetilde{w}_\tau||s^2\tilde{v}||s^2\widetilde{w}|\, ds \\ &\lesssim \left(\int_0^\infty |s^2\widetilde{w}_\tau|^2\, ds\right)^{\frac{1}{2}} \left(\int_0^\infty \frac{1}{\sinh^2 s}|s^2\widetilde{w}|^3\, ds\right)^{\frac{1}{3}} \\ &\qquad\qquad \times \left(\int_0^\infty |s^2\tilde{v}(\tau,s)|^6 \frac{1}{\sinh^8 s}\, ds\right)^{\frac{1}{6}} \\ &\lesssim e^{-\tau/2}E_h(\widetilde{w}(\tau))^{\frac{5}{6}}\left(\int_0^\infty \frac{1}{\sinh^2 s}|s^2\tilde{v}|^3\, ds\right)^{\frac{1}{6}}.\end{aligned} \tag{5-64}$$

Combining (5-63) with (5-64) and employing Hölder again, we have

$$\frac{d}{d\tau}E_h(\widetilde{w}(\tau) \lesssim e^{-\tau/2}\left[E_h(\widetilde{w}(\tau) + \int_0^\infty \frac{1}{\sinh^2 s}|s^2\tilde{v}|^3\, ds\right]. \tag{5-65}$$

On the other hand, by a change of variables, we have

$$\begin{aligned}(5\text{-}66)\qquad \int_0^\infty\int_0^\infty \frac{1}{\sinh^2 s}|s^2\tilde{v}|^3\,ds\,d\tau &\le \int_{\delta_1}^\infty\int_0^\infty |v(t,r)|^3 r^4\,dr\,dt\\ &\le \|v\|^3_{L^3_{t,x}([\delta_1,\infty)\times\mathbb{R}^5)}\lesssim 1.\end{aligned}$$

Hence, by Gronwall's inequality, (5-65) and (5-66) yield that $E_h(\widetilde{w})(\tau)$ is uniformly bounded in $\mathbb{R}_+$.

5E. *Conclusion of the proof of Proposition 1.3.* We complete the proof by studying the Morawetz action in hyperbolic coordinates.

Proposition 5.1. *Let $\widetilde{w}$ be defined in (5-5), then*

$$(5\text{-}67)\qquad \iint_{\Omega_2}|w(t,r)|^3r^4\,dt\,dr=\int_0^\infty\int_0^\infty\frac{|s^2\widetilde{w}|^3}{\sinh^2 s}\,ds\,d\tau\lesssim 1.$$

Proof. Define the Morawetz action by

$$(5\text{-}68)\qquad M(\tau)=\int_0^\infty (s^2\widetilde{w})_\tau(s^2\widetilde{w})_s\,ds=\int_0^\infty \widetilde{w}_\tau\Big(\widetilde{w}_s+\frac{2}{s}\widetilde{w}\Big)s^4\,ds.$$

One can easily find that $|M(\tau)|\le E_h(\widetilde{w})(\tau)$. By (5-8), we have

$$\begin{aligned}(5\text{-}69)\qquad \frac{d}{d\tau}M(\tau)&=-\int_0^\infty\frac{2s^2\widetilde{w}}{\sinh^2 s}(s^2\widetilde{w})_s\ ds-\int_0^\infty\frac{s^4}{\sinh^2 s}|\tilde{u}|\tilde{u}s^2\widetilde{w}_s\ ds\\ &=-2\int_0^\infty\frac{|s^2\widetilde{w}|^2}{\sinh^2 s}\frac{\cosh s}{\sinh s}\,ds-\frac{2}{3}\int_0^\infty\frac{|s^2\widetilde{w}|^3}{\sinh^2 s}\frac{\cosh s}{\sinh s}\,ds\\ &\quad+\int_0^\infty\frac{s^4}{\sinh^2 s}(|\widetilde{w}|\widetilde{w}-|\tilde{u}|\tilde{u})s^2\widetilde{w}_s\ ds.\end{aligned}$$

By Hölder, the estimate (5-62), and the fact that $E_h(\widetilde{w}(\tau))$ is uniformly bounded for $\tau\ge 0$, we have

$$\begin{aligned}(5\text{-}70)\qquad &\Big|\int_0^\infty\int_0^\infty\frac{s^4}{\sinh^2 s}(|\widetilde{w}|\widetilde{w}-|\tilde{u}|\tilde{u})s^2\widetilde{w}_s\ ds\,d\tau\Big|\\ &\lesssim\int_0^\infty\int_0^\infty\frac{1}{\sinh^2 s}|s^2\widetilde{w}_\tau||s^2\tilde{v}|^2\,ds\,d\tau\\ &\quad+\int_0^\infty\int_0^\infty\frac{1}{\sinh^2 s}|s^2\widetilde{w}_\tau||s^2\tilde{v}||s^2\widetilde{w}|\,ds\,d\tau\end{aligned}$$

$$
\begin{aligned}
&\lesssim \int_0^\infty e^{-\tau} E_h(\widetilde{w}(\tau))^{\frac{1}{2}} \left(\int_0^\infty \frac{1}{\sinh^2 s} |s^2 \tilde{v}|^3 \, ds \right)^{\frac{1}{2}} \tau \\
&\qquad + \int_0^\infty e^{-\tau} E_h(\widetilde{w}(\tau))^{\frac{5}{6}} \left(\int_0^\infty \frac{1}{\sinh^2 s} |s^2 \tilde{v}|^3 \, ds \right)^{\frac{1}{6}} d\tau \\
&\lesssim \|v\|^3_{L^3_{t,x}(\mathbb{R}\times\mathbb{R}^5)} \lesssim 1.
\end{aligned}
$$

This together with the equality (5-69) and the fact $M(\tau)$ is uniformly bounded for $\tau \geq 0$, implies that

$$
\int_0^\infty \int_0^\infty \frac{|s^2 \widetilde{w}|^3}{\sinh^2 s} \frac{\cosh s}{\sinh s} \, ds \, d\tau \lesssim 1. \tag{5-71}
$$

Thus, we have

$$
\int_0^\infty \int_0^\infty \frac{|s^2 \widetilde{w}|^3}{\sinh^2 s} \, ds \, d\tau \lesssim 1. \tag{5-72}
$$

This yields (5-67) by the definition of $\widetilde{w}$. □

Acknowledgments

Miao is supported in part by the National Natural Science Foundation of China under grant No. 11831004, No. 11926303 and No. 11671047. Yang is supported by the NNSF of China under grant No. 11901032 and the Labex MME-DII. Zhao is supported in part by the Chinese Postdoc Foundation Grant 2019M650457 and National Natural Science Foundation of China (NSAF-U1930402).

References

[Bahouri and Gérard 1999] H. Bahouri and P. Gérard, "High frequency approximation of solutions to critical nonlinear wave equations", *Amer. J. Math.* **121**:1 (1999), 131–175. MR Zbl

[Bahouri and Shatah 1998] H. Bahouri and J. Shatah, "Decay estimates for the critical semilinear wave equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **15**:6 (1998), 783–789. MR Zbl

[Christ and Weinstein 1991] F. M. Christ and M. I. Weinstein, "Dispersion of small amplitude solutions of the generalized Korteweg–de Vries equation", *J. Funct. Anal.* **100**:1 (1991), 87–109. MR Zbl

[Colzani et al. 2002] L. Colzani, A. Cominardi, and K. Stempak, "Radial solutions to the wave equation", *Ann. Mat. Pura Appl.* (4) **181**:1 (2002), 25–54. MR Zbl

[Dodson 2019] B. Dodson, "Global well-posedness and scattering for the radial, defocusing, cubic wave equation with initial data in a critical Besov space", *Anal. PDE* **12**:4 (2019), 1023–1048. MR Zbl

[Ginibre and Velo 1985] J. Ginibre and G. Velo, "The global Cauchy problem for the nonlinear Klein–Gordon equation", *Math. Z.* **189**:4 (1985), 487–505. MR Zbl

[Ginibre and Velo 1989] J. Ginibre and G. Velo, "The global Cauchy problem for the nonlinear Klein–Gordon equation, II", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**:1 (1989), 15–35. MR Zbl

[Ginibre and Velo 1995] J. Ginibre and G. Velo, "Generalized Strichartz inequalities for the wave equation", pp. 153–160 in *Partial differential operators and mathematical physics* (Holzhau, Germany, 1994), edited by M. Demuth and B.-W. Schulze, Oper. Theory Adv. Appl. **78**, Birkhäuser, Basel, 1995. MR Zbl

[Grillakis 1990] M. G. Grillakis, "Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity", *Ann. of Math.* (2) **132**:3 (1990), 485–509. MR Zbl

[Grillakis 1992] M. G. Grillakis, "Regularity for the wave equation with a critical nonlinearity", *Comm. Pure Appl. Math.* **45**:6 (1992), 749–774. MR Zbl

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR Zbl

[Kenig 2015] C. E. Kenig, *Lectures on the energy critical nonlinear wave equation*, CBMS Regional Conference Series in Mathematics **122**, American Mathematical Society, Providence, RI, 2015. MR Zbl

[Kenig and Merle 2006] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case", *Invent. Math.* **166**:3 (2006), 645–675. MR Zbl

[Kenig and Merle 2008] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation", *Acta Math.* **201**:2 (2008), 147–212. MR Zbl

[Killip and Vişan 2013] R. Killip and M. Vişan, "Nonlinear Schrödinger equations at critical regularity", pp. 325–437 in *Evolution equations*, edited by D. Ellwood et al., Clay Math. Proc. **17**, American Mathematical Society, Providence, RI, 2013. MR Zbl

[Killip and Visan 2011] R. Killip and M. Visan, "The defocusing energy-supercritical nonlinear wave equation in three space dimensions", *Trans. Amer. Math. Soc.* **363**:7 (2011), 3893–3934. MR Zbl

[Koch et al. 2014] H. Koch, D. Tataru, and M. Vişan, *Dispersive equations and nonlinear waves: generalized Korteweg–de Vries, nonlinear Schrödinger, wave and Schrödinger maps*, Oberwolfach Seminars **45**, Springer, 2014. MR Zbl

[Lindblad and Sogge 1995] H. Lindblad and C. D. Sogge, "On existence and scattering with minimal regularity for semilinear wave equations", *J. Funct. Anal.* **130**:2 (1995), 357–426. MR Zbl

[Lindblad and Sogge 1996] H. Lindblad and C. D. Sogge, "Long-time existence for small amplitude semilinear wave equations", *Amer. J. Math.* **118**:5 (1996), 1047–1135. MR Zbl

[Morawetz 1968] C. S. Morawetz, "Time decay for the nonlinear Klein–Gordon equations", *Proc. Roy. Soc. London Ser. A* **306** (1968), 291–296. MR Zbl

[Nakanishi 1999] K. Nakanishi, "Unique global existence and asymptotic behaviour of solutions for wave equations with non-coercive critical nonlinearity", *Comm. Partial Differential Equations* **24**:1-2 (1999), 185–221. MR Zbl

[Rammaha 1987] M. A. Rammaha, "Finite-time blow-up for nonlinear wave equations in high dimensions", *Comm. Partial Differential Equations* **12**:6 (1987), 677–700. MR Zbl

[Ramos 2012] J. Ramos, "A refinement of the Strichartz inequality for the wave equation with applications", *Adv. Math.* **230**:2 (2012), 649–698. MR Zbl

[Rodriguez 2017] C. Rodriguez, "Scattering for radial energy-subcritical wave equations in dimensions 4 and 5", *Comm. Partial Differential Equations* **42**:6 (2017), 852–894. MR Zbl

[Shatah and Struwe 1993] J. Shatah and M. Struwe, "Regularity results for nonlinear wave equations", *Ann. of Math.* (2) **138**:3 (1993), 503–518. MR Zbl

[Sogge 1995] C. D. Sogge, *Lectures on nonlinear wave equations*, Monographs in Analysis **II**, International Press, Boston, 1995. MR Zbl

[Taggart 2010] R. J. Taggart, "Inhomogeneous Strichartz estimates", *Forum Math.* **22**:5 (2010), 825–853. MR Zbl

Received June 25, 2018.

CHANGXING MIAO
INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS
BEIJING
CHINA
miao_changxing@iapcm.ac.cn

JIANWEI YANG
BEIJING INSTITUTE OF TECHNOLOGY
BEIJING
CHINA
jw-urbain.yang@bit.edu.cn

TENGFEI ZHAO
BEIJING COMPUTATIONAL SCIENCE RESEARCH CENTER
BEIJING
CHINA
zhao_tengfei@csrc.ac.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.291

LIOUVILLE-TYPE THEOREMS FOR WEIGHTED p-HARMONIC 1-FORMS AND WEIGHTED p-HARMONIC MAPS

KEOMKYO SEO AND GABJIN YUN

In this paper, we obtain Bochner–Weitzenböck formulas for the weighted Hodge Laplacian operator acting on differential forms and more generally on vector bundle-valued weighted p-harmonic forms. Applying these formulas, we prove Liouville-type theorems for weighted L^q p-harmonic 1-forms and for weighted p-harmonic maps in a weighted complete noncompact manifold with nonnegative Bakry–Émery Ricci curvature, where $q = 2p - 2$ or $q = p$.

1. Introduction

The celebrated Liouville theorem states that every positive harmonic function on $\mathbb{R}^n$ is constant. There have been a lot of effort over the years to generalize the classical Liouville theorem into complete noncompact Riemannian manifolds. Huber [1957] proved that any negative subharmonic function on a complete surface with nonnegative curvature is constant. Yau [1975] proved that any positive harmonic function on a noncompact Riemannian manifold with nonnegative Ricci curvature is constant. See also [Greene and Wu 1979; Hildebrandt 1982; Karp 1982] for further related results. Moreover, Yau [1976] obtained an L^p-Liouville type theorem. More precisely, he proved that, for $1 < p < \infty$, any L^p harmonic function on a complete Riemannian manifold is constant. Given a harmonic function f on a Riemannian manifold M, we note that the differential df is obviously a harmonic 1-form on M. In the case where M is a complete noncompact Riemannian manifold, it is natural to consider L^2 harmonic forms on M because L^2-Hodge theory remains valid in complete noncompact manifolds as classical Hodge theory works well in compact manifolds. It turned out that the theory of L^2 harmonic 1-forms is useful to investigate the geometry and topology at infinity. For example, Li and Tam [1992] proved that if the space of L^2 harmonic 1-forms on a complete Riemannian manifold M is trivial, then M must have at most one nonparabolic end. Cao, Shen,

Yun is the corresponding author.

MSC2010: primary 53C20; secondary 58A10, 58E20.

Keywords: harmonic form, harmonic map, Liouville-type theorem, weighted manifold.

and Zhu [Cao et al. 1997] also obtained an interesting topological result which says that if M is a complete Riemannian manifold with all ends of infinite volume supporting a Sobolev inequality and if the space of L^2 harmonic 1-forms is trivial, then M must have only one end. Their argument using the space of L^2 harmonic 1-forms to study the geometry and topology at infinity has been extended in various ways. We refer the readers to [Dung and Seo 2012; 2017; Li and Wang 2002; 2004; Lin 2015; Pigola et al. 2005; Seo 2010; 2014; Vieira 2016; Yun 2002] for recent developments on this topic.

In this paper, we study Liouville-type properties on p-harmonic 1-forms and p-harmonic maps in weighted manifolds. Given a smooth Riemannian manifold (M, g) and a smooth function $f : M \to \mathbb{R}$, a *weighted manifold* (or a smooth metric measure space, also known as a manifold with density) is a triple $M_f := (M, g, e^{-f} dv_g)$, where dv_g is the volume form induced by the metric g. Since the geometry of weighted manifolds were developed by Bakry and Émery [1985], it has been intensively studied by many authors (for instance, see [Lott 2003; Lott and Villani 2009; Sturm 2006a; 2006b; Wei and Wylie 2009]). Moreover, it turned out that the study of weighted manifolds is closely related with that of self-shrinkers and gradient Ricci solitons.

An important geometric quantity on a weighted manifold M_f known as *Bakry–Émery Ricci curvature* is defined by

$$\mathrm{Ric}_f^M = \mathrm{Ric} + \mathrm{Hess}(f),$$

where $\mathrm{Hess}(f)$ denotes the Hessian of f. Obviously, the Bakry–Émery Ricci curvature is a generalization of Ricci curvature. In a weighted manifold, there is a useful elliptic differential operator, the so-called *f-Laplacian*, Δ_f which is defined by

$$\Delta_f u = \Delta u - \langle \nabla f, \nabla u \rangle.$$

The f-Laplacian is a natural generalization of the Laplace–Beltrami operator Δ as it is self-adjoint with respect to the weighted measure $e^{-f} dv_g$, i.e.,

$$\int_M v \Delta_f u e^{-f}\, dv_g = \int_M (u \Delta_f v) e^{-f}\, dv_g$$

and

$$\int_M (v \Delta_f u)\, e^{-f}\, dv_g = -\int_M \langle \nabla u, \nabla v \rangle e^{-f}\, dv_g$$

for $u, v \in C_0^\infty(M)$.

On the other hand, for a smooth map $\varphi : (M^n, g, e^{-f} dv_g) \to (N^m, h)$ from an f-weighted manifold into a Riemannian manifold, and for a bounded domain

$\Omega \subset M$, the *f-weighted p-energy* $\mathcal{E}_{f,p}(\varphi; \Omega)$ with $p > 1$ of φ over Ω is defined by

$$\mathcal{E}_{f,p}(\varphi; \Omega) = \frac{1}{p} \int_{\Omega} |d\varphi|^p e^{-f} \, dv_g, \tag{1-1}$$

where $|d\varphi|$ denotes the Hilbert–Schmidt norm of $d\varphi$ induced by the metrics g and h. Namely, if $\{e_i\}$ is a local frame on M, $|d\varphi|$ is given by

$$|d\varphi|^2 = \sum_{i=1}^{n} \langle d\varphi(e_i), d\varphi(e_i)\rangle \tag{1-2}$$

so that

$$|d\varphi|^2 = \mathrm{tr}_g \, \varphi^* h = \langle g, \varphi^* h\rangle.$$

A smooth map $\varphi : (M, g, e^{-f} dv_g) \to (N, h)$ is called *f-weighted p-harmonic* if it is a critical point of the f-weighted p-energy functional $\mathcal{E}_{f,p}(\varphi; \Omega)$ for any bounded domain $\Omega \subset M$. It can be easily shown that when φ is C^2-regular, the Euler–Lagrange equation for the f-weighted p-energy $\mathcal{E}_{f,p}$ is the f-weighted p-harmonic map equation

$$\tau_{f,p}(\varphi) = -\delta_f(|d\varphi|^{p-2} d\varphi) = |d\varphi|^{p-2}\tau_f(\varphi) + d\varphi(\nabla |d\varphi|^{p-2}) = 0. \tag{1-3}$$

Here $\delta_f = \delta + i_{\nabla f}$ is the adjoint operator of the exterior derivative d with respect to the measure $e^{-f} dv_g$, $i_{\nabla f}$ denotes the interior product with the vector ∇f, $\tau_f(\varphi) = \tau(\varphi) - i_{\nabla f} d\varphi$ and $\tau(\varphi)$ is the classical tension field of φ. In the case where $p = 2$ and f is a constant function, Schoen and Yau [1976] obtained the following well-known Liouville-type theorem for harmonic maps between complete Riemannian manifolds.

Theorem [Schoen and Yau 1976]. *Let M be a complete Riemannian manifold of nonnegative Ricci curvature and let N be a complete Riemannian manifold of nonpositive sectional curvature. Then, for any constant function f, every harmonic map $u : M \to N$ with finite 2-energy $\mathcal{E}_{f,2}(u)$ must be constant.*

Recently, Rimoldi and Veronelli [2013] generalized Schoen and Yau's Liouville-type theorem for harmonic maps into f-weighted 2-harmonic maps between complete Riemannian manifolds. More precisely, they showed that if

$$u : (M^n, g, e^{-f} dv_g) \to (N^m, h)$$

is an f-weighted 2-harmonic map from a complete Riemannian manifold M with nonnegative Bakry–Émery Ricci curvature into a complete Riemannian manifold with nonpositive sectional curvature and if the f-weighted 2-energy $\mathcal{E}_{f,2}(u)$ is finite, then the harmonic map u must be constant. See also [Hua et al. 2017; Nakauchi 1998; Takeuchi 1991; Zhang and Wang 2016] for related previous results. In this paper, we extend their result into f-weighted p-harmonic maps.

The organization of this paper is the following. In Section 2 we derive a Bochner–Weitzenböck formula for the weighted Hodge Laplacian Δ_f on differential forms. Applying this formula, we are able to show a Liouville-type property of weighted L^q p-harmonic 1-forms on a complete noncompact weighted manifold with nonnegative Bakry–Émery Ricci curvature (see Theorem 2.4 for $q = 2p - 2$ and Theorem 2.5 for $q = p$). In Section 3 we obtain a Bochner–Weitzenböck formula for vector bundle-valued weighted p-harmonic forms (Lemma 3.1), which is an extension of our previous results in Section 2. In Section 4 we prove Liouville-type theorems for weighted p-harmonic maps. In fact, we prove that if u is a weighted p-harmonic map from a complete noncompact weighted manifold with nonnegative Bakry–Émery Ricci curvature into a Riemannian manifold with nonpositive sectional curvature and if u has finite weighted q-energy, then u must be constant (see Theorem 4.1 for $q = 2p - 2$ and Theorem 4.2 for $q = p$).

2. Weighted p-harmonic forms

Let (M^n, g) be an n-dimensional complete noncompact Riemannian manifold and let $f : M \to \mathbb{R}$ be a smooth function on M. We consider differential forms on the f-weighted manifold $(M, g, e^{-f} dv_g)$ and derive a Bochner–Weitzenböck formula for the weighted Hodge Laplacian. Recall that the formal adjoint of the exterior derivative d with respect to the measure $e^{-f} dv_g$ is given by the formula

$$\delta_f = \delta + i_{\nabla f}.$$

Then the f-Hodge Laplacian Δ_f on differential forms is defined by

$$\Delta_f = -(d\delta_f + \delta_f d).$$

Lemma 2.1 (Bochner–Weitzenböck formula). *Let $(M, g, e^{-f} dv_g)$ be an f-weighted manifold. If ω is a differential 1-form on M, then*

$$\text{(2-1)} \quad \tfrac{1}{2}\Delta_f|\omega|^{2p-2} = \langle |\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 + |\omega|^{2p-4}\mathrm{Ric}_f^M(\omega^\sharp, \omega^\sharp).$$

Here $\omega^\sharp$ is the dual vector field to ω.

Proof. It is well-known (see [Chang and Sung 2011]) that

$$\begin{aligned}\tfrac{1}{2}\Delta|\omega|^{2p-2} &= \tfrac{1}{2}\Delta||\omega|^{p-2}\omega|^2 \\ &= \langle |\omega|^{p-2}\omega, \Delta(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 + |\omega|^{2p-4}\mathrm{Ric}(\omega^\sharp, \omega^\sharp).\end{aligned}$$

Using the definition of the f-weighted Laplacian $\Delta_f = \Delta - \langle \nabla f, \nabla\cdot\rangle$, we have

$$\tfrac{1}{2}\Delta_f|\omega|^{2p-2} = \tfrac{1}{2}\Delta|\omega|^{2p-2} - \tfrac{1}{2}\langle \nabla f, \nabla|\omega|^{2p-2}\rangle.$$

Since $\mathrm{Ric}_f^M = \mathrm{Ric} + \mathrm{Hess}(f)$ and $\Delta_f = \Delta - di_{\nabla f} - i_{\nabla f}d$, we get

$$\begin{aligned}\tfrac{1}{2}\Delta_f|\omega|^{2p-2} = \langle |\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 + |\omega|^{2p-4}\mathrm{Ric}_f^M(\omega^\sharp, \omega^\sharp)\\ + \langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle + \langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle\\ - |\omega|^{2p-4}\mathrm{Hess}(f)(\omega^\sharp, \omega^\sharp) - \tfrac{1}{2}\langle\nabla f, \nabla(|\omega|^{2p-2})\rangle.\end{aligned}$$

We claim that

$$\begin{aligned}\text{(2-2)}\quad \langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle + \langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle\\ - |\omega|^{2p-4}\mathrm{Hess}(f)(\omega^\sharp, \omega^\sharp) - \tfrac{1}{2}\langle\nabla f, \nabla(|\omega|^{2p-2})\rangle = 0.\end{aligned}$$

Let $\{e_1, \ldots, e_n\}$ be a local geodesic frame at a point p in M and $\{\theta_1, \ldots, \theta_n\}$ its dual coframe. Let $\{\theta_{ij}\}$ be the connection 1-form vanishing at the point p. Writing $\omega = \omega^i\theta_i$ with Einstein convention, we have

$$|\omega|^{2p-4}\mathrm{Hess}(f)(\omega^\sharp, \omega^\sharp) = |\omega|^{2p-4}\omega^i\omega^j f_{ij}.$$

Since

$$di_{\nabla f}(|\omega|^{p-2}\omega) = f_i\omega^i d|\omega|^{p-2} + |\omega|^{p-2}\omega^i f_{ij}\theta_j + |\omega|^{p-2} f_i\omega^i_{;j}\theta_j,$$

we have

$$\begin{aligned}\langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle\\ = |\omega|^{p-2} f_i\omega^i\langle\omega,\ d|\omega|^{p-2}\rangle + |\omega|^{2p-4}\omega^i\omega^j f_{ij} + |\omega|^{2p-4}\omega^j f_i\omega^i_{;j}.\end{aligned}$$

Here the semicolon means the covariant differentiation. Moreover,

$$\begin{aligned}d(|\omega|^{p-2}\omega) &= d|\omega|^{p-2}\wedge\omega + |\omega|^{p-2}d\omega\\ &= d|\omega|^{p-2}\wedge\omega + |\omega|^{p-2}\,d\omega^i\wedge\theta_i\end{aligned}$$

which gives

$$\begin{aligned}\text{(2-3)}\quad & i_{\nabla f}d(|\omega|^{p-2}\omega)\\ &= d|\omega|^{p-2}(\nabla f)\omega - \omega^i f_i d|\omega|^{p-2} + |\omega|^{p-2}\omega^i_{;j} f_j\theta_i - |\omega|^{p-2}\omega^i_{;j} f_i\theta_j.\end{aligned}$$

Thus

$$\begin{aligned}\langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle = |\omega|^p d|\omega|^{p-2}(\nabla f) - |\omega|^{p-2}\omega^i f_i\langle\omega, d|\omega|^{p-2}\rangle\\ + |\omega|^{2p-4}\omega^i_{;j}\omega^i f_j - |\omega|^{2p-4}\omega^i_{;j} f_i\omega^j.\end{aligned}$$

Next we have

$$\begin{aligned}\tfrac{1}{2}\langle\nabla f, \nabla|\omega|^{2p-2}\rangle &= \tfrac{1}{2}\langle\nabla f, \nabla(|\omega|^p\cdot|\omega|^{p-2})\rangle\\ &= \tfrac{1}{2}|\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle + \tfrac{1}{2}|\omega|^{p-2}\langle\nabla f, \nabla|\omega|^p\rangle.\end{aligned}$$

Since

$$\nabla|\omega|^2 = 2\omega^i\omega^i_{;j}e_j \quad\text{and}\quad \nabla|\omega|^p = \nabla(|\omega|^2)^{p/2} = \frac{p}{2}|\omega|^{p-2}\nabla|\omega|^2,$$

we get

$$\tfrac{1}{2}|\omega|^{p-2}\langle\nabla f, \nabla|\omega|^p\rangle = \frac{p}{4}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle = \frac{p}{2}|\omega|^{2p-4} f_j\omega^i\omega^i_{;j}$$

and

$$|\omega|^{2p-4} f_j\omega^i_{;j}\omega^i = \tfrac{1}{2}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle.$$

Thus the left-hand side of (2-2) becomes

$$|\omega|^p d|\omega|^{p-2}(\nabla f)+|\omega|^{2p-4} f_j\omega^i_{;j}\omega^i-\tfrac{1}{2}|\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle-\tfrac{1}{2}|\omega|^{p-2}\langle\nabla f, \nabla|\omega|^p\rangle$$
$$= \tfrac{1}{2}|\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle + \frac{2-p}{4}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle.$$

Since

$$\nabla|\omega|^{p-2} = \nabla(|\omega|^2)^{(p-2)/2} = \frac{p-2}{2}(|\omega|^2)^{(p-2)/2-1}\nabla|\omega|^2 = \frac{p-2}{2}|\omega|^{p-4}\nabla|\omega|^2,$$

the left-hand side of (2-2) vanishes, which completes the proof of Lemma 2.1. □

As a consequence of Lemma 2.1, we have the following.

Corollary 2.2. *Let ω be a differential 1-form on a weighted manifold $(M, g, e^{-f}dv_g)$. Then*

$$|\omega|^{p-1}\Delta_f|\omega|^{p-1} \geq \langle|\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\omega|^{2p-4}\operatorname{Ric}_f^M(\omega^\sharp, \omega^\sharp).$$

Proof. Since

(2-4) $$\tfrac{1}{2}\Delta_f|\omega|^{2p-2} = |\omega|^{p-1}\Delta_f|\omega|^{p-1} + |\nabla|\omega|^{p-1}|^2,$$

it follows from Lemma 2.1 that

$$|\omega|^{p-1}\Delta_f|\omega|^{p-1} + |\nabla|\omega|^{p-1}|^2$$
$$= \langle|\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 + |\omega|^{2p-4}\operatorname{Ric}_f^M(\omega^\sharp, \omega^\sharp).$$

From the generalized Kato type inequality, we have

$$|\nabla|\omega|^{p-1}|^2 = \big|\nabla|\,|\omega|^{p-2}\omega|\big|^2 \leq |\nabla(|\omega|^{p-2}\omega)|^2.$$

Thus we get

$$|\omega|^{p-1}\Delta_f|\omega|^{p-1} \geq \langle|\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\omega|^{2p-4}\operatorname{Ric}_f^M(\omega^\sharp, \omega^\sharp). \qquad \square$$

Let $\phi : M \to \mathbb{R}$ be a harmonic function. Since

$$d(d\phi) = 0 \quad \text{and} \quad \Delta\phi = \delta(d\phi) = 0,$$

the differential $d\phi$ is a harmonic 1-form. Similarly, if $\phi : M \to \mathbb{R}$ is a p-harmonic function, then

$$\Delta_p\phi = \operatorname{div}(|\nabla\phi|^{p-2}\nabla\phi) = 0,$$

which is equivalent to the equation

$$\delta(|d\phi|^{p-2}d\phi)=0.$$

In fact, this is the Euler–Lagrange equation of the p-energy functional $\mathcal{E}_p(\phi)=\frac{1}{p}\int_M |d\phi|^p\,dv_g$. Using this observation, one can define a p-harmonic form ω on M as follows [Chang and Sung 2011]:

$$d\omega=0 \quad \text{and} \quad \delta(|\omega|^{p-2}\omega)=0,$$

which shows that, for any p-harmonic function ϕ on M, its differential $d\phi$ is a p-harmonic 1-form. Motivated by this notion of p-harmonic differential forms in [Chang and Sung 2011] and weighted harmonic forms in [Vieira 2013], we give the definition of weighted p-harmonic forms on a weighted manifold.

Definition 2.3. A differential form ω on M is *f-weighted p-harmonic* if ω satisfies

$$d\omega=0 \quad \text{and} \quad \delta_f(|\omega|^{p-2}\omega)=0.$$

When f is constant, we note that the above definition of f-weighted p-harmonic forms is equivalent to the definition of p-harmonic forms in the sense of [Chang and Sung 2011]. Consider an f-weighted L_f^{2p-2} p-harmonic 1-form ω on a weighted manifold M_f with nonnegative Bakry–Émery Ricci curvature, where the L_f^{2p-2} norm of ω is given by

$$\int_M |\omega|^{2p-2}e^{-f}\,dv_g<\infty.$$

Then we have the following Liouville-type theorem for weighted p-harmonic 1-forms.

Theorem 2.4. *Let $(M,g,e^{-f}dv_g)$ be a complete noncompact f-weighted manifold with nonnegative Bakry–Émery Ricci tensor*, $\mathrm{Ric}_f^M\geq 0$. *Suppose that f is a bounded function. If ω is an f-weighted L_f^{2p-2} p-harmonic 1-form on M for $p>1$, then ω vanishes.*

Proof. Since ω is an f-weighted p-harmonic 1-form, we have

$$\delta_f(|\omega|^{p-2}\omega)=0.$$

Thus Corollary 2.2 together with curvature condition implies

$$|\omega|^{p-1}\Delta_f|\omega|^{p-1}\geq\langle|\omega|^{p-2}\omega,\,\delta_f d(|\omega|^{p-2}\omega)\rangle. \tag{2-5}$$

Fix a point $p\in M$ and choose a cut-off function η satisfying

$$0\leq\eta\leq 1,\quad \eta=1 \text{ on } B_p(r),\quad \mathrm{supp}(\eta)\subset B_p(2r),\quad \text{and}\quad |\nabla\eta|\leq\frac{1}{r}. \tag{2-6}$$

Here $B_p(r)$ denotes the geodesic ball of radius r centered at p. Multiplying (2-5) by η^2 and integrating it over M with respect to the measure $e^{-f}\,dv_g$, we obtain

$$\begin{aligned}
&\int_M \eta^2|\omega|^{p-1}\Delta_f|\omega|^{p-1}e^{-f}\,dv_g\\
&=-\int_M \eta^2|\nabla|\omega|^{p-1}|^2e^{-f}\,dv_g-2\int_M \eta|\omega|^{p-1}\langle\nabla\eta,\nabla|\omega|^{p-1}\rangle e^{-f}\,dv_g\\
&\le-\int_M \eta^2|\nabla|\omega|^{p-1}|^2e^{-f}\,dv_g+\tfrac{1}{2}\int_M \eta^2|\nabla|\omega|^{p-1}|^2e^{-f}\,dv_g\\
&\qquad\qquad+2\int_M |\omega|^{2p-2}|\nabla\eta|^2e^{-f}\,dv_g\\
&=-\tfrac{1}{2}\int_M \eta^2|\nabla|\omega|^{p-1}|^2e^{-f}\,dv_g+2\int_M |\omega|^{2p-2}|\nabla\eta|^2e^{-f}\,dv_g.
\end{aligned}$$

Moreover

$$\begin{aligned}
&\int_M \eta^2\langle|\omega|^{p-2}\omega,\delta_f d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g\\
&=\int_M \langle d(\eta^2|\omega|^{p-2}\omega),d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g\\
&=2\int_M \eta|\omega|^{p-2}\langle d\eta\wedge\omega,d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g+\int_M \eta^2|d(|\omega|^{p-2}\omega)|^2e^{-f}\,dv_g\\
&\ge-\int_M \eta^2|d(|\omega|^{p-2}\omega)|^2e^{-f}\,dv_g-\int_M |\nabla\eta|^2|\omega|^{2p-2}e^{-f}\,dv_g\\
&\qquad\qquad+\int_M \eta^2|d(|\omega|^{p-2}\omega)|^2e^{-f}\,dv_g\\
&=-\int_M |\nabla\eta|^2|\omega|^{2p-2}e^{-f}\,dv_g.
\end{aligned}$$

Therefore

$$\tfrac{1}{2}\int_M \eta^2|\nabla|\omega|^{p-1}|^2e^{-f}\,dv_g\le 3\int_M |\nabla\eta|^2|\omega|^{2p-2}e^{-f}\,dv_g\le\frac{3}{r^2}\int_M |\omega|^{2p-2}e^{-f}\,dv_g.$$

Since ω is an f-weighted L^{2p-2} harmonic 1-form, we obtain

$$\nabla|\omega|^{p-1}=0$$

by letting $r\to\infty$. Hence $|\omega|^{p-1}$ is constant. Since $\mathrm{Ric}_f^M\ge 0$ and f is bounded, the f-volume of (M,g) is infinite (see [Wei and Wylie 2009], for example). Therefore we see that $\omega=0$. □

Using the Bochner–Weitzenböck formula, we can also prove the following.

Theorem 2.5. *Let $(M,g,e^{-f}\,dv_g)$ be a complete noncompact f-weighted manifold with nonnegative Bakry–Émery Ricci tensor. Suppose that f is a bounded function. For $p\ge 2$, if ω is an f-weighted L_f^p p-harmonic 1-form on M, then $\omega=0$.*

Proof. Since $\delta_f(|\omega|^{p-2}\omega) = 0$, Corollary 2.2 and the curvature condition implies

$$\begin{aligned} |\omega|\Delta_f|\omega|^{p-1} &\geq \langle \omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\omega|^{2p-4}\operatorname{Ric}_f^M(\omega^\sharp, \omega^\sharp) \\ &\geq \langle \omega, \delta_f d(|\omega|^{p-2}\omega)\rangle. \end{aligned} \tag{2-7}$$

Fix a point $p \in M$ and choose a cut-off function η satisfying (2-6). Multiplying (2-7) by η^2 and integrating it over M with respect to the measure $e^{-f}dv_g$, we obtain

$$\int_M \eta^2|\omega|\Delta_f|\omega|^{p-1}e^{-f}\,dv_g \geq \int_M \eta^2\langle \omega, \delta_f d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g. \tag{2-8}$$

Then the left-hand side of (2-8) is given by

$$\begin{aligned} &\int_M \eta^2|\omega|\Delta_f|\omega|^{p-1}e^{-f}\,dv_g \\ &= -\int_M \eta^2\langle \nabla|\omega|, \nabla|\omega|^{p-1}\rangle e^{-f}\,dv_g - 2\int_M \eta|\omega|\langle \nabla\eta, \nabla|\omega|^{p-1}\rangle e^{-f}\,dv_g \\ &= -(p-1)\int_M \eta^2|\omega|^{p-2}\big|\nabla|\omega|\big|^2 e^{-f}\,dv_g \\ &\qquad - 2(p-1)\int_M \eta|\omega|^{p-1}\langle \nabla\eta, \nabla|\omega|\rangle e^{-f}\,dv_g. \end{aligned} \tag{2-9}$$

Note that

$$|\omega|^{p-2}\big|\nabla|\omega|\big|^2 = \frac{4}{p^2}\big|\nabla|\omega|^{p/2}\big|^2$$

and

$$|\omega|^{p-1}\nabla|\omega| = |\omega|^{p/2}\cdot|\omega|^{p/2-1}\nabla|\omega| = \frac{2}{p}|\omega|^{p/2}\nabla|\omega|^{p/2}. \tag{2-10}$$

Substituting these two identities into (2-9), we obtain

$$\begin{aligned} &\int_M \eta^2|\omega|\Delta_f|\omega|^{p-1}e^{-f}\,dv_g \\ &= -\frac{4(p-1)}{p^2}\int_M \eta^2\big|\nabla|\omega|^{p/2}\big|^2 e^{-f}\,dv_g \\ &\qquad - \frac{4(p-1)}{p}\int_M \eta|\omega|^{p/2}\langle \nabla\eta, \nabla|\omega|^{p/2}\rangle e^{-f}\,dv_g \\ &\leq -\frac{4(p-1)}{p^2}\int_M \eta^2\big|\nabla|\omega|^{p/2}\big|^2 e^{-f}\,dv_g \\ &\quad + \frac{2(p-1)}{p}\left\{\varepsilon\int_M \eta^2\big|\nabla|\omega|^{p/2}\big|^2 e^{-f}\,dv_g + \frac{1}{\varepsilon}\int_M |\omega|^p|\nabla\eta|^2 e^{-f}\,dv_g\right\}, \end{aligned} \tag{2-11}$$

where we used Young's inequality in the last inequality for arbitrary $\varepsilon > 0$.

On the other hand, applying the divergence theorem with respect to the measure $e^{-f}\,dv_g$, the right-hand side of (2-8) becomes

$$\int_M \eta^2\langle\omega, \delta_f d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g = \int_M \langle d(\eta^2\omega), d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g.$$

Since

$$|d(\varphi\omega)| = |d\varphi\wedge\omega| \le |d\varphi|\,|\omega|$$

for any smooth function $\varphi : M \to \mathbb{R}$ and any closed 1-form ω (see Lemma 13 in [Pigola et al. 2008]), using (2-10) and Young's inequality again gives

$$\begin{aligned}
\text{(2-12)}\qquad \big|\langle d(\eta^2\omega), d(|\omega|^{p-2}\omega)\rangle\big| &\le |d(\eta^2\omega)|\big|d(|\omega|^{p-2}\omega)\big| \\
&\le |d\eta^2|\,|\omega|^2\big|d|\omega|^{p-2}\big| \\
&= 2\eta|\omega|^2|\nabla\eta|\big|\nabla|\omega|^{p-2}\big| \\
&= 2(p-2)\eta|\nabla\eta|\,|\omega|^{p-1}\big|\nabla|\omega|\big| \\
&= \frac{4(p-2)}{p}\eta|\nabla\eta|\,|\omega|^{p/2}\big|\nabla|\omega|^{p/2}\big| \\
&\le \frac{2(p-2)}{p}\Big(\delta\eta^2\big|\nabla|\omega|^{p/2}\big|^2 + \frac{1}{\delta}|\nabla\eta|^2|\omega|^p\Big)
\end{aligned}$$

for any $\delta > 0$. Therefore

$$\begin{aligned}
\text{(2-13)}\qquad \int_M \eta^2\langle\omega, \delta_f d(|\omega|^{p-2}\omega)\rangle e^{-f}\,dv_g \ge &-\frac{2(p-2)}{p}\delta\int_M \eta^2\big|\nabla|\omega|^{p/2}\big|^2 e^{-f}\,dv_g \\
&-\frac{2(p-2)}{p}\frac{1}{\delta}\int_M |\nabla\eta|^2|\omega|^p e^{-f}\,dv_g.
\end{aligned}$$

Combining (2-8), (2-11) and (2-13), we obtain

$$\begin{aligned}
\Big(\frac{4(p-1)}{p^2} - \frac{2(p-1)}{p}\varepsilon - \frac{2(p-2)}{p}\delta\Big)&\int_M \eta^2\big|\nabla|\omega|^{p/2}\big|^2 e^{-f}\,dv_g \\
&\le \Big(\frac{2(p-1)}{p}\frac{1}{\varepsilon} + \frac{2(p-2)}{p}\frac{1}{\delta}\Big)\int_M |\nabla\eta|^2|\omega|^p e^{-f}\,dv_g.
\end{aligned}$$

Choose ε and δ sufficiently small so that

$$\frac{4(p-1)}{p^2} - \frac{2(p-1)}{p}\varepsilon - \frac{2(p-2)}{p}\delta > 0.$$

Since ω is an L^p_f p-harmonic 1-form, as r tends to infinity, we see

$$\nabla|\omega|^{p/2} = 0,$$

which implies that $\omega \equiv 0$ as in the proof of Theorem 2.4. □

Remark 2.6. In Theorems 2.4 and 2.5, the boundedness on the weighted function f is only needed to guarantee that the weighted volume of $(M, g, e^{-f} dv_g)$ is infinite. In fact, we prove that any f-weighted L_f^{2p-2} p-harmonic 1-form with $p > 1$ or L_f^p p-harmonic 1-form with $p \geq 2$ on a complete noncompact f-weighted manifold with nonnegative Bakry–Émery Ricci tensor has constant length, which implies that ω is f-harmonic. Thus applying the standard Bochner formula for f-harmonic 1-forms (see Lemma 2.1 with $p = 2$, [Lott 2003] or [Vieira 2013]), one can see that ω is parallel without the assumption that f is bounded. This result leads to applications in gradient steady Ricci solitons or, more generally, to applications in weighted manifolds with infinite weighted volumes (see [Vieira 2013]). Recall that a gradient steady Ricci soliton is a manifold (M, g) together with a smooth function f satisfying $\mathrm{Ric}_f^M = 0$.

Furthermore, if we assume that Ric_f^M is nonnegative and positive at a point, it is easy to see, from Corollary 2.2, that ω vanishes without assuming the boundedness of f. This property leads to applications in gradient shrinking Ricci solitons satisfying $\mathrm{Ric}_f^M = \lambda g$ for some positive constant λ as follows.

Corollary 2.7. *Let $(M, g, e^{-f} dv_g)$ be a complete gradient shrinking Ricci soliton satisfying* $\mathrm{Ric} + \mathrm{Hess}(f) = \lambda g$ *with* $\lambda > 0$, *constant. Then if ω is an L_f^{2p-2} $(p > 1)$ or L_f^p $(p \geq 2)$ p-harmonic 1-form on M, then $\omega = 0$.*

Proof. The proof follows from the argument in Remark 2.6. □

In case of gradient steady Ricci solitons, we also have the following same vanishing property.

Corollary 2.8. *Let $(M, g, e^{-f} dv_g)$ be a complete gradient steady Ricci soliton satisfying* $\mathrm{Ric} + \mathrm{Hess}(f) = 0$. *Then if ω is an L_f^{2p-2} $(p > 1)$ or L_f^p $(p \geq 2)$ p-harmonic 1-form on M, then $\omega = 0$.*

Proof. For $q = 2p - p$ or $q = p$, applying the same argument as in the proofs of Theorems 2.4 and 2.5, we see that $|\omega| \equiv C$ for some constant C. Thus

$$\int_M |\omega|^q e^{-f} dv_g = C^q \,\mathrm{Vol}_f(M),$$

where $\mathrm{Vol}_f(M)$ denotes the f-weighted volume of M.

On the other hand, it is well-known that the scalar curvature of a gradient steady Ricci soliton is nonnegative and $|\nabla f|$ is bounded by a positive constant (see [Cao 2010] for example). Moreover, Munteanu and Wang [2011] proved that the first eigenvalue of f-Laplacian Δ_f on the nontrivial gradient steady Ricci solitons is positive. Therefore, applying the result by Vieira [2013], we get $\mathrm{Vol}_f(M) = \infty$. This shows that $\omega = 0$. □

3. Vector bundle-valued weighted p-harmonic forms

In this section, we extend the notions discussed in Section 2 including the Bochner–Weitzenböck formula to vector bundles over a weighted manifold.

Let $\pi : E \to M$ be a vector bundle of rank m over a smooth oriented Riemannian manifold (M^n, g). We denote by $\Gamma(E)$ the vector space of smooth sections of E over M. A Riemannian structure on the bundle E is a pair (∇^E, ρ), where ρ is a Riemannian metric on E, ∇^E a connection and $\nabla^E \rho = 0$. Denoting $\rho = \langle \cdot, \cdot \rangle$, the condition $\nabla^E \rho = 0$ means that, for each $X \in \Gamma(TM)$ and $s_1, s_2 \in \Gamma(E)$, we have

$$X \cdot \langle s_1, s_2 \rangle = \langle \nabla^E s_1, s_2 \rangle + \langle s_1, \nabla^E s_2 \rangle.$$

The curvature of the connection ∇^E is the map $R^E : \Lambda^2 TM \otimes \Gamma(E) \to \Gamma(E)$ defined by

$$R^E(X, Y)s = -\nabla^E_X \nabla^E_Y s + \nabla^E_Y \nabla^E_X s + \nabla^E_{[X,Y]} s.$$

Let ω be an l-form on M with values in the vector bundle $\pi : E \to M$. Then, choosing a (local) frame $s_1, \ldots, s_m$ on E, for each $X_1, \ldots, X_l \in \Gamma(TM)$, we can write

$$\omega(X_1, \ldots, X_l) = \sum_{\alpha=1}^{m} a_\alpha s_\alpha$$

for some local smooth functions a_α on M. For the Levi–Civita connection $D^M = D$ on (M, g), the induced connection ∇ on $\Gamma(\Lambda^l T^* M \otimes E)$, the space of smooth l-forms on M with values in the vector bundle $\pi : E \to M$, is given by

$$(\nabla_X \omega)(X_1, \ldots, X_l) = \nabla^E_X(\omega(X_1, \ldots, X_l)) - \sum_{i=1}^{l} \omega(X_1, \ldots, D_X X_i, \ldots, X_l)$$

and its associated curvature is given by

$$\begin{aligned}(R(X, Y)\omega)&(X_1, \ldots X_l) \\ &= R^E(X, Y)(\omega(X_1, \ldots, X_l)) - \sum_{i=1}^{l} \omega(X_1, \ldots, R^M(X, Y)X_i, \ldots, X_l).\end{aligned}$$

For the induced connection ∇, the exterior differential operator

$$d : \Gamma(\Lambda^l T^* M \otimes E) \to \Gamma(\Lambda^{l+1} T^* M \otimes E)$$

is given by

$$(d\omega)(X_1, \ldots, X_{l+1}) = \sum_{i=1}^{l+1} (-1)^{i+1} (\nabla_{X_i} \omega)(X_1, \ldots, \widehat{X_i}, \ldots, X_{l+1}),$$

where the symbol covered by $\widehat{X_i}$ is omitted. The codifferential operator δ is given by

$$(\delta\omega)(X_1,\dots,X_{l-1}) = -\sum_{i=1}^{n}(\nabla_{e_i}\omega)(e_i, X_1,\dots,X_{l-1}),$$

where $\{e_i\}$ is a local frame on M. Finally the Laplacian Δ and the f-weighted Laplacian Δ_f are defined on E-valued differential forms by

$$\Delta = -(d\delta+\delta d) \quad \text{and} \quad \Delta_f = -(d\delta_f+\delta_f d),$$

respectively.

For a vector bundle $\pi : E \to M$ over a weighted manifold $(M, g, e^{-f}dv_g)$, we have the following Bochner–Weitzenböck formula for differential 1-forms on M with values in E.

Lemma 3.1 (Bochner–Weitzenböck formula). *Let $\pi : E \to M$ be a vector bundle of rank m over a smooth oriented Riemannian manifold (M, g), and let $f : M \to \mathbb{R}$ be a smooth function. If ω is an E-valued 1-form on M, then*

$$\begin{aligned}(3\text{-}1)\quad \tfrac{1}{2}\Delta_f|\omega|^{2p-2} = \langle|\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 \\ + |\omega|^{2p-4}\sum_{i=1}^{n}\langle\omega(\mathrm{Ric}_f^M(e_i)), \omega(e_i)\rangle \\ - |\omega|^{2p-4}\sum_{i,j}\langle R^E(e_i,e_j)\omega(e_i), \omega(e_j)\rangle,\end{aligned}$$

where $\{e_i\}$ is a local frame on M and $\mathrm{Ric}_f^M(e_i)$ is a vector given by

$$\mathrm{Ric}_f^M(e_i) = \sum_{j=1}^{n}\mathrm{Ric}_f^M(e_i,e_j)e_j = \sum_{j=1}^{n}\big[\mathrm{Ric}^M(e_i,e_j)+\mathrm{Hess}(f)(e_i,e_j)\big]e_j.$$

Proof. It is well-known (see [Eells and Lemaire 1983]) that

$$\begin{aligned}(3\text{-}2)\quad \tfrac{1}{2}\Delta|\omega|^{2p-2} = \langle|\omega|^{p-2}\omega, \Delta(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2 \\ + |\omega|^{2p-4}\sum_{i=1}^{n}\langle\omega(\mathrm{Ric}^M(e_i)), \omega(e_i)\rangle \\ - |\omega|^{2p-4}\sum_{i,j}\langle R^E(e_i,e_j)\omega(e_i), \omega(e_j)\rangle.\end{aligned}$$

By definition of weighted Laplacian $\Delta_f = \Delta - \langle\nabla f, \nabla\cdot\rangle$, we have

$$\tfrac{1}{2}\Delta_f|\omega|^{2p-2} = \tfrac{1}{2}\Delta|\omega|^{2p-2} - \tfrac{1}{2}\langle\nabla f, \nabla|\omega|^{2p-2}\rangle.$$

Since

$$\mathrm{Ric}_f^M = \mathrm{Ric}^M + \mathrm{Hess}(f) \quad \text{and} \quad \Delta_f = \Delta - di_{\nabla f} - i_{\nabla f}d,$$

we get

$$\begin{aligned}
&\tfrac{1}{2}\Delta_f|\omega|^{2p-2}\\
&= \langle |\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + \langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle\\
&\qquad + \langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle + |\nabla(|\omega|^{p-2}\omega)|^2\\
&\qquad + |\omega|^{2p-4}\sum_{i=1}^{n}\langle \omega(\mathrm{Ric}_f^M(e_i)), \omega(e_i)\rangle - |\omega|^{2p-4}\langle \omega(\mathrm{Hess}(f)(e_i)), \omega(e_i)\rangle\\
&\qquad - |\omega|^{2p-4}\sum_{i,j}\langle R^E(e_i,e_j)\omega(e_i), \omega(e_j)\rangle - \tfrac{1}{2}\langle \nabla f, \nabla(|\omega|^{2p-2})\rangle.
\end{aligned}$$

We claim that

$$\begin{aligned}
\text{(3-3)}\quad &\langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle + \langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle\\
&\qquad - |\omega|^{2p-4}\langle \omega(\mathrm{Hess}(f)(e_i)), \omega(e_i)\rangle - \tfrac{1}{2}\langle \nabla f, \nabla(|\omega|^{2p-2})\rangle = 0.
\end{aligned}$$

Let $\{e_1, \dots, e_n\}$ be a local geodesic frame at a point p in M, and $\{\theta_1, \dots, \theta_n\}$ be its dual coframe. Let $\{\theta_{ij}\}$ be the connection 1-form vanishing at the point p. Let $\{s_1, \dots, s_m\}$ be a local frame on E such that

$$\nabla^E s_\alpha|_p = 0.$$

Then ω can be expressed as

$$\omega = \sum_{\alpha=1}^{m}\sum_{i=1}^{n} a_{i\alpha}\theta_i \otimes s_\alpha$$

so that

$$\omega(e_j) = \sum_\alpha a_{j\alpha}s_\alpha \quad \text{and} \quad |\omega|^2 = \sum_{i,\alpha} a_{i\alpha}^2.$$

Since

$$\begin{aligned}
di_{\nabla f}(|\omega|^{p-2}\omega) &= di_{\nabla f}(|\omega|^{p-2}a_{i\alpha}\theta_i \otimes s_\alpha)\\
&= d(|\omega|^{p-2}f_i a_{i\alpha}s_\alpha)\\
&= f_i a_{i\alpha}d|\omega|^{p-2} \otimes s_\alpha + |\omega|^{p-2}a_{i\alpha}f_{ij}\theta_j \otimes s_\alpha\\
&\qquad + |\omega|^{p-2}f_i a_{i\alpha;j}\theta_j \otimes s_\alpha + |\omega|^{p-2}f_i a_{i\alpha}\nabla^E s_\alpha,
\end{aligned}$$

we have

$$\begin{aligned}
\text{(3-4)}\quad \langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle &= |\omega|^{p-2}\langle a_{j\alpha}\theta_j \otimes s_\alpha, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle\\
&= |\omega|^{p-2}a_{j\alpha}f_i a_{i\alpha}d|\omega|^{p-2}(e_j)\\
&\qquad + |\omega|^{2p-4}a_{j\alpha}a_{i\alpha}f_{ij} + |\omega|^{2p-4}a_{j\alpha}f_i a_{i\alpha;j}.
\end{aligned}$$

Moreover

$$\begin{aligned} d(|\omega|^{p-2}\omega) &= a_{i\alpha}(d|\omega|^{p-2}\wedge\theta_i)\otimes s_\alpha + |\omega|^{p-2}(da_{i\alpha}\wedge\theta_i)\otimes s_\alpha \\ &\qquad + |\omega|^{p-2}a_{i\alpha}\theta_{ij}\wedge\theta_j\otimes s_\alpha - |\omega|^{p-2}a_{i\alpha}\theta_i\wedge\nabla^E s_\alpha \\ &= a_{i\alpha}(d|\omega|^{p-2}\wedge\theta_i)\otimes s_\alpha + |\omega|^{p-2}a_{i\alpha;j}(\theta_j\wedge\theta_i)\otimes s_\alpha \end{aligned}$$

which gives

$$\begin{aligned} i_{\nabla f}d(|\omega|^{p-2}\omega) &= d|\omega|^{p-2}(\nabla f)\omega - a_{i\alpha}f_i d|\omega|^{p-2}\otimes s_\alpha \\ &\qquad + |\omega|^{p-2}a_{i\alpha;j}f_j\theta_i\otimes s_\alpha - |\omega|^{p-2}a_{i\alpha;j}f_i\theta_j\otimes s_\alpha. \end{aligned}$$

Thus

$$\begin{aligned} &\langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle \\ &= |\omega|^{p-2}\langle a_{j\beta}\theta_j\otimes s_\beta, i_{\nabla f}d(|\omega|^{p-2}\omega\rangle \\ &= |\omega|^p d|\omega|^{p-2}(\nabla f) - |\omega|^{p-2}a_{i\alpha}a_{j\alpha}f_i d|\omega|^{p-2}(e_j) \\ &\qquad + |\omega|^{2p-4}a_{i\alpha;j}a_{i\alpha}f_j - |\omega|^{2p-4}a_{i\alpha;j}f_i a_{j\alpha} \\ &= |\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle - |\omega|^{p-2}a_{i\alpha}a_{j\alpha}f_i d|\omega|^{p-2}(e_j) \\ &\qquad + \tfrac{1}{2}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle - |\omega|^{2p-4}a_{i\alpha;j}f_i a_{j\alpha}. \end{aligned} \tag{3-5}$$

Note that

$$\langle\omega(\mathrm{Hess}(f)(e_i)), \omega(e_i)\rangle = f_{ij}\langle\omega(e_j), \omega(e_i)\rangle = f_{ij}a_{j\alpha}a_{i\alpha}. \tag{3-6}$$

From (3-4), (3-5), and (3-6), it follows that

$$\begin{aligned} \langle |\omega|^{p-2}\omega, di_{\nabla f}(|\omega|^{p-2}\omega)\rangle + \langle |\omega|^{p-2}\omega, i_{\nabla f}d(|\omega|^{p-2}\omega)\rangle \\ - |\omega|^{2p-4}\langle\omega(\mathrm{Hess}(f)(e_i)), \omega(e_i)\rangle \\ = |\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle + \tfrac{1}{2}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle. \end{aligned}$$

We observe that

$$\begin{aligned} \tfrac{1}{2}\langle\nabla f, \nabla|\omega|^{2p-2}\rangle &= \tfrac{1}{2}\langle\nabla f, \nabla(|\omega|^p\cdot|\omega|^{p-2})\rangle \\ &= \tfrac{1}{2}|\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle + \tfrac{1}{2}|\omega|^{p-2}\langle\nabla f, \nabla|\omega|^p\rangle. \end{aligned}$$

Since

$$\nabla|\omega|^p = \nabla(|\omega|^2)^{p/2} = \frac{p}{2}|\omega|^{p-2}\nabla|\omega|^2,$$

we have

$$\tfrac{1}{2}|\omega|^{p-2}\langle\nabla f, \nabla|\omega|^p\rangle = \frac{p}{4}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle.$$

Thus the left-hand side of (3-3) becomes

$$\tfrac{1}{2}|\omega|^p\langle\nabla f, \nabla|\omega|^{p-2}\rangle + \frac{2-p}{4}|\omega|^{2p-4}\langle\nabla f, \nabla|\omega|^2\rangle.$$

Using

$$\begin{aligned}\nabla|\omega|^{p-2} &= \nabla(|\omega|^2)^{(p-2)/2}\\ &= \frac{p-2}{2}(|\omega|^2)^{(p-2)/2-1}\nabla|\omega|^2\\ &= \frac{p-2}{2}|\omega|^{p-4}\nabla|\omega|^2,\end{aligned}$$

we see that the left-hand side of (3-3) vanishes, which completes the proof of Lemma 3.1. □

As in the proof of Corollary 2.2, we can easily show the following by using Lemma 3.1.

Corollary 3.2. *Let $\pi : E \to M$ be a vector bundle of rank m over a smooth oriented Riemannian manifold (M, g), and let $f : M \to \mathbb{R}$ be a smooth function. If ω is an E-valued 1-form on M, then*

$$\begin{aligned}|\omega|^{p-1}\Delta_f|\omega|^{p-1} \geq \langle|\omega|^{p-2}\omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\omega|^{2p-4}\sum_{i=1}^{n}\langle\omega(\mathrm{Ric}_f^M(e_i)), \omega(e_i)\rangle\\ - |\omega|^{2p-4}\sum_{i,j}\langle R^E(e_i, e_j)\omega(e_i), \omega(e_j)\rangle.\end{aligned}$$

4. Weighted p-harmonic maps

In this section, we obtain some Liouville-type theorems for weighted p-harmonic maps as an application of the Bochner–Weitzenböck formula stated in Section 3. The following theorem shows that the same result holds for f-weighted p-harmonic maps with L_f^{2p-2}-finite energy for $p > 1$ as in the case of f-weighted L_f^{2p-2} p-harmonic 1-forms.

Theorem 4.1. *Let $u : (M, g, e^{-f}dv_g) \to (N, h)$ be an f-weighted p-harmonic map from an oriented complete noncompact f-weighted manifold into a Riemannian manifold for $p > 1$. Suppose that f is bounded. Assume that the Bakry–Émery Ricci curvature of M is nonnegative, $\mathrm{Ric}_f^M \geq 0$, and the sectional curvature of N is nonpositive, $K^N \leq 0$. If u has finite f-weighted $(2p-2)$-energy, i.e.,*

$$\int_M |du|^{2p-2}e^{-f}\,dv_g < \infty,$$

then u must be a constant map.

Proof. Let $du = \omega$. Then ω is an f-weighted p-harmonic 1-form with values in the pull-back bundle $u^{-1}TN$. In particular,

$$\delta_f(|\omega|^{p-2}\omega) = 0.$$

From Corollary 3.2 together with curvature conditions, it follows that

$$|\omega|^{p-1}\Delta_f|\omega|^{p-1} \geq \langle |\omega|^{p-2}\omega, \delta_f d(|\omega|^{p-2}\omega)\rangle. \tag{4-1}$$

From this, we can see that the same argument as in the proof of Theorem 2.4 shows $\omega = 0$. □

From Corollary 3.2, it follows that

$$|\omega|\Delta_f|\omega|^{p-1} \geq \langle \omega, \Delta_f(|\omega|^{p-2}\omega)\rangle + |\omega|^{p-2}\sum_{i=1}^{n}\langle \omega(\mathrm{Ric}_f^M(e_i)), \omega(e_i)\rangle - |\omega|^{p-2}\sum_{i,j}\langle R^E(e_i, e_j)\omega(e_i), \omega(e_j)\rangle. \tag{4-2}$$

Applying the same argument as in Theorem 2.5 to (4-2), we are able to prove the following theorem.

Theorem 4.2. *Let $u : (M, g, e^{-f}dv_g) \to (N, h)$ be an f-weighted p-harmonic map from an oriented complete noncompact f-weighted manifold into a Riemannian manifold. Suppose that f is bounded, and $\mathrm{Ric}_f^M \geq 0$ and $K^N \leq 0$. For $p \geq 2$, if u has finite f-weighted p-energy, then u must be a constant map.*

Remark 4.3. In Theorems 4.1 and 4.2, without the boundedness of f, if we assume that Ric_f is nonnegative and positive at a point, we can conclude that any f-weighted p-harmonic map $u : (M, g, e^{-f}dv_g) \to (N, h)$ with finite f-weighted $(2p-2)$-energy or p-energy for $p > 1$ from an oriented complete noncompact f-weighted manifold into a Riemannian manifold of nonpositive sectional curvature, $K^N \leq 0$, must be constant.

Applying the argument in Remark 4.3 to gradient shrinking Ricci solitons, we have the following as in the case of L_f^p p-harmonic 1-forms.

Corollary 4.4. *Let $(M, g, e^{-f}dv_g)$ be a complete noncompact gradient shrinking Ricci soliton satisfying $\mathrm{Ric} + \mathrm{Hess}(f) = \lambda g$ with $\lambda > 0$, constant. If $u : (M, g, e^{-f}dv_g) \to (N, h)$ is an f-weighted p-harmonic map into a Riemannian manifold of nonpositive sectional curvature $K^N \leq 0$ with finite f-weighted $(2p-2)$-energy or p-energy for $p > 1$, then u must be a constant map.*

Acknowledgments

The authors would like to express their gratitude to the referees for valuable comments. In particular, their suggestions for vanishing properties of p-harmonic forms and p-harmonic maps on gradient Ricci solitons make this paper more fruitful. Seo was supported by the National Research Foundation of Korea (NRF-2016R1C1B2009778) and Yun was supported by the National Research Foundation of Korea (NRF-2016R1D1A1A09916749).

References

[Bakry and Émery 1985] D. Bakry and M. Émery, "Diffusions hypercontractives", pp. 177–206 in *Séminaire de probabilités, XIX, 1983/84*, edited by J. Azéma and M. Yor, Lecture Notes in Math. **1123**, Springer, 1985. MR Zbl

[Cao 2010] H.-D. Cao, "Recent progress on Ricci solitons", pp. 1–38 in *Recent advances in geometric analysis*, edited by Y.-I. Lee et al., Adv. Lect. Math. (ALM) **11**, International Press, Somerville, MA, 2010. MR Zbl

[Cao et al. 1997] H.-D. Cao, Y. Shen, and S. Zhu, "The structure of stable minimal hypersurfaces in $\mathbb{R}^{n+1}$", *Math. Res. Lett.* **4**:5 (1997), 637–644. MR Zbl

[Chang and Sung 2011] L.-C. Chang and C.-J. A. Sung, "A note on p-harmonic l-forms on complete manifolds", *Pacific J. Math.* **254**:2 (2011), 295–307. MR Zbl

[Dung and Seo 2012] N. T. Dung and K. Seo, "Stable minimal hypersurfaces in a Riemannian manifold with pinched negative sectional curvature", *Ann. Global Anal. Geom.* **41**:4 (2012), 447–460. MR Zbl

[Dung and Seo 2017] N. T. Dung and K. Seo, "p-harmonic functions and connectedness at infinity of complete submanifolds in a Riemannian manifold", *Ann. Mat. Pura Appl.* (4) **196**:4 (2017), 1489–1511. MR Zbl

[Eells and Lemaire 1983] J. Eells and L. Lemaire, *Selected topics in harmonic maps*, CBMS Regional Conference Series in Mathematics **50**, American Mathematical Society, Providence, RI, 1983. MR Zbl

[Greene and Wu 1979] R. E. Greene and H. Wu, *Function theory on manifolds which possess a pole*, Lecture Notes in Mathematics **699**, Springer, 1979. MR Zbl

[Hildebrandt 1982] S. Hildebrandt, "Liouville theorems for harmonic mappings, and an approach to Bernstein theorems", pp. 107–131 in *Seminar on Differential Geometry*, edited by S. T. Yau, Ann. of Math. Stud. **102**, Princeton Univ. Press, 1982. MR Zbl

[Hua et al. 2017] B. Hua, S. Liu, and C. Xia, "Liouville theorems for f-harmonic maps into Hadamard spaces", *Pacific J. Math.* **290**:2 (2017), 381–402. MR Zbl

[Huber 1957] A. Huber, "On subharmonic functions and differential geometry in the large", *Comment. Math. Helv.* **32** (1957), 13–72. MR Zbl

[Karp 1982] L. Karp, "Subharmonic functions on real and complex manifolds", *Math. Z.* **179**:4 (1982), 535–554. MR Zbl

[Li and Tam 1992] P. Li and L.-F. Tam, "Harmonic functions and the structure of complete manifolds", *J. Differential Geom.* **35**:2 (1992), 359–383. MR Zbl

[Li and Wang 2002] P. Li and J. Wang, "Minimal hypersurfaces with finite index", *Math. Res. Lett.* **9**:1 (2002), 95–103. MR Zbl

[Li and Wang 2004] P. Li and J. Wang, "Stable minimal hypersurfaces in a nonnegatively curved manifold", *J. Reine Angew. Math.* **566** (2004), 215–230. MR Zbl

[Lin 2015] H. Lin, "L^2 harmonic forms on submanifolds in a Hadamard manifold", *Nonlinear Anal.* **125** (2015), 310–322. MR Zbl

[Lott 2003] J. Lott, "Some geometric properties of the Bakry–Émery–Ricci tensor", *Comment. Math. Helv.* **78**:4 (2003), 865–883. MR Zbl

[Lott and Villani 2009] J. Lott and C. Villani, "Ricci curvature for metric-measure spaces via optimal transport", *Ann. of Math.* (2) **169**:3 (2009), 903–991. MR Zbl

[Munteanu and Wang 2011] O. Munteanu and J. Wang, "Smooth metric measure spaces with non-negative curvature", *Comm. Anal. Geom.* **19**:3 (2011), 451–486. MR Zbl

[Nakauchi 1998] N. Nakauchi, "A Liouville type theorem for p-harmonic maps", *Osaka J. Math.* **35**:2 (1998), 303–312. MR Zbl

[Pigola et al. 2005] S. Pigola, M. Rigoli, and A. G. Setti, "Vanishing theorems on Riemannian manifolds, and geometric applications", *J. Funct. Anal.* **229**:2 (2005), 424–461. MR Zbl

[Pigola et al. 2008] S. Pigola, M. Rigoli, and A. G. Setti, "Constancy of p-harmonic maps of finite q-energy into non-positively curved manifolds", *Math. Z.* **258**:2 (2008), 347–362. MR Zbl

[Rimoldi and Veronelli 2013] M. Rimoldi and G. Veronelli, "Topology of steady and expanding gradient Ricci solitons via f-harmonic maps", *Differential Geom. Appl.* **31**:5 (2013), 623–638. MR Zbl

[Schoen and Yau 1976] R. Schoen and S. T. Yau, "Harmonic maps and the topology of stable hypersurfaces and manifolds with non-negative Ricci curvature", *Comment. Math. Helv.* **51**:3 (1976), 333–341. MR Zbl

[Seo 2010] K. Seo, "L^2 harmonic 1-forms on minimal submanifolds in hyperbolic space", *J. Math. Anal. Appl.* **371**:2 (2010), 546–551. MR Zbl

[Seo 2014] K. Seo, "L^p harmonic 1-forms and first eigenvalue of a stable minimal hypersurface", *Pacific J. Math.* **268**:1 (2014), 205–229. MR Zbl

[Sturm 2006a] K.-T. Sturm, "On the geometry of metric measure spaces, I", *Acta Math.* **196**:1 (2006), 65–131. MR Zbl

[Sturm 2006b] K.-T. Sturm, "On the geometry of metric measure spaces, II", *Acta Math.* **196**:1 (2006), 133–177. MR Zbl

[Takeuchi 1991] H. Takeuchi, "Stability and Liouville theorems of p-harmonic maps", *Japan. J. Math.* (*N.S.*) **17**:2 (1991), 317–332. MR Zbl

[Vieira 2013] M. Vieira, "Harmonic forms on manifolds with non-negative Bakry–Émery–Ricci curvature", *Arch. Math.* (*Basel*) **101**:6 (2013), 581–590. MR Zbl

[Vieira 2016] M. Vieira, "Vanishing theorems for L^2 harmonic forms on complete Riemannian manifolds", *Geom. Dedicata* **184** (2016), 175–191. MR Zbl

[Wei and Wylie 2009] G. Wei and W. Wylie, "Comparison geometry for the Bakry–Emery Ricci tensor", *J. Differential Geom.* **83**:2 (2009), 377–405. MR Zbl

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR Zbl

[Yau 1976] S. T. Yau, "Some function-theoretic properties of complete Riemannian manifold and their applications to geometry", *Indiana Univ. Math. J.* **25**:7 (1976), 659–670. MR Zbl

[Yun 2002] G. Yun, "Total scalar curvature and L^2 harmonic 1-forms on a minimal hypersurface in Euclidean space", *Geom. Dedicata* **89** (2002), 135–141. MR Zbl

[Zhang and Wang 2016] J.-F. Zhang and Y. Wang, "A theorem of Liouville type for p-harmonic maps in weighted Riemannian manifolds", *Kodai Math. J.* **39**:2 (2016), 354–365. MR Zbl

Received April 30, 2018. Revised February 17, 2019.

KEOMKYO SEO
DEPARTMENT OF MATHEMATICS
SOOKMYUNG WOMEN'S UNIVERSITY
YONGSAN-GU
SEOUL
SOUTH KOREA
kseo@sookmyung.ac.kr

GABJIN YUN
DEPARTMENT OF MATHEMATICS
MYONG JI UNIVERSITY
CHEOIN-GU
YONG-IN
SOUTH KOREA
gabjin@mju.ac.kr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.311

REMARKS ON THE HÖLDER-CONTINUITY OF SOLUTIONS TO PARABOLIC EQUATIONS WITH CONIC SINGULARITIES

YUANQI WANG

This is a note on work of Ladyzhenskaja et al. (AMS 1968) and of Ferretti and Safonov (2013). Using their work line by line, we prove the Hölder-continuity of solutions to linear parabolic equations of mixed type, assuming the coefficient of $\frac{\partial}{\partial t}$ has time-derivative bounded from above. On a Kähler manifold, this Hölder estimate works when the metrics possess conic singularities along a normal crossing divisor.

1. Introduction

Historically, Hölder-continuity of solutions to linear elliptic and parabolic equations (in various cases) has been proved and extensively studied by De Giorgi [1957], Nash [1958], Moser [1964], Krylov and Safonov [1980]. Many other experts have contributed to this topic as well, for example, see [Caffarelli and Cabré 1995; Ferretti and Safonov 2013; Gilbarg and Trudinger 1977; Ladyzhenskaja et al. 1968]. Ferretti and Safonov [2013] give a unified proof of the Hölder-continuity in both the divergence case and nondivergence case. The key is to establish growth properties for the (sub- and super-) level sets of the solutions.

We focus on divergence-form equations. The operator in our main parabolic equation (3) below is exactly the one considered in [Ferretti and Safonov 2013]. The little difference is: the a_0 $\left(\text{coefficient of } \frac{\partial}{\partial t}\right)$ in [Ferretti and Safonov 2013, line 11, page 89] is not allowed to depend on time, but here we allow a_0 in (3) to depend on time.

Our motivation is to study the heat equation associated to a Ricci flow. The Ricci flow is a special time-parametrized family of Riemannian metrics $g(t)$. Given a time-family of Riemannian metrics $g(t)$ over a Euclidean ball B, the associated heat equation reads as

$$\frac{\partial u}{\partial t} - \Delta_g u \triangleq \frac{\partial u}{\partial t} - \frac{1}{\sqrt{\det g_{ij}}} \frac{\partial}{\partial x_i}\left(g^{ij}\sqrt{\det g_{ij}}\,\frac{\partial u}{\partial x_j}\right) = f, \tag{1}$$

MSC2010: 35K10.
Keywords: Hölder estimate, parabolic equations.

where the x_i are the Euclidean coordinates. To estimate the Hölder norm of u, we only care about the L^∞-norm of f, though we can assume that everything involved has higher derivatives. Multiplying (1) by $\sqrt{\det g_{ij}}$, we get

$$\sqrt{\det g_{ij}}\frac{\partial u}{\partial t}-\frac{\partial}{\partial x_i}\left(g^{ij}\sqrt{\det g_{ij}}\frac{\partial u}{\partial x_j}\right)=F\triangleq f\sqrt{\det g_{ij}}. \tag{2}$$

Let $a_0\triangleq\sqrt{\det g_{ij}}$ and $a^{ij}\triangleq g^{ij}\sqrt{\det g_{ij}}$; (1) is a special case of (3) and [Ferretti and Safonov 2013, Equation (D), page 89]. Suppose $\det g_{ij}$ is uniformly bounded; the L^∞-norm of f is equivalent to the L^∞-norm of F, thus it makes no difference for the Hölder estimate.

Our main observation (and a one sentence proof of Theorem 1.1) is that when a_0 depends on time and $\frac{\partial \log a_0}{\partial t}$ is bounded from above, the general energy estimates are still true (Lemma 4.5). By the proof in [Ferretti and Safonov 2013], these energy estimates imply the main growth theorem [Ferretti and Safonov 2013, Theorem 5.3]. Moreover, by an idea in [Ladyzhenskaja et al. 1968], [Ferretti and Safonov 2013, Theorem 5.3] directly implies the Hölder continuity of solutions, without involving the Harnack inequality in [Ferretti and Safonov 2013, Theorem 1.5]. We believe these are known by experts. Let

- $Y=(y,s)$ be a space-time point, and $C_r(Y)=B_y(r)\times(s-r^2,s)$ be the parabolic cylinder centered at Y with radius r, where $B_y(r)$ is the usual m-dimensional Euclidean ball.
- Let $[\,\cdot\,]_\alpha$ denote the usual parabolic Hölder seminorm of exponent α (for example, see [Lieberman 1996, (4.1)] for a definition).

The simplest version of our main theorem can be stated as follows.

Theorem 1.1. *Suppose $u\in C^\infty[C_r(Y)]$ solves the following equation* (or the metric heat equation (1) via the correspondence in (2)) *in the classical sense*

$$a_0\frac{\partial u}{\partial t}-\frac{\partial}{\partial x_j}\left(a^{ij}\frac{\partial u}{\partial x_i}\right)=f, \tag{3}$$

where a_0, a^{ij} $(1\le i,j\le m)$ are space-time smooth functions. Suppose

$$\frac{1}{K}\le a_0\le K,\qquad \frac{\partial\log a_0}{\partial t}\le K,\qquad \frac{I}{K}\le a^{ij}\le KI. \tag{4}$$

Then there exist constants $\alpha(m,K)\in(0,1)$ and $N(m,K)$ such that

$$r^\alpha[u]_{\alpha,C_{r/2}(Y)}+|u|_{L^\infty[C_{r/2}(Y)]}\le N\left(\frac{|u|_{L^1[C_r(Y)]}}{r^{m+2}}+r^2|f|_{L^\infty[C_r(Y)]}\right).$$

Remark 1.2. When $\frac{\partial\log a_0}{\partial t}$ is not uniformly bounded from above (while the other conditions in Theorem 1.1 hold true), the above uniform Hölder estimate does not hold in general. We refer the interested readers to the beautiful example constructed by Chen and Safonov [2017, Theorems 4.1 and 4.2].

Remark 1.3. In view of the Riemannian geometry setting in (2) and the line right below it, when g_t is a Ricci flow, the upper bound on $\frac{\partial \log a_0}{\partial t}$ means

$$\frac{\partial}{\partial t} d\mathrm{vol}_{g_t} \leq K\, d\mathrm{vol}_{g_t}, \tag{5}$$

where $d\mathrm{vol}_{g_t}$ is the evolving volume form. The K is actually a lower bound for the scalar curvature of g_t. Fortunately, the scalar curvature is usually bounded from below along Ricci flows without any additional condition, see [Cao and Chen 2012, page 5; Hamilton 1995].

Therefore, when the scalar curvature of a Kähler–Ricci flow is not (assumed to be) bounded from above, Theorems 1.1 and 2.2 (as well as the ideas in the proof) might help in obtaining higher-order estimates (convergence) of the Kähler metric as t approaches a singular time or as $t \to \infty$.

Theorem 1.1 can be generalized to heat equations of Kähler-metrics with conic singularities along normal-crossing divisors (Theorem 2.2). We only prove Theorem 2.2, the proof of Theorem 1.1 is the same (by discarding the necessary techniques for the conic singularities, see Claim 4.7 for example).

This note is organized as follows. In Section 2 we define Kähler metrics with conic singularities, and state Theorem 2.2. In Section 3 we prove Theorems 1.1 and 2.2 assuming the main growth theorem (Theorem 5.7) on sublevel sets of subsolutions. In Section 4, we define weak subsolutions and prove energy inequalities for them. This is a preparation for the main growth theorem (Theorem 5.7). In Section 5, we prove the main growth theorem using the energy inequalities and the measure theoretic arguments in [Ferretti and Safonov 2013].

2. The more general version of Theorem 1.1 in Kähler geometry involving conic singularity

In Kähler geometry setting, Theorem 1.1 holds even when the metrics possess conic singularities along analytic hypersurfaces. To state the result, we first give a geometric formulation following [Guenancia and Păun 2016]. Given a closed Kähler manifold M and a divisor $D = \sum_{j=1}^{N} 2\pi(1-\beta_j)D_j$, where each D_j is an irreducible hypersurface and may have self-intersection, suppose D has (no worse than) normal crossing singularities, i.e., there is an open cover of supp D by neighborhoods $\mathcal{U}_i$ such that in each $\mathcal{U}_i$, supp $D \cap \mathcal{U}_i = \{z_1 z_2 z_3 \cdots z_k = 0\}$, where $k \leq n$ and $z_1 \cdots z_n$ are holomorphic coordinate functions in $\mathcal{U}_i$.

Definition 2.1. A Kähler metric g (defined away from supp D) is said to be a weak-conic metric with quasi-isometric constant K, if and only if it's Hölder-continuous away from supp D and in each $\mathcal{U}_i$,

$$\begin{aligned} &\frac{g_\beta^k}{K} \le g \le K g_\beta^k \quad \text{(quasi-isometric)}, \\ &g_\beta^k = \sum_{j=1}^{k} \frac{\beta_j^2}{|z_j|^{2-2\beta_j}} \, dz_j \otimes d\bar{z}_j + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j. \end{aligned} \tag{6}$$

g_β^k is one of the 2 model metrics on $\mathbb{C}^n$ we work with, and in this local setting we abuse notation by denoting supp D as D.

Similarly, a Kähler metric g is called a ϵ-nearly-conic metric with quasi-isometric constant K, if and only if it's Hölder-continuous over the whole M (across supp D) and in each $\mathcal{U}_i$, for $\epsilon > 0$,

$$\begin{aligned} &\frac{g_{\beta,\epsilon}^k}{K} \le g \le K g_{\beta,\epsilon}^k, \\ &g_{\beta,\epsilon}^k = \sum_{j=1}^{k} \frac{\beta_j^2}{(|z_j|^2 + \epsilon^2)^{1-\beta_j}} \, dz_j \otimes d\bar{z}_j + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j, \end{aligned} \tag{7}$$

We recall the well known intrinsic polar coordinates of g_β^k. Let

$$\xi_j = r_j e^{\sqrt{-1}\theta_j}, \quad r_j = |z_j|^{\beta_j}, \quad 1 \le j \le k.$$

In these polar coordinates the model cone g_β^k is equal to

$$g_\beta^k = \sum_{j=1}^{k} (dr_j^2 + \beta_j^2 r_j^2 \, d\theta_j^2) + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j,$$

and it's quasi-isometric to the Euclidean metric, i.e.,

$$\bigl(\min_j \beta_j^2\bigr) g_E \le g_\beta^k \le g_E, \quad g_E = \sum_{j=1}^{k} (dr_j^2 + r_j^2 \, d\theta_j^2) + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j. \tag{8}$$

This is important because we want to take advantage of the rescaling and translation invariance of the Euclidean metric.

Similarly, we also have intrinsic polar coordinates for $g_{\beta,\epsilon}^k$. Let s_j be the solution to

$$\frac{ds_j}{d\rho_j} = \frac{\beta_j}{(\rho_j^2 + \epsilon^2)^{(1-\beta_j)/2}}, \quad s_j(0) = 0, \quad \rho_j = |z_j|. \tag{9}$$

Then $\xi_j = s_j e^{\sqrt{-1}\theta_j}$, $1 \le j \le k$ defines the polar coordinates of $g^k_{\beta,\epsilon}$. By [Wang 2016, Lemma 4.3], in these coordinates we have

$$g^k_{\beta,\epsilon} = \sum_{j=1}^{k}(ds_j^2 + a_{j,\epsilon}s_j^2\,d\theta_j^2) + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j, \quad \beta_j^2 < a_{j,\epsilon} \le 1. \tag{10}$$

Hence $g^k_{\beta,\epsilon}$ is also quasi-isometric to the Euclidean metric in its polar coordinate, i.e.,

$$\big(\min_j \beta_j^2\big) g_E \le g^k_{\beta,\epsilon} \le g_E, \quad g_E = \sum_{j=1}^{k}(ds_j^2 + s_j^2\,d\theta_j^2) + \sum_{j=k+1}^{n} dz_j \otimes d\bar{z}_j. \tag{11}$$

Unless specified (via a parentheses or a subsymbol), the constants N and C in this article depend on (at most) n, K, M, D, $\beta_j's$, and the open cover $\bigcup_i \mathcal{U}_i$. They don't depend on ϵ. Different $N's$ could be different. The real dimension is $m = 2n$ in the Kähler setting.

Theorem 2.2. *Let* $\epsilon \in [0, 1]$.

Part I (local estimate). *Suppose g_t is a time-differentiable family of weak-conic Kähler metrics or of ϵ-nearly-conic metrics, which is defined over a parabolic cylinder $C_r(Y)$ in $\mathbb{C}^n$ under a polar coordinate as below* (7) *or* (9), *respectively. Suppose the quasi-isometric constant of g_t is K*,

$$\frac{\partial}{\partial t}\, d\mathrm{vol}_t \le K\, d\mathrm{vol}_t. \tag{12}$$

and u is a bounded weak solution (in the sense of Definition 4.2) *to*

$$\frac{\partial u}{\partial t} = \Delta_{g_t} u + f \quad \text{over } C_r(Y). \tag{13}$$

Then there exists $\alpha(n, \beta, K) \in (0, 1)$ and $N(n, \beta, K)$ such that

$$r^\alpha [u]_{\alpha, C_{r/2}(Y)} + |u|_{L^\infty, C_{r/2}(Y)} \le N\left(\frac{|u|_{L^1[C_r(Y)]}}{r^{2n+2}} + r^2 |f|_{L^\infty[C_r(Y)]}\right).$$

Part II (global estimate). *In the setting of* (6) *and paragraph above it, suppose all the conditions in* Part I *hold globally on $M \times [0, T]$. Then for all $t_0 \in (0, T)$, there exists an $\alpha(n, \beta, K)$ and $C_{t_0}(n, \beta, K)$ such that*

$$[u]_{\alpha, M\times[t_0,T]} + |u|_{L^\infty(M\times[t_0,T])} \le C_{t_0}(|u|_{L^1(M\times[0,T])} + |f|_{L^\infty(M\times[0,T])}).$$

Remark 2.3. When the divisor is smooth, a weaker version of this Hölder estimate is in section 4 of [Wang 2016]. We hope it's still somewhat valuable to present the proof separately here. The $[u]_\alpha$ is the usual parabolic Hölder seminorm with respect to g^k_β ($g^k_{\beta,\epsilon}$) (see (8)). An important point is that Hölder continuity with respect to the distance of g^k_β ($g^k_{\beta,\epsilon}$) is equivalent to Hölder continuity in the usual sense in holomorphic coordinates (apart from a difference of Hölder exponents). We refer interested readers to [Wang 2016, Lemma 4.4]. Please see Definition 4.2 for the definition of weak solutions (replace SC_r by the underlying domain).

Remark 2.4. Using the Kähler structure, Equation (13) can be written in both divergence and nondivergence form. In this case, we expect that Theorem 2.2 still holds without condition (12).

3. Proof of the main results assuming Theorem 5.7

From now on (and in the subsequent sections), we work in the polar coordinates in (8) and (11). In this coordinate, we do not "see" the conic singularity (except that the coefficients of the equations and solutions are not defined on D).

Definition 3.1. Let C_r^0 denote $C_r(y, s-3r^2)$ (see above Theorem 1.1 on definition of parabolic cylinders). We note that C_r^0 is "earlier" in time than C_r.

Let $|\Omega|$ denote the Euclidean measure of a set $\Omega \in \mathbb{R}^m \times \mathbb{R}$, where $\mathbb{R}^m$ is the spatial direction, and $\mathbb{R}$ is the time direction, i.e., $\mathbb{R}^m \times \mathbb{R}$ is the full parabolic cylinder.

Given a function u on an arbitrary set $\Omega \in \mathbb{R}^m \times \mathbb{R}$, we define

$$\mathrm{osc}_\Omega\, u \triangleq \sup_\Omega u - \inf_\Omega u.$$

Roughly speaking, to show parabolic Hölder continuity of a function u, it suffices to show that comparing to the oscillation in an arbitrary parabolic cylinder (in the domain of u), the oscillation in the concentric subcylinder of half radius decreases by a fixed amount.

The idea in proving Theorems 1.1 and 2.2 can be explained as follows. For simplicity, we assume $f = 0$ in the main heat equation (3).

For a subsolution to the homogeneous version of (3), the role of the main growth theorem (Theorem 5.7) is to improve a bound on the measure of the superlevel set in an "earlier" parabolic cylinder to a pointwise bound in a "later" cylinder.

The crucial observation is that for any solution u, either u or $-u$ (adjusted by proper constants) admits a bound on the measure of the superlevel set in the "earlier" parabolic cylinder (see Cases 1 and 2 below (15)), hence the condition in the main growth theorem is always satisfied, and it yields the desired decrease of oscillation.

Proof of Theorems 1.1 and 2.2. We only prove (Part I of) Theorem 2.2 as mentioned at the end of the introduction. Notice that y does not have to be in supp D (as long as integration by parts is true, see proof of Lemma 4.5). By the interior L^∞-estimate in Proposition 5.3 which holds for every cylinder and every subsolution, it suffices to show the Hölder norm is bounded by the L^∞-norm, i.e.,

$$r^\alpha [u]_{\alpha, C_{r/2}(Y)} \le N(|u|_{L^\infty[C_r(Y)]} + r^2 |f|_{L^\infty[C_r(Y)]}). \tag{14}$$

By [Lieberman 1996, Lemma 4.6], it suffices to show the oscillation decays for every cylinder C_{2r} and every subsolution u, i.e.,

$$\operatorname{osc}_{C_r} u \leq (1-b)\operatorname{osc}_{C_{2r}} u + 4r^2|f|_{0,C_{2r}}, \quad b = b(n,\beta,K) > 0. \tag{15}$$

By rescaling and translation invariance, it suffices to assume $r=1$, $s=0$. By adding a constant, it suffices to assume $0 \leq u \leq h$, where $h \triangleq \operatorname{osc}_{C_2} u$. As in [Ladyzhenskaja et al. 1968], one of the following must hold:

$$\text{Case 1:} \quad \left|\left\{u > \frac{h}{2}\right\} \cap C_1^0\right| \leq \frac{|C_1^0|}{2}; \qquad \text{Case 2:} \quad \left|\left\{u < \frac{h}{2}\right\} \cap C_1^0\right| \leq \frac{|C_1^0|}{2}.$$

We only prove (15) in Case 1 in detail. Case 2 is similar by applying the proof in Case 1 to $h-u$. Consider $\bar{u} = u - t|f|_{0,C_{2r}}$ $(t \in (-4,0))$. Then

$$\begin{gathered} \frac{\partial \bar{u}}{\partial t} - \Delta_g \bar{u} \leq 0, \quad 0 \leq \bar{u} \leq h + 4|f|_{0,C_2}. \\ \text{Moreover,} \quad \bar{u} > \frac{h}{2} + 4|f|_{0,C_2} \Rightarrow u > \frac{h}{2}. \end{gathered} \tag{16}$$

Hence the assumption of Case 1 implies

$$\left|\left\{\bar{u} \geq \frac{h}{2} + 4|f|_{0,C_2}\right\} \cap C_1^0\right| \leq \left|\left\{u > \frac{h}{2}\right\} \cap C_1^0\right| \leq \frac{|C_1^0|}{2}.$$

Then Theorem 5.7 $\big($applied to $\bar{u} - \frac{h}{2} - 4|f|_{0,C_2}\big)$, (16), and the above inequality imply that there exists $a(n,\beta,K) > 0$ such that

$$\sup_{C_1}\left(\bar{u} - \frac{h}{2} - 4|f|_{0,C_2}\right) \leq (1-a)\sup_{C_2}\left(\bar{u} - \frac{h}{2} - 4|f|_{0,C_2}\right) \leq \frac{(1-a)h}{2}.$$

Then $\operatorname{osc}_{C_1} \leq \sup_{C_1} u \leq \sup_{C_1} \bar{u} \leq \big(1 - \frac{a}{2}\big)h + 4|f|_{0,C_2}$. The proof of (15) (under the normalization conditions below it) is complete. □

4. Energy inequalities

We follow closely the definitions and tricks in [Ferretti and Safonov 2013]; the point is that they work equally well in the presence of conic singularity (Definitions 4.1, 4.2, and 4.4). The functions and integrations are all defined away from the divisor D (see the content below (6)). If the notation of a function space does not involve D, we mean the space satisfies the indicated asymptotic property (which should be clear from the context). The sets and (slant) cylinders are standard ones minus D. This does not affect any measure theory, integration, or technique in this article, because the spacewise codimension of D is 2. For the proof of Theorem 1.1, we don't have to chop off any singularity.

Definition 4.1. A slant cylinder $\mathrm{SC}_r(y_0, y_1, T_0, T_1)$, which we abbreviate in most the time as SC_r, is the following set:

$$\mathrm{SC}_r \triangleq \{x \mid |x - y(t)| < r,\ T_0 < t \le T_1\}, \tag{17}$$

where $y(t) = y_0 + (t - T_0)(y_1 - y_0)/(T_1 - T_0)$. When $y_1 = y_0$, SC_r is just the usual cylinder C_r defined above Theorem 1.1. We define $l \triangleq r(y_1 - y_0)/(T_1 - T_0)$ as the parabolic slope of SC_r. The parabolic slope l is invariant under

- the usual parabolic rescaling (linear multiplication on y_0, y_1, r and quadratic multiplication on T_0, T_1 by the same factor),
- the spacewise translation (on y_0, y_1 by the same displacement),
- and the timewise translation (on T_0 and T_1 by the same displacement).

Definition 4.2. We say u is a weak subsolution to

$$\frac{\partial u}{\partial t} - \Delta_g u \le 0, \tag{18}$$

in a slant cylinder SC_r if

(1) $u \in C^{2+\alpha,1+\alpha/2}\{\mathrm{SC}_r \setminus D \times [T_0, T_1]\} \cap L^\infty(\mathrm{SC}_r)$;

(2) Inequality (18) holds over $\mathrm{SC}_r \setminus D \times [T_0, T_1]$ in the classical sense.

We call a function η (defined in any bounded space-time domain $\Omega \in \mathbb{C}^n \times (-\infty, \infty)$) tame if $\eta \in C^{1,1}\{\Omega \setminus D \times [T_0, T_1]\} \cap L^\infty(\Omega)$ and the following holds:

$$\frac{\partial \eta}{\partial t} \in L^1(\Omega), \qquad \nabla \eta \in L^2(\Omega). \tag{19}$$

Remark 4.3. The $L^\infty(\mathrm{SC}_r)$-requirement in Definition 4.2 is crucial, and is the only global condition. It guarantees (18) holds across the singularity in the sense of integration by parts.

Definition 4.4. Exactly as in [Ferretti and Safonov 2013, Corollary 2.3], we define the cutoff function of u as

$$u_\epsilon = G(u), \tag{20}$$

where G is a function with one variable such that $G(u) = 0$ when $u \le \epsilon$, $G(u) = u + G(2\epsilon) - 2\epsilon$ when $u \ge 2\epsilon$, and $G, G', G'' \ge 0$. Consequently, we have

$$G(2\epsilon) \le \epsilon \quad \text{and} \quad \max\{u - 2\epsilon, 0\} \le u_\epsilon \le \max\{u - \epsilon, 0\}. \tag{21}$$

The most important feature of u_ϵ is that, if u is a solution to (18), so is u_ϵ, i.e.,

$$\frac{\partial u_\epsilon}{\partial t} - \Delta_g u_\epsilon \le 0. \tag{22}$$

The cutoff function u_ϵ can be understood as the smoothing of u^+ (nonnegative part of u). We note that in the classical case, u^+ is a subsolution (in proper sense) if u is. The above smoothing is pointwise, thus works in the presence of conic singularities.

Lemma 4.5. *Under the same assumptions in* Part I *of Theorem 2.2* (for any r), *suppose u is a nonnegative weak solution to (18) in the sense of Definition 4.2 in a slant cylinder* SC_r, $r \le \frac{1}{100n}$. *Then for any nonnegative tame function η which is compactly supported in* SC_r *spacewisely*, *we have*

$$\text{(23)}\quad \int_{\mathbb{C}^n} u\eta^2\, dV_g|_{t_1}^{t_2} + \int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g u, \nabla_g \eta^2\rangle\, dV_g\, ds \le \int_{t_1}^{t_2}\int_{\mathbb{C}^n} u\frac{\partial \eta^2}{\partial t}\, dV_g\, ds + K\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u\eta^2\, dV_g\, ds.$$

Moreover, *we have*

$$\text{(24)}\quad \int_{\mathbb{C}^n} u^2\eta^2\, dV_g|_{t_1}^{t_2} + \tfrac{1}{2}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2\eta^2\, dV_g\, ds \le \int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\frac{\partial \eta^2}{\partial t}\, dV_g\, ds + (2K+200)\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2(\eta^2 + |\nabla_g \eta|^2)\, dV_g\, ds,$$

and therefore

$$\text{(25)}\quad \int_\Omega |\nabla_g u|^2\, dV_g\, ds < +\infty, \text{ *for any* (parabolic) *compact subdomain* } \Omega \text{ *of* } \mathrm{SC}_r.$$

Remark 4.6. By the same proof, the energy estimate of (3) is similar.

Proof of Lemma 4.5. Let r_i be the distance function to the smooth hypersurface D_i. We consider Berndtsson's cutoff function $\psi_{i,\epsilon} = \psi(\epsilon\log(-\log r_i))$, ψ is the standard cutoff function such that $\psi(x) \equiv 1$ when $x \le \frac{1}{2}$, and $\psi(x) \equiv 0$ when $x \ge \frac{4}{5}$. Then

$$\text{(26)}\qquad \psi_{i,\epsilon} \equiv 0 \quad \text{when } r_i \le e^{-e^{4/(5\epsilon)}}; \qquad \psi_{i,\epsilon} \equiv 1 \quad \text{when } r_i \ge e^{-e^{1/(2\epsilon)}}.$$

Let $\psi_\epsilon = \prod_{i=1,\dots,n} \psi_{i,\epsilon}$.

Claim 4.7. *We have*

$$\text{(27)}\qquad \lim_{\epsilon\to 0} |\nabla_E \psi_\epsilon|_{L^2(B(1/2))} = 0.$$

The proof of Claim 4.7 is elementary. We only verify it for $\partial\psi_\epsilon/\partial r_1$, the other directional derivatives are similar. We compute

$$\frac{\partial \psi_\epsilon}{\partial r_1} = -\psi' \frac{\epsilon}{r_1 \log r_1}\prod_{i\ne 1}\psi_{i,\epsilon}.$$

Hence in polycylindrical coordinates we find

$$\int_{B(1/2)} |\frac{\partial \psi_\epsilon}{\partial r_1}|^2 \, d\mathrm{vol}_E \le C\epsilon^2 \int_0^{1/2} \frac{1}{r_1 (\log r_1)^2} \, dr_1 \le C\epsilon^2.$$

We first prove (24). By definition we have $\lim_{\epsilon\to 0} \psi_\epsilon = 1$ everywhere except on supp D. We multiply both hand sides of (18) by $u\eta^2\psi_\epsilon^2$, then integrate by parts and integrate with respect to time. We obtain

$$\begin{aligned}(28)\quad &\tfrac{1}{2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \, dV_g|_{t_1}^{t_2} + \int_{t_1}^{t_2} |\nabla_g u|^2 \eta^2 \psi_\epsilon^2 \, dV_g \, ds \\ &\le \tfrac{1}{2}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \frac{\partial dV_g}{\partial t} \, ds - 2\int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g u, \nabla_g \eta\rangle u\eta\psi_\epsilon^2 \, dV_g \, ds \\ &\quad + \tfrac{1}{2}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2 \frac{\partial \eta^2}{\partial t}\psi_\epsilon^2 dV_g \, ds - 2\int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g u, \nabla_g \psi_\epsilon\rangle u\eta^2\psi_\epsilon \, dV_g \, ds.\end{aligned}$$

Using the Cauchy–Schwartz inequality we deduce that

$$\begin{aligned}(29)\quad &\left| 2\int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g u, \nabla_g \psi_\epsilon\rangle u\eta^2\psi_\epsilon \, dV_g \, ds \right| \\ &\le \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2 \psi_\epsilon^2\eta^2 \, dV_g \, ds + 100\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2 |\nabla_g \psi_\epsilon|^2 \, dV_g \, ds.\end{aligned}$$

Similarly we have

$$\begin{aligned}(30)\quad &\left| 2\int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g u, \nabla_g \eta\rangle u\eta\psi_\epsilon^2 \, dV_g \, ds \right| \\ &\le \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2 \psi_\epsilon^2\eta^2 \, dV_g \, ds + 100\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\psi_\epsilon^2 |\nabla_g \eta|^2 \, dV_g \, ds.\end{aligned}$$

Notice that by (5) we have

$$(31)\quad \int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \frac{\partial dV_g}{\partial t} \, ds \le K\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \, dV_g \, ds.$$

Then

$$\begin{aligned}(32)\quad &\tfrac{1}{2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \, dV_g|_{t_1}^{t_2} + \int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2 \eta^2 \psi_\epsilon^2 \, dV_g \, ds \\ &\le K\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2\psi_\epsilon^2 \, dV_g \, ds + \tfrac{1}{2}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2 \frac{\partial \eta^2}{\partial t}\psi_\epsilon^2 \, dV_g \, ds \\ &\quad + \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2 \psi_\epsilon^2\eta^2 \, dV_g \, ds + 100\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\psi_\epsilon^2 |\nabla_g \eta|^2 \, dV_g \, ds \\ &\quad + \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g u|^2 \psi_\epsilon^2\eta^2 \, dV_g \, ds + 100\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2 |\nabla_g \psi_\epsilon|^2 \, dV_g \, ds.\end{aligned}$$

We note that Definition 4.2 requires $u \in L^\infty$, then (27) implies

$$\lim_{\epsilon\to 0} 100\int_{t_1}^{t_2}\int_{\mathbb{C}^n} u^2\eta^2|\nabla_g\psi_\epsilon|^2\, dV_g\, ds = 0. \tag{33}$$

Let $\epsilon \to 0$ in (32), the proof of (24) and (25) is complete.

Multiplying both hand sides of (18) by $\eta\psi_\epsilon$ and integrating by parts over space-time, (23) is proved similarly. □

By the same proof as for Lemma 4.5 (with Berndtsson's cutoff function as in (27)), the Sobolev embedding theorem is true.

Lemma 4.8 (Sobolev embedding). *Given a function $u \in C^1\{B \setminus D\} \cap L^\infty(B)$, for any cutoff function $\eta \in C_0^1(B)$, the following holds*:

$$\left(\int_B |\eta u|^{\frac{2n}{2n-1}}\, dV_E\right)^{\frac{2n-1}{2n}} \le N(\beta, n)\int_B |\nabla(\eta u)|\, dV_E.$$

Proof. It's true when $\int_B |\nabla(\eta u)|\, dV_E = \infty$. When $\int_B |\nabla(\eta u)|\, dV_E < \infty$, using Berndtsson's cutoff function ψ_ϵ, Claim 4.7, and the same proof as for Lemma 4.5, ηu belongs to $W^{1,1}(B)$ in the usual sense. Then it follows from the usual Sobolev-inequality. □

Remark 4.9. The $N(\beta, n)$ above does not depend on the radius or center of the ball. The only place where we use the Sobolev embedding is (41).

5. Proof of Theorem 5.7 by energy inequalities

Lemma 5.1 (Growth Lemma). *Suppose u is a weak subsolution to (18) in a cylinder $C_{2r}(Y)$. Then there exists a $\mu_2(n, \beta, K) > 0$ such that*

$$\frac{|\{u > 0\} \cap C_{2r}(Y)|}{|C_{2r}(Y)|} \le \mu_2 \quad \textit{implies} \quad \sup_{C_r} u \le \tfrac{1}{2}\sup_{C_{2r}} u^+. \tag{34}$$

Proof. The proof is formally the same as for [Ferretti and Safonov 2013, Lemma 4.1]. Since condition (5) is involved, we still give a detailed proof for the reader's convenience. The point is to show that we don't need more on the equation than the energy estimates of subsolutions (Lemma 4.5 and the proof of it). The constants N in this proof only depend on n, β, K.

By rescaling invariance of the subequation (18), it suffices to assume $r = 1$ and $\sup_{C_r} u = 1$. We let μ_2 be small enough. It suffices to prove that for all $Z \notin D$ and $Z \in C_1(Y) = C_1$, under the condition

$$\frac{|\{u > 0\} \cap C_1(Z)|}{|C_1(Z)|} \le \frac{2^{2n+2}|\{u > 0\} \cap C_2(Z)|}{|C_2(Z)|} \le 2^{2n+2}\mu_2 \triangleq \mu_1, \tag{35}$$

the following estimate holds

$$u(Z) \le \tfrac{1}{2}. \tag{36}$$

We only need to apply Lemma 5.2 (see [Ferretti and Safonov 2013, (3.8), page 33]). Using exactly the induction argument [Ferretti and Safonov 2013, from the last line of page 99 to line 16 of page 100] (only involving Lemma 5.2), we deduce for any integer $j \ge 0$, for some $N(n, \beta, K)$, the following estimate holds when $N\mu_1^{1/(2n+2)} < \frac{1}{2}$.

$$\left|\left\{u > \frac{1-\rho}{2}\right\} \cap C_\rho(Z)\right| \le \mu_1 \rho^{2n+2} |C_\rho(Z)|, \quad \rho = 2^{-j}. \tag{37}$$

Since $Z \notin D$, (37) directly implies that $u(Z) \le \frac{1}{2}$. Were this not true, $u(Z) > \frac{1}{2}$ implies that there exists dyadic ρ_0 small enough such that $C_{\rho_0}(Z)$ does not touch the singularity D, and $u > \frac{1}{2}$ over $C_{\rho_0}(Z)$. This contradicts (37). □

Lemma 5.2. *Under the same setting as in Lemma 5.1 and its proof above, for any constant* $A \ge 0$, *we have*

$$\int_{C_{\rho/2}(Z)} (u - A)_+ \, dV_E \, ds \le \frac{N}{\rho} |\{u > A\} \cap C_\rho(Z)|^{1+1/(2n+2)}.$$

Proof. By linearity and rescaling invariance of the subequation (18), without loss of generality we can assume $A = 0$ and $\rho = 1$ (note $u \le 1$). Denote the set $\{(u > 0) \cap C_1(Z)\}$ as E_u, and the spacewise set $\{x \mid (x, t) \in (u > 0) \cap C_1(Z)\}$ as $Q(t)$. Hence $|E_u| = \int_0^1 |Q(t)| dt$. We need to prove

$$\int_{C_{1/2}(Z)} u_+ \, dV_E \, ds \le N |E_u|^{1+1/(2n+2)} \tag{38}$$

To show (38) is true, it suffices to show that for any ϵ small enough, u_ϵ satisfies

$$\int_{C_{1/2}(Z)} u_\epsilon \, dV_E \, ds \le N |E_{u_\epsilon}|^{1+1/(2n+2)}. \tag{39}$$

The advantage of u_ϵ is that it's supported in $Q(t)$, and $0 \le u_\epsilon \le 1$. Then integration by parts implies the energy estimates in Lemma 4.5 holds true over $Q(t)$. Let η be the standard cut-off function in $C_1(Z)$ which vanishes near the parabolic boundary; Hölder's inequality and Lemma 4.5 imply

$$\int_B \eta u_\epsilon \, dV_E|_t \le |Q(t)|^{1/2} \left(\int_B \eta^2 u_\epsilon^2 \, dV_E \right)^{1/2} \Big|_t \le N E_{u_\epsilon}^{1/2} |Q(t)|^{1/2}. \tag{40}$$

We also have the following bootstrapping estimate on the same term.

$$
\begin{aligned}
(41)\quad \int_B \eta u_\epsilon \, dV_E &\le \left(\int_B |\eta u_\epsilon|^{2n/(2n-1)} \, dV_E\right)^{(2n-1)/(2n)} |Q(t)|^{1/(2n)} \\
&\le N\left(\int_B |\nabla(\eta u_\epsilon)| \, dV_E\right)|Q(t)|^{1/(2n)} \\
&\le N\left(\int_B |\nabla(\eta u_\epsilon)|^2 dV_E\right)^{1/2} |Q(t)|^{1/(2n)+1/2}, \\
&\qquad \text{since } \operatorname{supp} \nabla(\eta u_\epsilon) \subset \{u > 0\} \cap B.
\end{aligned}
$$

By (40), (41), Lemma 4.5, and the Fubini theorem, with the help of (25),

$$
\begin{aligned}
(42)\quad & \int_{-1}^{0} \int_B \eta u_\epsilon \, dV_E \, ds \\
&= \int_{-1}^{0} \left(\int_B \eta u_\epsilon \, dV_E\right)^{1/(n+1)} \left(\int_B \eta u_\epsilon \, dV_E\right)^{n/(n+1)} ds \\
&\le N E_{u_\epsilon}^{1/(2n+2)} \int_{-1}^{0} |Q(t)|^{(n+2)/(2n+2)} \left(\int_B |\nabla(\eta u_\epsilon)|^2 \, dV_E\right)^{n/(2n+2)} dt \\
&\le N|E_{u_\epsilon}|^{1/(2n+2)} \left(\int_{-1}^{0} |Q(t)| \, dt\right)^{(n+2)/(2n+2)} \left(\int_{-1}^{0} \int_B |\nabla(\eta u_\epsilon)|^2 \, dV_E \, dt\right)^{n/(2n+2)} \\
&\le N|E_{u_\epsilon}|^{1+1/(2n+2)}
\end{aligned}
$$

Since $\eta \equiv 1$ over $C_{1/2}(Z)$, the proof is complete. As we've seen, nothing in this proof involves more than Lemma 4.5 on the subsolutions. □

Proposition 5.3. *Suppose u is a weak subsolution to (18) in a cylinder $C_r(Y)$, $y \notin D$. Then*

$$
(43)\qquad u(Y) \le \frac{N}{|C_r|} \int_{C_r(Y)} u_+ \, dV_E
$$

Proof. The proof is exactly as in [Ferretti and Safonov 2013, Theorem 3.4]. The only thing worth mentioning is that we should deal with the singularity D. In [Ferretti and Safonov 2013], they consider the maximal point of $d^\gamma u$, where $\gamma = (2n+2)/p$ and d is the parabolic distance to the parabolic boundary of $C_r(Y)$. However, when a singularity is present, $d^\gamma u$ might not attain maximum away from D. To overcome this, we simply assume $u(Y) > 0$, and use the fact that there exists an almost maximal point away from D. Namely, there exists $X_0 = (x_0, t_0)$ such that $x_0 \notin D$ and

$$
(44)\qquad d^\gamma(X_0)u(X_0) \ge \frac{M}{2}, \qquad M \triangleq \sup_{C_r} d^\gamma u
$$

(we can assume $M > 0$ with out loss of generality). Then the rest of the proof is line by line as [Ferretti and Safonov 2013, line 13 to proof end, page 101], except the μ_1 on line 19 should correspond to $\beta_1 = 2^{-\gamma-2}$, because we have an additional $\frac{1}{2}$ in (44). □

Remark 5.4. As mentioned in [Ferretti and Safonov 2013, Remark 3.5], this proof does *not* involve explicitly the subequation (18). Instead, it only requires the growth lemma (Lemma 5.1). Thus the condition (12) is not involved explicitly in this proof.

Lemma 5.5 (slant cylinder lemma). *Suppose u is a weak subsolution to* (18) *in a slant cylinder* SC_r. *Suppose $u \le 0$ in $B_r \times \{T_0\}$. Then*

$$u(Y) \le (1-\lambda)\sup_{\mathrm{SC}_r} u^+, \tag{45}$$

where $\lambda \in (0, 1)$ depends on n, β, an upper bound, $|l|$, on $r|y_1 - y_0|/(T_1 - T_0)$, and upper bounds on $(T_1 - T_0)/r^2$, and K.

Proof. The first paragraph in the proof of Lemma 5.1 also applies here. By translation and rescaling (see Definition 4.1), without changing the parabolic slope, we can transform SC_r to a slant cylinder SC_1 with $r = 1$, $T_0 = 0$, $T_1 = T$, $y_0 = \{0\}$, and $y_1 = y$. We then pull back u and the matrix of the metric g on SC_r to "u" (by abuse of notation) and $\hat{g}$ on SC_1. Thus, u satisfies in SC_1 the following:

$$\frac{\partial u}{\partial t} - \Delta_{\hat{g}} u \le 0 \quad \text{in the sense of Definition 4.2, and} \tag{46}$$

$$\frac{g_E}{K} \le \hat{g} \le K g_E \quad \text{in } \mathrm{SC}_1 . \tag{47}$$

It suffices to prove (45) for u_ϵ. By rescaling, we can assume $u \le 1$ and $\sup_{\mathrm{SC}_1} u = 1$. Then $0 \le u_\epsilon \le 1 - \epsilon$ and $\sup_{\mathrm{SC}_1} u_\epsilon \ge 1 - 3\epsilon$. It suffices to derive an estimate for $v = -\log(1 - u_\epsilon)$ which is independent of ϵ. Since u_ϵ satisfies (46), v satisfies

$$\frac{\partial v}{\partial t} - \Delta_{\hat{g}} v \le -|\nabla_{\hat{g}} v|^2 \tag{48}$$

in the sense of Definition 4.2. Let $\underline{\eta}$ be the standard cut-off function in the Euclidean unit ball $B(1)$ which only depends on $|x|^2$. By (the proof of) Lemma 4.5 (replace the 0 on the right hand side of (18) by $-|\nabla_{\hat{g}} v|^2$), using $u_\epsilon \ge 0$, $u_\epsilon|_{t=0} = 0$, by abuse of notation with Lemma 4.5, we consider $\eta = \underline{\eta}[x - y(t)]$ and obtain (similarly to (23))

$$\int_{\mathbb{C}^n} v\eta^2 \, dV_g \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2}\int_{\mathbb{C}^n} \langle \nabla_g v, \nabla_g \eta^2 \rangle \, dV_g \, ds + \int_{t_1}^{t_2}\int_{\mathbb{C}^n} |\nabla_g v|^2 \eta^2 \, dV_g \, ds \\ \le \int_{t_1}^{t_2}\int_{\mathbb{C}^n} v \frac{\partial \eta^2}{\partial t} \, dV_g \, ds + K \int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\eta^2 \, dV_g \, ds. \tag{49}$$

We first estimate the term $\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\frac{\partial\eta^2}{\partial t}\,dV_g\,ds$. It's the same as in [Ferretti and Safonov 2013]. We note that $|\partial\eta^2/\partial t| \le |l||\nabla_E\eta^2|$ (Definition 4.1). Then

$$\left|\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\frac{\partial\eta^2}{\partial t}\,dV_g\,ds\right| \le |l|K^{2n}\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v|\nabla_E\eta^2|\,dV_E\,ds. \tag{50}$$

Using [Ferretti and Safonov 2013, lines 14–23, page 103], we obtain

$$\int_{\mathbb{C}^n} v|\nabla_E\eta^2|\,dV_E \le N\int_{\mathbb{C}^n}(|v|+|\nabla_E v|)\eta^2\,dV_E. \tag{51}$$

Then the Cauchy–Schwartz inequality and the quasi-isometric condition (47) imply

$$\begin{aligned}\left|\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\frac{\partial\eta^2}{\partial t}\,dV_g\,ds\right|& \\ \le \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n}&|\nabla_g v|^2\eta^2\,dV_g\,ds + N + N\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\eta^2\,dV_g\,ds.\end{aligned} \tag{52}$$

For the same reason we have

$$\left|\int_{t_1}^{t_2}\int_{\mathbb{C}^n}\langle\nabla_g v,\nabla_g\eta^2\rangle\,dV_g\,ds\right| \le \tfrac{1}{100}\int_{t_1}^{t_2}\int_{\mathbb{C}^n}|\nabla_g v|^2\eta^2\,dV_g\,ds + N. \tag{53}$$

Then (49), (52), and (53) imply

$$\int_{\mathbb{C}^n} v\eta^2\,dV_g\Big|_{t_1}^{t_2} \le N + N\int_{t_1}^{t_2}\int_{\mathbb{C}^n} v\eta^2\,dV_g\,ds. \tag{54}$$

Let $\int_{\mathbb{C}^n} v\eta^2\,dV_g|_t = I(t)$. Since $I(0)=0$, (54) implies $I(t)$ satisfies the assumption in Lemma 5.6. Hence Lemma 5.6 implies $I(t)\le N$ for all $t\in[0,T]$. Then Proposition 5.3 implies $v(Y)\le N$. Hence for some λ (as in Lemma 5.5) which is independent of ϵ, $u_\epsilon(Y)\le 1-2\lambda \le (1-\lambda)\sup_{SC_1} u_\epsilon$ when ϵ is small enough. Let $\epsilon\to 0$; the proof of (45) is complete. Again, nothing in this proof involves more than the energy estimates of the subsolutions. □

Lemma 5.6. *Suppose $I(t)$, $t\in[T_0,T_1]$ is an everywhere defined L^∞ function. Suppose $I(t)\ge 0$ for all t, $I(T_0)=0$, and*

$$I(t)\le I(t_1) + N_1\int_{t_1}^{t_2} I(s)\,ds + N_2, \quad \textit{for all } t_1, t_2 \textit{ and } t\in[t_1,t_2]. \tag{55}$$

Then there exists N depending on N_1, N_2, and T_1-T_0, such that $I(t)\le N$.

Proof. Choose a such that $a\le 1/(100N_1)$ and $(T_1-T_0)/a = k_0$ is an integer. Then for $k\le k_0-1$, we deduce

$$\max_{ka\le t\le(k+1)a} I(t) \le \tfrac{1}{2}\max_{ka\le t\le(k+1)a} I(t) + N_2 + I(ka),$$

then

$$\max_{ka \le t \le (k+1)a} I(t) \le 2N_2 + 2I(ka). \tag{56}$$

Since $I(T_0) = 0$, the proof is complete by induction. □

Theorem 5.7 (main growth theorem). *Suppose u is a weak subsolution to (18) in a cylinder $C_{2r}(Y)$. Suppose*

$$\frac{|\{u > 0\} \cap C_r(y, s - 3r^2)|}{|C_r(y, s - 3r^2)|} \le \frac{1}{2}. \tag{57}$$

Then

(58) $\sup_{C_r(Y)} u \le (1 - \lambda) \sup_{C_{2r}(Y)} u^+$, *where $\lambda \in (0, 1)$ depends on n, β, K.*

Assuming the growth lemma and slant cylinder lemma, the proof in [Ferretti and Safonov 2013] goes through without any change. We observe that

> except measure theory which does not involve the subequation (18), the proof of [loc. cit., Theorem 5.3] only depends on the fact that [loc. cit., Theorem 3.3] (Lemma 5.1) and [loc. cit., Lemma 4.1] (Lemma 5.5) hold true for any subsolution (with suitable conditions on initial value or level sets) in any scale.

Therefore, instead of directly quoting [loc. cit.], we sketch the proof of [loc. cit., Theorem 5.3] and Theorem 5.7 for the reader's convenience.

Proof sketch. We screen the center of a parabolic cylinder, i.e., we denote $C_r(Y)$ by C_r. Again, by the rescaling and translation invariance, it suffices to prove it assuming $s = 0$ and $r = 1$, i.e., we shall prove

$$\sup_{C_1} u \le (1 - \lambda) \sup_{C_2} u^+. \tag{59}$$

- In view of Definition 3.1, let C_1^0 denote the cylinder $C_1(y, -3)$ which is "earlier" in t than the target cylinder $C_1 = C_1(y, 0)$ on which we want to prove the estimate.
- Let $\Gamma_u \triangleq \{u \le 0\} \cap C_1^0$, i.e., the subset of C_1^0 on which u is nonpositive.
- Let $\mathcal{A}$ denote the set of all cylinders $C \subset C_2$ such that $|(C \cap \Gamma)/C| \ge 1 - \mu_2$, where μ_2 is the one in Lemma 5.1 (i.e., any such C has a "sufficient large portion" contained in Γ).

The dimension n, cone angle β, K, and the μ_2 in Lemma 5.1 determines 3 numbers $\epsilon_0 > 0$, $R_0 > 0$, and $0 < \beta_2 < 1$ with the following properties (which hold without condition (57)). If there is a cylinder $C_{r_0} \in \mathcal{A}$ with radius $r_0 \ge R_0$, then the growth lemma and slant cylinder lemma routinely yield the desired estimate (59).

If the radius of any cylinder $C \in \mathcal{A}$ is $< R_0$, then measure theory and still the two lemmas imply that

$$\left|\left\{u \leq (1-\beta_2)\sup_{C_2} u^+\right\} \cap C_1^0\right| \geq (1+\epsilon_0)|\Gamma_u|. \tag{60}$$

This means that increasing the level by a fixed amount enlarges the sublevel set by a fixed amount of measure.

Then we do an induction argument using the above paragraph successively. We define u_k inductively by

$$u_k = u_{k-1} - (1-\beta_2)\sup_{C_2} u_{k-1}, \quad u_0 = u \text{ (consequently } \Gamma_{u_0} = \Gamma_u).$$

Then $u_k = u - (1-\beta_2^k)\sup_{C_2} u^+$. We note that for any $k \geq 0$, u_k is also a subsolution, and the two alternative possibilities above (and including) (60) apply to any subsolution, particularly to u_k. Therefore for any integer k_0 such that $(1+\epsilon_0)^{k_0}/2 > 1$, either

(1) $\Gamma_{u_k} \geq (1+\epsilon_0)|\Gamma_{u_{k-1}}|$ for all $k \leq k_0$, or

(2) there exists a $k \leq k_0$ such that u_k satisfies the desired estimate (59).

In case 2 above, u also satisfies the desired estimate (59). In case 1 above, note that condition (57) says $|\Gamma_u| \geq |C_1^0|/2$; we get a contradiction because $\Gamma_{u_{k_0}}$ is a subset of C_1^0, but

$$|\Gamma_{u_{k_0}}| \geq (1+\epsilon_0)|\Gamma_{u_{k_0-1}}| \geq \cdots \geq (1+\epsilon_0)^{k_0}|\Gamma_u| \geq (1+\epsilon_0)^{k_0}\frac{|C_1^0|}{2} > |C_1^0|. \quad \square$$

Acknowledgement

The author is grateful to Gong Chen for valuable communications.

References

[Caffarelli and Cabré 1995] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, American Mathematical Society Colloquium Publications **43**, American Mathematical Society, Providence, RI, 1995. MR Zbl

[Cao and Chen 2012] H.-D. Cao and Q. Chen, "On locally conformally flat gradient steady Ricci solitons", *Trans. Amer. Math. Soc.* **364**:5 (2012), 2377–2391. MR Zbl

[Chen and Safonov 2017] G. Chen and M. Safonov, "On second order elliptic and parabolic equations of mixed type", *J. Funct. Anal.* **272**:8 (2017), 3216–3237. MR Zbl

[De Giorgi 1957] E. De Giorgi, "Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari", *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* (3) **3** (1957), 25–43. MR Zbl

[Ferretti and Safonov 2001] E. Ferretti and M. V. Safonov, "Growth theorems and Harnack inequality for second order parabolic equations", pp. 87–112 in *Harmonic analysis and boundary value problems* (Fayetteville, AR, 2000), edited by L. Capogna and L. Lanzani, Contemporary Mathematics **277**, American Mathematical Society, Providence, RI, 2001. Zbl

[Ferretti and Safonov 2013] E. Ferretti and M. V. Safonov, "Growth theorems and Harnack inequality for second order parabolic equations", updated version of [Ferretti and Safonov 2001], 2013, Available at http://www-users.math.umn.edu/∼safon002/NOTES/FS/GT1.pdf.

[Gilbarg and Trudinger 1977] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der Mathematischen Wissenschaften **224**, Springer, 1977. MR Zbl

[Guenancia and Păun 2016] H. Guenancia and M. Păun, "Conic singularities metrics with prescribed Ricci curvature: general cone angles along normal crossing divisors", *J. Differential Geom.* **103**:1 (2016), 15–57. MR Zbl

[Hamilton 1995] R. S. Hamilton, "The formation of singularities in the Ricci flow", pp. 7–136 in *Surveys in differential geometry* (Cambridge, MA, 1993), vol. 2, edited by S.-T. Yau, Int. Press, Cambridge, MA, 1995. MR Zbl

[Krylov and Safonov 1980] N. V. Krylov and M. V. Safonov, "A property of the solutions of parabolic equations with measurable coefficients", *Izv. Akad. Nauk SSSR Ser. Mat.* **44**:1 (1980), 161–175. In Russian; translated in *Math. USSR-Izv.* **16**:1 (1981), 151–164. MR Zbl

[Ladyzhenskaja et al. 1968] O. A. Ladyzhenskaja, V. A. Solonnikov, and N. N. Uralceva, *Linear and quasilinear equations of parabolic type*, Translations of Mathematical Monographs **23**, American Mathematical Society, Providence, RI, 1968. MR Zbl

[Lieberman 1996] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific, River Edge, NJ, 1996. MR Zbl

[Moser 1964] J. Moser, "A Harnack inequality for parabolic differential equations", *Comm. Pure Appl. Math.* **17** (1964), 101–134. MR Zbl

[Nash 1958] J. Nash, "Continuity of solutions of parabolic and elliptic equations", *Amer. J. Math.* **80** (1958), 931–954. MR Zbl

[Wang 2016] Y. Wang, "Smooth approximations of the conical Kähler–Ricci flows", *Math. Ann.* **365**:1-2 (2016), 835–856. MR Zbl

Received August 6, 2016. Revised January 8, 2019.

YUANQI WANG
DEPARTMENT OF MATHEMATICS
STONY BROOK UNIVERSITY
STONY BROOK, NY
UNITED STATES
ywang@scgp.stonybrook.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.329

DEFORMATION OF MILNOR ALGEBRAS

ZHENJIAN WANG

We investigate deformations of Milnor algebras of smooth homogeneous polynomials, and prove in particular that any smooth degree d homogeneous polynomial in $n+1$ variables that is not of Sebastiani–Thom type is determined by the degree k homogeneous component of its Jacobian ideal for any $d-1 \le k \le (n+1)(d-2)$. Our results generalize the previous result on the reconstruction of a homogeneous polynomial from its Jacobian ideal.

1. Introduction

The classical theory of variation of Hodge structures for smooth hypersurfaces in a complex projective space gives a variation of Milnor algebras of homogeneous polynomials. The celebrated generic Torelli theorem for hypersurfaces is almost reduced to the study of injectivity of some mappings concerning the deformation of Milnor algebras, see [Voisin 2003, Subsection 6.3.2, p. 179; Donagi 1983]. Nevertheless, the homogeneous components of the Milnor algebra involved there are of specific degrees, namely degrees of the form $pd-n-1$. In this note, we will investigate homogeneous components of all degrees of the Milnor algebra.

To fix notation, let $S=\mathbb{C}[x_0,\dots,x_n]$ be the homogeneous coordinate ring of the complex projective space $\mathbb{P}^n$,

$$S=\bigoplus_{d=0}^{\infty} S_d,$$

where S_d is the vector space of homogeneous polynomials of degree d. Given a homogeneous polynomial $f\in S_d$, denote by $J(f)$ the Jacobian ideal of f:

$$J(f)=(\partial f/\partial x_0, \partial f/\partial x_1, \dots, \partial f/\partial x_n).$$

Set $M(f)=S/J(f)$, known as the Milnor algebra of f. The algebra $M(f)$ has the natural grading

$$M(f)=\bigoplus_{k=0}^{\infty} M(f)_k,$$

MSC2010: primary 14A25; secondary 14C34, 14J70.
Keywords: Homogeneous polynomials, Milnor algebras, Macaulay inverse system.

where $M(f)_k = S_k/(J(f) \cap S_k)$.

We say that $f \in \mathbb{P}(S_d)$ is a *smooth* polynomial if the hypersurface $V_f : f = 0$ in $\mathbb{P}^n$ is a smooth hypersurface. The discriminant defines a divisor $\Delta \subset \mathbb{P}(S_d)$ such that the complement $\mathbb{P}(S_d)_\Delta$ parametrizes smooth homogeneous polynomials of degree d.

We say that a polynomial $f \in S_d$ is of *Sebastiani–Thom type* (ST type) or a *direct sum* if f can be represented as

$$f(x_0, \dots, x_n) = f_1(x_0, \dots, x_\ell) + f_2(x_{\ell+1}, \dots, x_n) \tag{1}$$

for a choice of homogeneous coordinates $\{x_i\}_{i=0}^n$ of $\mathbb{P}^n$ and some $0 \le \ell < n$; see [Ueda and Yoshinaga 2009; Wang 2015; Buczyńska et al. 2015]. For various characterizations of polynomials of ST type, we refer to [Fedorchuk 2020]. Denote by $\mathscr{U} \subset \mathbb{P}(S_d)_\Delta$ the set of all smooth homogeneous polynomials that are *not* of ST type.

It is well-known that $\dim M(f)_k$ does not depend on the concrete equation of f for smooth f (see for instance [Dimca 1987, Proposition 7.22, p. 108]); we denote this dimension by $a_{n,d}(k)$. Let $\mathrm{Grass}(S_k, a_{n,d}(k))$ be the Grassmannian parametrizing all $a_{n,d}(k)$ dimensional *quotient spaces* of S_k, then we have the following map

$$\varphi_k : \mathbb{P}(S_d)_\Delta \to \mathrm{Grass}(S_k, a_{n,d}(k)), \tag{2}$$

defined by $f \mapsto M(f)_k$.

More generally, denote $\mathrm{Grass}(n+1, S_{d-1})$ the Grassmannian of *linear subspaces* of dimension $n+1$ of the space of degree $d-1$ homogeneous polynomials S_{d-1}. Following [Fedorchuk 2017, Subsection 1.2], given $W \in \mathrm{Grass}(n+1, S_{d-1})$, we form the ideal $I_W := (W)$ and the quotient algebra $M(W) = S/I_W$. Let $g_0, \dots, g_n$ be a basis of W, then the sequence $g_0, \dots, g_n$ is a regular sequence if and only if I_W is a complete intersection ideal if and only if $M(W)$ is a standard local Artinian Gorenstein algebra of socle degree $T := (n+1)(d-2)$ if and only if the resultant of $g_0, \dots, g_n$ is nonzero; we refer to [Gelfand et al. 1994, Chapter 13] for the definition and basic properties of the resultant. Therefore, there exists a divisor $\mathrm{Res} \subset \mathrm{Grass}(n+1, S_{d-1})$ parametrizing all W such that I_W is not a complete intersection ideal. We denote by $\mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$ the affine complement of Res. For more discussions about ideals of the form I_W, see [Fedorchuk 2017, Subsection 1.2] .

For $W \in \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$, we have $\dim M(W)_k = a_{n,d}(k)$ by [Dimca 1987, Proposition 7.22, p. 108]. Hence the assignment $W \mapsto M(W)_k$ defines a map

$$\Phi_k : \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}} \to \mathrm{Grass}(S_k, a_{n,d}(k)). \tag{3}$$

Our first result is the following theorem.

Theorem 1.1. *For any $d-1 \leq k \leq T = (n+1)(d-2)$, the map Φ_k is an immersion, that is, it is injective and the differential* $\mathrm{d}\Phi_k$ *is also injective at any point of* $\mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$.

Using our previous result on determination of a polynomial by its Jacobian ideal (see [Wang 2015, Theorem 1.1; Ueda and Yoshinaga 2009, Lemma 3]), we further prove the following result.

Theorem 1.2. *For $d-1 \leq k \leq T = (n+1)(d-2)$, the restriction of the map φ_k* (defined in (2)) *to $\mathscr{U}$,*

$$\varphi_k : \mathscr{U} \to \mathrm{Grass}(S_k, a_{n,d}(k)),$$

is an immersion.

In particular, we have that a smooth homogeneous polynomial $f \in \mathscr{U}$ can be reconstructed from the degree k homogeneous component of its Jacobian ideal $J(f)$ for any k satisfying $d-1 \leq k \leq T = (n+1)(d-2)$. This gives a generalization of the previous results, in the case of smooth polynomials, in [Wang 2015; Ueda and Yoshinaga 2009].

We will also investigate the map φ_k defined in (2) and discuss its fibers over $\varphi_k(f)$ for homogeneous polynomials f that are of ST type, see Section 4.

Our results are related to the problem of characterizing the hypersurface singularity $\widehat{V}_f = \{x \in \mathbb{C}^{n+1} \mid f(x) = 0\}$ at the origin 0 of $\mathbb{C}^{n+1}$ using the Milnor algebra $M(f)$. In fact, the characterization problem of a singularity by its algebraic data can be proposed and solved in a much more general setting, see [Gaffney and Hauser 1985]. As a general philosophy in singularity theory, the Milnor algebra $M(f)$ is closely connected to the topology and geometry of the hypersurface singularity $(\widehat{V}_f, 0) \subset (\mathbb{C}^{n+1}, 0)$. Instead of giving characterizations of a singularity by *algebras* or *modules* derived from it, as in [Gaffney and Hauser 1985], here using Theorem 1.2, we can give a characterization of the *isolated* hypersurface singularity $(\widehat{V}_f, 0)$ just by a single homogeneous component $M(f)_k$ of the Minor algebra $M(f)$ for any $d-1 \leq k \leq (n+1)(d-2)$ with $d = \deg f$. This conclusion can obviously be extended to an isolated complete intersection singularity by using Theorem 1.1.

Of course, our results concern only the case when the hypersurface $(\widehat{V}_f, 0)$ is an isolated singularity. It is natural to extend these results to the case where the singularities of the hypersurface $\widehat{V}_f$ have positive dimension. However, the tools used in this note cannot be directly applied in the extended case because they depend heavily on the condition that $M(f)$ is a local Artinian Gorenstein algebra which holds only when 0 is an isolated singularity of $\widehat{V}_f$. In addition, heuristically, the results in [Gaffney and Hauser 1985] also show that any possible extension must be more complicated and more technical than our results above; see also the results in [Wang 2015] concerning the nonsmooth homogeneous polynomials.

We hope the results in this note can be applied to the study of Lefschetz properties for Milnor algebras. In fact, this is an important impetus to our present work. As is well-known, the strong Lefschetz property holds for $M(f)$ for a generic f. Our naïve idea is to investigate the Lefschetz properties by deforming the Milnor algebras. For an excellent exposition for the Lefschetz properties, we refer to [Harima et al. 2003; Migliore and Nagel 2013; Harima et al. 2013]. In addition, the strong Lefschetz property for $M(f)$ where f is of ST type can be reduced to that where f is not of ST type (see [Harima and Watanabe 2007, Theorem 3.10; Harima et al. 2013, Proposition 3.77, p. 137]), since $M(f)$ is the tensor product of $M(f_1)$ and $M(f_2)$ when f is represented as in (1). This is an important reason why we specifically investigate the set $\mathscr{U}$ in this paper; another reason is about the determination of a homogeneous polynomial by its Jacobian ideal, see the proof of Corollary 3.2 below.

2. Polar paring and Macaulay inverse systems

2.1. ***Polar pairing.*** Let $S = \mathbb{C}[x_0, \dots, x_n]$ and $R = \mathbb{C}[z_0, \dots, z_n]$ be two polynomial rings. There is a natural action of S on R by the "polar paring"

$$S \times R \to R$$

defined by

$$\big(f(x_0, \dots, x_n), Q(z_0, \dots, z_n)\big) \mapsto f \cdot Q := f(\partial/\partial z_0, \dots, \partial/\partial z_n) Q(z_0, \dots, z_n).$$

It induces perfect parings $S_\rho \times R_\rho \to \mathbb{C}$ for every $\rho \in \mathbb{N}$. In particular, for $f \in S_\rho$ written as

$$f(x_1, \dots, x_n) = \sum_{|\alpha|=\rho} a_\alpha x^\alpha$$

and $Q \in R_\rho$ written as

$$Q(z_0, \dots, z_n) = \sum_{|\alpha|=\rho} b_\alpha z^\alpha$$

we have

$$f \cdot Q = \sum_{|\alpha|=\rho} \alpha! a_\alpha b_\alpha.$$

Define the polynomial $q = \sum_{|\alpha|=\rho} b_\alpha x^\alpha$, or equivalently

$$q(x_0, \dots, x_n) = Q(x_0, \dots, x_n),$$

and define the inner product of f and q by

$$\langle f, q\rangle = \sum_{|\alpha|=\rho} \alpha! a_\alpha b_\alpha = f \cdot Q. \tag{4}$$

For any linear space $E \subset S_\rho$, with respect to the above inner product $\langle \cdot, \cdot \rangle$, we have its orthogonal complement, denoted by $E^\perp$.

2.2. ***Macaulay inverse system.*** Let $I \subset S$ be a Gorenstein ideal and ν the socle degree of the algebra $\mathcal{A} = S/I$. Recall that a (homogeneous) Macaulay inverse system of $\mathcal{A}$ is an element $Q_{\mathcal{A}} \in \mathbb{P}(R_\nu)$ such that I is equal to the apolar ideal $Q_{\mathcal{A}}^\perp$, namely,

$$I = \{ f \in S \mid f \cdot Q_{\mathcal{A}} = 0 \}$$

(see [Iarrobino and Kanev 1999, Lemma 2.12] or [Eisenbud 1995, Exercise 2.17]).

Let $W = \operatorname{span}\langle g_0, \dots, g_n \rangle$ such that $W \in \operatorname{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$, the *associated form* $\boldsymbol{A}_W := \boldsymbol{A}(g_0, \dots, g_n) \in \mathbb{P}(R_T)$ (recall that $T = (n+1)(d-2)$) gives the Macaulay inverse system for $S_W = S/I_W$; see [Alper and Isaev 2018, Proposition 2.1]. We write

$$\boldsymbol{A}_W = \sum_{|\alpha|=T} c_\alpha z^\alpha.$$

In this case, define $\boldsymbol{B}_W \in \mathbb{P}(S_T)$ by

$$\boldsymbol{B}_W = \sum_{|\alpha|=T} c_\alpha x^\alpha.$$

The polynomial $\boldsymbol{B}_W$, by definition, determines and is determined by $\boldsymbol{A}_W$. Moreover, by the definition of Macaulay inverse systems, we have that $(I_W)_T^\perp = \mathbb{C}\boldsymbol{B}_W$, namely, the line $\mathbb{C}\boldsymbol{B}_W$ is exactly the orthogonal complement of $(I_W)_T$ with respect to the inner product $\langle \cdot, \cdot \rangle$ on S_T. Therefore, $\boldsymbol{A}_W \in \mathbb{P}(R_T)$ is uniquely determined by $(I_W)_T$.

Lemma 2.3. *For two points $U, W \in \operatorname{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$, the following statements are equivalent*:

(1) $U = W$.

(2) $I_U = I_W$.

(3) *For any k satisfying $d-1 \le k \le T = (n+1)(d-2)$, we have $(I_U)_k = (I_W)_k$.*

(4) *For some k satisfying $d-1 \le k \le T$, we have $(I_U)_k = (I_W)_k$.*

(5) $(I_U)_T = (I_W)_T$.

Proof. It is obvious that (1), (2), (3) are all equivalent and (3) implies (4).

(4)⇒(5): Since I_U is generated by polynomials all of which have degree $d-1$, we have that $(I_U)_T$ is the image of $S_{T-k} \times (I_U)_k$ under the multiplication map $S_{T-k} \times S_k \to S_T$. Hence $(I_U)_T = (I_W)_T$ whenever $(I_U)_k = (I_V)_k$ for $d-1 \le k \le T$.

(5)⇒(1): This is clear once we note that A_U can be uniquely determined by $(I_U)_T$, and I_U is the apolar ideal $A_U^{\perp}$. □

Recall that as it is shown in the introduction, $a_{n,d}(k) = \dim M(f)_k$ for any $f \in \mathbb{P}(S_d)_\Delta$, which is also the dimension of $S_k/(I_W)_k$ for any $W \in \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$. Set $b_{n,d}(k) = \dim S_k - a_{n,d}(k)$ and let $\mathrm{Grass}(b_{n,d}(k), S_k)$ be the Grassmannian parametrizing all $b_{n,d}(k)$ dimensional linear *subspaces* of S_k. For a subspace $E \subset S_k$ of dimension $b_{n,d}(k)$, we obtain the quotient space S_k/E of dimension $a_{n,d}(k)$; and the mapping $S_k \supset E \mapsto S_k/E$ clearly defines an isomorphism between the Grassmannians $\mathrm{Grass}(b_{n,d}(k), S_k)$ and $\mathrm{Grass}(S_k, a_{n,d}(k))$. Then to prove Theorem 1.1, it suffices to prove the following theorem.

Theorem 2.4. *For any $d-1 \le k \le T$, the assignment $W \mapsto (I_W)_k$ defines an immersion*

$$\Psi_k : \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}} \to \mathrm{Grass}(b_{n,d}(k), S_k), \tag{5}$$

that is Ψ_k is injective and the differential $\mathrm{d}\Psi_k$ is also injective at any point of $\mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$.

Proof. The injectivity of Ψ_k follows from the equivalence (1)⇔(4) in Lemma 2.3.

Given $W \in \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$ such that $W = \mathrm{span}\langle g_0, \ldots, g_n\rangle$. For any $h \in T_W(\mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}) \simeq \mathrm{Hom}(W, S_{d-1}/W)$, choose $h_i \in S_{d-1}$ such that $h(g_i) = h_i \mod W$ for $i = 0, \ldots, n$. Then if $h \in \mathrm{Ker}(\mathrm{d}\Psi_k)_W$, we have $(\mathrm{d}\Psi_k)_W(h) = 0$ as an element in $\mathrm{Hom}((I_W)_k, S_k/(I_W)_k)$. A direct computation gives that

$$(\mathrm{d}\Psi_k)_W(h)((I_W)_k) = (I_H)_k + (I_W)_k \mod (I_W)_k,$$

where $H = \mathrm{span}\langle h_0, \ldots, h_n\rangle$. It follows from $(\mathrm{d}\Psi_k)_W(h) = 0$ that $(I_H)_k \subset (I_W)_k$.

For $t \in \mathbb{C}^*$ and $|t|$ sufficiently small, we have that

$$W_t := \mathrm{span}\langle g_0 + th_0, \ldots, g_n + h_n\rangle$$

satisfies $W_t \in \mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}}$. It then follows from $(I_H)_k \subset (I_W)_k$ that $(I_{W_t})_k \subset (I_W)_k$, hence $(I_{W_t})_k = (I_W)_k$ because $\dim(I_{W_t})_k = b_{n,d}(k) = \dim(I_W)_k$. Therefore $W_t = W$ by (4)⇒(1) in Lemma 2.3. It follows that $h_i \in W$ for $i = 0, \ldots, n$ and thus $h = 0$ as an element of $T_W(\mathrm{Grass}(n+1, S_{d-1})_{\mathrm{Res}})$.

Since h can be arbitrarily chosen, $(\mathrm{d}\Psi_k)_W$ is injective. We are done. □

3. Variation of Milnor algebras

3.1. ***Polynomials not of ST type.*** Recall that $\mathscr{U} \subset \mathbb{P}(S_d)$ denotes the space of smooth homogeneous polynomials of degree d that are not of ST type, or equivalently, the space of smooth hypersurfaces whose defining equations are not of ST type. From the proof of [Wang 2015, Corollary 6.1], we have that $\mathscr{U}$ is a Zariski open subset of $\mathbb{P}(S_d)_\Delta$.

For $f \in \mathscr{U}$, recall that $J(f)$ denotes the Jacobian ideal of f and $M(f) = S/J(f)$ the Milnor algebra. For $k \geq d-1$, we denote by $E_k(f) = J(f) \cap S_k$. Then $\dim E_k(f) = b_{n,d}(k) = \dim S_k - a_{n,d}(k)$ is independent of $f \in \mathscr{U}$. Moreover, since $\partial f/\partial x_0, \ldots, \partial f/\partial x_n$ form a regular sequence and $J(f) = I_{E_{d-1}(f)}$, from Lemma 2.3, we immediately get the following corollary.

Corollary 3.2. *Given $f, g \in \mathbb{P}(S_d)$ and $f \in \mathscr{U}$, the following conditions are equivalent*:

(1) $E_{d-1}(f) = E_{d-1}(g)$.

(2) $J(f) = J(g)$.

(3) *For any k satisfying $d-1 \leq k \leq T = (n+1)(d-2)$, we have $E_k(f) = E_k(g)$.*

(4) *For some k satisfying $d-1 \leq k \leq T$, we have $E_k(f) = E_k(g)$.*

(5) $E_T(f) = E_T(g)$.

(6) $f = g$.

Proof. The equivalences among the first five statements follow from Lemma 2.3; we here just note that any one of these conditions implies that $E_{T+1}(g) = S_{T+1}$, hence g is also smooth and thus $J(g)$ is a complete intersection ideal.

The equivalence (1)$\Leftrightarrow$(6) follows from [Wang 2015, Theorem 1.1] or [Ueda and Yoshinaga 2009, Lemma 3]. □

Now we are ready to prove Theorem 1.2. Similar to the proof of Theorem 1.1, it is sufficient to prove the following theorem.

Theorem 3.3. *For any $d-1 \leq k \leq T$, the assignment $f \mapsto E_k(f)$ defines an immersion*

$$\psi_k : \mathscr{U} \to \mathrm{Grass}(b_{n,d}(k), S_k),$$

Namely, ψ_k is injective and its differential $\mathrm{d}\psi_k$ is also injective at any point $f \in \mathscr{U}$.

Proof. By the equivalence of (4) and (6) in Corollary 3.2, we have that ψ_k is injective.

We will not distinguish an element $f \in \mathbb{P}(S_d)$ and its lifting in S_d. For $f \in \mathscr{U}$, we have $T_f\mathscr{U} = T_f\mathbb{P}(S_d) \simeq \mathrm{Hom}(\mathbb{C}f, S_d/\mathbb{C}f)$. The mapping $\mathrm{Hom}(\mathbb{C}f, S_d/\mathbb{C}f) \ni$

$\eta \mapsto \eta(f) \in S_d/\mathbb{C}f$ then gives an identification $T_f\mathscr{U} \simeq S_d/\mathbb{C}f$. With the help of this identification, the differential of ψ_k at $f \in \mathscr{U}$ is given by

$$(\mathrm{d}\psi_k)_f : T_f\mathscr{U} \simeq S_d/\mathbb{C}f \to \mathrm{Hom}(E_k(f), S_k/E_k(f)).$$

Therefore, we have $(\mathrm{d}\psi_k)_f(h) = 0$ as an element of $\mathrm{Hom}(E_k(f), S_k/E_k(f))$ for any $h \in \mathrm{Ker}(\mathrm{d}\psi_k)_f$. Represent h by an element in $S_d/\mathbb{C}f$, and lift it to an element in S_d which is still denoted by h. A direct computation gives that

$$(\mathrm{d}\psi_k)_f(h)(E_k(f)) = E_k(h) + E_k(f) \mod E_k(f).$$

Hence it follows from $(\mathrm{d}\psi_k)_f(h) = 0$ that $E_k(h) \subset E_k(f)$.

From the semicontinuity of the dimension of $E_k(f)$ with respect to $f \in S_d$, we obtain that for a small positive number $\epsilon > 0$ and for any $t \in \mathbb{C}$ such that $|t| < \epsilon$, the following hold:

(i) $\dim E_k(f + th) = b_{n,d}(k) = \dim E_k(f)$;

(ii) $E_k(f + th) \subset E_k(f)$.

Hence $E_k(f+th) = E_k(f)$ for any $|t| < \epsilon$. In particular, choosing $t_0 \neq 0$ satisfying $|t_0| < \epsilon$, we have $E_k(f + t_0h) = E_k(f)$. Using (4)$\Leftrightarrow$(6) in Corollary 3.2 again, we deduce that $f + t_0h = f$ in $\mathbb{P}(S_d)$, hence $h = f$ in $\mathbb{P}(S_d)$ which implies that the chosen tangent vector $h \in \mathrm{Ker}(\mathrm{d}\psi_k)_f$ is equal to zero. Therefore $(\mathrm{d}\psi_k)_f$ is also injective. □

The above proof also gives the following corollary, which is interesting in its own right; compare with Corollary 3.2.

Corollary 3.4. *Given $f \in \mathscr{U}$ and $h \in \mathbb{P}(S_d)$. Suppose $E_k(h) \subset E_k(f)$ for some $d - 1 \le k \le T$, then $h = f$.*

4. Polynomials of Sebastiani–Thom type

In this section, we give a brief discussion about the fibers of the map φ_k in (2) over $\varphi_k(f)$ for a polynomial f of ST type.

By [Fedorchuk 2020, Proposition 4.8 or Corollary 3.15], a smooth homogeneous polynomial $f \in S_d$ admits a unique maximally fine "direct sum decomposition"

$$\text{(6)}\quad f(x_0, \dots, x_n) = f_1(x_0, \dots, x_{n_1-1}) + f_2(x_{n_1}, \dots, x_{n_2}) + \cdots + f_s(x_{n_{s-1}}, \dots, x_n),$$

for a choice of linear coordinates $\{x_i\}_{i=0}^n$, where $0 \le n_1 \le n_2 \le \cdots \le n_{s-1} \le n$ and none of the f_j are of ST type. In addition, if $g \in S_d$ satisfies $E_{d-1}(g) \subset E_{d-1}(f)$, then necessarily, g is of the following form

$$\text{(7)}\quad g = \lambda_1 f_1 + \cdots + \lambda_s f_s, \quad \lambda_i \in \mathbb{C},$$

see [Fedorchuk 2020, Corollary 3.12]. In particular, if g is also smooth, then all the λ_j in (7) are nonzero. With these results at hand, we prove the following theorem.

Theorem 4.1. *For any $d-1 \le k \le T = (n+1)(d-1)$ and any $f \in \mathbb{P}(S_d)_\Delta$, the fiber over $\varphi_k(f)$ of φ_k defined in (2), namely,*

$$\varphi_k : \mathbb{P}(S_d)_\Delta \to \mathrm{Grass}(S_k, a_{n,d}(k)),$$

is

$$\varphi_k^{-1}(\varphi_k(f)) = \{\lambda_1 f_1 + \cdots + \lambda_s f_s \mid \lambda_i \in \mathbb{C}^*, i = 1, \ldots, s\}.$$

Proof. It is obvious that for the λ_j nonzero, the polynomial $\lambda_1 f_1 + \cdots + \lambda_s f_s$ is smooth and is mapped under φ_k to $\varphi_k(f)$.

Conversely, if $g \in \mathbb{P}(S_d)_\Delta$ satisfies $\varphi_k(g) = \varphi_k(f)$, then we have $E_k(g) = E_k(f)$. It follows by (4)$\Rightarrow$(1) in Lemma 2.3 that $E_{d-1}(g) = E_{d-1}(f)$. Hence by [Fedorchuk 2020, Corollary 3.12], we have that g is of the form $\lambda_1 f_1 + \cdots + \lambda_s f_s$ for nonzero λ_j's. □

In conclusion, for the map φ_k, we can explicitly and completely determine all the fibers.

Acknowledgements

We would like to thank Professor Herwig Hauser for the reference [Gaffney and Hauser 1985]. We thank an anonymous referee for useful remarks and suggestions. We also thank the Yau Mathematical Sciences Center for financial support and stimulating working atmosphere.

References

[Alper and Isaev 2018] J. Alper and A. Isaev, "Associated forms and hypersurface singularities: the binary case", *J. Reine Angew. Math.* **745** (2018), 83–104. MR Zbl

[Buczyńska et al. 2015] W. Buczyńska, J. Buczyński, J. Kleppe, and Z. Teitler, "Apolarity and direct sum decomposability of polynomials", *Michigan Math. J.* **64**:4 (2015), 675–719. MR Zbl

[Dimca 1987] A. Dimca, *Topics on real and complex singularities: an introduction*, Vieweg, Braunschweig, 1987. MR Zbl

[Donagi 1983] R. Donagi, "Generic Torelli for projective hypersurfaces", *Compositio Math.* **50**:2-3 (1983), 325–353. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR Zbl

[Fedorchuk 2017] M. Fedorchuk, "GIT semistability of Hilbert points of Milnor algebras", *Math. Ann.* **367**:1-2 (2017), 441–460. MR Zbl

[Fedorchuk 2020] M. Fedorchuk, "Direct sum decomposability of polynomials and factorization of associated forms", *Proc. Lond. Math. Soc.* (3) **120**:3 (2020), 305–327. MR

[Gaffney and Hauser 1985] T. Gaffney and H. Hauser, "Characterizing singularities of varieties and of mappings", *Invent. Math.* **81**:3 (1985), 427–447. MR Zbl

[Gelfand et al. 1994] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Birkhäuser, Boston, 1994. MR Zbl

[Harima and Watanabe 2007] T. Harima and J. Watanabe, "The strong Lefschetz property for Artinian algebras with non-standard grading", *J. Algebra* **311**:2 (2007), 511–537. MR Zbl

[Harima et al. 2003] T. Harima, J. C. Migliore, U. Nagel, and J. Watanabe, "The weak and strong Lefschetz properties for Artinian K-algebras", *J. Algebra* **262**:1 (2003), 99–126. MR Zbl

[Harima et al. 2013] T. Harima, T. Maeno, H. Morita, Y. Numata, A. Wachi, and J. Watanabe, *The Lefschetz properties*, Lecture Notes in Mathematics **2080**, Springer, 2013. MR Zbl

[Iarrobino and Kanev 1999] A. Iarrobino and V. Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics **1721**, Springer, 1999. MR Zbl

[Migliore and Nagel 2013] J. Migliore and U. Nagel, "Survey article: a tour of the weak and strong Lefschetz properties", *J. Commut. Algebra* **5**:3 (2013), 329–358. MR Zbl

[Ueda and Yoshinaga 2009] K. Ueda and M. Yoshinaga, "Logarithmic vector fields along smooth divisors in projective spaces", *Hokkaido Math. J.* **38**:3 (2009), 409–415. MR Zbl

[Voisin 2003] C. Voisin, *Hodge theory and complex algebraic geometry*, vol. II, Cambridge Studies in Advanced Mathematics **77**, Cambridge University Press, 2003. MR Zbl

[Wang 2015] Z. Wang, "On homogeneous polynomials determined by their Jacobian ideal", *Manuscripta Math.* **146**:3-4 (2015), 559–574. MR Zbl

Received July 17, 2019. Revised October 4, 2019.

ZHENJIAN WANG
YAU MATHEMATICAL SCIENCES CENTER
TSINGHUA UNIVERSITY
BEIJING
CHINA
wzhj@mail.tsinghua.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.339

PRESERVATION OF LOG-SOBOLEV INEQUALITIES UNDER SOME HAMILTONIAN FLOWS

BO XIA

We prove that the probability measure induced by the BBM flow satisfies a logarithmic Sobolev type inequality. Precisely, we suppose the initial data u_0 induces a Gaussian measure on H^s with $s \in \left[1-\frac{\gamma}{2}, \frac{\gamma}{2}\right]$ for $\gamma \in \left(\frac{3}{2}, 2\right]$. Then the induced measure ν under BBM flow satisfies, for any ε small enough,

$$\mathbb{E}_\nu\left[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\right] \le C_\varepsilon\left(\mathbb{E}_\nu\left[|\nabla f|^{2+\varepsilon}_{H^{s+\gamma/2}}\right]\right)^{\frac{2}{2+\varepsilon}},$$

where C_ε is an ε-dependent constant.

1. Introduction

(Gaussian) Logarithmic Sobolev inequalities were first introduced by Gross [1975] for Gaussian measures on finite-dimensional spaces. They turned out to be effective tools for analysis on manifolds. For infinite-dimensional manifolds, thanks to its dimensionless character, logarithmic Sobolev inequalities seem to be similar to these classical ones. Indeed, logarithmic Sobolev inequalities were proved for infinite-dimensional spaces equipped with Gaussian measures [Da Prato 2006], for some infinite-dimensional spaces equipped with certain weighted Gaussian measures [Ledoux 2001], and even for some measures induced by certain transformations [Üstünel 2010] (these inequalities were also established on path spaces [Hsu 1997] and loop groups [Gross 1991]). Here we establish logarithmic Sobolev-type inequalities for measures induced by some flows associated to BBM equation.

Consider the generalized BBM model equation

$$\text{(1-1)}\quad \begin{cases}\partial_t u + \partial_t |\partial_x|^\gamma u + \partial_x(u+u^2) = 0,\\ u(0) = u_0,\end{cases} \qquad \text{where } u:(t,x)\in\mathbb{R}\times\mathbb{T}\mapsto u(t,x)\in\mathbb{R}.$$

One can find in Section 4A that (1-1) is quasiglobally well-posed. Precisely, for fixed $\gamma \in \left(\frac{3}{2}, 2\right]$ and $s \in \left[1-\frac{\gamma}{2}, \frac{\gamma}{2}\right]$, if $u_0 \in H^s$ and $\widetilde{T} \in (0,\infty)$, then (1-1) has a

I am grateful to Nicolas Burq for his encouragement on this subject. I am also grateful to Chenmin Sun and Jiqiang Zheng for discussions concerning this subject. I thank the anonymous referee for useful suggestions.
MSC2010: 35B99.
Keywords: Hamiltonian flows, BBM equation, log-Sobolev inequalities.

solution in $C([0,\widetilde{T}];H^s)$. We denote by $\Phi(t)$ the flow associated to (1-1). Now suppose that the initial data u_0 is given by

$$u_0=\phi_s(\omega,x)=\sum_{n\neq 0}\frac{g_n(\omega)}{|n|^{s+\gamma/2}}e^{inx},$$

where $g_n=\overline{g_{-n}}$ and $(g_n)_{n>0}$ is a sequence of independent standard complex Gaussian random variables on some proper probability space $(\Omega,\mathcal{A},\mathbb{P})$. Then the map $\omega\mapsto u_0$ induces a Gaussian measure on H^s, which we denote by μ_s. The classical theory asserts that μ_s satisfies a logarithmic Sobolev inequality (see [Üstünel 2010] for a proof in the case of Wiener space)

$$\mathbb{E}_{\mu_s}\left[f^2\log\frac{f^2}{\mathbb{E}_{\mu_s}[f^2]}\right]\le C\mathbb{E}_{\mu_s}[|\nabla f|^2_{H^{s+\gamma/2}}]$$

for any $f\in W^{1,2}(H^s,\mathbb{R})$, where C is a universal constant. In this manuscript, we consider whether or not the measure $\nu:=(\Phi(t))_*\mu_s$ satisfies inequalities of this type. We here give some partial answer to this question and our main result is:

Theorem 1.1. *Let $\gamma\in\left(\frac{3}{2},2\right]$ and $s\in\left[1-\frac{\gamma}{2},\frac{\gamma}{2}\right]$. Assume also $t\in[0,\widetilde{T}]$ is fixed and denote $T:=\Phi(t)$. Then there exists some constant C, such that the induced measure $\nu=T_*\mu_s$ satisfies for functions defined on H^s*

$$\mathbb{E}_{\nu}\left[f^2\log\frac{f^2}{\mathbb{E}_{\nu}[f^2]}\right]\le C\mathbb{E}_{\mu_s}\big[|\nabla f\circ T|^2_{H^{s+\gamma/2}}(1+\|\cdot\|^2_{H^s})\big].$$

Furthermore, by invoking Fernique's theorem, for $\varepsilon>0$, there exists $C_\varepsilon>0$ such that ν satisfies a log-Sobolev type inequality with a loss of integrability

$$\mathbb{E}_{\nu}\left[f^2\log\frac{f^2}{\mathbb{E}_{\nu}[f^2]}\right]\le C_\varepsilon\big(\mathbb{E}_{\nu}\big[|\nabla f|^{2+\varepsilon}_{H^{s+\gamma/2}}\big]\big)^{\frac{2}{2+\varepsilon}},$$

where $f\in W^{1,2+\varepsilon}(H^s,\mathbb{R})$.

The article is organized as follows. In Section 2 and Section 3, we prove that the logarithmic Sobolev inequality is preserved under the flows generated by certain ODEs in finite and infinite-dimensional spaces respectively. That is, the induced measures still satisfy logarithmic Sobolev inequalities both in finite and infinite dimensional cases. Then in Section 4A we prove the existence of the dynamics of BBM equation, and in Section 4B we prove Theorem 1.1.

2. The flow generated by vector fields in the finite-dimensional case

Let $\left(\mathbb{R}^d, d\mu(x)=\frac{1}{(2\pi)^{d/2}}e^{-x^2/2}\,dx\right)$ be the Gaussian space, then there holds the classical logarithmic Sobolev inequality

$$\int f^2\log\frac{f^2}{\int f^2\,d\mu}d\mu\le 2\int|\nabla f|^2\,d\mu. \tag{2-1}$$

Suppose $T : \mathbb{R}^d \to \mathbb{R}^d$ is an invertible diffeomorphism; then it induces a new probability measure $\nu = T_*\mu$ on $\mathbb{R}^d$. By denoting $\mathbb{E}_\mu[\,\cdot\,] = \int \cdot \, d\mu$, and applying (2-1) with $f \circ T$, we have

$$\begin{aligned}\mathbb{E}_\nu\left[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\right] &= \mathbb{E}_\mu\left[(f\circ T)^2 \log \frac{(f\circ T)^2}{\mathbb{E}_\mu[(f\circ T)^2]}\right] \\ &\le 2\mathbb{E}_\mu[|\nabla(f\circ T)|^2] \\ &\le 2\mathbb{E}_\mu[|\nabla f\circ T|^2 \cdot |\nabla T|^2] \\ &\le 2c\mathbb{E}_\nu[|\nabla f|^2],\end{aligned}$$

provided that, for some constant $c > 0$,

$$|\nabla T| \le c, \quad \mu - a.e. \tag{2-2}$$

We are now in a position to claim:

Proposition 2.1. *Suppose $T : \mathbb{R}^d \to \mathbb{R}^d$ is an invertible diffeomorphism and it satisfies the assumption (2-2). Then for the induced probability measure $\nu = T_*\mu$ there holds, for some other constant C,*

$$\mathbb{E}_\nu\left[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\right] \le C\mathbb{E}_\nu[|\nabla f|^2].$$

Next we suppose that the transformation $T_t : \mathbb{R}^d \ni x \mapsto U_t(x) \in \mathbb{R}^d$ is the flow map associated to the ODE

$$\begin{cases}\dfrac{d}{dt}U_t(x) = B(U_t(x)), \\ U_0(x) = x\end{cases} \tag{2-3}$$

under the condition that B is C^1 and also globally Lipschitzian. In the following, we are going to seek some condition on the vector field B such that $|\nabla T|$ is $(T_t)_*\mu$-almost surely bounded by some constant C. By differentiating the flow equation (2-3) in the space variable, we arrive at

$$\begin{cases}\dfrac{d}{dt}\nabla U_t(x) = \nabla B(U_t(x)) \cdot \nabla U_t(x), \\ \nabla U_0(x) = \mathrm{Id}.\end{cases} \tag{2-4}$$

By the assumption that B is globally Lipschitzian, the ODE system (2-4) is globally well-posed and its solution can be written as

$$\nabla U_t(x) = \mathrm{Id} + \int_0^t \nabla B(U_\tau(x)) \cdot \nabla U_\tau(x)\, d\tau. \tag{2-5}$$

Hence we have

$$\|\nabla U_t\| \le 1 + \left| \int_0^t \|\nabla B\| \times \|\nabla U_\tau(x)\| \, d\tau \right|.$$

With L being the Lipschitz constant of B, we have

$$\|\nabla U_t\| \le 1 + \left| \int_0^t L \|\nabla U_\tau(x)\| \, d\tau \right|.$$

By Gronwall's inequality, we get

$$\|\nabla U_t\| \le e^{L|t|}.$$

This last estimate verifies the assumption (2-2), with an upper bound depending on the time. Thus we can state:

Proposition 2.2. *Let $T_t : x \mapsto T_t(x) = U_t(x)$, where $U_t(x)$ is the flow map defined by (2-3). Assume that B is a C^1 Lipschitz vector field with Lipschitz constant L. Then for any $t \in (-\infty, +\infty)$, the induced measure $\nu_t = (T_t)_* \mu$ satisfies a logarithmic Sobolev inequality with some constant depending on c and t. In particular, for some given time T, and for any $t \in [-|T|, |T|]$, the measure ν_t satisfies a logarithmic Sobolev inequality with a uniform constant $C = C(T)$.*

3. The flow generated by vector fields in the infinite dimensional case

Suppose that $W = C_0([0, 1]; \mathbb{R})$ is the space of continuous functions vanishing at 0. We equip W with supremum norm $\|\cdot\|_{C_0}$, then $(W, \|\cdot\|_{C_0})$ is a Banach space. Consider its Cameron-Martin space H defined by

$$H = \left\{ u \in W : u' \text{ exists and } \int_0^1 |u'(\tau)|^2 \, d\tau < \infty \right\}.$$

We supply H an inner product $(u, v) = \int_0^1 u'(\tau) v'(\tau) \, d\tau$ for $u, v \in H$. We select an orthonormal basis $\{e_1, \dots, e_k, \dots\}$ in H, such that all these e_k's are from a subspace H_0 of H

$$H_0 = \{h \in H : h'' \text{ is a signed measure}\}.$$

For example we can take this orthonormal basis to be the Faber–Schauder system. Then the linear continuous functional $x \mapsto (x, e_k)$ for any k defined on H can be extended as a continuous functional defined on W. We also supply a Gaussian or Wiener measure μ on W via the formula

$$\int_W e^{i(x,h)} \, d\mu(x) = e^{-\frac{1}{2}|h|_H^2} \qquad \text{for all } h \in H.$$

For any $n \geq 1$, denote by V_n the linear envelope of $\{e_1, e_2, \ldots, e_n\}$. Suppose that we are given a function $f(x) = F(x_1, x_2, \ldots, x_k, \ldots)$,[1] where $x_i = (e_i, x)$ for all $i \geq 1$, we define its restriction to V_n by

$$f_n(x) = F(x_1, x_2, \ldots, x_n).$$

An equivalent way to define this restriction is via the Hermite polynomials. For any $k \geq 0$, the Hermite polynomial $H_k(y)$ on $\mathbb{R}$ is defined by

$$H_k(y) = \frac{(-1)^k}{\sqrt{k!}} e^{y^2/2} \frac{d^k}{dy^k} e^{-y^2/2}, \quad y \in \mathbb{R}.$$

We define the Hermite polynomials on W by

$$H_k(x) = \Pi_i H_{k_i}((e_i, x)),$$

where $k = (k_1, \ldots, k_n, \ldots)$, $k_i \geq 0$, $|k| = \sum k_i < \infty$. Then $\{H_k(x)\}_{k \in \mathbb{N}^{\mathbb{N}}}$ is an orthonormal basis of $L^2(W, \mathbb{R})$. We denote

$$C_n = \{k = (k_1, \ldots, k_n, \ldots) | k_q = 0, \text{ for all } q > n\}.$$

Then the restriction of f to V_n can be expressed as $f_n(x) = \sum_{k \in C_n} c_k H_k(x)$.

For any $\phi \in L^2(W, \mu)$, we define its H-derivative $\nabla_h \phi$ for any $h \in H$ provided that the following limit exists

$$\frac{d}{d\varepsilon}|_{\varepsilon=0}\, \phi(x + \varepsilon h) =: \nabla_h \phi(x) = \langle \nabla \phi, h \rangle.$$

We can see that $\nabla \phi(x)$ is actually in $H^* = H$, that is, it is an H-valued random variable. More generally, if X is another Banach space, then $\nabla \phi(x)$ is indeed an element in $\mathcal{L}(H, X)$, the space of linear operators from H to X. We next define the Sobolev spaces $W^{1,p}$ on W as the collection of these functions f on W such that $f \in L^p(W)$ and their derivatives satisfy $\|\nabla f\|_H \in L^p(W, d\mu)$.

With these notions and notations, we are ready to study the following infinite-dimensional ODE

$$\begin{cases} \dfrac{d}{dt} U_t(x) = B(U_t(x)), \\ U_0(x) = x, \end{cases} \tag{3-1}$$

where B is a vector field over W. We can write the equation in the integral form

$$U_t(x) = x + \int_0^t B(U_\tau(x))\, d\tau.$$

Hence, by the Cameron–Martin theorem, it is expected that the induced measure $\nu_t = (U_t)_* \mu$ is absolutely continuous with respect to μ if B is Cameron–Martin space-valued. Indeed this was studied by Cruzeiro [1983], and he obtained this

[1] This expression is unique due to the fact that $\{e_1, e_2, \ldots\}$ is a Schauder basis.

absolute continuity under some exponential integrability condition, which gives a sense of the Radon–Nikodym derivative. The successive question is whether or not there exist some other properties of the measure, which remain (almost) invariant under the flow associated to (3-1)? Notice that initially the measure is a Gaussian one, which satisfies a logarithmic Sobolev inequality. We are thus lead to consider the preservation of inequalities of this type.

To address this question, we begin with the statement of well-posedness of the flow associated to (3-1), whose proof is standard.

Proposition 3.1. *Let $B : x \in W \mapsto B(x) \in W$ be C^1, in the general sense rather than H-derivative, and globally Lipschitzian. Then (3-1) defines an invertible global flow on W. Moreover, for fixed t, the flow map is actually a C^1-diffeomorphism, and so is its inverse.*

Under the conditions of the above proposition, the induced measure $\nu_t = (U_t)_*\mu$ is also a probability measure on W, but it may not be absolutely continuous with respect to μ. For example, we take $B(x) = x$ on W, and the flow U_t is

$$U_t(x) = e^t x,$$

and hence the induced measure is just a scaling of μ. In the infinite-dimensional case, it is well-known that ν_t is singular to μ for any $t \neq 0$! But ν_t is still a Gaussian measure, and hence it still satisfies a log-Sobolev type inequality.

In the rest of this section, B is assumed to be Lipschitzian and H-valued. In this case, we can perform the trick used in the finite-dimensional ODE case. Substituting the composed function $f \circ T$ in the log-Sobolev inequality for μ leads us to

$$\begin{aligned}\mathbb{E}_\nu\left[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\right] &= \mathbb{E}_\mu\left[(f\circ T)^2 \log \frac{(f\circ T)^2}{\mathbb{E}_\mu[(f\circ T)^2]}\right] \\ &\le C\mathbb{E}_\mu[|\nabla f|_H^2 \cdot |\nabla T|_H^2],\end{aligned}$$

where $T := U_t$ for fixed t. Then we are going to estimate the upper bound of the operator norm of ∇T. For any $h \in H$, differentiating the ODE (3-1) in the direction h, we get

$$\begin{cases} \dfrac{d}{dt}\nabla_h U_t(x) = (\nabla B(U_t(x)), \nabla_h U_t(x)), \\ \nabla_h U_0(x) = h, \end{cases}$$

or, the equivalent form in the space $\mathcal{L}(H, H)$,

$$\begin{cases} \dfrac{d}{dt}\nabla_. U_t(x) = (\nabla B(U_t(x)), \nabla_. U_t(x)), \\ \nabla_. U_0(x) = \mathrm{Id}. \end{cases}$$

The succeeding estimates are the same as that in the finite-dimensional ODE case. Finally we get the desired result.

4. The flow generated by the BBM model equation

4A. ***Existence of the global transformation.*** Consider the generalized BBM model equation

$$\begin{cases} \partial_t u + \partial_t |\partial_x|^\gamma u + \partial_x(u+u^2) = 0, \\ u(0) = u_0, \end{cases} \tag{4-1}$$

where u is a real-valued function defined on $\mathbb{R}_t \times \mathbb{T}^1$. Observing that $\int_{\mathbb{T}} u\,dx$ is preserved, we thus assume $\int_{\mathbb{T}} u\,dx = 0$ and work on the space H^s of functions of Sobolev regularity s with mean zero. By denoting $\varphi(\partial_x) = \partial_x/(1+|\partial_x|^\gamma)$, we rewrite (4-1) as

$$\partial_t u = -\varphi(\partial_x)(u+u^2). \tag{4-2}$$

By integrating on the time interval $[0,t]$, we can also write the equation in the integral form

$$u(t) = u_0 - \int_0^t \varphi(\partial_x)(u(\tau)+u^2(\tau))d\tau. \tag{4-3}$$

We notice that, if we are using fixed point argument to solve (4-3), the main obstacle is the one brought by the nonlinear term u^2. To deal with this nonlinearity, we invoke the following lemma.

Lemma 4.1. *Fix $\gamma > \frac{3}{2}$. Let $0 \le \alpha, \beta \le s$ such that $2s - \alpha - \beta < \gamma - \frac{3}{2}$. Then for any $u \in H^\alpha$, $v \in H^\beta$, we have*

$$\|\varphi(\partial_x)(uv)\|_{H^s} \le C\|u\|_{H^\alpha}\,\|v\|_{H^\beta},$$

where C is a finite positive constant, depending on s, α and β.

Proof. We follow the ideas in [Bona and Tzvetkov 2009; Roumégoux 2010] to prove this lemma. Indeed, it suffices to prove, for any $w \in L^2$, there exists some universal constant C such that

$$\left|\sum_k \frac{\langle k\rangle^s k}{1+|k|^\gamma} \sum_l \hat{u}(k-l)\hat{v}(l)\overline{\hat{w}(k)}\right| \le C\|u\|_{H^\alpha}\,\|v\|_{H^\beta}\,\|w\|_{L^2}. \tag{4-4}$$

Denoting the left-hand side of (4-4) by I, we then do the following calculations:

$$\begin{aligned} I &= \left|\sum_l \langle l\rangle^\beta \hat{v}(l) \sum_k \frac{1}{\langle k\rangle^{-s+\alpha+\beta}} \frac{\langle k\rangle^s k}{(1+|k|^\gamma)} \frac{\langle k\rangle^{-s+\alpha+\beta}}{\langle k-l\rangle^\alpha \langle l\rangle^\beta}\big(\langle k-l\rangle^\alpha \hat{u}(k-l)\big)\overline{\hat{w}(k)}\right| \\ &\le \|v\|_{H^\beta}\left\|\sum_k \frac{1}{\langle k\rangle^{-s+\alpha+\beta}} \frac{\langle k\rangle^s k}{(1+|k|^\gamma)} \frac{\langle k\rangle^{-s+\alpha+\beta}}{\langle k-l\rangle^\alpha \langle l\rangle^\beta}\big(\langle k-l\rangle^\alpha \hat{u}(k-l)\big)\overline{\hat{w}(k)}\right\|_{\ell^2_l} \end{aligned}$$

$$\leq C\|u\|_{H^\alpha}\,\|v\|_{H^\beta}\left\|\frac{1}{\langle k\rangle^{\gamma-1-2s+\alpha+\beta}}\overline{\hat{w}(k)}\right\|_{\ell^1_k},$$

where in the first inequality we used Hölder's inequality, and in the second one we used Young's inequality and the fact that the quantity $(\langle k\rangle^{-s+\alpha+\beta})/(\langle k-l\rangle^\alpha\langle l\rangle^\beta)$ is bounded by some constant C on $(k,l)\in\mathbb{Z}^2$ provided $\alpha,\beta\in[0,s]$. To finish the proof, we use Hölder's inequality again

$$\left\|\frac{1}{\langle k\rangle^{\gamma-1-2s+\alpha+\beta}}\overline{\hat{w}(k)}\right\|_{\ell^1_k}\leq\|w\|_{L^2}\left\|\frac{1}{\langle k\rangle^{\gamma-1-2s+\alpha+\beta}}\right\|_{\ell^2_k}\leq\widetilde{C}\|w\|_{L^2},$$

where $\widetilde{C}=\|1/\langle k\rangle^{\gamma-1-2s+\alpha+\beta}\|_{\ell^2_k}$, which is finite provided that $\gamma-1-2s+\alpha+\beta>\frac{1}{2}$, i.e., $\gamma-\frac{3}{2}>2s-\alpha-\beta$. This last condition is just the assumption in the statement of the lemma. Thus this completes the proof of Lemma 4.1. □

Remark 4.2. In particular, if we take $\alpha=\beta=s$, the inequality in the lemma reads

$$\|\varphi(\partial_x)(uv)\|_{H^s}\leq C\|u\|_{H^s}\,\|v\|_{H^s}.$$

Here we need γ to be strictly bigger than $\frac{3}{2}$. Furthermore, this inequality cannot hold for $\gamma\leq\frac{3}{2}$ and $s=0$, as shown by Chenmin Sun [2015]. In a later paper, we will talk more about this inequality in more general cases.

With the help of these bilinear estimates, we prove the following local well-posedness result.

Theorem 4.3 (local well-posedness). *Fix $\gamma>\frac{3}{2}$ and $s\geq 0$. Then for any $u_0\in H^s$, (4-1) is locally well-posed in $X^s_T:=C(0,T;H^s)$ provided T is sufficiently small.*

Proof. Taking X^s_T-norm on both sides of (4-3), we get the following estimates by using Lemma 4.1 with $\alpha=\beta=s$:

$$\|u\|_{X^s_T}\leq\|u_0\|_{H^s}+T(\|u\|_{X^s_T}+\|u\|^2_{X^s_T}). \tag{4-5}$$

By taking $R=2\|u_0\|_{H^s}$, the map defined by the right-hand side of (4-3) is onto $B_R\subset X^s_T$ for T sufficiently small. Suppose v is another solution with the same initial data, we estimate

$$\|u-v\|_{X^s_T}\leq T(1+\|u+v\|_{X^s_T})\|u-v\|_{X^s_T}\leq T(1+2R)\|u-v\|_{X^s_T}. \tag{4-6}$$

Choosing $T=C(1+\|u_0\|_{H^s})^{-1}$ with C a small constant, we see that the solution map defined on the right-hand side of (4-3) is a contraction map. Therefore, by the contraction mapping principle, there is a solution u to (4-3) in X^s_T for T sufficiently small. □

Remark 4.4. We see that the length of the time interval is just of size $(1+\|u_0\|_{H^s})^{-1}$. This contradicts the general expectation that we can get the solution on any long time interval if we let the size of the initial data be sufficiently small. To remedy this expectation, we need to rewrite (4-2) in the Duhamel form as

$$u(t) = e^{-t\varphi(\partial_x)}u_0 - \int_0^t e^{-(t-\tau)\varphi(\partial_x)}\varphi(\partial_x)(u^2(\tau))\,d\tau.$$

Doing the same estimates as above, we have the following estimates analogous to estimates (4-5) and (4-6):

$$\|u\|_{X_T^s} \le \|u_0\|_{H^s} + T\|u\|_{X_T^s}^2,$$
$$\|u-v\|_{X_T^s} \le T\big(\|u\|_{X_T^s} + \|v\|_{X_T^s}\big)\|u-v\|_{X_T^s}.$$

Taking $R = 2\|u_0\|_{H^s}$ and then playing the fixed-point argument, we get the desired size of the existence time interval. In particular, for any $T > 0$, there exists $\delta > 0$, such that for any data $u_0 \in H^s$, as long as $\|u_0\|_{H^s} \le \delta$, there exists solution $u(t)$ to (4-1) up to time T.

Next we are going to study the large time existence. For any $u_0 \in H^s$ and any $T > 0$, we take N a sufficiently large positive integer such that

$$\sum_{|k|\ge N} \langle k\rangle^{2s}|\widehat{u_0}(k)|^2 \le T^{-2}.$$

Denote $v_0 = \sum_{|k|\ge N} \widehat{u_0}(k)e^{ikx}$. Then by Remark 4.4, there exists a unique solution v in X_T^s to (4-1), issued from v_0. Furthermore, the solution v is of size $\frac{1}{T}$ in X_T^s. Decompose $u_0 = v_0 + w_0$. Then if we want to solve (4-1), we only need to solve

$$\begin{cases} \partial_t w = -\varphi(\partial_x)(w + 2vw + w^2), \\ w(0) = w_0. \end{cases} \tag{4-7}$$

Suppose w solves (4-7) up to time T. Then $u := v + w$ solves (4-1) on the time interval $[0, T]$. As w_0 consists of only the first N-frequencies, it belongs to H^r for any $r > 0$. But we only treat this in $H^{\frac{\gamma}{2}}$, which is just what we need. In this case (4-7) is locally well-posed. Indeed, by writing (4-7) in its Duhamel form

$$w(t) = e^{-t\varphi(\partial_x)}w_0 - \int_0^t e^{-(t-\tau)\varphi(\partial_x)}\phi(\partial_x)(2v(\tau)w(\tau) + w^2(\tau))\,d\tau =: L(w), \tag{4-8}$$

we play the fixed point argument as follows. Before doing this, we need the following lemma, which allows us to deal with the nonlinear term $\varphi(\partial_x)(vw)$.

Lemma 4.5. *Fix $\alpha > \frac{1}{2}$. Let $s \in [0, \alpha]$. Then for any $u \in H^\alpha$ and $v \in H^s$, we have for some positive constant $C(\alpha, s)$*

$$\|\varphi(\partial_x)(uv)\|_{H^{s+\gamma-1}} \le C(\alpha, s)\|u\|_{H^\alpha}\,\|v\|_{H^s}.$$

Proof. We prove this lemma essentially along the same lines as in [Bona and Tzvetkov 2009]. By the smoothing effect of $\varphi(\partial_x)$ in the L^2-based Sobolev spaces, we only need to show

$$\|uv\|_{H^s} \le C(\alpha, s)\|u\|_{H^\alpha}\|v\|_{H^s},$$

for $u \in H^\alpha$ and $v \in H^s$. This last inequality follows exactly from the fact that elements of H^α are multipliers in H^s for $\alpha > \frac{1}{2}$ and $0 \le s \le \alpha$. □

Suppose that S is a positive time to be selected. We estimate for $t \in [0, S]$

$$\begin{aligned}\|L(w)\|_{X_S^{\gamma/2}} &\le \|w_0\|_{H^{\gamma/2}} + \int_0^t \|\varphi(\partial_x)(2vw + w^2)\|_{H^{\gamma/2}}\, d\tau \\ &\le \|w_0\|_{H^{\gamma/2}} + \int_0^t (\|\varphi(\partial_x)(wv)\|_{H^{\gamma/2}} + \|w\|^2_{X_S^{\gamma/2}})\, d\tau \\ &\le \|w_0\|_{H^{\gamma/2}} + S(\|w\|^2_{X_S^{\gamma/2}} + \|w\|_{X_S^{\gamma/2}}\|v\|_{X_S^{1-\gamma/2}}),\end{aligned}$$

where in the last inequality, we have used Lemma 4.5. Under the assumption that $s \ge 1 - \gamma/2$, we have

$$\|L(w)\|_{X_S^{\gamma/2}} \le \|w_0\|_{H^{\gamma/2}} + S(\|w\|^2_{X_S^{\gamma/2}} + \|w\|_{X_S^{\gamma/2}}\|v\|_{X_S^s}). \tag{4-9}$$

A similar argument gives us the estimate

$$\|L(w_1) - L(w_2)\|_{X_S^{\gamma/2}} \le CS(\|v\|_{X_S^s} + \|w_1\|_{X_S^{\gamma/2}} + \|w_2\|_{X_S^{\gamma/2}})\|w_1 - w_2\|_{X_S^{\gamma/2}}. \tag{4-10}$$

Thus by selecting

$$R = 2\|w_0\|_{H^{\gamma/2}} \sim N^{\gamma/2-s}\|u_0\|_{H^s},$$

and choosing $S \sim c/(\|w_0\|_{H^{\gamma/2}} + 1/T)$ with c sufficiently small, the map L is a contraction map onto $B_R \subset X_S^{\gamma/2}$. Hence it has a solution w to (4-7) up to time S.

In order to establish the large time existence of the solution of (4-7), we need to establish the following *a priori* estimate, which can be viewed as an almost conservation law.

Multiplying the first equation in (4-7) by $(1 + |\partial_x|^\gamma)w$ and integrating on the circle, we get

$$\frac{d}{dt}\frac{1}{2}\|w\|^2_{H^{\gamma/2}} = -\int_{\mathbb{T}} \varphi(\partial_x)(w + 2wv + w^2)(1 + |\partial_x|^\gamma)w\, dx.$$

On one hand, by the definition of $\varphi(\partial_x)$ and integration by parts, we have that

$$\int_{\mathbb{T}} \varphi(\partial_x)(w^j)(1 + |\partial_x|^\gamma)w\, dx = 0 \quad \text{for } j = 1, 2.$$

On the other hand, by the self-adjointness of $(1+|\partial_x|^\gamma)^{1/2}$ and the Cauchy–Schwartz inequality, we have

$$\left|\int_{\mathbb{T}} \varphi(\partial_x)(wv)(1+|\partial_x|^\gamma)w\,dx\right| \le \|\varphi(\partial_x)(vw)\|_{H^{\gamma/2}}\,\|w\|_{H^{\gamma/2}}.$$

By using Lemma 4.5 with $s = 1-\frac{\gamma}{2}$ and $\alpha = \frac{\gamma}{2}$, we have

$$\left|\int_{\mathbb{T}} \varphi(\partial_x)(wv)(1+|\partial_x|^\gamma)w\,dx\right| \le \|v\|_{H^{1-\gamma/2}}(\|w\|_{H^{\gamma/2}})^2.$$

Thus, by combining these two points, and using the assumption that $s \ge 1-\frac{\gamma}{2}$, we get

$$\frac{d}{dt}\frac{1}{2}\|w\|^2_{H^{\gamma/2}} \le \|w\|_{H^{\gamma/2}} + \|w\|^2_{H^{\gamma/2}}\,\|v\|_{H^s}.$$

A usage of Gronwall's inequality gives us, for any $t \in [0, T]$,

$$\|w\|_{H^{\gamma/2}} \le \|w_0\|_{H^{\gamma/2}}\, e^{\int_0^T (1+\|v(\tau)\|_{H^s})\,d\tau} \le e^{1+T}\,\|w_0\|_{H^{\gamma/2}}. \tag{4-11}$$

Therefore, by the *a priori* estimate (4-11), we can solve (4-7) on the interval $[0, S]$, with S of size

$$\frac{c}{\|w_0\|_{H^{\gamma/2}}e^{1+T} + 1/T}.$$

Thanks to this *a priori* bound, we can solve (4-7) on the succeeding interval $[S, 2S]$ with initial data $w(S)$ obtained in the previous step. Because S does not depend at which step we solve the equation, we can repeat the above procedure until we arrive at some interval $[kS, (k+1)S]$ such that $(k+1)S \ge T$. That is to say, we can solve (4-1) up to time T and also validate the estimate by using the assumption $s \le \frac{\gamma}{2}$

$$\|u\|_{X^s_T} \le \frac{1}{T} + N^{\gamma/2-s}e^{1+T} \tag{4-12}$$

up to some constants. Therefore, we are in a position to state:

Theorem 4.6. *Fix* $\gamma \in \left(\frac{3}{2}, 2\right]$. *Let* $1-\frac{\gamma}{2} \le s \le \frac{\gamma}{2}$. *Then for any* $u_0 \in H^s$ *and any* $T > 0$, *there exists a unique solution* u *to* (4-1) *in* $C([0, T]; H^s)$. *Furthermore, there exists some* $N_0 \in \mathbb{N}$ *such that for all* $t \in [0, T]$, *we have*

$$\|u(t)\|_{H^s} \le \frac{1}{T} + e^{1+T}N_0^{\gamma/2-s}.$$

Remark 4.7. To get long time existence, we used the local well-posedness and Lemma 4.5. To establish the local well-posed result we need to use the assumption $\gamma > \frac{3}{2}$, while we do not need the assumption in Lemma 4.5. This assumption just arises when we use Lemma 4.1 to deal with the nonlinearity. So this motivates us to seek some other conditions that are sufficient to achieve local well-posedness.

There are two such conditions:

(1) $\gamma \geq 1$ and $s > \frac{1}{2}$, which works due to the fact that H^s is an algebra when $s > \frac{1}{2}$ and $\varphi(\partial_x)$ is a smoothing operator;

(2) $\gamma > \frac{5}{4}$ and $s > \frac{1}{4}$, which works thanks to the fact that, in this case, $u \in H^s$ implies $u^2 \in H^{s-1/4}$ and hence $\varphi(\partial_x)(u^2) \in H^s$.

But both of the above conditions cannot guarantee the large time well-posedness for the regularity $s < (\gamma - 1)/2$, so we omit the detailed discussion here.

4B. ***Some kind of LSI.*** In the following, we consider the special random initial data

$$u_0 = \phi_s(\omega, x) = \sum_{n \neq 0} \frac{g_n(\omega)}{|n|^{s+\gamma/2}} e^{inx},$$

where $g_n = \overline{g_{-n}}$ and $(g_n)_{n>0}$ is a sequence of independent standard complex Gauss random variables on a given probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then one can show that the map $\omega \mapsto \phi_s(\omega, x)$ induces a Gaussian measure μ_s on H^s, and the space $H^{s+\gamma/2}$ with $\gamma > 1$ is the Cameron–Martin space of the Gaussian probability space (H^s, μ_s). In this case, the triple $(H^{s+\gamma/2}, H^s, \mu_s)$ is called an *abstract Wiener space*. Furthermore, one can also define the H-derivative as in the case of classical Wiener space. Under these notions, we actually have an infinite-dimensional log-Sobolev type inequality with some constant c

$$\mathbb{E}_{\mu_s}\left[f^2 \log \frac{f^2}{\mathbb{E}_{\mu_s}[f^2]} \right] \leq c \mathbb{E}_{\mu_s}[|\nabla f|^2_{H^{s+\gamma/2}}], \tag{4-13}$$

for any $f \in W^{1,2}(H^s, \mathbb{R})$, the space of real functions on H^s, which is $L^2(\mu_s)$-integrable together with its $H^{s+\gamma/2}$-derivatives.

As $u_0 = \phi_s(\omega, x)$ lies in H^s almost surely, the flow $\Phi(t)$ is defined almost surely everywhere. In order to study the preservation of log-Sobolev type inequalities, we are going to study the linearization, in the direction v_0, of the solution $\Phi(t)(u_0)$ as follows.[2]

For brevity, we denote the solution $\Phi(t)(u_0) := S(t)(u_0) + K(t)_{u_0}$. Then

$$(D\Phi(t))_{u_0}(v_0) = S(t)(v_0) + (DK(t))_{u_0}(v_0),$$

where

$$(DK(t))_{u_0}(v_0) = -2 \int_0^t S(t-\tau)\big((1 + |\partial_x|^\gamma)^{-1} \partial_x(\Phi(\tau)(u_0) v(\tau))\big)\, d\tau,$$

where $v(t)$ solves the linearized equation

$$\begin{cases} \partial_t v + \partial_t |\partial_x|^\gamma v + \partial_x v + 2\partial_x(\Phi(t)(u_0) v) = 0, \\ v|_{t=0} = v_0. \end{cases} \tag{4-14}$$

[2]Here we follow [Tzvetkov 2015].

Proposition 4.8. *Let γ, s and u_0 be as in Theorem 4.6, then for any $v_0 \in H^{s+\gamma/2}$, we have that $v(t)$ is also in $H^{s+\gamma/2}$ for any t in the life span of $u(t)$. Furthermore, we have the bound*

$$\|v(t)\|_{H^{s+\gamma/2}} \leq C(t)\|v_0\|_{H^{s+\gamma/2}}.$$

Next we are going to show that, as an operator parametrized by u_0, $(DK(t))$ is bounded from $H^{s+\gamma/2}$ into itself.

$$\begin{aligned}\|(DK(t))_{u_0}v_0\|_{H^{s+\gamma/2}} &\leq \Big\|\int_0^t S(t-\tau)\varphi(\partial_x)\big(\Phi(\tau)(u_0)v(\tau)\big)\,d\tau\Big\|_{H^{s+\gamma/2}} \\ &\leq \int_0^t \|\varphi(\partial_x)(\Phi(\tau)(u_0)v(\tau))\|_{H^{s+1-\gamma/2+\gamma-1}}\,d\tau \\ &\leq \int_0^t \|\Phi(\tau)(u_0)\|_{H^s}\,\|v(\tau)\|_{H^{s+1-\gamma/2}}\,d\tau.\end{aligned}$$

Under the condition that $\gamma > \frac{3}{2}$ and hence $s - \frac{\gamma}{2} + 1 \leq \frac{\gamma}{2} + s$, we have

$$\|DK(t)_{u_0}\|_{H^{s+\gamma/2}\to H^{s+\gamma/2}} \leq C\|u_0\|_{H^s}. \tag{4-15}$$

By denoting $T := \Phi(t) : u \to \Phi(t)(u)$ for fixed t, we do the following calculations

$$\begin{aligned}\mathbb{E}_\nu\Big[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\Big] &= \mathbb{E}_\mu\Big[(f\circ T)^2 \log \frac{(f\circ T)^2}{\mathbb{E}_\mu[(f\circ T)^2]}\Big] \\ &\leq c\mathbb{E}_\mu[|\nabla(f\circ T)|^2_{H^{s+\gamma/2}}] \leq c\mathbb{E}_\mu\big[|\nabla f\circ T|^2_{H^{s+\gamma/2}}(1+\|\cdot\|^2_{H^s})\big].\end{aligned}$$

Then thanks to Fernique's theorem, we arrive at the following log-Sobolev type inequality, with a loss of integrability:

Proposition 4.9. *Let $\gamma \in \big(\frac{3}{2}, 2\big]$ and $s \geq 1 - \frac{\gamma}{2}$. Then for any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that the induced measure $\nu = (\Phi(t))_*\mu_s$ satisfies a log-Sobolev type inequality*

$$\mathbb{E}_\nu\Big[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\Big] \leq C_\varepsilon(\mathbb{E}_\nu[|\nabla f|^{2+\varepsilon}_{H^{s+\gamma/2}}])^{\frac{2}{2+\varepsilon}}. \tag{4-16}$$

Proof. Given $\varepsilon > 0$, we first estimate

$$\begin{aligned}\mathbb{E}_\nu\Big[f^2 \log \frac{f^2}{\mathbb{E}_\nu[f^2]}\Big] &\leq c\mathbb{E}_\mu\big[|\nabla f\circ T|^2_{H^{s+\gamma/2}}(1+\|\cdot\|^2_{H^s})\big] \\ &\leq c\big(\mathbb{E}_\mu[|\nabla f\circ T|^{2+\varepsilon}_{H^{s+\gamma/2}}]\big)^{\frac{2}{2+\varepsilon}}\big(\mathbb{E}_\nu[(1+\|\cdot\|^2_{H^s})^{\frac{2+\varepsilon}{\varepsilon}}]\big)^{\frac{\varepsilon}{2+\varepsilon}},\end{aligned}$$

where in the last inequality, we have used Hölder's inequality. Noticing that

$$\big(\mathbb{E}_\mu[|\nabla f\circ T|^{2+\varepsilon}_{H^{s+\gamma/2}}]\big)^{\frac{2}{2+\varepsilon}} = \big(\mathbb{E}_\nu[|\nabla f|^{2+\varepsilon}_{H^{s+\gamma/2}}]\big)^{\frac{2}{2+\varepsilon}},$$

we only need to control

$$\big(\mathbb{E}_\nu[(1+\|\cdot\|^2_{H^s})^{\frac{2+\varepsilon}{\varepsilon}}]\big)^{\frac{\varepsilon}{2+\varepsilon}}.$$

Recall that Fernique's theorem [Üstünel 2010, Chapter 9] states that for some small positive $c > 0$,

$$\mathbb{E}_\mu[e^{c\|\cdot\|^2_{H^s}}] < +\infty.$$

Consequently, for any positive integer k, we have

$$\mathbb{E}_\mu[\|\cdot\|^k_{H^s}] < +\infty. \tag{4-17}$$

Letting k be the smallest integer that is bigger than $(2+\varepsilon)/\varepsilon$, we have

$$\mathbb{E}_\mu\big[(1+\|\cdot\|^2_{H^s})^{\frac{2+\varepsilon}{\varepsilon}}\big] \le \mathbb{E}_\mu\big[(1+\|\cdot\|^2_{H^s})^k\big] = \sum_{j=0}^{k}\binom{2k}{2j}\mathbb{E}_\mu[\|\cdot\|^{2j}_{H^s}],$$

which is finite thanks to (4-17). This completes the proof of Proposition 4.9. □

References

[Bona and Tzvetkov 2009] J. L. Bona and N. Tzvetkov, "Sharp well-posedness results for the BBM equation", *Discrete Contin. Dyn. Syst.* **23**:4 (2009), 1241–1252. MR Zbl

[Cruzeiro 1983] A. B. Cruzeiro, "Équations différentielles sur l'espace de Wiener et formules de Cameron–Martin non-linéaires", *J. Funct. Anal.* **54**:2 (1983), 206–227. MR Zbl

[Da Prato 2006] G. Da Prato, *An introduction to infinite-dimensional analysis*, Revised ed., Springer, 2006. MR Zbl

[Gross 1975] L. Gross, "Logarithmic Sobolev inequalities", *Amer. J. Math.* **97**:4 (1975), 1061–1083. MR Zbl

[Gross 1991] L. Gross, "Logarithmic Sobolev inequalities on loop groups", *J. Funct. Anal.* **102**:2 (1991), 268–313. MR Zbl

[Hsu 1997] E. P. Hsu, "Logarithmic Sobolev inequalities on path spaces over Riemannian manifolds", *Comm. Math. Phys.* **189**:1 (1997), 9–16. MR Zbl

[Ledoux 2001] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89**, American Mathematical Society, Providence, RI, 2001. MR Zbl

[Roumégoux 2010] D. Roumégoux, "A symplectic non-squeezing theorem for BBM equation", *Dyn. Partial Differ. Equ.* **7**:4 (2010), 289–305. MR Zbl

[Sun 2015] C. Sun, personal communication, 2015.

[Tzvetkov 2015] N. Tzvetkov, "Quasiinvariant Gaussian measures for one-dimensional Hamiltonian partial differential equations", *Forum Math. Sigma* **3** (2015), art. id. e28, 35 pp. MR Zbl

[Üstünel 2010] A. S. Üstünel, "Analysis on Wiener space and applications", preprint, 2010. arXiv

Received August 20, 2018. Revised March 8, 2019.

BO XIA
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI
CHINA
xaboustc@hotmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 305, No. 1, 2020
dx.doi.org/10.2140/pjm.2020.305.353

GROUND STATE SOLUTIONS OF POLYHARMONIC EQUATIONS WITH POTENTIALS OF POSITIVE LOW BOUND

CAIFENG ZHANG, JUNGANG LI AND LU CHEN

The purpose of this paper is threefold. First, we establish the critical Adams inequality on the whole space with restrictions on the norm

$$(\|\nabla^m u\|_{\frac{n}{m}}^{\frac{n}{m}} + \tau \|u\|_{\frac{n}{m}}^{\frac{n}{m}})^{\frac{m}{n}}$$

for any $\tau > 0$. Second, we prove a sharp concentration-compactness principle for singular Adams inequalities and a new Sobolev compact embedding in $W^{m,2}(\mathbb{R}^{2m})$. Third, based on the above results, we give sufficient conditions for the existence of ground state solutions to the following polyharmonic equation with singular exponential nonlinearity

$$(-\Delta)^m u + V(x)u = \frac{f(x,u)}{|x|^\beta} \quad \text{in } \mathbb{R}^{2m}, \tag{0-1}$$

where $0 < \beta < 2m$, $V(x)$ has a positive lower bound and $f(x,t)$ behaves like $\exp(\alpha|t|^2)$ as $t \to +\infty$. Furthermore, when $\beta = 0$, in light of the principle of the symmetric criticality and the radial lemma, we also derive the existence of nontrivial weak solutions by assuming $f(x,t)$ and $V(x)$ are radially symmetric with respect to x and $f(x,t) = o(t)$ at origin. Thus our main theorems extend the recent results on bi-Laplacian in $\mathbb{R}^4$ by Chen, Li, Lu and Zhang (2018) to $(-\Delta)^m$ in $\mathbb{R}^m$.

1. Introduction and main results

The standard Sobolev space $W_0^{k,p}(\Omega)$ is defined by the completion of $C_c^\infty(\Omega)$ equipped with the norm

$$\|u\|_{W^{k,p}} = \left(\|u\|_p^p + \sum_{j=1}^{k} \|\nabla^j u\|_p^p\right)^{\frac{1}{p}},$$

Lu Chen is the corresponding author.
MSC2010: 26D10, 35A23, 46E35.
Keywords: ground state solutions, Adams inequality, concentration-compactness principle, exponential growth.

where Ω denotes a smooth bounded domain in $\mathbb{R}^n$. Basically, the Sobolev continuous embeddings state that

$$W_0^{k,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{for } 1 \le q \le \frac{np}{n-kp},\ kp < n.$$

However, in the limiting case $kp = n$, many examples show that $W_0^{k,\frac{n}{k}}(\Omega) \not\subseteq L^\infty(\Omega)$. In this case, the Trudinger–Moser inequality and the Adams inequality serve as appropriate replacements. Research concerning the sharp constant for the Trudinger–Moser inequality could be traced back to the 1960s and 1970s. Trudinger [1967] proved there exists a constant $\alpha > 0$ such that the following inequality holds (also see [Pohozaev 1965; Yudovich 1961]):

$$\text{(1-1)} \qquad \sup_{u \in W_0^{1,n}(\Omega),\, \|\nabla u\|_n \le 1} \frac{1}{|\Omega|} \int_\Omega e^{\alpha |u|^{\frac{n}{n-1}}}\, dx \le C_0.$$

Nevertheless, the best constant for (1-1) is unknown. Later, Moser [1971] established the sharp version of inequality (1-1) which can be stated as follows:

$$\text{(1-2)} \qquad \sup_{u \in W_0^{1,n}(\Omega),\, \|\nabla u\|_n \le 1} \frac{1}{|\Omega|} \int_\Omega e^{\alpha_n |u|^{\frac{n}{n-1}}}\, dx \le C_0,$$

where $\alpha_n = n\omega_{n-1}^{1/(n-1)}$ is the sharp constant in the sense that if α_n is replaced by any larger number, the supremum would become infinity. ω_{n-1} denotes the area of the surface of the unit ball in $\mathbb{R}^n$. Estimate (1-2) is now referred as the Trudinger–Moser inequality and plays an important role in geometric analysis and partial differential equations (e.g., see [Moser 1973]). For more results of Trudinger–Moser inequalities on compact Riemannian manifolds, one can refer to [Li 2001; 2006; Li and Ndiaye 2007]. If we replace Ω with $\mathbb{R}^n$, the Trudinger–Moser inequality (1-2) makes no sense. Instead, a subcritical Trudinger–Moser type inequality was proved by Adachi and Tanaka [2000]. By replacing the Dirichlet norm with the standard Sobolev norm in $W^{1,n}(\mathbb{R}^n)$, Cao [1992] (for $n = 2$), Panda [1996] and J. M. do Ó [2014] (for general n) constructed the Trudinger–Moser inequality in the whole space which states that for any $\alpha < \alpha_n$,

$$\text{(1-3)} \qquad \sup_{u \in W^{1,n}(\mathbb{R}^n),\, ||u||_{W^{1,n}(\mathbb{R}^n)} \le 1} \int_{\mathbb{R}^n} \Phi_{n,1}(\alpha |u(x)|^{\frac{n}{n-1}})\, dx \le C_n,$$

$$\text{where } \Phi_{n,1}(t) := e^t - \sum_{j=0}^{n-2} \frac{t^j}{j!}.$$

However, they did not prove the criticality of this inequality. Later, Ruf [2005] (for the case $n = 2$), Li and Ruf [2008] (for the general case $n \ge 3$) proved that Trudinger–Moser inequality (1-3) still holds in the critical case $\alpha = \alpha_n$ by using the symmetrization argument and the blow-up procedure. Both the critical and subcritical Trudinger–Moser inequalities on $\mathbb{R}^n$ given in the aforementioned works are based on the Pólya–Szegő inequality and symmetrization argument which is not available in other non-Euclidean settings. Lam and Lu [2012c] developed a symmetrization-free argument on the Heisenberg group and established the critical Trudinger–Moser inequality (see also Lam, Lu and Tang [Lam et al. 2014] for the subcritical Trudinger–Moser inequality without using symmetrization argument). In fact, the critical and subcritical Trudinger–Moser inequalities are proved equivalent by Lam, Lu and Zhang [Lam et al. 2017b], where they also establish relationships between supremums of the critical and subcritical Trudinger–Moser inequalities. Such a relationship has been used to establish the existence of extremal functions for subcritical Trudinger–Moser inequalities on the entire space $\mathbb{R}^n$ [Lam et al. 2019].

The above Trudinger–Moser inequalities and its generalizations are often applied to derive the existence of weak solutions for the following n-Laplacian equations:

$$-\operatorname{div}(|\nabla u|^{n-2}\nabla u)+V(x)|u|^{n-2}u=\frac{f(x,u)}{|x|^{\beta}}+\varepsilon h(x), \tag{1-4}$$

where $f:\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ is continuous and behaves like $\exp(\alpha|t|^{\frac{n}{n-1}})$ as $t\to+\infty$, $h(x)$ belongs to the dual space of $W^{1,n}(\mathbb{R}^n)$. Adding some appropriate assumptions on $V(x)$, one can see that the compact embedding

$$E=\left\{u:\int_{\mathbb{R}^n}|\nabla u|^n+V(x)|u|^n dx<+\infty\right\}\hookrightarrow L^p(\mathbb{R}^n)\quad\text{for } p\ge n$$

becomes admissible. The authors of [Adimurthi and Yang 2010; Alves and Figueiredo 2009; do Ó et al. 2014; Lam and Lu 2013a; Yang 2012b] carried out the standard mountain-pass procedure to obtain nontrivial weak solutions of (1-4). When $V(x)$ is constant, there is a long way to go yet. In order to overcome the possible failure of the Palais–Smale compactness condition which is caused by the absence of a compact embedding $W^{1,n}(\mathbb{R}^n)\hookrightarrow L^n(\mathbb{R}^n)$, Masmoudi and Sani [2015] applied a method involved with a constrained minimization argument and the sharp Trudinger–Moser inequality with the exact growth condition to investigate the existence of ground state solutions for (1-4) in the case of $V(x)=1$, $f(x,u)=f(u)$, $\beta=\varepsilon=0$. By assuming $f(x,t)=o(t)$, J. M. do Ó et al. [2014] employed a modified form of the Trudinger–Moser inequality and rearrangement inequalities to give sufficient conditions for the existence of ground state solutions. We also note that Lam and Lu [2014; 2013a] investigated the n-Laplacian equation and polyharmonic operators without the Ambrosetti–Rabinowitz condition. For more results about the Trudinger–Moser inequality and its application, we refer the reader to [Adimurthi and Sandeep 2007; Adimurthi and Yang 2010; Atkinson and

Peletier 1986; Carleson and Chang 1986; de Figueiredo et al. 2002; do Ó 1996; Lam and Lu 2012a; Panda 1996; Silva and Soares 1999].

D. Adams [1988] established the sharp Trudinger–Moser inequality with higher order derivatives. More precisely, he proved that

$$\sup_{u\in W_0^{m,\frac{n}{m}}(\mathbb{R}^n),\,\|\nabla^m u\|_{\frac{n}{m}}\le 1}\int_\Omega e^{\beta_{n,m}|u|^{\frac{n}{n-m}}}\,dx<\infty,$$

where

$$\beta_{n,m}=\begin{cases}\dfrac{n}{\omega_{n-1}}\left[\dfrac{\pi^{\frac{n}{2}}2^m\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{n-m+1}{2}\right)}\right]^{\frac{n}{n-m}}, & \text{if } m \text{ is odd,}\\[2ex] \dfrac{n}{\omega_{n-1}}\left[\dfrac{\pi^{\frac{n}{2}}2^m\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{n-m}{2}\right)}\right]^{\frac{n}{n-m}}, & \text{if } m \text{ is even.}\end{cases}$$

and

$$\nabla^m u=\begin{cases}\Delta^{\frac{m}{2}}, & \text{if } m \text{ is even,}\\ \nabla\Delta^{\frac{m-1}{2}}, & \text{if m is odd.}\end{cases}$$

The above inequality was extended by Tarsi [2012] to a larger space $W_N^{m,n/m}(\Omega)$ containing the Sobolev space $W_0^{n,n/m}(\Omega)$ as a closed subspace, where $W_N^{m,n/m}(\Omega)$ is given by

$$W_N^{m,\frac{n}{m}}(\Omega):=\left\{u\in W^{n,\frac{n}{m}}(\Omega)\mid \Delta^j u=0 \text{ on } \partial\Omega \text{ for } 0\le j\le\left[\tfrac{m-1}{2}\right]\right\}.$$

Sharp singular Adams inequalities on $W_N^{m,n/m}(\Omega)$ were also established by Lam and Lu [2012d]. We also mention that existence results concerning extremals of the Adams inequality in the case $n=2m=4$ were established by Lu and Yang [2009]. Li, Lu and Q. Yang [Li et al. 2018a; Lu and Yang 2017] proved the Hardy–Adams inequalities on hyperbolic spaces as a borderline case of the higher order Hardy–Sobolev–Mazya inequalities established by Lu and Q. Yang [2019] on upper half spaces.

After Adams, establishing Adams type inequalities in higher order Sobolev space $W^{m,n/m}(\mathbb{R}^n)$ has attracted much attention. Ogawa and Ozawa [1991] (for $n=2m$) and Ozawa [1995] (for general n, m) proved that there exist positive constants α and C_α such that

$$\int_{\mathbb{R}^n}\Phi_{n,m}(\alpha|u|^{\frac{n}{n-m}})\,dx\le C_\alpha,\quad \text{for all } u\in W^{m,\frac{n}{m}}(\mathbb{R}^n),\ |u|_{m,n}\le 1,$$

where

$$\Phi_{n,m}(t)=e^t-\sum_{j=0}^{j_{\frac{n}{m}}-2}\frac{t^j}{j!},\quad j_{\frac{n}{m}}=\min\left\{j\in\mathbb{N}: j\ge\tfrac{n}{m}\right\}$$

and $|u|_{m,n}$ is given by $|u|_{m,n}=\|(I-\Delta)^{m/2}u\|_{n/m}$. Kozono et al. [2006] studied the sharp constant problem by applying O'Neil's results on the rearrangement of convolution functions. In fact, they proved that there exists a constant $\beta^*_{n.m} \le \beta(n,m)$, particularly, $\beta^*_{2m.m} = \beta(2m,m)$ such that if $\beta < \beta^*_{n.m}$, then

$$\sup_{u\in W_0^{m,\frac{n}{m}}(\mathbb{R}^n),\,|u|_{m,n}\le 1} \int_{\mathbb{R}^n} \Phi_{n,m}(\beta|u|^{\frac{n}{n-m}})\,dx < \infty.$$

Ruf and Sani [2013] established the sharp Adams type inequality for the critical case $\beta = \beta_{n,m}$ when m is an even integer, where the Talenti's comparison principle plays an important role in their proof. Lam and Lu [2012b; 2012d] proved the above inequality for all integers m (including fractional order γ). More precisely, they showed

$$\sup_{\|(I-\Delta)^{\frac{m-1}{2}}u\|^{\frac{n}{m}}_{\frac{n}{m}}+\|\nabla(I-\Delta)^{\frac{m-1}{2}}u\|^{\frac{n}{m}}_{\frac{n}{m}}\le 1} \int_{\mathbb{R}^n} \Phi_{n,m}(\beta_{n,m}|u|^{\frac{n}{n-m}})\,dx \le C(m,n),$$

for any odd integer m. Lam and Lu [2013b] further developed a rearrangement-free approach. This method can help us to get rid of the symmetrization or the comparison principle argument. Using this method, Lam and Lu established the sharp Adams inequality which can be stated as follows:

$$\sup_{u\in W^{\gamma,p}(\mathbb{R}^n),\,\|(\tau I-\Delta)^{\frac{\gamma}{2}}u\|_p\le 1} \int_{\mathbb{R}^n} \exp(\beta_0(n,\gamma)|u|^{p'})\,dx \le C(n,\gamma), \tag{1-5}$$

$$\text{where}\quad 0<\gamma<n,\quad p=\frac{n}{\gamma}\quad\text{and}\quad \beta_0(n,\gamma)=\frac{n}{\omega_{n-1}}\left(\frac{\pi^{\frac{n}{2}}2^{\gamma}\Gamma\left(\frac{\gamma}{2}\right)}{\Gamma\left(\frac{n-\gamma}{2}\right)}\right)^{p'}.$$

Recently, Lam and Lu [2013b] obtained the Adams inequality involved with the norm $(\|\Delta u\|^{n/2}_{n/2} + \|u\|^{n/2}_{n/2})^{2/n}$. Later, Fontana and Morpurgo [2015] extended Lam and Lu's results to higher order derivatives. They proved that there exists some constant $C_{m,n}$ such that

$$\sup_{\|\nabla^m u\|^{\frac{n}{m}}_{\frac{n}{m}}+\|u\|^{\frac{n}{m}}_{\frac{n}{m}}\le 1} \int_{\mathbb{R}^n} \Phi_{n,m}(\beta_{n,m}|u|^{\frac{n}{n-m}})\,dx \le C_{m,n}. \tag{1-6}$$

Note that in (0-1), we assume $V(x)$ has a positive lower bound, thus we need an Adams inequality involved with the norm $(\|\nabla^m u\|^{n/m}_{n/m} + \tau\|u\|^{n/m}_{n/m})^{m/n}$. We utilize the change of variable to obtain the following result.

Theorem 1.1. *For any $\tau > 0$ and $0 < \alpha \le \beta_{n,m}$, there exists some constant $C_{m,n}$ such that for $u \in W^{m,n/m}(\mathbb{R}^n)$ with $\|\nabla^m u\|^{n/m}_{n/m} + \tau\|u\|^{n/m}_{n/m} \le 1$,*

$$\int_{\mathbb{R}^n} \Phi_{n,m}(\alpha|u|^{\frac{n}{n-m}})\,dx \le C_{m,n}. \tag{1-7}$$

Theorem 1.2. *For any $\tau > 0$, $0 \le t < n$ and $0 < \alpha < \beta_{n,m}$, there exists some constant $C_{m,n}$ such that for $u \in W^{m,n/m}(\mathbb{R}^n)$ with $\|\nabla^m u\|_{n/m}^{n/m} + \tau\|u\|_{n/m}^{n/m} \le 1$,*

$$\int_{\mathbb{R}^n} \frac{\Phi_{n,m}\big(\alpha\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx \le C_{m,n}. \tag{1-8}$$

Remark 1.3. In fact, the inequality (1-8) still holds in the case of $\alpha = \beta_{n,m}$. However, in order to prove the concentration-compactness principle for the Adams inequality, we only need inequality (1-5). For convenience, we also give the proof of the critical case of Theorem 1.2.

The purpose of proving such inequalities is to prove the following sharp version of concentration-compactness principle for weighted Adams inequalities in $W^{m,2}(\mathbb{R}^{2m})$. For simplicity, we define a new function space

$$E = \left\{u \in W^{m,2}(\mathbb{R}^{2m}) : \|u\|_E^2 = \int_{\mathbb{R}^{2m}} |\nabla^m u|^2 + V(x)|u|^2\,dx < \infty\right\},$$

where $V(x) \ge c_0$ $(c_0 > 0)$.

Theorem 1.4. *For $0 \le t < 2m$, assume $\{u_k\}_k$ is a sequence in E satisfying $\|u_k\|_E^2 = 1$ and $u_k \rightharpoonup u \not\equiv 0$ in E. If*

$$0 < p < p_{2m,m}(u) := \frac{1}{1-\|u\|_E^2},$$

then

$$\sup_k \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)pu_k^2\big)}{|x|^t}\,dx < \infty, \tag{1-9}$$

$$\text{where}\quad \Phi_{2m,m}(t) = e^t - 1 \quad\text{and}\quad \beta_{2m,m} = \frac{2m}{\omega_{2m-1}}\pi^{2m}2^{2m}.$$

Furthermore, for any positive constant c, if $V(x) = c$, the constant $p_{2m,m}(u)$ is sharp in the sense that if $p \ge p_{2m,m}(u)$, the supremum will become infinite.

Remark 1.5. Theorem 1.4 is an extension of do Ó's result [2014] which relies heavily on the Pólya–Szegő inequality. Therefore, the methods they used cannot be applied to obtain the concentration-compactness principle of the Adams inequality on $\mathbb{R}^n$ or the Trudinger–Moser inequality in settings where a rearrangement argument fails such as the Heisenberg group $\mathbb{H}^n$. Recently, Li, Lu and Zhu [Li et al. 2018b] developed a symmetrization-free approach and established Lions concentration-compactness of the singular Trudinger–Moser inequality on the Heisenberg group $\mathbb{H}$. The method is rearrangement-free and can be easily applied to other settings. In the present paper, we use a different approach to prove Theorem 1.4. This is due to the fact that the Sobolev space $W^{m,2}(\mathbb{R}^n)$ we are dealing with is a Hilbert space. Analyzing the energy loss when taking the weak limit is an essential part in proving concentration-compactness principle and for the Hilbert space, the weak limit is relatively simple and with the help of the Brezis–Lieb lemma (Lemma 3.1), we are able to develop a different proof.

Remark 1.6. Nguyen [2016] took advantage of the Talenti comparison theorem to obtain inequality (1-9) in the case of $V(x) = 1$ and $t = 0$. However, they did not verify the sharpness of $p_{2m,m}(u)$. By constructing a proper sequence, we also verify that the supremum in (1-9) becomes infinite if $p \geq p_{2m,m}(u)$.

Recently, another improved version of the sharp Adams inequality was investigated by Lam, Lu and Tang [Lam et al. 2017a] in the spirit of Lions' work [1985]. Their result can be stated as follows:

$$\sup_{\substack{u\in W^{2,m}(\mathbb{R}^{2m}) \\ \int_{\mathbb{R}^{2m}} |\Delta u|^m+\tau|u|^m\,dx\leq 1}} \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,2}\Big(\frac{1/(2^{m-1}\alpha)}{(1+\|\Delta u\|_m^m)^{\frac{1}{m-1}}}|u|^{\frac{m}{m-1}}\Big)}{|x|^\beta}\,dx \leq C(m,\beta,\tau) \tag{1-10}$$

for $0 \leq \beta < 2m,\ \tau > 0$ and $0 \leq \alpha \leq \big(1 - \frac{\beta}{2m}\big)\beta_{2m,2}$. It is easy to verify that the above inequality is stronger than the general Adams inequalities in $W^{2,m}(\mathbb{R}^{2m})$.

Adams inequalities (1-5) and (1-10) are often used to study nonlinear equations related to the Bessel potential. Bao, Lam and Lu [Bao et al. 2016] considered polyharmonic equations of the form

$$(I-\Delta)^m u = f(x,u) \quad \text{in } \mathbb{R}^{2m}. \tag{1-11}$$

Yang [2012a] exploited the following bi-Laplacian equation with small perturbation

$$\Delta^2 u - \operatorname{div}(a(x)\nabla u) + b(x)u = \frac{f(x,u)}{|x|^\beta} + \varepsilon h(x) \quad \text{in } \mathbb{R}^4, \tag{1-12}$$

where $f(x,u)$ has exponential growth and $V(x)$ satisfies $\lim_{|x|\to+\infty} V(x) = +\infty$.

Recently, Chen, Li, Lu and Zhang [Chen et al. 2018] considered the following equation in $\mathbb{R}^4$:

$$(-\Delta)^2 u + V(x)u = \frac{f(x,u)}{|x|^\beta} \quad \text{in } \mathbb{R}^4, \quad 0<\beta<4, \tag{1-13}$$

where $V(x) \geq c_0$ and $f(x,t)$ satisfies some critical exponential growth. They established the existence of the ground state solutions.

Motivated by the work [Chen et al. 2018], we will study the existence of ground state solutions for the following polyharmonic equations with singular nonlinear term

$$(-\Delta)^m u + V(x)u = \frac{f(x,u)}{|x|^\beta} \quad \text{in } \mathbb{R}^{2m}, \quad 0<\beta<2m, \tag{1-14}$$

where $V(x) \geq c_0$ and $f(x,t)$ has critical exponential growth. Furthermore, we assume that $f(x,t)$ satisfies the following conditions.

(H_0) The nonlinearity $f(x,t) : \mathbb{R}^{2m} \times \mathbb{R} \to \mathbb{R}$ is continuous, $f(x,t) = 0$ at $(x,0)$, and has exponential growth as $t \to +\infty$, which means there exists a constant $\alpha_0 > 0$ such that

$$\lim_{t\to+\infty} f(x,t)e^{-\alpha|t|^2} = \begin{cases} 0 & \text{for all } \alpha > \alpha_0, \\ +\infty & \text{for all } \alpha < \alpha_0, \end{cases}$$

uniformly in $x \in \mathbb{R}^{2m}$.

(H_1) There exist constants $\alpha_0, b_1, b_2 > 0$ such that for any $(x,t) \in \mathbb{R}^{2m} \times (0,+\infty)$,

$$0 < f(x,t) \leq b_1 t + b_2 \Phi_{2m,m}(\alpha_0 t^2), \quad \text{where } \Phi_{2m,m}(t) = e^t - 1.$$

(H_2) There exist constants t_0 and $M_0 > 0$ such that

$$0 < F(x,t) := \int_0^t f(x,s)\,ds \leq M_0 f(x,t) \quad \text{for all } (x,t) \in \mathbb{R}^{2m} \times [t_0,+\infty).$$

(H_3) There exists a constant $\theta > 2$ such that for all $x \in \mathbb{R}^{2m}$ and $t > 0$,

$$0 < \theta F(x,t) \leq f(x,t)t.$$

(H_4) $\displaystyle\limsup_{t\to 0+} \frac{2F(x,t)}{|t|^2} < \lambda_\beta$ uniformly in $x \in \mathbb{R}^{2m}$,

$$\text{where} \quad \lambda_\beta = \inf_{u\in E} \frac{\int_{\mathbb{R}^{2m}} |\nabla^m u|^2 + V(x)|u|^2\,dx}{\int_{\mathbb{R}^{2m}} |u|^2/|x|^\beta\,dx}.$$

(H_5) There exist constants $p > 2$ and C_p such that for all $(x,t) \in \mathbb{R}^{2m} \times (0,+\infty)$,

$$f(x,t) \geq C_p t^{p-1},$$

$$\text{where} \quad C_p > \left(\frac{\beta_{2m,m}\left(1-\frac{\beta}{2m}\right)}{\alpha_0}\right)^{\frac{(2-p)}{2}} \left(\frac{p-2}{p}\right)^{\frac{p-2}{2}} S_p^p$$

$$\text{and} \quad S_p^2 := \inf_{u\in E} \frac{\int_{\mathbb{R}^{2m}} |\nabla^m u|^2 + V(x)|u|^2\,dx}{\left(\int_{\mathbb{R}^{2m}} |u|^p/|x|^\beta\,dx\right)^{\frac{2}{p}}}.$$

(H_6) The function $\frac{f(x,t)}{t}$ is increasing for $t > 0$.

By (H_2) and (H_3), we can get that for all $(x,t) \in \mathbb{R}^{2m} \times [0,+\infty)$, there exists $\mu > 0$ such that

$$0 < F(x,t) \leq \mu f(x,t).$$

This result together with (H_1) and the singular Adams inequality in $W^{m,2}(\mathbb{R}^{2m})$ yields the boundedness of $F(x,u)$ and $f(x,u)v$ in $L^1(\mathbb{R}^{2m}, |x|^{-\beta}dx)$ for any

$u, v \in E$. Hence, one can easily find the functional related with polyharmonic equation (1-14), given by

$$I_\beta(u) = \tfrac{1}{2}\|u\|_E^2 - \int_{\mathbb{R}^{2m}} \frac{F(x,u)}{|x|^\beta}\,dx,$$

is well defined. With standard calculations, it is easy to obtain that $I_\beta \in C^1(E, \mathbb{R})$ and

$$I'_\beta(u)v = \int_{\mathbb{R}^{2m}} (\nabla^m u \nabla^m v + V(x)uv)\,dx - \int_{\mathbb{R}^{2m}} \frac{f(x,u)v}{|x|^\beta}\,dx, \quad u, v \in E.$$

Since the weak solutions of (1-14) are equivalent to the critical points of functional I_β, we focus our attention on critical points of functional I_β. Equation (1-14) is different from Equations (1-11) and (1-12). Unlike Bao, Lam and Lu's result [Bao et al. 2016], we do not necessarily assume that $f(x,t)$ satisfies some periodicity conditions. Moreover, the presence of potential $V(x)$ of (1-14) makes it difficult to directly apply Yang's argument in [Yang 2012a]. Thus a new compactness embedding in $W^{m,2}(\mathbb{R}^{2m})$ must be established and we observe that the weight term $1/|x|^\beta$ provides a good control to the integral away from zero, which enables us to establish the following compactness result.

Theorem 1.7. *The Sobolev space $W^{m,2}(\mathbb{R}^{2m})$ can be compactly embedded into $L^q(\mathbb{R}^{2m}, |x|^{-s}\,dx)$ when $q \geq 2$ and $0 < s < 2m$.*

Remark 1.8. In view of $E \hookrightarrow W^{m,2}(\mathbb{R}^{2m})$ and Theorem 1.7, we can derive that E can be compactly embedded into $L^q(\mathbb{R}^{2m}, |x|^{-s}\,dx)$ for $q \geq 2$ and $0 < s < 2m$.

With the help of Theorem 1.7, our next result will concern the existence of the ground state solution of polyharmonic equation (1-14).

Theorem 1.9. *Assume $f(x,t)$ satisfies (H_1)–(H_6), then (1-14) has a ground state solution.*

In the case of $\beta = 0$, (1-14) becomes the following nonsingular polyharmonic equation

$$(-\triangle)^m u + V(x)u = f(x,u) \quad \text{in } \mathbb{R}^{2m}. \tag{1-15}$$

The existence of the ground state solution of (1-15) cannot be obtained immediately from Theorem 1.9 due to the absence of compactness embedding. There is a common constrained minimization theory to deal with this problem. Unfortunately, this method crucially depends on the rearrangement inequality which is not obvious available in $W^{m,2}(\mathbb{R}^{2m})$. In order to overcome this difficulty, we use the principle of the symmetric criticality of the Hilbert space. By assuming $f(x,t)$ and $V(x)$ are radially symmetric with respect to x, one can carry out the same process as what we do in Theorem 1.9 to derive a nontrivial weak solution of the polyharmonic equation with nonsingular nonlinearity (1-15). However, whether there exists a ground state solution to (1-15) is still open. In a very recent work of Chen, Lu and Zhu [Chen et al. 2019], they made the first attempt in this direction. They derive the existence of ground state solutions to (1-15) when $m=2$, V is a trapping potential and

$$f(x,u)=u\exp(2u^2).$$

Theorem 1.10. *Under the assumptions of Theorem 1.9, if we additionally assume that $V(x)$ and $f(x,t)$ are radially symmetric in x, $f(x,t)=o(t)$ at origin, then polyharmonic equation with nonsingular linearity (1-15) has a nontrivial weak solution.*

The plan of the paper is as follows. In Section 2, we employ the change of variable to establish some weighted Adams inequalities in $W^{m,2}(\mathbb{R}^{2m})$ involved with the norm $(\|\nabla^m u\|_{n/m}^{n/m}+\tau\|u\|_{n/m}^{n/m})^{m/n}$ for any $\tau>0$. Sections 3 and 4 are devoted to the concentration-compactness principle for the weighted Adams inequality and a new compactness embedding in $W^{m,2}(\mathbb{R}^{2m})$. As an immediate application of Theorem 1.4, in Section 5, we give sufficient conditions to guarantee the existence of ground state solutions for the polyharmonic equation with singular exponential nonlinearity term. Finally, in Section 6, we also derive the existence of a nontrivial weak solution for the polyharmonic equation (1-15) through the principle of the symmetric criticality.

2. Proof of Theorems 1.1 and 1.2

In this section, we will utilize a change of variable to establish Adams inequality (1-7).

Proof. For any $\tau>0$, $0<\alpha\le\beta_{n,m}$ and $u\in W^{m,2}(\mathbb{R}^{2m})$ with

$$\int_{\mathbb{R}^n}|\nabla^m u|^{\frac{n}{m}}+\tau|u|^{\frac{n}{m}}\,dx\le 1,$$

we denote a new function $v(x)$ given by $v(\tau^{1/n}x)=u(x)$. Then direct computations yield that

$$\int_{\mathbb{R}^n}|\nabla^m v|^{\frac{n}{m}}+|v|^{\frac{n}{m}}\,dx=\int_{\mathbb{R}^n}|\nabla^m u|^{\frac{n}{m}}+\tau|u|^{\frac{n}{m}}\,dx\le 1$$

and

$$\int_{\mathbb{R}^n}\Phi_{n,m}(\alpha|u|^{\frac{n}{n-m}})\,dx=\frac{1}{\tau}\int_{\mathbb{R}^n}\Phi_{n,m}(\alpha|v|^{\frac{n}{n-m}})\,dx.$$

Combining this with inequality (1-6), we obtain

$$\sup_{\|\nabla^m u\|_{\frac{n}{m}}^{\frac{n}{m}}+\tau\|u\|_{\frac{n}{m}}^{\frac{n}{m}}\le 1}\int_{\mathbb{R}^n}\Phi_{n,m}(\alpha|u|^{\frac{n}{n-m}})\,dx \le \frac{1}{\tau}\sup_{\|\nabla^m v\|_{\frac{n}{m}}^{\frac{n}{m}}+\|v\|_{\frac{n}{m}}^{\frac{n}{m}}\le 1}\int_{\mathbb{R}^n}\Phi_{n,m}(\alpha|v|^{\frac{n}{n-m}})\,dx\le C(m,n).$$

Next, it suffices to show that inequality (1-8) still holds for any $\tau>0$ and $0\le t<n$. In fact, we have

$$\int_{\mathbb{R}^n}\frac{\Phi_{n,m}\big(\alpha\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx \le \int_{|x|\le 1}\frac{\Phi_{n,m}\big(\alpha\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx+\int_{|x|\ge 1}\Phi_{n,m}\big(\alpha(1-\tfrac{t}{n})|u|^{\frac{n}{n-m}}\big)\,dx. \tag{2-1}$$

This together with (1-7) and the Hölder inequality leads to

$$\sup_{\|\nabla^m u\|_{\frac{n}{m}}^{\frac{n}{m}}+\tau\|u\|_{\frac{n}{m}}^{\frac{n}{m}}\le 1}\int_{\mathbb{R}^n}\frac{\Phi_{n,m}\big(\alpha\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx\lesssim 1$$

for any $0<\alpha<\beta_{n,m}$.

Finally, we prove that inequality (1-8) still holds in the case of $\alpha=\beta_{n,m}$. Following the same line of the proof of Theorem 1 in [Fontana and Morpurgo 2015], we can obtain that for any $u\in W^{m,n/m}(\mathbb{R}^n)$ satisfying $\int_{\mathbb{R}^n}|\nabla^m u|^{\frac{n}{m}}+|u|^{\frac{n}{m}}\,dx\le 1$,

$$\int_{\Omega}\frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx\lesssim(1+|\Omega|^{1-\frac{t}{n}}), \tag{2-2}$$

where Ω is any bounded domain of $\mathbb{R}^n$. Let

$$A:=\{x\in\mathbb{R}^n:|u(x)|\ge 1\}.$$

Since $u\in W^{m,n/m}(\mathbb{R}^n)$, it is obvious that A is a bounded domain. Now, we split the integral over $\mathbb{R}^n$ into two parts:

$$\begin{aligned}\int_{\mathbb{R}^n}&\frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx\\ &=\int_A\frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx+\int_{\mathbb{R}^n\setminus A}\frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx\\ &=I_1+I_2.\end{aligned}$$

For I_1, using the estimate (2-2), we obtain that

$$I_1 = \int_A \frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx \lesssim 1.$$

For I_2, since

$$\int_{\mathbb{R}^n\setminus A} \frac{\Phi_{n,m}\big(\beta_{n,m}\big(1-\frac{t}{n}\big)|u|^{\frac{n}{n-m}}\big)}{|x|^t}\,dx \lesssim \int_{\{|x|\le 1,\,|u|\le 1\}} \frac{|u|^{\frac{n}{m}}}{|x|^t}\,dx + \int_{\{|x|\ge 1,\,|u|\le 1\}} \frac{|u|^{\frac{n}{m}}}{|x|^t}\,dx \lesssim 1,$$

we derive that $I_2 \lesssim 1$. Combining the above estimates, we derive inequality (1-8) in the case of $\alpha = \beta_{n,m},\ \tau = 1$. Carrying out the same procedure as the proof of Theorem 1.1, one can conclude that inequality (1-8) still holds for any $\tau > 0$. □

3. The proof of Theorem 1.4

Our purpose in this section is to prove Theorem 1.4. Namely, we will give the proof of the concentration-compactness principle for weighted Adams inequalities. Our proof relies on the following lemmas.

Lemma 3.1 [Brézis and Lieb 1983]. *Let Ω be an open subset of $\mathbb{R}^n$ and $\{u_k\}_k \subseteq L^p(\Omega)$ $(1 \le p < \infty)$. If $\{u_k\}_k$ satisfies the following conditions*:

(i) *$\{u_k\}_k$ is bounded in $L^p(\Omega)$,*

(ii) *$u_k \to u$ almost everywhere in Ω,*

then

$$\lim_{k\to\infty} (\|u_k\|_p^p - \|u_k - u\|_p^p) = \|u\|_p^p.$$

Lemma 3.2. *Let $\Omega \subseteq \mathbb{R}^n$ be an open domain and $\{f_k\}_k \subseteq W^{m,n/m}(\Omega)$ that strongly converges to f in $W^{m,n/m}(\Omega)$. Then there exists a subsequence $\{f_{k_j}\}_j$ and a positive function $g \in W^{m,n/m}(\Omega)$ such that*

$$f_{k_j}(x) \to f(x) \quad \text{a.e. in } \Omega \text{ as } j \to +\infty,$$

and

$$|f_{k_j}(x)| \le g(x) \quad \text{a.e. in } \Omega \text{ for all } j.$$

Remark 3.3. Since the proof of Lemma 3.2 is similar to that of Proposition 1 in [do Ó et al. 2009], we omit the details.

Proof of Theorem 1.4. At first, we show the proof of inequality (1-9). It follows from the semicontinuity of the norm in E that

$$\|u\|_E^2 \le \liminf_k \|u_k\|_E^2 = 1.$$

We carry out the process by considering the following two cases.

Case 1. $\|u\|_E^2 = 1$. Applying the Brezis–Lieb lemma (Lemma 3.1) on the Hilbert space E, one can show that $u_k \to u$ strongly in E. In light of Lemma 3.2, we can find a subsequence $\{u_{k_j}\}_j$ and a positive function $v \in E$ such that $|u_{k_j}(x)| \le v(x)$. Then it follows that

$$\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)pu_{k_j}^2\big)}{|x|^t}\,dx \le \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)pv^2\big)}{|x|^t}\,dx < \infty. \tag{3-1}$$

Case 2. $0 < \|u\|_E^2 < 1$. Defining $\Psi(X) = \Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)pX\big)$ for notational convenience, one can write that

$$\begin{aligned} \sup_k \int_{\mathbb{R}^{2m}} \frac{\Psi(u_k^2)}{|x|^t}\,dx &\le \sup_k \int_{\mathbb{R}^{2m}} \frac{\Psi\big((1+\varepsilon)(u_k-u)^2 + C_\varepsilon u^2\big)}{|x|^t}\,dx \\ &= \sup_k \int_{\mathbb{R}^{2m}} \frac{\Psi((1+\varepsilon)(u_k-u)^2)\Psi(C_\varepsilon u^2)}{|x|^t}\,dx \\ &\qquad + \sup_k \int_{\mathbb{R}^{2m}} \frac{\Psi((1+\varepsilon)(u_k-u)^2)}{|x|^t}\,dx + \sup_k \int_{\mathbb{R}^{2m}} \frac{\Psi(C_\varepsilon u^2)}{|x|^t}\,dx \\ &=: I_1 + I_2 + I_3, \end{aligned} \tag{3-2}$$

where we use the elementary inequality which states

$$(a+b)^2 \le (1+\varepsilon)a^2 + C_\varepsilon b^2 \quad \text{for } a, b \ge 0 \text{ and } \varepsilon > 0.$$

For I_1, as an immediately consequence of the Hölder inequality and the singular Adams inequality, we can derive that

$$I_1 \lesssim \left(\sup_k \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)pr(1+\varepsilon)(u_k-u)^2\big)}{|x|^t}\,dx\right)^{\frac{1}{r}}, \tag{3-3}$$

where r is sufficiently close to 1. Noting that $u_k \rightharpoonup u$ weakly in E and E is a Hilbert space, one can apply the Brezis–Lieb lemma to derive that

$$\|u_k - u\|_E^2 = \|u_k\|_E^2 - \|u\|_E^2 = 1 - \|u\|_E^2,$$

which yields that

$$\beta_{2m,m}\big(1-\tfrac{t}{2m}\big)pr(1+\varepsilon)(\|\nabla^m(u_k-u)\|_2^2 + c_0\|u_k-u\|_2^2) < \beta_{2m,m}\big(1-\tfrac{t}{2m}\big).$$

Combining this with Theorem 1.2 with $\tau = c_0$, we conclude that $I_1 < +\infty$. Similarly, we can obtain that $I_2 < +\infty$. Thus, we accomplish the proof of inequality (1-9).

Next, we are ready to show that $p_{2m,m}(u)$ is sharp when $V(x)$ is constant. Without loss of generality, we assume $V(x) = 1$. The idea of proving this sharpness follows from the result of do Ó et al. [2014]. Similarly, we construct a sequence $\{u_k\}_k \subseteq W^{m,2}(\mathbb{R}^{2m})$ and a function $u \in W^{m,2}(\mathbb{R}^{2m})$ such that

$$\|u_k\| = 1, \quad u_k \rightharpoonup u \not\equiv 0 \text{ in } W^{m,2}(\mathbb{R}^{2m}), \quad \|u\| = \delta < 1,$$

but

$$\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)p_{2m,m}(u)u_k^2\big)}{|x|^t}\,dx \to \infty.$$

We denote a sequence $\{w_k\}_k \subseteq W^{m,2}(\mathbb{R}^{2m})$ by

$$w_k(x) = \begin{cases} \frac{1}{(2m-2)!!(2m)^{\frac{1}{2}}}\omega_{2m-1}^{-\frac{1}{2}}k^{\frac{1}{2}} & \text{if } |x| \in [0, re^{\frac{-k}{2m}}], \\ \omega_{2m-1}^{-\frac{1}{2}}\frac{(2m)^{\frac{1}{2}}}{(2m-2)!!}\ln\big(\frac{r}{|x|}\big)k^{-\frac{1}{2}} & \text{if } |x| \in [re^{\frac{-k}{2m}}, r], \\ 0 & \text{if } |x| \in [r, +\infty), \end{cases}$$

where $r > 0$ to be chosen later. Simple calculations show that

$$w_k \rightharpoonup 0 \text{ in } W^{m,2}(\mathbb{R}^{2m}), \quad \|\nabla^m w_k\|_2^2 = 1, \quad \|w_k\|_2^2 = O(k^{-1}).$$

Next, we define a new function $u : \mathbb{R}^{2m} \to \mathbb{R}$ given by

$$u(x) = \begin{cases} A & \text{if } |x| \in [0, \frac{2R}{3}], \\ \big(1-\big(\frac{2}{3}\big)^m\big)^{-1}\big(A - \frac{A}{R^m}|x|^m\big) & \text{if } |x| \in [\frac{2R}{3}, R], \\ 0 & \text{if } |x| \in [R, +\infty), \end{cases}$$

where $R = 3r$ and A is a positive constant which needs to be chosen later. Then

$$\begin{aligned} \|u\|^2 &= \|u\|_2^2 + \|\nabla^m u\|_2^2 \\ &= \frac{\omega_{2m-1}}{2m}\big(\tfrac{2}{3}R\big)^{2m}A^2 \\ &\quad + \omega_{2m-1}\int_{\frac{2R}{3}}^{R}\Big(\big(1-\big(\tfrac{2}{3}\big)^m\big)^{-1}\Big(A - \frac{A}{R^m}r^m\Big)\Big)^2 r^{2m-1}\,dr \\ &\qquad + \big(1-\big(\tfrac{2}{3}\big)^m\big)^{-2}\Big(\frac{A}{R^m}\Big)^2\omega_{2m-1}\int_{\frac{2R}{3}}^{R}\frac{m!!(3m-2)!!}{(2m-2)!!}r^{2m-1}\,dr \\ &= CA^2. \end{aligned} \tag{3-4}$$

Picking A satisfying $\|u\| = \delta < 1$, a direct application of the Hölder inequality yields that

$$\begin{aligned}\|v_k\|_2^2 &:= \|u + (1-\delta^2)^{\frac{1}{2}} w_k\|_2^2 \\ &= \int_{\mathbb{R}^{2m}} |u + (1-\delta^2)^{\frac{1}{2}} w_k|^2 dx \\ &= \int_{\mathbb{R}^{2m}} u^2 + 2(1-\delta^2)^{\frac{1}{2}} u w_k + (1-\delta^2) w_k^2 dx \\ &= \|u\|_2^2 + \eta_k,\end{aligned} \tag{3-5}$$

where

$$\eta_k = \left(\frac{1}{m} A (1-\delta^2)^{\frac{1}{2}} \frac{(2m)^{\frac{1}{2}}}{(2m-2)!!} \omega_{2m-1}^{\frac{1}{2}} \left(r^{2m} e^{-k} \frac{k}{2m} + \frac{r^{2m}}{2m} - \frac{r^{2m}}{2m} e^{-k}\right)\right) k^{-\frac{1}{2}} + O(k^{-1}).$$

It is clear that $\nabla^m u$ and $\nabla^m w_k$ have disjoint supports, so

$$\|\nabla^m v_k\|_2^2 = \|\nabla^m u\|_2^2 + (1-\delta^2) \quad \text{and} \quad \|v_k\|^2 = 1 + \eta_k. \tag{3-6}$$

Let $u_k = v_k/(1+\eta_k)^{1/2}$; one can easily see that

$$\|u_k\| = 1 \quad \text{and} \quad u_k \rightharpoonup u \text{ in } W^{m,2}(\mathbb{R}^{2m}).$$

Consequently,

$$\begin{aligned}&\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big) p_{2m,m}(u) u_k^2\big)}{|x|^t} dx \\ &\geq \int_{B_{re^{\frac{-k}{2m}}}} \frac{\exp\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)(1-\delta^2)^{-1} u_k^2\big)}{|x|^t} dx - \int_{B_{re^{\frac{-k}{2m}}}} \frac{1}{|x|^t} dx \\ &= \int_{B_{re^{\frac{-k}{2m}}}} \frac{\exp\big(\beta_{2m,m}\big(1-\frac{t}{2m}\big)((1+\eta_k)^{-\frac{1}{2}} (A + (1-\delta^2)^{\frac{1}{2}} w_k))^2 (1-\delta^2)^{-1}\big)}{|x|^t} dx + C \\ &= \int_{B_{re^{\frac{-k}{2m}}}} \frac{\exp\Big(\big(1-\frac{t}{2m}\big)\Big((1+\eta_k)^{-\frac{1}{2}} \big(\frac{\beta_{2m,m}^{1/2} A}{(1-\delta^2)^{1/2}} + k^{\frac{1}{2}}\big)\Big)^2\Big)}{|x|^t} dx + C \\ &\gtrsim \exp\left(\left(1-\frac{t}{2m}\right)\left(\left((1+\eta_k)^{-\frac{1}{2}} \left(\frac{\beta_{2m,m}^{\frac{1}{2}} A}{(1-\delta^2)^{\frac{1}{2}}} + k^{\frac{1}{2}}\right)\right)^2 - k\right)\right) r^{2m-t} + C \to +\infty,\end{aligned} \tag{3-7}$$

where $r < 1$ is selected in such a way that $\eta_k < \frac{\beta_{2m,m}^{1/2} A}{(1-\delta^2)^{1/2}} k^{-\frac{1}{2}}$. Then Theorem 1.4 is completed. □

4. the proof of Theorem 1.7

In this section, we begin with a simple fact that the norm $(\|\nabla^m u\|_2^2+\|u\|_2^2)^{1/2}$ and the standard Sobolev norm given by

$$\|u\|_{W^{m,2}}=\left(\sum_{j=0}^{j=m}\|\nabla^j u\|_2^2\right)^{\frac{1}{2}}$$

is equivalent. In fact, for any $u\in C_c^\infty(\mathbb{R}^{2m})$, through Caffarelli–Kohn–Nirenberg inequalities [Lin 1986], one can derive that

$$(*)\qquad \int_{\mathbb{R}^{2m}}|\nabla^j u|^2\,dx\le\left(\int_{\mathbb{R}^{2m}}|u|^2\,dx\right)^{\frac{j}{m}}\left(\int_{\mathbb{R}^{2m}}|\Delta u|^2\,dx\right)^{1-\frac{j}{m}}.$$

Then a simple density argument implies that $(*)$ also holds for $u\in W^{m,2}(\mathbb{R}^{2m})$. Now, we are in a position to show that a Sobolev space equipped with the norm $(\|\nabla^m u\|_2^2+\|u\|_2^2)^{1/2}$ can be compactly embedded into $L^p(\mathbb{R}^{2m},|x|^{-\beta}dx)$ for any $p\ge 2$ and $0<\beta<2m$.

Proof. Continuous embedding is an easy consequence of the Adams inequality (1-5). Our aim is to show that the above continuous embedding is compact. In light of $W^{m,2}(\mathbb{R}^{2m})\hookrightarrow\hookrightarrow L^q_{\mathrm{loc}}(\mathbb{R}^{2m})$ for $q\ge 1$, one can find a subsequence $\{u_{k_j}\}_j$ such that

$$u_{k_j}(x)\to u(x),\quad \text{strongly in } L^q(B_R(0)) \text{ for any } R>0,$$
$$u_{k_j}(x)\to u(x),\quad \text{for almost every } x\in\mathbb{R}^{2m}.$$

Therefore, our purpose is to show that

$$(4\text{-}1)\qquad u_{k_j}\to u\quad\text{in } L^q(\mathbb{R}^{2m},|x|^{-s}dx).$$

Applying the Egorov theorem, we obtain that for any $B_R(0)$ and $\delta>0$,

$$\text{there exists } E_\delta\subset B_R(0) \text{ satisfying } m(E_\delta)<\delta,$$

such that

$$u_{k_j} \text{ uniformly converges to } u \text{ in } B_R(0)\setminus E_\delta.$$

Hence, it follows that

$$(4\text{-}2)\quad \lim_{R\to+\infty}\lim_{\delta\to 0}\lim_{j\to+\infty}\int_{\mathbb{R}^{2m}}\frac{|u_{k_j}-u|^q}{|x|^s}\,dx$$
$$=\lim_{R\to+\infty}\lim_{\delta\to 0}\lim_{j\to+\infty}\int_{E_\delta}\frac{|u_{k_j}-u|^q}{|x|^s}\,dx$$
$$+\lim_{R\to+\infty}\lim_{\delta\to 0}\lim_{j\to+\infty}\int_{B_R(0)\setminus E_\delta}\frac{|u_{k_j}-u|^q}{|x|^s}\,dx$$
$$+\lim_{R\to+\infty}\lim_{\delta\to 0}\lim_{j\to+\infty}\int_{\mathbb{R}^{2m}\setminus B_R(0)}\frac{|u_{k_j}-u|^q}{|x|^s}\,dx$$
$$=: I_1+I_2+I_3.$$

By the Hölder inequality and the Sobolev continuous embedding, one can derive that

$$I_1 \le \lim_{\delta\to 0}\lim_{j\to+\infty}\left(\int_{E_\delta} 1\,dx\right)^{\frac{1}{t}}\left(\int_{E_\delta}\frac{|u_{k_j}-u|^{qt'}}{|x|^{st'}}\,dx\right)^{\frac{1}{t'}} \lesssim \lim_{\delta\to 0}\sup_j \|u_{k_j}\|^q (m(E_\delta))^{\frac{1}{t}} = 0, \tag{4-3}$$

where $t>1$ and $st'<2m$. For I_2, the uniform convergence of u_{k_j} in $B_R(0)\backslash E_\delta$ yields that $I_2=0$. For I_3, the Sobolev continuous embedding $W^{m,2}(\mathbb{R}^{2m})\hookrightarrow L^q(\mathbb{R}^{2m})$ for $q\ge 2$ yields that

$$I_3 \le \lim_{R\to+\infty}\lim_{\delta\to 0}\lim_{j\to+\infty}\frac{1}{R^s}\int_{\mathbb{R}^{2m}\backslash B_R(0)}|u_{k_j}-u|^q\,dx \lesssim \lim_{R\to+\infty}\sup_j\|u_{k_j}\|^q\frac{1}{R^s} = 0. \tag{4-4}$$

Thus, we have accomplished the proof of Theorem 1.7. □

As a direct result of Theorem 1.7 and Remark 1.8, we can easily see that the best constant S_p $(p\ge 2)$ in (H_3) could be achieved (one can refer to [Zhang and Chen 2018] for details).

5. The proof of Theorem 1.9

This section is devoted to the proof of Theorem 1.9. We carry out the proof in three parts. In Part 1, we use the mountain-pass theorem without the Palais–Smale compactness condition to derive the existence of weak solutions of (1-14) satisfying hypotheses (H_1)–(H_4). Therefore, in Part 2, we utilize the method combining the concentration-compactness principle and the new compactness theorem in $W^{m,2}(\mathbb{R}^{2m})$ to verify that the functional I_β satisfies the Palais–Smale compactness condition. Part 3 is devoted to showing that the critical point of the functional I_β is actually a ground state solution of polyharmonic equation (1-14). Before starting the proof, we need a couple of important lemmas for which we omit the proofs.

Lemma 5.1 [Badiale and Serra 2011]. *Let X be a Hilbert space, $\varphi\in C^2(X,\mathbb{R})$, $e\in X$ and $r>0$ such that $\|e\|>r$ and $b:=\inf_{\|u\|=r}\varphi(u)>\varphi(0)\ge\varphi(e)$. Define*

$$c=\inf_{g\in\Gamma}\max_{s\in[0,1]}\varphi(g(s)),$$

where

$$\Gamma:=\{g\in C([0,1],X): g(0)=0,\ g(1)=e\}.$$

Then there exists a sequence $\{u_k\}_k \in X$ *such that* $\varphi(u_k) \to c,\ \varphi'(u_k) \to 0$ *as* $k \to +\infty$.

Remark 5.2. In the case of $p = 2$, one can use the property of the Hilbert space to replace $u_k \to u$ almost everywhere in Ω with $u_k \rightharpoonup u$.

Now, we are ready to start the proof of Theorem 1.9.

Part 1. In this part, we first check that $I_\beta(u)$ satisfies geometric conditions without the Palais–Smale compactness condition.

Lemma 5.3. *Assume* (H$_1$)–(H$_4$) *hold. Then*

(i) *there exist constants* $\delta, \rho > 0$ *such that* $I_\beta(u) \geq \delta$ *for any* $\|u\|_E = \rho$,

(ii) *there exists* $e \in E$ *such that* $\|e\|_E > \rho$, *but* $I_\beta(e) < 0$.

Proof. According to (H$_4$), there exist positive constants ε, δ such that for any $|t| \leq \delta$,

$$F(x,t) \leq \tfrac{1}{2}(\lambda_\beta - \varepsilon)|t|^2 \quad \text{for } x \in \mathbb{R}^{2m}. \tag{5-1}$$

Moreover, by (H$_1$), we derive that for any $|t| \geq \delta$ and $x \in \mathbb{R}^{2m}$, there exists constants c_1, c_2 such that

$$F(x,t) \leq c_1|t|^2 + c_2|t|\Phi_{2m,m}(\alpha_0|t|^2) \leq C_\delta |t|^3 \Phi_{2m,m}(\alpha_0|t|^2), \quad \text{where} \quad C_\delta = \frac{c_1}{\delta \Phi_{2m,m}(\alpha_0|\delta|^2)} + \frac{c_2}{\delta^2}. \tag{5-2}$$

Then it follows from (5-1) and (5-2) that

$$F(x,t) \leq \tfrac{1}{2}(\lambda_\beta - \varepsilon)|t|^2 + C|t|^3\Phi_{2m,m}(\alpha_0|t|^2) \quad \text{for all } (x,t) \in \mathbb{R}^{2m} \times \mathbb{R}. \tag{5-3}$$

For sufficiently small $\|u\|_E$, we claim that the following inequality holds:

$$\int_{\mathbb{R}^{2m}} |u|^3 \frac{\Phi_{2m,m}(\alpha_0|u|^2)}{|x|^\beta}\, dx \leq C\|u\|_E^3. \tag{5-4}$$

For the continuity of our work, let us postpone the proof of (5-4).

Suppose (5-4) holds, we can combine (5-3) and (5-4) to arrive at

$$\begin{aligned} I_\beta(u) &= \tfrac{1}{2}\|u\|_E^2 - \int_{\mathbb{R}^{2m}} \frac{F(x,u)}{|x|^\beta}\, dx \\ &\geq \tfrac{1}{2}\|u\|_E^2 - \tfrac{1}{2}(\lambda_\beta - \varepsilon)\int_{\mathbb{R}^{2m}} \frac{|u|^2}{|x|^\beta}\, dx - C\int_{\mathbb{R}^4} |u|^3 \frac{\Phi_{2m,m}(\alpha_0|u|^2)}{|x|^\beta}\, dx \\ &\geq \tfrac{1}{2}\|u\|_E^2 - \tfrac{1}{2}\frac{\lambda_\beta - \varepsilon}{\lambda_\beta}\|u\|_E^2 - C\|u\|_E^3 \\ &= \|u\|_E^2\Big(\frac{\varepsilon}{2\lambda_\beta} - C\|u\|_E\Big). \end{aligned} \tag{5-5}$$

When $\|u\|_E \leq \varepsilon/(2C\lambda_\beta)$, inequality (i) holds.

Now, we give the proof of inequality (5-4). By applying the Hölder inequality and considering the level sets of the function, one can obtain that for $p>1$ and $\frac{1}{p}+\frac{1}{p'}=1$,

$$\begin{aligned}(5\text{-}6)\quad \int_{\mathbb{R}^{2m}} |u|^3 \frac{\Phi_{2m,m}(\alpha_0|u|^2)}{|x|^\beta}\,dx &\lesssim \left(\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(p\alpha_0|u|^2)}{|x|^\beta}\,dx\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^{2m}} \frac{|u|^{3p'}}{|x|^\beta}\,dx\right)^{\frac{1}{p'}} \\ &\lesssim \left(\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(p\alpha_0|u|^2)}{|x|^\beta}\,dx\right)^{\frac{1}{p}} \|u\|_E^3,\end{aligned}$$

where the last inequality comes from the Sobolev continuous embedding $E\hookrightarrow L^q(\mathbb{R}^{2m},|x|^{-\beta}dx)$. Pick $p>1$ sufficiently close to 1 such that

$$p\alpha_0\|u\|^2 \le \beta_{2m,m}\big(1-\tfrac{\beta}{2m}\big)$$

due to the fact that $\|u\|\le\|u\|_E$ is sufficiently small. The singular Adams inequalities in $\mathbb{R}^{2m}$ yield that

$$(5\text{-}7)\qquad \int_{\mathbb{R}^{2m}} |u|^3 \frac{\Phi_{2m,m}(\alpha_0|u|^2)}{|x|^\beta}\,dx \le C\|u\|_E^3.$$

For (ii), it suffices to show that for a fixed $u\in E$,

$$I_\beta(su)\to-\infty \quad \text{as } s\to+\infty.$$

Without loss of generality, we may assume u has bounded support Ω. Through (H_3), one finds that for any $t>0$,

$$\frac{\partial}{\partial t}(\ln F(x,t)) \ge \frac{\theta}{t},$$

which leads to the result $F(x,t)\ge F(x,t_0)t_0^{-\theta}t^\theta$ for some $t_0>0$. Therefore, there exist positive constants c_1, c_2 such that

$$F(x,t)\ge c_1t^\theta - c_2 \quad \text{for } (x,t)\in\Omega\times[0,\infty).$$

Then,

$$\begin{aligned}(5\text{-}8)\qquad I_\beta(su) &= \frac{s^2}{2}\|u\|_E^2 - \int_\Omega \frac{F(x,su)}{|x|^\beta}\,dx \\ &\le \frac{s^2}{2}\|u\|_E^2 - c_1s^\theta \int_\Omega \frac{|u|^\theta}{|x|^\beta}\,dx + c_3|\Omega|^{1-\frac{\beta}{2m}}.\end{aligned}$$

This inequality together with $\theta>2$ implies that

$$I_\beta(su)\to-\infty \quad \text{as } s\to+\infty.$$

The proof of Lemma 5.3 is finished. □

Lemma 5.3 shows that the functional I_β satisfies geometric conditions of the mountain-pass theorem which yields that there exists a Palais–Smale sequence $\{u_k\}_k$ which satisfies $I_\beta(u_k) \to c_\beta$ and $I'_\beta(u_k) \to 0$ as $k \to +\infty$, where

$$c_\beta = \inf_{g \in \Gamma} \max_{s \in [0,1]} I_\beta(g(s)), \quad \Gamma := \big\{ g \in C([0,1], E) : g(0) = 0,\ I(g(1)) < 0 \big\}.$$

Lemma 5.4. *Assume* (H$_1$), (H$_2$) *and* (H$_3$) *hold. Let* $\{u_k\}_k \subset E$ *be an arbitrary Palais–Smale sequence*, *i.e.*,

$$I_\beta(u_k) \to c_\beta, \quad I'_\beta(u_k) \to 0, \quad \text{as } k \to +\infty.$$

Then there exists a subsequence of $\{u_k\}_k$ (still denoted by $\{u_k\}_k$) *and* $u \in E$ *such that*

$$\begin{cases} \dfrac{f(x,u_k)}{|x|^\beta} \to \dfrac{f(x,u)}{|x|^\beta} & \text{strongly in } L^1_{\mathrm{loc}}(\mathbb{R}^{2m}), \\ \dfrac{F(x,u_k)}{|x|^\beta} \to \dfrac{F(x,u)}{|x|^\beta} & \text{strongly in } L^1(\mathbb{R}^{2m}). \end{cases}$$

Furthermore, *u is a weak solution of* (1-14).

Proof. At first, we prove that

$$\int_{\mathbb{R}^{2m}} \frac{F(x,u_k)}{|x|^\beta}\,dx \le C \quad \text{and} \quad \int_{\mathbb{R}^{2m}} \frac{f(x,u_k)u_k}{|x|^\beta}\,dx \le C. \tag{5-9}$$

Let $\{u_k\}_k$ denote a Palais–Smale sequence of the function I_β, i.e.,

$$\tfrac{1}{2}\|u_k\|_E^2 - \int_{\mathbb{R}^{2m}} \frac{F(x,u_k)}{|x|^\beta}\,dx \to c_\beta \quad \text{as } k \to \infty \tag{5-10}$$

and

$$|I'(u_k)v| \le \tau_k \|v\|_E \quad \text{for all } v \in E, \tag{5-11}$$

where $\tau_k \to 0$ as $k \to \infty$. Moreover, taking $v = u_k$ in (5-11), we get

$$\int_{\mathbb{R}^{2m}} \frac{f(x,u_k)u_k}{|x|^\beta}\,dx - \|u_k\|_E^2 \le \tau_k \|u_k\|_E. \tag{5-12}$$

This together with (5-10) and (H$_4$) leads to

$$\begin{aligned} \theta c_\beta + \tau_k \|u_k\|_E &\ge \Big(\frac{\theta}{2} - 1\Big)\|u_k\|_E^2 - \int_{\mathbb{R}^{2m}} \frac{[\theta F(x,u_k) - f(x,u_k)u_k]}{|x|^\beta}\,dx \\ &\ge \Big(\frac{\theta - 2}{2}\Big)\|u_k\|_E^2. \end{aligned}$$

Thus, $\|u_k\|_E$ is bounded. Combine this with (5-10) and (5-12), we can get (5-9). Since $\|u_k\|_E$ is bounded. Thanks to Theorem 1.7, we can assume that up to a sequence,

$$\begin{aligned} u_k &\rightharpoonup u, && \text{weakly in } E, \\ u_k &\to u, && \text{strongly in } L^q(\mathbb{R}^{2m}, |x|^{-\beta}dx) \text{ for all } q \geq 2, \\ u_k(x) &\to u(x), && \text{for almost every } x \in \mathbb{R}^{2m}. \end{aligned}$$

By hypothesis (H_1), through similar arguments to Lemma 2.1 in [de Figueiredo et al. 1995], we derive that

$$\frac{f(x,u_k)}{|x|^\beta} \to \frac{f(x,u)}{|x|^\beta} \quad \text{strongly in} \quad L^1_{\mathrm{loc}}(\mathbb{R}^{2m}). \tag{5-13}$$

To show the convergence of $\int_{\mathbb{R}^{2m}} \frac{F(x,u_k)}{|x|^\beta}\,dx$, one can write

$$\int_{\mathbb{R}^{2m}} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx = \int_{B_R} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx + \int_{\mathbb{R}^{2m}\setminus B_R} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx.$$

According to (H_2) and (H_3), there exists a positive constant R_0 such that

$$\frac{F(x,u_k)}{|x|^\beta} \leq \frac{R_0 f(x,u_k)}{|x|^\beta} \quad \text{for all} \ x \in \mathbb{R}^{2m}. \tag{5-14}$$

Together with the generalized Lebesgue dominated convergence theorem, we can get that

$$\lim_{R\to+\infty} \lim_{k\to+\infty} \int_{B_R} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx = 0. \tag{5-15}$$

Thus, it suffices to check that

$$\lim_{R\to+\infty} \lim_{k\to+\infty} \int_{\mathbb{R}^{2m}\setminus B_R} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx = 0. \tag{5-16}$$

By dividing the integral into two parts, we arrive at

$$\begin{aligned} \int_{\mathbb{R}^{2m}\setminus B_R} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx &= \int_{\{|x|\geq R\}\bigcap\{|u_k|>A\}} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx \\ &\quad + \int_{\{|x|\geq R\}\bigcap\{|u_k|\leq A\}} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx \\ &=: I_A + II_A. \end{aligned}$$

For I_A, it follows from (5-9) that

$$\int_{\{|x|\geq R\}\bigcap\{|u_k|>A\}} \frac{|F(x,u_k)|}{|x|^\beta}\,dx \leq \frac{R_0}{A}\int_{\{|x|\geq R\}\bigcap\{|u_k|>A\}} \frac{|f(x,u_k)u_k|}{|x|^\beta}\,dx \lesssim \frac{R_0}{A}.$$

Thus, $\lim_{A\to+\infty}\lim_{R\to+\infty}\lim_{k\to+\infty} I_A = 0$.

For II_A, applying hypothesis (H$_1$) and Theorem 1.7, one can derive that

$$\begin{aligned}\lim_{A\to+\infty}\lim_{R\to+\infty}\lim_{k\to+\infty} II_A &\leq \lim_{A\to+\infty}\lim_{R\to+\infty}\lim_{k\to+\infty} C(\alpha_0,A)\int_{\{|x|\geq R\}\bigcap\{|u_k|\leq A\}} \frac{|u_k|^2}{|x|^\beta}\,dx \\ &\leq \lim_{A\to+\infty}\lim_{R\to+\infty}\lim_{k\to+\infty} \frac{C(\alpha_0,A)}{R^{\beta/2}}\sup_k \|u_k\|_E^2 \\ &= 0.\end{aligned} \tag{5-17}$$

Hence,

$$\lim_{k\to+\infty}\int_{\mathbb{R}^{2m}} \frac{|F(x,u_k)-F(x,u)|}{|x|^\beta}\,dx = 0. \tag{5-18}$$

A simple application of (5-13) shows that

$$\int_{\mathbb{R}^{2m}} (\nabla^m u\nabla^m\varphi + u\varphi)\,dx - \int_{\mathbb{R}^{2m}} \frac{f(x,u)}{|x|^\beta}\varphi\,dx = 0, \quad \text{for all } \varphi\in C_0^\infty(\mathbb{R}^{2m}).$$

Thus, u is a weak solution of polyharmonic equation (1-14). □

Part 2. This part is devoted to showing that the Palais–Smale sequence $\{u_k\}_k$ satisfies the Palais–Smale condition in light of the concentration-compactness principle. We begin with a crucial fact:

$$0 < c_\beta < \left(1-\frac{\beta}{2m}\right)\frac{\beta_{2m,m}}{2\alpha_0}.$$

Recall that we have shown the attainability of S_p in Section 4, so there exists a function u such that

$$\int_{\mathbb{R}^{2m}} \frac{|u|^p}{|x|^\beta}\,dx = 1 \quad\text{and}\quad \|u\|_E = S_p.$$

Through the definition of c_β, we get

$$0 < c_\beta \leq \max_{t\geq 0} I_\beta(tu) = \max_{t\geq 0}\left(\frac{t^2}{2}S_p^2 - \int_{\mathbb{R}^{2m}} \frac{F(x,tu)}{|x|^\beta}\,dx\right).$$

According to the definition of C_p, we can obtain that

$$c_\beta \le \max_{t\ge 0}\left(\frac{t^2}{2}S_p^2 - t^p\frac{C_p}{p}\right) = \frac{(p-2)}{2p}\frac{S_p^{2p/(p-2)}}{C_p^{2/(p-2)}} < \frac{\beta_{2m,m}\left(1-\frac{\beta}{2m}\right)}{2\alpha_0}. \tag{5-19}$$

Now, we are in a position to verify that $\{u_k\}_k$ satisfies the Palais–Smale compactness condition. We discuss this by the following two cases.

Case 1. ($c_\beta \ne 0, u = 0$). We first claim that there exists some $q > 1$ such that

$$\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(\alpha_0|u_k|^2)^q}{|x|^\beta}\,dx \le C \quad \text{for all } k \in \mathbb{N}.$$

Since $u = 0$, one can employ Lemma 5.4 to drive that

$$\int_{\mathbb{R}^{2m}} \frac{F(x,u_k)}{|x|^\beta}\,dx \to \int_{\mathbb{R}^{2m}} \frac{F(x,u)}{|x|^\beta}\,dx = 0. \tag{5-20}$$

Together with (5-10), we obtain that

$$\|u_k\|_E^2 \to 2c_\beta \quad \text{as } k \to \infty. \tag{5-21}$$

Take $q > 1$ sufficiently close to 1 such that

$$\alpha_0 q\|u_k\|^2 \le \alpha_0 q\|u_k\|_E^2 \le \beta_0 < \left(1-\frac{\beta}{2m}\right)\beta_{2m,m}. \tag{5-22}$$

Then, it follows that

$$\begin{aligned}\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(\alpha_0|u_k|^2)^q}{|x|^\beta}\,dx &\le \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(\alpha_0|u_k|^2)^q}{|x|^\beta}\,dx \\ &\lesssim \int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\left(\beta_0\left(\frac{u_k}{\|u_k\|}\right)^2\right)}{|x|^\beta}\,dx \\ &\lesssim 1.\end{aligned} \tag{5-23}$$

Combining hypothesis (H$_1$), the Hölder inequality and (5-23), one can derive that

$$\begin{aligned}&\left|\int_{\mathbb{R}^{2m}} \frac{f(x,u_k)u_k}{|x|^\beta}\,dx\right| \\ &\lesssim \int_{\mathbb{R}^{2m}} \frac{|u_k|^2}{|x|^\beta}\,dx + \left(\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}(\alpha_0|u_k|^2)^q}{|x|^\beta}\,dx\right)^{\frac{1}{q}}\left(\int_{\mathbb{R}^{2m}} \frac{|u_k|^{q'}}{|x|^\beta}\,dx\right)^{\frac{1}{q'}} \\ &\lesssim \left(\int_{\mathbb{R}^{2m}} \frac{|u_k-u|^2}{|x|^\beta}\,dx\right)^{\frac{1}{2}} + \left(\int_{\mathbb{R}^{2m}} \frac{|u_k-u|^{q'}}{|x|^\beta}\,dx\right)^{\frac{1}{q'}},\end{aligned} \tag{5-24}$$

where $q > 1$ close enough to 1 and $\frac{1}{p} + \frac{1}{p'} = 1$.

Thanks to Theorem 1.7 again, we arrive at

$$\int_{\mathbb{R}^{2m}} \frac{f(x,u_k)u_k}{|x|^\beta}\,dx \to 0 \quad \text{as } k\to\infty.$$

Taking $I'_\beta(u_k)u_k \to 0$ into consideration, we get $\lim_{k\to\infty}\|u_k\|_E \to 0$, which is a contradiction with $c_\beta > 0$.

Case 2. $(c_\beta \neq 0, u \neq 0)$. We claim that $\lim_{k\to\infty}\|u_k\|_E = \|u\|_E$. We argue this by contradiction. Suppose $\lim_{k\to\infty}\|u_k\|_E > \|u\|_E$, and define

$$v_k := \frac{u_k}{\|u_k\|_E} \quad \text{and} \quad v_0 := \frac{u}{\lim_{k\to\infty}\|u_k\|_E}.$$

We claim that for $q>1$ sufficiently close to 1, there exists a constant $\beta_0>0$ such that the following inequality holds.

$$q\alpha_0\|u_k\|_E^2 \le \beta_0 < \frac{\beta_{2m,m}\big(1-\frac{\beta}{2m}\big)}{1-\|v_0\|_E^2}. \tag{5-25}$$

Indeed,

$$\begin{aligned}\lim_{k\to\infty}\|u_k\|_E^2(1-\|v_0\|_E^2) &= \lim_{k\to\infty}\|u_k\|_E^2\bigg(1-\frac{\|u\|_E^2}{\lim_{k\to\infty}\|u_k\|_E^2}\bigg)\\ &= 2c_\beta + 2\int_{\mathbb{R}^{2m}}\frac{F(x,u)}{|x|^\beta}\,dx - 2I_\beta(u) - 2\int_{\mathbb{R}^{2m}}\frac{F(x,u)}{|x|^\beta}\,dx\\ &< \frac{\beta_{2m,m}\big(1-\frac{\beta}{2m}\big)}{\alpha_0},\end{aligned} \tag{5-26}$$

where we apply $I_\beta(u)\ge 0$. Then it follows from the above estimate and Theorem 1.4 that

$$\int_{\mathbb{R}^{2m}} \frac{(\Phi_{2m,m}(\alpha_0|u_k|^2))^q}{|x|^\beta}\,dx \le C\int_{\mathbb{R}^{2m}} \frac{\Phi_{2m,m}\big(\beta_0\big|\frac{u_k}{\|u_k\|_E}\big|^2\big)}{|x|^\beta}\,dx \lesssim 1. \tag{5-27}$$

Under hypothesis (H_1), the Hölder inequality gives that

$$\begin{aligned}\bigg|\int_{\mathbb{R}^{2m}}\frac{f(x,u_k)(u_k-u)}{|x|^\beta}\,dx\bigg| &\le b_1\bigg(\int_{\mathbb{R}^{2m}}\frac{|u_k|^2}{|x|^\beta}\,dx\bigg)^{\frac12}\bigg(\int_{\mathbb{R}^{2m}}\frac{|u_k-u|^2}{|x|^\beta}\,dx\bigg)^{\frac12}\\ &\quad + b_2\bigg(\int_{\mathbb{R}^{2m}}\frac{|u_k-u|^{q'}}{|x|^\beta}\,dx\bigg)^{\frac{1}{q'}}\bigg(\int_{\mathbb{R}^{2m}}\frac{(\Phi_{2m,m}(\alpha_0|u_k|^2))^q}{|x|^\beta}\,dx\bigg)^{\frac1q}.\end{aligned} \tag{5-28}$$

Thanks to Theorem 1.7, we derive the following conclusion with inequalities (5-27) and (5-28):

$$\int_{\mathbb{R}^{2m}} \frac{f(x,u_k)(u_k-u)}{|x|^\beta}\,dx \to 0.$$

Together with $I'_\beta(u_k)(u_k-u)\to 0$, we get

$$\int_{\mathbb{R}^{2m}} \nabla^m u_k(\nabla^m u_k - \nabla^m u)\,dx + \int_{\mathbb{R}^{2m}} V(x)u_k(u_k-u)\,dx \to 0.$$

Since $u_k \rightharpoonup u$ in E, we have

$$\int_{\mathbb{R}^{2m}} \nabla^m u(\nabla^m u_k - \nabla^m u)\,dx \to 0 \quad \text{and} \quad \int_{\mathbb{R}^{2m}} V(x)u(u_k-u)\,dx \to 0.$$

Therefore

$$\begin{aligned}(5\text{-}29)\quad \lim_{k\to+\infty}\|u_k-u\|_E^2 &= \lim_{k\to+\infty}\int_{\mathbb{R}^{2m}}(\nabla^m u_k-\nabla^m u)(\nabla^m u_k-\nabla^m u)\,dx\\ &\qquad + \lim_{k\to+\infty}\int_{\mathbb{R}^{2m}} V(x)(u_k-u)(u_k-u)\,dx\\ &=0,\end{aligned}$$

which arrives at a contradiction with $\lim_{k\to\infty}\|u_k\|_E > \|u\|_E$.

Part 3. In this part, we show that the critical point of functional I_β is actually a ground state solution for the singular polyharmonic equation (1-14). Define

$$m = \inf_{u\in s} I_\beta(u) \quad \text{and} \quad S := \{u\in E : u \not\equiv 0 \text{ and } I'_\beta(u)=0\}.$$

For all $w\in S$, pick t_0 sufficiently large such that $I_\beta(t_0 w) < 0$. Denote $h : (0,+\infty)\to\mathbb{R}$ by $h(t) = I_\beta(tw)$ and $g:[0,1]\to E$ by $g(t) = tt_0 w$. It is easy to check that

$$h'(t) = I'_\beta(tw)w = t\|w\|_E^2 - \int_{\mathbb{R}^{2m}} \frac{f(x,tw)w}{|x|^\beta}\,dx, \quad \text{for all } t>0.$$

Combine this with $I'_\beta(w)w = 0$, we easily see that

$$h'(t) = t\int_{\mathbb{R}^{2m}} \left(\frac{f(x,w)}{w} - \frac{f(x,tw)}{tw}\right)\frac{w^2}{|x|^\beta}\,dx,$$

which implies that $h'(t) > 0$ for $t\in(0,1)$ and $h'(t) < 0$ for $t > 1$ under hypothesis (H_6). Thus,

$$c_\beta \le \max_{t\in[0,1]} I_\beta(g(t)) \le \max_{t\ge 0} I_\beta(tw) = I_\beta(w),$$

which concludes the proof of Theorem 1.9.

6. The proof of Theorem 1.10

In this section, we will investigate the existence of the nontrivial weak solutions for nonsingular polyharmonic equation (1-15). The presence of the constant potential $V(x)$ makes it hard to follow the same line of reasoning as for Theorem 1.9. In order to overcome this difficulty, we need to use the principle of symmetric criticality. We first introduce some background knowledge about the principle of symmetric criticality.

Definition 6.1. The action of a topological group G on a normed space X is a continuous map

$$G \times X \to X : \quad [g, u] \mapsto gu$$

such that

$$1 \cdot u = u, \qquad (gh)u = g(hu), \qquad u \longmapsto gu \quad \text{is linear.}$$

The action is isometric if

$$\|gu\| = \|u\|$$

The space of invariant points is defined by

$$Fix(G) := \{u \in X : gu = u, \forall g \in G\}.$$

A function $\varphi : X \to \mathbb{R}$ is invariant if $\varphi \circ g = \varphi$ for every $g \in G$.

Lemma 6.2 (principle of symmetric criticality [Badiale and Serra 2011]). *Assume that the action of the topological group G on the Hilbert space X is isometric. If $\varphi \in C^1(X, \mathbb{R})$ is invariant and if u is a critical point of φ restricted to* $\mathrm{Fix}(G)$, *then u is also a critical point of φ.*

Lemma 6.3. *For $q \geq 2$, $W_r^{m,2}(\mathbb{R}^{2m})$ can be compactly embedded into $L^q(\mathbb{R}^{2m})$ for any $q > 2$.*

Remark 6.4. Through applying the radial lemma, one can easily get Lemma 6.3 with a slight modification of the proof of Theorem 1.7.

Now, we are in a position to prove Theorem 1.10. The functional related with (1-15) is given by $I(u) = \frac{1}{2}\|u\|_E^2 - \int_{\mathbb{R}^{2m}} F(x, u)\, dx$. Based on Lemma 6.2, we can restrict the functional I to the subspace E_r of E, where E_r is the set of all radial functions in E. It follows from same reasoning as for Lemma 5.3 that functional I satisfies the geometric conditions which imply that there exists a sequence $\{u_k\}_k \in E_r$ such that $I(u_k) \to c_0$, $I'(u_k) \to 0$ as $k \to +\infty$. Furthermore, we also can obtain

$$\begin{aligned} u_k &\rightharpoonup u_0, && \text{in } E_r, \\ u_k &\to u_0, && \text{in } L^q(\mathbb{R}^{2m}) \text{ for all } q > 2, \\ u_k(x) &\to u_0(x), && \text{almost everywhere in } \mathbb{R}^{2m}. \end{aligned}$$

We will use a new method based on Lemma 6.3 to prove that

$$\int_{\mathbb{R}^{2m}} F(x, u_k)\, dx \to \int_{\mathbb{R}^{2m}} F(x, u)\, dx.$$

By splitting the integral into three parts, we have

$$\begin{aligned} &\lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{\mathbb{R}^{2m}} |F(x, u_k) - F(x, u)|\, dx \\ &= \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{B_R} |F(x, u_k) - F(x, u)|\, dx \\ &\qquad + \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|>A} |F(x, u_k) - F(x, u)|\, dx \\ &\qquad\qquad + \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|\le A} |F(x, u_k) - F(x, u)|\, dx \\ &=: I_1 + I_2 + I_3. \end{aligned} \tag{6-1}$$

For I_1, it directly follows from (5-13), (5-14) for the case $\beta = 0$. For I_2, in view of hypotheses (H_2) and (H_3), we have

$$\begin{aligned} I_2 &= \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|>A} |F(x, u_k) - F(x, u)|\, dx \\ &\lesssim \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|>A} |F(x, u_k)|\, dx \\ &\lesssim \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \frac{1}{A} \int_{|x|>R,\, |u_k|>A} |f(x, u_k) u_k|\, dx \\ &= 0. \end{aligned} \tag{6-2}$$

For I_3, combining the hypothesis $f(x,t) = o(t)$ and Lemma 6.3, one can obtain that for any $\varepsilon > 0$,

$$\begin{aligned} I_3 &= \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|\le A} |F(x, u_k) - F(x, u)|\, dx \\ &\lesssim \lim_{R\to\infty} \lim_{k\to\infty} \lim_{A\to\infty} \int_{|x|>R,\, |u_k|\le A} |F(x, u_k)|\, dx \\ &\lesssim \varepsilon \|u_k\|_E^2 + \lim_{R\to\infty} \lim_{k\to\infty} \int_{|x|>R} |u_k|^3\, dx \\ &\lesssim \varepsilon \|u_k\|_E^2, \end{aligned} \tag{6-3}$$

which leads to $I_3=0$. Carrying out similar steps as we did in Section 4 (Part 1), one can easily see that u is a weak solution of (1-15).

Next, we show u_k satisfies the Palais–Smale compactness condition and u is a critical point of functional I restricted in E_r. The process of proof follows from the similar argument of Section 4 (Part 2) as long as we can verify that

$$\left|\int_{\mathbb{R}^{2m}} f(x,u_k)(u_k-u)\,dx\right| \to 0.$$

Since $f(x,t)=o(t)$ at the origin, through hypothesis (H_1) and the Hölder inequality, we derive that for any $\varepsilon>0$, it holds that

$$\begin{aligned}\left|\int_{\mathbb{R}^{2m}} f(x,u_k)(u_k-u)\,dx\right| &\le \varepsilon\Big(\int_{\mathbb{R}^{2m}} |u_k|^2\,dx\Big)^{\frac{1}{2}}\Big(\int_{\mathbb{R}^{2m}} |u_k-u|^2\,dx\Big)^{\frac{1}{2}} \\ &\quad + C_\varepsilon\Big(\int_{\mathbb{R}^{2m}} |u_k-u|^{q'}\,dx\Big)^{\frac{1}{q'}}\Big(\int_{\mathbb{R}^{2m}} (\Phi_{2m,m}(\alpha_0|u_k|^2))^q\,dx\Big)^{\frac{1}{q}}.\end{aligned} \tag{6-4}$$

Letting $k\to\infty$ and $\varepsilon\to 0$, we arrives at the desired conclusion. Finally applying the principle of symmetric criticality again, we see that u is also a critical point of I in E.

Acknowledgements

Zhang and Chen were partly supported by a grant from the NNSF of China (No. 11901031). Chen was also supported by the Beijing Institute of Technology Research Fund Program for Young Scholars (No.3170012221903). Li was partly supported by a Simons Collaboration Grant from the Simons Foundation and a US NSF grant.

References

[Adachi and Tanaka 2000] S. Adachi and K. Tanaka, "Trudinger type inequalities in $\mathbb{R}^N$ and their best exponents", *Proc. Amer. Math. Soc.* **128**:7 (2000), 2051–2057. MR Zbl

[Adams 1988] D. R. Adams, "A sharp inequality of J. Moser for higher order derivatives", *Ann. of Math.* (2) **128**:2 (1988), 385–398. MR Zbl

[Adimurthi and Sandeep 2007] Adimurthi and K. Sandeep, "A singular Moser–Trudinger embedding and its applications", *NoDEA Nonlinear Differential Equations Appl.* **13**:5-6 (2007), 585–603. MR Zbl

[Adimurthi and Yang 2010] Adimurthi and Y. Yang, "An interpolation of Hardy inequality and Trundinger–Moser inequality in $\mathbb{R}^N$ and its applications", *Int. Math. Res. Not.* **2010**:13 (2010), 2394–2426. MR Zbl

[Alves and Figueiredo 2009] C. O. Alves and G. M. Figueiredo, "On multiplicity and concentration of positive solutions for a class of quasilinear problems with critical exponential growth in $\mathbb{R}^N$", *J. Differential Equations* **246**:3 (2009), 1288–1311. MR Zbl

[Atkinson and Peletier 1986] F. V. Atkinson and L. A. Peletier, "Ground states and Dirichlet problems for $-\Delta u = f(u)$ in $\mathbb{R}^2$", *Arch. Rational Mech. Anal.* **96**:2 (1986), 147–165. MR Zbl

[Badiale and Serra 2011] M. Badiale and E. Serra, *Semilinear elliptic equations for beginners: existence results via the variational approach*, Springer, 2011. MR Zbl

[Bao et al. 2016] J. Bao, N. Lam, and G. Lu, "Polyharmonic equations with critical exponential growth in the whole space $\mathbb{R}^n$", *Discrete Contin. Dyn. Syst.* **36**:2 (2016), 577–600. MR Zbl

[Brézis and Lieb 1983] H. Brézis and E. Lieb, "A relation between pointwise convergence of functions and convergence of functionals", *Proc. Amer. Math. Soc.* **88**:3 (1983), 486–490. MR Zbl

[Cao 1992] D. M. Cao, "Nontrivial solution of semilinear elliptic equation with critical exponent in $\mathbb{R}^2$", *Comm. Partial Differential Equations* **17**:3-4 (1992), 407–435. MR Zbl

[Carleson and Chang 1986] L. Carleson and S.-Y. A. Chang, "On the existence of an extremal function for an inequality of J. Moser", *Bull. Sci. Math.* (2) **110**:2 (1986), 113–127. MR Zbl

[Chen et al. 2018] L. Chen, J. Li, G. Lu, and C. Zhang, "Sharpened Adams inequality and ground state solutions to the bi-Laplacian equation in $\mathbb{R}^4$", *Adv. Nonlinear Stud.* **18**:3 (2018), 429–452. MR Zbl

[Chen et al. 2019] L. Chen, G. Lu, and M. Zhu, "Ground states of bi-harmonic equations with critical exponential growth involving constant and trapping potentials", preprint, 2019. arXiv

[de Figueiredo et al. 1995] D. G. de Figueiredo, O. H. Miyagaki, and B. Ruf, "Elliptic equations in $\mathbb{R}^2$ with nonlinearities in the critical growth range", *Calc. Var. Partial Differential Equations* **3**:2 (1995), 139–153. MR Zbl

[de Figueiredo et al. 2002] D. G. de Figueiredo, J. M. do Ó, and B. Ruf, "On an inequality by N. Trudinger and J. Moser and related elliptic equations", *Comm. Pure Appl. Math.* **55**:2 (2002), 135–152. MR Zbl

[Fontana and Morpurgo 2015] L. Fontana and C. Morpurgo, "Sharp Adams and Moser–Trudinger inequalities on $\mathbb{R}^n$ and other spaces of infinite measure", preprint, 2015. arXiv

[Kozono et al. 2006] H. Kozono, T. Sato, and H. Wadade, "Upper bound of the best constant of a Trudinger–Moser inequality and its application to a Gagliardo–Nirenberg inequality", *Indiana Univ. Math. J.* **55**:6 (2006), 1951–1974. MR Zbl

[Lam and Lu 2012a] N. Lam and G. Lu, "Existence and multiplicity of solutions to equations of N-Laplacian type with critical exponential growth in $\mathbb{R}^N$", *J. Funct. Anal.* **262**:3 (2012), 1132–1165. MR Zbl

[Lam and Lu 2012b] N. Lam and G. Lu, "Sharp Adams type inequalities in Sobolev spaces $W^{m,\frac{n}{m}}(\mathbb{R}^n)$ for arbitrary integer m", *J. Differential Equations* **253**:4 (2012), 1143–1171. MR Zbl

[Lam and Lu 2012c] N. Lam and G. Lu, "Sharp Moser–Trudinger inequality on the Heisenberg group at the critical case and applications", *Adv. Math.* **231**:6 (2012), 3259–3287. MR Zbl

[Lam and Lu 2012d] N. Lam and G. Lu, "Sharp singular Adams inequalities in high order Sobolev spaces", *Methods Appl. Anal.* **19**:3 (2012), 243–266. MR Zbl

[Lam and Lu 2013a] N. Lam and G. Lu, "N-Laplacian equations in $\mathbb{R}^N$ with subcritical and critical growth without the Ambrosetti–Rabinowitz condition", *Adv. Nonlinear Stud.* **13**:2 (2013), 289–308. MR Zbl

[Lam and Lu 2013b] N. Lam and G. Lu, "A new approach to sharp Moser–Trudinger and Adams type inequalities: a rearrangement-free argument", *J. Differential Equations* **255**:3 (2013), 298–325. MR Zbl

[Lam and Lu 2014] N. Lam and G. Lu, "Elliptic equations and systems with subcritical and critical exponential growth without the Ambrosetti–Rabinowitz condition", *J. Geom. Anal.* **24**:1 (2014), 118–143. MR Zbl

[Lam et al. 2014] N. Lam, G. Lu, and H. Tang, "Sharp subcritical Moser–Trudinger inequalities on Heisenberg groups and subelliptic PDEs", *Nonlinear Anal.* **95** (2014), 77–92. MR Zbl

[Lam et al. 2017a] N. Lam, G. Lu, and H. Tang, "Sharp affine and improved Moser–Trudinger–Adams type inequalities on unbounded domains in the spirit of Lions", *J. Geom. Anal.* **27**:1 (2017), 300–334. MR Zbl

[Lam et al. 2017b] N. Lam, G. Lu, and L. Zhang, "Equivalence of critical and subcritical sharp Trudinger–Moser–Adams inequalities", *Rev. Mat. Iberoam.* **33**:4 (2017), 1219–1246. MR Zbl

[Lam et al. 2019] N. Lam, G. Lu, and L. Zhang, "Existence and nonexistence of extremal functions for sharp Trudinger–Moser inequalities", *Adv. Math.* **352** (2019), 1253–1298. MR Zbl

[Li 2001] Y. Li, "Moser–Trudinger inequality on compact Riemannian manifolds of dimension two", *J. Partial Differential Equations* **14**:2 (2001), 163–192. MR Zbl

[Li 2006] Y. X. Li, "Remarks on the extremal functions for the Moser–Trudinger inequality", *Acta Math. Sin.* (*Engl. Ser.*) **22**:2 (2006), 545–550. MR Zbl

[Li and Ndiaye 2007] Y. Li and C. B. Ndiaye, "Extremal functions for Moser–Trudinger type inequality on compact closed 4-manifolds", *J. Geom. Anal.* **17**:4 (2007), 669–699. MR Zbl

[Li and Ruf 2008] Y. Li and B. Ruf, "A sharp Trudinger–Moser type inequality for unbounded domains in $\mathbb{R}^n$", *Indiana Univ. Math. J.* **57**:1 (2008), 451–480. MR Zbl

[Li et al. 2018a] J. Li, G. Lu, and Q. Yang, "Fourier analysis and optimal Hardy–Adams inequalities on hyperbolic spaces of any even dimension", *Adv. Math.* **333** (2018), 350–385. MR Zbl

[Li et al. 2018b] J. Li, G. Lu, and M. Zhu, "Concentration-compactness principle for Trudinger–Moser inequalities on Heisenberg groups and existence of ground state solutions", *Calc. Var. Partial Differential Equations* **57**:3 (2018), art. id. 84, 26 pp. MR Zbl

[Lin 1986] C. S. Lin, "Interpolation inequalities with weights", *Comm. Partial Differential Equations* **11**:14 (1986), 1515–1538. MR Zbl

[Lions 1985] P.-L. Lions, "The concentration-compactness principle in the calculus of variations. The limit case, II", *Rev. Mat. Iberoamericana* **1**:2 (1985), 45–121. MR Zbl

[Lu and Yang 2009] G. Lu and Y. Yang, "Adams' inequalities for bi-Laplacian and extremal functions in dimension four", *Adv. Math.* **220**:4 (2009), 1135–1170. MR Zbl

[Lu and Yang 2017] G. Lu and Q. Yang, "Sharp Hardy–Adams inequalities for bi-Laplacian on hyperbolic space of dimension four", *Adv. Math.* **319** (2017), 567–598. MR Zbl

[Lu and Yang 2019] G. Lu and Q. Yang, "Paneitz operators on hyperbolic spaces and high order Hardy–Sobolev–Maz'ya inequalities on half spaces", *Amer. J. Math.* **141**:6 (2019), 1777–1816. MR Zbl

[Masmoudi and Sani 2015] N. Masmoudi and F. Sani, "Trudinger–Moser inequalities with the exact growth condition in $\mathbb{R}^N$ and applications", *Comm. Partial Differential Equations* **40**:8 (2015), 1408–1440. MR Zbl

[Moser 1971] J. Moser, "A sharp form of an inequality by N. Trudinger", *Indiana Univ. Math. J.* **20** (1971), 1077–1092. MR Zbl

[Moser 1973] J. Moser, "On a nonlinear problem in differential geometry", pp. 273–280 in *Dynamical systems* (Salvador, 1971), edited by M. M. Peixoto, 1973. MR Zbl

[Nguyen 2016] V. H. Nguyen, "An improvement for the sharp Adams inequalities in bounded domains and whole space $\mathbb{R}^n$", preprint, 2016. arXiv

[do Ó 1996] J. M. B. do Ó, "Semilinear Dirichlet problems for the N-Laplacian in $\mathbb{R}^N$ with nonlinearities in the critical growth range", *Differential Integral Equations* **9**:5 (1996), 967–979. MR Zbl

[do Ó et al. 2009] J. M. do Ó, E. Medeiros, and U. Severo, "On a quasilinear nonhomogeneous elliptic equation with critical growth in $\mathbb{R}^N$", *J. Differential Equations* **246**:4 (2009), 1363–1386. MR Zbl

[do Ó et al. 2014] J. M. do Ó, M. de Souza, E. de Medeiros, and U. Severo, "An improvement for the Trudinger–Moser inequality and applications", *J. Differential Equations* **256**:4 (2014), 1317–1349. MR Zbl

[Ogawa and Ozawa 1991] T. Ogawa and T. Ozawa, "Trudinger type inequalities and uniqueness of weak solutions for the nonlinear Schrödinger mixed problem", *J. Math. Anal. Appl.* **155**:2 (1991), 531–540. MR Zbl

[Ozawa 1995] T. Ozawa, "On critical cases of Sobolev's inequalities", *J. Funct. Anal.* **127**:2 (1995), 259–269. MR Zbl

[Panda 1996] R. Panda, "On semilinear Neumann problems with critical growth for the n-Laplacian", *Nonlinear Anal.* **26**:8 (1996), 1347–1366. MR Zbl

[Pohozaev 1965] S. I. Pohozaev, "On S. L. Sobolev's embedding theorem in the case $pl = n$", pp. 158–170 in *Proceedings of the Technical Scientific Conference on Advances of Scientific Research*, Moscov. Energet. Inst., Moscow, 1965. In Russian.

[Ruf 2005] B. Ruf, "A sharp Trudinger–Moser type inequality for unbounded domains in $\mathbb{R}^2$", *J. Funct. Anal.* **219**:2 (2005), 340–367. MR Zbl

[Ruf and Sani 2013] B. Ruf and F. Sani, "Sharp Adams-type inequalities in $\mathbb{R}^n$", *Trans. Amer. Math. Soc.* **365**:2 (2013), 645–670. MR Zbl

[Silva and Soares 1999] E. A. d. B. Silva and S. H. M. Soares, "Liouville–Gelfand type problems for the N-Laplacian on bounded domains of $\mathbb{R}^N$", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **28**:1 (1999), 1–30. MR Zbl

[Tarsi 2012] C. Tarsi, "Adams' inequality and limiting Sobolev embeddings into Zygmund spaces", *Potential Anal.* **37**:4 (2012), 353–385. MR Zbl

[Trudinger 1967] N. S. Trudinger, "On imbeddings into Orlicz spaces and some applications", *J. Math. Mech.* **17** (1967), 473–483. MR Zbl

[Yang 2012a] Y. Yang, "Adams type inequalities and related elliptic partial differential equations in dimension four", *J. Differential Equations* **252**:3 (2012), 2266–2295. MR Zbl

[Yang 2012b] Y. Yang, "Existence of positive solutions to quasi-linear elliptic equations with exponential growth in the whole Euclidean space", *J. Funct. Anal.* **262**:4 (2012), 1679–1704. MR Zbl

[Yudovich 1961] V. I. Yudovich, "Some estimates connected with integral operators and with solutions of elliptic equations", *Dokl. Akad. Nauk SSSR* **138** (1961), 805–808. In Russian; translated in *Soviet Math. Dokl.* **2** (1961), 746–749. MR Zbl

[Zhang and Chen 2018] C. Zhang and L. Chen, "Concentration-compactness principle of singular Trudinger–Moser inequalities in $\mathbb{R}^n$ and n-Laplace equations", *Adv. Nonlinear Stud.* **18**:3 (2018), 567–585. MR Zbl

Received May 31, 2018. Revised September 13, 2018.

CAIFENG ZHANG
DEPARTMENT OF APPLIED MATHEMATICS, SCHOOL OF MATHEMATICS AND PHYSICS
UNIVERSITY OF SCIENCE AND TECHNOLOGY BEIJING
CHINA
zhangcaifeng1991@mail.bnu.edu.cn

JUNGANG LI
DEPARTMENT OF MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RI
UNITED STATES
jungang.2.li@uconn.edu

LU CHEN
SCHOOL OF MATHEMATICS AND STATISTICS
BEIJING INSTITUTE OF TECHNOLOGY
CHINA
luchen2015@mail.bnu.edu.cn

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 305 No. 1 March 2020
