Pacific Journal of Mathematics

Volume 301 No. 2 August 2019

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 2019 is US $490/year for the electronic version, and $665/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. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.385

NEW APPLICATIONS OF EXTREMELY REGULAR FUNCTION SPACES

TROND A. ABRAHAMSEN, OLAV NYGAARD AND MÄRT PÕLDVERE

Let L be an infinite locally compact Hausdorff topological space. We show that extremely regular subspaces of $C_0(L)$ have very strong diameter 2 properties and, for every real number ε with $0 < \varepsilon < 1$, contain an ε-isometric copy of c_0. If L does not contain isolated points they even have the Daugavet property, and thus contain an asymptotically isometric copy of ℓ_1.

1. Introduction

Throughout, let L be an infinite locally compact Hausdorff topological space, and denote as usual by $C_0(L)$ the Banach space of continuous $\mathbb{K}$-valued functions on L that "vanish at infinity", where $\mathbb{K}$ is the field of either real or complex numbers.

Definition 1.1 [Cengiz 1973a]. An *extremely regular function space* is a linear subspace $\mathcal{A}$ of $C_0(L)$ such that for every $x_0 \in L$, every real number ε with $0 < \varepsilon < 1$, and every open neighbourhood V of x_0, there exists an $f \in \mathcal{A}$ such that

$$\|f\| = 1 = f(x_0) > \varepsilon > \sup_{x \in L \setminus V} |f(x)|.$$

The interest in extremely regular function spaces came from their importance in Banach–Stone type theorems. An example, also due to Cengiz [1973a], is as follows: *If L_1 and L_2 are locally compact Hausdorff topological spaces such that there exists a linear isomorphism φ from an extremely regular subspace of $C_0(L_1)$ onto such a subspace of $C_0(L_2)$ with $\|\varphi\|\|\varphi^{-1}\| < 2$, then L_1 and L_2 are homeomorphic* (here $C_0(L_1)$ and $C_0(L_2)$ are complex spaces). Properties of extremely regular function spaces were studied in [Cengiz 1973b].

In this paper, we demonstrate that extremely regular function spaces play a role in a quite recent subfield of the theory of Banach spaces, namely that involving *Daugavet spaces, diameter 2 spaces*, and *octahedral spaces*. Let us briefly explain

Research by the third-named author supported, in part, by institutional research funding IUT20-57 of the Estonian Ministry of Education and Research.
MSC2010: 46B20, 46B22, 46B04 .
Keywords: extremely regular function space, strong diameter 2 property, almost square space, octahedrality, Daugavet property.

some main lines of this theory before returning to extremely regular function spaces and our results.

Let X be a Banach space and B_X its unit ball. By a *slice* of B_X we mean a set of the form $S(x^*, \varepsilon) := \{x \in B_X : \operatorname{Re} x^*(x) > 1-\varepsilon\}$, where x^* is in the unit sphere S_{X^*} of X^* and $\varepsilon > 0$. A *finite convex combination of slices* of B_X is a set S of the form $S = \sum_{i=1}^n \lambda_i S(x_i^*, \varepsilon_i)$ where $n \in \mathbb{N}$, $\lambda_i > 0$, $\sum_{i=1}^n \lambda_i = 1$, $x_i^* \in S_{X^*}$, and $\varepsilon_i > 0$.

Definition 1.2. A Banach space X has the *strong diameter* 2 *property* (SD2P) if every finite convex combination of slices of B_X has diameter 2.

A lemma by Bourgain [Ghoussoub et al. 1987, page 26, Lemma II.1] (and independently rediscovered by Shvydkoy [2000]) says that *every nonempty relatively weakly open subset of B_X contains a finite convex combination of slices*. Thus the SD2P implies that every nonempty relatively weakly open subset of B_X has diameter 2, which in turn implies that every slice of B_X has diameter 2. None of these implications is reversible ([Becerra Guerrero et al. 2015; Haller and Langemets 2014]).

It is an important observation of Deville and Godefroy from the late 1980s, stated without proof in [Godefroy 1989], that X having SD2P is equivalent to X^* being *octahedral*. A Banach space Z is *octahedral* if, for every finite-dimensional subspace F of Z and every $\varepsilon > 0$, there exists a $y \in S_Z$ such that

$$\|x + ty\| \geq (1-\varepsilon)(\|x\| + |t|) \quad \text{for every } x \in F \text{ and every } t \in \mathbb{K}.$$

A complete proof of this equivalence can be found in [Becerra Guerrero et al. 2014, Corollary 2.2] (the proof is carried out for the real case, but it is not too hard to see that the result holds also in the complex case). More on the history of the equivalence can be found in [Abrahamsen et al. 2017, Remark 1.4].

Definition 1.3. A Banach space X

(1) is *almost square* (ASQ) if whenever $n \in \mathbb{N}$ and $x_1, \dots, x_n \in S_X$, there exists a sequence (y_k) in B_X such that $\|x_i \pm y_k\| \xrightarrow[k\to\infty]{} 1$ for every $i \in \{1, \dots, n\}$ and $\|y_k\| \xrightarrow[k\to\infty]{} 1$.

(2) has the *symmetric strong diameter* 2 *property* (SSD2P) if whenever $n \in \mathbb{N}$, $S_1, \dots, S_n$ are slices of B_X, and $\varepsilon > 0$, there exist $x_i \in S_i$, $i = 1, \dots, n$, and $y \in B_X$ such that $x_i \pm y \in S_i$ for every $i \in \{1, \dots, n\}$ and $\|y\| > 1-\varepsilon$.

ASQ Banach spaces were studied in [Abrahamsen et al. 2016]. The SSD2P has not been fully explored yet, but can be found in [Abrahamsen et al. 2013, Lemma 4.1], where it is observed that the SD2P is implied by the SSD2P. In turn, it is not too hard to show that ASQ Banach spaces have the SSD2P. On the other hand, the space $L_1[0, 1]$ has the SD2P, but not the SSD2P (see [Haller et al. 2018, Remark 3.3]), and the space $C[0, 1]$ has the SSD2P, but is not ASQ.

The ASQ property of a Banach space, and also the SSD2P, are rather strong. The widest class of spaces known to be ASQ are nonreflexive M-embedded spaces [Abrahamsen et al. 2016, Corollary 4.3]. Also, $c_0(X)$ is ASQ for any Banach space X. The widest class of spaces known to have the SSD2P are uniform algebras [Abrahamsen et al. 2013, Theorem 4.2]. Also, $\ell_\infty(X)$ has the SSD2P for any Banach space X (see [Haller et al. 2018, Proposition 3.4]).

Let us also relate the *Daugavet property* to the diameter 2 properties. Recall that a bounded linear operator T on a Banach space X is said to satisfy the *Daugavet equation* if $\|I+T\| = 1+\|T\|$. Daugavet [1963] discovered that every compact operator on $C[0, 1]$ satisfies this equation, thus initiating a very important topic in the theory of Banach spaces. Lozanovskii [1966] obtained the analogous result for $L_1[0, 1]$.

Definition 1.4. A Banach space X has the *Daugavet property* if every rank 1 operator on X satisfies the Daugavet equation.

Note that if a Banach space X has the Daugavet property, then, in fact, every weakly compact operator on X satisfies the Daugavet equation (see, e.g., [Kadets et al. 2000, Theorem 2.3] or [Werner 2001, Theorem 2.7]).

Towards the end of the 1990s, the Daugavet property was described in geometrical terms (see [Kadets et al. 2000, Lemmas 2.1 and 2.2; Shvydkoy 2000, Lemmas 2 and 3; Werner 2001, Lemmas 2.2–2.4]). Spaces with the Daugavet property have the SD2P [Abrahamsen et al. 2013, Theorem 4.4] and are octahedral [Becerra Guerrero et al. 2014, Corollary 2.5].

Finally, we can announce our main results: *An extremely regular subspace of* $C_0(L)$

- *has the SSD2P* (Theorem 2.2);
- *is ASQ whenever L is noncompact* (Theorem 2.5);
- *has the Daugavet property whenever L does not contain isolated points* (Theorem 2.6);
- *contains an ε-isometric copy of c_0 whenever* $0 < \varepsilon < 1$ (Theorem 3.1).

In fact, we prove these results for a wider class of subspaces of $C_0(L)$ than extremely regular ones, that we call *somewhat regular* subspaces (see Definition 2.1). Throughout the paper, it should not cause any confusion to denote, for a functional $\mu \in C_0(L)^*$, its representing (regular) Borel measure also by μ.

2. Diameter 2 properties for subspaces of $C_0(L)$

Definition 2.1. We call a linear subspace $\mathcal{A}$ of $C_0(L)$ *somewhat regular,* if, whenever V is a nonempty open subset of L and $0 < \varepsilon < 1$, there is an $f \in \mathcal{A}$ such that

$$\|f\| = 1 \quad \text{and} \quad |f(x)| \leq \varepsilon \qquad \text{for every } x \in L \setminus V. \tag{2-1}$$

Notice that, in this case, $|f(x_0)| = 1$ for some $x_0 \in V$, thus one may choose an $f \in \mathcal{A}$ satisfying (2-1) so that $f(x_0) = 1$ for some $x_0 \in V$.

It is clear that extremely regular subspaces of $C_0(L)$ are somewhat regular. On the other hand, whenever x_1 and x_2 are different accumulation points of L, the subspace $\{f \in C_0(L) : f(x_1) = 2f(x_2)\}$ of $C_0(L)$ is somewhat regular by courtesy of Urysohn's lemma, but fails to be extremely regular. Thus the class of somewhat regular subspaces of $C_0(L)$ is strictly larger than that of extremely regular ones.

Theorem 2.2. *Somewhat regular linear subspaces of* $C_0(L)$ *have the* SSD2P.

Theorem 2.2 is a corollary of the following lemma.

Lemma 2.3. *Let* $\mathcal{A}$ *be a somewhat regular linear subspace of* $C_0(L)$, *and let* $n, m \in \mathbb{N}$, $f_1, \dots, f_n \in B_{\mathcal{A}}$, $\mu_1, \dots, \mu_m \in B_{C_0(L)^*}$, *and* $\varepsilon > 0$. *Then there are* $g_1, \dots, g_n$, $\phi \in B_{\mathcal{A}}$ *such that, for every* $j \in \{1, \dots, n\}$,

(1) $|\mu_i(f_j - g_j)| < \varepsilon$ *for every* $i \in \{1, \dots, m\}$;

(2) $|\mu_i(\phi)| < \varepsilon$ *for every* $i \in \{1, \dots, m\}$;

(3) $\|\phi\| > 1 - \varepsilon$;

(4) $\|g_j \pm \phi\| \le 1$.

When dealing with subspaces of $C_0(L)$, the main challenge is often to find a substitute for Urysohn's lemma (see, e.g., [Cascales et al. 2013, Section 2]). The following lemma — which the proofs of both Lemma 2.3 and Theorem 2.6 rely on — is a "Urysohn's lemma" for somewhat regular subspaces of $C_0(L)$. The lemma is inspired by [Nygaard and Werner 2001, proof of Theorem 1].

Lemma 2.4 (cf. [Nygaard and Werner 2001, proof of Theorem 1]). *Let* $\mathcal{A}$ *be a somewhat regular linear subspace of* $C_0(L)$, *let* V *be a nonempty open subset of* L, *and let* $0 < \varepsilon < 1$. *Then there are an* $x_0 \in V$ *and an* $f \in \mathcal{A}$ *such that*

(1) $f(x_0) = 1 \le \|f\| \le 1 + \varepsilon$;

(2) $|1 - f(x)| \le 1 + \varepsilon$ *for every* $x \in V$;

(3) $|f(x)| \le \varepsilon$ *for every* $x \in L \setminus V$.

Proof. Let $0 < \delta < 1$ and let $n \in \mathbb{N}$ satisfy $2/n < \delta$. Putting $V_0 := V$, by courtesy of the somewhat regularity of $\mathcal{A}$, one can recursively find points $x_1, \dots, x_n \in V$, functions $g_1, \dots, g_n \in \mathcal{A}$, and nonvoid open subsets $V_0 \supset V_1 \supset \cdots \supset V_n$ such that, for every $j \in \{1, \dots, n\}$,

$$x_j \in V_{j-1}, \qquad g_j(x_j) = \|g_j\| = 1, \qquad |g_j(x)| \le \delta \text{ for every } x \in L \setminus V_{j-1},$$

and $V_j = \{x \in V_{j-1} : |g_j(x) - 1| < \delta\}$; thus, in fact, $x_j \in V_j$. Defining $x_0 := x_n$ and

$$g := \frac{g_1 + \cdots + g_n}{n},$$

one has $\|g\| \leq 1$, $|g(x)| \leq \delta$ for every $x \in L \setminus V$, and

$$|1 - g(x)| \leq \frac{1}{n} \sum_{j=1}^{n} |1 - g_j(x)| \quad \text{for every } x \in L.$$

Now let $x \in V$. Put $k := \max\{j \in \{0, 1, \dots, n\} : x \in V_j\}$. For $1 \leq j \leq k$, one has $|1 - g_j(x)| < \delta$; $\delta \leq |1 - g_{k+1}(x)| \leq 2$; and, for $k+2 \leq j \leq n$, one has $|g_j(x)| \leq \delta$ and hence $|1 - g_j(x)| \leq 1 + \delta$. Thus

$$|1 - g(x)| \leq \frac{(n-1)(1+\delta)+2}{n} < 1 + \delta + \frac{2}{n} < 1 + 2\delta.$$

Since $x_0 = x_n \in V_n$, one has $|g_j(x_0) - 1| < \delta$ for every $j \in \{1, \dots, n\}$ and thus $|g(x_0) - 1| < \delta$. Defining $f := (1/g(x_0))g$, it remains to observe that, taking, from the very beginning, δ to be "small enough", the conditions (1)–(3) hold, because, since $|g(x_0)| > 1 - \delta$,

$$1 = f(x_0) \leq \|f\| = \frac{\|g\|}{|g(x_0)|} < \frac{1}{1-\delta},$$

for every $x \in V$,

$$|1 - f(x)| = \frac{|g(x_0) - g(x)|}{|g(x_0)|} < \frac{|g(x_0) - 1| + |1 - g(x)|}{1 - \delta} < \frac{1 + 3\delta}{1 - \delta},$$

and, for every $x \in L \setminus V$,

$$|f(x)| = \frac{|g(x)|}{|g(x_0)|} < \frac{\delta}{1 - \delta}. \qquad \square$$

Proof of Lemma 2.3. Let $0 < \delta < 1/2$. Since L is infinite, there is a point $y \in L$ such that $\max_{1 \leq i \leq m} |\mu_i|(\{y\}) < \delta$; hence, by the regularity of $\mu_1, \dots, \mu_m$ and the continuity of $f_1 \dots, f_n$, there is a nonempty open subset V of L such that

$$\max_{1 \leq i \leq m} |\mu_i|(V) < \delta \quad \text{and} \quad \max_{1 \leq j \leq n} \sup_{x,z \in V} |f_j(x) - f_j(z)| < \delta.$$

Since, by our assumption, $\mathcal{A}$ is somewhat regular, there are $x_0 \in V$ and $f \in \mathcal{A}$ satisfying the conditions (1)–(3) of Lemma 2.4 with ε replaced by δ. For every $j \in \{1, \dots, n\}$, defining $\alpha_j := f_j(x_0)$ and $h_j := f_j - \alpha_j f \in \mathcal{A}$, one has $h_j(x_0) = 0$ and $\|h_j\| \leq 1 + 2\delta$, because

$$|h_j(x)| \leq \begin{cases} |f_j(x) - \alpha_j| + |\alpha_j|\,|1 - f(x)| \leq 1 + 2\delta & \text{if } x \in V; \\ |f_j(x)| + |\alpha_j|\,|f(x)| \leq 1 + \delta & \text{if } x \in L \setminus V. \end{cases}$$

For every $j \in \{1, \dots, n\}$, defining $g_j := (1 - 2\delta)h_j$, one has $\|g_j\| \leq 1 - 4\delta^2$ and, for every $i \in \{1, \dots, m\}$, since

$$|\mu_i(f)| \leq \int_{L \setminus V} |f|\, d|\mu_i| + \int_V |f|\, d|\mu_i| \leq \delta\, |\mu_i|(L \setminus V) + 2|\mu_i|(V) < 3\delta,$$

one also has

$$|\mu_i(f_j - g_j)| \le 2\delta|\mu_i(f_j)| + (1-2\delta)|\alpha_j|\,|\mu_i(f)| < 5\delta.$$

Choose an open neighbourhood $U \subset V$ of x_0 such that

$$\max_{1\le j\le n} \sup_{x\in U} |g_j(x)| \le \delta.$$

Since $\mathcal{A}$ is somewhat regular, there is a $\psi \in \mathcal{A}$ such that

$$\|\psi\| = 1 \quad \text{and} \quad |\psi(x)| \le 4\delta^2 \ \text{ for every } x \in L \setminus U.$$

Put $\phi := (1-\delta)\psi$. Then, for every $j \in \{1, \dots, m\}$, one has $\|g_j \pm \phi\| \le 1$, i.e., the condition (4) holds, and, for every $i \in \{1, \dots, m\}$,

$$|\mu_i(\phi)| \le \int_{L\setminus V} |\phi|\, d|\mu_i| + \int_V |\phi|\, d|\mu_i| \le 4\delta^2\, |\mu_i|(L\setminus V) + |\mu_i|(V) < 5\delta.$$

Thus, one observes that taking, from the very beginning, δ to be "small enough", the conditions (1)–(3) also hold. □

If the space L is noncompact, a stronger statement than that of Theorem 2.2 is true.

Theorem 2.5. *Assume that L is noncompact. Then every somewhat regular linear subspace of $C_0(L)$ is* ASQ.

Proof. Let $\mathcal{A}$ be a somewhat regular linear subspace of $C_0(L)$, and let $n \in \mathbb{N}$, $f_1, \dots, f_n \in S_{\mathcal{A}}$, and $\varepsilon > 0$. It suffices to find an $f \in \mathcal{A}$ such that $\|f\| = 1$ and

$$\|f_j \pm f\| \le 1 + \varepsilon \quad \text{for every } j \in \{1, \dots, n\}. \tag{2-2}$$

To this end, observe that the sets $K_j := \{x \in L : |f_j(x)| \ge \varepsilon\}$, $j = 1, \dots, n$, are compact; thus also their union $K := \bigcup_{j=1}^n K_j$ is compact, and its complement $V := L \setminus K$ is nonempty and open. By the somewhat regularity of $\mathcal{A}$, there is an $f \in \mathcal{A}$ satisfying (2-1). This f also satisfies (2-2). □

Our next result produces examples of spaces with the Daugavet property.

Theorem 2.6. *Assume that L does not contain isolated points. Then every somewhat regular linear subspace of $C_0(L)$ has the Daugavet property.*

Proof. Let $\mathcal{A}$ be a somewhat regular linear subspace of $C_0(L)$, let $g \in S_{\mathcal{A}}$, let $\mu \in S_{C_0(L)^*}$ be such that $\|\mu|_{\mathcal{A}}\| = 1$, and let $\alpha, \varepsilon > 0$. In order for $\mathcal{A}$ to have the Daugavet property, by [Werner 2001, Lemma 2.2] (or [Kadets et al. 2000, Lemma 2.2]), it suffices to find a $\psi \in S_{\mathcal{A}}$ satisfying

$$\operatorname{Re} \mu(\psi) > 1 - \alpha \quad \text{and} \quad \|g + \psi\| > 2 - \varepsilon. \tag{2-3}$$

To this end, let $\delta \in (0, 1/3)$, let $\phi \in S_{\mathcal{A}}$ be such that $\operatorname{Re} \mu(\phi) > 1 - \delta$, let $y_0 \in L$

be such that $|g(y_0)| = 1$, and let an open neighbourhood U of y_0 be such that

$$|g(x) - g(y_0)| < \delta \quad \text{and} \quad |\phi(x) - \phi(y_0)| < \delta \qquad \text{for all } x \in U.$$

Since y_0 is not an isolated point, the set U is infinite; thus there is a point $z_0 \in U$ such that $|\mu|(\{z_0\}) < \delta$. By the regularity of μ, there is an open neighbourhood V of z_0 such that $|\mu|(V) < \delta$. One may assume that $V \subset U$ and thus

$$|\phi(x) - \phi(z)| < 2\delta \quad \text{for all } x, z \in V.$$

Since $\mathcal{A}$ is somewhat regular, there are $x_0 \in V$ and $f \in \mathcal{A}$ satisfying the conditions (1)–(3) of Lemma 2.4 with ε replaced by δ. Put $h := \phi - \phi(x_0) f$; then $h(x_0) = 0$ and $\|h\| \le 1 + 3\delta$ (this can be shown as in the proof of Theorem 2.2 for h_j); indeed,

$$|h(x)| \le \begin{cases} |\phi(x) - \phi(x_0)| + |\phi(x_0)|\,|1 - f(x)| \le 1 + 3\delta & \text{if } x \in V; \\ |\phi(x)| + |\phi(x_0)|\,|f(x)| \le 1 + \delta & \text{if } x \in L \setminus V. \end{cases}$$

Since $h(x_0) = 0$, there is an open neighbourhood W of x_0 such that

$$|h(x)| < \delta \quad \text{for all } x \in W.$$

One may assume that $W \subset V$. Since $\mathcal{A}$ is somewhat regular, there are $w_0 \in W$ and $\hat{f} \in \mathcal{A}$ such that

$$\hat{f}(w_0) = \|\hat{f}\| = 1 \quad \text{and} \quad |\hat{f}(x)| \le \delta \ \text{ for every } x \in L \setminus W.$$

Putting $\hat{\psi} := h + g(w_0)\hat{f}$, one has $\|\hat{\psi}\| \le 1 + 4\delta$, because

$$|\hat{\psi}(x)| \le |h(x)| + |\hat{f}(x)| \le \begin{cases} \delta + 1 & \text{if } x \in W; \\ (1 + 3\delta) + \delta = 1 + 4\delta & \text{if } x \in L \setminus W. \end{cases}$$

Since

$$\begin{aligned} \|\hat{\psi} + g\| &\ge |\hat{\psi}(w_0) + g(w_0)| \ge 2|g(w_0)| - |h(w_0)| \\ &\ge 2|g(y_0)| - 2|g(y_0) - g(w_0)| - |h(w_0)| \\ &> 2 - 2\delta - \delta = 2 - 3\delta, \end{aligned}$$

one has $\|\hat{\psi}\| > 1 - 3\delta$; thus, for $\psi := \hat{\psi}/\|\hat{\psi}\|$, one has

$$\|\hat{\psi} - \psi\| = \left|1 - \frac{1}{\|\hat{\psi}\|}\right| \|\hat{\psi}\| = \big|\|\hat{\psi}\| - 1\big| \le 4\delta,$$

and hence

$$\|g + \psi\| \ge \|g + \hat{\psi}\| - \|\hat{\psi} - \psi\| > 2 - 3\delta - 4\delta = 2 - 7\delta.$$

One has

$$\begin{aligned} \operatorname{Re}\mu(\hat{\psi}) &= \operatorname{Re}\mu(h) + \operatorname{Re} g(w_0)\mu(\hat{f}) = \operatorname{Re}\mu(\phi) - \operatorname{Re}\phi(x_0)\mu(f) + \operatorname{Re} g(w_0)\mu(\hat{f}) \\ &> 1 - \delta - |\mu(f)| - |\mu(\hat{f})|. \end{aligned}$$

Since

$$|\mu(f)| \le \left|\int_L f\,d\mu\right| \le \int_L |f|\,d|\mu| = \int_V |f|\,d|\mu| + \int_{L\setminus V} |f|\,d|\mu|$$
$$\le (1+\delta)|\mu|(V) + \delta|\mu|(L\setminus V) < (1+\delta)\delta + \delta = (2+\delta)\delta < 3\delta,$$

and, similarly, $|\mu(\hat{f})| < 2\delta$, it follows that $\operatorname{Re}\mu(\hat{\psi}) > 1 - 6\delta$, and thus

$$\operatorname{Re}\mu(\psi) = \frac{\operatorname{Re}\mu(\hat{\psi})}{\|\hat{\psi}\|} \ge \frac{\operatorname{Re}\mu(\hat{\psi})}{1+4\delta} > \frac{1-6\delta}{1+4\delta}.$$

Hence one observes that, choosing, from the very beginning, δ to be "small enough", the function ψ meets the conditions (2-3). □

3. Containment of c_0 and ℓ_1

Let X and Y be normed spaces, and let $0 < \varepsilon < 1$. Recall that a linear surjection $T : X \to Y$ is called an *ε-isometry* if

$$(1-\varepsilon)\|x\| \le \|Tx\| \le (1+\varepsilon)\|x\| \quad \text{for every } x \in X.$$

It is well known that $C_0(L)$ contains isometric copies of c_0 (see, e.g., [Albiac and Kalton 2006, Proposition 4.3.11]), and the same is true for many of its subspaces. For the somewhat regular linear subspaces of $C_0(L)$ we have the following theorem.

Theorem 3.1. *Let $\mathcal{A}$ be a somewhat regular closed linear subspace of $C_0(L)$. Then, whenever $0 < \varepsilon < 1$, there is an ε-isometry from c_0 onto a closed linear subspace of $\mathcal{A}$.*

Proof. Let $0 < \varepsilon < 1$. Choose pairwise disjoint nonvoid open subsets U_j, $j \in \mathbb{N}$, of L. Since $\mathcal{A}$ is somewhat regular, for every $j \in \mathbb{N}$, there are an $x_j \in U_j$ and an $f_j \in \mathcal{A}$ such that

$$f_j(x_j) = 1 \quad \text{and} \quad |f_j(x)| \le \frac{\varepsilon}{2^j} \ \text{ for every } x \in L \setminus V_j.$$

Denoting by c_{00} the linear subspace of finitely supported sequences in c_0, let $S_0 : c_{00} \to \mathcal{A}$ be the linear operator satisfying $S_0 e_j = f_j$ for every $j \in \mathbb{N}$ where e_j are the standard unit vectors in c_0. Observing that, whenever $a = \sum_{j=1}^n \alpha_j e_j \in S_{c_{00}}$ and $x \in L$, one has

$$|(S_0 a)(x)| = \left|\sum_{j=1}^n \alpha_j f_j(x)\right| \le \sum_{j=1}^n |f_j(x)| \le 1 + \sum_{j=1}^n \frac{\varepsilon}{2^j} < 1 + \varepsilon$$

(because $|f_j(x)| \le \varepsilon/2^j$ whenever $x \notin U_j$, and there is at most one $j \in \mathbb{N}$ such that $x \in U_j$ (in which case $|f_j(x)| \le 1$)), thus S_0 is bounded and $\|S_0\| \le 1 + \varepsilon$. Letting $S : c_0 \to \mathcal{A}$ be the bounded linear extension of S_0, one has $\|S\| \le 1 + \varepsilon$ as well, and it remains to observe that, whenever $a = (\alpha_j)_{j=1}^{\infty} \in c_0$, picking $k \in \mathbb{N}$ such that $|\alpha_k| = \|a\|$, one has

$$\|Sa\| = \left\| \sum_{j=1}^{\infty} \alpha_j f_j \right\| \geq \left| \sum_{j=1}^{\infty} \alpha_j f_j(x_k) \right| \geq |\alpha_k| \, |f_k(x_k)| - \sum_{j=1;\, j \neq k}^{\infty} |\alpha_j| \, |f_j(x_k)|$$

$$\geq \|a\| - \|a\| \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = (1-\varepsilon)\|a\|,$$

because, for $j \neq k$, one has $x_k \notin V_j$ and thus $|f_j(x_k)| \leq \varepsilon$. □

It is natural to ask about containment of ℓ_1 in somewhat regular linear subspaces of $C_0(L)$. If L does not contain isolated points, we have from Theorem 2.6 and [Kadets et al. 2000, Theorem 2.9] that all somewhat regular linear subspaces of $C_0(L)$ contain ℓ_1 (even asymptotically isometric copies of ℓ_1). But, if L contains isolated points, the picture is not so clear. In this case there might be somewhat regular subspaces of $C_0(L)$ which contain ℓ_1 and other such subspaces which do not. For an example, take $C(\beta\mathbb{N})$ and its subspaces $X = \{f \in C(\beta\mathbb{N}) : f(x) = 0 \text{ for every } x \in \beta\mathbb{N} \setminus \mathbb{N}\}$ and $Y = \{f \in C(\beta\mathbb{N}) : f(y) = 0\}$ where $y \in \beta\mathbb{N} \setminus \mathbb{N}$ is a fixed element. It is straightforward to show that both these subspaces are somewhat regular. Moreover, X is isometrically isomorphic to c_0 and Y is isomorphic to $C(\beta\mathbb{N})$.

Acknowledgement

We thank the anonymous referee for useful comments and suggestions that improved the paper.

References

[Abrahamsen et al. 2013] T. Abrahamsen, V. Lima, and O. Nygaard, "Remarks on diameter 2 properties", *J. Convex Anal.* **20**:2 (2013), 439–452. MR Zbl

[Abrahamsen et al. 2016] T. A. Abrahamsen, J. Langemets, and V. Lima, "Almost square Banach spaces", *J. Math. Anal. Appl.* **434**:2 (2016), 1549–1565. MR Zbl

[Abrahamsen et al. 2017] T. A. Abrahamsen, A. Leraand, A. Martiny, and O. Nygaard, "Two properties of Müntz spaces", *Demonstr. Math.* **50**:1 (2017), 239–244. MR Zbl

[Albiac and Kalton 2006] F. Albiac and N. J. Kalton, *Topics in Banach space theory*, Graduate Texts in Mathematics **233**, Springer, 2006. MR Zbl

[Becerra Guerrero et al. 2014] J. Becerra Guerrero, G. López-Pérez, and A. Rueda Zoca, "Octahedral norms and convex combination of slices in Banach spaces", *J. Funct. Anal.* **266**:4 (2014), 2424–2435. MR Zbl

[Becerra Guerrero et al. 2015] J. Becerra Guerrero, G. López-Pérez, and A. Rueda Zoca, "Big slices versus big relatively weakly open subsets in Banach spaces", *J. Math. Anal. Appl.* **428**:2 (2015), 855–865. MR Zbl

[Cascales et al. 2013] B. Cascales, A. J. Guirao, and V. Kadets, "A Bishop–Phelps–Bollobás type theorem for uniform algebras", *Adv. Math.* **240** (2013), 370–382. MR Zbl

[Cengiz 1973a] B. Cengiz, "A generalization of the Banach–Stone theorem", *Proc. Amer. Math. Soc.* **40** (1973), 426–430. MR Zbl

[Cengiz 1973b] B. Cengiz, "On extremely regular function spaces", *Pacific J. Math.* **49** (1973), 335–338. MR Zbl

[Daugavet 1963] I. K. Daugavet, "A property of completely continuous operators in the space C", *Uspehi Mat. Nauk* **18**:5 (1963), 157–158. In Russian. MR Zbl

[Ghoussoub et al. 1987] N. Ghoussoub, G. Godefroy, B. Maurey, and W. Schachermayer, *Some topological and geometrical structures in Banach spaces*, Mem. Amer. Math. Soc. **378**, 1987. MR Zbl

[Godefroy 1989] G. Godefroy, "Metric characterization of first Baire class linear forms and octahedral norms", *Studia Math.* **95**:1 (1989), 1–15. MR Zbl

[Haller and Langemets 2014] R. Haller and J. Langemets, "Two remarks on diameter 2 properties", *Proc. Estonian Acad. Sci.* **63**:1 (2014), 2–7. Zbl

[Haller et al. 2018] R. Haller, J. Langemets, V. Lima, and R. Nadel, "Symmetric strong diameter two property", preprint, 2018. To appear in *Mediterr. J. Math.* arXiv

[Kadets et al. 2000] V. M. Kadets, R. V. Shvidkoy, G. G. Sirotkin, and D. Werner, "Banach spaces with the Daugavet property", *Trans. Amer. Math. Soc.* **352**:2 (2000), 855–873. MR Zbl

[Lozanovskii 1966] G. J. Lozanovskii, "On almost integral operators in *KB*-spaces", *Vestnik Leningrad. Univ.* **21**:7 (1966), 35–44. In Russian. MR Zbl

[Nygaard and Werner 2001] O. Nygaard and D. Werner, "Slices in the unit ball of a uniform algebra", *Arch. Math.* (*Basel*) **76**:6 (2001), 441–444. MR Zbl

[Shvydkoy 2000] R. V. Shvydkoy, "Geometric aspects of the Daugavet property", *J. Funct. Anal.* **176**:2 (2000), 198–212. MR Zbl

[Werner 2001] D. Werner, "Recent progress on the Daugavet property", *Irish Math. Soc. Bull.* **46** (2001), 77–97. MR Zbl

Received November 2, 2017. Revised September 25, 2018.

TROND A. ABRAHAMSEN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF AGDER
KRISTIANSAND
NORWAY
trond.a.abrahamsen@uia.no

OLAV NYGAARD
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF AGDER
KRISTIANSAND
NORWAY
olav.nygaard@uia.no

MÄRT PÕLDVERE
INSTITUTE OF MATHEMATICS AND STATISTICS
UNIVERSITY OF TARTU
TARTU
ESTONIA
mart.poldvere@ut.ee

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.395

REGULARITY AND UPPER SEMICONTINUITY OF PULLBACK ATTRACTORS FOR A CLASS OF NONAUTONOMOUS THERMOELASTIC PLATE SYSTEMS

FLANK D. M. BEZERRA, VERA L. CARBONE,
MARCELO J. D. NASCIMENTO AND KARINA SCHIABEL

We study the long-time dynamics, in the sense of pullback attractors, of solutions for semilinear nonautonomous thermoelastic plate systems in a bounded smooth domain in $\mathbb{R}^N$, $N \geqslant 2$. Using the theory of uniform sectorial operators, in the sense of P. Sobolevskiĭ (1961), we will prove existence, uniform boundedness, regularity and upper semicontinuity of pullback attractors for the evolution system

$$\begin{cases} u_{tt} + \Delta^2 u + a\Delta\theta = f(u), & t > \tau,\ x \in \Omega, \\ \theta_t - \kappa(t)\Delta\theta - a\Delta u_t = 0, & t > \tau,\ x \in \Omega, \end{cases}$$

subject to boundary conditions

$$u = \Delta u = \theta = 0, \quad t > \tau, \quad x \in \partial\Omega,$$

with respect to the functional parameter κ.

1. Introduction

In this paper we study a model that describes the small vibrations of a homogeneous, elastic and thermal isotropic Euler–Bernoulli plate. In fact we consider the initial-boundary value problem

$$\begin{cases} u_{tt} + \Delta^2 u + a\Delta\theta = f(u), & t > \tau,\ x \in \Omega, \\ \theta_t - \kappa(t)\Delta\theta - a\Delta u_t = 0, & t > \tau,\ x \in \Omega, \end{cases} \tag{1-1}$$

subject to boundary conditions

$$\begin{cases} u = \Delta u = 0,\ t > \tau,\ x \in \partial\Omega, \\ \theta = 0,\ t > \tau,\ x \in \partial\Omega, \end{cases} \tag{1-2}$$

Bezerra's research was partially supported by FAPESP #2014/03686-3, Brazil. Nascimento's research was partially supported by FAPESP #2017/06582-2, Brazil.
MSC2010: 35B40, 35B41, 35B65, 37B55.
Keywords: pullback attractors, regularity, thermoelastic plate systems, upper semicontinuity.

and initial conditions

$$u(\tau,x)=u_0(x),\ u_t(\tau,x)=v_0(x)\ \text{ and }\ \theta(\tau,x)=\theta_0(x),\quad x\in\Omega,\ \tau\in\mathbb{R}, \tag{1-3}$$

where Ω is a bounded domain in $\mathbb{R}^N$ with $N \geqslant 2$, where the boundary $\partial\Omega$ is assumed to be regular enough and $a > 0$.

Next we exhibit conditions under which the nonautonomous problem (1-1)–(1-3) is locally and globally well posed in some appropriate space that we will specify later.

We assume that κ is continuously differentiable in $\mathbb{R}$ and satisfies

$$0<\kappa_0\leqslant\kappa(t),\kappa'(t)\leqslant\kappa_1\quad\text{for all } t\in\mathbb{R}, \tag{1-4}$$

for some positive constants κ_0 and κ_1.

Suppose that $f:\mathbb{R}\to\mathbb{R}$ is locally Lipschitz satisfying

$$\limsup_{|s|\to\infty}\frac{f(s)}{s}<\lambda_1 \tag{1-5}$$

uniformly in $t\in\mathbb{R}$, where $\lambda_1>0$ is the first eigenvalue of negative Laplacian operator with homogeneous Dirichlet boundary condition. Furthermore, the function f satisfies the subcritical growth condition; that is,

$$|f'(s)|\leqslant C(1+|s|^{\rho-1})\quad\text{for all } s\in\mathbb{R}, \tag{1-6}$$

where $1\leqslant\rho<\frac{N}{N-4}$, with $N\geqslant 5$, and $C>0$ independent of $t\in\mathbb{R}$. In this case, the embedding $H^2(\Omega)\cap H_0^1(\Omega)\hookrightarrow L^{2N/(N-4)}(\Omega)$ is compact and this will be used in analysis of the energy functionals. We will justify these restrictions later in the paper. If $N=2,3,4$, we suppose the growth condition (1-6) with $\rho\geqslant 1$.

Using the theory of uniform sectorial operators, in the sense of [Sobolevskiĭ 1961], the authors proved in [Bezerra et al. 2018] the local and global well-posedness of the nonautonomous problem (1-1)–(1-3) (under conditions (1-5) and (1-6)), the existence of pullback attractors and uniform bounds for these pullback attractors when $\kappa(t)\equiv\kappa$.

The main goal of this paper is to prove the regularity of the pullback attractors and their upper semicontinuity with respect to the functional parameter κ. For completeness, under the additional condition (1-4) we prove the local and global posedness for (1-1)–(1-3) as well the existence and uniform boundedness of pullback attractors for this problem.

We emphasize that no additional damping in first evolution equation in (1-1) is required in the present work.

To formulate the nonautonomous problem (1-1)–(1-3) in the nonlinear evolution process setting, we introduce some notation. Here, we denote $X=L^2(\Omega)$ and

$\Lambda : D(\Lambda) \subset X \to X$ to be the biharmonic operator defined by

$$D(\Lambda) = \{u \in H^4(\Omega);\ u = \Delta u = 0 \text{ on } \partial\Omega\}$$

and

$$\Lambda u = (-\Delta)^2 u \quad \text{for all } u \in D(\Lambda). \tag{1-7}$$

Then Λ is a positive self-adjoint operator in X with compact resolvent and therefore $-\Lambda$ generates a compact analytic semigroup on X (that is, Λ is a sectorial operator, in the sense of [Henry 1981]). Denote by X^α, $\alpha > 0$, the fractional power spaces associated with the operator Λ; that is, $X^\alpha = D(\Lambda^\alpha)$ endowed with the graph norm. With this notation, we have $X^{-\alpha} = (X^\alpha)'$ for all $\alpha > 0$, (see [Amann 1995]). Of special interest is the case $\alpha = \frac{1}{2}$, since $-\Lambda^{\frac{1}{2}}$ is the Laplacian operator with homogeneous Dirichlet boundary conditions.

If we denote $v = u_t$, then we can rewrite the nonautonomous problem (1-1)–(1-3) in the abstract form

$$\begin{cases} w_t = A_{(\kappa)}(t)w + F(w),\ t > \tau, \\ w(\tau) = w_0,\ \tau \in \mathbb{R}, \end{cases} \tag{1-8}$$

where $w = w(t)$ for all $t \in \mathbb{R}$, and $w_0 = w(\tau)$ are given by

$$w = \begin{bmatrix} u \\ v \\ \theta \end{bmatrix}, \quad \text{and} \quad w_0 = \begin{bmatrix} u_0 \\ v_0 \\ \theta_0 \end{bmatrix}, \tag{1-9}$$

and, for each $t \in \mathbb{R}$, the unbounded linear operator $A_{(\kappa)}(t) : D(A_{(\kappa)}(t)) \subset Y \to Y$ is defined by

$$A_{(\kappa)}(t) \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ -\Lambda & 0 & -a\Lambda^{\frac{1}{2}} \\ 0 & a\Lambda^{\frac{1}{2}} & \kappa(t)\Lambda^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} = \begin{bmatrix} v \\ -\Lambda u - a\Lambda^{\frac{1}{2}}\theta \\ a\Lambda^{\frac{1}{2}}v + \kappa(t)\Lambda^{\frac{1}{2}}\theta \end{bmatrix}, \tag{1-10}$$

where

$$Y = (H^2(\Omega) \cap H^1_0(\Omega)) \times L^2(\Omega) \times L^2(\Omega)$$

is the phase space of the problem (1-1)–(1-3) and the domain of the operator $A_{(\kappa)}(t)$ is defined by the space

$$D(A_{(\kappa)}(t)) = X^1 \times X^{\frac{1}{2}} \times X^{\frac{1}{2}}, \tag{1-11}$$

with $X^1 = \{u \in H^4(\Omega);\ u = \Delta u = 0 \text{ on } \partial\Omega\}$ and $X^{\frac{1}{2}} = H^2(\Omega) \cap H^1_0(\Omega)$.

The nonlinearity F is given by

$$F(w) = \begin{bmatrix} 0 \\ f^e(u) \\ 0 \end{bmatrix}, \tag{1-12}$$

where $f^e(u)$ is the Nemytskii operator associated with $f(u)$; that is,

$$f^e(u)(x) := f(u(x)) \quad \text{for all } x \in \Omega.$$

This paper is organized as follows: in Section 2 we recall concepts and results about singularly nonautonomous problems. Section 3 is devoted to studying the existence of local and global solutions in some appropriate space, as well as the existence of pullback attractors for (1-1)–(1-3). In Section 4 we present results on regularity of the pullback attractors, following Carvalho, Langa, Robinson [Carvalho et al. 2013]. Finally, in Section 5 we prove that the family of pullback attractors behave upper semicontinuously with respect to the functional parameter κ.

2. Abstract linear problem

Throughout the paper, $L(\mathcal{Z})$ will denote the space of linear and bounded operators defined in a Banach space $\mathcal{Z}$. Let $\mathcal{B}(t)$, $t \in \mathbb{R}$, be a family of unbounded closed linear operators defined on a fixed dense subspace D of $\mathcal{Z}$.

2A. ***Nonautonomous abstract linear problem.*** Consider the singularly nonautonomous abstract linear parabolic problem of the form

$$\begin{cases} \dfrac{du}{dt} = -\mathcal{B}(t)u, \ t > \tau, \\ u(\tau) = u_0 \in D. \end{cases}$$

We assume that:

(a) The family of operators $\mathcal{B}(t) : D \subset \mathcal{Z} \to \mathcal{Z}$ is *uniformly sectorial*, that is, $\mathcal{B}(t)$ is closed densely defined (the domain D is fixed) and there is a constant $C > 0$ (independent of $t \in \mathbb{R}$) such that

$$\|(\mathcal{B}(t) + \lambda I)^{-1}\|_{L(\mathcal{Z})} \leqslant \frac{C}{|\lambda| + 1} \quad \text{for all } \lambda \in \mathbb{C} \text{ with } \operatorname{Re} \lambda \geqslant 0.$$

(b) The map $\mathbb{R} \ni t \mapsto \mathcal{B}(t)$ is *uniformly Hölder continuous*, that is, there are constants $C > 0$ and $\varepsilon_0 > 0$ such that, for any $t, \tau, s \in \mathbb{R}$,

$$\|[\mathcal{B}(t) - \mathcal{B}(\tau)]\mathcal{B}^{-1}(s)\|_{L(\mathcal{Z})} \leqslant C(t - \tau)^{\varepsilon_0}, \quad \varepsilon_0 \in (0, 1].$$

Denote by $\mathcal{B}_0$ the operator $\mathcal{B}(t_0)$ for some $t_0 \in \mathbb{R}$ fixed. If $\mathcal{Z}^\alpha$ denotes the domain of $\mathcal{B}_0^\alpha$, $\alpha > 0$, with the graph norm and $\mathcal{Z}^0 := \mathcal{Z}$, denote by $\{\mathcal{Z}^\alpha; \alpha \geqslant 0\}$ the fractional power scale associated with $\mathcal{B}_0$.

From (a), $-\mathcal{B}(t)$ is the generator of an analytic semigroup $\{e^{-\tau\mathcal{B}(t)} \in L(\mathcal{Z}) : \tau \geqslant 0\}$. Using this and the fact that $0 \in \rho(\mathcal{B}(t))$, it follows that

$$\|e^{-\tau\mathcal{B}(t)}\|_{L(\mathcal{Z})} \leqslant C, \quad \tau \geqslant 0, \ t \in \mathbb{R},$$

and

$$\|\mathcal{B}(t)e^{-\tau\mathcal{B}(t)}\|_{L(\mathcal{Z})} \leqslant C\tau^{-1}, \quad \tau > 0, \ t \in \mathbb{R}.$$

It follows from (b) that $\|\mathcal{B}(t)\mathcal{B}^{-1}(\tau)\|_{L(\mathcal{Z})} \leqslant C$, for all $(t, \tau) \in I$, for some $I \subset \mathbb{R}^2$ bounded. Also, the semigroup $e^{-\tau\mathcal{B}(t)}$ generated by $-\mathcal{B}(t)$ satisfies the estimate

$$\|e^{-\tau\mathcal{B}(t)}\|_{L(\mathcal{Z}^\beta,\mathcal{Z}^\alpha)} \leqslant M\tau^{\beta-\alpha}, \tag{2-1}$$

where $0 \leqslant \beta \leqslant \alpha < 1 + \varepsilon_0$.

Next we recall the definition of a linear evolution process associated with a family of operators $\{\mathcal{B}(t) : t \in \mathbb{R}\}$.

Definition 2.1. A family $\{L(t, \tau) : t \geqslant \tau \in \mathbb{R}\} \subset L(\mathcal{Z})$ satisfying

(1) $L(\tau, \tau) = I$,

(2) $L(t, \sigma)L(\sigma, \tau) = L(t, \tau)$ for any $t \geqslant \sigma \geqslant \tau$,

(3) $\mathcal{P} \times \mathcal{Z} \ni ((t, \tau), u_0) \mapsto L(t, \tau)v_0 \in \mathcal{Z}$ is continuous, where $\mathcal{P} = \{(t, \tau) \in \mathbb{R}^2 : t \geqslant \tau\}$

is called a *linear evolution process* (*process* for short) or *family of evolution operators*.

If the operator $\mathcal{B}(t)$ is uniformly sectorial and uniformly Hölder continuous, then there exists a linear evolution process $\{L(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ associated with $\mathcal{B}(t)$, which is given by

$$L(t, \tau) = e^{-(t-\tau)\mathcal{B}(\tau)} + \int_\tau^t L(t, s)[\mathcal{B}(\tau) - \mathcal{B}(s)]e^{-(s-\tau)\mathcal{B}(\tau)}\, ds.$$

The evolution process $\{L(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ satisfies the condition

$$\|L(t, \tau)\|_{\mathcal{L}(\mathcal{Z}^\beta,\mathcal{Z}^\alpha)} \leqslant C(\alpha, \beta)(t - \tau)^{\beta-\alpha}, \tag{2-2}$$

where $0 \leqslant \beta \leqslant \alpha < 1 + \varepsilon_0$. For more details see [Carvalho and Nascimento 2009] and [Sobolevskiĭ 1961].

2B. ***Abstract results on pullback attractors.*** In this subsection we will present basic definitions and results of the theory of pullback attractors for nonlinear evolution processes. For more details, we refer to [Caraballo et al. 2010], [Carvalho et al. 2013] and [Chepyzhov and Vishik 2002].

We consider the singularly nonautonomous abstract parabolic problem

$$\begin{cases} \dfrac{du}{dt} = -\mathcal{B}(t)u + g(u), \ t > \tau, \\ u(\tau) = u_0 \in D, \end{cases} \tag{2-3}$$

where the operator $\mathcal{B}(t)$ is uniformly sectorial and uniformly Hölder continuous and the nonlinearity g satisfies conditions which will be specified later. The nonlinear evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ associated with $\mathcal{B}(t)$ is given by

$$S(t,\tau)u_0 = L(t,\tau)u_0 + \int_\tau^t L(t,s)g(S(s,\tau))\,ds \quad \text{for all } t \geqslant \tau.$$

Definition 2.2. Let $g : \mathbb{R} \times X^\alpha \to X^\beta$, $\alpha \in [\beta, \beta+1)$, be a continuous function. We say that a continuous function $u : [\tau, \tau+t_0] \to X^\alpha$ is a (*local*) *solution* of (2-3) starting in $u_0 \in X^\alpha$ if $u \in C([\tau, \tau+t_0], X^\alpha) \cap C^1((\tau, \tau+t_0], X^\alpha)$, $u(\tau) = u_0$, $u(t) \in D(\mathcal{B}(t))$ for all $t \in (\tau, \tau+t_0]$ and (2-3) is satisfied for all $t \in (\tau, \tau+t_0)$.

We can now state the following result, from [Caraballo et al. 2011]. We also refer to [Carvalho and Nascimento 2009] for a more general version that includes the critical growth case.

Theorem 2.3. *Suppose that the family of operators $\mathcal{B}(t)$ is uniformly sectorial and uniformly Hölder continuous in X^β. If $g : X^\alpha \to X^\beta$, $\alpha \in [\beta, \beta+1)$, is a Lipschitz continuous map in bounded subsets of X^α, then, given $r > 0$, there is a time $t_0 > 0$ such that for all $u_0 \in B_{X^\alpha}(0; r)$ (open ball of radius r centered at the origin of X^α) there exists a unique solution of the problem* (2-3) *starting in u_0 and defined in* $[\tau, \tau+t_0]$. *Moreover, such solutions are continuous with respect the initial data in* $B_{X^\alpha}(0; r)$.

Next we present several definitions from the theory of pullback attractors, which can be found in [Caraballo et al. 2010; 2013; Chepyzhov and Vishik 2002].

We begin by recalling the definition of Hausdorff semidistance between two subsets A and B of a metric space (X, d):

$$\operatorname{dist}_H(A, B) = \sup_{a\in A} \inf_{b\in B} d(a, b).$$

Definition 2.4. Let $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ be an evolution process in a metric space X. Given A and B subsets of X, we say that A *pullback attracts* B at time t if

$$\lim_{\tau\to-\infty} \operatorname{dist}_H(S(t,\tau)B, A) = 0,$$

where $S(t,\tau)B := \{S(t,\tau)x \in X : x \in B\}$.

Definition 2.5. The *pullback orbit* of a subset $B \subset X$ relative to the evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ in the time $t \in \mathbb{R}$ is defined by $\gamma_p(B, t) := \bigcup_{\tau\leqslant t} S(t,\tau)B$.

Definition 2.6. An evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ in X is *pullback strongly bounded* if, for each $t \in \mathbb{R}$ and each bounded subset B of X, $\bigcup_{\tau\leqslant t} \gamma_p(B, \tau)$ is bounded.

Definition 2.7. An evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ in X is *pullback asymptotically compact* if, for each $t \in \mathbb{R}$, each sequence $\{\tau_n\}$ in $(-\infty, t]$ with $\tau_n \to -\infty$

as $n \to \infty$ and each bounded sequence $\{x_n\}$ in X such that $\{S(t, \tau_n)x_n\} \subset X$ is bounded, the sequence $\{S(t, \tau_n)x_n\}$ is relatively compact in X.

Definition 2.8. We say that a family of bounded subsets $\{B(t) : t \in \mathbb{R}\}$ of X is *pullback absorbing* for the evolution process $\{S(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ if, for each $t \in \mathbb{R}$ and for any bounded subset B of X, there exists $\tau_0(t, B) \leqslant t$ such that

$$S(t, \tau)B \subset B(t) \quad \text{for all } \tau \leqslant \tau_0(t, B).$$

Definition 2.9. A family of subsets $\{\mathcal{A}(t) : t \in \mathbb{R}\}$ of X is called a *pullback attractor* for the evolution process $\{S(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ if it is invariant (that is, $S(t, \tau)\mathcal{A}(\tau) = \mathcal{A}(t)$, for any $t \geqslant \tau$), $\mathcal{A}(t)$ is compact for all $t \in \mathbb{R}$, and pullback attracts bounded subsets of X at time t, for each $t \in \mathbb{R}$.

In applications, to prove a process has a pullback attractor, we use Theorem 2.11, proved in [Caraballo et al. 2010], which gives a sufficient condition for existence of a compact pullback attractor. For this, we will need the concept of pullback strongly bounded dissipativeness.

Definition 2.10. An evolution process $\{S(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ in X is *pullback strongly bounded dissipative* if, for each $t \in \mathbb{R}$, there is a bounded subset $B(t)$ of X which pullback absorbs bounded subsets of X at time s for each $s \leqslant t$; that is, given a bounded subset B of X and $s \leqslant t$, there exists $\tau_0(s, B)$ such that $S(s, \tau)B \subset B(t)$ for all $\tau \leqslant \tau_0(s, B)$.

Now we can present the result which guarantees the existence of pullback attractors for nonautonomous problems; see [Caraballo et al. 2010].

Theorem 2.11. *If an evolution process $\{S(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ in the metric space X is pullback strongly bounded dissipative and pullback asymptotically compact, then $\{S(t, \tau) : t \geqslant \tau \in \mathbb{R}\}$ has a pullback attractor $\{\mathcal{A}(t) : t \in \mathbb{R}\}$ with the property that $\bigcup_{\tau \leqslant t} \mathcal{A}(\tau)$ is bounded for each $t \in \mathbb{R}$.*

The next result gives sufficient conditions for pullback asymptotic compactness, and its proof can be found in [Caraballo et al. 2010].

Theorem 2.12. *Let $\{S(t, s) : t \geqslant s \in \mathbb{R}\}$ be a pullback strongly bounded evolution process such that $S(t, s) = L(t, s) + U(t, s)$, where there exists a nonincreasing function $k : [0, +\infty) \times [0, +\infty) \to \mathbb{R}$, with $k(\sigma, r) \to 0$ when $\sigma \to \infty$, and for all $s \leqslant t$ and $x \in X$ with $\|x\| \leqslant r$,*

$$\|L(t, s)x\| \leqslant k(t - s, r),$$

and $U(t, s)$ is compact. Then, the family of evolution process $\{S(t, s) : t \geqslant s \in \mathbb{R}\}$ is pullback asymptotically compact.

3. Existence results

In this section we study the existence of global solutions for (1-8). For this, we consider the linear problem associated with (1-1)–(1-3),

$$\begin{cases} w_t = A_{(\kappa)}(t)w, \ t > \tau, \\ w(\tau) = w_0, \ \tau \in \mathbb{R}, \end{cases}$$

where w and w_0 are defined in (1-9) and the linear unbounded operator $A_{(\kappa)}$ is defined by (1-10) and (1-11).

We use the term singularly nonautonomous to express the fact that the unbounded operator $A_{(\kappa)}(t)$ is time-dependent and generates a semigroup that satisfies an estimate as in (2-1).

It is not difficult to see that $0 \in \rho(A_{(\kappa)}(t))$ for any $t \in \mathbb{R}$. Moreover, the operator $A_{(\kappa)}^{-1}(t) : D(A_{(\kappa)}^{-1}(t)) \subset Y \to Y$ is defined by

$$D(A_{(\kappa)}^{-1}(t)) = L^2(\Omega) \times H^{-2}(\Omega) \times H^{-2}(\Omega),$$

where $H^{-2}(\Omega)$ denotes the dual $X^{-\frac{1}{2}}$ of $X^{\frac{1}{2}}$ and

$$A_{(\kappa)}^{-1}(t) \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} = \begin{bmatrix} \frac{a^2}{\kappa(t)}\Lambda^{-\frac{1}{2}} & -\Lambda^{-1} & -\frac{a}{\kappa(t)}\Lambda^{-1} \\ I & 0 & 0 \\ -\frac{a}{\kappa(t)}I & 0 & \frac{1}{\kappa(t)}\Lambda^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} u \\ v \\ \theta \end{bmatrix}$$
$$= \begin{bmatrix} \frac{a^2}{\kappa(t)}\Lambda^{-\frac{1}{2}}u - \Lambda^{-1}v - \frac{a}{\kappa(t)}\Lambda^{-1}\theta \\ u \\ -\frac{1}{\kappa(t)}au + \frac{1}{\kappa(t)}\Lambda^{-\frac{1}{2}}\theta \end{bmatrix}.$$

Proposition 3.1. *Denote by Y_{-1} the extrapolation space of $Y = X^{\frac{1}{2}} \times X \times X$ generated by operator $A_{(\kappa)}^{-1}(t)$. The following equality holds*:

$$Y_{-1} = X \times X^{-\frac{1}{2}} \times X^{-\frac{1}{2}}.$$

Proof. This proof follows the same ideas of the proof in [Bezerra et al. 2018, Proposition 3.1]. □

Remark. Following the same ideas from [Baroun et al. 2009] and [Lasiecka and Triggiani 1998], we conclude that for all t, there exists a positive constant M (independent of t), such that

$$\|(\lambda I + A_{(\kappa)}(t))^{-1}\|_{L(Y)} \leqslant \frac{M}{1+|\lambda|} \quad \text{for all } \lambda \in \mathbb{C} \text{ with } \operatorname{Re}\lambda \geqslant 0.$$

From this we can conclude that $A_{(\kappa)}(t)$ is uniformly sectorial (in Y).

Note that the operator $A_{(\kappa)}(t)$ can be extended to its closed Y_{-1}-realization (see [Amann 1995, p. 262]), which we will still denote by the same symbol so that $A_{(\kappa)}(t)$ considered in Y_{-1} is then the sectorial positive operator (see [Carvalho and Cholewa 2002]). Our next concern will be to obtain embedding of the spaces from the fractional powers scale $Y_{\alpha-1},\ \alpha \geqslant 0$, generated by $(A_{(\kappa)}(t), Y_{-1})$.

Theorem 3.2. *The operators $A_{(\kappa)}(t)$ are uniformly sectorial and the map $\mathbb{R} \ni t \mapsto A_{(\kappa)}(t) \in \mathcal{L}(Y, Y_{-1})$ is uniformly Hölder continuous. Then, there exists a process*

$$\{L(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$$

(*or simply $L(t,\tau)$*) *associated with the operator $A_{(\kappa)}(t)$, that is given by*

$$L(t,\tau)=e^{-(t-\tau)A_{(\kappa)}(\tau)}+\int_\tau^t L(t,s)[A_{(\kappa)}(\tau)-A_{(\kappa)}(s)]e^{-(s-\tau)A_{(\kappa)}(\tau)}\,ds \quad \textit{for all } t\geqslant\tau.$$

The linear evolution operator $\{L(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ satisfies the condition (2-2).

Proof. Following the same ideas from [Carvalho and Cholewa 2002] and [Lasiecka and Triggiani 1998], we can conclude that the operator $A_{(\kappa)}(t)$ is a sectorial positive operator in Y_{-1}. It is not difficult to see that it is also closed and densely defined. Note that for $[u\ \ v\ \ \theta]^T \in X^{\frac{1}{2}} \times X \times X$, and $t, s \in \mathbb{R}$, we can estimate the norm $\|[(A_{(\kappa)}(t) - A_{(\kappa)}(s))[u\ \ v\ \ \theta]^T\|_{X\times X^{-\frac{1}{2}}\times X^{-\frac{1}{2}}}$ using (1-4). In fact,

$$\left\|(A_{(\kappa)}(t)-A_{(\kappa)}(s))\begin{bmatrix}u\\ v\\ \theta\end{bmatrix}\right\|_{Y_{-1}}=|\kappa(t)-\kappa(s)|\,\|(-\Delta)\theta\|_{X^{-\frac{1}{2}}}\leqslant c|t-s|^\beta\left\|\begin{bmatrix}u\\ v\\ \theta\end{bmatrix}\right\|_{X^{\frac{1}{2}}\times X\times X}$$

for any $t, s \in \mathbb{R}$; hence the application $\mathbb{R} \ni t \mapsto A_{(\kappa)}(t) \in \mathcal{L}(Y)$ is uniformly Hölder continuous, and this argument shows that

$$\|A_{(\kappa)}(t) - A_{(\kappa)}(s)\|_{\mathcal{L}(Y,Y_{-1})} \leqslant c|t-s|^\beta.$$

Therefore, there exists a linear evolution process $\{L(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ associated with the operator $A_{(a)}(t)$, that is given by

$$L(t,\tau)=e^{-(t-\tau)A_{(\kappa)}(\tau)}+\int_\tau^t L(t,s)[A_{(\kappa)}(\tau)-A_{(\kappa)}(s)]e^{-(s-\tau)A_{(\kappa)}(\tau)}\,ds \quad \text{for all } t\geqslant\tau.$$

Furthermore, the process $\{L(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ satisfies the condition (2-2). □

The following result is a direct consequence of (1-6), see [Carbone et al. 2011, Lemma 2.4].

Lemma 3.3. *Let $f \in C^1(\mathbb{R})$ be a function such that the condition* (1-6) *holds. Then*

$$|f(s_1) - f(s_2)| \leqslant 2^{\rho-1}c|s_1 - s_2|(1 + |s_1|^{\rho-1} + |s_2|^{\rho-1}) \quad \textit{for all } s_1, s_2 \in \mathbb{R}.$$

Lemma 3.4 [Carbone et al. 2011]. *Assume that $1 < \rho < \frac{N+4}{N-4}$ and let $f \in C^1(\mathbb{R})$ be a function such that*

$$|f'(s)| \leqslant C(1+|s|^{\rho-1}) \quad \textit{for all } s \in \mathbb{R}.$$

Then there exists $s \in (0,1)$ such that the Nemytskii operator $f^e : X^{\frac{1}{2}} \to X^{-\frac{s}{2}}$ is Lipschitz continuous in bounded subsets of $X^{\frac{1}{2}}$ uniformly in $t \in \mathbb{R}$.

Remark. Since $L^{2N/(N-4)}(\Omega) \hookrightarrow L^2(\Omega)$, it follows from the proof of [Carbone et al. 2011, Lemma 2.5] that $f^e : X^{\frac{1}{2}} \to L^2(\Omega)$ is Lipschitz continuous in bounded subsets; that is,

$$\|f^e(u) - f^e(v)\|_{L^2(\Omega)} \leqslant \tilde{c}\, \|f^e(u) - f^e(v)\|_{L^{\frac{2N}{(N-4)\rho}}(\Omega)} \leqslant \tilde{\tilde{c}}\|u - v\|_{X^{1/2}},$$

with $\tilde{\tilde{c}} = \tilde{\tilde{c}}(\|u\|_{X^{1/2}}, \|v\|_{X^{1/2}})$. The scheme below describes this situation:

$$X^{\frac{1}{2}} \hookrightarrow H^2(\Omega) \hookrightarrow L^{\frac{2N}{N-4}}(\Omega) \xmapsto{f(u)\approx u^\rho} L^{\frac{2N}{(N-4)\rho}}(\Omega) \hookrightarrow L^2(\Omega),$$

where in the last inclusion we use that $\rho < \frac{N}{N-4}$.

Proposition 3.5. *The operator $A_{(\kappa)}(t)$ given in* (1-10) *is maximal accretive.*

Proof. This proof is analogous to the proof [Bezerra et al. 2018, Proposition 4.3], and so we omit it. □

Remark. Below we have a partial description of the fractional power spaces scale for $A_{(\kappa)}(t)$. For convenience we denote Y by Y_0, then

$$Y_0 \hookrightarrow Y_{\alpha-1} \hookrightarrow Y_{-1} \quad \text{for all } 0 < \alpha < 1,$$

where

$$Y_{\alpha-1} = [Y_{-1}, Y_0]_\alpha = X^{\frac{\alpha}{2}} \times X^{\frac{\alpha-1}{2}} \times X^{\frac{\alpha-1}{2}},$$

where $[\,\cdot\,,\cdot\,]_\alpha$ denotes the complex interpolation functor (see [Triebel 1978]). The first equality follows from Proposition 3.5 (since $0 \in \rho(A_{(\kappa)}(t))$) (see [Amann 1995, Example 4.7.3(b)]) and the second equality follows from [Carvalho and Cholewa 2002, Proposition 2].

Corollary 3.6. *If f is as in Lemma 3.4, then the function $F : Y \to Y_{\alpha-1}$ ($\alpha \in (0,1)$), given by* (1-12), *is Lipschitz continuous in bounded subsets of Y.*

Now, Theorem 2.3 guarantees local well posedness for the problem (1-8) in the energy space Y.

Corollary 3.7. *If f and F are as in Corollary 3.6, then given $r > 0$, there is a time $\tau = \tau(r) > 0$, such that for all $w_0 \in B_Y(0; r)$ there exists a unique solution $w : [t_0, t_0 + \tau] \to Y$ of the problem* (1-8) *starting in w_0. Moreover, such solutions are continuous with respect the initial data in $B_Y(0; r)$.*

Since τ can be chosen uniformly in bounded subsets of Y, the solutions which do not blow up in Y must exist globally. Alternatively, we obtain a uniform in time estimate of $\|(u(t), \partial_t u(t), \theta(t))\|_Y$; such an estimate is needed to justify global solvability of the problem (1-8) in Y.

The total energy of the system $\mathcal{E}(t)$ associated with the solution $(u(t), \partial_t u(t), \theta(t))$ of (1-1)–(1-3) in Y is defined by

$$\mathcal{E}(t) = \tfrac{1}{2}\|u(t)\|^2_{X^{1/2}} + \tfrac{1}{2}\|u_t(t)\|^2_X + \tfrac{1}{2}\|\theta(t)\|^2_X - \int_\Omega \int_0^u f(s)\,ds\,dx. \tag{3-1}$$

It is not difficult to see that the function $t \mapsto \mathcal{E}(t)$ is monotone decreasing along solutions. In fact, using (1-1), we can show that there exists a positive constant c such that

$$\mathcal{E}'(t) \leqslant 0.$$

We obtain (from (1-5)) that for each $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that if

$$\int_\Omega \int_0^{u(\cdot\,,t)} f(s)\,ds\,dx \leqslant \varepsilon \|u(\cdot\,,t)\|^2_X + C_\varepsilon, \tag{3-2}$$

the property $\mathcal{E}(t) \leqslant \mathcal{E}(\tau)$ offers an a priori estimate of the solution $(u(t), \partial_t u(t), \theta(t))$ in Y. In fact,

$$\frac{1}{2}\left\| \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} \right\|^2_Y \leqslant c\mathcal{E}(\tau) + c\varepsilon_0 \|u(\cdot\,,t)\|^2_X + C_{\varepsilon_0} \leqslant c\mathcal{E}(\tau) + c\varepsilon_0 \left\| \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} \right\|^2_Y + C_{\varepsilon_0},$$

and, if we choose $0 < \varepsilon_0 < \frac{1}{2c}$, we get boundedness as desired; that is,

$$\limsup_{t\to+\infty} \left\| \begin{bmatrix} u \\ v \\ \theta \end{bmatrix} \right\|_Y < +\infty.$$

With this, we ensure that there exists a global solution $w(t)$ for Cauchy problem (1-8) in Y and it defines an evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$, that is,

$$S(t,\tau)w_0 = w(t) \quad \text{for all } t \geqslant \tau \in \mathbb{R}.$$

According to [Carvalho and Nascimento 2009],

$$S(t,\tau)w_0 = L(t,\tau)w_0 + \int_\tau^t L(t,s)F(s, S(s,\tau)w_0)\,ds \quad \text{for all } t \geqslant \tau \in \mathbb{R}, \tag{3-3}$$

where $\{L(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ is the linear evolution process associated with the homogeneous problem (1-8).

In order to prove the existence of pullback attractors for (1-1)–(1-3) we use the modified energy method.

Theorem 3.8. *Let $\mathcal{L}$ be the energy functional associated to* (1-1)–(1-3) *given by*

$$\mathcal{L}(t) = M\mathcal{E}(t) + \delta_1 \int_\Omega u u_t \, dx - \delta_2 \int_\Omega u_t \Delta^{-1}\theta \, dx,$$

where $\mathcal{E}$ is defined in (3-1), *and $0 < \delta_1 < \delta_2 < 1$ and $M > 0$ are appropriate constants.*

(a) *There exist constants $M_1, M_2 > 0$ such that*

$$\mathcal{L}'(t) \leqslant -M_1\mathcal{E}(t) + M_2 \tag{3-4}$$

for any $t \geqslant 0$.

(b) *For $M > 0$ sufficiently large, there exist constants $\beta_1, \beta_2, \beta_3 > 0$ and $\beta_4 > 0$ such that*

$$\beta_3\mathcal{E}(t) - \beta_4 \leqslant \mathcal{L}(t) \leqslant \beta_1\mathcal{E}(t) + \beta_2 \tag{3-5}$$

for any $t \geqslant 0$.

Proof. See [Bezerra et al. 2018, Theorems 5.1 and 5.2]. □

Remark. For every $t \in \mathbb{R}$, from (3-2) we have

$$\begin{aligned}\mathcal{E}(t) &= \tfrac{1}{2}\|u(t)\|^2_{X^{1/2}} + \tfrac{1}{2}\|u_t(t)\|^2_X + \tfrac{1}{2}\|\theta(t)\|^2_X - \int_\Omega \int_0^u f(t,s)\, ds\, dx \\ &\geqslant \big(\tfrac{1}{2} - \tfrac{1}{2}\varepsilon C_0\big)\|u(t)\|^2_{X^{1/2}} + \tfrac{1}{2}\|u_t(t)\|^2_X + \tfrac{1}{2}\|\theta(t)\|^2_X - C_\varepsilon,\end{aligned}$$

where ε is such that $\varepsilon < 1/C_0$; that is

$$\|\Delta u(t)\|^2_X + \|u_t(t)\|^2_X + \|\theta(t)\|^2_X \leqslant C_1\mathcal{E}(t) + C'_\varepsilon,$$

where $C_1^{-1} = \min\big\{\big(\tfrac{1}{2} - \tfrac{1}{2}\varepsilon C_0\big), \tfrac{1}{2}\big\}$.

Corollary 3.9. *Under the same conditions as in Theorem 3.8, if $B \subset Y$ is a bounded set, and $(u, v, \theta) : [\tau, \tau + T] \to Y$, $T > 0$, is the solution of* (1-1)–(1-3) *starting in $(u_0, v_0, \theta_0) \in B$, there exist positive constants $\bar{\omega}$, $\gamma_1 = \gamma_1(B)$ and γ_2, such that*

$$\|\Delta u(t)\|^2_X + \|u_t(t)\|^2_X + \|\theta(t)\|^2_X \leqslant \gamma_1 e^{-\bar{\omega}(t-\tau)} + \gamma_2 \tag{3-6}$$

for any $t \in [\tau, \tau + T]$.

Proof. From (3-4) and (3-5), we obtain

$$\mathcal{L}'(t) \leqslant -\sigma_1\mathcal{L}(t) + \sigma_2,$$

where $\sigma_1 = M_1/\beta_1$ and $\sigma_2 = M_1\beta_2/\beta_1 + M_2$, and thus

$$\mathcal{L}(t) \leqslant \mathcal{L}(\tau)e^{-\sigma_1(t-\tau)} + \sigma_2 e^{-\sigma_1 t}\int_\tau^t e^{\sigma_1 s}\, ds \leqslant \mathcal{L}(\tau)e^{-\sigma_1(t-\tau)} + \frac{\sigma_2}{\sigma_1}.$$

Again, by (3-5) together with the remark on page 406, we conclude that

$$\|\Delta u(t)\|_X^2 + \|u_t(t)\|_X^2 + \|\theta(t)\|_X^2 \leqslant \gamma_1 e^{-\sigma_1(t-\tau)} + \gamma_2,$$

where $\gamma_1 = \gamma_1(\mathcal{L}(\tau)) > 0$ and $\gamma_2 > 0$. □

Theorem 3.10. *Under the same conditions as in Theorem 3.8, the problem* (1-1)–(1-3) *has a pullback attractor* $\{\mathcal{A}(t) : t \in \mathbb{R}\}$ *in* Y *and*

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \subset Y.$$

Proof. From estimate (3-6), it is easy to check that the evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ associated with (1-1)–(1-3) is pullback strongly bounded dissipative in Y.

Hence, applying the same ideas of the proofs of [Bezerra et al. 2018, Theorems 5.1 and 5.2], we conclude that the family of evolution process $\{S(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ is pullback asymptotically compact (see Theorem 2.12). In fact, from (3-3) we write

$$S(t,\tau)w_0 = L(t,\tau)w_0 + U(t,\tau)w_0,$$

where

$$U(t,\tau)w_0 := \int_\tau^t L(t,s)F(S(t,s)w_0)\,ds \tag{3-7}$$

for any initial condition $w_0 \in Y$.

With the same arguments used in [Bezerra et al. 2018, Theorem 5.1] with $f \equiv 0$ in (1-1) and with the functionals

$$\mathcal{E}(t) = \tfrac{1}{2}\|u(t)\|_{X^{1/2}}^2 + \tfrac{1}{2}\|u_t(t)\|_X^2 + \tfrac{1}{2}\|\theta(t)\|_X^2$$

and

$$\mathcal{L}(t) = M\mathcal{E}(t) + \delta_1\langle u, u_t\rangle_X - \delta_2\langle u_t, (\Delta^{-1}\theta)\rangle_X,$$

we get from (3-4) that there exists $c_1 > 0$ such that

$$\mathcal{L}'(t) \leqslant -c_1\mathcal{E}(t).$$

From arguments used in the proof of [Bezerra et al. 2018, Theorem 5.2] with $f \equiv 0$ in (1-1), by (3-5) we get $c_2, c_3 > 0$ such that

$$c_2\mathcal{E}(t) \leqslant \mathcal{L}(t) \leqslant c_3\mathcal{E}(t) \tag{3-8}$$

and hence

$$\mathcal{L}'(t) \leqslant -c_0\mathcal{L}(t)$$

for some $c_0 > 0$. From this, we obtain

$$\mathcal{L}(t) \leqslant \mathcal{L}(\tau)e^{-c_0(t-\tau)},$$

and thanks to (3-8) we get

$$\mathcal{E}(t) \leqslant \frac{c_3}{c_2}\mathcal{E}(\tau)e^{-c_0(t-\tau)},$$

for some $c_0 > 0$. This ensures that there exist constants $K, \alpha > 0$ such that

$$\|L(t,\tau)\|_{\mathcal{L}(Y)} \leqslant Ke^{-\alpha(t-\tau)} \quad \text{for all } t \geqslant \tau. \tag{3-9}$$

The family of evolution processes $\{U(t,\tau) : t \geqslant \tau \in \mathbb{R}\}$ is compact from Y into Y. In fact, the compactness of $U(t,\tau)$ follows easily from

$$X^{1/2} \xrightarrow{f^e} X^{-s/2} \hookrightarrow X^{-1/2},$$

being the last inclusion compact (since $s < 1$; see Lemma 3.4). Thanks to the assumptions on the nonlinearity of f, it follows that f^e is compact from $X^{\frac{1}{2}}$ into $X^{-\frac{1}{2}}$. Taking into account that F is given by (1-12), compactness of f^e implies that F is also compact from Y into Y_{-1}, and since $L(t,\tau)$ is a bounded linear operator from Y_{-1} to Y, the operator $U(t,\tau)$ is compact from Y into Y (see [Hale 1988, Theorem 4.6.1]).

Now, applying Theorem 2.11, we get that the problem (1-1)–(1-3) has a pullback attractor $\{\mathcal{A}(t) : t \in \mathbb{R}\}$ in Y and that $\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \subset Y$ is bounded. □

4. Regularity of the pullback attractors

In this section we investigate the regularity of the pullback attractors; in fact, we prove that $\bigcup_{t\in\mathbb{R}} \mathcal{A}(t)$ is a bounded subset of Y^1.

Theorem 4.1. *The pullback attractor* $\{\mathcal{A}(t) : t \in \mathbb{R}\}$ *for the problem* (1-1)–(1-3), *obtained in Theorem 3.8, lies in a more regular space than Y; in fact,*

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t)$$

is a bounded subset of Y^1.

Proof. The main idea is to use the argument of progressive increases of regularity, following Babin and Vishik [1992] (see also [Carvalho et al. 2013, Chapter 15]).

Let $\xi : \mathbb{R} \to Y$ be a global bounded solution of (1-1). Then, the set $\{\xi(t);\ t \in \mathbb{R}\}$ is a bounded subset of Y. First, observe that we already know that

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \text{ is bounded in } Y.$$

Hence, if $\xi(\cdot) = (u(\cdot), u_t(\cdot), \theta(\cdot)) : \mathbb{R} \to Y$ is such that $\xi(t) \in \mathcal{A}(t)$ for all $t \in \mathbb{R}$, then

$$\xi(t) = L(t,s)\xi(s) + \int_s^t L(t,\theta)F(\xi(\theta))\,d\theta,$$

and, using the decay of $L(t,s)$ in (3-9) and letting $s \to -\infty$ it follows that

$$\xi(t) = \int_{-\infty}^{t} L(t,\theta) F(\xi(\theta))\, d\theta. \tag{4-1}$$

Now fix $s \in \mathbb{R}$, set $(\mu_0, \mu_1, \vartheta_0) = \xi(s)$, and consider

$$\begin{bmatrix} \mu(t) \\ \mu_t(t) \\ \vartheta(t) \end{bmatrix} = U(t,s) \begin{bmatrix} \mu_0 \\ \mu_1 \\ \vartheta_0 \end{bmatrix} = \int_s^t L(t,\theta) F\left(S(\theta,s) \begin{bmatrix} \mu_0 \\ \mu_1 \\ \vartheta_0 \end{bmatrix} \right) d\theta,$$

where $U(\,\cdot\,,\cdot\,)$ is defined in (3-7). Note that $(\mu(\cdot), \vartheta(\cdot))$ solves the system

$$\begin{cases} \mu_{tt} + \Delta^2\mu + a\Delta\vartheta = f(\mu(t,s;\mu_0)), & t > s,\ x \in \Omega, \\ \vartheta_t - \kappa(t)\Delta\vartheta - a\Delta\mu_t = 0, & t > s,\ x \in \Omega, \end{cases} \tag{4-2}$$

with

$$\mu(s,x) = \mu_t(s,x) = 0 \quad \text{and} \quad \vartheta(s,x) = 0,\ x \in \Omega. \tag{4-3}$$

This happens inside the pullback attractor $\mathcal{A}(\cdot)$.

To estimate the solution of (4-2)–(4-3) for $(\mu_0, \mu_1, \vartheta_0)$ in a bounded subset B of Y, we again consider the energy functional

$$\begin{aligned} \mathcal{L}_\delta(t) = \tfrac{1}{2}M\|\mu(t)\|^2_{X^{1/2}} + \tfrac{1}{2}M\|\mu_t(t)\|^2_X + \tfrac{1}{2}M\|\vartheta(t)\|^2_X \\ + \langle \mu(t), \mu_t(t)\rangle_X - \delta_2\langle \mu_t(t), \Delta^{-1}\vartheta(t)\rangle_X, \end{aligned}$$

to obtain (we omitted t on the right side in order to simplify the notation)

$$\begin{aligned} \mathcal{L}'_\delta(t) = M\langle \mu, \mu_t\rangle_X + M\langle \mu_t, \mu_{tt}\rangle_X + M\langle \vartheta, \vartheta_t\rangle_X + \|\mu_t\|^2_X \\ + \langle \mu, \mu_{tt}\rangle_X - \delta_2\langle \mu_{tt}, \Delta^{-1}\vartheta\rangle_X - \delta_2\langle \mu_t, \Delta^{-1}\vartheta_t\rangle_X, \end{aligned}$$

and by (4-2) we get

$$\begin{aligned} \mathcal{L}'_\delta(t) = M\langle \mu, \mu_t\rangle_X - M\langle \mu_t, \Delta^2\mu\rangle_X + M\langle \mu_t, f(\mu)\rangle_X - \kappa(t)M\|\vartheta\|^2_{H^1_0(\Omega)} \\ + (1 - a\delta_2)\|\mu_t\|^2_X - \|\mu\|^2_{X^{1/2}} + \langle \mu, f(\mu)\rangle_X + (\delta_2 - a)\langle \Delta\mu, \vartheta\rangle_X \\ + a\delta_2\|\vartheta\|^2_X - \delta_2\langle f(\mu), \Delta^{-1}\vartheta\rangle_X - \delta_2\kappa(t)\langle \mu_t, \vartheta\rangle_X. \end{aligned}$$

From Poincaré and Young inequalities

$$\begin{aligned} \mathcal{L}'_\delta(t) \leqslant \big(\tfrac{1}{2}\nu_0 M + \tfrac{1}{2}\nu_1 C_a\big)\|\mu\|^2_{X^{1/2}} + \Big(\frac{M}{2\nu_0} + C_a + \tfrac{1}{2}\nu_2\delta_2\kappa_1 + \tfrac{1}{2}M\Big)\|\mu_t\|^2_X \\ + \Big(+\frac{C_a}{2\nu_1} + a\delta_2\Big)\|\vartheta\|^2_X - \kappa_0\lambda_1 M\|\vartheta\|^2_{H^1_0(\Omega)} + \tfrac{1}{2}\delta_2 \int_\Omega |\Delta^{-1}\vartheta|^2\, dx \\ + \big(\tfrac{1}{2}M + \tfrac{1}{2}\delta_2\big)\int_\Omega |f(\mu)|^2\, dx + \int_\Omega f(\mu)\mu\, dx, \end{aligned}$$

where $C_a = 1 - a\delta_2$ and $C_a = \delta_2 - a$.

To deal with the integral terms, just notice that from dissipativeness condition (1-5), for each $\nu > 0$ there exists $C_\nu > 0$ such that

$$\int_\Omega f(\mu)\mu\,dx \leqslant \nu\|\mu\|_X^2 + C_\nu \leqslant m_0\nu\|\mu\|_{X^{1/2}}^2 + C_\nu,$$

where $m_0 > 0$ is the embedding constant for $\|\cdot\|_X \leqslant m_0\|\cdot\|_{X^{1/2}}$.

From (1-6), there exists $C > 0$ such that

$$\int_\Omega |f(\mu)|^2\,dx \leqslant C\|\mu\|_X^2 + C\|\mu\|_{L^{2\rho}(\Omega)}^{2\rho}.$$

Since the condition $1 \leqslant \rho < \frac{N}{N-4}$ implies $X^{\frac{1}{2}} \hookrightarrow L^{2\rho}(\Omega)$, we get

$$\int_\Omega |f(\mu)|^2\,dx \leqslant C\|\mu\|_X^2 + \overline{C} \leqslant \overline{C}_1\|\mu\|_{X^{1/2}}^2 + \overline{C}_2,$$

whenever $\|\mu\|_{X^{1/2}} \leqslant r$ (see [Carbone et al. 2011] and [Carvalho et al. 2009]).

From this it follows that

$$\bigcup_{s\leqslant\tau\leqslant t} U(\tau,s)B \text{ is a bounded subset of } Y. \tag{4-4}$$

Hence $(\varpi,\zeta) = (\mu_t,\vartheta_t)$ solves the system

$$\begin{cases} \varpi_{tt} + \Delta^2\varpi + a\Delta\zeta = f'(\mu(t,s;\mu_0))\varpi(t,s;\mu_0), & t > s,\ x\in\Omega,\\ \zeta_t - \kappa(t)\Delta\zeta - \kappa'(t)\Delta\vartheta - a\Delta\varpi_t = 0, & t > s,\ x\in\Omega, \end{cases} \tag{4-5}$$

with $\varpi(s) = 0$, $\varpi_t(s) = f(\mu_0)$, and $\zeta(s) = 0$.

Finally, now we would like to estimate (ϖ,ϖ_t,ζ) in Y, but solutions are not regular enough to allow this directly. Instead we work "towards" Y by progressive increases of regularity. For $\alpha > 0$, we define the fractional power spaces $X^\alpha = D(\Lambda^\alpha)$ with the graph norm, and let $X^{-\alpha} = (X^\alpha)'$; see (1-7).

For

$$(\varpi,\varpi_t,\zeta) \in X^{\frac{1-\alpha}{2}} \times X^{-\frac{\alpha}{2}} \times X^{-\frac{\alpha}{2}},$$

we define

$$\begin{aligned} \mathcal{L}_\alpha(t) = \tfrac{1}{2}M\big(\|\varpi(t)\|^2_{X^{\frac{1-\alpha}{2}}} + \|\varpi_t(t)\|^2_{X^{-\frac{\alpha}{2}}} + \|\zeta(t)\|^2_{X^{-\frac{\alpha}{2}}}\big)&\\ + \delta_1\langle\varpi,\varpi_t\rangle_{X^{-\frac{\alpha}{2}}} - \delta_2\langle\varpi_t,\Delta^{-1}\zeta\rangle_{X^{-\frac{\alpha}{2}}}&. \end{aligned} \tag{4-6}$$

$$\begin{aligned} \mathcal{L}'_\alpha(t) = M\langle\varpi_t,\varpi\rangle_{X^{\frac{1-\alpha}{2}}} + M\langle\varpi_{tt},\varpi_t\rangle_{X^{-\frac{\alpha}{2}}} + M\langle\zeta_t,\zeta\rangle_{X^{-\frac{\alpha}{2}}} + \delta_1\langle\varpi_t,\varpi_t\rangle_{X^{-\frac{\alpha}{2}}}&\\ + \delta_1\langle\varpi,\varpi_{tt}\rangle_{X^{-\frac{\alpha}{2}}} - \delta_2\langle\varpi_{tt},\Delta^{-1}\zeta\rangle_{X^{-\frac{\alpha}{2}}}&\\ - \delta_2\langle\varpi_t,\Delta^{-1}\zeta_t\rangle_{X^{-\frac{\alpha}{2}}}&. \end{aligned}$$

Note that from (1-2)–(1-7) (that is, $\Lambda^{\frac{1}{2}} = -\Delta$), (4-5) and (4-6) we have

$$\begin{aligned}\mathcal{L}'_\alpha(t) = {} & M\langle\varpi_t, \varpi\rangle_{X^{\frac{1-\alpha}{2}}} - M\langle\varpi_t, \varpi\rangle_{X^{\frac{1-\alpha}{2}}} - Ma\langle\varpi_t, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & + M\langle\varpi_t, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} + M\kappa'(t)\langle\zeta, \Delta\vartheta\rangle_{X^{-\frac{\alpha}{2}}} + M\kappa(t)\langle\zeta, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & + Ma\langle\zeta, \Delta\varpi_t\rangle_{X^{-\frac{\alpha}{2}}} + \delta_1\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} - \delta_1\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} - a\delta_1\langle\varpi, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & + \delta_1\langle\varpi, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} + \delta_2\langle\zeta, \Lambda^{\frac{1}{2}}\varpi\rangle_{X^{-\frac{\alpha}{2}}} + a\delta_2\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} \\ & - \delta_2\langle\Lambda^{-\frac{1}{2}}\zeta, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} - \delta_2\kappa(t)\langle\varpi_t, \zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & - \delta_2\kappa'(t)\langle\varpi_t, \vartheta\rangle_{X^{-\frac{\alpha}{2}}} - a\delta_2\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}};\end{aligned}$$

in other words,

$$\begin{aligned}\mathcal{L}'_\alpha(t) = {} & M\langle\varpi_t, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} + M\kappa'(t)\langle\zeta, \Delta\vartheta\rangle_{X^{-\frac{\alpha}{2}}} + M\kappa(t)\langle\zeta, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & + (\delta_1 - a\delta_2)\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} - \delta_1\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} - a\delta_1\langle\varpi, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ & + \delta_1\langle\varpi, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} + \delta_2\langle\zeta, \Lambda^{\frac{1}{2}}\varpi\rangle_{X^{-\frac{\alpha}{2}}} \\ & + a\delta_2\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} - \delta_2\langle\Lambda^{-\frac{1}{2}}\zeta, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} \\ & - \delta_2\kappa(t)\langle\varpi_t, \zeta\rangle_{X^{-\frac{\alpha}{2}}} - \delta_2\kappa'(t)\langle\varpi_t, \vartheta\rangle_{X^{-\frac{\alpha}{2}}}.\end{aligned} \tag{4-7}$$

Next, we collect estimates of the terms that appear on the right side of (4-7). First, we deal with the three terms in which the nonlinearity of f' appears explicitly.

Let

$$\alpha_1 := \frac{(\rho-1)(N-4)}{4}.$$

Note that since $\rho < \frac{N}{N-4}$, we obtain $\alpha_1 < 1$. We observe that

$$\langle\varpi_t, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} \leqslant \|\varpi_t\|_{X^{-\frac{\alpha}{2}}}\|f'(\mu)\varpi\|_{X^{-\frac{\alpha}{2}}},$$

and using the embedding $X^{\frac{\alpha}{2}} = H^{2\alpha}(\Omega) \hookrightarrow L^p(\Omega)$ (or equivalently $L^{\frac{p}{p-1}}(\Omega) \hookrightarrow X^{-\frac{\alpha}{2}}$) for any $1 < p \leqslant \frac{2N}{N-4\alpha}$ $(0 < \alpha \leqslant \alpha_1)$ and (1-6), we have for some $c_4 > 0$,

$$\begin{aligned}\|f'(\mu)\varpi\|_{X^{-\frac{\alpha}{2}}} & \leqslant c_4\|f'(\mu)\varpi\|_{L^{\frac{2N}{N+4\alpha}}(\Omega)} \\ & \leqslant c_4 C\|\varpi(1+|\mu|^{\rho-1})\|_{L^{\frac{2N}{N+4\alpha}}(\Omega)} \\ & \leqslant c_4 C\|\varpi\|_X\|1+|\mu|^{\rho-1}\|_{L^{\frac{N}{2\alpha}}(\Omega)}\end{aligned} \tag{4-8}$$

and so

$$\|f'(\mu)\varpi\|^2_{X^{-\frac{\alpha}{2}}} \leqslant c_4^2 C^2\|\varpi\|^2_X\|1+|\mu|^{\rho-1}\|^2_{L^{\frac{N}{2\alpha}}(\Omega)}$$

and from (4-4) μ remains in a bounded subset of $X^{\frac{1}{2}}$ and $X^{\frac{1}{2}} \hookrightarrow L^{\frac{(\rho-1)N}{2\alpha}}(\Omega)$ for any $1<\rho<\frac{N-4+4\alpha}{N-4}$. This implies

$$\begin{aligned}\int_\Omega (1+|\mu|^{\rho-1})^{\frac{N}{2\alpha}}\,dx &\leqslant |\Omega|+\|\mu\|_{L^{\frac{(\rho-1)N}{2\alpha}}(\Omega)}^{\frac{(\rho-1)N-2\alpha}{N(\rho-1)}}\\ &\leqslant |\Omega|+c_5\|\mu\|_{X^{1/2}}^{\frac{(\rho-1)N-2\alpha}{N(\rho-1)}}\\ &\leqslant c_5,\end{aligned}$$

for some $c_5>0$. From this, there exists a positive constant $C_{f,1}>0$ such that

$$\|f'(\mu)\varpi\|^2_{X^{-\frac{\alpha}{2}}}\leqslant C_{f,1}. \tag{4-9}$$

With this we have

$$\begin{aligned}M\langle\varpi_t, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} &\leqslant \frac{\varepsilon_0}{2}\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}}+\frac{M^2}{2\varepsilon_0}\|f'(\mu)\varpi\|^2_{X^{-\frac{\alpha}{2}}}\\ &\leqslant \frac{\varepsilon_0}{2}\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}}+\frac{C_{f,1}M^2}{2\varepsilon_0}\end{aligned} \tag{4-10}$$

for some $\varepsilon_0>0$.

Again, from (4-9) we obtain

$$\begin{aligned}\delta_1\langle\varpi, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} &\leqslant \tfrac{1}{2}\big(\|\varpi\|^2_{X^{-\frac{\alpha}{2}}}+\delta_1^2\|f'(\mu)\varpi\|^2_{X^{-\frac{\alpha}{2}}}\big)\\ &\leqslant \frac{\varepsilon_1}{2}\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}}+\frac{C_{f,1}\delta_1^2}{2\varepsilon_1}\end{aligned} \tag{4-11}$$

for all $\varepsilon_1>0$.

We have the embedding $X^\alpha=H^{4\alpha}(\Omega)\hookrightarrow L^p(\Omega)$ (or equivalently $L^{\frac{p}{p-1}}(\Omega)\hookrightarrow X^{-\alpha}$) for any $1<p\leqslant\frac{2N}{N-4(1+\alpha)}$. From this and using (1-6), it follows that

$$\begin{aligned}\|f'(\mu)\varpi\|_{X^{-\frac{1+\alpha}{2}}} &\leqslant c_6\|f'(\mu)\varpi\|_{L^{\frac{2N}{N+4(1+\alpha)}}(\Omega)}\\ &\leqslant Cc_6\|\varpi(1+|\mu|^{\rho-1})\|_{L^{\frac{2N}{N+4(1+\alpha)}}(\Omega)}\\ &\leqslant Cc_6\|\varpi\|_X\|1+|\mu|^{\rho-1}\|_{L^{\frac{N}{2(1+\alpha)}}(\Omega)},\end{aligned} \tag{4-12}$$

and so

$$\|f'(\mu)\varpi\|^2_{X^{-\frac{1+\alpha}{2}}}\leqslant C^2c_6^2\|\varpi\|_X^2\|1+|\mu|^{\rho-1}\|^2_{L^{\frac{N}{2(1+\alpha)}}(\Omega)},$$

where

$$\int_\Omega(1+|\mu|^{\rho-1})^{\frac{N}{2(1+\alpha)}}\,dx\leqslant|\Omega|+\|\mu\|_{L^{\frac{(\rho-1)N}{2(1+\alpha)}}(\Omega)}^{\frac{2(1+\alpha)}{N(\rho-1)}}\leqslant|\Omega|+c_7\|\mu\|_{X^{1/2}}^{\frac{2(1+\alpha)}{N(\rho-1)}}. \tag{4-13}$$

In the last estimate we used the embedding

$$X^{\frac{1}{2}}\hookrightarrow L^{\frac{(\rho-1)N}{2(1+\alpha)}}(\Omega),\qquad 1<\rho<\frac{N+4\alpha}{N-4}.$$

From this, there exists a positive constant c_8 such that

$$\|1+|\mu|^{\rho-1}\|^2_{L^{\frac{N}{2(1+\alpha)}}(\Omega)} \leqslant c_8.$$

From (4-12) and (4-13), we have

$$\begin{aligned} -\delta_2\langle \Lambda^{-\frac{1}{2}}\zeta, f'(\mu)\varpi\rangle_{X^{-\frac{\alpha}{2}}} &= -\delta_2\langle \Lambda^{-\frac{1}{2}-\frac{\alpha}{2}}\zeta, \Lambda^{-\frac{\alpha}{2}} f'(\mu)\varpi\rangle_X \\ &\leqslant \delta_2\|\zeta\|_{X^{-\frac{\alpha}{2}}}\|f'(\mu)\varpi\|_{X^{-\frac{1+\alpha}{2}}} \\ &\leqslant \tfrac{1}{2}\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} + \tfrac{1}{2}\delta_2^2 C_{f,2} \end{aligned} \tag{4-14}$$

for some $C_{f,2} > 0$.

Finally, we consider the last term:

$$-\delta_2\kappa(t)\langle\varpi_t, \zeta\rangle_{X^{-\frac{\alpha}{2}}} \leqslant \delta_2\kappa_1\|\varpi_t\|_{X^{-\frac{\alpha}{2}}}\|\zeta\|_{X^{-\frac{\alpha}{2}}} \leqslant \frac{\delta_2\kappa_1}{2}\big(\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} + \|\zeta\|^2_{X^{-\frac{\alpha}{2}}}\big).$$

Since ϑ remains in a bounded subset of X (see (4-4)), for $\frac{1}{2} \leqslant \alpha < 1$ we have the embedding $L^2 = X^0 \hookrightarrow X^{\frac{1}{4}-\frac{\alpha}{2}}$ $\big(\frac{1}{4} - \frac{\alpha}{2} \leqslant 0\big)$ and

$$\begin{aligned} M\kappa'(t)\langle\zeta, \Delta\vartheta\rangle_{X^{-\frac{\alpha}{2}}} &= -M\kappa'(t)\langle\zeta, \Lambda^{\frac{1}{2}}\vartheta\rangle_{X^{-\frac{\alpha}{2}}} \\ &= -M\kappa'(t)\langle\Lambda^{\frac{1}{4}-\frac{\alpha}{2}}\zeta, \Lambda^{\frac{1}{4}-\frac{\alpha}{2}}\vartheta\rangle_X \\ &\leqslant M\kappa_1\|\Lambda^{\frac{1}{4}-\frac{\alpha}{2}}\zeta\|_X\|\Lambda^{\frac{1}{4}-\frac{\alpha}{2}}\vartheta\|_X \\ &= M\kappa_1\|\zeta\|_{X^{\frac{1}{4}-\frac{\alpha}{2}}}\|\vartheta\|_{X^{\frac{1}{4}-\frac{\alpha}{2}}} \\ &\leqslant \tfrac{1}{2}M\kappa_1 c(\|\zeta\|^2_X + \|\vartheta\|^2_X) \leqslant C_3 \end{aligned}$$

for some $c > 0$ and $C_3 > 0$.

It is not difficult to see that

$$\langle\zeta, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} = -\|\zeta\|^2_{X^{\frac{1-2\alpha}{4}}}.$$

From this we conclude that

$$M\kappa(t)\langle\zeta, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} \leqslant -M\kappa_0\|\zeta\|^2_{X^{\frac{1-2\alpha}{4}}} \leqslant -Mc_2\kappa_0\|\zeta\|^2_{X^{-\frac{\alpha}{2}}}.$$

Using Cauchy and Young inequalities we obtain

$$\begin{aligned} -a\delta_1\langle\varpi, \Delta\zeta\rangle_{X^{-\frac{\alpha}{2}}} &= a\delta_1\langle\varpi, \Lambda^{\frac{1}{2}}\zeta\rangle_{X^{-\frac{\alpha}{2}}} = a\delta_1\langle\Lambda^{\frac{1}{2}}\varpi, \zeta\rangle_{X^{-\frac{\alpha}{2}}} \\ &\leqslant a\delta_1\|\varpi\|_{X^{\frac{1-\alpha}{2}}}\|\zeta\|_{X^{-\frac{\alpha}{2}}} \leqslant \tfrac{1}{2}a\delta_1\varepsilon_2\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} + \frac{a\delta_1}{2\varepsilon_2}\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} \end{aligned}$$

for all $\varepsilon_2 > 0$, and

$$\delta_2\langle\zeta, \Lambda^{\frac{1}{2}}\varpi\rangle_{X^{-\frac{\alpha}{2}}} \leqslant \delta_2\|\varpi\|_{X^{\frac{1-\alpha}{2}}}\|\zeta\|_{X^{-\frac{\alpha}{2}}} \leqslant \tfrac{1}{2}\delta_2\varepsilon_3\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} + \frac{\delta_2}{2\varepsilon_3}\|\zeta\|^2_{X^{-\frac{\alpha}{2}}}$$

for all $\varepsilon_3 > 0$.

Finally, from $X \hookrightarrow X^{-\frac{\alpha}{2}}$ we obtain

$$
\begin{aligned}
-\delta_2\kappa'(t)\langle\varpi_t,\vartheta\rangle_{X^{-\frac{\alpha}{2}}} &\leqslant \delta_2\kappa_0\|\varpi_t\|_{X^{-\frac{\alpha}{2}}}\|\vartheta\|_{X^{-\frac{\alpha}{2}}} \\
&\leqslant \tfrac{1}{2}\delta_2\kappa_0\big(\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}}+\|\vartheta\|^2_{X^{-\frac{\alpha}{2}}}\big) \\
&\leqslant \tfrac{1}{2}\delta_2\kappa_0\big(\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}}+c\|\vartheta\|^2_{X}\big) \\
&\leqslant \tfrac{1}{2}\delta_2\kappa_0\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}}+c
\end{aligned}
\tag{4-15}
$$

for some $c > 0$.

Now combining (4-7) with (4-10), (4-11) and (4-14)–(4-15) we conclude that

$$
\begin{aligned}
\mathcal{L}'_\alpha(t) \leqslant{}& -\tfrac{1}{2}\delta_1\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} - \big(\tfrac{1}{2}\delta_1-\tfrac{1}{2}\varepsilon_1-\tfrac{1}{2}a\delta_1\varepsilon_2-\tfrac{1}{2}\delta_2\varepsilon_3\big)\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} \\
&-\big(a\delta_2-\tfrac{1}{2}\varepsilon_0-\delta_1-\tfrac{1}{2}\delta_2\kappa_1-\tfrac{1}{2}\delta_2\kappa_0\big)\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} \\
&-\Big(Mc_2\kappa_0-\tfrac{1}{2}-\frac{a\delta_1}{2\varepsilon_2}-a\delta_2-\frac{\delta_2}{2\varepsilon_3}-\tfrac{1}{2}\delta_2\kappa_1\Big)\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} \\
&+\frac{C_{f,1}M^2}{2\varepsilon_0}+\frac{C_{f,1}\delta_1^2}{2\varepsilon_1}+\tfrac{1}{2}\delta_2^2C_{f,2}+C_3+c.
\end{aligned}
$$

In other words,

$$
\begin{aligned}
\mathcal{L}'_\alpha(t) \leqslant{}& -\tfrac{1}{2}\delta_1\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} - \big(\tfrac{1}{2}\delta_1-\tfrac{1}{2}\varepsilon_1-\tfrac{1}{2}a\delta_1\varepsilon_2-\tfrac{1}{2}\delta_2\varepsilon_3\big)\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} \\
&-\big(a\delta_2-\tfrac{1}{2}\varepsilon_0-\delta_1-\delta_2\kappa_1\big)\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} \\
&-\Big(Mc_2\kappa_0-\tfrac{1}{2}-\frac{a\delta_1}{2\varepsilon_2}-a\delta_2-\frac{\delta_2}{2\varepsilon_3}-\tfrac{1}{2}\delta_2\kappa_1\Big)\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} \\
&+\frac{C_{f,1}M^2}{2\varepsilon_0}+\frac{C_{f,1}\delta_1^2}{2\varepsilon_1}+\tfrac{1}{2}\delta_2^2C_{f,2}+C_3+c.
\end{aligned}
$$

Now, it is enough to choose $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ and $\varepsilon_3 > 0$, respectively, such that

$$
\varepsilon_1=\frac{\delta_1}{3},\quad \varepsilon_2=\frac{1}{3a},\quad \text{and}\quad \varepsilon_3=\frac{\delta_1}{3\delta_2},
$$

and so

$$
\begin{aligned}
\mathcal{L}'_\alpha(t) \leqslant{}& -\tfrac{1}{2}\delta_1\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} - \big(a\delta_2-\tfrac{1}{2}\varepsilon_0-\delta_1-\delta_2\kappa_1\big)\|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} \\
&-\Big(Mc_2\kappa_0-\tfrac{1}{2}-\frac{a\delta_1}{2\varepsilon_2}-a\delta_2-\frac{\delta_2}{2\varepsilon_3}-\tfrac{1}{2}\delta_2\kappa_1\Big)\|\zeta\|^2_{X^{-\frac{\alpha}{2}}} \\
&+\frac{C_{f,1}M^2}{2\varepsilon_0}+\frac{C_{f,1}\delta_1^2}{2\varepsilon_1}+\tfrac{1}{2}\delta_2^2C_{f,2}+C_3+c.
\end{aligned}
$$

Since $\delta_1 < \delta_2$, if we assume that $a < 1+\kappa_1$, then choosing $\varepsilon_0 > 0$ such that

$$
\varepsilon_0 < 2(a-\kappa_1)\delta_2-2\delta_1 < 2(\delta_2-\delta_1),
$$

we conclude that

$$
a\delta_2-\tfrac{1}{2}\varepsilon_0-\delta_1-\delta_2\kappa_1 > 0.
$$

Now, it is enough to choose $M > 0$ sufficiently large such that

$$Mc_2\kappa_0 - \tfrac{1}{2} - \frac{a\delta_1}{2\varepsilon_2} - a\delta_2 - \frac{\delta_2}{2\varepsilon_3} - \tfrac{1}{2}\delta_2\kappa_1 > 0,$$

and so there exist $\ell_1 > 0$ and $\ell_2 > 0$ such that

$$\mathcal{L}'_\alpha(t) \leqslant -\ell_1(\|\varpi\|^2_{X^{\frac{1-\alpha}{2}}} + \|\varpi_t\|^2_{X^{-\frac{\alpha}{2}}} + \|\zeta\|^2_{X^{-\frac{\alpha}{2}}}) + \ell_2.$$

From this, (4-1), and the fact $\mathcal{A}(t) = \{\xi(t); \xi(t) \text{ is a global bounded solution}\}$ we obtain

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \text{ is bounded in } X^{\frac{2-\alpha_1}{2}} \times X^{\frac{1-\alpha_1}{2}} \times X^{\frac{1-\alpha_1}{2}}. \tag{4-16}$$

Using (4-16) and restarting from (4-8) and (4-12) with $\alpha_2 = (1+\rho)\alpha_1 - \rho$ it follows that

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \text{ is bounded in } X^{\frac{2-\alpha_2}{2}} \times X^{\frac{1-\alpha_2}{2}} \times X^{\frac{1-\alpha_2}{2}}.$$

Iterating this procedure a finite number of times, we can now show that

$$\bigcup_{t\in\mathbb{R}} \mathcal{A}(t) \text{ is bounded in } X^1 \times X^{\frac{1}{2}} \times X^{\frac{1}{2}},$$

which implies

$$\sup_{\xi\in\mathcal{A}} \sup_{t\in\mathbb{R}} \{\|\xi(t)\|_Y, \|\xi(t)\|_{Y^1}, \|\xi_t(t)\|_Y\} < \infty,$$

where $\mathcal{A}$ is the set of global bounded solutions for (1-8). □

5. Upper semicontinuity of the pullback attractors

From the results obtained in the previous section, we can prove a result on upper semicontinuity of the pullback attractors with respect to the functional parameter κ. Let $\{\kappa_\varepsilon : \varepsilon \in [0,1]\}$ be the family of real valued functions of one real variable satisfying (1-4), and denote by $S_{(\kappa_\varepsilon)}(\cdot,\cdot)$ and $\{\mathcal{A}_{(\kappa_\varepsilon)}(t) : t \in \mathbb{R}\}$, respectively, the evolution process and its pullback attractor associated with problem (1-1)–(1-3).

Moreover, we will assume that

$$\|\kappa_\varepsilon - \kappa_0\|_{L^\infty(\mathbb{R})} \to 0 \quad \text{as } \varepsilon \to 0^+.$$

Now we are able to present the main result of this section.

Theorem 5.1. *For each $a > 0$ and $\varepsilon \in [0,1]$, let $w^{(\varepsilon)}(\cdot) = S_{(\kappa_\varepsilon)}(\cdot,\tau)w_0$ be the solution of* (1-8) *in Y. Then, for each $T > 0$, $w^{(\varepsilon)}$ converges to $w^{(0)}$ in $C([0,T];Y)$ as $\varepsilon \to 0^+$. Moreover, the family of pullback attractors $\{\mathcal{A}_{(\kappa_\varepsilon)}(t) : t \in \mathbb{R}\}$ is upper semicontinuous in $\varepsilon = 0$.*

Proof. For each $w_0 \in Y$, consider $w^{(\varepsilon)} = S_{(\kappa_\varepsilon)}(t,\tau)w_0$ and $w^{(0)} = S_{(\kappa_0)}(t,\tau)w_0$. Let $w = w^{(\varepsilon)} - w^{(0)}$, with $w^{(\varepsilon)} = (u^{(\varepsilon)}, {u^{(\varepsilon)}}_t, \theta^{(\varepsilon)})$ and $w^{(0)} = (u^{(0)}, {u^{(0)}}_t, \theta^{(0)})$ ($u = u^{(\varepsilon)} - u^{(0)}$ and $\theta = \theta^{(\varepsilon)} - \theta^{(0)}$). Then, for all $t > \tau$ and $x \in \Omega$,

$$\begin{cases} u_{tt} + \Delta^2 u + a\Delta\theta = f(u^{(\varepsilon)}) - f(u^{(0)}), \\ \theta_t - \kappa_\varepsilon(t)\Delta\theta^{(\varepsilon)} + \kappa_0(t)\Delta\theta^{(0)} - a\Delta u_t = 0. \end{cases}$$

Multiplying the first equation by u_t and multiplying the second equation by θ, we get

$$\tfrac{1}{2}\frac{d}{dt}\int_\Omega |u_t|^2\,dx + \tfrac{1}{2}\frac{d}{dt}\int_\Omega |\Delta u|^2\,dx + a\int_\Omega \Delta\theta u_t\,dx = \int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx,$$

$$\tfrac{1}{2}\frac{d}{dt}\int_\Omega |\theta|^2\,dx + \kappa_\varepsilon(t)\int_\Omega |\nabla\theta|^2\,dx - (\kappa_\varepsilon - \kappa_0)(t)\int_\Omega \Delta\theta^{(0)}\theta\,dx - a\int_\Omega \Delta u_t\theta\,dx = 0.$$

Since

$$\int_\Omega \Delta\theta u_t\,dx = \int_\Omega \Delta u_t\theta\,dx,$$

it follows that

$$\text{(5-1)}\quad \frac{d}{dt}\left(\tfrac{1}{2}\int_\Omega |u_t|^2\,dx + \tfrac{1}{2}\int_\Omega |\Delta u|^2\,dx + \tfrac{1}{2}\int_\Omega |\theta|^2\,dx\right)$$
$$= -\kappa_\varepsilon(t)\int_\Omega |\nabla\theta|^2\,dx$$
$$+ (\kappa_\varepsilon - \kappa_0)(t)\int_\Omega \Delta\theta^{(0)}\theta\,dx + \int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx$$
$$\leqslant (\kappa_\varepsilon - \kappa_0)(t)\int_\Omega \Delta\theta^{(0)}\theta\,dx + \int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx.$$

Using Young's inequality

$$\int_\Omega \nabla\theta^{(0)}\nabla\theta\,dx \leqslant \tfrac{1}{2}\int_\Omega |\nabla\theta^{(0)}|^2\,dx + \tfrac{1}{2}\int_\Omega |\nabla\theta|^2\,dx,$$

by (5-1) we conclude that

$$\text{(5-2)}\quad \frac{d}{dt}\big(\|u_t\|_X^2 + \|u\|_{X^{1/2}}^2 + \|\theta\|_X^2\big)$$
$$\leqslant \|\theta^{(0)}\|^2_{H_0^1(\Omega)}\|\kappa_\varepsilon - \kappa_0\|_{L^\infty(\mathbb{R})} + \|\theta\|^2_{H_0^1(\Omega)}\|\kappa_\varepsilon$$
$$-\kappa_0\|_{L^\infty(\mathbb{R})} + 2\int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx,$$

and from Section 4 we have $w = w^{(\varepsilon)} - w^{(0)}$, with $w^{(\varepsilon)} = (u^{(\varepsilon)}, u_t^{(\varepsilon)}, \theta^{(\varepsilon)})$ and $w^{(0)} = (u^{(0)}, {u^{(0)}}_t, \theta^{(0)})$ ($u = u^{(\varepsilon)} - u^{(0)}$ and $\theta = \theta^{(\varepsilon)} - \theta^{(0)}$) bounded in $X^1 \times X^{\frac{1}{2}} \times X^{\frac{1}{2}}$.

Hence there exists $C > 0$ independent of $\varepsilon > 0$ such that

$$\|\theta\|_{H_0^1(\Omega)} \leqslant C \quad (\text{and } \|\theta^{(0)}\|_{H_0^1(\Omega)} \leqslant C) \tag{5-3}$$

for any $\varepsilon \in [0, 1)$.

Combining (5-2) and (5-3) we conclude that

$$\frac{d}{dt}(\|u_t\|_X^2+\|u\|_{X^{1/2}}^2+\|\theta\|_X^2)\leqslant C\|\kappa_\varepsilon-\kappa_0\|_{L^\infty(\mathbb{R})}+2\int_\Omega [f(u^{(\varepsilon)})-f(u^{(0)})]u_t\,dx, \tag{5-4}$$

where $C > 0$ is independent of ε.

From the mean value theorem, assumption (1-6) and $\frac{(\rho-1)}{2\rho} + \frac{1}{2\rho} + \frac{1}{2} = 1$, we obtain

$$\begin{aligned}\int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx \\ &\leqslant \|f'(\xi u^{(\varepsilon)} + (1-\xi)u^{(0)})\|_{L^{\frac{2\rho}{\rho-1}}(\Omega)} \|u^{(\varepsilon)} - u^{(0)}\|_{L^{2\rho}(\Omega)} \|u_t\|_{L^2(\Omega)} \\ &\leqslant C_{0,f}\|u^{(\varepsilon)} - u^{(0)}\|_{L^{2\rho}(\Omega)}\|u_t\|_X,\end{aligned}$$

and so

$$\int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx \leqslant C_{0,f}\|u\|_{L^{2\rho}(\Omega)}\|u_t\|_X \leqslant C_{0,f}\|u\|_{X^{1/2}}\|u_t\|_X$$

for some $\xi \in [0, 1]$ and such that $C_{0,f} > 0$ is a constant depending on the initial data.

Hence, from Young's inequality,

$$\int_\Omega [f(u^{(\varepsilon)}) - f(u^{(0)})]u_t\,dx \leqslant C'(\|u\|_{X^{1/2}}^2 + \|u_t\|_X^2 + \|\theta\|_X^2) \tag{5-5}$$

for some $C' > 0$ independent of ε.

Therefore, by (5-4) and (5-5)

$$\frac{d}{dt}(\|u\|_{X^{1/2}} + \|u_t\|_X + \|\theta\|_X) \leqslant C\|\kappa_\varepsilon - \kappa_0\|_{L^\infty(\mathbb{R})} + C''(\|u\|_{X^{1/2}}^2 + \|u_t\|_X^2 + \|\theta\|_X^2),$$

and consequently

$$\|u_t\|_X^2 + \|u\|_{X^{1/2}}^2 + \|\theta\|_X^2 \leqslant C\|\kappa_\varepsilon - \kappa_0\|_{L^\infty(\mathbb{R})}(t-\tau)e^{C''(t-\tau)}, \quad t > \tau, \tag{5-6}$$

that is, $w^{(\varepsilon)} (= S_{(\kappa_\varepsilon)}(t, \tau)w_0)$ goes to $w^{(0)}$ $(= S_{(\kappa_0)}(t, \tau)w_0)$ as $\varepsilon \to 0^+$ in compact subsets of $\mathbb{R}$ uniformly for w_0 in bounded subsets of Y.

For $\delta > 0$ given, let $\tau \in \mathbb{R}$ be such that $\operatorname{dist}(S_{(\kappa_0)}(t, \tau)B, \mathcal{A}_{(\kappa_0)}(t)) < \frac{\delta}{2}$ for all $t \in \mathbb{R}$, $B \supset \bigcup_{s\leqslant t} \mathcal{A}_{(\kappa_\varepsilon)}(s)$, is a bounded set in Y whose existence is guaranteed by Theorem 2.11.

Using (5-6), there exists $\varepsilon_0 > 0$ such that

$$\sup_{u_\varepsilon\in\mathcal{A}_{(\kappa_\varepsilon)}(\tau)} \|S_{(\kappa_\varepsilon)}(t, \tau)u_\varepsilon - S_{(\kappa_0)}(t, \tau)u_\varepsilon\|_Y < \frac{\delta}{2}$$

for all $\varepsilon < \varepsilon_0$. Finally,

$$\begin{aligned}
&\operatorname{dist}_H(\mathcal{A}_{(\kappa_\varepsilon)}(t), \mathcal{A}_{(\kappa_0)}(t)) \\
&\qquad \leqslant \operatorname{dist}_H(S_{(\kappa_\varepsilon)}(t,\tau)\mathcal{A}_{(\kappa_\varepsilon)}(\tau), S_{(\kappa_0)}(t,\tau)\mathcal{A}_{(\kappa_\varepsilon)}(\tau)) \\
&\qquad\qquad\qquad + \operatorname{dist}_H(S_{(\kappa_0)}(t,\tau)\mathcal{A}_{(\kappa_\varepsilon)}(\tau), S_{(\kappa_0)}(t,\tau)\mathcal{A}_{(\kappa_0)}(\tau)) \\
&\qquad \leqslant \sup_{u_\varepsilon \in \mathcal{A}_{(\kappa_\varepsilon)}(\tau)} \operatorname{dist}_H(S_{(\kappa_\varepsilon)}(t,\tau)u_\varepsilon, S_{(\kappa_0)}(t,\tau)u_\varepsilon) \\
&\qquad\qquad\qquad + \operatorname{dist}_H(S_{(\kappa_0)}(t,\tau)\mathcal{A}_{(\kappa_\varepsilon)}(\tau), \mathcal{A}_{(\kappa_0)}(t)) \\
&\qquad < \frac{\delta}{2} + \frac{\delta}{2} = \delta,
\end{aligned}$$

which proves the upper semicontinuity of the family of attractors. □

Remark. Observe that, if we assume that a is continuously differentiable in $\mathbb{R}$, and there exist positive constants a_0 and a_1 such that

$$0 < a_0 \leqslant a(t), a'(t) \leqslant a_1 \quad \text{for all } t \in \mathbb{R},$$

then all the calculations in this paper remain valid for $a(t)$ instead of a.

Acknowledgments

The authors would like to thank the referees for their constructive comments and suggestions which helped us to improve the original manuscript considerably.

References

[Amann 1995] H. Amann, *Linear and quasilinear parabolic problems, I: Abstract linear theory*, Monographs in Math. **89**, Birkhäuser, Basel, 1995. MR Zbl

[Babin and Vishik 1992] A. V. Babin and M. I. Vishik, *Attractors of evolution equations*, Studies in Math. Appl. **25**, North-Holland, Amsterdam, 1992. MR Zbl

[Baroun et al. 2009] M. Baroun, S. Boulite, T. Diagana, and L. Maniar, "Almost periodic solutions to some semilinear non-autonomous thermoelastic plate equations", *J. Math. Anal. Appl.* **349**:1 (2009), 74–84. MR Zbl

[Bezerra et al. 2018] F. D. M. Bezerra, V. L. Carbone, M. J. D. Nascimento, and K. Schiabel, "Pullback attractors for a class of non-autonomous thermoelastic plate systems", *Discrete Contin. Dyn. Syst. B* **23**:9 (2018), 3553–3571. Zbl

[Caraballo et al. 2010] T. Caraballo, A. N. Carvalho, J. A. Langa, and F. Rivero, "Existence of pullback attractors for pullback asymptotically compact processes", *Nonlinear Anal.* **72**:3-4 (2010), 1967–1976. MR Zbl

[Caraballo et al. 2011] T. Caraballo, A. N. Carvalho, J. A. Langa, and F. Rivero, "A non-autonomous strongly damped wave equation: existence and continuity of the pullback attractor", *Nonlinear Anal.* **74**:6 (2011), 2272–2283. MR Zbl

[Carbone et al. 2011] V. L. Carbone, M. J. D. Nascimento, K. Schiabel-Silva, and R. P. da Silva, "Pullback attractors for a singularly nonautonomous plate equation", *Electron. J. Differential Equations* **2011** (2011), art. id. 77. MR Zbl

[Carvalho and Cholewa 2002] A. N. Carvalho and J. W. Cholewa, "Local well posedness for strongly damped wave equations with critical nonlinearities", *Bull. Austral. Math. Soc.* **66**:3 (2002), 443–463. MR Zbl

[Carvalho and Nascimento 2009] A. N. Carvalho and M. J. D. Nascimento, "Singularly non-autonomous semilinear parabolic problems with critical exponents", *Discrete Contin. Dyn. Syst. S* **2**:3 (2009), 449–471. MR Zbl

[Carvalho et al. 2009] A. N. Carvalho, J. W. Cholewa, and T. Dlotko, "Damped wave equations with fast growing dissipative nonlinearities", *Discrete Contin. Dyn. Syst. A* **24**:4 (2009), 1147–1165. MR Zbl

[Carvalho et al. 2013] A. N. Carvalho, J. A. Langa, and J. C. Robinson, *Attractors for infinite-dimensional non-autonomous dynamical systems*, Appl. Math. Sci. **182**, Springer, 2013. MR Zbl

[Chepyzhov and Vishik 2002] V. V. Chepyzhov and M. I. Vishik, *Attractors for equations of mathematical physics*, Amer. Math. Soc. Colloquium Publ. **49**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Hale 1988] J. K. Hale, *Asymptotic behavior of dissipative systems*, Math. Surv. Monographs **25**, Amer. Math. Soc., Providence, RI, 1988. MR Zbl

[Henry 1981] D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Math. **840**, Springer, 1981. MR Zbl

[Lasiecka and Triggiani 1998] I. Lasiecka and R. Triggiani, "Analyticity, and lack thereof, of thermoelastic semigroups", pp. 199–222 in *Control and partial differential equations* (Marseille-Luminy, 1997), edited by C. Fabre et al., ESAIM Proc. **4**, Soc. Math. Appl. Indust., Paris, 1998. MR Zbl

[Sobolevskiĭ 1961] P. E. Sobolevskiĭ, "Equations of parabolic type in a Banach space", *Trudy Mosk. Mat. Obs.* **10** (1961), 297–350. In Russian; translated in *Amer. Math. Soc. Trans.* **49** (1966), 1–62. MR

[Triebel 1978] H. Triebel, *Interpolation theory, function spaces, differential operators*, VEB Deutscher, Berlin, 1978. MR Zbl

Received September 27, 2018. Revised January 12, 2019.

FLANK D. M. BEZERRA
DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DA PARAÍBA
JOÃO PESSOA
BRAZIL
flank@mat.ufpb.br

VERA L. CARBONE
carbone@dm.ufscar.br

MARCELO J. D. NASCIMENTO
marcelo@dm.ufscar.br

KARINA SCHIABEL
schiabel@dm.ufscar.br

(*all*)
DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DE SÃO CARLOS
SÃO CARLOS
BRAZIL

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.421

VARIATIONS OF PROJECTIVITY FOR C^*-ALGEBRAS

DON HADWIN AND TATIANA SHULMAN

We consider various lifting problems for C^*-algebras. As an application of our results we show that any commuting family of order zero maps from matrices to a von Neumann central sequence algebra can be lifted to a commuting family of order zero maps to the C^*-central sequence algebra.

1. Introduction

Many important properties of C^*-algebras are formulated in terms of liftings.

By a lifting property we mean the following. Suppose we are given a surjective $*$-homomorphism $\mathcal{B} \twoheadrightarrow \mathcal{M}$. We will say that a C^*-algebra $\mathcal{A}$ has the lifting property corresponding to this surjection if for any $*$-homomorphism $\phi : \mathcal{A} \to \mathcal{M}$ there is a $*$-homomorphism $\psi : \mathcal{A} \to \mathcal{B}$ such that the following diagram commutes.

$$\begin{tikzcd} & \mathcal{B} \arrow[d, two heads] \\ \mathcal{A} \arrow[ur, dashed, "\psi"] \arrow[r, "\phi"'] & \mathcal{M} \end{tikzcd}$$

In other words any $*$-homomorphism from $\mathcal{A}$ to the C^*-algebra $\mathcal{M}$ "downstairs" "lifts" to a $*$-homomorphism to the C^*-algebra $\mathcal{B}$ "upstairs".

C^*-algebras which have the lifting property with respect to any surjection are called *projective*. They were introduced by B. Blackadar [1985].

Many problems that arise in C^*-algebras reduce to the question of the existence of liftings in various special situations. Here are some examples:

(1) Problems about approximation of almost commuting matrices by commuting ones and, more generally, matricial weak semiprojectivity for C^*-algebras [Loring 1997; Lin 1997; Friis and Rørdam 1996; Eilers et al. 1998], is expressed as the lifting property corresponding to the surjection $\prod_{n\in\mathbb{N}} M_n \twoheadrightarrow \prod_{n\in\mathbb{N}} M_n / \oplus_{n\in\mathbb{N}} M_n$ (here M_n is the C^*-algebra of all n-by-n matrices).

(2) Stability of C^*-algebraic relations under small Hilbert–Schmidt perturbations in matrices is expressed as the lifting property corresponding to the surjection

MSC2010: 46L05.

Keywords: projective C^*-algebras, RFD C^*-algebras, almost commuting matrices, tracial ultraproducts, order zero maps.

$\prod_{n\in\mathbb{N}} M_n \twoheadrightarrow \prod_{n\in\mathbb{N}}^{\alpha}(M_n, \mathrm{tr}_n)$ (here α is a nontrivial ultrafilter on $\mathbb{N}$ and the C^*-algebra $\prod_{n\in\mathbb{N}}^{\alpha}(M_n, \mathrm{tr}_n)$ "downstairs" is the tracial ultraproduct of the matrix algebras) [Hadwin and Shulman 2018b]. Stability under small tracial perturbations in II_1-factors is expressed as the lifting property corresponding to the surjection $\prod_{n\in\mathbb{N}} N_n \twoheadrightarrow \prod_{n\in\mathbb{N}}^{\alpha}(N_n, \tau_n)$ (here N_n is a II_1-factor with a faithful trace τ_n) [Hadwin and Shulman 2018b]. Similar problems for groups are discussed in [Hadwin and Shulman 2018a] and [Arzhantseva and Păunescu 2015].

(3) The property of a C^*-algebra to be residually finite-dimensional (RFD) was proved in [Hadwin 2014] to be the lifting property corresponding to the surjection $\mathcal{B} \twoheadrightarrow B(H)$, where $\mathcal{B} \subseteq \prod M_n$ is defined as the C^*-algebra of all $*$-strongly convergent sequences of matrices and the surjection $\mathcal{B} \twoheadrightarrow B(H)$ is defined by sending each sequence to its $*$-strong limit. Here we identify M_n with $B(l^2\{1, \dots, n\})$ naturally included in $B(l^2\{\mathbb{N}\}) = B(H)$.

(4) The famous Brown–Douglas–Fillmore theory deals with the lifting of injective $*$-homomorphisms from $C(X)$ to the Calkin algebra $\mathcal{C}(H)$ with respect to the surjection $B(H) \twoheadrightarrow \mathcal{C}(H)$.

(5) In the classification program for C^*-algebras one sometimes has to deal with liftings of $*$-homomorphisms to a von Neumann central sequence algebra $N^{\omega} \cap N'$ to $*$-homomorphisms to the C^*-central sequence algebra $A_{\omega} \cap A'$ (see for instance [Toms et al. 2015]). More details on this and on the surjection $A_{\omega} \cap A' \twoheadrightarrow N^{\omega} \cap N'$ are given in Section 3.

We see that in the examples above the corresponding surjections sometimes have a von Neumann algebra "upstairs", sometimes "downstairs", sometimes at both places. This leads us to introducing the following more general notions.

We say that a C^*-algebra A is *C^*-W^*-projective* if it has the lifting property with respect to any surjection $\mathcal{B} \twoheadrightarrow \mathcal{M}$ with $\mathcal{M}$ being a von Neumann algebra; in a similar way *W^*-C^*-projectivity* and *W^*-W^*-projectivity* are defined. In this terminology the usual projectivity may be called C^*-C^*-projectivity.

Dealing with specific lifting problems, one has to look at liftability of projections, isometries, matrix units, various commutational relations, etc. So it is natural to explore whether and which of those basic relations have the more general property of being C^*-W^*, W^*-W^*, W^*-C^*-projective, and we do it in this paper. The main focus is given to commutational relations, that is, to the C^*-W^*, W^*-W^* and W^*-C^*-projectivity of commutative C^*-algebras, but we consider basic noncommutative relations here as well. Note that for the usual projectivity a characterization of when a separable commutative C^*-algebra is projective is obtained in [Chigogidze and Dranishnikov 2010] and is the following: $C(K)$ is projective if and only if K is a compact absolute retract of covering dimension not larger than 1.

In Section 2 we give necessary definitions and discuss a relation between unital and nonunital cases.

In Section 3 we study C^*-W^*-projectivity. The main result of the section is a characterization of when a separable unital commutative C^*-algebra is C^*-W^*-projective: $C(K)$ is C^*-W^*-projective if and only if K is connected and locally path-connected (Theorem 5). Thus for commutative C^*-algebras C^*-W^*-projectivity is very different from the usual projectivity. We also give restrictions on a C^*-algebra to be C^*-W^*-projective, namely it has to be RFD and cannot have nontrivial projections (Propositions 3 and 4); furthermore we prove that tensoring a separable nonunital commutative C^*-W^*-projective C^*-algebra with matrices preserves C^*-W^*-projectivity (Theorem 6). These results are applied to certain lifting problems for order zero maps (completely positive maps preserving orthogonality). A commonly used tool in the classification of C^*-algebras is the fact that an order zero map from the matrix algebra M_n to any quotient C^*-algebra lifts (the so-called projectivity of order zero maps). In particular, the possibility of lifting an order zero map from M_n to a von Neumann central sequence algebra $N^\omega \cap N'$ to an order zero map to the C^*-central sequence algebra $A_\omega \cap A'$ is a key ingredient in obtaining uniformly tracially large order zero maps [Toms et al. 2015]. As an application of our results we prove a stronger statement: one can lift any commuting family of order zero maps $M_n \to N^\omega \cap N'$ to a commuting family of order zero maps $M_n \to A_\omega \cap A'$ (Theorem 8).

In Section 4 we study W^*-C^*-projectivity. This seems to be the most intractable case. We don't have a characterization of W^*-C^*-projectivity for commutative C^*-algebras, we only have a sufficient condition (Corollary 12) and, in the case when the spectrum is a Peano continuum, a necessary condition (Proposition 15). We prove basic noncommutative results such as lifting projections and partial isometries, we consider W^*-C^*-projectivity of matrix algebras, Toeplitz algebra, Cuntz algebras and we discuss a relation with extension groups Ext. Some of the techniques developed in this section are applied in Section 5.

In Section 5 we study W^*-W^*-projectivity. The main result here is that all separable subhomogeneous C^*-algebras are W^*-W^*-projective (Theorem 31). In particular, all separable commutative C^*-algebras are W^*-W^*-projective. We discuss also a relation between W^*-W^*-projectivity and the RFD property. It is easy to show that if a C^*-algebra $\mathcal{A}$ is separable nuclear W^*-W^*- projective and has a faithful trace, then it must be RFD. Moreover if Connes' embedding problem has an affirmative answer, then every unital W^*-W^*-projective C^*-algebra with a faithful trace is RFD. The converse to this statement is not true. Indeed in [Hadwin and Shulman 2018b] we constructed a nuclear C^*-algebra which is RFD (hence has a faithful trace) but is not matricially tracially stable (that is not stable under small Hilbert–Schmidt perturbations in matrices) and hence is not W^*-W^*-projective. In this paper we give an example which is not only nuclear but is even AF (Theorem 33). Our arguments of why it is not matricially tracially stable are much simpler than the ones in [Hadwin and Shulman 2018b].

2. Definitions

Definition 1. Suppose $\mathcal{X}$ and $\mathcal{Y}$ are classes of unital C^*-algebras that are closed under isomorphism. We say that a unital C^*-algebra $\mathcal{A}$ is *$\mathcal{X}$-$\mathcal{Y}$ projective* if, for every $\mathcal{B} \in \mathcal{X}$, $\mathcal{M} \in \mathcal{Y}$ and unital surjective $*$-homomorphisms $\pi : \mathcal{B} \to \mathcal{M}$ and every unital $*$-homomorphism $\phi : \mathcal{A} \to \mathcal{M}$, there is a unital $*$-homomorphism $\psi : \mathcal{A} \to \mathcal{B}$ such that $\pi \circ \psi = \phi$.

$$\begin{tikzcd} & \mathcal{B} \arrow[d, two heads, "\pi"] \\ \mathcal{A} \arrow[ur, dashed, "\psi"] \arrow[r, "\phi"'] & \mathcal{M} \end{tikzcd}$$

The same conditions with all the words "unital" omitted define *$\mathcal{X}$-$\mathcal{Y}$ projectivity in the nonunital category*.

We use the term *C^*-W^*-projective* when $\mathcal{X}$ is the class of all unital C^*-algebras and $\mathcal{Y}$ is the class of all von Neumann algebras. We use the term *C^*-W^*-projective in the nonunital category* when $\mathcal{X}$ is the class of all C^*-algebras and $\mathcal{Y}$ is the class of all von Neumann algebras.

The terms *W^*-C^*-projective (in the nonunital category)*, *W^*-W^*-projective (in the nonunital category)* and *C^*-C^*-projective (in the nonunital category)* are defined similarly. The usual notion of projectivity defined by Blackadar [1985] is the C^*-C^*-projectivity in the nonunital category.

The term *RR0-projectivity* is used when $\mathcal{X} = \mathcal{Y}$ is the class of unital real rank zero C^*-algebras.

Thus in the introduction in the formulation of some of our results we should have added "in the nonunital category". We did not do this to avoid confusing the readers.

We will work mostly with the unital category, but with some exceptions. Namely, in Section 3 dealing with order zero maps one has to consider the nonunital case, and in Sections 4 and 5 proving stability of the class of W^*-W^* and W^*-C^*-projective C^*-algebras under tensoring with matrices and taking direct sums, one has to deal with the nonunital category.

In fact the relation between the unital and nonunital cases is simple. For a C^*-algebra $\mathcal{A}$, let $\widetilde{\mathcal{A}} = \mathcal{A}^+$ if $\mathcal{A}$ is nonunital and $\widetilde{\mathcal{A}} = \mathcal{A} \oplus \mathbb{C}$ if $\mathcal{A}$ is unital.

Proposition 2. *Let $\mathcal{A}$ be a C^*-algebra. Then $\mathcal{A}$ is W^*-C^*-projective in the nonunital category* (W^*-W^*, C^*-W^*, *C^*-C^*-projective in the nonunital category respectively*) *if and only if $\widetilde{\mathcal{A}}$ is W^*-C^*-projective* (W^*-W^*, C^*-W^*, *C^*-C^*-projective respectively*).

3. C^*-W^*-projectivity

The following result puts a severe restriction on being C^*-W^*-projective. In particular, if $C(K)$ is C^*-W^*-projective, then K must be connected.

Proposition 3. *Let $\mathcal{A}$ be a unital C^*-algebra. If $\mathcal{A}$ is C^*-W^*-projective, then $\mathcal{A}$ is $*$-isomorphic to a unital C^*-subalgebra of the unitization of the cone of $\mathcal{A}^{**}$. In particular, $\mathcal{A}$ has no nontrivial projections.*

Proof. Let $\mathcal{B}$ be the unitization of the cone of $\mathcal{A}^{**}$. We know that there is a unital $*$-homomorphism $\pi : \mathcal{B} \to \mathcal{A}^{**}$ and there is a faithful unital $*$-homomorphism $\rho : \mathcal{A} \to \mathcal{A}^{**}$. If $\mathcal{A}$ is C^*-W^*-projective, then there must be a unital $*$-homomorphism $\tau : \mathcal{A} \to \mathcal{B}$ such that $\rho = \pi \circ \tau$. Since ρ is faithful, τ is an embedding. However, $\mathcal{B}$ has no nontrivial projections, so $\mathcal{A}$ has no nontrivial projections. □

Proposition 4. *Let $\mathcal{A}$ be a separable C^*-algebra. If $\mathcal{A}$ is C^*-W^*-projective in either the unital or nonunital category, then $\mathcal{A}$ is RFD.*

Proof. Let $H = l^2(\mathbb{N})$. We will identify the algebra M_n of n-by-n matrices with

$$B(l^2\{1, \dots, n\}) \subseteq B(H).$$

Let $\mathcal{B} \subseteq \prod M_n$ be the C^*-algebra of all $*$-strongly convergent sequences and let I be the ideal of all sequences $*$-strongly convergent to zero. Then one can identify $\mathcal{B}/I$ with $B(H)$ by sending each sequence to its $*$-strong limit. In [Hadwin 2014] it was proved (answering a question of Loring) that a separable C^*-algebra $\mathcal{A}$ is RFD if and only if each $*$-homomorphism from $\mathcal{A}$ to $\mathcal{B}/I$ lifts to a $*$-homomorphism from $\mathcal{A}$ to $\mathcal{B}$. Since $\mathcal{B}/I = B(H)$ is a von Neumann algebra, the result follows. □

We next characterize C^*-W^*- projectivity and C^*-AW^*-projectivity for separable commutative C^*-algebras. Recall that a C^*-algebra is AW^* [Kaplansky 1951] if every set of projections has a least upper bound and every maximal abelian selfadjoint subalgebra is the C^*-algebra generated by its projections.

Theorem 5. *Let K be a compact metric space. Then the following are equivalent*:

(1) *$C(K)$ is C^*-W^*-projective.*

(2) *$C(K)$ is C^*-AW^*-projective.*

(3) *$C(K)$ is C^*-$\mathcal{Y}$ projective, where $\mathcal{Y}$ is the class of all unital C^*-algebras in which every commutative separable C^*-subalgebra is contained in a commutative C^*-algebra generated by projections.*

(4) *K is a continuous image of* [0, 1].

(5) *K is connected and locally path-connected.*

(6) *Every continuous function from a closed subset of* [0, 1] *into K can be extended to a continuous function from* [0, 1] *into K.*

(7) *Every continuous function from a closed subset of the Cantor set into K can be extended to a continuous function from* [0, 1] *into K.*

Proof. The implications (3) $\Rightarrow$ (2) $\Rightarrow$ (1) and (6) $\Rightarrow$ (7) are clear.

(4) $\Leftrightarrow$ (5). This is the Hahn–Mazurkiewicz theorem, see [Hocking and Young 1961, Theorems 3–30].

(1) $\Rightarrow$ (4). Suppose $C(K)$ is C^*-W^*-projective. Then there is a compact Hausdorff space X such that $C(K)^{**} = C(X)$. We know that there is an embedding $\varphi : X \to \prod_{i\in I}[0, 1]$ (product topology). Hence there is a surjective $*$-homomorphism from $C\bigl(\prod_{i\in I}[0, 1]\bigr)$ onto $C(K)^{**}$. Since the canonical embedding from $C(K)$ to $C(K)^{**}$ is an injective $*$-homomorphism, we know from (1) that there is an injective unital $*$-homomorphism from $C(K)$ to $C\bigl(\prod_{i\in I}[0, 1]\bigr)$. Hence there is a continuous surjective map $\beta : \prod_{i\in I}[0, 1] \to K$. Suppose D is a countable dense subset of K and E is the set of all elements of $\prod_{i\in I}[0, 1]$ with finite support, i.e., only finitely many nonzero coordinates. If $x \in D$, then there is a countable set E_x of E such that $x \in \overline{\beta(E_x)}$. Hence

$$\beta\Bigl(\overline{\bigcup_{x\in D} E_x}\Bigr) = K,$$

since $\beta\bigl(\overline{\bigcup_{x\in D} E_x}\bigr)$ is compact and contains D. It follows that there is a countable subset $J \subseteq I$ such that

$$\Bigl(\overline{\bigcup_{x\in D} E_x}\Bigr) \subseteq \prod_{i\in J}[0, 1] \times \prod_{i\in I\setminus J}\{0\};$$

whence,

$$\beta\Bigl(\prod_{i\in J}[0, 1] \times \prod_{i\in I\setminus J}\{0\}\Bigr) = K.$$

Hence K is a continuous image of $\prod_{i\in J}[0, 1]$, which in turn is a continuous image of $[0, 1]$. Thus K is a continuous image of $[0, 1]$.

(5) $\Rightarrow$ (6). Suppose K is connected and locally path connected and E is any closed subset of $[0, 1]$, and $f : E \to K$ is continuous. For $x, y \in K$, let $\Delta(x, y)$ be the infimum of the diameters of every path in K from x to y. We first note that

$$\lim_{d(x,y)\to 0} \Delta(x, y) = 0.$$

If this is not true, then there is an $\varepsilon > 0$ and sequences $\{x_n\}$ and $\{y_n\}$ in K such that $d(x_n, y_n) \to 0$ and $\Delta(x_n, y_n) \geq \varepsilon$ for every $n \in \mathbb{N}$. Since K is compact we can replace $\{x_n\}$ and $\{y_n\}$ with subsequences that converge to x and y, respectively. Since $d(x, y) = \lim_{n\to\infty} d(x_n, y_n) = 0$, we see that $x = y$. Since X is locally connected, there is path connected neighborhood U of x such that U is contained in the ball centered at x with radius $\varepsilon/3$. There must be an n such that $x_n, y_n \in U$, and there must be a path γ in U from x_n to y_n. Since γ is in the ball centered at x with radius $\varepsilon/3$, the diameter of γ is at most $2\varepsilon/3$, which implies $\Delta(x_n, y_n) < \varepsilon$, which is a contradiction.

We can clearly add 0 and 1 to E and extend f so that it is still continuous. Hence we can assume that $0, 1 \in E$. We can write

$$[0, 1]\backslash E = \bigcup_{n \in I} (a_n, b_n),$$

where $\{(a_n, b_n) : n \in I\}$ is a disjoint set of open intervals with $I \subseteq \mathbb{N}$. For each $n \in I$ we chose a path $\gamma_n : [a_n, b_n] \to K$ from $f(a_n)$ to $f(b_n)$ so that the diameter of γ_n is less than $2\Delta(f(a_n), f(b_n))$.

(7) $\Rightarrow$ (3). Suppose (7) is true. Suppose $\mathcal{B}$ is a unital C^*-algebra, $\mathcal{M} \in \mathcal{Y}$ and $\pi : \mathcal{B} \to \mathcal{M}$ is a unital surjective $*$-homomorphism. Let $\rho : C(K) \to \mathcal{M}$ be a unital $*$-homomorphism. Since K is metrizable, $C(K)$ is separable, and since $\mathcal{M} \in \mathcal{Y}$, there is a countable commuting family $\{p_1, p_2, \ldots\}$ of projections in $\mathcal{M}$ such that $\rho(C(K)) \subseteq \mathcal{C}^*(p_1, p_2, \ldots)$. Let E be the maximal ideal space of $\mathcal{C}^*(p_1, p_2, \ldots)$. Since $\mathcal{C}^*(p_1, p_2, \ldots)$ is generated by countably many projections, E is a totally disconnected compact metric space and is therefore homeomorphic to a subset of the Cantor set. Hence there is an $a = a^* \in \mathcal{C}^*(p_1, p_2, \ldots)$ such that $\mathcal{C}^*(p_1, p_2, \ldots) = C^*(a)$ and $\sigma(a)$ (homeomorphic to E) is a subset of the Cantor set. Let $\Gamma : C^*(a) \to C(\sigma(a))$ be the Gelfand map. Then $\Gamma \circ \rho : C(K) \to C(\sigma(a))$ is a unital $*$-homomorphism, so there is a continuous function $\psi : \sigma(a) \to K$ so that

$$\Gamma(\rho(f)) = f \circ \psi$$

for every $f \in C(K)$. By applying Γ^{-1}, we get

$$\rho(f) = (f \circ \psi)(a)$$

for every $f \in C(K)$.

By (7), we can assume (by extending) that $\psi : [0, 1] \to K$. We can find $A \in \mathcal{B}$ with $0 \le A \le 1$ such that $\pi(A) = a$. We now define $\nu : C(K) \to \mathcal{B}$ by

$$\nu(f) = (f \circ \psi)(A).$$

Then ν is a unital $*$-homomorphism and, for every $f \in C(K)$, we have

$$\pi(\nu(f)) = \pi((f \circ \psi)(A)) = (f \circ \psi)(a) = \rho(f).$$

Hence, $\rho = \pi \circ \nu$. This proves (3). □

Let K be a compact metric space and let x_0 be any point in K. Let us denote by $C_0(K \setminus \{x_0\})$ the C^*-algebra of all continuous functions on K vanishing at x_0.

Theorem 6. *Let K be a continuous image of* $[0, 1]$, $n \in \mathbb{N}$. *Then the C^*-algebra $C_0(K \setminus \{x_0\}) \otimes M_n$ is C^*-W^*-projective in the nonunital category.*

Proof. Since $C_0(K \setminus \{x_0\})^{**}$ is a commutative von Neumann algebra,

$$C_0(K \setminus \{x_0\})^{**} \cong C(X),$$

for some extremally disconnected space X. Let $i_* : C_0(K \setminus \{x_0\}) \to C(X)$ be the canonical embedding into the bidual. It is induced by some surjective continuous map $i : X \twoheadrightarrow K$. Since K is a continuous image of $[0, 1]$, there is a surjective continuous map $\alpha : [0, 1] \twoheadrightarrow K$. By the universal property of extremally disconnected spaces (Gleason's theorem), i factorizes through a continuous map $\beta : X \to [0, 1]$.

$$\begin{tikzcd}
{[0,1]} \arrow[d, two heads, "\alpha"] & \\
K & X \arrow[l, two heads, "i"'] \arrow[ul, bend right, "\beta"']
\end{tikzcd}$$

Let

$$\alpha_* : C_0(K \setminus \{x_0\}) \to C_0(0, 1], \quad \beta_* : C_0(0, 1] \to C_0(X) \subset C(X)$$

be the $*$-homomorphisms induced by α and β. Let $\mathcal{M}$ be a von Neumann algebra, let $\mathcal{B}$ be a C^*-algebra, and let $q : \mathcal{B} \to \mathcal{M}$ be a surjective $*$-homomorphism. Furthermore, let $\phi : C_0(K \setminus \{x_0\}) \otimes M_n \to \mathcal{M}$ be a $*$-homomorphism. By the universal property of double duals we can extend ϕ to a $*$-homomorphism $\tilde{\phi} : C(X) \otimes M_n \cong (C_0(K \setminus \{x_0\}) \otimes M_n)^{**} \to \mathcal{M}$.

$$\begin{tikzcd}
\mathcal{B} \arrow[d, "q"] & & C_0(0,1] \otimes M_n \arrow[ll, dotted, "\psi"'] \arrow[dr, "\beta_* \otimes \mathrm{id}_{M_n}"] & \\
\mathcal{M} & C_0(K \setminus \{x_0\}) \otimes M_n \arrow[l, "\phi"'] \arrow[ur, "\alpha_* \otimes \mathrm{id}_{M_n}"] \arrow[rr, hook, "i_* \otimes \mathrm{id}_{M_n}"] & & C(X) \otimes M_n \arrow[lll, bend left, "\tilde{\phi}"]
\end{tikzcd}$$

Since $C_0(0, 1] \otimes M_n$ is projective [Loring 1997, Theorem 10.2.1], there is a $*$-homomorphism $\psi : C_0(0, 1] \otimes M_n \to \mathcal{B}$ such that

$$q \circ \psi = \tilde{\phi} \circ (\beta_* \otimes \mathrm{id}_{M_n}).$$

Then

$$q \circ \psi \circ (\alpha_* \otimes \mathrm{id}_{M_n}) = \tilde{\phi} \circ (\beta_* \otimes \mathrm{id}_{M_n}) \circ (\alpha_* \otimes \mathrm{id}_{M_n}) = \tilde{\phi} \circ (i_* \otimes \mathrm{id}_{M_n}) = \phi.$$

Thus $\psi \circ (\alpha_* \otimes \mathrm{id}_{M_n})$ is a lift of ϕ. $\square$

We don't know if the previous result can be generalized to get the following: if A is C^*-W^*-projective in the nonunital category, then so is $A \otimes M_n$.

Recent developments in the classification of C^*-algebras show the importance of the analysis of central sequence algebras. Let ω be a free ultrafilter on $\mathbb{N}$. With any C^*-algebra A with a faithful tracial state τ one can associate its C^*-central sequence algebra $A_\omega \cap A'$ and its W^*-central sequence algebra $N^\omega \cap N'$, where N is the weak closure of A under the GNS representation π_τ of A. There is a natural $*$-homomorphism $\gamma : A_\omega \cap A' \to N^\omega \cap N'$; it was proved in [Sato 2011] and [Kirchberg and Rørdam 2014] that γ is surjective.

A commonly used tool in the classification of C^*-algebras is the fact that an order zero map from the matrix algebra M_n to any quotient C^*-algebra lifts (the so-called projectivity of order zero maps). In particular, the possibility of lifting an order zero map $M_n \to N^\omega \cap N'$ to an order zero map $M_n \to A_\omega \cap A'$ is a key ingredient in obtaining uniformly tracially large order zero maps [Toms et al. 2015]. Below we prove a stronger statement: one can lift any commuting family of order zero maps $M_n \to N^\omega \cap N'$ to a commuting family of order zero maps $M_n \to A_\omega \cap A'$.

Lemma 7. *Let $d \in \mathbb{N}$, $k \in \mathbb{N} \cup \{\infty\}$. The C^*-algebra $C_0([0, 1]^k) \otimes M_{kd}$ is isomorphic to the universal C^*-algebra with generators $e_{ij}^{n,l}$, $l \le k$, $i, j \le d$, $n \in \mathbb{N}$ and relations*

$$\begin{aligned}
0 \le e_{ii}^{n,l} &\le 1, \\
(e_{ij}^{n,l})^* &= e_{ji}^{n,l}, \\
e_{ij}^{m,l} e_{ks}^{n,l} &= \delta_{jk} e_{is}^{m,l}, \quad m < n, \\
e_{ij}^{n,l} e_{ks}^{n,l} &= \delta_{jk} e_{ii}^{n,l} e_{is}^{n,l}, \\
[e_{ij}^{n,l}, e_{i'j'}^{n',l'}] = [e_{ij}^{n,l}, (e_{i'j'}^{n',l'})^*] &= 0 \quad \textit{when } l \ne l'.
\end{aligned}$$

The proof of the lemma is analogous to the proof of Lemma 6.2.4 in [Loring 1997].

Theorem 8. *Any commuting family of order zero maps from M_n to $N^\omega \cap N'$ lifts to a commuting family of order zero maps from M_n to $A_\omega \cap A'$.*

Proof. As was proved in [Winter and Zacharias 2009], with any order zero map $\phi : B \to D$ one can associate a $*$-homomorphism $f : CB \to D$ and vice versa. Here $CB = C_0(0, 1] \otimes B$ is the cone over B. Moreover, it follows from the construction in [Winter and Zacharias 2009] that order zero maps ϕ_i's have commuting ranges if and only if the corresponding f_i's have commuting ranges. It also follows from the construction that if for an order zero map $\phi : A \to B/I$ the corresponding $f : CA \to B/I$ lifts to a $*$-homomorphism $\tilde{f} : CA \to B$, then ϕ lifts to the order zero map $\tilde{\phi}$ corresponding to $\tilde{f}$. Thus we need to prove that any family of $*$-homomorphisms $f_i : CM_n \to N^\omega \cap N'$ with pairwise commuting ranges lifts to a family of $*$-homomorphisms $\tilde{f}_i : CM_n \to A_\omega \cap A'$ with pairwise commuting ranges. It follows from Lemma 7 that it is equivalent to lifting one $*$-homomorphism from $C_0([0, 1]^k) \otimes M_{kd}$ to $N^\omega \cap N'$, where k is the number of $*$-homomorphisms in the family. Since $[0, 1]^k$ is a continuous image of $[0, 1]$, the statement follows from Theorem 6. □

4. W^*-C^*-projectivity and RR0-projectivity

Recall that a C^*-algebra has *real rank zero* (RR0) if each of its self-adjoint elements can be approximated by self-adjoint elements with finite spectra. Since a $*$-homomorphic image of a real rank zero C^*-algebra is real rank zero, and since every von Neumann algebra has real rank zero [Brown and Pedersen 1991], it follows that every RR0-projective C^*-algebra is W^*-C^*-projective.

We first prove that in the W^*-C^* case we can lift projections. This was proved in the RR0 case by L. G. Brown and G. Pederson [1991]. We include the proof of the W^*-C^* case because it is much shorter.

Proposition 9. *$\mathbb{C} \oplus \mathbb{C}$ is W^*-C^*-projective.*

Proof. Suppose $\mathcal{B}$ is a von Neumann algebra, $\mathcal{M}$ is a C^*-algebra, and $\pi : \mathcal{B} \to \mathcal{M}$ is a unital $*$-homomorphism. Since $\mathbb{C} \oplus \mathbb{C}$ is the unital universal C^*-algebra of one projection, suppose $q \in \mathcal{M}$ is a projection. By [Loring 1997, Lemma 10.1.12], we can lift q to b_1 and $1-q$ to b_2 in $\mathcal{B}$ such that $0 \le b_j \le 1$, $\pi(b_1) = q$, $\pi(b_2) = 1-q$ and $b_1 b_2 = b_2 b_1 = 0$. Let $p \in \mathcal{B}$ be the range projection for b_1. Then $pb_2 = b_2 p = 0$. Thus $b_1 \le p$ and $b_2 \le 1-p$. Hence $q \le \pi(p)$ and $1-q \le 1-\pi(p)$. Hence, $\pi(p) = q$. □

Definition 10. Suppose B, $\mathcal{M}$ are unital C^*-algebras and $\pi : \mathcal{B} \to \mathcal{M}$ is a surjective $*$-homomorphism, and suppose $\mathcal{S}$ is a C^*-subalgebra of $\mathcal{M}$. We say that γ *is a $*$-cross section for π on S* if $\gamma : \mathcal{S} \to \mathcal{B}$ is a $*$-homomorphism and $\pi \circ \gamma$ is the identity on $\mathcal{S}$. Clearly, every such γ is injective.

Theorem 11. *Suppose $\mathcal{B}$ is a real rank zero C^*-algebra, $\mathcal{M}$ is a C^*-algebra, and $\pi : \mathcal{B} \to \mathcal{M}$ is a unital surjective $*$-homomorphism. Suppose $\{p_1, p_2, \ldots\} \subseteq \mathcal{M}$ is a commuting family of projections. Then there is a unital $*$-cross section γ for π on $C^*(p_1, p_2, \ldots)$. Moreover, if $\gamma_n : C^*(p_1, p_2, \ldots, p_n) \to \mathcal{M}$ is a unital $*$-cross section for π on $C^*(p_1, p_2, \ldots, p_n)$, then there is a unital $*$-cross section γ_{n+1} for π on $C^*(p_1, p_2, \ldots, p_{n+1})$ whose restriction to $C^*(p_1, \ldots, p_n)$ is γ_n.*

Proof. We know that if $0 \ne p \ne 1$ is a projection in $\mathcal{M}$, there is a projection $P \in \mathcal{B}$ with $\pi(P) = p$. Clearly $0 \ne P \ne 1$. Suppose $\gamma_n : C^*(p_1, p_2, \ldots, p_n) \to \mathcal{M}$ is a unital $*$-cross section for π on $C^*(p_1, p_2, \ldots, p_n)$. We know $C^*(p_1, p_2, \ldots, p_n)$ is generated by an orthogonal family of projections $\{q_1, \ldots, q_m\}$ whose sum is 1. For $1 \le k \le m$, let $Q_k = \gamma_n(q_k)$. Now p_{n+1} commutes with $\{q_1, \ldots, q_m\}$, so $C^*(p_1, \ldots, p_{n+1})$ is generated by the orthogonal family

$$\bigcup_{k=1}^{m} \{q_k p_{n+1} q_k,\ q_k - q_k p_{n+1} q_k\}.$$

If $q_k p_{n+1} q_k = 0$ or $q_k - q_k p_{n+1} q_k = 0$, there is nothing new to lift. If $q_k p_{n+1} q_k \ne 0$ and $q_k - q_k p_{n+1} q_k \ne 0$, we need to find a lifting of $q_k p_{n+1} q_k$ in $Q_k \mathcal{B} Q_k$. However, it was proved in [Brown and Pedersen 1991] that $Q_k \mathcal{B} Q_k$ has real rank 0. Thus such a lifting is possible. □

As a corollary we get a sufficient condition for $C(K)$ to be RR0-projective:

Corollary 12. *If K is a totally disconnected compact metric space, then $C(K)$ is* RR0*-projective, and hence W^*-C^*-projective.*

The following corollary uses the fact that if $\gamma : \mathcal{A} \to \mathcal{B}$ is a unital $*$-homomorphism and $\mathcal{B}$ is a von Neumann algebra, then there is an extension to a weak*-weak* continuous $*$-homomorphism $\hat{\gamma} : \mathcal{A}^{**} \to \mathcal{B}$.

Corollary 13. *Let K be a compact metric space. The following are equivalent:*

(1) *$C(K)$ is W^*-C^*-projective.*

(2) *Whenever $\mathcal{B}$ is a von Neumann algebra, $\mathcal{M}$ is a C^*-algebra, $\pi : \mathcal{B} \to \mathcal{M}$ is a surjective $*$-homomorphism and $\rho : C(K) \to \mathcal{M}$ is a unital $*$-homomorphism, there is a commutative C^*-subalgebra $\mathcal{D}$ of $\mathcal{M}$ that contains $\rho(C(K))$ such that the maximal ideal space of $\mathcal{D}$ is totally disconnected.*

(3) *Whenever $\mathcal{B}$ is a von Neumann algebra, $\mathcal{M}$ is a C^*-algebra, $\pi : \mathcal{B} \to \mathcal{M}$ is a surjective $*$-homomorphism and $\rho : C(K) \to \mathcal{M}$ is a unital $*$-homomorphism, ρ extends to a $*$-homomorphism $\hat{\rho} : C(K)^{**} \to \mathcal{M}$.*

Remark 14. Without the separability assumption on $C(K)$ (i.e., the metrizability of K), it is not generally true that $C(K)$ is W^*-C^*-projective whenever K is compact and totally disconnected. For example, let $\mathcal{A}$ be the universal C^*-algebra generated by a mutually orthogonal family $\{P_t : t \in [0, 1]\}$ of projections. The maximal ideal space of $\mathcal{A}$ is the one-point compactification K of the discrete space $[0, 1]$. If we let $\mathcal{B} = B(\ell^2)$ and $\mathcal{M} = B(\ell^2)/\mathcal{K}(\ell^2)$, and let $\pi : \mathcal{B} \to \mathcal{M}$ be the quotient map, then there is an injective unital $*$-homomorphism $\rho : \mathcal{A} \to \mathcal{M}$, but there is no injective unital $*$-homomorphism from $\mathcal{A}$ to $\mathcal{B}$ since $B(\ell^2)$ does not contain an uncountable orthogonal family of nonzero projections. Hence $C(K)$ is not W^*-C^*-projective although K is totally disconnected. This shows that an attempt at a transfinite inductive version of the proof of Theorem 11 is doomed to failure. This also shows that being W^*-C^*-projective is not closed under arbitrary direct limits, since $\mathcal{A}$ is the direct limit of the family $\{C^*(\{P_t : t \in E\}) : E \subseteq [0, 1] \text{ is countable}\}$. We doubt that ρ can be extended to a $*$-homomorphism from $C(K)^{**}$ to $\mathcal{M}$, so the equivalence of (1) and (3) in Corollary 13 may conceivably be true.

In the case when K is a Peano continuum (that is, a nonempty compact connected metric space which is locally connected at each point) there is a necessary condition for $C(K)$ to be W^*-C^*-projective.

Proposition 15. *Suppose K is a Peano continuum and $C(K)$ is W^*-C^*-projective. Then* $\dim_{\text{cov}}(K) \le 1$ (*here* $\dim_{\text{cov}}$ *is the covering dimension*).

Proof. Suppose $\dim_{\text{cov}} K > 1$. Then by [Chigogidze and Dranishnikov 2010, Proposition 3.1], K contains a circle, S^1. Let $j : C(K) \to C(S^1)$ be the restriction map. Let $\pi : B(H) \to B(H)/K(H)$ be the canonical surjection. Let $T \in B(H)$ be the unilateral shift. Define a $*$-homomorphism $\rho : C(S^1) \to B(H)/K(H)$ by sending the identity function z to $\pi(T)$. We claim that $\rho \circ j$ is not liftable. Indeed suppose it lifts to a $*$-homomorphism $\gamma : C(K) \to B(H)$. Let $f \in C(K)$ be any preimage of $z \in C(S^1)$ under the map j. Since f is normal, $\gamma(f) \in B(H)$ is a normal preimage of $\pi(T)$. Since any preimage of $\pi(T)$ has Fredholm index -1 and since any normal Fredholm operator has Fredholm index zero, we come to a contradiction. $\square$

Below we give a few noncommutative examples and nonexamples of W^*-C^*-projective C^*-algebras. The following lemma shows that Murray–von Neumann equivalent projections can be lifted to Murray–von Neumann equivalent projections in the W^*-C^* case. More simply, it states that partial isometries can be lifted.

Lemma 16. *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{M}$ be a von Neumann algebra and $\pi : \mathcal{M} \to \mathcal{B}$ be a surjective $*$-homomorphism. Let $v \in \mathcal{B}$ be a partial isometry with $v^*v = p$ and $vv^* = q$. Let $X \in \mathcal{M}$, $\|X\| \le 1$, and let P, Q be projections in $\mathcal{M}$ such that $\pi(P) = p$, $\pi(Q) = q$ and $\pi(X) = v$. If V is the partial isometry in the polar decomposition of PXQ, then $\pi(V) = v$.*

Proof. Let $Y = PXQ$. Then the range projection P_1 of Y is less than or equal to P and the range projection of Y^* is less than or equal to Q. Moreover, $(YY^*)^{1/2} \le P_1 \le P$. Also,

$$p = \pi((Y^*Y)^{1/2}) \le \pi(P_1) \le \pi(P) = p.$$

If $Y = (YY^*)^{1/2}V$ is the polar decomposition, then

$$v = p\pi(V) = \pi(P_1 V) = \pi(V). \qquad \square$$

Corollary 17. *Let $\mathcal{T}$ be the Toeplitz algebra. Then $\mathcal{T} \oplus \mathbb{C}$ is W^*-C^*-projective.*

Proof. $\mathcal{T} \oplus \mathbb{C}$ is the universal unital C^*-algebra generated by v with the relation that v is a partial isometry. $\square$

Corollary 18. *$\mathcal{M}_n(\mathbb{C}) \oplus \mathbb{C}$ is W^*-C^*-projective.*

Proof. It would be equivalent to prove that $\mathcal{M}_n(\mathbb{C})$ is W^*-C^*-projective in the nonunital category. We will use induction on n. The case when $n = 1$ amounts to lifting a projection. Assume the theorem is true for n. Suppose $\mathcal{M}$ is a von Neumann algebra, $\mathcal{B}$ is a C^*-algebra and $\pi : \mathcal{M} \to \mathcal{B}$ is a surjective $*$-homomorphism. Suppose $\rho : \mathcal{M}_n(\mathbb{C}) \to \mathcal{B}$ is a $*$-homomorphism. It follows from the induction assumption that there is a $*$-homomorphism

$$\gamma_0 : C^*(\{e_{1k} : 2 \le k \le n\}) \to \mathcal{M}$$

so that $\gamma_0(e_{jk}) = E_{jk}$ and $(\pi \circ \gamma_0)(e_{1k}) = \rho(e_{1k})$ for $2 \le k \le n$. We then have $E_{1k}E_{1k}^* = \gamma_0(e_{11})$ for $2 \le k \le n$. Choose $X \in \mathcal{M}$ so that $\pi(X) = \rho(e_{1,n+1})$. If we replace X with $E_{11}X\big(1-\sum_{1\le k\le n}E_{kk}i\big)$, and let V be the partial isometry in the polar decomposition of $E_{11}X\big(1-\sum_{1\le k\le n}E_{kk}\big)$, we have from Lemma 16 that $\pi(V)=\rho(e_{1,n+1})$, and $VV^*\le E_{11}$ and $\pi(VV^*)=\rho(e_{11})$. If we replace E_{1k} with $F_{1k}=VV^*E_{1k}$ for $2\le k\le n$, and define $F_{1,n+1}=V$, we obtain a representation $\gamma:\mathcal{M}_{n+1}(\mathbb{C})\to\mathcal{M}$ with $\gamma(e_{1k})=F_{1k}$ for $1\le k\le n+1$ such that $\pi\circ\gamma=\rho$. □

Lemma 19. *Suppose C^*-algebras A and D are unital W^*-C^*-projective (W^*-W^*-projective respectively) in the nonunital category. Then $A\oplus D$ is W^*-C^*-projective (W^*-W^*-projective respectively) in the nonunital category.*

Proof. Suppose $\mathcal{B}$ and $\mathcal{M}$ are von Neumann algebras ($\mathcal{B}$ is a von Neumann algebra in the W^*-W^* case respectively) and $\pi:\mathcal{B}\to\mathcal{M}$ is a surjective $*$-homomorphism. Let $\phi:A\oplus D\to\mathcal{M}$ be a $*$-homomorphism. Let

$$p=\phi(1_A\oplus 0),\quad q=\phi(0\oplus 1_D).$$

Define $\phi_A:A\to pMp$ and $\phi_D:D\to qMq$ by $\phi_A(a)=\phi(a\oplus 0)$ and $\phi_D(d)=\phi(0\oplus d)$, respectively. It follows from Corollary 12 that we can lift p,q to projections P, Q in B which are orthogonal to each other. Let $\psi_A:A\to PBP$ and $\psi_D:D\to QBQ$ be lifts of ϕ_A and ϕ_D. Define lift ψ of ϕ by

$$\psi(a\oplus d)=\psi_A(a)+\psi_D(d).$$ □

Combining this lemma with Corollary 18 we obtain the following result:

Corollary 20. *If $\mathcal{A}$ is a finite-dimensional C^*-algebra, then $\mathcal{A}\oplus\mathbb{C}$ is W^*-C^*-projective.*

Remark 21. The result in Corollary 20 cannot be extended to AF-algebras even in the W^*-W^* case. Indeed the tracial ultraproduct $\prod_{n\in\mathbb{N}}^{\alpha}(\mathcal{M}_{2^n}(\mathbb{C}),\tau_{2^n})$ with respect to a free ultrafilter α, where τ_{2^n} is the normalized trace on $\mathcal{M}_{2^n}(\mathbb{C})$, is a von Neumann algebra. Thus $\pi:\prod_{n\in\mathbb{N}}\mathcal{M}_{2^n}(\mathbb{C})\to\prod_{n\in\mathbb{N}}^{\alpha}(\mathcal{M}_{2^n}(\mathbb{C}),\tau_{2^n})$ is a unital surjective $*$-homomorphism and the domain and range are both von Neumann algebras. If $\mathcal{A}$ is the CAR algebra, then clearly there is an embedding $\rho:\mathcal{A}\to\prod_{n\in\mathbb{N}}^{\alpha}(\mathcal{M}_{2^n}(\mathbb{C}),\tau_{2^n})$. However, $\mathcal{A}$ is simple and infinite-dimensional, so there is no embedding from $\mathcal{A}\oplus\mathbb{C}$ into $\prod_{n\in\mathbb{N}}\mathcal{M}_{2^n}(\mathbb{C})$ such that $\rho=\pi\circ\tau$.

A trace ψ on a unital MF-algebra $\mathcal{A}$ is called an *MF-trace* if there is a free ultrafilter α on $\mathbb{N}$ and a unital $*$-homomorphism $\pi:\mathcal{A}\to\prod^{\alpha}M_k(\mathbb{C})$ to the C^*-ultraproduct of matrices, such that $\psi=\tau_\alpha\circ\pi$, where $\tau_\alpha(\{A_k\}_\alpha)=\lim_{k\to\alpha}\tau_k(A_k)$ [Li et al. 2014, Proposition 4].

The ideas in the preceding remark easily extend to the following result:

Proposition 22. *If $\mathcal{A}$ is a unital MF C^*-algebra and is W^*-C^*-projective, then $\mathcal{A}$ must be RFD. If the MF-traces are a faithful set on $\mathcal{A}$, i.e., $\tau(a^*a)=0$ for every MF trace implies $a=0$, and if $\mathcal{A}$ is W^*-W^*-projective, then $\mathcal{A}$ must be RFD.*

The following result shows that without adding $\mathbb{C}$ as a direct summand Corollaries 17 and 18 no longer hold:

Proposition 23. *$\mathcal{T}$ and $M_n(\mathbb{C})$ are not W^*-C^*-projective.*

Proof. The Toeplitz algebra is not W^*-C^*-projective, since an isometry in Calkin algebra need not lift to an isometry in $B(H)$. $M_n(\mathbb{C})$ is not W^*-C^*-projective, because $M_n(\mathbb{C})$ is a quotient of $M_n(\mathbb{C}) \oplus \mathbb{C}$ and since $M_n(\mathbb{C})$ does not admit any unital $*$-homomorphisms to $\mathbb{C}$, the identity map on $M_n(\mathbb{C})$ is not liftable. □

Proposition 24. *Suppose $\mathcal{A}$ is a separable unital C^*-algebra.*

(1) *If* Ext($\mathcal{A}$) *is not trivial, then $\mathcal{A}$ is not W^*-C^*-projective*

(2) *If* $\mathrm{Ext}_w(\mathcal{A})$ *is not trivial, then $\mathbb{C} \oplus \mathcal{A}$ is not W^*-C^*-projective.*

Proof. (1) This is obvious.

(2) Suppose that $\mathrm{Ext}_w(\mathcal{A})$ is not trivial. Then there is an injective unital $*$-homomorphism $\rho : \mathcal{A} \to B(\ell^2)/\mathcal{K}(\ell^2)$ that is not weakly equivalent to the trivial element in $\mathrm{Ext}_w(\mathcal{A})$. We prove this by contradiction. Assume there is a nonunital $*$-homomorphism $\gamma : \mathcal{A} \to \mathcal{B}(\ell^2)$ such that $\pi \circ \gamma = \rho$. Then

$$\pi(1 - \gamma(1)) = 1 - \rho(1) = 0.$$

Thus $1 - \gamma(1)$ is a finite-rank projection, and if $\gamma_0(A) = \gamma(A)|_{\gamma(1)(\ell^2)}$, we have $\gamma = 0 \oplus \gamma_0$ relative to $\ell^2 = \ker \gamma(1) \oplus \gamma(1)(\ell^2)$. Since $\rho = \pi \circ \gamma$ is injective, γ_0 must be injective. Choose an isometry V in $B(\ell^2)$ whose range is $\gamma(1)(\ell^2)$. Then $V^*\gamma(\cdot)V$ is unitarily equivalent to γ_0. Thus $\pi(V)$ is unitary in $B(\ell^2)/\mathcal{K}(\ell^2)$ and $\pi(V^*)\rho(\cdot)\pi(V)$ lifts to $V^*\gamma(\cdot)V = U^*\gamma_0(\cdot)U$ for some unitary U. This means ρ is weakly equivalent to the trivial element in $\mathrm{Ext}_w(\mathcal{A})$, which is a contradiction. □

Corollary 25. *If $n \geq 2$, the Cuntz algebra $\mathcal{O}_n$ is not W^*-C^*-projective. If $n \geq 3$, $\mathbb{C} \oplus O_n$ is not W^*-C^*-projective.*

Proof. By Theorem V.6.5 of [Davidson 1996], $\mathrm{Ext}(O_n) \cong \mathbb{Z}$, and by Theorem V.6.6 of the same work, $\mathrm{Ext}_w(O_n) \cong \mathbb{Z}_{n-1}$, when $n \geq 2$. □

Remark 26. By Corollary 18 and Proposition 23, if $n \geq 2$, $\mathcal{M}_n(\mathbb{C})$ is not W^*-C^*-projective, but $\mathbb{C} \oplus \mathcal{M}_n(\mathbb{C})$ is W^*-C^*-projective, and this happily coincides with the fact that $\mathrm{Ext}(\mathcal{M}_n(\mathbb{C}))$ is not trivial and $\mathrm{Ext}_w(\mathcal{M}_n(\mathbb{C}))$ is trivial.

The following is a consequence of the proof of Theorem 9 in [Loring and Shulman 2014]. It generalizes Olsen's structure theorem [1971] for polynomially compact operators.

Theorem 27. *Let $R \geq 0$ and $p \in \mathbb{C}[x]$. The universal C^*-algebra generated by a such that $\|a\| \leq R$ and $p(a) = 0$ is* RR0*-projective and hence W^*-C^*-projective.*

5. W^*-W^*-projectivity

We begin with the separable unital commutative C^*-algebras.

Theorem 28. *Every separable unital commutative C^*-algebra is* RR0-AW^*- *projective. In particular every separable unital commutative C^*-algebra is W^*-W^*-projective.*

Proof. Suppose $\mathcal{B}$ is a real rank zero C^*-algebra, $\mathcal{M}$ is an AW^*-algebra and $\pi : \mathcal{B} \to \mathcal{M}$ is a surjective unital $*$-homomorphism, and suppose that $\mathcal{A}$ is a separable unital commutative C^*-subalgebra of $\mathcal{M}$. Since $\mathcal{M}$ is an AW^*-algebra, every maximal abelian selfadjoint C^*-subalgebra of $\mathcal{M}$ is the C^*-algebra generated by its projections. Since $\mathcal{A}$ is contained in such a maximal algebra and $\mathcal{A}$ is separable, it follows that there is a countable commuting family $\{p_1, p_2, \ldots\}$ of projections in $\mathcal{M}$ such that $\mathcal{A} \subset C^*(p_1, p_2, \ldots)$. By Theorem 11 there is a $*$-cross section γ for π on $C^*(p_1, p_2, \ldots)$. Clearly, the restriction of g to $\mathcal{A}$ is a $*$-cross section of π for $\mathcal{A}$. □

Theorem 29. *Let A be a unital C^*-algebra. If A is W^*-C^*-projective* (W^*-W^*-*projective respectively*) *in the nonunital category, then for each $n \in \mathbb{N}$, $M_n(A)$ is W^*-C^*-projective* (W^*-W^*-*projective respectively*) *in the nonunital category.*

Proof. Our proof is a modification of Loring's proof [1997] of the fact that the class of projective C^*-algebras is closed under tensoring with matrices.

Let $\phi : M_n \otimes A \to B/I$ be a $*$-homomorphism and B (and B/I, for the W^*-W^*-projectivity case) be a von Neumann algebra and let $\pi : B \to B/I$ denote the canonical surjection. We need to prove that ϕ lifts. Define $j : M_n \to M_n \otimes A$ by

$$j(T) = T \otimes 1_A$$

and let $\phi_2 = \phi \circ j$. Since by Corollary 18 M_n is W^*-C^*-projective in the nonunital category, ϕ_2 lifts to $\psi : M_n \to B$.

$$\begin{tikzcd} M_n \arrow[r, "\psi"] \arrow[d, "j"] \arrow[dr, "\phi_2"] & B \arrow[d, "\pi"] \\ M_n \otimes A \arrow[r, "\phi"'] & B/I \end{tikzcd}$$

Let (e_{ij}) be a matrix unit in M_n. Define a $*$-homomorphism

$$i : M_n \otimes \phi(e_{11} \otimes A) \to \phi(M_n \otimes A)$$

by $i(T \otimes \phi(e_{11} \otimes a)) = \phi(T \otimes a)$. It is obviously surjective. To see that it is injective, we will use the fact that an ideal in a tensor product C^*-algebra is a tensor product of ideals. Hence the kernel of i is either 0 or of the form $M_n \otimes J$, where J is an ideal in $\phi(e_{11} \otimes A)$. Let $\phi(e_{11} \otimes a) \in J$. Then for each $T \in M_n$,

$\phi(T\otimes a)=T\otimes\phi(e_{11}\otimes a)=0$. In particular, $\phi(e_{11}\otimes a)=0$. Thus $\phi(e_{11}\otimes a)=0$ and $J=0$. So i is an isomorphism. Let

$$p=\phi_2(e_{11}),\qquad P=\psi(e_{11}),$$

and let i_1 be the inclusion $\phi(e_{11}\otimes A)\subseteq pB/Ip=PBP/PIP$. Then the composition $(\mathrm{id}_{M_n}\otimes i_1)\circ i^{-1}\circ\phi : M_n\otimes A\to M_n\otimes PBP/PIP$ is of the form $\mathrm{id}_{M_n}\otimes\gamma$, where $\gamma : A\to PBP/PIP$ is defined by

$$\gamma(a)=p\phi(e_{11}\otimes a)p.$$

Since PBP (and pB/Ip, for the W^*-W^*-projectivity case) is a von Neumann algebra, by W^*-C^*-projectivity (W^*-W^*-projectivity) of A, it can be lifted to

$$\mathrm{id}_{M_n}\otimes\tilde{\gamma} : M_n\otimes A\to M_n\otimes PBP.$$

$$\begin{tikzcd}
 & & & M_n\otimes PBP \arrow[d, "\mathrm{id}_{M_n}\otimes\pi|_{PBP}"] \\
M_n\otimes A \arrow[r, "\phi"'] \arrow[urrr, "\mathrm{id}_{M_n}\otimes\tilde{\gamma}"] & \phi(M_n\otimes A) \arrow[r, "i^{-1}"'] & M_n\otimes\phi(e_{11}\otimes A) \arrow[r, "\mathrm{id}_{M_n}\otimes i_1"'] & M_n\otimes pB/Ip
\end{tikzcd}$$

Now we are going to embed $M_n\otimes PBP$ back into B and $M_n\otimes pB/Ip$ back into B/I. Define

$$\tilde{\alpha} : M_n\otimes PBP\to B\quad\text{and}\quad\alpha : M_n\otimes pB/Ip\to B/I$$

by

$$\tilde{\alpha}(e_{ij}\otimes PbP)=\psi(e_{i1})b\psi(e_{1j})\quad\text{and}\quad\alpha(e_{ij}\otimes p\pi(b)p)=\phi_2(e_{i1})\pi(b)\phi_2(e_{1j})$$

respectively, for each $b\in B$. It is straightforward to check that

$$\alpha\circ(\mathrm{id}_{M_n}\otimes i_1)\circ i^{-1}\circ\phi=\phi$$

and that the diagram

$$\begin{tikzcd}
M_n\otimes PBP \arrow[r, "\tilde{\alpha}"] \arrow[d, "\mathrm{id}_{M_n}\otimes\pi|_{PBP}"] & B \arrow[d, "\pi"] \\
M_n\otimes pB/Ip \arrow[r, "\alpha"'] & B/I
\end{tikzcd}$$

commutes. It follows that $\tilde{\alpha}\circ(\mathrm{id}_{M_n}\otimes\tilde{\gamma})$ is a lift of ϕ. □

Remark 30. We did not consider the C^*-W^* case in the theorem because no unital C^*-algebra is C^*-W^*-projective in the nonunital category. Otherwise $A\oplus\mathbb{C}$ would be unital and C^*-W^*-projective, which would contradict Proposition 3 since $A\oplus\mathbb{C}$ has a nontrivial projection.

Recall that a C^*-algebra is *subhomogeneous* if there is an upper bound for the dimensions of its irreducible representations.

Theorem 31. *Let A be a separable subhomogeneous C^*-algebra. Then A is W^*-W^*-projective in the nonunital category.*

Proof. Suppose $\mathcal{B}$ and $\mathcal{M}$ are von Neumann algebras and $\pi : \mathcal{B} \to \mathcal{M}$ is a surjective $*$-homomorphism. Let $\phi : A \to \mathcal{M}$ be a $*$-homomorphism. If A is nonunital, we can extend ϕ to a homomorphism from the unitization of A to $\mathcal{M}$. It implies that it will be sufficient to prove the theorem under the assumption that A is unital. Since $\mathcal{M} \subseteq B(H)$, by the universal property of the second dual there exists $\tilde{\phi} : A^{**} \to B(H)$ such that $\tilde{\phi}|_A = \phi$ and $\tilde{\pi}(A^{**}) = \pi(A)''$. Hence $\tilde{\phi}$ is a $*$-homomorphism from A^{**} to M. It can be easily deduced from some well-known properties of subhomogeneous algebras (see for instance [Shulman and Uuye 2012], Lemmas 2.3 and 2.4) that A^{**} can be written as

$$A^{**} = \oplus_{k=1}^{n} M_k(D_k),$$

where $D_k, k \leq n$, are abelian von Neumann algebras. Let $\pi_k : A^{**} \to M_k(D_k)$ be the projection on the k-th summand. Let

$$F_k = \{b \in D_k \mid \text{ there exists } a \in A \text{ such that b is a matrix element of } \pi_k(a)\},$$

for each $k \leq n$. Let E_k denote the C^*-subalgebra of D_k generated by F_k, for each $k \leq n$. Then each E_k is separable and $A \subseteq \oplus_{k=1}^{n} M_k(E_k) \subseteq A^{**}$. By Theorems 28, 29 and Lemma 19, $\tilde{\phi}|_{\oplus_{k=1}^{n} M_k(E_k)}$ lifts to some $*$-homomorphism $\psi : \oplus_{k=1}^{n} M_k(E_k) \to \mathcal{B}$. The restriction of ψ onto A is a lift of ϕ. □

The following are easy observations.

Proposition 32. *Suppose $\mathcal{A}$ is a separable unital W^*-W^*- projective C^*-algebra.*

(1) *If $\mathcal{A}$ is nuclear and has a faithful trace, then it must be RFD.*

(2) *If Connes' embedding problem has an affirmative answer, then every unital W^*-W^*-projective C^*-algebra with a faithful trace is RFD.*

The converse of the previous proposition is not true. Indeed in [Hadwin and Shulman 2018b] we constructed a nuclear RFD C^*-algebra which is not matricially tracially stable and hence is not W^*-W^*-projective. Below we give an example which is not only nuclear but is even AF. Our arguments of why it is not matricially tracially stable are much simpler than the ones in [Hadwin and Shulman 2018b].

Theorem 33. *There exists an AF RFD C^*-algebra which is not matricially tracially stable and hence is not W^*-W^*-projective (in both unital and nonunital categories).*

Proof. Suppose $\mathcal{A}$ and $\mathcal{B}$ are separable unital AF-C^*-algebras. Suppose $\mathcal{A} = C^*(a_1, a_2, \ldots)$ and $\mathcal{B} = C^*(b_1, b_2, \ldots)$ with each a_n and b_n selfadjoint. We can assume that $\sigma(a_1) \subset [0,1]$ and $\sigma(b_1) \subset [4,5]$. Then we can find, for each $n \in \mathbb{N}$, a finite-dimensional C^*-subalgebra $\mathcal{A}_n$ of $\mathcal{A}$ and elements $a_{1,n}, \ldots, a_{n,n} \in \mathcal{A}_n$ such that $\|a_k - a_{k,n}\| < 1/n$ for $1 \le k \le n$. Similarly, we can find, for each $n \in \mathbb{N}$ a finite-dimensional C^*-subalgebra $\mathcal{B}_n$ of $\mathcal{B}$ and elements $b_{1,n}, \ldots, b_{n,n} \in \mathcal{B}_n$ such that $\|b_k - b_{k,n}\| < 1/n$ for $1 \le k \le n$. We can also assume that $\sigma(a_{1,n}) \subset [-1,2]$ and $\sigma(b_{1,n}) \subset [3,6]$ for every $n \in \mathbb{N}$. We can assume, for each $n \in \mathbb{N}$, that $\mathcal{A}_n, \mathcal{B}_n \subset \mathcal{M}_{s_n}(\mathbb{C})$ (unital embeddings). For each $1 \le k \le n < \infty$ define

$$c_{k,n} = a_{k,n}^{(n)} \oplus b_{k,n} \in \mathcal{M}_{(n+1)s_n}(\mathbb{C}).$$

Define $c_{k,n} = 0$ when $1 \le n < k < \infty$. Let $C_k = \sum_{n\in\mathbb{N}}^{\oplus} c_{k,n} \in \prod_{n\in\mathbb{N}} \mathcal{M}_{(n+1)s_n}(\mathbb{C})$ and define the C^*-algebra C as the C^*-algebra generated by $C_1, C_2, \ldots$ and $\mathcal{J} = \sum_{n\in\mathbb{N}}^{\oplus} \mathcal{M}_{(n+1)s_n}(\mathbb{C})$. Clearly, $\mathcal{C}$ is RFD and

$$\mathcal{C}/\mathcal{J} \cong C^*(a_1 \oplus b_1, a_2 \oplus b_2, \ldots) \subseteq \mathcal{A} \oplus \mathcal{B}.$$

However, if $f : \mathbb{R} \to \mathbb{R}$ is continuous and $f = 0$ on $[0,1]$ and $f = 1$ on $[2,3]$, we have $f(a_1 \oplus b_1) = 0 \oplus 1$. Thus $0 \oplus 1 \in C^*(a_1 \oplus b_1, a_2 \oplus b_2, \ldots)$ and hence $C^*(a_1 \oplus b_1, a_2 \oplus b_2, \ldots) = \mathcal{A} \oplus \mathcal{B}$. Since $\mathcal{J}$ and $\mathcal{C}/\mathcal{J}$ are AF, $\mathcal{C}$ must be AF. Now, to get an example that we wanted, suppose $\mathcal{A} = \mathcal{B} = \mathcal{M}_{2^\infty}$ with trace τ. Let $\pi = \pi_1 \oplus \pi_2 : \mathcal{C} \to \mathcal{A} \oplus \mathcal{B}$ be the map whose kernel is $\mathcal{J}$. Then $\rho = \tau \circ \pi_2$ is a tracial state on $\mathcal{C}$. Note that since $\mathcal{J} \subset \mathcal{C}$, the only irreducible finite-dimensional representations of $\mathcal{C}$ are (unitarily equivalent to) the coordinate projections onto $\mathcal{M}_{(n+1)s_n}(\mathbb{C})$ and, for each of these representations the trace of the image of $f(C_1) = \sum_{n\in\mathbb{N}}^{\oplus} 0^{(n)} \oplus 1$ is at most $1/2$. However, $\rho(f(C_1)) = 1$. Thus ρ is not a weak*-limit of finite-dimensional traces. By [Hadwin and Shulman 2018b, Theorem 3.10], $\mathcal{C}$ is not matricially tracially stable. □

Acknowledgements

Hadwin was supported by a Collaboration Grant from the Simons Foundation. The research of Shulman was supported by the Polish National Science Centre grant under the contract number DEC-2012/06/A/ST1/00256, by the grant H2020-MSCA-RISE-2015-691246-QUANTUM DYNAMICS and Polish Government grant 3542/H2020/2016/2, and from the Eric Nordgren Research Fellowship Fund at the University of New Hampshire.

References

[Arzhantseva and Păunescu 2015] G. Arzhantseva and L. Păunescu, "Almost commuting permutations are near commuting permutations", *J. Funct. Anal.* **269**:3 (2015), 745–757. MR Zbl

[Blackadar 1985] B. Blackadar, "Shape theory for C^*-algebras", *Math. Scand.* **56**:2 (1985), 249–275. MR Zbl

[Brown and Pedersen 1991] L. G. Brown and G. K. Pedersen, "C^*-algebras of real rank zero", *J. Funct. Anal.* **99**:1 (1991), 131–149. MR Zbl

[Chigogidze and Dranishnikov 2010] A. Chigogidze and A. N. Dranishnikov, "Which compacta are noncommutative ARs?", *Topology Appl.* **157**:4 (2010), 774–778. MR Zbl

[Davidson 1996] K. R. Davidson, *C^*-algebras by example*, Fields Institute Monographs **6**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Eilers et al. 1998] S. Eilers, T. A. Loring, and G. K. Pedersen, "Stability of anticommutation relations: an application of noncommutative CW complexes", *J. Reine Angew. Math.* **499** (1998), 101–143. MR Zbl

[Friis and Rørdam 1996] P. Friis and M. Rørdam, "Almost commuting self-adjoint matrices — a short proof of Huaxin Lin's theorem", *J. Reine Angew. Math.* **479** (1996), 121–131. MR Zbl

[Hadwin 2014] D. Hadwin, "A lifting characterization of RFD C*-algebras", *Math. Scand.* **115**:1 (2014), 85–95. MR Zbl

[Hadwin and Shulman 2018a] D. Hadwin and T. Shulman, "Stability of group relations under small Hilbert–Schmidt perturbations", *J. Funct. Anal.* **275**:4 (2018), 761–792. MR Zbl

[Hadwin and Shulman 2018b] D. Hadwin and T. Shulman, "Tracial stability for C^*-algebras", *Integral Equations Operator Theory* **90**:1 (2018), Art. 1, 35. MR Zbl

[Hocking and Young 1961] J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, Reading, MA, 1961. MR Zbl

[Kaplansky 1951] I. Kaplansky, "Projections in Banach algebras", *Ann. of Math.* (2) **53** (1951), 235–249. MR Zbl

[Kirchberg and Rørdam 2014] E. Kirchberg and M. Rørdam, "Erratum to Central sequence C^*-algebras and tensorial absorption of the Jiang–Su algebra", *J. Reine Angew. Math.* **695** (2014), 215–216. MR Zbl

[Li et al. 2014] Q. Li, D. Hadwin, W. Li, and J. Shen, "MF-traces and a lower bound for the topological free entropy dimension in unital C^*-algebras", *J. Operator Theory* **71**:1 (2014), 15–44. MR Zbl

[Lin 1997] H. Lin, "Almost commuting selfadjoint matrices and applications", pp. 193–233 in *Operator algebras and their applications* (Waterloo, ON, 1994/1995), edited by P. A. Fillmore and J. A. Mingo, Fields Inst. Commun. **13**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Loring 1997] T. A. Loring, *Lifting solutions to perturbing problems in C^*-algebras*, Fields Institute Monographs **8**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Loring and Shulman 2014] T. Loring and T. Shulman, "Lifting algebraic contractions in C^*-algebras", pp. 85–92 in *Algebraic methods in functional analysis*, edited by I. G. Todorov and L. Turowska, Oper. Theory Adv. Appl. **233**, Springer, 2014. MR Zbl

[Olsen 1971] C. L. Olsen, "A structure theorem for polynomially compact operators", *Amer. J. Math.* **93** (1971), 686–698. MR Zbl

[Sato 2011] Y. Sato, "Discrete amenable group actions on von Neumann algebras and invariant nuclear C^*-subalgebras", preprint, 2011. arXiv

[Shulman and Uuye 2012] T. Shulman and O. Uuye, "Approximations of subhomogeneous algebras", preprint, 2012. arXiv

[Toms et al. 2015] A. S. Toms, S. White, and W. Winter, "$\mathcal{Z}$-stability and finite-dimensional tracial boundaries", *Int. Math. Res. Not.* **2015**:10 (2015), 2702–2727. MR Zbl

[Winter and Zacharias 2009] W. Winter and J. Zacharias, "Completely positive maps of order zero", *Münster J. Math.* **2** (2009), 311–324. MR Zbl

Received September 13, 2017.

DON HADWIN
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW HAMPSHIRE
DURHAM
UNITED STATES
operatorguy@gmail.com

TATIANA SHULMAN
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSAW
POLAND
tshulman@impan.pl

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.441

LOWER SEMICONTINUITY OF THE ADM MASS IN DIMENSIONS TWO THROUGH SEVEN

JEFFREY L. JAUREGUI

The semicontinuity phenomenon of the ADM mass under pointed (i.e., local) convergence of asymptotically flat metrics is of interest because of its connections to nonnegative scalar curvature, the positive mass theorem, and Bartnik's mass-minimization problem in general relativity. We extend a previously known semicontinuity result in dimension three for C^2 pointed convergence to higher dimensions, up through seven, using recent work of S. McCormick and P. Miao (which itself builds on the Riemannian Penrose inequality of H. Bray and D. Lee). For a technical reason, we restrict to the case in which the limit space is asymptotically Schwarzschild. In a separate result, we show that semicontinuity holds under weighted, rather than pointed, C^2 convergence, in all dimensions $n \geq 3$, with a simpler proof independent of the positive mass theorem. Finally, we also address the two-dimensional case for pointed convergence, in which the asymptotic cone angle assumes the role of the ADM mass.

1. Introduction

Motivated by the Bartnik minimal mass extension conjecture in general relativity [1989; 1997; 2002], as well as the study of Ricci flow on asymptotically flat manifolds [Dai and Ma 2007; Oliynyk and Woolgar 2007], in [Jauregui 2018] the author established the following result regarding how the ADM mass behaves under pointed convergence of a sequence of asymptotically flat 3-manifolds of nonnegative scalar curvature. Briefly, the ADM mass cannot increase in a local C^2 limit:

Theorem 1 [Jauregui 2018]. *Let (M_i, g_i, p_i) be a sequence of pointed asymptotically flat 3-manifolds without boundary, such that each (M_i, g_i) has nonnegative scalar curvature and contains no compact minimal surfaces. If (M_i, g_i, p_i) converges in the pointed C^2 Cheeger–Gromov sense to a pointed asymptotically flat 3-manifold (N, h, q), then*

$$m_{\mathrm{ADM}}(N, h) \leq \liminf_{i \to \infty} m_{\mathrm{ADM}}(M_i, g_i). \tag{1}$$

MSC2010: 53C20, 53C80, 83C99.
Keywords: scalar curvature, mass in general relativity.

We recall the relevant definitions in Section 3; for now we note that pointed C^k Cheeger–Gromov convergence essentially means C^k convergence of the metric tensors on compact subsets, modulo diffeomorphisms. Examples are given in [Jauregui 2018] in which strictness holds in (1).

Theorem 1 is intimately connected to scalar curvature and to the positive mass theorem (PMT) [Schoen and Yau 1979a; Witten 1981]. In [Jauregui 2018] it was shown that (1) can fail without assuming nonnegative scalar curvature (and the absence of compact minimal surfaces). Somewhat surprisingly, a simple blow-up example in [Jauregui 2018] shows that Theorem 1 actually implies the PMT. However, to prove Theorem 1, either the PMT itself, or a stronger result, is required. The key estimate in the proof of Theorem 1 was the lower bound

$$m_{\mathrm{ADM}} \geq m_H(\Sigma) \tag{2}$$

of the ADM mass in terms of the Hawking mass of an outward-minimizing surface Σ, established by G. Huisken and T. Ilmanen [2001]. Note that it is well known that (2) implies the PMT.

Two major questions were left unsettled in [Jauregui 2018]. First, to what extent does this lower semicontinuity property of the ADM mass hold for weaker convergence than C^2? Subsequently the author and D. Lee proved in [Jauregui and Lee 2017] that the theorem continues to hold if only pointed C^0 convergence is assumed. Second, does Theorem 1 generalize to higher dimensions? The primary concern of the present paper is to address the latter question.

Unfortunately, a bound directly analogous to (2) is unknown beyond dimension three: Huisken–Ilmanen's proof in $n = 3$ uses "Geroch monotonicity" of the Hawking mass, which crucially relies on the Gauss–Bonnet theorem in one dimension lower. Generally, the missing link in establishing Theorem 1 in higher dimensions has been a useful quantitative lower bound for the ADM mass in terms of the geometry of an outward-minimizing surface. Fortunately, a recent result of S. McCormick and P. Miao [2017] provides such an estimate (see Theorem 7 below) that is sufficient for our purposes. Their work uses the Riemannian Penrose inequality in higher dimensions, due to Bray and Lee [2009] (which itself was a generalization of Bray's original proof in dimension three [2001]). Our main result is:

Theorem 2. *Theorem 1 is true with "3" replaced by "n", where $3 \leq n \leq 7$, provided the limit (N, h) is asymptotically Schwarzschild.*

The Riemannian manifolds (M_i, g_i) need not be asymptotically Schwarzschild even if their limit (N, h) is.

The restriction in Theorem 2 of $n \leq 7$ is primarily due to the fact that it is the highest dimension in which the Riemannian Penrose inequality is currently known. It was pointed out in [Bray and Lee 2009] that even the positive mass theorem for $n \geq 8$ is insufficient to automatically extend the Riemannian Penrose inequality to $n \geq 8$. We strongly conjecture that the $n \leq 7$ restriction is unnecessary, and that the asymptotically Schwarzschild hypothesis can be replaced with asymptotic flatness; see Remark D.

Remark A. It is reasonable to attempt to extend Theorem 1 to spin manifolds in higher dimensions using Witten's spinor technique in his proof of the PMT [1981]. However, as pointed out to the author by Bray, it is not clear how to make effective use of the hypothesis of no compact minimal surfaces in the spinor argument, and it was shown in [Jauregui 2018] that (1) can fail without this hypothesis.

For the purpose of telling a more complete story, we also include two other related results. First, assuming weighted (rather than pointed) C^2 convergence, we prove lower semicontinuity of the ADM mass in all dimensions $n \geq 3$ (Theorem 13 below). Weighted convergence assumes global control on the asymptotics of the metrics, in contrast to pointed convergence. In this case, with a stronger hypothesis than in Theorems 1 and 2, the absence of compact minimal surfaces is unnecessary and the proof is easier. However, the weighted result does not recover nor rely on the positive mass theorem. Prior results for weighted convergence were known; see Section 6 (in particular Remark E) for details.

Second, it was suggested by E. Woolgar that the author investigate the lower semicontinuity of "mass" in dimension two. This is carried out in Section 7 for pointed C^2 convergence, where the asymptotic cone angle replaces the ADM mass; see Theorem 14.

2. Motivation and examples

In this section we describe several examples to motivate the lower semicontinuity phenomenon for the ADM mass.

2.1. ***Lower semicontinuity of mass in Newtonian gravity.*** We begin here with a general discussion of why lower semicontinuity of the total mass is plausible from the point of view of Newtonian gravity. Consider a matter distribution on $\mathbb{R}^n$ described by a continuous, integrable mass density function $\rho \geq 0$. The total Newtonian mass is simply given by the integral

$$m(\rho) = \int_{\mathbb{R}^n} \rho \, dx^1 \cdots dx^n.$$

Now, if $\{\rho_i\}_{i=1}^{\infty}$ is a sequence of such matter distributions that converges pointwise to ρ, then by Fatou's lemma,

$$\liminf_{i\to\infty} m(\rho_i) \geq m(\rho).$$

Any drop in the total Newtonian mass can be viewed as mass escaping out to infinity in the limit. Such an argument does not apply to the context of general relativity, because the ADM mass is not known (or expected) to be given as the integral of a locally defined, nonnegative, geometric/physical quantity.

Convergence of ρ_i to ρ in Newtonian gravity is analogous to C^2 convergence of the Riemannian metrics in general relativity, as the scalar curvature represents energy density and is given by two derivatives of the metric. The C^0 convergence in [Jauregui and Lee 2017] can then be viewed as a general relativistic analog of convergence of the Newtonian gravitational potentials $u_i \to u$, where $\Delta u_i = 4\pi\rho_i$ and $\Delta u = 4\pi\rho$.

2.2. ***Blow-up example.*** In [Jauregui 2018], the author gave the example of a fixed asymptotically flat n-manifold (M, g) of nonnegative scalar curvature and considered the sequence of homothetic rescalings $\{(M, i^2 g, p)\}$ for $p \in M$ fixed and $i = 1, 2, \dots$. This sequence converges in the pointed C^2 Cheeger–Gromov sense to Euclidean $\mathbb{R}^n$ (which has zero ADM mass), and indeed the statement of lower semicontinuity of mass implies that the ADM mass of (M, g) is nonnegative. In other words, the positive mass theorem is recovered.

The example in Section 2.1 suggests that from a Newtonian point of view, the mass-drop phenomenon can be completely accounted for by matter escaping off to infinity. But by choosing (M, g) here to be scalar-flat (i.e., vacuum) with positive ADM mass, the example of $\{(M, i^2 g, p)\}$ converging to Euclidean space shows that the mass can drop by an infinite amount in the limit with no matter fields present. This can be interpreted as the energy of the gravitational field escaping to infinity.

2.3. ***Escaping point example.*** Similar to the previous example, begin with a fixed asymptotically flat n-manifold (M, g). Now consider a sequence of points $\{p_i\}$ in M escaping to infinity. By asymptotic flatness, the sequence $\{(M, g, p_i)\}$ converges in the pointed C^2 Cheeger–Gromov sense to Euclidean $\mathbb{R}^n$. Again the statement of lower semicontinuity of ADM mass here recovers the positive mass theorem; and again by choosing (M, g) to be scalar-flat with positive ADM mass we can interpret the mass drop as gravitational energy escaping to infinity.

2.4. ***Lower semicontinuity of mass and Ricci flow.*** To the author's knowledge, the ADM mass drop phenomenon under pointed convergence was first observed by T. Oliynyk and Woolgar in their study of Ricci flow on rotationally symmetric, asymptotically flat spaces [2007]; see also the work of X. Dai and L. Ma, who first showed that the ADM mass is constant along Ricci flow, thereby arguing an asymptotically flat Ricci flow cannot converge uniformly to Euclidean space [Dai and Ma 2007]. Under natural hypotheses, Oliynyk and Woolgar proved the long-time existence of Ricci flow on asymptotically flat, rotationally symmetric spaces, with pointed C^k Cheeger–Gromov convergence to Euclidean space as $t \to \infty$.

Moreover, the ADM mass is not only monotone but is in fact constant along the Ricci flow. In particular, if the initial space has positive ADM mass, then the ADM mass must drop to zero in the limit.

In light of this discussion, the author suggested in [Jauregui 2018] that using Theorem 1 (or its higher-dimensional analog) would be necessary in any proof of the PMT that involved convergence of the Ricci flow to Euclidean space. Since Theorem 1 already subsumes the PMT, this seemed to suggest that an independent Ricci flow proof of the PMT was unlikely. Nevertheless, such a proof has very recently been given by Y. Li in [2018]. His argument circumvents this apparent circular logic by establishing lower semicontinuity of the ADM mass directly for the case of a convergent Ricci flow (i.e., the technique does not apply to general pointed C^2 Cheeger–Gromov convergence). We generalize Li's argument to weighted C^2 convergence in Section 6.

3. Background

We begin with the definition of an asymptotically flat manifold (with one end). Many slight variants appear in the literature; the version below is commonly used.

Definition 3. A smooth, connected Riemannian n-manifold (M, g), with $n \geq 3$, possibly with compact boundary, is *asymptotically flat (AF)* if there exists a compact set $K \subset M$ and a diffeomorphism $\Phi : M \setminus K \to \mathbb{R}^n \setminus B$, for a closed ball B, such that in the "asymptotically flat" coordinates $x = (x^1, \dots, x^n)$ given by Φ, we have

$$g_{ij} = \delta_{ij} + O(|x|^{-\tau}), \quad \partial_k g_{ij} = O(|x|^{-\tau-1}), \quad \partial_k \partial_\ell g_{ij} = O(|x|^{-\tau-2}), \tag{3}$$

for some constant $\tau > \frac{n-2}{2}$ (the *order*), and the scalar curvature of g is integrable. (Indices i, j, k, ℓ above run from 1 to n, and ∂ denotes partial differentiation in the coordinate chart.)

For example, for a real number $m > 0$, the Schwarzschild metric

$$g_{ij} = \left(1 + \frac{m}{2|x|^{n-2}}\right)^{\frac{4}{n-2}} \delta_{ij}$$

on $\mathbb{R}^n$ minus a ball about the origin is asymptotically flat of order $n-2$.

We will also need two classes of asymptotically flat manifolds with more restricted asymptotics at infinity:

Definition 4. An asymptotically flat Riemannian n-manifold (M, g) is *asymptotically Schwarzschild* if there exists an "asymptotically Schwarzschild coordinate system" $(x^1, \dots, x^n)$ on $M \setminus K$, i.e.,

$$g_{ij} = \left(1 + \frac{m}{2|x|^{n-2}}\right)^{\frac{4}{n-2}} \delta_{ij} + h_{ij}, \tag{4}$$

for some real constant m, where

$$h_{ij} = O(|x|^{1-n}), \quad \partial_k h_{ij} = O(|x|^{-n}), \quad \partial_k \partial_\ell h_{ij} = O(|x|^{-n-1}). \tag{5}$$

Note that an asymptotically Schwarzschild Riemannian n-manifold is AF of order $n-2$.

Definition 5. An asymptotically flat Riemannian n-manifold (M, g) is *harmonically flat at infinity (HF)* if there exists a "harmonically flat coordinate system" $(x^1, \dots, x^n)$ on $M \setminus K$, i.e.,

$$g_{ij} = U^{\frac{4}{n-2}} \delta_{ij},$$

on $M \setminus K$ for some function U, where $\Delta U = 0$ and $U(x) \to 1$ as $|x| \to \infty$. (Here Δ is the Euclidean Laplacian on $\mathbb{R}^n$.)

It is well known that the harmonic function U appearing in Definition 5 admits an expansion at infinity of the form

$$U(x) = 1 + \frac{a}{|x|^{n-2}} + O_\infty(|x|^{-n+1}), \tag{6}$$

where the notation $O_k(|x|^\ell)$ denotes an expression that is $O(|x|^\ell)$ for $|x|$ large and for which the γth partial derivative (γ being a multi-index with $|\gamma| \le k$) is $O(|x|^{\ell - |\gamma|})$. The fact that $\Delta U = 0$ implies that g as above has zero scalar curvature outside of K. Note that HF manifolds are necessarily asymptotically Schwarzschild, and that the Schwarzschild metric itself is HF.

Next, we recall the definition of ADM mass.

Definition 6. The *ADM mass* [Arnowitt et al. 1961] (cf. [Bartnik 1986; Chruściel 1986]) of an asymptotically flat manifold (M, g) of dimension n is the real number

$$m_{\mathrm{ADM}}(M, g) = \frac{1}{2(n-1)\omega_{n-1}} \lim_{r \to \infty} \int_{S_r} \sum_{i,j=1}^n (\partial_i g_{ij} - \partial_j g_{ii}) \frac{x^j}{r} \, dA,$$

where dA is the induced volume form on the coordinate sphere

$$S_r = \{|x| = r\}$$

with respect to the Riemannian metric δ_{ij}, all in an AF coordinate chart.

It is straightforward to verify that for an HF manifold, the ADM mass is given by the value $2a$, where a is the constant appearing in (6), and for an asymptotically Schwarzschild manifold, the ADM mass is given by the constant m appearing in (5).

Recall that if (M, g) is asymptotically flat with boundary ∂M, then we say ∂M is *outward-minimizing* if

$$|S| \ge |\partial M|$$

for all surfaces S enclosing ∂M, where $|\cdot|$ denotes the hypersurface area (with respect to g). The following theorem was recently proved by McCormick and Miao [2017].

Theorem 7 [McCormick and Miao 2017]. *Let (M, g) be an AF manifold of dimension $3 \leq n \leq 7$, with compact, connected boundary Σ that is outward-minimizing. Assume that the scalar curvature of (M, g) is nonnegative. Let $H \geq 0$ be the mean curvature of Σ (in the direction pointing into M), let ρ be the scalar curvature of Σ with respect to the induced Riemannian metric, and suppose that*

$$\min_{\Sigma} \rho > \frac{n-2}{n-1} \max_{\Sigma} H^2.$$

Then

$$m_{\mathrm{ADM}}(M, g) \geq \frac{1}{2}\left(\frac{|\Sigma|}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}}\left(1 - \frac{n-2}{n-1}\frac{\max_{\Sigma} H^2}{\min_{\Sigma} \rho}\right). \tag{7}$$

To simplify notation later, we make the following definition.

Definition 8. Let S be a smooth, compact hypersurface in a Riemannian manifold (M, g) of dimension $n \geq 3$. Define

$$F_g(S) = \frac{1}{2}\left(\frac{|S|}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}}\left(1 - \frac{n-2}{n-1} \cdot \frac{\max_S H^2}{\min_S \rho}\right),$$

where $|S|$, H, and ρ are the area, mean curvature, and scalar curvature of Σ with respect to the Riemannian metric induced by g.

We conclude this section with the definition of convergence used in Theorems 1 and 2.

Definition 9. Fix a nonnegative integer ℓ. A sequence of complete, connected, pointed Riemannian n-manifolds (M_i, g_i, p_i) converges in the *pointed C^ℓ Cheeger–Gromov sense* to a complete, connected, pointed Riemannian n-manifold (N, h, q) if for every $r > 0$ there exists a domain Ω containing the metric ball $B_h(q, r)$ in (N, h), and there exist (for all i sufficiently large) smooth embeddings

$$\Phi_i : \Omega \to M_i$$

such that $\Phi_i(\Omega)$ contains the metric ball $B_{g_i}(p_i, r)$, and the Riemannian metrics $\Phi_i^* g_i$ converge in C^ℓ norm to h as tensors on Ω.

Note that no M_i need be diffeomorphic to N in the above definition, and that the asymptotics of M_i can be wildly different from those of N in the noncompact case.

4. The mass of asymptotically Schwarzschild metrics

In this section we prove that the ADM mass of an asymptotically Schwarzschild manifold can be recovered from the $r \to \infty$ limit of the expression $F_g(S_r)$, a key ingredient in the proof of Theorem 2. Before doing so (in Lemma 11), we first verify this for HF metrics in Lemma 10.

Remark B. For an asymptotically flat manifold (M, g) of dimension $3 \le n \le 7$, the inequality

$$m_{\mathrm{ADM}}(M, g) \ge \limsup_{r\to\infty} F_g(S_r)$$

follows from Theorem 7. However, equality need not hold. Such an example, pointed out to the author by McCormick, can be found by considering an AF manifold (M, g) of nonnegative scalar curvature and strictly positive ADM mass that contains an isometric copy of half of a Euclidean space. Such spaces were constructed by Carlotto and Schoen [2016]. For r sufficiently large, S_r intersects the Euclidean region in M, which gives $F_g(S_r) \le 0$.

Lemma 10. *If (M, g) is an HF manifold, then*

$$m_{\mathrm{ADM}}(M, g) = \lim_{r\to\infty} F_g(S_r), \tag{8}$$

where F_g is given in Definition 8, and S_r is the coordinate sphere $\{|x| = r\}$ in a harmonically flat coordinate system.

Except for the calculations (9) at the end of the following proof, the proof of Lemma 11 will be independent of Lemma 10.

Proof. The proof involves straightforward computations of the asymptotic behavior, for large r, of the area, mean curvature, and scalar curvature of S_r. Let U be the harmonic function as in Definition 5, with expansion (6).

First we compute the area of S_r:

$$\begin{aligned}
|S_r|_g &= \int_{S_r} U^{\frac{2(n-1)}{n-2}}\, dA \\
&= \int_{S_r} \left(1 + \frac{2a(n-1)}{(n-2)r^{n-2}} + O(r^{1-n})\right) dA \\
&= \omega_{n-1} r^{n-1}\left(1 + \frac{2a(n-1)}{(n-2)r^{n-2}}\right) + O(1),
\end{aligned}$$

where dA is the area form on S_r induced by δ. In particular,

$$\frac{1}{2}\left(\frac{|S_r|_g}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}} = \frac{1}{2} r^{n-2}\left(1 + \frac{2a}{r^{n-2}}\right) + O(r^{-1}).$$

Second we compute the mean curvature. Recall that the mean curvature of S_r with respect to δ_{ij} is $\frac{n-1}{r}$. From a well-known formula relating the mean curvatures of conformally related Riemannian metrics, letting H_r represent the mean curvature of S_r with respect to g, we have

$$\begin{aligned}
H_r &= U^{-\frac{2}{n-2}} \cdot \frac{n-1}{r} + \frac{2(n-1)}{n-2} \cdot U^{-\frac{n}{n-2}} \cdot \nu(U) \\
&= \left(1 + \frac{a}{r^{n-2}} + O(r^{1-n})\right)^{-\frac{2}{n-2}} \cdot \frac{n-1}{r} \\
&\quad + \frac{2(n-1)}{n-2}\left(1 + \frac{a}{r^{n-2}} + O(r^{1-n})\right)^{-\frac{n}{n-2}}\left(-\frac{a(n-2)}{r^{n-1}} + O(r^{-n})\right) \\
&= \left(1 - \frac{2a}{(n-2)r^{n-2}} + O(r^{1-n})\right) \cdot \frac{n-1}{r} \\
&\quad + \frac{2(n-1)}{n-2}\left(1 - \frac{an}{(n-2)r^{n-2}} + O(r^{1-n})\right)\left(-\frac{a(n-2)}{r^{n-1}} + O(r^{-n})\right) \\
&= \frac{n-1}{r} - \frac{2a(n-1)^2}{(n-2)r^{n-1}} + O(r^{-n}),
\end{aligned}$$

where we used the fact that the δ-unit normal ν to S_r equals $\frac{\partial}{\partial r}$. Thus,

$$H_r^2 = \frac{(n-1)^2}{r^2} - \frac{4a(n-1)^3}{(n-2)r^n} + O(r^{-n-1}).$$

Third, we compute the scalar curvature of S_r with respect to $g|_{TS_r}$. Recall that if $g_2 = e^{2\psi} g_1$ are conformally related Riemannian metrics on a manifold of dimension $n-1$, then their scalar curvatures are related by

$$R_{g_2} = e^{-2\psi}(R_{g_1} - 2(n-2)\Delta_{g_1}\psi - (n-3)(n-2)|d\psi|^2_{g_1}).$$

In particular, with $g_2 = g|_{TS_r},\ \ g_1 = \delta|_{TS_r}$, and $U^{\frac{4}{n-2}} = e^{2\psi}$ on S_r, we have

$$\rho = U^{-\frac{4}{n-2}}\left(\frac{(n-1)(n-2)}{r^2} - \frac{4\Delta_r U}{U} + \frac{4|\nabla_r U|^2}{(n-2)U^2}\right),$$

where Δ_r and ∇_r are the Laplacian and (tangential) gradient on S_r with the Riemannian metric induced from δ, and $|\cdot|^2$ is taken with respect to δ. Now, we address the Laplacian term. A well known formula for smooth functions f on $\mathbb{R}^n$ is

$$\Delta f = \Delta_\Sigma f + \mathrm{Hess}(f)(\nu, \nu) + H\partial_\nu(f),$$

where Σ is a smooth hypersurface with unit normal ν, mean curvature H in the direction of ν, and induced Laplacian Δ_Σ. Applying this to $f = U$ and $\Sigma = S_r$, we have

$$0 = \Delta_r U + \mathrm{Hess}(U)(\partial_r, \partial_r) + \frac{n-1}{r} \cdot \frac{\partial U}{\partial r}.$$

By explicit calculation, the leading (i.e., $O(r^{-n})$) terms of $\mathrm{Hess}(U)(\partial_r, \partial_r)$ and $\frac{n-1}{r} \cdot \frac{\partial U}{\partial r}$ cancel, implying that

$$\Delta_r U = O(r^{-n-1}).$$

Next, for the term $|\nabla_r U|$, since $1 + a/(r^{n-2})$ is constant on S_r, we see from the expansion of U that

$$|\nabla_r U|^2 = O(r^{-2n}).$$

Using these expansions, along with the expansion for U, we arrive at

$$\begin{aligned}\rho &= \left(1 + \frac{a}{r^{n-2}} + O(r^{1-n})\right)^{-\frac{4}{n-2}} \left(\frac{(n-1)(n-2)}{r^2} + O(r^{-n-1})\right)\\ &= \frac{(n-1)(n-2)}{r^2} - \frac{4a(n-1)}{r^n} + O(r^{-n-1}).\end{aligned}$$

Putting it all together, we have

$$\begin{aligned}F_g(S_r) &= \left(\frac{1}{2} r^{n-2}\left(1 + \frac{2a}{r^{n-2}}\right) + O(r^{-1})\right)\\ &\qquad\times\left(1 - \frac{n-2}{n-1} \cdot \frac{\frac{(n-1)^2}{r^2} - \frac{4a(n-1)^3}{(n-2)r^n} + O(r^{-n-1})}{\frac{(n-1)(n-2)}{r^2} - \frac{4a(n-1)}{r^n} + O(r^{-n-1})}\right)\\ &= \left(\frac{1}{2} r^{n-2} + a + O(r^{-1})\right)\left(1 - \frac{1 - \frac{4a(n-1)}{(n-2)r^{n-2}} + O(r^{-n+1})}{1 - \frac{4a}{(n-2)r^{n-2}} + O(r^{-n+1})}\right)\\ &= \left(\frac{1}{2} r^{n-2} + a + O(r^{-1})\right)\left(\frac{4a}{r^{n-2}} + O(r^{-n+1})\right)\\ &= 2a + O(r^{-1}).\end{aligned} \tag{9}$$

Since the ADM mass of g equals $2a$, the proof is complete. □

The next lemma is a generalization of the previous one:

Lemma 11. *If (M, g) is an asymptotically Schwarzschild manifold, then*

$$m_{\mathrm{ADM}}(M, g) = \lim_{r\to\infty} F_g(S_r),$$

where S_r is the coordinate sphere $\{|x| = r\}$ in an asymptotically Schwarzschild coordinate system.

Proof. This follows from Lemma 15 in the Appendix and (9). □

5. Proof of Theorem 2

The method of proof of Theorem 2 is similar to the proof of Theorem 1 in [Jauregui 2018].

Let $m_i = m_{\text{ADM}}(M_i, g_i)$, and note that $m_i \geq 0$ by the positive mass theorem in dimension $3 \leq n \leq 7$ ([Schoen and Yau 1979a; 1979b], cf. Section 4 of [Schoen 1989]). If $m_{\text{ADM}}(N, h) = 0$, the claim (1) follows trivially, so we may assume it is strictly positive.

Let $\epsilon > 0$. Fix an asymptotically Schwarzschild coordinate system $(x^1, \ldots, x^n)$ on (N, h), and let S_r denote the coordinate sphere $\{|x| = r\}$, a smooth, compact hypersurface in N for r sufficiently large. Let B_r denote the bounded open region in N that S_r encloses.

By Lemma 11 and the hypothesis that (N, h) is asymptotically Schwarzschild of positive ADM mass, we may choose a number $r_1 > 0$ sufficiently large so that

$$m_{\text{ADM}}(N, h) < F_h(S_{r_1}) + \frac{\epsilon}{2}, \text{ and} \tag{10}$$

$$F_h(S_{r_1}) > 0. \tag{11}$$

By asymptotic flatness of h, we may increase r_1 if necessary, preserving (10) and (11), to arrange that the mean curvature of S_r with respect to h is strictly positive for all $r \geq r_1$, and that hypersurface areas measured with respect to h and the Euclidean metric δ differ by at most a factor of 2 on $N \setminus B_{r_1}$ (i.e., the respective Hausdorff $(n-1)$-measures are uniformly equivalent by factors of 2).

We apply the definition of pointed C^2 Cheeger–Gromov convergence. First, take a number $r_2 > 0$ so that the metric ball $B_h(q, r_2)$ contains B_{33r_1}. (The value $33r_1$ is chosen because later we will need a point in $B_{33r_1} \setminus B_{r_1}$ that is distance $16r_1$ from both the inner and outer boundary.) Then there exists a domain $U \subset N$, with $U \supset B_h(q, r_2) \supset B_{33r_1}$, and smooth embeddings $\Phi_i : U \to M_i$, for $i \geq$ some i_0, with $\Phi_i(U) \supset B_{g_i}(p_i, r_2)$, such that

$$h_i := \Phi_i^* g_i \to h \text{ in } C^2 \text{ on } U. \tag{12}$$

(Below, we will repeatedly use the fact that $\Phi_i : (U, h_i) \to (\Phi_i(U), g_i)$ is trivially an isometry.) Taking i to be at least some $i_1 \geq i_0$, we can be sure that hypersurface areas measured with respect to h_i and h differ by at most a factor of 2 on U, by C^0 convergence. Taking i to be at least some $i_2 \geq i_1$, we can arrange that the mean curvatures of S_r with respect to h_i are strictly positive for all $r \in [r_1, 33r_1]$, using C^1 convergence of h_i to h on U.

Next, let $S_i = \Phi_i(S_{r_1})$, a smooth compact hypersurface in M_i. We want to apply Theorem 7 to the AF manifold-with-boundary obtained by removing $\Phi_i(B_{r_1})$ from M_i (whose boundary is S_i). To do so, we must verify that S_i is outward-minimizing in (M_i, g_i). (This is not at all obvious, since S_i need not even lie in the asymptotically flat end of (M_i, g_i).) This issue was handled in [Jauregui 2018] via a monotonicity formula for minimal surfaces in a Riemannian manifold. However, we will instead use the more robust argument in [Jauregui and Lee 2017], using the notion of almost-minimizing currents.

Lemma 12. *For $i \geq i_2$, S_i is (strictly) outward-minimizing in (M_i, g_i).*

Proof of Lemma 12. It is well known from standard results in geometric measure theory (see [Huisken and Ilmanen 2001] for instance) that there exists a compact hypersurface $\widetilde{S}_i$ enclosing S_i that has the least hypersurface area (with respect to g_i) among all compact hypersurfaces in M_i enclosing S_i. Moreover, $\widetilde{S}_i$ has at least $C^{1,1}$ regularity, and $\widetilde{S}_i \setminus S_i$, if nonempty, is a smooth minimal hypersurface. (This uses $n \leq 7$.) We complete the proof of the lemma by arguing that $\widetilde{S}_i = S_i$, assuming henceforth that $i \geq i_2$.

If $\widetilde{S}_i$ were to possess a connected component disjoint from S_i, then that component would be a compact minimal hypersurface in (M_i, g_i), contrary to the hypothesis of Theorem 2. Thus, every connected component of $\widetilde{S}_i$ intersects S_i.

Next, if $\widetilde{S}_i$ happens to be contained in the compact region $\Phi_i(\bar{B}_{33r_1})$ and hence in $\Phi_i(\bar{B}_{33r_1} \setminus B_{r_1})$, there exists some point $p \in \widetilde{S}_i$ at which the function

$$r \circ \Phi^{-1}|_{\widetilde{S}_i}$$

achieves its maximum on $\widetilde{S}_i$. Say this maximum value is $r^* \in [r_1, 33r_1]$. If $r^* > r_1$, then $\widetilde{S}_i$ is smooth and minimal (with respect to g_i) near p and is tangent to $\Phi_i(S_{r^*})$. However, this contradicts the standard comparison principle for mean curvature, as $\Phi_i(S_{r^*})$ has strictly positive mean curvature with respect to g_i (because S_{r^*} has strictly positive mean curvature with respect to h_i). Thus, $r^* = r_1$, and so $\widetilde{S}_i = S_i$, as claimed.

The only remaining case is that $\widetilde{S}_i$ possesses a connected component, say $\widetilde{S}_i'$, that is not contained in $\Phi_i(\bar{B}_{33r_1} \setminus B_{r_1})$, but that intersects $S_i = \Phi_i(S_{r_1})$. Let

$$T_i = \Phi_i^{-1}(\widetilde{S}_i' \cap \Phi_i(B_{33r_1} \setminus \bar{B}_{r_1})) \subset B_{33r_1} \setminus \bar{B}_{r_1} \subset N.$$

Note that T_i is a smooth hypersurface in the AF end of N, so that we may regard $T_i \subset \mathbb{R}^n$ with $\partial T_i \subset S_{r_1} \cup S_{33r_1}$. By the connectedness of $\widetilde{S}_i'$ and the continuity of r, there exists some point $q_i \in T_i \cap S_{17r_1}$, and the Euclidean distance from q_i to ∂T_i is $16r_1$. Viewing T_i naturally as an $(n-1)$-dimensional integral current in $\mathbb{R}^n$, we claim that T_i is *γ-almost-minimizing* for $\gamma = 16$ (and will verify this later). Recall this means that given any ball B in $\mathbb{R}^n$ that does not intersect ∂T_i, and any integral current T with the same boundary as the restriction $T_i \llcorner B$, we have

$$|T_i \llcorner B|_\delta \leq \gamma |T|_\delta$$

for some constant $\gamma \geq 1$. (Here we are using $|\cdot|_\delta$ to denote both the Euclidean hypersurface area and the more general current mass.) The following fact is a natural generalization of the classical monotonicity formula for minimal surfaces to the class of γ-almost-minimizing currents (see [Bray and Lee 2009] for instance): for $0 \leq s < \operatorname{dist}(q_i, \partial T_i) = 16r_1$,

$$|T_i \llcorner B(q_i, s)|_\delta \geq \gamma^{2-n} \omega_{n-1} s^{n-1}.$$

Taking the limit $s \nearrow 16r_1$, we have

$$|T_i \llcorner B(q_i, 16r_1)|_\delta \geq \gamma^{2-n}\omega_{n-1}(16r_1)^{n-1} = 16\omega_{n-1}(r_1)^{n-1},$$

taking $\gamma = 16$. Using the factor-of-two area comparisons between δ and h and between h and h_i on $U \setminus B_{r_1}$ for $i \geq i_2$, we then have

$$|T_i \llcorner B(q_i, 16r_1)|_{h_i} \geq \tfrac{1}{4} \cdot 16\omega_{n-1}(r_1)^{n-1}.$$

Applying Φ_i, it follows that $|\widetilde{S}_i' \cap \Phi_i(\bar{B}_{33r_1})|_{g_i} \geq 4\omega_{n-1}(r_1)^{n-1}$. Since $\widetilde{S}_i$ leaves $\Phi_i(\bar{B}_{33r_1})$, we obtain a strict inequality below:

$$|\widetilde{S}_i|_{g_i} \geq |\widetilde{S}_i'|_{g_i} > 4\omega_{n-1}(r_1)^{n-1}. \tag{13}$$

On the other hand, since $\widetilde{S}_i$ by definition has at most as much g_i-area as S_i,

$$|\widetilde{S}_i|_{g_i} \leq |S_i|_{g_i} = |S_{r_1}|_{h_i} \leq 4|S_{r_1}|_\delta = 4\omega_{n-1}(r_1)^{n-1},$$

producing a contradiction with (13).

We now prove that T_i is γ-almost-minimizing in $\mathbb{R}^n$ with $\gamma = 16$, which will complete the proof of Lemma 12. Since T_i is area-minimizing with respect to h_i in $B_{33r_1} \setminus \bar{B}_{r_1}$, we know that

$$|T_i \llcorner B|_{h_i} \leq |T|_{h_i}$$

for any integral current T supported in $B_{33r_1} \setminus \bar{B}_{r_1}$, with $\partial T = \partial(T_i \llcorner B)$, where B is a Euclidean ball in $B_{33r_1} \setminus \bar{B}_{r_1}$. For $i \geq i_2$, since the Hausdorff $(n-1)$-measures of h and h_i are uniformly equivalent by factors of two on U, this implies

$$|T_i \llcorner B|_h \leq 4|T|_h$$

for such B and T. Since T_i is contained outside S_{r_1}, we can use the comparison of areas between δ and h to see that

$$|T_i \llcorner B|_\delta \leq 16|T|_\delta$$

for such B and T. However, in the definition of γ-almost-minimizing, one may without loss of generality consider competitors T supported in $\bar{B}$, since $\bar{B}$ is convex. It follows that T_i is 16-almost-minimizing, and the proof of Lemma 12 is complete. □

We continue with the proof of Theorem 2. Observe that $F_g(S)$ varies continuously with respect to C^2 perturbations of g on any neighborhood of S, since the area, mean curvature, and scalar curvature depend continuously on g and its first and second derivatives. Then by the C^2 convergence in (12), we may restrict to i at least as large as some $i_3 \geq i_2$ so that

$$F_h(S_{r_1}) \leq F_{h_i}(S_{r_1}) + \frac{\epsilon}{2} \tag{14}$$

and that

$$F_{h_i}(S_{r_1}) > 0 \tag{15}$$

(since $F_h(S_{r_1}) > 0$ by (11)). Lemma 12 and (15) show that Theorem 7 may be applied to M_i minus the open region $\Phi_i(B_{r_1})$, which has (connected) boundary S_i. Thus:

$$F_{h_i}(S_{r_1}) = F_{g_i}(S_i) \leq m_i. \tag{16}$$

Then for all $i \geq i_3$, we may combine (10), (14), and (16) to arrive at

$$m_{\mathrm{ADM}}(N, h) < m_i + \epsilon.$$

Now, taking $\liminf_{i\to\infty}$ proves Theorem 2, since $\epsilon > 0$ was arbitrary.

Remark C. The above proof generalizes the C^2 lower semicontinuity result from $n = 3$ in [Jauregui 2018] to $3 \leq n \leq 7$. By contrast, extending the C^0 lower semicontinuity result in [Jauregui and Lee 2017] to higher dimensions would be much more difficult. In the C^0 case, the dimension three hypothesis is relied on to a greater extent. First, the Hawking mass estimate (2) of Huisken and Ilmanen, valid only in dimension three, is used to ensure monotonicity under mean curvature flow of a certain quantity (whose details we omit here) defined by Huisken. The author is not aware of such a monotone quantity in higher dimensions. Second, in [Jauregui and Lee 2017], use is made of B. White's regularity theory for the weak (level set) version of mean curvature flow that is especially nice in ambient dimension three [2000].

Remark D. As mentioned in the introduction, we strongly conjecture that the hypothesis that the limit (N, h) is asymptotically Schwarzschild in Theorem 2 (as opposed to asymptotically flat) is unnecessary. We note this generalization would follow by establishing a density result of the following form: Given $\epsilon > 0$ and a sequence (M_i, g_i, p_i) of AF manifolds of nonnegative scalar curvature converging in the pointed C^2 Cheeger–Gromov sense to an AF manifold (N, h, q), construct an HF perturbation $\bar{h}$ of h (with $|m_{\mathrm{ADM}}(N, \bar{h}) - m_{\mathrm{ADM}}(N, h)| < \epsilon$) and AF metrics $\bar{g}_i$ on M_i of nonnegative scalar curvature, with $|m_{\mathrm{ADM}}(M_i, \bar{g}_i) - m_{\mathrm{ADM}}(M_i, g_i)| < \epsilon$, such that $(M_i, \bar{g}_i, p_i) \to (N, \bar{h}, q)$ in the pointed C^2 Cheeger–Gromov sense. Such a result would immediately generalize Theorem 2 to remove the restriction that (N, h) is asymptotically Schwarzschild, since HF manifolds are such.

6. Lower semicontinuity for weighted C^2 convergence in all dimensions

In this section we study the behavior of the ADM mass under *weighted* C^2 convergence. This corresponds to a finer topology than that of pointed C^2 Cheeger–Gromov convergence. In particular it is easier here to establish semicontinuity of the ADM mass and to obtain a stronger result: Theorem 13 is valid in all dimensions $n \geq 3$, requires no hypothesis on minimal surfaces, and does not rely on (nor recover) the positive mass theorem.

To describe the setup, let M be a smooth n-manifold that admits an AF metric. Fix a compact set $K \subset M$ and an AF coordinate system on $M \setminus K$ (for some AF metric). For an integer $k \geq 0$ and a real number $\tau > 0$, let $C^k_{-\tau}(M \setminus K)$ denote the class of C^k functions $f : M \setminus K \to \mathbb{R}$ for which the quantity

$$\|f\|_{C^k_{-\tau}(M\setminus K)} = \sum_{0\leq|\gamma|\leq k} \sup_{x\in M\setminus K} |x|^{|\gamma|+\tau} |\partial^\gamma f(x)|$$

is finite, where the partial derivatives are taken with respect to the coordinate chart, and γ represents multi-indices. Thus, functions in $C^k_{-\tau}(M \setminus K)$ decay as $O(r^{-\tau})$ or faster as $r \to \infty$, with successively faster decay up through k-th-order derivatives. Define $C^k_{-\tau}(M)$ to be the set of C^k functions $f : M \to \mathbb{R}$ with $f|_{M\setminus K} \in C^k(M \setminus K)$, equipped with the norm given as the sum of $\|f\|_{C^k_{-\tau}(M\setminus K)}$ and the C^k norm of $f|_K$.

Note that if g is an AF metric on g of order τ obeying the decay conditions (3) in the fixed coordinate chart, then

$$g_{ij} - \delta_{ij} \in C^2_{-\tau}(M \setminus K). \tag{17}$$

For $k \geq 2$ and $\tau > 0$, we let $\mathrm{Met}^k_{-\tau}(M)$ denote the set of C^k Riemannian metrics g on M satisfying (17) in the fixed coordinate chart. (The ADM mass of $g \in \mathrm{Met}^k_{-\tau}(M)$ is well defined if $\tau > \frac{n-2}{2}$ and the scalar curvature of g is integrable [Bartnik 1986; Chruściel 1986].) We say a sequence of Riemannian metrics $\{g^\ell\}_{\ell=1}^\infty$ in $\mathrm{Met}^k_{-\tau}(M)$ converges to $g \in \mathrm{Met}^k_{-\tau}(M)$ as $\ell \to \infty$ if $\|g^\ell_{ij} - g_{ij}\|_{C^k_{-\tau}(M\setminus K)} \to 0$ for all i and j and the tensors $g^\ell|_K$ converge in C^k to $g|_K$ as $\ell \to \infty$.

Theorem 13. *Suppose $\{g^\ell\}_{\ell=1}^\infty$ converges to g as asymptotically flat Riemannian metrics in $\mathrm{Met}^2_{-\tau}(M)$, where $\tau > \frac{n-2}{2}$. Then*

$$\lim_{\ell\to\infty}\left(m_{\mathrm{ADM}}(M, g^\ell) - \frac{1}{2(n-1)\omega_{n-1}} \int_M R(g^\ell)\, dV_{g^\ell}\right) = m_{\mathrm{ADM}}(M, g) - \frac{1}{2(n-1)\omega_{n-1}} \int_M R(g)\, dV_g, \tag{18}$$

where dV_{g^ℓ} and dV_g are the volume measures of g^ℓ and g. Moreover, if there exists a compact set $K \subset M$ such that $R(g^\ell) \geq 0$ on $M \setminus K$ for all ℓ, then

$$\liminf_{\ell\to\infty} m_{\mathrm{ADM}}(M, g_\ell) \geq m_{\mathrm{ADM}}(M, g). \tag{19}$$

Remark E. Our (18) is well known to experts as the statement of the continuity of the Regge–Teitelboim Hamiltonian [1974]. This is related to Lemma 9.4 in [Lee and

Parker 1987], which gives continuity of the ADM under weighted $C^{1,\alpha}$ convergence if the scalar curvatures converge in L^1. After posting this paper we became aware of Theorem 14 of [McFeron and Székelyhidi 2012], which implies Theorem 13; this result of D. McFeron and G. Székelyhidi requires local C^2 convergence and a uniform weighted $C^{1,\alpha}$ bound. Our proof below is a generalization of that of Y. Li [2018] (see the proof of Theorem 1.2 therein), who studied the behavior of the ADM mass and integral of scalar curvature in the case of a convergent Ricci flow.

Proof. Let g_0 be a background Riemannian metric on M whose expression in $M \setminus K$ in the given AF coordinate chart is δ_{ij}. Let div_0 be the divergence operator on tensors and Δ_0 the Laplacian on functions with respect to g_0. Define the continuous operator $\mathcal{D} : \mathrm{Met}^2_{-\tau}(M) \to C^0_{-\tau-2}(M)$ by

$$\mathcal{D}(g) = \operatorname{div}_0(\operatorname{div}_0 g) - \Delta_0(\operatorname{tr}_{g_0}(g)).$$

The significance of $\mathcal{D}$ is the formula for the ADM mass of $g \in \mathrm{Met}^2_{-\tau}(M)$ (provided $\tau > \frac{n-2}{2}$ and the scalar curvature of g is integrable):

$$m_{\mathrm{ADM}}(g) = \frac{1}{2(n-1)\omega_{n-1}} \int_M \mathcal{D}(g)\, dV_0, \tag{20}$$

which follows immediately from the divergence theorem. Here, dV_0 is the volume measure of g_0.

By the $\mathrm{Met}^2_{-\tau}(M)$ convergence of g^ℓ to g, we have $\mathcal{D}(g^\ell) \to \mathcal{D}(g)$ in $C^0_{-\tau-2}(M)$. However, since $\tau+2$ is generally less than the $O(r^{-n})$ threshold for integrability, we cannot immediately apply the dominated convergence theorem. (And since we have no control on the sign of $\mathcal{D}(g^\ell)$, we cannot apply Fatou's lemma.)

We proceed instead by considering the difference between $\mathcal{D}(\cdot)$ and $R(\cdot)$ (a well-known trick), where $R : \mathrm{Met}^2_{-\tau}(M) \to C^0_{-\tau-2}(M)$ is the scalar curvature operator. Working in the fixed chart on $M \setminus K$, for any Riemannian metric $h \in \mathrm{Met}^2_{-\tau}(M)$ with Christoffel symbols Γ^k_{ij}, we have

$$\begin{aligned} \mathcal{D}(h) &= \partial_i \partial_j h_{ij} - \partial_j \partial_j h_{ii}, \\ \mathcal{R}(h) &= h^{jk}(\partial_i \Gamma^i_{jk} - \partial_k \Gamma^i_{ij} + \Gamma^m_{jk}\Gamma^i_{im} - \Gamma^m_{ij}\Gamma^i_{km}). \end{aligned}$$

By direct computation, $\mathcal{D}(g^\ell) - R(g^\ell)$ is $O(r^{-2-2\tau})$, where $O(r^{-2-2\tau})$ here is uniform in ℓ and moreover goes to zero in $C^0_{-2-2\tau}(M)$ as $\ell \to \infty$. Since $2+2\tau > n$, this $O(r^{-2-2\tau})$ error term is uniformly bounded by an integrable function on M. Then by the dominated convergence theorem and the pointwise convergence of $\mathcal{D}(g^\ell) - R(g^\ell)$ to $\mathcal{D}(g) - R(g)$,

$$\lim_{\ell \to \infty} \int_M (\mathcal{D}(g^\ell) - R(g^\ell))\, dV_{g^\ell} = \int_M (\mathcal{D}(g) - R(g))\, dV.$$

Together with (20), this proves (18).

For the last claim, assume $R(g^\ell) \geq 0$ for all ℓ on $M \setminus K$, and let

$$\mu = \liminf_{\ell\to\infty} m_{\mathrm{ADM}}(M, g_\ell).$$

If $\mu = +\infty$, the claim follows trivially. Suppose μ is finite. Pass to a subsequence $\{(M, g^{\ell(k)})\}_k$ for which

$$\lim_{k\to\infty} m_{\mathrm{ADM}}(M, g^{\ell(k)}) = \mu.$$

By the first part of the theorem, the sequence

$$\int_M R(g^{\ell(k)})\, dV_{g^{\ell(k)}}$$

then converges, and moreover

$$\mu = m_{\mathrm{ADM}}(M, g) + \frac{1}{2(n-1)\omega_{n-1}} \left(\lim_{k\to\infty} \int_M R(g^{\ell(k)})\, dV_{g^{\ell(k)}} - \int_M R(g)\, dV_g \right). \tag{21}$$

By the (weighted) C^2 convergence of $g^{\ell(k)}$ to g as $k \to \infty$, we have pointwise convergence of the scalar curvatures and volume forms. In particular,

$$\int_K R(g^{\ell(k)})\, dV_{g^{\ell(k)}} \to \int_K R(g)\, dV_g$$

as $k \to \infty$. Then, by Fatou's lemma and the hypothesis $R(g^{\ell(k)}) \geq 0$ on $M \setminus K$, the expression in parentheses in (21) is nonnegative. This completes the proof if μ is finite.

Finally, suppose $\mu = -\infty$. Then by (18), a subsequence $\{(M, g^{\ell(k)})\}_k$ has its integral of scalar curvature converging to $-\infty$. Since the scalar curvatures are nonnegative outside the compact set K, the integrals of the scalar curvatures on K also converge to $-\infty$. This contradicts the fact that these integrals converge to $\int_K R(g)\, dV_g$. $\square$

Remark F. Interestingly, Theorem 13 implies that for the case of weighted C^2 convergence, the mass drop is accounted for completely by the total matter (i.e., the integral of scalar curvature) escaping off to infinity, much like in the example in Section 2.1 from Newtonian gravity. This contrasts with the case of pointed C^2 Cheeger–Gromov convergence, in which the ADM mass can drop within the class of scalar-flat metrics (e.g., the examples in Sections 2.2 or 2.3, choosing (M, g) to be scalar-flat with positive ADM mass).

Remark G. Note that the lower semicontinuity of the ADM mass with respect to weighted C^2 convergence does not imply the positive mass theorem as in the blow-up example or escaping point example with pointed convergence in Section 2. In those cases, the metrics do not converge to Euclidean space in a weighted sense.

6.1. ***Example: mass drop with weighted convergence.*** We conclude this section by describing an example of AF metrics g_i, with nonnegative scalar curvature, converging in $\mathrm{Met}^2_{-\tau}(M)$ with $\tau > \frac{n-2}{2}$ for which the ADM mass drops. Physically, the construction involves considering a sequence of shells of matter, of fixed total mass, at progressively larger radii. For $n \geq 3$, let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a smooth, radially symmetric, nonnegative function supported in the annulus between radii $\frac{1}{2}$ and 1, with $\int_{\mathbb{R}^n} \rho = 1$. For $i = 1, 2, \ldots$, define a sequence of smooth functions

$$\rho_i(x) = i^{-n} \rho(x/i),$$

which also satisfy $\int_{\mathbb{R}^n} \rho_i = 1$ and are supported in the annulus between radii $\frac{i}{2}$ and i.

By elliptic PDE theory (or ODE theory), there exists a unique smooth solution (for each i) to the linear elliptic problem:

$$\begin{cases} -\Delta v_i = \rho_i & \text{on } \mathbb{R}^n \\ v_i \to 0 & \text{at infinity.} \end{cases}$$

Recognizing $v_i(x) = i^{2-n} v_1(x/i)$, it is easy to see that $v_i \to 0$ in $C^2_{-\tau}(M)$ for any $\tau < n-2$ as $i \to \infty$. Fix $\tau \in \left(\frac{n-2}{2}, n-2\right)$.

For i sufficiently large, $u_i := 1 + v_i$ is positive, and the Riemannian metric $g_i := u_i^{4/(n-2)} \delta$ is asymptotically flat. Note that the scalar curvature of g_i

$$R_i = -\frac{4(n-1)}{n-2} u_i^{-\frac{n+2}{n-2}} \Delta u_i = \frac{4(n-1)}{n-2} u_i^{-\frac{n+2}{n-2}} \rho_i,$$

is integrable because it has compact support.

Now, g_i converges to the Euclidean metric in $\mathrm{Met}^2_{-\tau}(M)$ as $i \to \infty$, and each g_i has nonnegative scalar curvature. We show now (using the divergence theorem) that the ADM mass of g_i is a positive constant, independent of i:

$$\begin{aligned} m_{\mathrm{ADM}}(g_i) &= -\frac{2}{(n-2)\omega_{n-1}} \lim_{r \to \infty} \int_{S_r} \nu(u_i)\, dA = -\frac{2}{(n-2)\omega_{n-1}} \int_{\mathbb{R}^n} \Delta u_i \, dV \\ &= \frac{2}{(n-2)\omega_{n-1}} \int_{\mathbb{R}^n} \rho_i \, dV = \frac{2}{(n-2)\omega_{n-1}}, \end{aligned}$$

where dA and dV are the hypersurface area and the volume forms with respect to the Euclidean metric. However, the ADM mass of the limit, Euclidean $\mathbb{R}^n$, vanishes.

7. Two-dimensional case of semicontinuity of mass

In two dimensions, a natural replacement for asymptotically flat manifolds is the class of asymptotically conical surfaces, with the asymptotic cone angle playing the role of mass. The author thanks Woolgar for his suggestion to investigate the semicontinuity of mass in this setting.

Following [Isenberg et al. 2013], for $\alpha > 0$, let

$$g_\alpha = dr^2 + \alpha^2 r^2 d\theta^2,$$

a smooth Riemannian metric on $\mathbb{R}^2 \setminus \{0\}$ describing a cone. Note that g_α has vanishing Gauss curvature. Define a connected two-dimensional Riemannian manifold (M, g) to be *asymptotically conical with cone angle* $2\pi\alpha > 0$ if there exists a compact set $C \subset M$ such that $M \setminus C$ is diffeomorphic to the complement of a closed ball in $\mathbb{R}^2$, on which $g - g_\alpha = O_2(r^{-\tau})$ for some constant $\tau > 0$. In particular, the Gauss curvature of g is $O(r^{-2-\tau})$ and hence integrable.

We recall here that the integral of the Gauss curvature captures the cone angle. To see this, let B_r be the compact region bounded by the coordinate circle Γ_r in (M, g) for r large. By the Gauss–Bonnet formula,

$$\int_{B_r} K\, dA = 2\pi\chi(B_r) - \int_{\Gamma_r} \kappa_g\, ds, \tag{22}$$

where κ_g is the geodesic curvature of Γ_r with respect to g. By the $O_2(r^{-\tau})$ decay of g to g_α, we have

$$\lim_{r\to\infty} \int_{\Gamma_r} \kappa_g\, ds = \lim_{r\to\infty} \int_{\Gamma_r} \kappa_{g_\alpha}\, ds = 2\pi\alpha,$$

the latter equality given by direct calculation, where κ_{g_α} is the geodesic curvature of Γ_r with respect to g_α. Taking the limit $r \to \infty$ in (22) (and noting that $\chi(B_r)$ is eventually a constant, $\chi(M)$), we have

$$\int_M K\, dA = 2\pi(\chi(M) - 1) + 2\pi(1 - \alpha). \tag{23}$$

Note that if $K \geq 0$, it follows that $\chi(M) > 0$, and using the fact that M is topologically the connect sum of $\mathbb{R}^2$ and a compact, connected surface, it follows that $\chi(M) = 1$ and that M itself is topologically $\mathbb{R}^2$.

We define the mass of an asymptotically conical surface to be

$$m_{\text{cone}}(M, g) = 1 - \alpha,$$

which we note is a dimensionless quantity. The positive mass theorem is then immediate: $K \geq 0$ implies $m_{\text{cone}} \geq 0$, and equality holds if and only if $K \equiv 0$ and M is homeomorphic to $\mathbb{R}^2$, which holds if and only if (M, g) is isometric to the Euclidean plane.

Below is the statement of C^2 pointed lower semicontinuity of the mass in two dimensions (i.e., upper semicontinuity of the cone angle). Note that no hypothesis on closed geodesics (the analogs of compact minimal hypersurfaces) is necessary.

Theorem 14. *Suppose (M_i, g_i, p_i) converges in the pointed C^2 Cheeger–Gromov sense to (N, h, q) as pointed asymptotically conical Riemannian 2-manifolds. Suppose each (M_i, g_i) has nonnegative Gauss curvature. Then*

$$m_{\text{cone}}(N, h) \leq \liminf_{i\to\infty} m_{\text{cone}}(M_i, g_i). \tag{24}$$

An example for which strict inequality holds in (24) can be found using the blow-up or escaping point examples in Sections 2.2 and 2.3, beginning with an asymptotically conical surface with nonnegative Gauss curvature and $\alpha < 1$.

Proof. Let $\epsilon > 0$. By the C^2 convergence, h itself has nonnegative Gauss curvature K_h, so in particular $\chi(N) = 1$. Then by (23),

$$m_{\text{cone}}(N, h) = \frac{1}{2\pi}\int_N K_h \, dA_h.$$

Since K_h is integrable, we may choose $r > 0$ sufficiently large so that the coordinate ball $B_r \subset N$ satisfies

$$m_{\text{cone}}(N, h) < \frac{1}{2\pi}\int_{B_r} K_h \, dA_h + \frac{\epsilon}{2}.$$

Choosing $U \supset B_r$ and obtaining appropriate embeddings $\Phi_i : U \to M_i$ such that $h_i := \Phi_i^* g_i$ converges in C^2 to h, we may take i sufficiently large so that

$$\frac{1}{2\pi}\int_{B_r} K_h \, dA_h - \frac{\epsilon}{2} < \frac{1}{2\pi}\int_{B_r} K_{h_i} \, dA_{h_i} = \frac{1}{2\pi}\int_{\Phi_i(B_r)} K_{g_i} \, dA_{g_i}. \tag{25}$$

Since (M_i, g_i) has nonnegative Gauss curvature, the right-hand side in (25) is an underestimate for $m_{\text{cone}}(M_i, g_i)$. Thus,

$$m_{\text{cone}}(N, h) < m_{\text{cone}}(M_i, g_i) + \epsilon$$

for i sufficiently large. From this, the result follows. □

We leave it as an open problem to study the behavior of the cone angle under weaker forms of convergence, such as pointed C^0 Cheeger–Gromov, pointed Gromov–Hausdorff, or pointed Sormani–Wenger intrinsic flat convergence [Sormani and Wenger 2011].

Appendix: Geometry of asymptotically Schwarzschild metrics

The purpose of this appendix is to prove the following asymptotic estimates for large coordinate spheres in an asymptotically Schwarzschild manifold. These were used in the proof of Lemma 11.

Lemma 15. *Let $(M, \tilde{g})$ be an asymptotically Schwarzschild manifold of dimension $n \geq 3$ and ADM mass m. Let S_r be the coordinate sphere of large radius r in M. Let*

$\tilde{\rho}$ be the scalar curvature of S_r with respect to the metric induced from $\tilde{g}$, and let $\widetilde{H}$ be the mean curvature of S_r with respect to $\tilde{g}$. Then:

$$\tilde{\rho} = \frac{(n-1)(n-2)}{r^2} - \frac{2(n-1)m}{r^n} + O(r^{-n-1}). \tag{26}$$

$$\widetilde{H} = \frac{n-1}{r} - \frac{(n-1)^2 m}{(n-2)r^{n-1}} + O(r^{-n}). \tag{27}$$

Proof. Let g be the Schwarzschild metric of mass m, and let h be as in (5), i.e.,

$$g = \left(1 + \frac{m}{2r^{n-2}}\right)^{\frac{4}{n-2}} \delta, \qquad \tilde{g} = g + h,$$

in the end of M.

We first address the scalar curvature of S_r. Let γ and $\tilde{\gamma}$ be the Riemannian metrics on S_r induced by g and $\tilde{g}$, respectively. The coordinate sphere S_r has constant scalar curvature with respect to the metric induced by δ equal to $(n-1)(n-2)/(r^2)$. Since the conformal factor relating g to δ is constant on S_r, the scalar curvature of (S_r, γ) can be found by rescaling:

$$\rho = \left(1 + \frac{m}{2r^{n-2}}\right)^{-\frac{4}{n-2}} \cdot \frac{(n-1)(n-2)}{r^2} = \frac{(n-1)(n-2)}{r^2} - \frac{2(n-1)m}{r^n} + O(r^{-2(n-1)}). \tag{28}$$

We proceed to estimate the scalar curvature of $(S_r, \tilde{\gamma})$ as follows. Introduce spherical coordinates $(r, \phi^1, \ldots, \phi^{n-1})$ on the asymptotically flat end of M:

$$\begin{aligned} x^1 &= r\cos(\phi^1), \\ x^2 &= r\sin(\phi^1)\cos(\phi^2), \\ &\vdots \\ x^{n-1} &= r\sin(\phi^1)\sin(\phi^2)\cdots\cos(\phi^{n-1}), \\ x^n &= r\sin(\phi^1)\sin(\phi^2)\cdots\sin(\phi^{n-1}). \end{aligned}$$

We use Greek indices for the directions tangent to S_r, i.e., ϕ^α for $\alpha = 1, \ldots, n-1$ for the coordinates on S_r and $\partial_\alpha = \frac{\partial}{\partial \phi^\alpha}$ for their derivatives. Note that $\delta(\partial_\alpha, \partial_\beta)$ is $O(r^2)$.

First, express γ and $\tilde{\gamma}$ in spherical coordinates on S_r:

$$\gamma_{\alpha\beta} = g(\partial_\alpha, \partial_\beta), \quad \tilde{\gamma}_{\alpha\beta} = \tilde{g}(\partial_\alpha, \partial_\beta) = \gamma_{\alpha\beta} + h_{\alpha\beta}, \tag{29}$$

where $h_{\alpha\beta} = h(\partial_\alpha, \partial_\beta)$ is $O(r^{3-n})$ by (5). Both $\gamma_{\alpha\beta}$ and $\tilde{\gamma}_{\alpha\beta}$ are $O(r^2)$. Also, we have the inverse metrics

$$\gamma^{\alpha\beta} = O(r^{-2}), \tag{30}$$

$$\tilde{\gamma}^{\alpha\beta} = \gamma^{\alpha\beta} + O(r^{-n-1}). \tag{31}$$

Note that the derivatives tangent to S_r satisfy

$$\partial_\mu \gamma_{\alpha\beta} = O(r^2), \tag{32}$$

$$\partial_\mu h_{\alpha\beta} = O(r^{3-n}),$$

$$\partial_\mu \tilde{\gamma}_{\alpha\beta} = O(r^2), \tag{33}$$

with the same orders for second derivatives. Similarly,

$$\partial_\mu \gamma^{\alpha\beta} = O(r^{-2}), \tag{34}$$

$$\partial_\mu \tilde{\gamma}^{\alpha\beta} = O(r^{-2}). \tag{35}$$

Next, let Γ and $\widetilde{\Gamma}$ denote the Christoffel symbols of (S_r, γ) and $(S_r, \tilde{\gamma})$, respectively, and define

$$\Psi^\mu_{\alpha\beta} = \widetilde{\Gamma}^\mu_{\alpha\beta} - \Gamma^\mu_{\alpha\beta}.$$

By (30) and (32), we have

$$\Gamma^\mu_{\alpha\beta} = O(1). \tag{36}$$

Using (34) as well,

$$\partial_\nu \Gamma^\mu_{\alpha\beta} = O(1). \tag{37}$$

Next, we need decay on Ψ and $\partial\Psi$. Using (29)–(31),

$$\begin{aligned}\Psi^\mu_{\alpha\beta} &= \tilde{\gamma}^{\mu\nu}(\partial_\beta \tilde{\gamma}_{\alpha\nu} + \partial_\alpha \tilde{\gamma}_{\beta\nu} - \partial_\nu \tilde{\gamma}_{\alpha\beta}) - \gamma^{\mu\nu}(\partial_\beta \gamma_{\alpha\nu} + \partial_\alpha \gamma_{\beta\nu} - \partial_\nu \gamma_{\alpha\beta}) \\ &= O(r^{-2})(\partial_\beta h_{\alpha\nu} + \partial_\alpha h_{\beta\nu} - \partial_\nu h_{\alpha\beta}) + O(r^{-n-1})(\partial_\beta \tilde{\gamma}_{\alpha\nu} + \partial_\alpha \tilde{\gamma}_{\beta\nu} - \partial_\nu \tilde{\gamma}_{\alpha\beta}).\end{aligned}$$

Since $h_{\alpha\beta}$ and $\partial_\mu h_{\alpha\beta}$ are $O(r^{3-n})$, and also by (33), we have

$$\Psi^\mu_{\alpha\beta} = O(r^{1-n}),$$

and a similar calculation, using (35), shows

$$\partial_\nu \Psi^\mu_{\alpha\beta} = O(r^{1-n}).$$

Finally:

$$\begin{aligned}\tilde{\rho} &= \tilde{\gamma}^{\beta\mu}\big(\partial_\alpha \widetilde{\Gamma}^\alpha_{\beta\mu} - \partial_\mu \widetilde{\Gamma}^\alpha_{\alpha\beta} + \widetilde{\Gamma}^\nu_{\beta\mu}\widetilde{\Gamma}^\alpha_{\alpha\nu} - \widetilde{\Gamma}^\nu_{\alpha\beta}\widetilde{\Gamma}^\alpha_{\mu\nu}\big) \\ &= (\gamma^{\beta\mu} + O(r^{-n-1})) \\ &\qquad \times \big[\partial_\alpha(\Gamma^\alpha_{\beta\mu} + \Psi^\alpha_{\beta\mu}) - \partial_\mu(\Gamma^\alpha_{\alpha\beta} + \Psi^\alpha_{\alpha\beta}) \\ &\qquad\qquad + (\Gamma^\nu_{\beta\mu} + \Psi^\nu_{\beta\mu})(\Gamma^\alpha_{\alpha\nu} + \Psi^\alpha_{\alpha\nu}) - (\Gamma^\nu_{\alpha\beta} + \Psi^\nu_{\alpha\beta})(\Gamma^\alpha_{\mu\nu} + \Psi^\alpha_{\mu\nu})\big] \\ &= \rho + O(r^{-n-1}),\end{aligned}$$

having used (30), (31), (36), and (37). Combining this with (28), (26) follows.

For the second part of the proof, we must compute the mean curvature of large coordinate spheres S_r with respect to $\tilde{g}$. We approach this through the first variation of area. Let ω_0, ω, and $\tilde{\omega}$ be the area forms of S_r induced by δ, g and $\tilde{g}$, respectively. The respective mean curvature vectors $\boldsymbol{H}_0$, $\boldsymbol{H}$, $\widetilde{\boldsymbol{H}}$ of S_r with respect to these metrics are characterized by the first variation of area formulas as follows:

$$D_X\omega_0 = \delta(X, -\boldsymbol{H}_0)\omega_0 = \delta\Big(X, \frac{n-1}{r}\cdot\partial_r\Big)\omega_0, \tag{38}$$

$$D_X\omega = g(X, -\boldsymbol{H})\omega, \tag{39}$$

$$D_X\tilde{\omega} = \tilde{g}(X, -\widetilde{\boldsymbol{H}})\tilde{\omega}, \tag{40}$$

where D_X denotes an infinitesimal deformation of S_r in the direction of X, where X is a tangent vector field to M along S_r.

We again use spherical coordinates as in the first part of the proof. Note that (ϕ^α) give coordinates on S_r that are orthogonal with respect to δ, and hence with respect to the conformal metric g. In addition to the estimates of $\gamma_{\alpha\beta}$, $h_{\alpha\beta}$ and their tangential derivatives used in the first part of the proof, we also need estimates on the radial derivatives. By the decay of g and h, as well as by (30), we obtain

$$\partial_r\gamma_{\alpha\beta} = O(r^1), \quad \partial_r\gamma^{\alpha\beta} = O(r^{-3}), \quad \partial_r h_{\alpha\beta} = O(r^{2-n}).$$

We begin by computing the mean curvature H of S_r with respect to g; this is well known, but we include it for completeness. The area forms ω_0 and ω on S_r are related by

$$\omega = \Big(1 + \frac{m}{2r^{n-2}}\Big)^{\frac{2(n-1)}{n-2}}\omega_0.$$

Then, using (38), elementary calculations show

$$\begin{aligned} D_r\omega &= \frac{2(n-1)}{n-2}\Big(1+\frac{m}{2r^{n-2}}\Big)^{\frac{2(n-1)}{n-2}-1}\cdot\Big(\frac{(2-n)m}{2r^{n-1}}\Big)\omega_0 + \Big(1+\frac{m}{2r^{n-2}}\Big)^{\frac{2(n-1)}{n-2}} D_r\omega_0 \\ &= \Big(1+\frac{m}{2r^{n-2}}\Big)^{\frac{2(n-1)}{n-2}}\cdot\Big[\frac{n-1}{r} - \frac{m(n-1)}{r^{n-1}}\Big(1+\frac{m}{2r^{n-2}}\Big)^{-1}\Big]\omega_0, \end{aligned} \tag{41}$$

where $D_r = D_{\partial_r}$. Now, using (39), we have

$$\begin{aligned} D_r\omega &= g(\partial_r, -\boldsymbol{H})\omega \\ &= \Big(1 + \frac{m}{2r^{n-2}}\Big)^{\frac{2n}{n-2}} H\omega_0, \end{aligned} \tag{42}$$

where $H = |\boldsymbol{H}|_g$. Now, combining (41) and (42), elementary calculations show

$$H = \frac{n-1}{r} - \frac{(n-1)^2 m}{(n-2)r^{n-1}} + O(r^{-n}). \tag{43}$$

Now, we proceed to estimate the mean curvature with respect to $\tilde{g}$. Define a function $\Phi > 0$ on the asymptotically flat end of M so that

$$\tilde{\omega} = \sqrt{\Phi}\omega \quad \text{on } S_r, \tag{44}$$

that is,

$$\Phi = \frac{\det(\tilde{\gamma}_{\alpha\beta})}{\det(\gamma_{\alpha\beta})}.$$

Using Jacobi's formula for the derivative of the determinant, along with the known decay of $\gamma_{\alpha\beta}$, and $\tilde{\gamma}_{\alpha\beta}$ and their derivatives, we have the following asymptotics of Φ:

$$\Phi = 1 + O(r^{1-n}), \tag{45}$$

$$\partial_\mu \Phi = O(r^{1-n}), \tag{46}$$

$$\partial_r \Phi = O(r^{-n}). \tag{47}$$

In order to compute $\widetilde{\boldsymbol{H}}$, we compute tangential and radial variations of $\tilde{\omega}$ beginning with (44):

$$\begin{aligned} D_\mu \tilde{\omega} &= \tfrac{1}{2}(\partial_\mu \Phi)\Phi^{-1/2}\omega + \sqrt{\Phi} D_\mu \omega \\ &= \tfrac{1}{2}(\partial_\mu \Phi)\Phi^{-1/2}\omega + \sqrt{\Phi} g(\partial_\mu, -\boldsymbol{H})\omega \\ &= O(r^{1-n})\omega, \end{aligned} \tag{48}$$

where we have used the fact that $\boldsymbol{H}$ is g-orthogonal to S_r, as well as (39) and (45)–(46). Next, for the radial directions:

$$\begin{aligned} D_r \tilde{\omega} &= \tfrac{1}{2}(\partial_r \Phi)\Phi^{-1/2}\omega + \sqrt{\Phi} g(\partial_r, -\boldsymbol{H})\omega \\ &= g(\partial_r, -\boldsymbol{H})\omega + O(r^{-n})\omega, \end{aligned} \tag{49}$$

having used (45), (47), and $H = O(r^{-1})$. The goal is to combine the last two statements with (40). Specifically, we estimate (40) as follows:

$$\begin{aligned} D_X \tilde{\omega} &= (g+h)(X, -\widetilde{\boldsymbol{H}})\sqrt{\Phi}\omega \\ &= g(X, -\widetilde{\boldsymbol{H}})\omega + O(r^{-n})|X|_g \omega, \end{aligned}$$

having used the decay of h, (45), and $|\widetilde{\boldsymbol{H}}|_g = O(r^{-1})$. Define $\boldsymbol{Y} = \widetilde{\boldsymbol{H}} - \boldsymbol{H}$. Then applying (49) and the last equation (with $X = \partial_r$) and applying (48) and the last equation (with $X = \partial_\mu$) produces

$$g(\partial_r, \boldsymbol{Y}) = O(r^{-n}), \tag{50}$$

$$g(\partial_\mu, \boldsymbol{Y}) = O(r^{1-n}). \tag{51}$$

By expanding $|\boldsymbol{Y}|_g^2$ in the g-orthogonal basis $(\partial_r, \partial_\mu)$ of TM along S_r, and using (50)–(51), we obtain

$$|\boldsymbol{Y}|_g^2 = O(r^{-2n}). \tag{52}$$

Finally, letting $\widetilde{H} = |\widetilde{\boldsymbol{H}}|_{\tilde{g}}$, we use the triangle inequality to show

$$\begin{aligned}
|\widetilde{H} - H| &\leq \big||\widetilde{\boldsymbol{H}}|_{\tilde{g}} - |\boldsymbol{H}|_{\tilde{g}}\big| + \big||\boldsymbol{H}|_{\tilde{g}} - |\boldsymbol{H}|_g\big| \\
&\leq |\boldsymbol{Y}|_{\tilde{g}} + \big||\boldsymbol{H}|_{\tilde{g}} - |\boldsymbol{H}|_g\big| \\
&= (g(\boldsymbol{Y}, \boldsymbol{Y}) + h(\boldsymbol{Y}, \boldsymbol{Y}))^{\frac{1}{2}} + |(g(\boldsymbol{H}, \boldsymbol{H}) + h(\boldsymbol{H}, \boldsymbol{H}))^{\frac{1}{2}} - g(\boldsymbol{H}, \boldsymbol{H})^{\frac{1}{2}}| \\
&= |\boldsymbol{Y}|_g + |\boldsymbol{Y}|_g O(r^{1-n}) + H \cdot O(r^{1-n}) \\
&= O(r^{-n}),
\end{aligned}$$

by (52). Combining this with (43), (27) follows. □

Acknowledgements

The author thanks J. Corvino, D. Lee, S. McCormick, and P. Miao for helpful discussions. The author acknowledges support from the Erwin Schrödinger Institute, where a portion of this work was completed in 2017.

References

[Arnowitt et al. 1961] R. Arnowitt, S. Deser, and C. W. Misner, "Coordinate invariance and energy expressions in general relativity", *Phys. Rev.* (2) **122** (1961), 997–1006. MR Zbl

[Bartnik 1986] R. Bartnik, "The mass of an asymptotically flat manifold", *Comm. Pure Appl. Math.* **39**:5 (1986), 661–693. MR Zbl

[Bartnik 1989] R. Bartnik, "New definition of quasilocal mass", *Phys. Rev. Lett.* **62**:20 (1989), 2346–2348. MR

[Bartnik 1997] R. Bartnik, "Energy in general relativity", pp. 5–27 in *Tsing Hua lectures on geometry & analysis* (Hsinchu, Taiwan, 1990/1991), edited by S.-T. Yau, Int. Press, Cambridge, 1997. MR Zbl

[Bartnik 2002] R. Bartnik, "Mass and 3-metrics of non-negative scalar curvature", pp. 231–240 in *Proceedings of the International Congress of Mathematicians, II* (Beijing, 2002), edited by T. Li, Higher Ed. Press, Beijing, 2002. MR Zbl arXiv

[Bray 2001] H. L. Bray, "Proof of the Riemannian Penrose inequality using the positive mass theorem", *J. Differential Geom.* **59**:2 (2001), 177–267. MR Zbl

[Bray and Lee 2009] H. L. Bray and D. A. Lee, "On the Riemannian Penrose inequality in dimensions less than eight", *Duke Math. J.* **148**:1 (2009), 81–106. MR Zbl

[Carlotto and Schoen 2016] A. Carlotto and R. Schoen, "Localizing solutions of the Einstein constraint equations", *Invent. Math.* **205**:3 (2016), 559–615. MR Zbl

[Chruściel 1986] P. Chruściel, "Boundary conditions at spatial infinity from a Hamiltonian point of view", pp. 49–59 in *Topological properties and global structure of space-time* (Erice, Italy, 1985), edited by P. G. Bergmann and V. De Sabbata, NATO Adv. Sci. Inst. Ser. B Phys. **138**, Plenum, New York, 1986. MR Zbl

[Dai and Ma 2007] X. Dai and L. Ma, "Mass under the Ricci flow", *Comm. Math. Phys.* **274**:1 (2007), 65–80. Zbl

[Huisken and Ilmanen 2001] G. Huisken and T. Ilmanen, "The inverse mean curvature flow and the Riemannian Penrose inequality", *J. Differential Geom.* **59**:3 (2001), 353–437. MR Zbl

[Isenberg et al. 2013] J. Isenberg, R. Mazzeo, and N. Sesum, "Ricci flow on asymptotically conical surfaces with nontrivial topology", *J. Reine Angew. Math.* **676** (2013), 227–248. MR Zbl arXiv

[Jauregui 2018] J. L. Jauregui, "On the lower semicontinuity of the ADM mass", *Comm. Anal. Geom.* **26**:1 (2018), 85–111. MR Zbl

[Jauregui and Lee 2017] J. L. Jauregui and D. A. Lee, "Lower semicontinuity of mass under C^0 convergence and Huisken's isoperimetric mass", *J. Reine Angew. Math.* (online publication May 2017).

[Lee and Parker 1987] J. M. Lee and T. H. Parker, "The Yamabe problem", *Bull. Amer. Math. Soc.* (*N.S.*) **17**:1 (1987), 37–91. MR Zbl

[Li 2018] Y. Li, "Ricci flow on asymptotically Euclidean manifolds", *Geom. Topol.* **22**:3 (2018), 1837–1891. MR Zbl

[McCormick and Miao 2017] S. McCormick and P. Miao, "On a Penrose-like inequality in dimensions less than eight", *Int. Math. Res. Not.* (online publication August 2017).

[McFeron and Székelyhidi 2012] D. McFeron and G. Székelyhidi, "On the positive mass theorem for manifolds with corners", *Comm. Math. Phys.* **313**:2 (2012), 425–443. MR Zbl

[Oliynyk and Woolgar 2007] T. A. Oliynyk and E. Woolgar, "Rotationally symmetric Ricci flow on asymptotically flat manifolds", *Comm. Anal. Geom.* **15**:3 (2007), 535–568. MR Zbl

[Regge and Teitelboim 1974] T. Regge and C. Teitelboim, "Role of surface integrals in the Hamiltonian formulation of general relativity", *Ann. Physics* **88** (1974), 286–318. MR Zbl

[Schoen 1989] R. M. Schoen, "Variational theory for the total scalar curvature functional for Riemannian metrics and related topics", pp. 120–154 in *Topics in calculus of variations* (Montecatini Terme, Italy, 1987), edited by M. Giaquinta, Lecture Notes in Math. **1365**, Springer, 1989. MR Zbl

[Schoen and Yau 1979a] 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

[Schoen and Yau 1979b] R. M. Schoen and S.-T. Yau, "Complete manifolds with nonnegative scalar curvature and the positive action conjecture in general relativity", *Proc. Nat. Acad. Sci. U.S.A.* **76**:3 (1979), 1024–1025. MR Zbl

[Sormani and Wenger 2011] C. Sormani and S. Wenger, "The intrinsic flat distance between Riemannian manifolds and other integral current spaces", *J. Differential Geom.* **87**:1 (2011), 117–199. MR Zbl

[White 2000] B. White, "The size of the singular set in mean curvature flow of mean-convex sets", *J. Amer. Math. Soc.* **13**:3 (2000), 665–695. 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 July 9, 2018.

JEFFREY L. JAUREGUI
DEPARTMENT OF MATHEMATICS
UNION COLLEGE
SCHENECTADY, NY
UNITED STATES
jaureguj@union.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.467

BOUNDARY REGULARITY FOR ASYMPTOTICALLY HYPERBOLIC METRICS WITH SMOOTH WEYL CURVATURE

XIAOSHANG JIN

We study the regularity of asymptotically hyperbolic metrics in general dimensions. By carefully constructing harmonic coordinates near the boundary at infinity, a method pioneered by Anderson, we show that, for $m \geq 3$, a $C^{m,\alpha}$ asymptotically hyperbolic metric that satisfies the asymptotic Einstein condition $\|E(g_+)\|_{g_+} = \|\mathrm{Ric}_{g_+} + ng_+\|_{g_+} = o(\rho^2)$ is in fact $C^{m+2,\alpha}$, provided that its Weyl curvature is $C^{m,\alpha}$ and the metric on the boundary that represents its conformal infinity is $C^{m+2,\alpha}$.

1. Introduction

It is well known that there are very close connections between the hyperbolic space $\mathbb{H}^{n+1}$ and its boundary. In recent years, however, many mathematicians tend to be interested in conformally compact Einstein manifolds with negative scalar curvature instead of hyperbolic space. The physics community has also become interested in the compact Einstein manifolds since the introduction of the AdS/CFT correspondence proposed by Maldacena in the theory of quantum gravity in theoretic physics. In this paper, we mainly discuss the boundary regularity problem when the Weyl curvature of the compactification has some regularity.

Let M be the interior of a compact $(n+1)$-dimensional manifold $\overline{M}$ with nonempty boundary ∂M. We say a complete metric g_+ on M is $C^{m,\alpha}$(or $W^{k,p}$) conformally compact if there exists a defining function ρ on $\overline{M}$ such that the conformally equivalent metric

$$g = \rho^2 g_+$$

can extend to a $C^{m,\alpha}$(or $W^{k,p}$) Riemannian metric on $\overline{M}$. The defining function is smooth on $\overline{M}$ and satisfies

$$\begin{cases} \rho > 0 & \text{in } M, \\ \rho = 0 & \text{on } \partial M, \\ d\rho \neq 0 & \text{on } \partial M. \end{cases} \tag{1-1}$$

Here $C^{m,\alpha}$ and $W^{k,p}$ are the usual Hölder space and the Sobolev space.

The author's research is partially supported by China Scholarship Council (File No.201706190079).
MSC2010: 53A30, 53C21, 53C25, 58J05.
Keywords: regularity, asymptotically hyperbolic, harmonic coordinates, Einstein, Weyl curvature.

The induced metric $h = g|_{\partial M}$ is called the boundary metric associated to the compactification g. The defining function is unique up to a multiplication by a positive function on $\overline{M}$. So the conformal class $[g]$ is uniquely determined by g_+, and the conformal class $[h]$ is uniquely determined by (M, g_+). We call $[h]$ the conformal infinity of g_+. We are interested in Einstein manifolds, which means the metric g_+ also satisfies

$$E(g_+) = \mathrm{Ric}_{g_+} + ng_+ = 0. \tag{1-2}$$

The boundary regularity problem was first raised by Fefferman and Graham in 1985. Namely, given a conformally compact Einstein manifold (M, g_+) and its compactification g, if the boundary metric h is in $C^{m,\alpha}$, is there a $C^{m,\alpha}$ compactification of g_+? In fact, Fefferman and Graham noticed that if $\dim M = n+1$ is odd, the boundary regularity in general breaks down at order n. When $\dim M = n+1$ is even, the $C^{m,\alpha}$ compactification should exist. M. T. Anderson [2003; 2008], solved the problem in dimension 4 by using the Bach equation in dimension 4. He only assumed the original compactification g is in $W^{2,p}$ for some $p > 4$. I'm not sure whether the $W^{2,p}$ condition is good enough for the manifold to improve the boundary regularity. Helliwell [2008] solved the issue in all even dimensions by following Anderson's method. He considered the Fefferman–Graham ambient obstruction tensor instead of the Bach tensor in higher dimensions. Helliwell assumed the compactification g is at least in $C^{n,\alpha}$ for a $(n+1)$ smooth manifold. It means the original compactification is $C^{3,\alpha}$ for a smooth manifold of dimension 4.

In this paper, we follow Anderson's approach to study the boundary regularity in general dimensions. As observed in [Fefferman and Graham 1985], if $\dim M = n+1$ is odd, there are log terms in the asymptotic expansion of g near ∂M at the order n. If we add a condition of g, the log term may be ruled out. By studying (3-8), we find that if the Weyl curvature and scalar curvature of g are smooth enough near boundary, the log term may not exist. In [Jin 2017] and [Shi and Tian 2005], (2-5) tells us that in harmonic coordinates, the regularity of a metric can be improved to two orders higher than the regularity of its Weyl curvature locally in the sense of conformal transformation. Our main aim is to extend the result to the manifolds with boundary.

At the end of this paper, we prove that the regularity of defining function is the same as the new structure on $\overline{M}$ for Einstein case. This extends the result in [Helliwell 2008], where it was proved that the regularity of defining functions is the same as the original compactification.

The equation (2-5) holds for all manifolds, not just for Einstein manifolds. We don't need to use the Einstein equation in the interior of $\overline{M}$. We focus on metrics that satisfy the condition of that Einstein equation vanishing to finite order near boundary, that is

$$\|E(g_+)\|_{g_+} = \|\mathrm{Ric}_{g_+} + ng_+\|_{g_+} = o(\rho^2). \tag{1-3}$$

The main result is as follows:

Theorem 1.1. *Let (M, g_+) be a conformally compact $(n+1)$-manifold with a $C^{m,\alpha}$ conformal compactification $g = \rho^2 g_+$ in a given C^∞ atlas $\{y^\beta\}_{\beta=0}^n$ of $\overline{M}$ near ∂M $(m \geq 3, 0 < \alpha < 1)$. $\|\mathrm{Ric}_{g_+} + ng_+\|_{g_+} = o(\rho^2)$. ρ is a C^∞ defining function of y^β. If the boundary metric $h = g|_{\partial M} \in C^{m+2,\alpha}(\partial M)$ and the Weyl curvature W of g is in $C^{m,\alpha}(\overline{M})$ in the atlas y^β, then there exists atlas $\{x^\beta\}$ of $\overline{M}$ near ∂M and in the atlas $\{x^\beta\}$, g_+ has a $C^{m+2,\alpha}$ compactification $\tilde{g} = \tilde{\rho}^2 g_+$ with boundary metric h. The atlas $\{x^\beta\}$ forms a $C^{m+3,\alpha}$ structure of $\overline{M}$. Furthermore, if g_+ is Einstein, $\tilde{\rho}$ is a $C^{m+2,\alpha}$ function in x-coordinates.*

Remark 1.2. If $\dim M = n+1$ is even, (M, g_+) is Einstein and $m+2 \geq n$, then the defining function in Theorem 1.1, $\tilde{\rho}$, is a $C^{m+3,\alpha}$ function in x-coordinates.

In [Chruściel et al. 2005], Chruściel, Delay, Lee and Skinner showed a good result of the boundary regularity of conformal compact Einstein manifolds. They proved that if the boundary metrics are smooth, the C^2 conformally compact Einstein metrics have conformal compactifications that are smooth up to the boundary in the sense of $C^{1,\lambda}$ diffeomorphism in dimension 3 and all even dimensions, and polyhomogeneous in odd dimensions greater than 3. The C^2 condition is of course weaker than the $C^{n,\alpha}$ condition in Helliwell's paper. I think the C^2 condition is probably sharp. However, their result only holds for the smooth case. It is unknown whether their method can be used to prove the finite boundary regularity. In this paper, by assuming a condition of Weyl tensors, we solve the finite regularity problem for a conformal compact manifold, which is not Einstein, only satisfies (1-3). Furthermore, by observing a calculation in Section 3B we find that the log term of the formal power series of a Weyl tensor vanishes if and only if the obstruction tensor of the metric vanishes. So if we assume the Weyl tensors are in C^{n-2} in Theorem A in [Chruściel et al. 2005] when $n+1$ is odd and greater than 3, we can obtain an extended smooth result. That is:

Remark 1.3. Let (M, g_+) be a conformally compact Einstein $(n+1)$-manifold with a C^2 conformal compactification $g = \rho^2 g_+$. If the boundary metric $h = g|_{\partial M}$ is smooth, then for any $\lambda > 0$, there exists $R > 0$ and a $C^{1,\lambda}$ diffeomorphism $\Phi : \overline{M}_R \to \overline{M}$ such that

$$\Phi^* g_+ = \rho^{-2}(d\rho^2 + G(\rho)),$$

where $\overline{M}_R = \partial M \times [0, R]$, $\{G(\rho) : 0 < \rho \leq R\}$ is a one-parameter family of smooth Riemannian metrics on ∂M.

If $\dim M$ is even or equal to 3, then $\Phi^* g_+$ is conformally compact of class C^∞.

If $\dim M$ is odd and greater than 3, the Weyl tensors of $\rho^2 \Phi^* g_+$ are of class C^{n-2}, then $\Phi^* g_+$ is conformally compact of class C^∞.

The outline of this paper is as follows. In Section 2, we introduce the constant scalar curvature compactification and construct a kind of harmonic coordinates near ∂M. The regularity of the metric does not change in the above two steps. We also study the relationship between Ricci curvature and Weyl curvature for constant scalar curvature.

In Section 3, we provide some literature reviews concerning the study of conformally compact Einstein manifolds and asymptotically hyperbolic metrics satisfying (1-3), including the change of curvature under conformal transformation, the existence and regularity of geodesic defining function. We also show the reason why the boundary regularity in general breaks down at order n when $\dim M = n+1$ is odd. Then after some simple computations, we get that the Weyl curvature has an influence on the regularity in geodesic coordinates. Besides, for conformally compact Einstein manifolds, by calculating the formal power series of the Weyl tensor in geodesic coordinates, we obtain that the obstruction tensor of the conformal metric vanishes if and only if the formal power series of the Weyl tensor doesn't contain an $x^{n-2}\log x$ term. This improves Theorem A in [Chruściel et al. 2005].

Section 4 is dedicated to the derivation of some boundary conditions, including the Dirichlet conditions for g_{ij} and Ricci curvature, and the Neumann conditions for $g^{0\beta}$. We use the geodesic defining function as a transition tool to calculate the Dirichlet conditions for Ricci curvature. The regularity will drop one order when we change the defining function into geodesic defining function. The geodesic compactification ought to be at least C^2 so that we can calculate some curvature tensor. That's why we need $m \geq 3$ in Theorem 1.1. Then with the property of harmonic coordinates, we obtain the Neumann condition for $g^{0\beta}$.

Finally, in Section 5, we use the theory of elliptic system to prove Theorem 1.1. We improve the regularity of the conformal metric and defining function in the new coordinates.

2. Basic geometry equations in harmonic coordinates

In this section, we discuss some basic geometry equations for the manifold $(\overline{M}, g)$. Before doing that, we choose an appropriate conformal compactification to make the scalar curvature constant near boundary and to construct harmonic coordinates for the metric.

2A. *Constant scalar curvature compactification.*

Lemma 2.1. *Let (M, g_+) be a conformally compact n-manifold. M has a $C^{2,\alpha}$ conformal compactification $g = \rho^2 g_+$ and $h = g|_{\partial M}$ is the boundary metric. Then there exists a $C^{2,\alpha}$ constant scalar curvature compactification $\hat{g} = \hat{\rho}^2 g_+$ with boundary metric h.*

Proof. We only need to solve a Yamabe problem with Dirichlet data. Let $\hat{g} = u^{4/(n-2)}g$; then we consider the equation

$$\begin{cases} \Delta_g u - \dfrac{n-2}{4(n-1)}Su + \dfrac{n-2}{4(n-1)}\lambda u^{(n+2)/(n-2)} = 0, \\ u > 0 \qquad \text{in } \overline{M}, \\ u \equiv 1 \qquad \text{on } \partial M. \end{cases} \tag{2-1}$$

Let $\lambda = -1$; from [Ma 1995] we know that the equation always has a $C^{2,\alpha}$ solution. Then $\hat{g} = u^{4/(n-2)}g$ is also in $C^{2,\alpha}$. The boundary metric h is not changed since $u \equiv 1$ on ∂M. □

By the standard theory for elliptic equations, if g is in $C^{m,\alpha}(\overline{M})$ for some $\alpha \in (0, 1)$, any such solution u is in $C^{m,\alpha}(\overline{M})$. Then $\hat{\rho} = u^{2/(n-2)}\rho$ and $\hat{g}$ are also in $C^{m,\alpha}$. The Weyl tensor $\hat{W} = u^{4/(n-2)}W$ of $\hat{g}$ are in $C^{m,\alpha}$. In the following, we don't distinguish g and $\hat{g}$. When we refer to the compactification g, we mean that the scalar curvature S of g is constant.

2B. ***The harmonic coordinates near boundary.*** We call the coordinates $\{x^\beta\}_{\beta=0}^n$ harmonic coordinates with respect to g when $\Delta_g x^\beta = 0$ for $0 \le \beta \le n$. We are now going to construct harmonic coordinates near ∂M which are also harmonic when restricted on $(\partial M, h)$. In the following, if there are no special instructions, any use of indices will follow the convention that Roman indices will range from 1 to n, while Greek indices range from 0 to n.

Firstly, for any point $p \in \partial M$, there is smooth structure $\{y^\beta\}$. It is easy to construct harmonic coordinates $\{x^i\}$ with respect to $(\partial M, h)$. When h is in $C^{m+2,\alpha}$, we have that x^i are $C^{m+3,\alpha}(\partial M)$ functions of y^β. Then

$$h_{ij} = h\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \in C^{m+2,\alpha}(\partial M).$$

Then by solving a local Dirichlet problem: $\Delta_g x^i = 0$, with the boundary condition x^i as above, we can extend x^i to $\overline{M}$. Similarly, we can choose a harmonic defining function x^0 which satisfies $\Delta_g x^0 = 0, x^0|_{\partial M} = 0$. Then $\{x^\beta\}_{\beta=0}^n$ form harmonic coordinates with respect to g in a neighborhood of ∂M. When g is in $C^{m,\alpha}$, x^β are $C^{m+1,\alpha}$ functions of y^β. Then

$$g_{\alpha\beta} = g\left(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\right) \in C^{m,\alpha}(\overline{M}).$$

Therefore when we change the coordinates $\{y^\beta\}$ to $\{x^\beta\}$, the regularities of g, the Weyl tensor on $\overline{M}$ and h on ∂M are unchanged.

2C. ***Geometry equations.*** Now we study the relationship between Weyl tensors and Ricci tensors. For any $(n+1)$-manifold, $0 \le i, j, k, l, h, m \le n$, we have

$$R_{ijkl} = \frac{1}{n-1}(R_{ik}g_{jl} + R_{jl}g_{ik} - R_{jk}g_{il} - R_{il}g_{jk}) - \frac{S}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}) + W_{ijkl}. \tag{2-2}$$

On the other hand, we have the second Bianchi identity for g.

$$0 = \mathrm{Bian}(g, \mathrm{Ric}) = g^{jh}\big(R_{jk,h} - \tfrac{1}{2}R_{jh,k}\big)$$

When the scalar curvature S is constant, we get

$$g^{jh}R_{jk,h} = 0. \tag{2-3}$$

From (2-2) and (2-3), we have

$$R_{ik,l} - R_{il,k} = \frac{n-1}{n-2}g^{jh}W_{ijkl,h}.$$

It follows

$$g^{lt}R_{ik,lt} - g^{lt}R_{il,kt} = \frac{n-1}{n-2}g^{lt}g^{jh}W_{ijkl,ht}.$$

Note that

$$R_{il,kt} = R_{il,tk} + \mathrm{Ric} * Rm, \quad g^{lt}R_{il,tk} = (g^{lt}R_{il,t})_k = 0.$$

Where $\mathrm{Ric} * Rm$ refers to a bilinear form of R_{ij} and R_{ijkl}. We finally get

$$g^{lt}R_{ik,lt} = g * \mathrm{Ric} * Rm + g * g * (\partial^2 W + \Gamma * \partial W + \Gamma * \Gamma * W + \partial\Gamma * W). \tag{2-4}$$

If $g \in C^{m,\alpha}$, in harmonic coordinates $\{x^\beta\}$, (2-4) can be written in the form:

$$g^{\gamma\tau}\frac{\partial}{\partial x^\gamma}\frac{\partial}{\partial x^\tau}R_{\alpha\beta} = Q(g, \partial g, \partial^2 g, W, \partial W, \partial^2 W), \tag{2-5}$$

where Q is a polynomial. Even when $m = 3$, we can define the first and second covariant derivatives of curvature in the sense of distribution (see [Jin 2017; Shi and Tian 2005]). In this case (2-5) still holds. For more details about geometric equations, see [Petersen 1998].

3. The conformal infinity of asymptotically hyperbolic metrics

We will discuss some background material for conformally compact metrics in this section. Just as in the definition in introduction, we let ρ be a defining function for M, and set

$$g = \rho^2 g_+.$$

We assume ρ is $C^{m,\alpha}$ on $\overline{M}$ in the initial atlas $\{y^\beta\}$, and that m, α are defined as in Theorem 1.1. We have that g_+ satisfies

$$E(g_+) = \mathrm{Ric}_{g_+} + ng_+, \quad \|E(g_+)\|_{g_+} = o(\rho^2).$$

The curvature of g can be expressed as the following formulas:

$$K_{ab}=\frac{K_{+ab}+|\nabla\rho|^2}{\rho^2}-\frac{1}{\rho}[D^2\rho(e_a,e_a)+D^2\rho(e_b,e_b)]. \tag{3-1}$$

$$\mathrm{Ric}=-(n-1)\frac{D^2\rho}{\rho}+\left[\frac{n(|\nabla\rho|^2-1)}{\rho^2}-\frac{\Delta\rho}{\rho}\right]g+E(g_+), \tag{3-2}$$

$$S=-2n\frac{\Delta\rho}{\rho}+n(n+1)\frac{|\nabla\rho|^2-1}{\rho^2}+\mathrm{tr}_g\,E(g_+). \tag{3-3}$$

Here D^2 is the Hessian. $E(g_+)$ satisfies $\|\mathrm{tr}_g\,E(g_+)\|=\|E(g_+)\|_g=o(1)$, which implies the tensor $E=E(g_+)$ vanishes on ∂M.

When g is at least C^2 in $\overline{M}$, $|\nabla\rho|\to 1$ as $\rho\to 0$, and $|K_{+ab}+1|=O(\rho^2)$. Hence a C^2 conformally compact Einstein manifold is asymptotically hyperbolic. On ∂M, we have $|\nabla\rho|=1$. Let $D^2\rho|_{\partial M}=A$ denote the second fundamental form of ∂M in $(\overline{M},g)$. The equation (3-2) further implies that ∂M is umbilic, i.e., $A=\lambda h$, for some function λ on ∂M, see Lemma 4.2.

3A. ***Geodesic conformal compactification.*** As noted in the introduction, defining functions are unique only up to multiplication by positive functions on $\overline{M}$. From the formulas (3-2) and (3-3), we can see that it is special if the defining function r and its compactification $\bar{g}=r^2g_+$ satisfies

$$|\overline{\nabla}r|_{\bar{g}}\equiv 1$$

in a neighborhood of ∂M. Such defining functions are called geodesic defining functions. In fact, the defining functions always exist and were first introduced by Graham and Lee [1991].

Lemma 3.1. *Let g be a C^2 conformal compactification of (M,g_+), $g=\rho^2g_+$. $h=g|_{\partial M}$ is the boundary metric. Then g_+ has a unique geodesic conformal compactification with the same boundary metric h.*

Proof. Let $r=u\rho$, $\bar{g}=r^2g_+$. The lemma is equivalent to the equation:

$$\begin{cases}2(\nabla\rho)(\log u)+\rho|\nabla\log u|_g^2=\frac{1-|\nabla\rho|_g^2}{\rho},\\ u\equiv 1\qquad\text{on }\partial M.\end{cases} \tag{3-4}$$

By general theory of first order partial differential equations, we know that it has a unique positive solution in a collar neighborhood U of ∂M. □

However, that $\bar{g}$ may not be as smooth as g. When g is in $C^{m,\alpha}$, we only have $u\in C^{m-1,\alpha}$, hence $\bar{g}$ is only a $C^{m-1,\alpha}$ conformal compactification. Here we refer the reader to [Lee 1995, Lemma 5.1] for more details concerning the Hölder regularity.

It is obvious that when $\|\mathrm{Ric}_{g_+}+ng_+\|_{g_+}=o(\rho^2)$ for the defining function ρ, we also have $\|\mathrm{Ric}_{g_+}+ng_+\|_{g_+}=o(r^2)$ for the geodesic defining function.

In the following of this section, we assume g_+ is Einstein, i.e.,

$$E(g_+) = \mathrm{Ric}_{g_+} + ng_+ = 0.$$

Under this condition, the terms $E(g_+)$ and $\mathrm{tr}_g\, E(g_+)$ in (3-2) and (3-3) do not exist.

When M has a C^2 geodesic conformal compactification $\bar{g}$, it is very convenient for us to do some calculations. From (3-2) and (3-3), we know that the second fundamental form $\bar{A}$ given by $\bar{A} = \bar{D}^2 r = 0$ on ∂M, i.e., ∂M is totally geodesic in $\overline{M}$. The Gauss lemma implies that $\bar{g}$ can split in U,

$$\bar{g} = dr^2 + g_r$$

for a 1-parameter family g_r of metrics on ∂M.

Now we choose the local coordinates $(r, x^1, x^2, \dots, x^n)$ near ∂M to study the regularity of $\bar{g}$ near the boundary ∂M. Using (3-2) and the Gauss–Codazzi equation, we finally get

$$\begin{aligned} r\partial_r^2 \bar{g}_{ij} - (n-1)\partial_r \bar{g}_{ij} - \bar{g}^{kl}\partial_r \bar{g}_{kl}\bar{g}_{ij} - r\bar{g}^{kl}\partial_r \bar{g}_{kl}\partial_r \bar{g}_{ij} \\ + \frac{r}{2}\bar{g}^{kl}\partial_r \bar{g}_{ik}\partial_r \bar{g}_{jl} - 2r\overline{\mathrm{Ric}}(g_r)_{ij} = 0. \end{aligned} \tag{3-5}$$

Here $\overline{\mathrm{Ric}}(g_r)_{ij}$ denotes the Ricci tensor for the induced metric on level sets of r. We assume that $\bar{g}$ is smooth enough so that we could make its expansion from (3-5). Let $r = 0$, we can derive that $\partial_r \bar{g}(0) = 0$. By using mathematical induction, we differentiate (3-5) $p-1$ times with respect to r

$$\begin{aligned} r\partial_r^{p+1}\bar{g}_{ij} + (p-n)\partial_r^p \bar{g}_{ij} - \bar{g}^{kl}(\partial_r^p \bar{g}_{kl})\bar{g}_{ij} \\ = Q_1(\bar{g}^{-1}, \partial_r^q \bar{g}, \partial_r^q \bar{g}\partial_s^2 \bar{g}) + rQ_2(\bar{g}^{-1}, \partial_r^q \bar{g}, \partial_r^p \bar{g}, \partial_r^{p-2}\partial_s^2 \bar{g}). \end{aligned} \tag{3-6}$$

Here $q < p$ and $\partial_s \bar{g}$ is the differential of $\bar{g}$ with respect to $x^i (1 \le i \le n)$. Q_1, Q_2 are the third polynomials.

Setting $r = 0$, (3-6) implies that $\partial_r^p \bar{g}(0) = 0$ when p is odd and that $\partial_r^p \bar{g}(0)$ is uniquely determined by each step when p is even. However, this will break down when $p = n$ and $n+1$ is odd. In this case we only have $\bar{g}^{kl}(\partial_r^n \bar{g}_{kl}) = 0$ at $r = 0$. It gives no further information at this order. To be specific, $\partial_r \bar{g}(0) = \partial_r^3 \bar{g}(0) = \cdots = \partial_r^{n-1}\bar{g}(0) = 0$, which means

$$\partial_r^{n-1}\bar{g}(r) = O(1).$$

Let's consider (3-6) when $p = n-1$,

$$\partial_r^n \bar{g}_{ij} = \frac{\partial_r^{n-1}\bar{g}_{ij} - \bar{g}^{kl}(\partial_r^{n-1}\bar{g}_{kl})\bar{g}_{ij} + Q_1}{r} + Q_2. \tag{3-7}$$

If $\partial_r^{n-1}\bar{g}(r) \neq O(r)$, there will be the log term in the expansion for g_r. Note that $\partial_r^{n-1}\bar{g}(r) = O(r)$ is the necessary condition to ensure $\bar{g}$ is n-th differentiable.

3B. ***Weyl tensor in geodesic coordinates.*** In this subsection we shall show that if the Weyl tensor $\overline{W}$ of $\bar{g}$ is $(n-2)$-th differentiable, then $\partial_r^{n-1}\bar{g}(r) = O(r)$ holds.

We begin with (2-4) by taking $x^l = r$, $1 \le i, k \le n$.

In the coordinates $(r, x^1, x^2, \dots, x^n)$, we have

$$\begin{aligned}\bar{R}_{ik,r} &= \partial_r^3\bar{g} + \partial_r\bar{g} * \partial_r^2\bar{g} * \bar{g} + \partial_r\bar{g} * \partial_r\bar{g} * \partial_r\bar{g} * \bar{g} + \partial_r P + \partial_r\bar{g} * P,\\ \bar{R}_{ir,k} &= \partial_r P * P + \partial_r\bar{g} * \partial_r^2\bar{g} * \bar{g} + \partial_r\bar{g} * \partial_r\bar{g} * \partial_r\bar{g} * \bar{g} + \partial_r\bar{g} * \bar{g} * \overline{\mathrm{Ric}}.\end{aligned}$$

Here $P = P(\bar{g}, \partial_{x^i}\bar{g}, \partial_{x^i}^2\bar{g})$ is a polynomial for $1 \le i \le n$. From (2-4) we have:

$$\partial_r^3\bar{g} = \partial_r\bar{g} * \partial_r^2\bar{g} * \bar{g} + \partial_r\bar{g} * \partial_r\bar{g} * \partial_r\bar{g} * \bar{g} + \partial_r P * P + \partial_r\bar{g} * \bar{g} * \overline{\mathrm{Ric}} + \bar{g} * \nabla\overline{W}. \tag{3-8}$$

We already know that $\partial_r g = \partial_r^3 g = \cdots \partial_r^{n-3} = O(r)$. Then differentiating (3-8) $n-4$ times with respect to r, we get that each term on the right-hand side is $O(r)$ apart from $\bar{g} * \nabla\overline{W}$. The left side is $\partial_r^{n-1}\bar{g} = O(1)$. Hence $\partial_r^{n-3}\overline{W} = O(1)$. When $\overline{W} \in C^{n-2}$, we get that $\partial_r^{n-3}\overline{W} = O(r)$. So the left-hand side is $\partial_r^{n-1}\bar{g} = O(r)$. This gives the necessary condition that $\bar{g}$ has C^n regularity or higher.

We calculate the formal power series of Weyl tensor in geodesic coordinates. We have $g = dr^2 + g_r$. From (3-6), (3-7) and [Graham 2000], when n is odd,

$$g_r = h + g^{(2)}r^2 + (\text{even powers of } r) + g^{(n-1)}r^{n-1} + g^{(n)}r^n + \cdots . \tag{3-9}$$

When n is even,

$$g_r = h + g^{(2)}r^2 + (\text{even powers of } r) + g^{(n)}r^n + f r^n \log r + \cdots , \tag{3-10}$$

where $g^{(2i)}$ and f are two tensors on ∂M and f is trace free and determined by h locally. We now consider the case when n is even. If $f = 0$, g is at least in C^n. Let $1 \le i, j \le n$.

$$W_{irjr} = R_{irjr} - \frac{1}{n-1}(R_{ik} + R_{rr}g_{ij}) + \frac{S}{n(n-1)}g_{ij} \tag{3-11}$$

The formal power series of the Weyl tensor contain r and $\log r$ and we only need to check the coefficients of $r^{n-2}\log r$. A short calculation reveals that the coefficients of $r^{n-2}\log r$ of R_{irjr} are

$$\operatorname{coeff}(R_{irjr}) = -\frac{n(n-1)}{2} f_{ij}.$$

And also

$$\begin{aligned}\operatorname{coeff}(R_{ij}) &= -\frac{n(n-1)}{2} f_{ij},\\ \operatorname{coeff}(R_{rr}) &= -\frac{n(n-1)}{2} h^{st} f_{st},\\ \operatorname{coeff}(S) &= -\frac{n(n-1)}{2} 2h^{st} f_{st}.\end{aligned}$$

When coeff$(W_{irjr}) = 0$, by (3-11) we finally derive

$$f_{ij} - \frac{1}{n}(h^{st} f_{st})h_{ij} = 0. \tag{3-12}$$

Since f is trace free, $f \equiv 0$. This gives the n-th regularity of g. From Theorem A in [Chruściel et al. 2005], we know the log term is the only obstruction of the smoothness, and hence Remark 1.3 holds.

4. The boundary condition

In this section, we derive a boundary problem for g and Ricci curvature of a conformal compact Einstein manifold in the harmonic coordinates as defined in Section 2. The compactification we use here will not be a geodesic compactification, but rather one which has constant scalar curvature. We do it locally, that is, for any $p \in \partial M$, there is a neighborhood V which contains p and local atlas $\{x^\beta\}$. Let $D = V \cap \partial M$ be the boundary portion. $W \in C^{m,\alpha}(V)$, $g \in C^{m,\alpha}(V)$, $h \in C^{m+2,\alpha}(D)$. We will give the Dirichlet and Neumann boundary conditions of g and $\mathrm{Ric}(g)$ on D.

4A. ***Dirichlet boundary conditions on g_{ij}.***

$$g_{ij} = h_{ij}. \tag{4-1}$$

4B. ***Dirichlet boundary conditions on R_{ij}.*** We claim that

$$R_{ij} = \frac{n-1}{n-2}(\mathrm{Ric}_h)_{ij} + \left(\frac{1}{2n} S - \frac{1}{2(n-2)} S_h\right) h_{ij} + \frac{n-1}{2n^2} H^2 h_{ij}. \tag{4-2}$$

Here Ric_h and S_h are the Ricci curvature and scalar curvature of $(\partial M, h)$, and H is the mean curvature, $H = g^{ij} A_{ij}$. We use the following three lemmas to prove (4-2).

Lemma 4.1. *Let $\bar{g} = r^2 g_+$ be a C^2 geodesic compactification of (M, g_+) with boundary metric h on ∂M. Then on ∂M,*

$$\bar{S} = \frac{n}{n-1} S_h. \tag{4-3}$$

$$\bar{R}_{ij} = \frac{n-1}{n-2}(\mathrm{Ric}_h)_{ij} - \frac{1}{2(n-1)(n-2)} S_h h_{ij}. \tag{4-4}$$

Proof. Since we only need to study the Ricci curvature on $T\partial M$, we can choose the coordinates $(r, x^1, \dots, x^n)$ in V where $(x^1, \dots, x^n)$ are harmonic with respect to h when restricted on D. We have $\bar{g} = dr^2 + g_r$, i.e.,

$$g_{ri} = g^{ri} = 0, \quad g_{rr} = g^{rr} = 1.$$

Since the second fundamental form $\bar{A} = \bar{D}^2 r$ vanishes on ∂M,

$$\begin{aligned}\bar{R}_{ij} &= \bar{g}^{\alpha\beta} \bar{R}_{i\alpha\beta j} \\ &= \bar{g}^{kl}((R_h)_{iklj} + \bar{A}_{il}\bar{A}_{kj} - \bar{A}_{ij}\bar{A}_{kl}) + \bar{R}_{irrj} \\ &= (R_h)_{ij} + \bar{R}_{irrj}.\end{aligned} \tag{4-5}$$

Taking the trace for i, j, we get

$$\bar{R}_{rr} = \tfrac{1}{2}(\bar{S} - S_h). \tag{4-6}$$

For $\bar{R}_{irrj}$, we have:

$$\begin{aligned}\bar{R}_{irrj} &= \bar{g}(\bar{\nabla}_{\partial_i}\bar{\nabla}_{\partial_r}\partial_r, \partial_j) - \bar{g}(\bar{\nabla}_{\partial_r}\bar{\nabla}_{\partial_i}\partial_r, \partial_j) - \bar{g}(\bar{\nabla}_{[\partial_r,\partial_i]}\partial_r, \partial_j) \\ &= -\partial_r \bar{g}(\bar{\nabla}_{\partial_i}\partial_r, \partial_j) + \bar{g}(\bar{\nabla}_{\partial_i}\partial_r, \bar{\nabla}_{\partial_r}\partial_j) \\ &= -\partial_r \bar{A}_{ij}.\end{aligned} \tag{4-7}$$

From (3-2) and (3-3),

$$\bar{R}_{ij} = -(n-1)\frac{\bar{A}_{ij}}{r} - \frac{\bar{\Delta} r}{r}\bar{g}_{ij} + E_{ij}, \tag{4-8}$$

$$\bar{S} = -2n\frac{\bar{\Delta} r}{r} + \operatorname{tr}_g E. \tag{4-9}$$

The tensor E and $\operatorname{tr}_g E$ vanish on the boundary, hence on ∂M,

$$\bar{S} = -2n\partial_r \bar{\Delta} r. \tag{4-10}$$

Taking the trace on (4-8),

$$\bar{S} - \bar{R}_{rr} = -(n-1)\frac{\bar{H}}{r} - n\frac{\bar{\Delta} r}{r} + \bar{g}^{ij}E_{ij}.$$

By l'Hôspital's rule, with $r \to 0$,

$$\partial_r \bar{A}_{ij}|_{r=0} = -\frac{1}{n-1}(\bar{R}_{ij} + \partial_r \bar{\Delta} r\,\bar{g}_{ij})|_{r=0}, \tag{4-11}$$

$$\partial_r \bar{\Delta} r|_{r=0} = -\frac{1}{n}(\bar{S} - \bar{R}_{rr} + (n-1)\partial_r \bar{H})|_{r=0}. \tag{4-12}$$

At last, we only need to calculate $\partial_r \bar{H}|_{r=0}$.

$$\begin{aligned}\partial_r \bar{H}|_{r=0} &= \partial_r(\bar{g}^{ij}\bar{A}_{ij})|_{r=0} = (\partial_r \bar{g}^{ij})\bar{A}_{ij}|_{r=0} + \bar{g}^{ij}(\partial_r \bar{A}_{ij})|_{r=0} \\ &= 0 + \bar{g}^{ij}(-\bar{R}_{irrj}) = -\bar{R}_{rr}.\end{aligned} \tag{4-13}$$

Combining the formulas above, we finally conclude (4-3) and (4-4). □

Lemma 4.2. *Let $g = \rho^2 g_+$ be a $C^{3,\lambda}$ conformal compactification and $\bar{g} = r^2 g_+$ be the $C^{2,\lambda}$ geodesic conformal compactification of (M, g_+) with the same boundary metric $g|_{\partial M} = \bar{g}|_{\partial M} = h$. If $r = u\rho$, $A = D^2\rho$, then $A|_{\partial M} = -u_r h$.*

Proof. We begin with the coordinates $(r, x^1, x^2, \dots, x^n)$. On ∂M, $0 = \bar{A}_{ij} = -\overline{\Gamma}^r_{ij}$. The connection ∇ and $\overline{\nabla}$ of g and $\bar{g}$ have the following relationship:

$$\Gamma^r_{ij} = \overline{\Gamma}^r_{ij} - \frac{1}{u}(\delta^r_j u_i + \delta^r_i u_j - g_{ij} u_r) = \frac{1}{u} u_r h_{ij}.$$

We have $g = u^{-2}\bar{g}$, $\operatorname{grad}_g = u^2 \operatorname{grad}_{\bar{g}}$.

$$\begin{aligned} A_{ij} &= D^2\rho(\partial_i, \partial_j) = g(\nabla_{\partial_i}\nabla\rho, \partial_j) = -g(\nabla\rho, \nabla_{\partial_i}\partial_j) \\ &= -\Gamma^r_{ij} g(\nabla\rho, \partial_r) = -\Gamma^r_{ij}\bar{g}(\overline{\nabla}\rho, \partial_r) \\ &= -\Gamma^r_{ij}\bar{g}\left(\overline{\nabla}\left(\frac{r}{u}\right), \partial_r\right) = -\Gamma^r_{ij}\bar{g}\left(\frac{u\overline{\nabla} r - r\overline{\nabla} u}{u^2}, \partial_r\right) \\ &= -\Gamma^r_{ij}\bar{g}(\overline{\nabla} r, \overline{\nabla} r) = -u_r h_{ij}. \end{aligned} \tag{4-14}$$

□

Lemma 4.2 tells us $u_r = -H/n$. Since $u|_{\partial M} \equiv 1$, we get

$$\overline{\nabla} u = -\frac{H}{n}\overline{\nabla} r.$$

Lemma 4.3. *We define g, $\bar{g}$ as in Lemma 4.2. Then on ∂M,*

$$\begin{aligned} R_{ir} &= -\frac{n-1}{n}\frac{\partial H}{\partial x_i}, \\ R_{rr} &= \tfrac{1}{2}(S - S_h) - \frac{n-1}{2n} H^2, \\ R_{ij} &= \bar{R}_{ij} + \left(\frac{1}{2n}(S - \bar{S})\right) h_{ij} + \frac{n-1}{2n^2} H^2 h_{ij}. \end{aligned} \tag{4-15}$$

Proof. By the standard formulas for conformal changes of the metric $g = u^{-2}\bar{g}$, the Ricci curvature of g and $\bar{g}$ are related by

$$\mathrm{Ric} = \overline{\mathrm{Ric}} + (n-1)\frac{\bar{D}^2 u}{u} + \left(\frac{\bar{\Delta} u}{u} + \frac{n|\overline{\nabla} u|^2_{\bar{g}}}{u^2}\right)\bar{g},$$

Since

$$\begin{aligned} \bar{\Delta} u &= \operatorname{div}\overline{\nabla} u = \operatorname{div}\left(-\frac{H}{n}\overline{\nabla} r\right) = -\frac{\partial_r H}{n}, \\ \bar{D}^2 u(\partial_i, \partial_j) &= 0, \\ \bar{D}^2 u(\partial_i, \partial_r) &= u_{ir} = -\frac{1}{n}\frac{\partial H}{\partial x_i}, \\ \bar{D}^2 u(\partial_r, \partial_r) &= -\frac{\partial_r H}{n} = \bar{\Delta} u; \end{aligned}$$

then on ∂M,

$$R_{ir} = \bar{R}_{ir} - \frac{n-1}{n}\frac{\partial H}{\partial x_i} = -\frac{n-1}{n}\frac{\partial H}{\partial x_i},$$

$$R_{rr} = \bar{R}_{rr} + n\bar{\Delta}u + \frac{H^2}{n},$$

$$R_{ij} = \bar{R}_{ij} + \left(\bar{\Delta}u + \frac{H^2}{n}\right)h_{ij}. \tag{4-16}$$

Therefore

$$S = \bar{S} + 2n\bar{\Delta}u + \frac{n+1}{n}H^2.$$

This implies

$$\bar{\Delta}u = \frac{1}{2n}\left(S - \bar{S} - \frac{n+1}{n}H^2\right). \tag{4-17}$$

Lemma 4.3 follows from (4-16) and (4-17). □

Finally, Lemmas 4.1 and 4.3 imply the formula (4-2).

4C. ***Neumann boundary conditions on $g^{0\alpha}$.*** We use the harmonic coordinates $\{x^\beta\}_{\beta=0}^n$ as defined in Section 2. Let $N = \nabla x_0/|\nabla x_0|$ be the unit norm vector on ∂M. In coordinates $\{x^\beta\}_{\beta=0}^n$,

$$N = (g^{00})^{-1/2} g^{0\beta}\partial_\beta.$$

These are the Neumann conditions:

$$N(g^{00}) = -2Hg^{00}. \tag{4-18}$$

$$N(g^{0i}) = -Hg^{0i} + \tfrac{1}{2}(g^{00})^{-1/2} g^{i\beta}\partial_\beta g^{00}. \tag{4-19}$$

Proof. Let $\{e_i\}_{i=1}^n$ be the orthonormal basis at a given point $p \in \partial M$.

$$0 = \Delta x^\alpha = \operatorname{div} \nabla x^\alpha = \sum_{i=1}^n g(\nabla_{e^i}\nabla x^\alpha, e_i) + g(\nabla_N \nabla x^\alpha, N) \tag{4-20}$$

Write $\nabla x^\alpha = (\nabla x^\alpha)^T + (\nabla x^\alpha)^N$, where

$$(\nabla x^\alpha)^T = \nabla_h x^\alpha, \quad (\nabla x^\alpha)^N = g(\nabla x^\alpha, N)N = (g^{00})^{-1/2} g^{0\alpha} N.$$

We have

$$0 = \Delta_h x^j = \operatorname{div} \nabla_h x^j = \sum_{i=1}^n h(\nabla^h_{e^i}(\nabla x^j)^T, e_i) = \sum_{i=1}^n g(\nabla_{e^i}(\nabla x^j)^T, e_i),$$

which also holds for $j = 0$ as $x^0 \equiv 0$ on ∂M. Then (4-20) turns into

$$\sum_{i=1}^n g(\nabla_{e^i}((g^{00})^{-1/2} g^{0\alpha} N), e_i) + g(\nabla_N \nabla x^\alpha, N) = 0 \tag{4-21}$$

The first term is just $(g^{00})^{-1/2}g^{0\alpha}\sum_{i=1}^{n}A(e_i,e_i)=H(g^{00})^{-1/2}g^{0\alpha}$ and the second term is

$$
\begin{aligned}
g(\nabla_N\nabla x^\alpha,N)&=\frac{1}{|\nabla x^0|^2}g(\nabla_{\nabla x^0}\nabla x^\alpha,\nabla x^0)\\
&=\frac{1}{|\nabla x^0|^2}[\nabla x^0(g(\nabla x^\alpha,\nabla x^0))-g(\nabla x^\alpha,\nabla_{\nabla x^0}\nabla x^0)]\\
&=\frac{1}{|\nabla x^0|^2}[\nabla x^0(g^{0\alpha})-g(\nabla x^0,\nabla_{\nabla x^\alpha}\nabla x^0)]\\
&=\frac{1}{|\nabla x^0|}N(g^{0\alpha})-\frac{1}{2|\nabla x^0|^2}\nabla x^\alpha(g^{00})
\end{aligned}
\tag{4-22}
$$

We know $|\nabla x^0|=(g^{00})^{1/2}$, then (4-21) is just

$$
Hg^{0\alpha}+N(g^{0\alpha})-\tfrac{1}{2}(g^{00})^{-1/2}\nabla x^\alpha(g^{00})=0.
\tag{4-23}
$$

Letting $\alpha=0$, we have $N(g^{00})=-2Hg^{00}$. Letting $\alpha=i$, we have

$$
N(g^{0i})=-Hg^{0i}+\tfrac{1}{2}(g^{00})^{-1/2}g^{i\beta}\partial_\beta g^{00}. \qquad \square
$$

4D. ***Dirichlet boundary conditions on $R_{\alpha\beta}$.*** We already know the formulas of R_{ij} on ∂M and the mixed components R_{ri} and R_{rr} of Ricci Curvature in the coordinates $(r,x^1,x^2,\dots,x^n)$ in Section 4B. That is,

$$
\begin{aligned}
R_{ir}&=-\frac{n-1}{n}\frac{\partial H}{\partial x_i},\\
R_{rr}&=\tfrac{1}{2}(S-S_h)-\frac{n-1}{2n}H^2.
\end{aligned}
$$

Now we study the the mixed components R_{0i} and R_{00} of Ricci Curvature in the harmonic coordinates $(x^0,x^1,\dots,x^n)$. In fact, as the vector $N=\nabla x_0/|\nabla x_0|$ is also the unit norm vector on ∂M with respect to g, we have $N=\overline{\nabla}r=\nabla r$. That is, $\nabla r=(g^{00})^{-1/2}g^{0\beta}\partial_\beta$. Then

$$
R_{0i}=-(g^{00})^{-1/2}\,\frac{n-1}{n}\frac{\partial H}{\partial x_i}-\frac{g^{0j}}{g^{00}}R_{ij}
\tag{4-24}
$$

$$
R_{00}=\frac{1}{(g^{00})^2}\left(g^{0i}g^{0j}R_{ij}+g^{00}\left(\tfrac{1}{2}(S-S_h)-\frac{n-1}{2n}H^2\right)\right)
\tag{4-25}
$$

4E. ***Neumann boundary conditions on R_{0i}.*** The Dirichlet conditions for R_{0i} in (4-24) are not good because there are still second order differential terms of metric on the right side. Now we consider the differential terms of Ricci curvature.

Since the scalar curvature is constant, with the second Bianchi identity, we have

$$
0=\tfrac{1}{2}S_{,\alpha}=g^{\eta\beta}R_{\eta\alpha,\beta}=g^{\eta\beta}\partial_\beta R_{\alpha\eta}-g^{\eta\beta}\Gamma^{\tau}_{\alpha\beta}R_{\eta\tau}.
$$

Then

$$g^{0\beta}\partial_\beta R_{0\alpha} = -g^{j\beta}\partial_\beta R_{j\alpha} + g^{\eta\beta}\Gamma^\tau_{\alpha\beta}R_{\eta\tau}.$$

Letting $\alpha = i$, we get

$$N(R_{0i}) = (g^{00})^{-1/2}(-g^{j\beta}\partial_\beta R_{ji} + g^{\eta\beta}\Gamma^\tau_{i\beta}R_{\eta\tau}). \tag{4-26}$$

5. Proof of the main theorem

In this section, we prove the main theorem. Suppose $m \geq 3$ and $\alpha \in (0, 1)$. For any point $p \in \partial M$, we select the harmonic coordinates $\{x^\beta\}_{\beta=0}^n$ in its neighborhood V. Let $D = V \cap \partial M$ be the boundary portion. Now we have $g \in C^{m,\alpha}(V)$, $h \in C^{m+2,\alpha}(D)$, the Weyl tensor $W \in C^{m,\alpha}(V)$.

5A. *Regularity of the metric.*

Step 1 (regularity of the Ricci curvature). We begin with the elliptic equation (2-5), with the right side of (2-5) in $C^{m-2,\alpha}$.

On ∂M, we already derived the formulas of $R_{\alpha\beta}$ in Section 3. We have $H = g^{ij}A_{ij}$, where $A_{ij} = \frac{1}{2}(g^{00})^{1/2}g^{0\beta}(\partial_\beta g_{ij} - \partial_i g_{\beta j} - \partial_j g_{\beta i}) \in C^{m-1,\alpha}(D)$. Then the Dirichlet condition of R_{ij} and R_{00} given by (4-4) and (4-25) shows that

$$R_{ij} \in C^{m-1,\alpha}(D), \quad R_{00} \in C^{m-1,\alpha}(D).$$

By standard elliptic regularity theory,

$$R_{ij} \in C^{m-1,\alpha}(V), \quad R_{00} \in C^{m-1,\alpha}(V).$$

Then by the Neumann boundary conditions on R_{0i} given by (4-26), we also have $R_{0i} \in C^{m-1,\alpha}(V)$ since $N(R_{0i}) \in C^{m-2,\alpha}(D)$.

In the following, we prove that $g_{\alpha\beta} \in C^{m+1,\alpha}(V)$ when $R_{\alpha\beta} \in C^{m-1,\alpha}(V)$. If it holds, we can get that $g_{\alpha\beta} \in C^{m+2,\alpha}(V)$ by repeating the steps.

Step 2 (regularity of g_{ij}). In harmonic coordinates, from [DeTurck and Kazdan 1981], we have

$$\Delta g_{ij} = -2R_{ij} + Q(g, \partial g).$$

Here Q contains a quadratic term of g and ∂g. Hence $\Delta g_{ij} \in C^{m-1,\alpha}(V)$, together with the boundary condition $g_{ij} = h_{ij} \in C^{m+2,\alpha}(D)$. From the standard elliptic theory in [Gilbarg and Trudinger 1977], we conclude that $g_{ij} \in C^{m+1,\alpha}(V)$.

Step 3 (regularity of $g_{0\beta}$). In Section 4, we obtain the Neumann boundary conditions of $g^{0\beta}$ which are expressed by mean curvature H. Since $H \in C^{m-1,\alpha}(D)$, we can't improve the regularity of $g^{0\beta}$ with this condition. Now we are going to calculate the oblique derivative of $g_{0\beta}$ on ∂M. As $g^{0\alpha}g_{\alpha\beta}=\delta^0_\beta$,

$$
\begin{aligned}
0 &= N(g^{0\alpha}g_{\alpha\beta}) = N(g^{00}g_{0\beta}) + N(g^{0j}g_{j\beta})\\
&= g^{00}N(g_{0\beta}) + g_{0\beta}N(g^{00}) + g_{j\beta}N(g^{0j}) + g^{0j}N(g_{j\beta})\\
&= g^{00}N(g_{0\beta}) + g_{0\beta}(-2Hg^{00})\\
&\qquad + g_{j\beta}\big(\tfrac{1}{2}(g^{00})^{-1/2}g^{j\tau}\partial_\tau g^{00} - Hg^{0j}\big) + g^{0j}N(g_{j\beta}) \qquad (5\text{-}1)\\
&= g^{00}N(g_{0\beta}) - 2Hg_{0\beta}g^{00}\\
&\qquad + \tfrac{1}{2}(g^{00})^{-1/2}(\delta^\tau_\beta - g_{0\beta}g^{0\tau})\partial_\tau g^{00} - H(-g_{0\beta}g^{00}) + g^{0j}N(g_{j\beta})\\
&= g^{00}N(g_{0\beta}) + \tfrac{1}{2}(g^{00})^{-1/2}\partial_\beta g^{00} + g^{0j}N(g_{j\beta}) - H\delta^0_\beta.
\end{aligned}
$$

Letting $\beta=0$, we get

$$g^{00}N(g_{00}) + \tfrac{1}{2}(g^{00})^{-1/2}\partial_0 g^{00} + g^{0j}N(g_{0j}) - H = 0. \tag{5-2}$$

Letting $\beta=i$, we get

$$g^{00}N(g_{0i}) + \tfrac{1}{2}(g^{00})^{-1/2}\partial_i g^{00} + g^{0j}N(g_{ij}) = 0. \tag{5-3}$$

Now we consider the elliptic system of g^{00}, $g_{01}, g_{02}, \dots, g_{0n}$:

$$
\begin{cases}
\Delta g^{00} = Q(g, \partial g, \mathrm{Ric}),\\
\Delta g_{01} = -\frac{1}{2}R_{01} + Q(g,\partial g),\\
\ \vdots\\
\Delta g_{0n} = -\frac{1}{2}R_{0n} + Q(g,\partial g).
\end{cases} \tag{5-4}
$$

Through the formulas (4-18), (5-3), the expression of H and the regularities of g_{ij}, we obtain the boundary condition:

$$
\begin{cases}
N(g^{00}) - 2Pg^{ij}\partial_i g_{0j} \in C^{m,\alpha}(D),\\
N(g_{01}) + \frac{1}{2P}\partial_1 g^{00} \in C^{m,\alpha}(D),\\
\ \vdots\\
N(g_{0n}) + \frac{1}{2P}\partial_n g^{00} \in C^{m,\alpha}(D),
\end{cases} \tag{5-5}
$$

where $P=(g^{00}(x))^{3/2}$. We are going to prove:

Lemma 5.1. *Let $u^0=g^{00}$, $u^i=g_{0i}$, $i=1,2,\dots,n$. The equations* (5-4) *can be written as the formula*

$$L_{\alpha\beta}u^\beta(x) = f_\alpha(x), \quad x\in V.$$

The boundary conditions (5-5) *can be written as*

$$B_{\alpha\beta}u^\beta(x) = g_\alpha(x), \quad x\in D.$$

Then the operator L is proper elliptic and the boundary operator B satisfies the complementing condition with respect to the system (L, B).

Proof. For any $n+1$ vector $\xi = (\xi_0, \xi_1, \ldots, \xi_n)$, we consider the principal part of L

$$L'_{\alpha\beta}(x,\xi) = \begin{bmatrix} g^{\alpha\beta}\xi_\alpha\xi_\beta & 0 & \cdots & 0 \\ 0 & g^{\alpha\beta}\xi_\alpha\xi_\beta & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g^{\alpha\beta}\xi_\alpha\xi_\beta \end{bmatrix} \tag{5-6}$$

Then

$$L_{\alpha\beta}(x,\xi) = \det(L'_{\alpha\beta}(x,\xi)) = |\xi|_g^{2(n+1)}.$$

For any $\xi \neq 0$, $L_{\alpha\beta}(x,\xi) \neq 0$, so L is elliptic.

For each pair of ξ and ξ' of linearly independent vectors, the equation

$$L_{\alpha\beta}(x, \xi + z\xi') = 0$$

is equivalent to

$$(z^2 \cdot |\xi'|_g^2 + 2 < \xi, \xi' >_g z + |\xi|_g^2)^{n+1} = 0.$$

It has $n+1$ roots with positive imaginary part and $n+1$ with negative imaginary part. So L is proper elliptic.

For any $x_0 \in D$, let $n = (1, 0, \ldots, 0)$ denote the unit normal at x_0 and $\xi = (0, \xi_1, \ldots, \xi_n)$ denote any nonzero real vector tangent to D at x_0. Let $z_s^+(x_0, \xi)$, $s = 0, 1, \ldots, n$ be the roots of $L_{\alpha\beta}(x_0, \xi + zn) = 0$ with positive imaginary part.

Define

$$L_0^+(x_0, \xi; z) = \prod_{s=0}^{n}(z - z_s^+(x_0, \xi).$$

Let $L^{\alpha\beta}(x_0, \xi + zn)$ be matrix adjoint to $L'_{\alpha\beta}(x_0, \xi + zn)$. Now we define

$$Q_{r\beta} = B'_{r\alpha}(x_0, \xi + zn) \cdot L^{\alpha\beta}(x_0, \xi + zn)$$

as polynomials in z, where $B'_{r\alpha}$ is the principal part of B.

Then B satisfies the complementing condition with respect to the system (L, B) if and only if the rows of the Q matrix are linearly independent modulo $L_0^+(x_0, \xi; z)$, that is, the polynomial

$$\sum_{r=0}^{n} C_r Q_{r\beta}(x_0, \xi; z) \equiv 0 \ (\mathrm{mod}\ L_0^+)$$

only if C_r are all 0.

By a direct computation, we can derive $z_s = \sqrt{-1}(|\xi|_h/\sqrt{g^{00}})$ for $s = 0, 1, \dots, n$. Then

$$L_0^+(x_0, \xi; z) = \left(z - \sqrt{-1}\frac{|\xi|_h}{\sqrt{g^{00}}}\right)^{n+1}.$$

From the above, we can get

$$L'_{\alpha\beta}(x_0, \xi + zn) = |\xi + zn|_g^2 \cdot \delta_{\alpha\beta} = (z^2 g^{00} + |\xi|_h^2) \cdot \delta_{\alpha\beta}.$$

Its adjoint matrix is

$$L^{\alpha\beta}(x_0, \xi + zn) = (z^2 g^{00} + |\xi|_h^2)^n \cdot \delta^{\alpha\beta}.$$

The principal part of B is

$$B'_{\alpha\beta}(x, \xi) = \begin{bmatrix} z & -2Pg^{i1}\xi_i & -2Pg^{i2}\xi_i & \cdots & -2Pg^{in}\xi_i \\ \frac{1}{2P}\xi_1 & z & 0 & \cdots & 0 \\ \frac{1}{2P}\xi_2 & 0 & z & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2P}\xi_n & 0 & 0 & \cdots & z \end{bmatrix} \tag{5-7}$$

Then

$$\begin{aligned} \sum_{r=0}^n C_r Q_{r\beta} &= \sum_{r=0}^n C_r B'_{r\alpha} \cdot L^{\alpha\beta} \\ &= \sum_{r=0}^n C_r B'_{r\alpha} \cdot (z^2 g^{00} + |\xi|_h^2)^n \cdot \delta^{\alpha\beta} \\ &= \sum_{r=0}^n C_r B'_{r\beta} \cdot (z^2 g^{00} + |\xi|_h^2)^n \\ &\equiv 0 \ \left(\mathrm{mod}\left(z - \sqrt{-1}\frac{|\xi|_h}{\sqrt{g^{00}}}\right)^{n+1}\right). \end{aligned} \tag{5-8}$$

This implies

$$z - \sqrt{-1}\frac{|\xi|_h}{\sqrt{g^{00}}} \,\Big|\, \sum_{r=0}^n C_r B'_{r\beta} \tag{5-9}$$

for any $0 \le \beta \le n$.

When $\beta \ge 1$, (5-9) implies

$$z - \sqrt{-1}\tfrac{|\xi|_h}{\sqrt{g^{00}}} \,\Big|\, \frac{C_0}{2P}\xi_\beta + C_\beta z.$$

Then

$$C_\beta = -\frac{C_0\sqrt{g^{00}}\xi_\beta}{2P|\xi|_h\cdot\sqrt{-1}}. \tag{5-10}$$

When $\beta = 0$, (5-9) implies

$$z - \sqrt{-1}\frac{|\xi|_h}{\sqrt{g^{00}}}\Big| C_0 z - 2PC_1 g^{i1}\xi_i - 2PC_2 g^{i2}\xi_i - \cdots - 2PC_n g^{in}\xi_i.$$

With (5-10), we have

$$z - \sqrt{-1}\frac{|\xi|_h}{\sqrt{g^{00}}}\Big| C_0 z + \frac{C_0\sqrt{g^{00}}|\xi|_h}{\sqrt{-1}}. \tag{5-11}$$

By a linear transformation, we can make $g^{00}(x_0) \neq 1$, for all $x_0 \in D$. Then (5-11) shows that $C_0 = 0$, and (5-10) implies $C_r = 0$, $r = 1, 2, \ldots, n$. □

Lemma 5.1 and Theorem 6.3.7 in [Morrey 1966] tell us that g^{00}, $g_{01}, \ldots, g_{0n}$ are all in $C^{m+1,\alpha}$. Then the boundary condition (5-2) can be written as

$$g^{00}N(g_{00}) = Q(g, \partial g^{00}, \partial g_{0i}, \partial g_{ij}) \in C^{m,\alpha}(D).$$

With the elliptic equation

$$\Delta g_{00} = -2R_{00} + Q(g, \partial g) \in C^{m-1,\alpha}(V),$$

we finally derive $g_{00} \in C^{m+1,\alpha}$.

Now we know that $g \in C^{m+1,\alpha}$. Back to step 1, we have $A_{ij} \in C^{m,\alpha}(D)$, thus $R_{\alpha\beta} \in C^{m,\alpha}(V)$. Repeating the steps above, we can get $g_{\alpha\beta} \in C^{m+2,\alpha}(V)$. Then we complete the proof.

5B. ***Regularity of the structure and the defining function.*** We already proved that g is in $C^{m+2,\alpha}$ in structure $\{x^\beta\}$. It is trivial to show that $\{x^\beta\}$ is a $C^{m+3,\alpha}$ structure of $\overline{M}$.

In Section 2, we obtain that u is in $C^{m,\alpha}(y)$ if the scalar curvature is constant. When we change the y-coordinates to harmonic coordinates x, we notice that x^β are $C^{m+1,\alpha}$ functions of y^β. So the defining function $\rho \in C^{m,\alpha}(x)$. Since the initial compactification g is smooth in M and the initial defining function is smooth in y-coordinates, then $\rho \in C^\infty(x)$ in M.

For any $p \in \partial M$, consider the neighborhood V of p and $D = \partial M \cap V$. By a linear transformation, we can assume that at p, $g_{\alpha\alpha} = 1$, $g_{ij} = g_{02} = g_{03} = \cdots = g_{0n} = 0 (i \neq j)$, $g_{01} = -\delta$ for some $\delta > 0$ and $1 - \delta$ is a very small positive number. When g_+ is Einstein, from (3-2) and (3-3),

$$\mathrm{Ric} - \frac{Sg}{n+1} = -(n-1)\frac{D^2\rho}{\rho} + \frac{n-1}{n+1}\frac{\Delta\rho}{\rho}g.$$

In local coordinates, acting on $(\partial/\partial x^0, \partial/\partial x^1)$, we have

$$\Delta\rho - (n+1)\cdot g_{01}^{-1}\cdot D^2\rho\left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}\right) = \frac{n+1}{n-1}\cdot g_{01}^{-1}\cdot\rho\left(\mathrm{Ric}_{01} - \frac{Sg_{01}}{n+1}\right) \tag{5-12}$$

If $1-\delta$ is small enough, we can find that the left-hand side of (5-12) is a uniformly elliptic operator on ρ locally. Since $\rho R_{01} \in C^{m,\alpha}(\overline{M})$, $\rho|_D \equiv 0$, so in V, $\rho \in C^{m+2,\alpha}(x)$.

To prove Remark 1.2, we only need to show that $\rho R_{01} \in C^{m+1,\alpha}(\overline{M})$. When $\dim M = n+1$ is even, we can define the obstruction tensor $\mathcal{O}_{ij}$, in local coordinates:

$$\mathcal{O}_{ij} = \Delta^{(n+1)/2-2}(P_{ij,k}{}^{k} - P_{ik,j}{}^{k}) + Q_n, \tag{5-13}$$

where P_{ij} is defined by $P_{ij} = \frac{1}{2}R_{ij} - \frac{S}{12}g_{ij}$ and Q_n denotes quadratic and higher terms in metric involving at most n-th derivatives. $\mathcal{O}_{ij}$ is conformally invariant of weight $2-n$ and if g_{ij} is conformal to an Einstein metric, then $\mathcal{O}_{ij} = 0$ (see more in [Graham and Hirachi 2005]).

Since the scalar curvature of $(\overline{M}, g)$ is constant, (5-13) can be written in the form:

$$\Delta^{(n+1)/2-1}R_{ij} = Q_n. \tag{5-14}$$

Now we consider the function ρR_{01}. Through a direct computation, $\Delta(\rho R_{ij}) = \rho\Delta R_{ij} + Q_3^2$. Here Q_3^2 denotes quadratic and higher terms in the metric involving at most 3-rd derivatives and in ρ involving at most 2-nd derivatives.

Then we use iterative method to obtain that $\Delta^k(\rho R_{ij}) = \rho\Delta^k R_{ij} + Q_{2k+1}^{2k}$ for $1 \le k \le (n+1)/2-1$. Letting $k = (n+1)/2-1$, we have an elliptic equation of second order with Dirichlet boundary condition:

$$\begin{cases} \Delta(\Delta^{(n+1)/2-2}(\rho R_{01})) = Q_n^{n-1} & \text{in } \overline{M}, \\ \Delta^{(n+1)/2-2}(\rho R_{01})|_{\partial M} = Q_{n-2}^{n-3}. \end{cases} \tag{5-15}$$

Since g and ρ are all in $C^{m+2,\alpha}$, $m+2 \ge n$, we have

$$\Delta^{(n+1)/2-2}(\rho R_{01}) \in C^{m+2-(n-2),\alpha}(\overline{M}).$$

Then we consider the equation

$$\begin{cases} \Delta(\Delta^{(n+1)/2-3}(\rho R_{01})) \in C^{m+2-(n-2),\alpha}(\overline{M}) \\ \Delta^{(n+1)/2-3}(\rho R_{01})|_{\partial M} = Q_{n-4}^{n-5} \in C^{m+2-(n-4),\alpha}(\partial M) \end{cases} \tag{5-16}$$

So we have $\Delta^{(n+1)/2-3}(\rho R_{01}) \in C^{m+2-(n-4),\alpha}(\overline{M})$

If we keep using the equation, we finally get $\rho R_{01} \in C^{m+1,\alpha}(\overline{M})$, which implies $\rho \in C^{m+3,\alpha}(\overline{M})$ by (5-12).

References

[Anderson 2003] M. T. Anderson, “Boundary regularity, uniqueness and non-uniqueness for AH Einstein metrics on 4-manifolds”, *Adv. Math.* **179**:2 (2003), 205–249. MR Zbl

[Anderson 2008] M. T. Anderson, “Einstein metrics with prescribed conformal infinity on 4-manifolds”, *Geom. Funct. Anal.* **18**:2 (2008), 305–366. MR Zbl

[Chruściel et al. 2005] P. T. Chruściel, E. Delay, J. M. Lee, and D. N. Skinner, “Boundary regularity of conformally compact Einstein metrics”, *J. Differential Geom.* **69**:1 (2005), 111–136. MR Zbl

[DeTurck and Kazdan 1981] D. M. DeTurck and J. L. Kazdan, “Some regularity theorems in Riemannian geometry”, *Ann. Sci. École Norm. Sup.* (4) **14**:3 (1981), 249–260. MR Zbl

[Fefferman and Graham 1985] C. Fefferman and C. R. Graham, “Conformal invariants”, pp. 95–116 in *The mathematical heritage of Élie Cartan* (Lyon, 1984), Astérisque **S131**, Soc. Math. France, Paris, 1985. MR Zbl

[Gilbarg and Trudinger 1977] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der Math. Wissenschaften **224**, Springer, 1977. MR Zbl

[Graham 2000] C. R. Graham, “Volume and area renormalizations for conformally compact Einstein metrics”, pp. 31–42 in *Proc. 19th Winter School: Geometry and Physics* (Srní, Czech Republic, 1999), edited by J. Slovák and M. Čadek, Rend. Circ. Mat. Palermo (2) Suppl. **63**, Circolo Matematico Palermo, 2000. MR Zbl

[Graham and Hirachi 2005] C. R. Graham and K. Hirachi, “The ambient obstruction tensor and Q-curvature”, pp. 59–71 in *AdS/CFT correspondence: Einstein metrics and their conformal boundaries*, edited by O. Biquard, IRMA Lect. Math. Theor. Phys. **8**, Eur. Math. Soc., Zürich, 2005. MR Zbl

[Graham and Lee 1991] C. R. Graham and J. M. Lee, “Einstein metrics with prescribed conformal infinity on the ball”, *Adv. Math.* **87**:2 (1991), 186–225. MR Zbl

[Helliwell 2008] D. W. Helliwell, “Boundary regularity for conformally compact Einstein metrics in even dimensions”, *Comm. Partial Differential Equations* **33**:5 (2008), 842–880. MR Zbl

[Jin 2017] X. S. Jin, “An extension of Weyl–Schouten theorem for Lipschitz and H^2 manifolds”, *Acta Math. Sin.* (*Eng.*) **33**:7 (2017), 926–932. MR Zbl

[Lee 1995] J. M. Lee, “The spectrum of an asymptotically hyperbolic Einstein manifold”, *Comm. Anal. Geom.* **3**:2 (1995), 253–271. MR Zbl

[Ma 1995] L. Ma, “The Yamabe problem with Dirichlet data”, *C. R. Acad. Sci. Paris Sér. I Math.* **320**:6 (1995), 709–712. MR Zbl

[Morrey 1966] C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, Grundlehren der Math. Wissenschaften **130**, Springer, 1966. MR Zbl

[Petersen 1998] P. Petersen, *Riemannian geometry*, Graduate Texts in Math. **171**, Springer, 1998. MR Zbl

[Shi and Tian 2005] Y. Shi and G. Tian, “Rigidity of asymptotically hyperbolic manifolds”, *Comm. Math. Phys.* **259**:3 (2005), 545–559. MR Zbl

Received January 16, 2018. Revised September 8, 2018.

XIAOSHANG JIN
BEIJING INTERNATIONAL CENTER FOR MATHEMATICAL RESEARCH
PEKING UNIVERSITY
HAIDIAN DISTRICT
BEIJING
CHINA
xsjin@bicmr.pku.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.489

GEOMETRIC TRANSITIONS AND SYZ MIRROR SYMMETRY

ATSUSHI KANAZAWA AND SIU-CHEONG LAU

We prove that generalized conifolds and orbifolded conifolds are mirror symmetric under the SYZ program with quantum corrections. Our work mathematically confirms the gauge-theoretic assertion of Aganagic–Karch–Lüst–Miemiec, and also provides a supportive evidence to Morrison's conjecture that geometric transitions are reversed under mirror symmetry.

1. Introduction

Morrison [1999] proposed that geometric transitions are reversed under mirror symmetry. A geometric transition is a birational contraction followed by a complex smoothing, or in the reverse way, applied to a Kähler manifold (see the nice review [Rossi 2006]) . We will denote a geometric transition by $\widehat{X} \dashrightarrow X \rightsquigarrow \tilde{X}$, where $\widehat{X} \dashrightarrow X$ is a birational contraction and $X \rightsquigarrow \tilde{X}$ is a smoothing. The conjecture can be formulated as follows.

Conjecture 1.1 [Morrison 1999]. *Let $\hat{X}$ and $\tilde{X}$ be Calabi–Yau manifolds*, *and suppose they are related by a geometric transition $\hat{X} \dashrightarrow X \rightsquigarrow \tilde{X}$. Suppose Y_1 and Y_2 are the mirrors of $\hat{X}$ and $\tilde{X}$ respectively. Then there exists a geometric transition $Y_2 \dashrightarrow Y \rightsquigarrow Y_1$ relating Y_1 and Y_2.*

The present paper investigates mirror symmetry for geometric transitions of two specific types of local singularities, namely generalized conifolds and orbifolded conifolds. Let us first recall mirror symmetry for a conifold. A conifold is an isolated singularity defined by $\{xy - zw = 0\} \subset \mathbb{C}^4$. It is an important singularity appearing in algebraic geometry and also plays a special role in superstring theory. A folklore mirror symmetry for the conifold [Morrison 1999; Szendrői 2004] that the deformed conifold is mirror symmetric to the resolved conifold can be refined in the framework of SYZ mirror symmetry as follows.

Theorem 1.2 (conifold case of Theorem 3.1). *Let $\hat{X} := \mathcal{O}_{\mathbb{P}^1}(-1)^{\oplus 2} \setminus D$ be the resolved conifold with a smooth anticanonical divisor D removed*, *and*

$$\tilde{X} := \{(x, y, z, w) \in \mathbb{C}^4 \mid xy - zw = 1\} \setminus (\{z = 1\} \cup \{w = 1\})$$

MSC2010: 14J33, 53D37.
Keywords: SYZ, mirror symmetry, conifold transition, open Gromov-Witten invariants.

the deformed conifold with the anticanonical divisor $\{z = 1\} \cup \{w = 1\}$ *removed. Then* $\hat{X}$ *and* $\tilde{X}$ *are SYZ mirror to each other.*

Although removing the divisors certainly does not affect the local geometry of the singularity, it is important when we discuss, for example, wrapped Fukaya categories and homological mirror symmetry.[1]

We now focus on two natural generalizations of the conifold: generalized conifolds and orbifolded conifolds. For integers $k, l \geq 1$, a generalized conifold is given by

$$G^{\sharp}_{k,l} := \{(x, y, z, w) \in \mathbb{C}^4 \mid xy - (1+z)^k (1+w)^l = 0\}$$

and an orbifolded conifold is given by

$$O^{\sharp}_{k,l} := \{(u_1, v_1, u_2, v_2, z) \in \mathbb{C}^5 \mid u_1 v_1 - (1+z)^k = u_2 v_2 - (1+z)^l = 0\}.$$

(We have made a change of coordinates, namely $z \mapsto 1 + z$ and $w \mapsto 1 + w$, for later convenience.) They reduce to the conifold when $k = l = 1$. The punctured generalized conifold is defined as $G_{k,l} := G^{\sharp}_{k,l} \setminus D_G$, where $D_G := \{z = 0\} \cup \{w = 0\}$ is a normal-crossing anticanonical divisor of $G^{\sharp}_{k,l}$, and the punctured orbifolded conifold as $O_{k,l} := O^{\sharp}_{k,l} \setminus D_O$, where $D_O := \{z = 0\}$ is a smooth anticanonical divisor of $O^{\sharp}_{k,l}$. As is the case of the conifold, their symplectic structures and complex structures are governed by the crepant resolutions and deformations respectively. The main theorem of the present paper is the following.

Theorem 3.1. *The punctured generalized conifold* $G_{k,l}$ *is mirror symmetric to the punctured orbifolded conifold* $O_{k,l}$ *in the sense that the deformed punctured generalized conifold* $\widetilde{G_{k,l}}$ *is SYZ mirror symmetric to the resolved punctured orbifolded conifold* $\widehat{O_{k,l}}$*, and the resolved punctured generalized conifold* $\widehat{G_{k,l}}$ *is SYZ mirror symmetric to the deformed punctured orbifolded conifold* $\widetilde{O_{k,l}}$*.*

$$\begin{tikzcd}
\widetilde{G_{k,l}} \arrow[d, leftrightarrow, "SYZ"'] & G_{k,l} \arrow[l, squiggly] \arrow[d, leftrightarrow, "MS"] & \widehat{G_{k,l}} \arrow[l, dotted] \arrow[d, leftrightarrow, "SYZ"] \\
\widehat{O_{k,l}} \arrow[r, dotted] & O_{k,l} \arrow[r, squiggly] & \widetilde{O_{k,l}}.
\end{tikzcd}$$

According to Theorem 3.1 the mirror duality of the conifold is purely caused by the fact that the conifold can be seen as either a generalized or an orbifolded conifold. The mirror duality of $G^{\sharp}_{k,l}$ and $O^{\sharp}_{k,l}$ has previously been studied by the physicists Aganagic, Karch, Lust and Miemiec in [Aganagic et al. 2000], where they use gauge theory and brane configurations. Our work mathematically confirms their gauge-theoretic assertion, and also provides a supportive evidence to Morrison's conjecture that geometric transitions are reversed under mirror symmetry.

[1] We are grateful to Murad Alim for informing us about the importance of this issue.

In the present paper, we use the framework introduced by the second author with Chan and Leung [Chan et al. 2012] for defining SYZ mirror pairs. Namely, generating functions of open Gromov–Witten invariants of fibers of a Lagrangian fibration were used to construct the complex coordinates of the mirror. The essential ingredient is wall-crossing of the generating functions, which was first studied by Auroux [2007]. We can also bypass symplectic geometry and employs the Gross–Siebert program [2011] which uses tropical geometry instead for defining mirror pairs. This tropical approach was taken by Castaño-Bernard and Matessi [2014] in the study of conifold transitions for compact Calabi–Yau varieties. Although our work has some overlap with theirs, the methods and interests are quite different. They deal with simultaneous multiple conifold transitions while we discuss orbifold/generalized conifolds transitions. In this paper tropical geometry is not directly needed since symplectic geometry can be handled directly.

One crucial feature of the present work is the involutive property of SYZ mirror symmetry. Namely, taking SYZ mirror twice gets back to itself, which we believe is an important point but often overlooked in literatures. We exhibit this feature by carrying out the SYZ construction for all the four directions in Theorem 3.1, namely from $\widehat{G_{k,l}}$ to $\widetilde{O_{k,l}}$, and from $\widetilde{O_{k,l}}$ back to $\widehat{G_{k,l}}$; from $\widetilde{G_{k,l}}$ to $\widehat{O_{k,l}}$, and from $\widehat{O_{k,l}}$ back to $\widetilde{G_{k,l}}$. The SYZ construction from $\widehat{G_{k,l}}$ to $\widetilde{O_{k,l}}$ is a bit tricky and we will discuss it in details. We will employ the various techniques developed in [Auroux 2007; Chan et al. 2012; Abouzaid et al. 2016; Lau 2014].

Another interesting feature is the dependence of the choice of a Lagrangian fibration, namely a Lagrangian fibration has to be "compatible" with the choice of an anticanonical divisor, in the sense that the anticanonical divisor is the preimage of the boundary of the base of the Lagrangian fibration, in order to obtain the corresponding mirror. For example, $\widehat{G_{k,l}}$ admits two different Lagrangian fibrations: the Gross fibration and a "doubled" Gross fibration. The former is not compatible with the anticanonical divisor D_G, and does not produce the orbifold conifold as its SYZ mirror. Choosing appropriate Lagrangian fibrations is a key step in our work.

Lastly, Theorem 3.1 not only unveils a connection between geometric transitions and SYZ mirror symmetry, but also yields many interesting problems and conjectures that naturally extend what is known for the conifolds. In fact, based on the local models studied in this paper, the second author recently confirmed Morrison's conjecture for a class of geometric transitions of the Schoen's Calabi–Yau threefolds [Lau 2018].

Structure of paper. Section 2 introduces generalized conifolds and orbifolded conifolds and basic properties thereof. Section 3 begins with a review on the Lagrangian torus fibrations and the SYZ program. Then we prove the main theorem

(Theorem 3.1) by carrying out the SYZ constructions. Section 4 discusses global geometric transitions and provides a few examples.

2. Generalized and orbifolded conifolds

In this section, we introduce two natural generalizations of the conifold, namely generalized conifolds and orbifolded conifolds. These two singularities possess interesting geometries and were studied by physicists in the context of gauge theory, for instance in [Katz et al. 1997; Aganagic et al. 2000; Miemiec 2000].

Generalized conifolds $G^\sharp_{k,l}$. A toric Calabi–Yau manifold is a semi-projective toric manifold with trivial canonical bundle. In dimension three, they can be described by a lattice polytope $\Delta \subset \mathbb{R}^2$ whose vertices lie in the lattice $\mathbb{Z}^2 \subset \mathbb{R}^2$. Its fan is produced by taking the cone over $\Delta \times \{1\} \subset \mathbb{R}^3$. A crepant resolution of a toric Calabi–Yau threefold corresponds to a subdivision of Δ into standard triangles,[2] which gives a refinement of the fan. For instance, the total space of the canonical bundle K_S of a smooth toric surface S is a toric Calabi–Yau threefold. In this situation, the surface S is the toric variety $\mathbb{P}_\Delta$ whose fan polytope is Δ.

The condition that a toric Calabi–Yau threefold contains no compact 4-cycles is equivalent to the condition that the polytope Δ contains no interior lattice points. The lattice polygons without interior lattice point are classified, up to the action of $GL(2, \mathbb{Z})$, into two types:

(1) triangle with vertices $(0, 0), (2, 0), (0, 2)$,

(2) trapezoid $\Delta_{k,l}$ with vertices $(0, 0), (0, 1), (k, 0), (l, 1)$ for $k \geq l \geq 0$ with $(k, l) \neq (0, 0)$ (Figure 1, left).

The former is the quotient of $\mathbb{C}^3$ by the subgroup $(\mathbb{Z}_2)^2 \subset SL(3, \mathbb{C})$ generated by the two elements $\mathrm{diag}(-1, -1, 1)$ and $\mathrm{diag}(1, -1, -1)$. In this paper, we are interested in the latter, which corresponds to the generalized conifold $G_{k,l}$ for $k \geq l \geq 1$. We do not consider the case $l = 0$, where the toric singularity essentially comes from the A_k-singularity in 2 dimensions.[3] The dual cone of the cone over

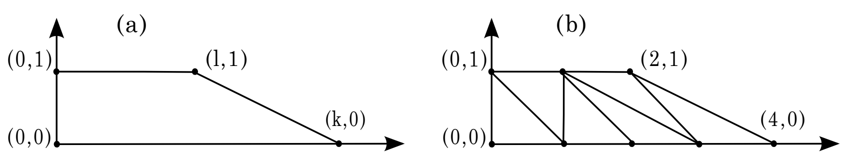


Figure 1. Left: trapezoid $\Delta_{k,l}$. Right: crepant resolution $\widehat{G^\sharp_{4,2}}$.

[2] A standard triangle in $\mathbb{R}^2$ is isomorphic to the convex hull of $(0, 0), (1, 0), (0, 1)$ under the $\mathbb{Z}^2 \rtimes GL(2, \mathbb{Z})$-transformation.

[3] Mirror symmetry of this class of singularities is discussed in [Szendrői 2004, Section 5].

the trapezoid $\Delta_{k,l}$ is spanned by the vectors

$$\nu_1 := (1,0,0),\ \nu_2 := (0,-1,1),\ \nu_3 := (-1,l-k,k),\ \nu_4 := (0,1,0) \tag{1}$$

with relation $\nu_1 - k\nu_2 + \nu_3 - l\nu_4 = 0$. In equation the generalized conifold $G^\sharp_{k,l}$ is given by

$$G^\sharp_{k,l} := \{xy - z^k w^l = 0\} \subset \mathbb{C}^4.$$

The coordinates x, y, z, w correspond to the dual lattice points $\nu_1, \nu_3, \nu_2, \nu_4$ respectively. For $(k,l) \neq (1,1)$, the generalized conifold $G^\sharp_{k,l}$ is a quotient of the conifold (which is given by $(k,l) = (1,1)$) and has a 1-dimensional singular locus. A punctured generalized conifold is $G_{k,l} := G^\sharp_{k,l} \setminus D_G$, where $D_G = \{z = 1\} \cup \{w = 1\}$ is an anticanonical divisor of $G^\sharp_{k,l}$. A crepant resolution $\widehat{G_{k,l}}$ of $G_{k,l}$ is called a resolved generalized conifold. We observe that

$$\widehat{G_{k,l}} = \widehat{G^\sharp_{k,l}} \setminus D_{\hat{G}},$$

where $D_{\hat{G}}$ is an anticanonical divisor of $\widehat{G^\sharp_{k,l}}$, and it uniquely corresponds to a maximal triangulation of the trapezoid $\Delta_{k,l}$ (Figure 1, right). The resolved generalized conifold $\widehat{G_{k,l}}$ is endowed with a natural symplectic structure as an open subset of a smooth toric variety $\widehat{G^\sharp_{k,l}}$.

Proposition 2.1. *There are $\binom{k+l}{k}$ distinct crepant resolutions of $G_{k,l}$ (or equivalently $G^\sharp_{k,l}$).*

Proof. There is a bijection between the crepant resolutions of $G_{k,l}$ and the maximal triangulations of $\Delta_{k,l}$. The assertion easily follows by induction with the relation $\binom{k+l+1}{k+1} = \binom{k+l}{k} + \binom{k+l}{k+1}$. □

We may also smooth out the punctured generalized conifold $G_{k,l}$ by deforming the equation. The deformed generalized conifold $\widetilde{G_{k,l}}$ is defined as

$$\widetilde{G_{k,l}} := \left\{(x,y,z,w) \in \mathbb{C}^2 \times (\mathbb{C} \setminus \{1\})^2 \;\middle|\; xy - \sum_{i=0}^{k}\sum_{j=0}^{l} a_{i,j} z^i w^j = 0\right\}$$

for generic $a_{i,j} \in \mathbb{C}$. The symplectic structure of $\widetilde{G_{k,l}}$ is given by the restriction of the standard symplectic structure on $\mathbb{C}^2 \times (\mathbb{C} \setminus \{1\})^2$. We observe that the complex deformation space has dimension $(k+1)(l+1) - 3$ because three of the parameters can be eliminated by rescaling z, w and rescaling the whole equation. On the other hand, the Kähler deformation space has dimension $(k+1) + (l+1) - 3$, the number of linearly dependent lattice vectors in the polytope. It is the number of the exceptional $\mathbb{P}^1$s' and a Kähler form is parametrized by the area of these.

Orbifolded conifolds $O^\sharp_{k,l}$. Let $X^\sharp$ be the conifold $\{xy - zw = 0\} \subset \mathbb{C}^4$. For $k \geq l \geq 1$, the orbifolded conifold $O^\sharp_{k,l}$ is the quotient of the conifold $X^\sharp$ by the abelian group $\mathbb{Z}_k \times \mathbb{Z}_l$, where $\mathbb{Z}_k$ and $\mathbb{Z}_l$ respectively act by

$$(x, y, z, w) \mapsto (\zeta_k x, \zeta_k^{-1} y, z, w), \text{ and } (x, y, z, w) \mapsto (x, y, \zeta_l z, \zeta_l^{-1} w)$$

where ζ_k, ζ_l are primitive k-th and l-th roots of unity respectively (assume k and l relatively prime for simplicity [Aganagic et al. 2000]). Alternatively the orbifolded conifold $O^\sharp_{k,l}$ is realized as a hypersurface in $\mathbb{C}^5$:

$$O^\sharp_{k,l} = \{u_1 v_1 = z^k,\ u_2 v_2 = z^l\} \subset \mathbb{C}^5.$$

The orbifolded conifold $O^\sharp_{k,l}$ is an example of a toric Calabi–Yau threefold and the corresponding polytope is given by the rectangle $\square_{k,l}$ with the vertices $(0, 0)$, $(k, 0)$, $(0, l)$, (k, l) (Figure 2, left).

The dual cone of the cone over the rectangle $\square_{k,l}$ is spanned by the following vectors

$$v_1 := (1, 0, 0),\ v_2 := (0, -1, l),\ v_3 := (-1, 0, k),\ v_4 := (0, 1, 0)$$

with relation $lv_1 - kv_2 + lv_3 - kv_4 = 0$.

A punctured orbifolded conifold is $O_{k,l} := O^\sharp_{k,l} \setminus D_O$, where $D_O = \{z = 1\}$ is a smooth anticanonical divisor of $O^\sharp_{k,l}$. Then a resolved orbifolded conifold $\widehat{O_{k,l}}$ is defined to be a crepant resolution of $O_{k,l}$. As before,

$$\widehat{O_{k,l}} = \widehat{O^\sharp_{k,l}} \setminus D_{\hat{O}},$$

where $D_{\hat{O}}$ is a smooth anticanonical divisor of the toric crepant resolution $\widehat{O^\sharp_{k,l}}$, and it corresponds to a maximal triangulation of the trapezoid $\square_{k,l}$ (Figure 2, right). It has a canonical symplectic structure as an open subset of a smooth toric variety $\widehat{O^\sharp_{k,l}}$. In contrast to Proposition 2.1, it is a famous open problem to find the number of the crepant resolutions of the orbifolded conifold $O^\sharp_{k,l}$ [Kaibel and Ziegler 2003]. The punctured orbifolded conifold $O_{k,l}$ can also be smoothed out by deforming the

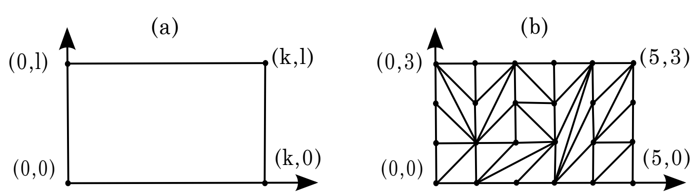


Figure 2. Left: rectangle $\square_{k,l}$. Right: crepant resolution $\widehat{O^\sharp_{5,3}}$.

defining equations. Thus the deformed orbifolded conifold $\widetilde{O_{k,l}}$ is given by

$$\widetilde{O_{k,l}} := \Big\{(u_1, v_1, u_2, v_2, z) \in \mathbb{C}^4 \times (\mathbb{C} \setminus \{1\}) \;\Big|\; u_1 v_1 = \sum_{i=0}^{k} a_i z^i,\ u_2 v_2 = \sum_{j=0}^{l} b_j z^j \Big\}$$

for generic coefficients $a_i, b_j \in \mathbb{C}$. The symplectic structure of $\widetilde{O_{k,l}}$ is the restriction of the standard symplectic structure on $\mathbb{C}^4 \times (\mathbb{C} \setminus \{1\})$. The complex deformation space of $\widetilde{O_{k,l}}$ has dimension $(k+1)+(l+1)-3$, while the Kähler deformation space has dimension $(k+1)(l+1)-3$. Therefore the naive dimension counting is compatible with our claim that these two classes of singularities are mirror symmetric. We will formulate this mirror duality in a rigorous manner by using SYZ mirror symmetry in the next section.

3. SYZ mirror construction

The Strominger–Yau–Zaslow (SYZ) conjecture [Strominger et al. 1996] provides a foundational geometric understanding of mirror symmetry. It asserts that, for a mirror pair of Calabi–Yau manifolds X and $X^\vee$, there exist Lagrangian torus fibrations $\pi : X \to B$ and $\pi^\vee : X^\vee \to B$ which are fiberwise-dual to each other. In particular, it suggests an intrinsic construction of the mirror $X^\vee$ by fiberwise dualizing a Lagrangian torus fibration on X. This is motivated by *T-duality* studied by string theorists.

The SYZ program has been carried out successfully in the semi-flat case [Leung 2005] in which the discriminant locus of the fibrations is empty. When singular fibers are present, quantum corrections by open Gromov–Witten invariants of the fibers are necessary, and they exhibit wall-crossing phenomenon. Wall-crossing of open Gromov–Witten invariants was first studied by Auroux [2007]. Later on Chan et al. [2012] gave an SYZ construction of mirrors with quantum corrections, which will be used in this paper. In algebro-geometric context, the Gross–Siebert program [2011] gives a reformulation of the SYZ program using tropical geometry, which provides powerful techniques to compute wall-crossing and scattering order-by-order. In this paper we will use the symplectic rather than the tropical approach.

We will shortly give a quick review of the setting of [Chan et al. 2012] for SYZ with quantum corrections. We say that X is *SYZ mirror symmetric* to Y if Y is produced from X as a SYZ mirror manifold by this SYZ mirror construction. The later parts of this section prove the following main theorem.

Theorem 3.1. *The punctured generalized conifold $G_{k,l}$ is mirror symmetric to the punctured orbifolded conifold $O_{k,l}$ in the sense that the deformed generalized conifold $\widetilde{G_{k,l}}$ is SYZ mirror symmetric to the resolved orbifolded conifold $\widehat{O_{k,l}}$, and the resolved generalized conifold $\widehat{G_{k,l}}$ is SYZ mirror symmetric to the deformed* *orbifolded conifold* $\widetilde{O_{k,l}}$:

$$\begin{tikzcd} \widetilde{G_{k,l}} \arrow[d, leftrightarrow, "SYZ"'] & G_{k,l} \arrow[l, leftsquigarrow] \arrow[d, leftrightarrow, "MS"] & \widehat{G_{k,l}} \arrow[l, dotted] \arrow[d, leftrightarrow, "SYZ"] \\ \widehat{O_{k,l}} \arrow[r, dotted] & O_{k,l} \arrow[r, rightsquigarrow] & \widetilde{O_{k,l}}. \end{tikzcd}$$

This mathematically confirms the gauge-theoretic assertion of the string theorists Aganagic, Karch, Lüst and Miemiec [Aganagic et al. 2000] and also provides a supportive evidence to Morrison's conjecture [1999] from the view point of SYZ mirror symmetry.

SYZ construction with quantum corrections. In this subsection we review the SYZ construction with quantum corrections given in [Chan et al. 2012]. We add a clarification that we only use transversal disc classes (Definition 3.5) in the definition of the mirror space.

Let $\pi : X \to B$ be a proper Lagrangian torus fibration of a Kähler manifold (X, ω) such that the base B is a compact manifold with corners, and the preimage of each codimension-one facet of B is a smooth irreducible divisor denoted as D_i for $1 \le i \le m$. We assume that the regular Lagrangian fibers of π are special with respect to a nowhere-vanishing meromorphic volume form Ω on X whose pole divisor is the boundary divisor $D := \sum_{i=1}^{m} D_i$ (and hence D is an anticanonical divisor). We denote by $B_0 \subset B$ the complement of the discriminant locus of π, and we assume that B_0 is connected.[4] We always denote by F_b a fiber of π at $b \in B_0$.

Lemma 3.2 (Maslov index of disc classes [Auroux 2007, Lemma 3.1]). *For a disc class $\beta \in \pi_2(X, F_b)$ where $b \in B_0$, the Maslov index of β is $\mu(\beta) = 2D \cdot \beta$.*

Definition 3.3 [Chan et al. 2012]. The wall H of a Lagrangian fibrartion $\pi : X \to B$ is the set of point $b \in B_0$ such that $F_b := \pi^{-1}(b)$ bounds non-constant holomorphic disks with Maslov index 0.

The complement of $H \subset B_0$ consists of several connected components, which we call chambers. Over different chambers the Lagrangian fibers behave differently in a Floer-theoretic sense. Away from the wall H, the one-point open Gromov–Witten invariants are well-defined using the machinery of Fukaya–Oh–Ohta–Ono:

Definition 3.4 (open Gromov–Witten invariants [Fukaya et al. 2009]). For b in $B_0 \setminus H$ and β in $\pi_2(X, F_b)$, let $\mathfrak{M}_1(\beta)$ be the moduli space of stable discs with one boundary marked point representing β, and $[\mathfrak{M}_1(\beta)]^{\mathrm{virt}}$ be the virtual fundamental class of $\mathfrak{M}_1(\beta)$. The open Gromov–Witten invariant associated to β

[4] When the discriminant locus has codimension-two, B_0 is automatically connected. Although the Lagrangian fibrations of $\widetilde{G_{k,l}}$ we study have codimension-one discriminant loci, B_0 is still connected.

is $n_\beta := \int_{[\mathfrak{M}_1(\beta)]^{\mathrm{virt}}} \mathrm{ev}^*[\mathrm{pt}]$, where $\mathrm{ev} : \mathfrak{M}_1(\beta) \to F_b$ is the evaluation map at the boundary marked point and [pt] is the Poincaré dual of the point class of F_b.

We will restrict to disc classes which are transversal to the boundary divisor D when we construct the mirror space (while for the mirror superpotential we need to consider all disc classes).

Definition 3.5 (transversal disc class). A disc class $\beta \in \pi_2(X, F_b)$ for $b \in B_0$ is said to be transversal to the boundary divisor D, which is denoted as $\beta \pitchfork D$, if all stable discs in $\mathfrak{M}_1(\beta)$ intersect transversely with the boundary divisor D.

Due to dimension reason, the open Gromov–Witten invariant n_β is non-zero only when the Maslov index $\mu(\beta) = 2$. When β is transversal to D or when X is semi-Fano, namely $c_1(\alpha) = D \cdot \alpha \geq 0$ for all holomorphic sphere classes α, the number n_β is invariant under small deformation of complex structure and under Lagrangian isotopy in which all Lagrangian submanifolds in the isotopy do not intersect D nor bound non-constant holomorphic disc with Maslov index $\mu(\beta) < 2$.

A procedure realizing the SYZ program based on symplectic geometry was proposed in [Chan et al. 2012] as follows:

(i) Construct the semi-flat mirror $X_0^\vee$ of $X_0 := \pi^{-1}(B_0)$ as the space of pairs (b, ∇) where $b \in B_0$ and ∇ is a flat $\mathrm{U}(1)$-connection on the trivial complex line bundle over F_b up to gauge. There is a natural map $\pi^\vee : X_0^\vee \to B_0$ given by forgetting the second coordinate. The semi-flat mirror $X_0^\vee$ has a canonical complex structure [Leung 2005] and the functions $\mathrm{e}^{-\int_\beta \omega} \mathrm{Hol}_\nabla(\partial\beta)$ on $X_0^\vee$ for disc classes $\beta \in \pi_2(X, F_b)$ are called semi-flat complex coordinates. Here $\mathrm{Hol}_\nabla(\partial\beta)$ denotes the holonomy of the flat $\mathrm{U}(1)$-connection ∇ along the path $\partial\beta \in \pi_1(F_b)$.

(ii) Define the generating functions of open Gromov–Witten invariants for $1 \leq i \leq m$

$$Z_i(b, \nabla) := \sum_{\substack{\beta \in \pi_2(X, F_b) \\ \beta \cdot D_i = 1, \beta \pitchfork D}} n_\beta \mathrm{e}^{-\int_\beta \omega} \mathrm{Hol}_\nabla(\partial\beta), \tag{2}$$

for $(b, \nabla) \in (\pi^\vee)^{-1}(B_0 \setminus H)$, which serve as quantum corrected complex coordinates. The function Z_i can be written in terms of the semi-flat complex coordinates, and hence they generate a subring $\mathbb{C}[Z_1, \ldots, Z_m]$ in the function ring[5] over $(\pi^\vee)^{-1}(B_0 \setminus H)$.

[5] In general we need to use the Novikov ring instead of $\mathbb{C}$ since Z_i could be a formal Laurent series. In the cases that we study later, Z_i are Laurent polynomials whose coefficients are convergent, and hence the Novikov ring is not necessary.

(iii) Define the SYZ mirror of X with respect to the Lagrangian torus fibration π to be the pair $(X^\vee, W)$ where $X^\vee := \operatorname{Spec}(\mathbb{C}[Z_1, \dots, Z_m])$ and

$$W = \sum_{\beta \in \pi_2(X, F_b)} n_\beta e^{-\int_\beta \omega} \operatorname{Hol}_\nabla(\partial\beta).$$

Moreover, $X^\vee$ is defined as the SYZ mirror of a non-compact Calabi–Yau manifold Y if it is obtained from the above construction for a compactification of Y. It is expected that different compactifications would result in the same SYZ mirror. In this paper we fix one compactification as an initial data for the SYZ construction.

In the following sections we will apply the above recipe to the generalized conifolds and orbifolded conifolds. We will carry out in detail the SYZ construction from $\widehat{G_{k,l}}$ to $\widetilde{O_{k,l}}$ which is the most interesting case (see immediately below), in which we construct a doubled version of the Gross fibration [Goldstein 2001; Gross 2001] and compute the open Gromov–Witten invariants. The other cases, namely the SYZ constructions from $\widetilde{O_{k,l}}$ to $\widehat{G_{k,l}}$, from $\widetilde{O_{k,l}}$ to $\widetilde{G_{k,l}}$, and from $\widetilde{G_{k,l}}$ to $\widetilde{O_{k,l}}$, are essentially obtained by applying the techniques developed in [Lau 2014; Chan et al. 2012; Abouzaid et al. 2016], and so we will be brief. In fact $G^\sharp_{k,l}$ and $O^\sharp_{k,l}$ are useful testing grounds for the SYZ program and we shall illustrate how these various important ideas fit together by examining them.

SYZ from $\widehat{G_{k,l}}$ to $\widetilde{O_{k,l}}$. We first construct the SYZ mirror of the resolved generalized conifold $\widehat{G_{k,l}}$. While the resolved generalized conifold $\widehat{G^\sharp_{k,l}}$ is a toric Calabi–Yau threefold, we will *not* use the Gross fibration because it is not compatible with the chosen anticanonical divisor D_G and hence do not produce the resolved orbifolded conifold $\widetilde{O_{k,l}}$ as the mirror. We will instead use a *doubled* version of the Gross fibration explained below.

The fan of $\widehat{G^\sharp_{k,l}}$ is given by the cone over a triangulation depicted in Figure 1, right. We label the divisors corresponding to the rays generated by $(i, 0, 1)$ to be D_i for $0 \le i \le k$, and those corresponding to the rays generated by $(j, 1, 1)$ to be D_{k+1+j} for $0 \le j \le l$. Each divisor D_i corresponds to a basic disc class $\beta_i \in \pi_2(X, L)$ where L denotes a moment-map fiber [Chan et al. 2012].

Let us first compactify the resolved generalized conifold $\widehat{G_{k,l}}$ as follows. We add the rays generated by $(0, 0, -1)$, $(0, -1, -1)$, $(1, 0, 0)$ and $(-1, 0, 0)$, and the corresponding cones, to the fan of $\widehat{G^\sharp_{k,l}}$. Let us denote the resulting toric variety by $\widehat{G_{k,l}}^*$ (and we fix a toric Kähler form on it). Let $\mathcal{D}_{z=\infty}$, $\mathcal{D}_{w=\infty}$, $\mathcal{D}_{\xi=0}$ and $\mathcal{D}_{\xi=\infty}$ be the corresponding additional toric prime divisors and $\beta_{z=\infty}$, $\beta_{w=\infty}$, $\beta_{\xi=0}$ and $\beta_{\xi=0}$ be the additional basic disc classes respectively.

Note that $\widehat{G_{k,l}}^*$ is in general not semi-Fano since there could be holomorphic spheres with the Chern class $c_1 < 0$ supported in the newly added divisors (or in other words the fan polytope of $\widehat{G_{k,l}}^*$ may contain a interior lattice point). However since we only need to consider transversal disc classes (Definition 3.5) in the definition of $X^\vee$, these holomorphic spheres do not enter into our constructions.

We now construct a special Lagrangian fibration and apply the SYZ construction on $\widehat{G_{k,l}}^*$. Consider the Hamiltonian T^1-action on $\widehat{G_{k,l}}^*$ corresponding to the vector $(1,0,0)$ in the vector space which supports the fan. Denote by $\pi_{T^1}:\widehat{G_{k,l}}^*\to\mathbb{R}$ the moment map associated to this Hamiltonian T^1-action, whose image is a closed interval I. Let θ be the angular coordinate corresponding to the Hamiltonian T^1-action. Recall that x, y, z, w are toric functions corresponding to the lattice points $\nu_1, \nu_3, \nu_2, \nu_4$ (in the vector space which supports the moment polytope) defined by (1). Note that $z=0$ on the toric divisors $D_0,\dots,D_k$, while $w=0$ on the toric divisors $D_{k+1},\dots,D_{k+l+1}$. Moreover the pole divisors of z and w are $\mathcal{D}_{z=\infty}$ and $\mathcal{D}_{w=\infty}$ respectively.

The toric Kähler form on $\widehat{G_{k,l}}^*$ can be written as

$$\omega := \mathrm{d}\pi_{T^1}\wedge \mathrm{d}\theta + \frac{\sqrt{-1}}{c_1(1+|z|^2)^2}\mathrm{d}z\wedge \mathrm{d}\bar{z} + \frac{\sqrt{-1}}{c_2(1+|w|^2)^2}\mathrm{d}w\wedge \mathrm{d}\bar{w}$$

for some $c_1, c_2\in\mathbb{R}_{>0}$. We define a T^3-fibration $\pi:\widehat{G_{k,l}}^*\to B:=[-\infty,\infty]^2\times I$ by

$$\pi(x,y,z,w) = (b_1,b_2,b_3) = (\log|z-1|, \log|w-1|, \pi_{T^1}(x,y,z,w)).$$

We also define a nowhere-vanishing meromorphic volume form by

$$\Omega := \mathrm{d}\log x\wedge \mathrm{d}\log(z-1)\wedge \mathrm{d}\log(w-1).$$

The pole divisor D of Ω is given by the union

$$D = \mathcal{D}_{z=1}\cup\mathcal{D}_{w=1}\cup\mathcal{D}_{z=\infty}\cup\mathcal{D}_{w=\infty}\cup\mathcal{D}_{\xi=0}\cup\mathcal{D}_{\xi=\infty}$$

whose image under π is the boundary of B ($\mathcal{D}_{z=1}$ and $\mathcal{D}_{w=1}$ denotes the divisors $\{z=1\}$ and $\{w=1\}$ respectively). Using the method of symplectic reductions [Goldstein 2001], we obtain the following.

Proposition 3.6. *The T^3-fibration π defined above is a special Lagrangian fibration with respect to ω and Ω.*

Proof. Consider the symplectic quotient of the Hamiltonian T^1-action: $\widetilde{M}:=\pi_{T^1}^{-1}(\{b_3\})/T^1$ for certain $b_3\in\mathbb{R}$. Since the toric coordinates z and w are invariant under the T^1-action, they descend to the quotient $\widetilde{M}$. This gives an identification of $\widetilde{M}$ with $\mathbb{P}^1\times\mathbb{P}^1$. The induced symplectic form on the quotient $\widetilde{M}$ is given by

$$\begin{aligned}\widetilde{\omega} &= \frac{\sqrt{-1}}{c_1(1+|z|^2)^2}\mathrm{d}z\wedge \mathrm{d}\bar{z} + \frac{\sqrt{-1}}{c_2(1+|w|^2)^2}\mathrm{d}w\wedge \mathrm{d}\bar{w}\\ &= \frac{\sqrt{-1}|z-1|^2}{c_1(1+|z|^2)^2}\mathrm{d}\log(z-1)\wedge \mathrm{d}\overline{\log(z-1)} + \frac{\sqrt{-1}|w-1|^2}{c_2(1+|w|^2)^2}\mathrm{d}\log(w-1)\wedge \mathrm{d}\overline{\log(w-1)},\end{aligned}$$

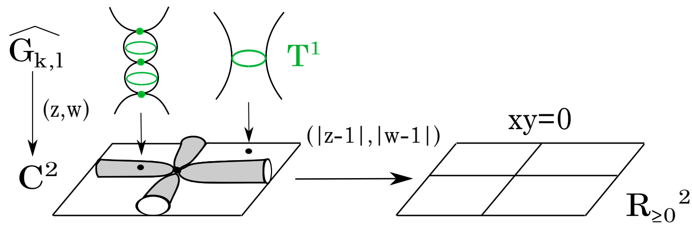


Figure 3. Conic fibration after resolution.

where $\sqrt{-1} = \sqrt{-1}$. The induced holomorphic volume form $\widetilde{\Omega}$, which is the contraction of Ω by the vector field induced from the T^1-action, equals to

$$\widetilde{\Omega} = \mathrm{d}\log(z-1) \wedge \mathrm{d}\log(w-1).$$

It is clear that $\widetilde{\omega}$ and $\mathrm{Re}(\widetilde{\Omega})$ restricted on each fiber of the fibration $(|z-1|, |w-1|)$ are both zero. Hence the fibers of the map $(|z-1|, |w-1|)$ are special Lagrangian in $\widetilde{M}$. By [Goldstein 2001, Lemma 2], we therefore conclude that the fibers of π are special Lagrangian. □

We may think of this fibration as the combination of a conic bundle (Figure 3) and the moment map π_{T^1} associated to the lift of (x, y)-coordinates (Figure 4, left). The latter measures the volumes of the exceptional curves $\mathbb{P}^1$ of the crepant resolution $\widehat{G_{k,l}} \to G_{k,l}$.

Proposition 3.7. *The discriminant locus of the fibration π is the union of the boundary ∂B together with the lines $\{b_1 = 0, b_3 = s_i\}_{i=1}^k \cup \{b_2 = 0, b_3 = t_j\}_{j=1}^l \subset B$ for $s_i, t_j \in \mathbb{R}$ with $\mathrm{Crit}(\pi_{T^1}) = \{s_1, \ldots, s_k, t_1, \ldots, t_l\}$.*

Proof. The first and second coordinates of π are $b_1 = \log|z-1|$ and $b_2 = \log|w-1|$ respectively, which degenerates over the boundaries $b_1 = \log|z-1| = \pm\infty$ or $b_2 = \log|w-1| = \pm\infty$. The third coordinate π_{T^1} degenerates at those codimension-two toric strata whose corresponding 2-dimensional cones in the fan contain the vector $(1, 0, 0)$. These cones are either $[i-1, i] \times \{0\} \times \mathbb{R}$ for $1 \leq i \leq k$ or $[j-1, j] \times \{1\} \times \mathbb{R}$ for $1 \leq j \leq l$. The corresponding images under π_{T^1} are

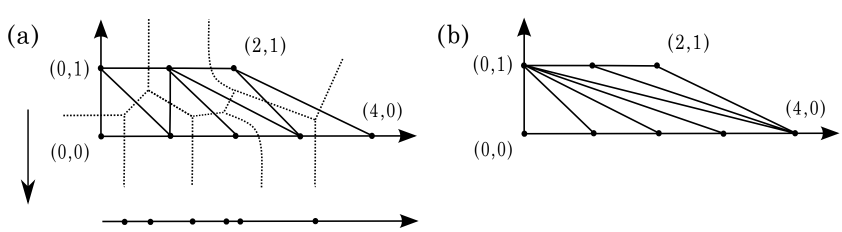


Figure 4. Left: moment map π_{T^1}. Right: crepant resolution.

isolated points $s_1, \dots, s_k$ or $t_1, \dots, t_l$ respectively. Moreover $z = 0$ on a toric strata corresponding to a cone $[i-1, i] \times \{0\} \times \mathbb{R}$, while $w = 0$ on a toric strata corresponding to a cone $[j-1, j] \times \{1\} \times \mathbb{R}$. Hence $b_1 = 1$ or $b_2 = 1$ respectively, and the discriminant locus is as stated above. □

Proposition 3.8. *The wall H of the fibration π is given by the union of two vertical planes given by $b_1 = 0$ and $b_2 = 0$.*

Proof. Suppose a fiber F_r bounds a non-constant holomorphic disc u of Maslov index 0. By the Maslov index formula in Lemma 3.2, the disc does not intersect the boundary divisors $\{z = 0\}$ nor $\{w = 0\}$. Thus the functions $(z-1)\circ u$ and $(w-1)\circ u$ can only be constants. If both the numbers $z \circ u$ and $w \circ u$ are non-zero, the fiber of (z, w) is just a cylinder, and a fiber of b_3 defines a non-contractible circle in this cylinder, which topologically does not bound any non-trivial disc. Thus either $z = 0$ or $w = 0$ on the disc, which implies that $b_1 = \log|z-1| = 0$ or $b_2 = \log|w-1| = 0$. In these cases F_r intersects a toric divisor and bounds holomorphic discs in the toric divisor. □

Figure 5, left, illustrates the wall stated in the above proposition.

From now on we fix the unique crepant resolution of $G_{k,l}$ such that $s_1 < \dots < s_k < t_1 \dots < t_l$ holds (Figure 4, right). For other crepant resolution the construction is similar (while the SYZ mirrors have different coefficients, namely the mirror maps are different). Such a choice is just for simplifying the notations and is not really necessary (see also Remark 3.13).

We then fix the basis

$$\{C_i\}_{i=1}^{k-1} \cup \{C_0\} \cup \{E_i\}_{i=1}^{l-1} \tag{3}$$

of $H_2(\widehat{G_{k,l}})$ (Figure 5, right), where C_i for $1 \le i \le k-1$ is the holomorphic sphere class represented by the toric 1-stratum corresponding to the 2-cone by $\{(0,1), (i,0)\}$; C_0 corresponds to the 2-cone generated by $\{(0,1), (k,0)\}$; E_i corresponds to the 2-cone generated by $\{(i,1), (k,0)\}$. The image of a holomorphic sphere in C_i under the fibration map lies in $\{0\} \times \mathbb{R} \times [s_i, s_{i+1}]$; the image of C_0 lies in $\{0\} \times \{0\} \times [s_k, t_1]$, and the image of E_i lies in $\mathbb{R} \times \{0\} \times [t_i, t_{i+1}]$.

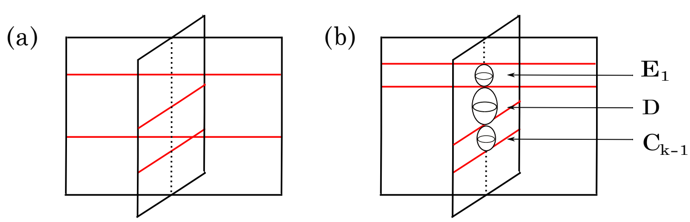


Figure 5. Left: walls. Right: holomorphic spheres.

Fix the contractible open subset

$$U := B_0 \setminus \{(b_1, b_2, b_3) \mid b_1 = 0 \text{ or } b_2 = 0, b_3 \in [s_1, +\infty)\} \subset B_0$$

over which the Lagrangian fibration π trivializes. For $b = (b_1, b_2, b_3)$ with $b_1 > 0$ and $b_2 > 0$, we use the Lagrangian isotopy

$$L^t = \{\log|z - t| = b_1, \log|w - t| = b_2, \pi_T(x, y, z, w) = b_3\} \tag{4}$$

for $t \in [0, 1]$ to link a moment-map fiber (when $t = 0$) with a Lagrangian torus fiber F_b of π (when $t = 1$). Then for a general base point $b' \in U$, we can link the fibers F_b and $F_{b'}$ by a Lagrangian isotopy induced by a path joining b and b' in the contractible set U (and the isotopy is independent of choice of the path). Through the isotopy disc classes bounded by a moment-map fiber L can be identified with those bounded by F_b, that is, $\pi_2(\widehat{G_{k,l}}^*, F_b) \cong \pi_2(\widehat{G_{k,l}}^*, L)$. Note that this identification depends on choice of trivialization, and henceforth we fix such a choice. The two vertical walls $\{b_1 = 0\}$ and $\{b_2 = 0\}$ divides the base $B = [-\infty, \infty]^2 \times I$ into four chambers. Lagrangian torus fibers over different chambers have different open Gromov–Witten invariants.

Theorem 3.9. *Denote by L a moment-map fiber of $\widehat{G_{k,l}}^*$ and by F_b a Lagrangian torus fiber of π at $b \in B_0$. Let $\beta \in \pi_2(\widehat{G_{k,l}}^*, F_b)$ with $\beta \pitchfork D$.*

(1) *Over the chamber $C_{++} := \{b_1 > 0, b_2 > 0\}$, we have $n_\beta^{F_b} = n_\beta^L$.*

(2) *Over the chamber $C_{+-} := \{b_1 > 0, b_2 < 0\}$, we have $n_\beta^{F_b} = 0$ unless $\beta = \beta_{k+1}$, $\beta_{\xi=0}$, $\beta_{\xi=\infty}$, $\beta_{z=\infty}$, $\beta_{w=\infty} + (\beta_j - \beta_{k+1}) + \alpha$ for $k+1 \le j \le k+l+1$ and $\alpha \in H_2^{c_1=0}$ being a class of rational curves which intersect the open toric orbit of the toric divisor D_j, or $\beta = \beta_i + \alpha$ for $0 \le i \le k$ and $\alpha \in H_2^{c_1=0}$ being a class of rational curves which intersect the open toric orbit of the toric divisor D_i. Moreover*

$$n^{F_b}_{\beta_{k+1}} = n^{F_b}_{\beta_{\xi=0}} = n^{F_b}_{\beta_{\xi=\infty}} = n^{F_b}_{\beta_{z=\infty}} = 1, \quad n^{F_b}_{\beta_{w=\infty}+(\beta_j-\beta_{k+1})+\alpha} = n^L_{\beta_j+\alpha}$$

for $k+1 \le j \le k+l+1$, and

$$n^{F_b}_{\beta_i+\alpha} = n^L_{\beta_i+\alpha}$$

for $0 \le i \le k$.

(3) *Over the chamber $C_{-+} := \{b_1 < 0, b_2 > 0\}$, we have $n_\beta^{F_b} = 0$ unless $\beta = \beta_0$, $\beta_{\xi=0}$, $\beta_{\xi=\infty}$, $\beta_{w=\infty}$, $\beta_{z=\infty} + (\beta_i - \beta_0) + \alpha$ for $0 \le i \le k$ and $\alpha \in H_2^{c_1=0}$ being a class of rational curves which intersect the open toric orbit of the toric divisor D_i, or $\beta = \beta_j + \alpha$ for $k+1 \le j \le k+l+1$ and α being a class of rational curves which intersect the open toric orbit of the toric divisor D_j. Moreover*

$$n^{F_b}_{\beta_0} = n^{F_b}_{\beta_{\xi=0}} = n^{F_b}_{\beta_{\xi=\infty}} = n^{F_b}_{\beta_{w=\infty}} = 1, \quad n^{F_b}_{\beta_{z=\infty}+(\beta_i-\beta_0)+\alpha} = n^L_{\beta_i+\alpha}$$

for $0 \leq i \leq k$, *and*

$$n^{F_b}_{\beta_j+\alpha} = n^L_{\beta_j+\alpha}$$

for $k+1 \leq j \leq k+l+1$.

(4) *Over the chamber* $C_{--} := \{b_1 < 0, b_2 < 0\}$, *we have* $n^{F_b}_\beta = 0$ *unless* $\beta = \beta_0$, β_{k+1}, $\beta_{\xi=0}$, $\beta_{\xi=\infty}$, $\beta_{z=\infty} + (\beta_i - \beta_0) + \alpha$ *for* $0 \leq i \leq k$ *and* $\alpha \in H_2^{c_1=0}$ *being a class of rational curves which intersect the open toric orbit of the toric divisor* D_i, *or* $\beta_{w=\infty} + (\beta_j - \beta_{k+1}) + \alpha$ *for* $k+1 \leq j \leq k+l+1$ *and* $\alpha \in H_2^{c_1=0}$ *being a class of rational curves which intersect the open toric orbit of the toric divisor* D_j. *Moreover*

$$n^{F_b}_{\beta_0} = n^{F_b}_{\beta_{k+1}} = n^{F_b}_{\beta_{\xi=0}} = n^{F_b}_{\beta_{\xi=\infty}} = 1, \quad n^{F_b}_{\beta_{z=\infty}+(\beta_i-\beta_0)+\alpha} = n^L_{\beta_i+\alpha}$$

for $0 \leq i \leq k$, *and*

$$n^{F_b}_{\beta_{w=\infty}+(\beta_j-\beta_{k+1})+\alpha} = n^L_{\beta_j+\alpha}$$

for $k+1 \leq j \leq k+l+1$.

Proof. The open Gromov–Witten invariants n_β is non-zero only when β has Maslov index 2, and so we can focus on $\mu(\beta) = 2$ with $\beta \pitchfork D$.

For a fiber $F_{(b_1,b_2,b_3)}$ with $b_1 > 0$ and $b_2 > 0$, $F_{(b_1,b_2,b_3)}$ is Lagrangian isotopic to a moment-map fiber L through L_t defined by (4). Moreover each L_t does not bound any holomorphic disc of Maslov index 0 because for every $t \in [0, 1]$, the circles $|z-t| = b_1$ and $|w-t| = b_2$ never pass through $z = 0$ and $w = 0$ respectively. Thus the open Gromov–Witten invariants of L and that of $F_{(b_1,b_2,b_3)}$ are the same.

Now consider the chamber C_{+-}. First we use the Lagrangian isotopy

$$L^{1,t} = \{\log|z-1| = b_1, \log|w-t| = b_2, \pi_T(x, y, z, w) = b_3\}$$

for $t \in [1, R]$ which identifies $F_{(b_1,b_2,b_3)}$ with $L^{1,R}$ for $R \gg 0$. $L^{1,t}$ never bounds any holomorphic disc of Maslov index 0, since for every $t \in [1, R]$, the circles $|z-1| = b_1$ and $|w-t| = b_2$ never pass through $z = 0$ and $w = 0$ respectively.

Then we take the involution $\iota : \widehat{G_{k,l}}^* \to \widehat{G_{k,l}}^*$ defined as identity on z, π_{T^1}, θ and mapping $w \mapsto \frac{Rw}{w-R}$. This involution maps the fiber $L^{1,R}$ to the Lagrangian

$$L' = \{\log|z-1| = b_1, \log|w-R| = 2(\log R) - b_2, \pi_T(x, y, z, w) = b_3\}$$

which can again be identified with the fiber $F_{(b_1,2(\log R)-b_2,b_3)}$ with $b_1 > 0$ and $2(\log R) - b_2 > 0$. Also ι tends to the negative identity map as R tends to infinity. Hence for $R \gg 0$, the pulled-back complex structure by ι is a small deformation of the original complex structure, and hence the open Gromov–Witten invariants of $F_{(b_1,b_2,b_3)}$ remain invariant. Now using Case 1 the open Gromov–Witten invariants of $F_{(b_1,2(\log R)-b_2,b_3)}$ can be identified with a moment-map fiber L.

By considering the intersection numbers of the disc classes and the divisors, one can check that the disc classes β_i for $0 \le i \le k$, β_{k+1}, $\beta_{\xi=0}$, $\beta_{\xi=\infty}$, $\beta_{z=\infty}$, $(\beta_j - \beta_{k+1})$ for $k+1 \le j \le k+l+1$, and all rational curve classes α are invariant under the involution. Moreover β_{k+1} and $\beta_{w=\infty}$ are switched under the involution. Putting all together, we obtain a complete relation between open Gromov–Witten invariants of $F_{(b_1,b_2,b_3)}$ and that of L, and this gives the formulae in Case 2.

Cases 3 and 4 are similar. For Case 3 we use the Lagrangian isotopy

$$L^{t,1} = \{\log|z-t| = b_1, \log|w-1| = b_2, \pi_T(x, y, z, w) = b_3\}$$

for $t \in [1, R]$ and the involution defined as identity on w, π_{T^1}, θ and mapping $z \mapsto \frac{Rz}{z-R}$. For Case 4 we use the Lagrangian isotopy

$$L^{t} = \{\log|z-t| = b_1, \log|w-t| = b_2, \pi_T(x, y, z, w) = b_3\}$$

for $t \in [1, R]$ and the involution defined as identity on π_{T^1}, θ and mapping $z \mapsto \frac{Rz}{z-R}$, $w \mapsto \frac{Rw}{w-R}$. □

We can compute the the open Gromov–Witten invariants of the moment-map fiber, using the open mirror theorem [Chan et al. 2016, Theorem 1.4 (1)]. The result is essentially the same as the one in [Lau et al. 2012, Theorem 4.2] for the minimal resolution of A_n-singularities as follows.

Theorem 3.10. *Let L be a regular moment-map fiber of $\widehat{G_{k,l}}$, and consider a disc class $\beta \in \pi_2(X, L)$. The open Gromov–Witten invariant n_β^L equals to 1 if $\beta = \beta_p + \alpha$ for $0 \le p \le k+l+1$, where α is a class of rational curves which takes the form*

$$\alpha = \begin{cases} \sum_{i=1}^{k-1} s_i C_i & \text{when } 1 \le p \le k-1, \\ \sum_{i=1}^{l-1} s_i E_i & \text{when } k+2 \le p \le k+l, \\ 0 & \text{when } p = 0, k, k+1 \text{ or } k+l+1, \end{cases}$$

and $\{s_i\}_{i=1}^{m-1}$ (where m equals to k in the first case and l in the second case) is an admissible sequence with center p in the first case, which means that

(1) *$s_i \ge 0$ for all i and $s_1, s_{m-1} \le 1$;*

(2) *$s_i \le s_{i+1} \le s_i + 1$ when $i < p$, and $s_i \ge s_{i+1} \ge s_i - 1$ when $i \ge p$,*

and with center $p-k-1$ in the second case. For any other β, $n_\beta^L = 0$.

Proof. We will prove the assertion by using the open mirror theorem. Recall the curve classes C_i, C_0 and E_j introduced in (3) for $1 \le i \le k-1$ and $1 \le j \le l-1$. C_i and E_i are $(-2, 0)$-curves, while C_0 is a $(-1, -1)$-curve. The intersection numbers with the toric prime divisors D_j are as follows:

(1) $C_i \cdot D_{i-1} = C_i \cdot D_{i+1} = 1$; $C_i \cdot D_i = -2$; and $C_i \cdot D_j = 0$ for all $j \ne i-1, i, i+1$;

(2) $C_0 \cdot D_k = C_0 \cdot D_{k+1} = -1$; $C_0 \cdot D_{k-1} = C_0 \cdot D_{k+2} = 1$; and $C_0 \cdot D_j = 0$ for all $j \neq k-1, k, k+1, k+2$;

(3) $E_i \cdot D_{k+i-1} = E_i \cdot D_{k+i+1} = 1$; $E_i \cdot D_{k+i} = -2$; and $E_i \cdot D_j = 0$ for all $j \neq k+i-1, k+i, k+i+1$.

Let $q^{C_i}, q^{C_0}, q^{E_i}$ be the corresponding Kähler parameters and $\check{q}^{C_i}, \check{q}^{C_0}, \check{q}^{E_i}$ be the corresponding complex parameters. They are related by the mirror map:

$$q^{C_i} = \check{q}^{C_i} \exp\left(-\sum_{j=0}^{k+1+l} (C_i \cdot D_j) g_j(\check{q})\right) = \check{q}^{C_i} \exp\bigl(-(g_{i-1}(\check{q}) + g_{i+1}(\check{q}) - 2g_i(\check{q}))\bigr),$$

$$q^{C_0} = \check{q}^{C_0} \exp\left(-\sum_{j=0}^{k+1+l} (C_0 \cdot D_j) g_j(\check{q})\right) = \check{q}^{C_0} \exp\bigl(-(g_{k-1}(\check{q}) + g_{k+1}(\check{q}))\bigr),$$

$$\begin{aligned} q^{E_i} &= \check{q}^{E_i} \exp\left(-\sum_{j=0}^{k+1+l} (E_i \cdot D_j) g_j(\check{q})\right) \\ &= \check{q}^{E_i} \exp\bigl(-(g_{k+i-1}(\check{q}) + g_{k+i+1}(\check{q}) - 2g_{k+i}(\check{q}))\bigr). \end{aligned}$$

The functions $g_i(\check{q})$ are attached to the toric prime divisor D_i for $0 \leq i \leq k+l+1$. We have $g_0 = g_k = g_{k+1} = g_{k+l+1} = 0$. Moreover for $1 \leq i \leq k-1$, g_i only depends on the variables $\check{q}^{C_r}$ for $1 \leq r \leq k-1$; for $k+2 \leq i \leq k+l$, the function g_i only depends on the variables $\check{q}^{E_r}$ for $1 \leq r \leq l-1$. Explicitly g_i is written in terms of hypergeometric series:

$$g_i(\check{q}) := \sum_{\substack{d \cdot D_i < 0 \\ d \cdot D_r \geq 0 \\ \text{for all } r \neq i}} \frac{(-1)^{(D_i \cdot d)}(-(D_i \cdot d) - 1)!}{\prod_{p \neq i}(D_p \cdot d)!} \check{q}^d$$

where for $1 \leq i \leq k-1$, the summation is over $d = \sum_{r=1}^{k-1} n_j C_j$ ($n_j \in \mathbb{Z}_{\geq 0}$) with $d \cdot D_i < 0$ and $d \cdot D_p \geq 0$ for all $p \neq i$; for $k+2 \leq i \leq k+l$, the summation is over $d = \sum_{r=1}^{l-1} n_j E_j$ ($n_j \in \mathbb{Z}_{\geq 0}$) with d satisfying the same condition. Then the open mirror theorem [Chan et al. 2016, Theorem 1.4 (1)] states that

$$\sum_{\alpha} n_{\beta_i + \alpha} q^{\alpha}(\check{q}) = \exp g_i(\check{q}).$$

Note that for $1 \leq i \leq k-1$, the function g_i takes exactly the same expression as that in the toric resolution of A_{k-1}-singularity; for $1 \leq i \leq l-1$, the function g_{i+k+1} takes exactly the same expression as that in the toric resolution of A_{l-1}-singularity. Thus the mirror maps for q^{C_i} and q^{E_i} coincide with that for A_{k-1}-resolution and A_{l-1}-resolution respectively. Moreover the above generating function of open Gromov–Witten invariants coincide. Then result follows from the formula for open

Gromov–Witten invariants in A_n-resolution given in [Lau et al. 2012, Theorem 4.2]. □

Remark 3.11. Theorem 3.10 can also be proved by comparing the disc moduli for $\widehat{G_{k,l}}$ and resolution of A_k- and A_l-singularities, which involves details of obstruction theory of disc moduli space. Here take the more combinatorial approach using open mirror theorems.

Theorem 3.12. *The SYZ mirror of the resolved generalized conifold $\widehat{G_{k,l}}$ is given by the deformed orbifolded conifold in $\mathbb{C}^4 \times \mathbb{C}^\times$ defined by the following equations, where $U_1, U_2, V_1, V_2 \in \mathbb{C}$ and $Z \in \mathbb{C}^\times$:*

$$U_1 V_1 = (1+Z)(1+q_1 Z)\dots(1+q_1\dots q_{k-1}Z),$$
$$U_2 V_2 = (1+cZ)(1+q'_1 cZ)\dots(1+q'_1\dots q'_{l-1}cZ),$$

and where $q_i = \mathrm{e}^{-\int_{C_i}\omega}$, $q'_j = \mathrm{e}^{-\int_{E_i}\omega}$, and $c = q_1\dots q_{k-1}\mathrm{e}^{-\int_{C_0}\omega}$.

Proof. Let $\tilde{Z}_\beta = \mathrm{e}^{-\int_\beta \omega}\,\mathrm{Hol}_\nabla(\partial\beta)$ be the semi-flat mirror complex coordinates corresponding to each disc class $\beta \in \pi_2(\widehat{G_{k,l}}^*, F_b)$ for $b \in B_0 \setminus H$. For simplicity we denote $\tilde{Z}_{\beta_{\xi=0}}$ by $\tilde{Z}$, $\tilde{Z}_{\beta_0}$ by $\tilde{U}_1$, $\tilde{Z}_{\beta_{k+1}}$ by $\tilde{U}_2$, $\tilde{Z}_{\beta_{z=\infty}}$ by $\tilde{V}_1$, and $\tilde{Z}_{\beta_{w=\infty}}$ by $\tilde{V}_2$. We have $\tilde{Z}\tilde{Z}_{\beta_{\xi=\infty}} = q^{\beta_{\xi=0}+\beta_{\xi=\infty}}$ is a constant (since $\beta_{\xi=0}+\beta_{\xi=\infty} \in H_2(\widehat{G_{k,l}}^*)$), and for simplicity we set the constant to be 1. Thus $\tilde{Z}_{\beta_{\xi=\infty}} = \tilde{Z}^{-1}$.

Let $Z_{\mathcal{D}}$ be the generating function of open Gromov–Witten invariants corresponding to a boundary divisor $\mathcal{D}$ (see (2)). By Theorem 3.9 there is no wall-crossing for the disc classes $\beta_{\xi=0}$ and $\beta_{\xi=\infty}$, and they are the only disc classes of Maslov index 2 and intersecting $\mathcal{D}_{\xi=0}$ and $\mathcal{D}_{\xi=\infty}$ exactly once respectively. Hence $Z_{\mathcal{D}_{\xi=0}} = \tilde{Z}$ and $Z_{\mathcal{D}_{\xi=\infty}} = \tilde{Z}^{-1}$. For simplicity we denote $Z_{\mathcal{D}_{\xi=0}}$ by Z, and hence $Z = \tilde{Z}$ (meaning that the coordinate Z does not need quantum corrections).

Theorem 3.10 gives the open Gromov–Witten invariants of moment-map tori, which then gives the open Gromov–Witten invariants of fibers of our Lagrangian fibration by Theorem 3.9. Then by some nice combinatorics which also appears in [Lau et al. 2012, Proof of Corollary 4.3], the generating functions factorizes as follows:

(1) $Z_{\mathcal{D}_{z=1}} = \tilde{U}_1$ and $Z_{\mathcal{D}_{w=1}} = \tilde{U}_2$ over C_{--},

(2) $Z_{\mathcal{D}_{z=1}} = \tilde{U}_1(1+Z)(1+q_1Z)\dots(1+q_1\dots q_{k-1}Z)$ and $Z_{\mathcal{D}_{w=1}} = \tilde{U}_2$ over C_{+-},

(3) $Z_{\mathcal{D}_{z=1}} = \tilde{U}_1$ and $Z_{\mathcal{D}_{w=1}} = \tilde{U}_2(1+cZ)(1+q'_1cZ)\dots(1+q'_1\dots q'_{l-1}cZ)$ over C_{-+},

(4) $Z_{\mathcal{D}_{z=1}} = \tilde{U}_1(1+Z)(1+q_1Z)\dots(1+q_1\dots q_{k-1}Z)$ and $Z_{\mathcal{D}_{w=1}} = \tilde{U}_2(1+cZ)(1+q'_1cZ)\dots(1+q'_1\dots q'_{l-1}cZ)$ over C_{++}.

(5) $Z_{\mathcal{D}_{z=\infty}} = \tilde{U}_1^{-1}$ over $C_{++} \cup C_{+-}$,

(6) $Z_{\mathcal{D}_{z=\infty}} = \tilde{U}_1^{-1}(1+Z)(1+q_1 Z)\dots(1+q_1\dots q_{k-1}Z)$ over $C_{-+}\cup C_{--}$,

(7) $Z_{\mathcal{D}_{w=\infty}} = \tilde{U}_2^{-1}$ over $C_{++}\cup C_{-+}$,

(8) $Z_{\mathcal{D}_{w=\infty}} = \tilde{U}_2^{-1}(1+cZ)(1+q_1'cZ)\dots(1+q_1'\dots q_{l-1}'cZ)$ over $C_{+-}\cup C_{--}$.

Therefore we conclude that the ring generated by the functions $Z = Z_{\mathcal{D}_{\xi=0}}$, $Z_{\mathcal{D}_{\xi=\infty}} = Z^{-1}$, $U_1 := Z_{\mathcal{D}_{z=1}}$, $V_1 := Z_{\mathcal{D}_{z=\infty}}$, $U_2 := Z_{\mathcal{D}_{w=1}}$, $V_2 := Z_{\mathcal{D}_{w=\infty}}$ is the polynomial ring $\mathbb{C}[U_1, U_2, V_1, V_2, Z, Z^{-1}]$ mod out by the relations

$$U_1V_1 = (1+Z)(1+q_1Z)\dots(1+q_1\dots q_{k-1}Z),$$
$$U_2V_2 = (1+cZ)(1+q_1'cZ)\dots(1+q_1'\dots q_{l-1}'cZ).$$

This completes the proof of the theorem. □

Remark 3.13. If a different crepant resolution is taken, the mirror takes the same form as above while the coefficients of the polynomials on the right hand side are different functions of the Kähler parameters q_i and q_i'. They correspond to different choices of limit points (and hence different flat coordinates) over the complex moduli.

SYZ from $\widetilde{G_{k,l}}$ to $\widehat{O_{k,l}}$. The deformed generalized conifold $\widetilde{G_{k,l}}$ is given by

$$\Big\{(x,y,z,w)\in\mathbb{C}^2\times(\mathbb{C}\setminus\{1\})^2 \;\Big|\; xy - \sum_{i=0}^{k}\sum_{j=0}^{l} a_{i,j}z^iw^j = 0\Big\}$$

for generic coefficients $a_{i,j}\in\mathbb{C}$. It is a conic fibration over the second factor $(\mathbb{C}\setminus\{1\})^2$ with discriminant locus being the Riemann surface $\Sigma_{k,l}\subset(\mathbb{C}\setminus\{1\})^2$ defined by the equation $\sum_{i=0}^{k}\sum_{j=0}^{l} a_{i,j}z^iw^j = 0$ which has genus kl and $(k+l)$ punctures. The SYZ construction for such a conic fibration follows from [Abouzaid et al. 2016]. Here we just give a brief description. We will use the standard symplectic form on $\mathbb{C}^2\times(\mathbb{C}\setminus\{1\})^2$ restricted to the hypersurface $\widetilde{G_{k,l}}$. First, $\widetilde{G_{k,l}}$ is naturally compactified in $\mathbb{P}^2\times(\mathbb{P}^1)^2$ as a symplectic manifold, and we denote the compactification by $\widetilde{G_{k,l}}^*$. There is a natural Hamiltonian T^1-action on $\widetilde{G_{k,l}}^*$ given by, for $t\in T^1\subset\mathbb{C}$

$$t\cdot(x,y,z,w) := (tx, t^{-1}y, z, w).$$

By carefully analyzing the symplectic reduction of this T^1-action, a Lagrangian torus fibration $\pi:\widetilde{G_{k,l}}^*\to B := [-\infty,\infty]^3$ was constructed in [Abouzaid et al. 2016, Section 4]. Topologically the fibration is the homeomorphic to the naive one given by

$$(x,y,z,w)\mapsto(b_1,b_2,b_3) = \big(\log|z|, \log|w|, \tfrac{1}{2}(|x|^2-|y|^2)\big).$$

However since the symplectic form induced on the symplectic quotient is not the standard one on $\mathbb{P}^2$, it has to be deformed to give a Lagrangian fibration.

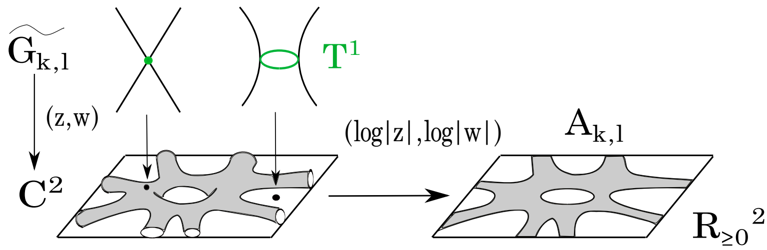


Figure 6. Conic fibration and amoeba $A_{k,l}$.

The discriminant locus of this fibration consists of the boundary of B and $\overline{A_{k,l}} \times \{0\}$, where $\overline{A_{k,l}} \subset [-\infty, \infty]^2$ is the compactification of the amoeba $A_{k,l} \subset \mathbb{R}^2_{\geq 0}$ of the Riemann surface $\Sigma_{k,l}$ (Figure 6), namely the image of $\Sigma_{k,l}$ under the map $(z, w) \mapsto (\log|z|, \log|w|)$.

The wall for open Gromov–Witten invariants is given by $H = \overline{A_{k,l}} \times [-\infty, \infty]$ [Abouzaid et al. 2016, Proposition 5.1]. The complement $B \setminus H$ consists of $(k+1) \times (l+1)$ chambers. In this specific case, we have a nice degeneration as follows. Let us consider a special point on the the complex moduli space of $\widetilde{G_{k,l}}^*$ where the defining equation of the Riemann surface $\Sigma_{k,l}$ factorizes as

$$\sum_{i=0}^{k} \sum_{j=0}^{l} a_{i,j} z^i w^j = f(z) g(w)$$

for polynomials $f(z)$ and $g(w)$ of degree k and l respectively (and we assume that their roots are all distinct and non-zero). At this point, $\widetilde{G_{k,l}}^*$ acquires kl conifold singularities. The wall becomes the union of vertical hyperplanes

$$\{b_2 = \log|r_i|\}_{i=1}^{k} \cup \{b_3 = \log|s_j|\}_{j=1}^{l} \subset \mathbb{R}^3,$$

where r_i and s_j are the roots of $f(z)$ and $g(w)$ respectively (Figure 7).

These hyperplanes divide the base into $(k+1)(l+1)$ chambers. We label the chambers by $C_{i,j}$ for $i = 0, \ldots, k$ and $j = 0, \ldots, l$ from left to right and from bottom to up.

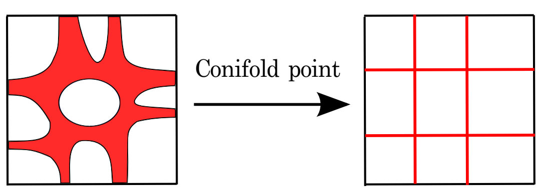


Figure 7. Amoeba around conifold locus.

Theorem 3.14. *The SYZ mirror of the deformed generalized conifold $\widetilde{G_{k,l}}$ is given by, for $U_i, V_i \in \mathbb{C}$ and $Z \in \mathbb{C}^\times$,*

$$U_1 V_1 = (1+Z)^k, \quad U_2 V_2 = (1+Z)^l,$$

which is the punctured orbifolded conifold $O_{k,l}$.

Proof. The wall-crossing of open Gromov–Witten invariants was deduced in [Abouzaid et al. 2016, Lemma 5.4], and we just sketch the result here. For $p = 1, 2$, let U_p be the generating function of open Gromov–Witten invariants for disc classes intersecting the boundary divisor $\pi^{-1}(\{b_p = 0\})$ once. Denote the semi-flat coordinates corresponding to the basic disc classes emanated from $\pi^{-1}(\{b_p = 0\})$ by $\tilde{Z}_i$, and denote by Z the semi-flat coordinate corresponding to the b_3-direction (which admits no quantum corrections). Then U_1 restricted to the chamber $C_{i,j}$ equals to the polynomial $\tilde{Z}_1(1+Z)^i$, and U_2 restricted to the chamber $C_{i,j}$ equals to $\tilde{Z}_2(1+Z)^j$. By gluing the various chambers together using the above wall-crossing factor $1 + Z$, we obtain the SYZ mirror as claimed. □

Remark 3.15. The equation in Theorem 3.14 defines a singular variety $O_{k,l}$. This is a typical feature of our SYZ construction, which produces a complex variety out of a symplectic manifold: we may obtain a singular variety as the SYZ mirror, and we need to take a crepant resolution to get a smooth mirror. Since we concern about the complex geometry of this variety, $O_{k,l}$ is not distinguishable from its crepant resolution $\widehat{O_{k,l}}$.

SYZ from $\widehat{O_{k,l}}$ to $\widetilde{G_{k,l}}$. The partial compactification $\widehat{O^\sharp_{k,l}}$ of $\widehat{O_{k,l}}$ is a toric Calabi–Yau threefold and its SYZ mirror was constructed in [Chan et al. 2012]. In this section we quote the relevant results, omitting the details. First, a crepant resolution $\widehat{O_{k,l}}$ of $\widehat{O_{k,l}}$ corresponds to a maximal triangulation of $\Box_{k,l}$ (Figure 2, right). We have the Lagrangian torus fibration

$$\pi : \widehat{O_{k,l}} \to B := \mathbb{R}^2 \times \mathbb{R}_{\geq 0}$$

constructed in [Goldstein 2001; Gross 2001], whose discriminant locus consists of two components. One is the boundary ∂B, and the other is topologically given by the dual graph of the maximal triangulation lying in the hyperplane $\{b_3 = 1\} \subset B$, where we denote the coordinates of B by $b = (b_1, b_2, b_3)$.

The wall is exactly the hyperplane $\{b_3 = 1\}$ containing one component of the discriminant locus. It is associated with a wall-crossing factor, which is a polynomial whose coefficients encode the information coming from holomorphic discs with Maslov index 0. The explicit formula for the coefficients were computed in [Chan et al. 2016]. Applying these results, we obtain the following:

Theorem 3.16. *The SYZ mirror of the resolved orbifolded conifold $\widehat{O_{k,l}}$ is a deformed generalized conifold $\widetilde{G_{k,l}}$ given as*

$$\Big\{(U, V, Z, W) \in \mathbb{C}^2 \times (\mathbb{C}^\times)^2 \;\Big|\; UV = \sum_{i=0}^{k} \sum_{j=0}^{l} q^{C_{ij}} (1 + \delta_{ij}(\check{q})) Z^i W^j \Big\}.$$

The notations are explained as follows. Let β_{ij} be the basic disc class corresponding to the toric divisor $D_{ij} \subset \widehat{O_{k,l}}$. Then C_{ij} denotes the curve class $\beta_{ij} - i(\beta_{10} - \beta_{00}) - j(\beta_{01} - \beta_{00}) - \beta_{00}$. The coefficients $1 + \delta_{ij}(\check{q})$ is given by $\exp(g_{ij}(\check{q}))$ where

$$g_{ij}(\check{q}) := \sum_{d} \frac{(-1)^{(D_{ij}\cdot d)}(-(D_{ij} \cdot d) - 1)!}{\prod_{(a,b)\neq(i,j)}(D_{ab} \cdot d)!} \check{q}^d,$$

and the summation is over all effective curve classes $d \in H_2^{\text{eff}}(O_{k,l})$ satisfying $D_{ij} \cdot d < 0$ and $D_p \cdot d \geq 0$ for all $p \neq (i, j)$. Lastly q and $\check{q}$ are related by the mirror map:

$$q^C = \check{q}^C \exp\Big(- \sum_{i,j} (D_{ij} \cdot C) g_{ij}(\check{q}) \Big).$$

It is worth noting that the above SYZ mirror manifold can be identified with the Hori–Iqbal–Vafa mirror manifold [Hori et al. 2000]. The former has the advantage that it is intrinsically expressed in terms of flat coordinates and contains the information about certain open Gromov–Witten invariants.

SYZ from $\widetilde{O_{k,l}}$ to $\widehat{G_{k,l}}$. Recall that the fan polytope of the orbifolded conifold $O^\sharp_{k,l}$ is the cone over a rectangle $[0, k] \times [0, l]$. Smoothings of $\widetilde{O_{k,l}}$ correspond to the Minkowski decompositions of $[0, k] \times [0, l]$ into k copies of $[0, 1] \times \{0\}$ and l copies of $\{0\} \times [0, 1]$ [Altmann 1997]. The SYZ mirrors for such smoothings were constructed in [Lau 2014]. Here we can write down the Lagrangian fibration more explicitly by realizing $\widetilde{O_{k,l}}$ as a double conic fibration. Recall that the deformed orbifolded conifold $\widetilde{O_{k,l}}$ is defined as

$$\widetilde{O_{k,l}} = \{(u_1, u_2, v_1, v_2, z) \in \mathbb{C}^4 \times \mathbb{C}^\times \mid u_1 v_1 = f(z),\; u_2 v_2 = g(z)\}$$

where $f(z)$ and $g(z)$ are generic polynomials of degree k and l respectively. We assume that all roots r_i and s_j of $f(z)$ and $g(z)$ respectively are distinct and non-zero. Moreover, we can naturally compactly $\widetilde{O_{k,l}}$ in $(\mathbb{P}^2)^2 \times \mathbb{P}^1$ to obtain $\widetilde{O_{k,l}}^* \subset (\mathbb{P}^2)^2 \times \mathbb{P}^1$ (where (u_1, u_2) and (v_1, v_2) above become inhomogeneous coordinates of the two $\mathbb{P}^2$ factors.). There is also a natural Hamiltonian T^2-action on $\widetilde{O_{k,l}}^*$ given by, for $(s, t) \in T^2 \subset \mathbb{C}^2$

$$(s, t) \cdot (u_1, v_1, u_2, v_2, z) := (su_1, s^{-1}v_1, tu_2, t^{-1}v_2, z).$$

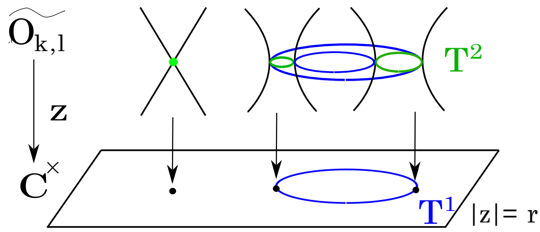


Figure 8. Double conic fibration.

On the other hand, $\widetilde{O_{k,l}}^*$ admits a double conic fibration $\pi_z : \widetilde{O_{k,l}}^* \to \mathbb{P}^1$ by the projection to the z-coordinate (Figure 8). In this situation, the base $\mathbb{P}^1$ of the fibration can be identified as the symplectic reduction of $\widetilde{O_{k,l}}^*$ by the Hamiltonian T^2-action.

As discussed in [Gross 2001], the Lagrangian fibration $|z| : \mathbb{P}^1 \to [0, \infty]$ gives rise to the Lagrangian torus fibration $\pi : \widetilde{O_{k,l}}^* \to B := [-\infty, \infty]^2 \times [0, \infty]$ given by

$$\pi(u_1, v_1, u_2, v_2, z) = \big(\tfrac{1}{2}(|u_1|^2 - |v_1|^2), \tfrac{1}{2}(|u_2|^2 - |v_2|^2), |z|\big).$$

The map to the first two coordinates is the moment map of the Hamiltonian T^2-action. We denote the coordinates of B by $b = (b_1, b_2, b_3)$.

Proposition 3.17. (1) *The discriminant locus of the fibration π is given by the disjoint union $\partial B \cup \big(\bigcup_{i=1}^k \{b_1 = 0, b_3 = |r_i|\}\big) \cup \big(\bigcup_{j=1}^l \{b_2 = 0, b_3 = |s_j|\}\big) \subset B$.*

(2) *The fibration π is special with respect to the nowhere-vanishing meromorphic volume form $\Omega := du_1 \wedge du_2 \wedge d \log z$ on $\widetilde{O_{k,l}}^*$.*

Proof. The fibration has tori T^3 as generic fibers. Over ∂B where $z = 0$, the fibers degenerate to T^2. Thus ∂B is a component of the discriminant locus. Away from $z = 0$, the map $z \to |z|$ is a submersion. Hence the discriminant locus of the fibration π comes from that of the moment map of the Hamiltonian T^2-action. This action has non-trivial stabilizers at $u_1 = v_1 = 0$ or $u_2 = v_2 = 0$, which implies $f(z) = 0$ or $g(z) = 0$ respectively. Their images under π are $\{b_1 = 0, b_3 = |r_i|\}$ or $\{b_2 = 0, b_3 = |s_j|\}$ respectively. □

Proposition 3.18. *A regular fiber of the fibration π bounds a holomorphic disc of Maslov index 0 only when $b_3 = |r_i|$ or $b_3 = |s_j|$. Thus the wall H of the fibration π is $H = \big(\bigcup_{i=1}^k \{b_3 = |r_i|\}\big) \cup \big(\bigcup_{j=1}^l \{b_3 = |s_j|\}\big) \subset B$.*

Proof. A singular fiber of the double conic fibration $\pi_z : \widetilde{O_{k,l}}^* \to \mathbb{P}^1$ bounds a holomorphic disc, which has Maslov index 0 by Lemma 3.2, and this happened only when $b_3 = |r_i|$ or $b_3 = |s_j|$. □

The wall components $\{b_3 = |r_i|\}$ and $\{b_3 = |s_j|\}$ correspond respectively to the pieces $[0, 1] \times \{0\}$ and $\{0\} \times [0, 1]$ of the Minkowski decomposition. Wall-crosing

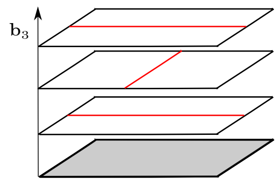


Figure 9. Walls: $(k, l) = (2, 1)$.

of open Gromov–Witten invariants in this case has essentially been studied in [Lau 2014] in details, and we will not repeat the details here. The key result is that each wall component contributes a linear factor: each component $\{b_3 = |r_i|\}$ contributes $1 + X$, and each component $\{b_3 = |s_j|\}$ contributes $1 + Y$. The SYZ mirror is essentially the product of all these factors, namely, we obtain the following.

Theorem 3.19. *The SYZ mirror of $\widetilde{O_{k,l}}$ is given by*

$$UV = (1 + X)^k (1 + Y)^l$$

for $U, V \in \mathbb{C}$ and $X, Y \in \mathbb{C}^\times$, which is the punctured generalized conifold $G_{k,l}$.

An almost the same remark as Remark 3.15 applies to Theorem 3.19 and thus we confirm the SYZ construction in this case.

4. Global geometric transitions: some discussion

We are now in position to turn to the global case. Let $\hat{X}$ and $\tilde{X}$ be compact Calabi–Yau threefolds. We call a geometric transition $\hat{X} \dashrightarrow X \rightsquigarrow \tilde{X}$ a generalized conifold transition if X has only generalized conifolds and orbifolded conifolds. The birational contraction appearing in the geometric transition of compact Calabi–Yau threefolds can be factorized into a sequence of primitive contractions of type I, type II and type III [Rossi 2006]. In general, type I and type III appear in the geometric transition of $G_{k,l}$ and all types appear in the geometric transition of $O_{k,l}$.

Motivated by the local case, we are tempted to propose that *generalized conifold transitions are reversed under mirror symmetry*. However, this naive conjecture does not hold because some global conifold transitions are mirror to hyperconifold transitions, which are not generalized conifold transitions [Davies 2011]. We expect that a generalized conifold transition is mirror to a reversed generalized conifold transition if the Calabi–Yau threefold has a Lagrangian torus fibration and the transition is locally modeled by those given in Section 3.

Example 4.1 (Schoen's CY threefold and its mirror [Schoen 1988; Lau 2018]). Using the methods of Castaño-Bernard and Matessi [2009; 2014] in the Gross–Siebert program, generalized conifold transitions and their mirrors are studied for the Schoen's Calabi–Yau threefold in [Lau 2018]. The threefold is a resolution of the fiber product of two rational elliptic surfaces over the base $\mathbb{P}^1$ [Schoen 1988]. It gives a global manifestation that orbifolded conifolds and generalized conifolds are mirror to each other.

For each pair of reflexive polygons (P_1, P_2) (*where P_1 and P_2 are not necessarily dual to each other*), we have an orbifolded conifold degeneration $O^{(P_1,P_2)}$ of a Schoen's Calabi–Yau threefold, and a generalized conifold degeneration $G^{(P_1,P_2)}$ of its mirror (in the sense of Legendre transform in Gross–Siebert program [2011]). A resolution of $O^{(P_1,P_2)}$ is mirror to a smoothing of $G^{(\check{P}_1,\check{P}_2)}$, and vice versa. $\check{P}$ denotes the dual polygon of P (see Figure 10 for an example). We refer the reader to [Lau 2018] for more details.

Example 4.2 (quintic threefold and its mirror). We set $(k,l) = (4,1)$ or $(3,2)$ in the following. Let $X \subset \mathbb{P}^4$ be the singular quintic threefold defined by

$$x_0 f(x_0, \dots, x_4) + x_1^k x_2^l = 0,$$

where $f(x)$ is a generic homogeneous polynomial of degree 4. The singular locus of X consists of 2 curves

$$\{x_0 = x_1 = f(x) = 0\} \cup \{x_0 = x_2 = f(x) = 0\} \subset X$$

of genus 3 intersecting at 4 points. The quintic threefold X has $G_{k,l}$ around the each intersection point. Successively blowing up X along the two curves followed by the blow-up along the divisor $\{x_0 = x_1 = 0\}$, we obtain a projective crepant resolution $\hat{X}$ of X. Thus a quintic threefold admits a generalized conifold transition.

On the other hand, the mirror quintic of a quintic threefold is defined as a crepant resolution of the orbifold

$$Y_\phi := \Big\{ \sum_{i=0}^{4} x_i^5 + \phi \prod_{i=0}^{4} x_i = 0 \Big\} \big/ G, \quad G := \Big\{ (a_i) \in (\mathbb{Z}_5)^5 \mid \sum_{i=0}^{4} a_i = 0 \Big\} \big/ \mathbb{Z}_5$$

for $\phi \in \mathbb{C}$. The orbifold Y_ϕ has A_4-singularities along ten curves

$$C_{ij} = \{x_i = x_j = 0\}/G \cong \mathbb{P}^1, \quad 0 \le i < j \le 4.$$

We can partially resolve Y_ϕ to obtain Y whose singular locus consists of A_k-singularities along C_{01} and A_l-singularities along C_{02} such that Y has $O_{k,l}$ around $C_{01} \cap C_{02}$ (Figure 11). It is interesting to ask whether or not Y admits any smoothing.

Although X and Y lie in the boundaries of the complex moduli space of the quintic and the Kähler moduli space of the mirror quintic respectively, we do not know whether or not they correspond each other under the mirror correspondence. This may be seen by the monomial-divisor correspondence in toric geometry, but it is possible that the mirror of X is a non-toric blow-down of the mirror quintic.

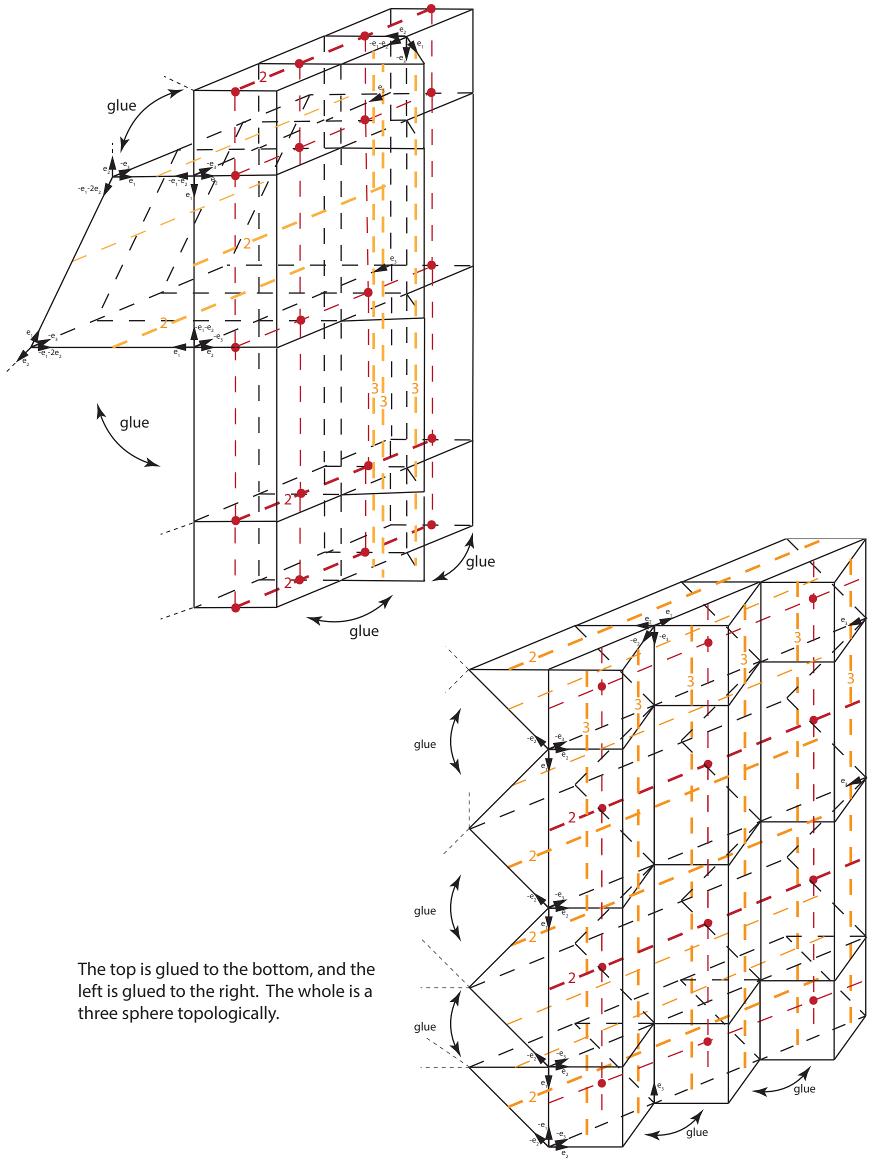


Figure 10. Base of Lagrangian fibrations on degenerations of a Schoen's Calabi–Yau threefold. Topologically they are $\mathbb{S}^3$ and the figures show polyhedral decompositions which are useful to describe the affine structures. The left shows an orbifolded conifold degeneration of Schoen's Calabi–Yau threefold. The right shows its mirror. Each thick dot represents an orbifolded conifold singularity on the left, and a generalized conifold singularity on the right. There are 24 orbifold singularities counted with multiplicities.

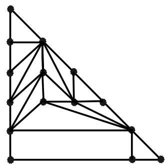


Figure 11. Two-dimensional faces of the polytope for $O_{4,1}$ and $O_{3,2}$.

It is straightforward to generalize this type of constructions to Calabi–Yau hypersurfaces in 4-dimensional weighted projective spaces.

We may also consider existence of generalized conifold transitions for compact Calabi–Yau geometries. Let $\hat{X}$ be a smooth threefold and $C_1, \dots, C_n$ be $(-1,-1)$-curves in $\hat{X}$. Let X be their contraction and $\tilde{X}$ be a smoothing of X. Small resolutions and deformations always exist topologically, but there are obstructions if we wish to preserve either the complex or symplectic structure:

Theorem 4.3 [Friedman 1986; Tian 1992]. *Assume that X satisfies the $\partial\bar{\partial}$-lemma (for example Kähler). Then a smoothing $\tilde{X}$ to exist if and only if there is a relation $\sum_{i=1}^{n} \lambda_i [C_i] = 0$ (with all the λ_i nonzero) in $H_2(X, \mathbb{Q})$.*

Theorem 4.4 [Smith et al. 2002]. *Let $\tilde{Y}$ be a symplectic sixfold with embedded Lagrangian S^3s', say $L_1, \dots, L_n$. Then there is a relation $\sum_{i=1}^{n} \lambda_i [L_i] = 0$ ($\lambda_i \neq 0\ \forall i$) in $H_3(\tilde{Y}, \mathbb{Q})$ if and only if there is a symplectic structure on one of 2^n choices of (reversed) conifold transitions of $\tilde{Y}$ in the Lagrnagians $L_1, \dots, L_n$, such that the resulting exceptional $\mathbb{P}^1$s' are symplectic.*

In our case, the contractions collapse 4-cycles as well as 2-cycles. On the other hand, smoothing $G_{k,l} \rightsquigarrow \widetilde{G_{k,l}}$ produces $(k+1)(l+1)-3$ vanishing S^3s' and smoothing $O_{k,l} \rightsquigarrow \widetilde{O_{k,l}}$ produces $k+l-2$ vanishing $S^1 \times S^2$s' and one vanishing S^3. The generators of these cycles can be found by considering the standard double Riemann surface fibrations [Feng et al. 2008] (Figure 12).

It is interesting to investigate the obstructions to the deformations/resolutions of the generalized and orbifolded conifolds in terms of these cycles. We hope to come back to these questions in future work.

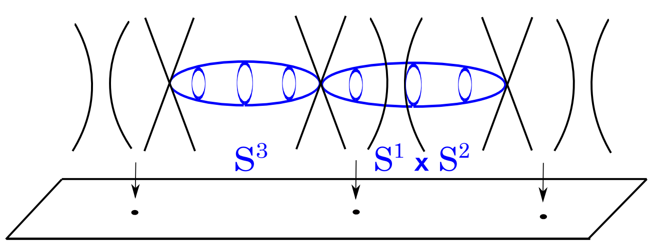


Figure 12. A double Riemann surface fibration.

Acknowledgement

The authors are grateful to Murad Alim, Kwokwai Chan, Yu-Wei Fan, Conan Leung, Shing-Tung Yau for useful discussions and encouragement. Kanazawa was supported by the Harvard CMSA. Lau appreciates the excellent environment provided by Harvard University in writing this paper. The Lau's work was supported by a Simons Collaboration Grant for Mathematicians at the time of publication.

References

[Abouzaid et al. 2016] M. Abouzaid, D. Auroux, and L. Katzarkov, "Lagrangian fibrations on blowups of toric varieties and mirror symmetry for hypersurfaces", *Publ. Math. Inst. Hautes Études Sci.* **123**:1 (2016), 199–282. MR Zbl

[Aganagic et al. 2000] M. Aganagic, A. Karch, D. Lüst, and A. Miemiec, "Mirror symmetries for brane configurations and branes at singularities", *Nuclear Phys. B* **569**:1-3 (2000), 277–302. MR Zbl

[Altmann 1997] K. Altmann, "The versal deformation of an isolated toric Gorenstein singularity", *Invent. Math.* **128**:3 (1997), 443–479. MR Zbl

[Auroux 2007] D. Auroux, "Mirror symmetry and T-duality in the complement of an anticanonical divisor", *J. Gökova Geom. Topol.* **1** (2007), 51–91. MR Zbl

[Castaño-Bernard and Matessi 2009] R. Castaño Bernard and D. Matessi, "Lagrangian 3-torus fibrations", *J. Differential Geom.* **81**:3 (2009), 483–573. MR Zbl

[Castaño-Bernard and Matessi 2014] R. Castaño-Bernard and D. Matessi, "Conifold transitions via affine geometry and mirror symmetry", *Geom. Topol.* **18**:3 (2014), 1769–1863. MR Zbl

[Chan et al. 2012] K. Chan, S.-C. Lau, and N. C. Leung, "SYZ mirror symmetry for toric Calabi–Yau manifolds", *J. Differential Geom.* **90**:2 (2012), 177–250. MR Zbl

[Chan et al. 2016] K. Chan, C.-H. Cho, S.-C. Lau, and H.-H. Tseng, "Gross fibrations, SYZ mirror symmetry, and open Gromov–Witten invariants for toric Calabi–Yau orbifolds", *J. Differential Geom.* **103**:2 (2016), 207–288. MR Zbl

[Davies 2011] R. Davies, "Hyperconifold transitions, mirror symmetry, and string theory", *Nuclear Phys. B* **850**:1 (2011), 214–231. MR Zbl

[Feng et al. 2008] B. Feng, Y.-H. He, K. D. Kennaway, and C. Vafa, "Dimer models from mirror symmetry and quivering amoebae", *Adv. Theor. Math. Phys.* **12**:3 (2008), 489–545. MR Zbl

[Friedman 1986] R. Friedman, "Simultaneous resolution of threefold double points", *Math. Ann.* **274**:4 (1986), 671–689. MR Zbl

[Fukaya et al. 2009] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, *Lagrangian intersection Floer theory: anomaly and obstruction* (2 vols.), AMS/IP Stud. Adv. Math. **46**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Goldstein 2001] E. Goldstein, "Calibrated fibrations on noncompact manifolds via group actions", *Duke Math. J.* **110**:2 (2001), 309–343. MR Zbl

[Gross 2001] M. Gross, "Examples of special Lagrangian fibrations", pp. 81–109 in *Symplectic geometry and mirror symmetry* (Seoul, 2000), edited by K. Fukaya et al., World Sci., River Edge, NJ, 2001. MR Zbl

[Gross and Siebert 2011] M. Gross and B. Siebert, "From real affine geometry to complex geometry", *Ann. of Math.* (2) **174**:3 (2011), 1301–1428. MR Zbl

[Hori et al. 2000] K. Hori, A. Iqbal, and C. Vafa, "D-branes and mirror symmetry", preprint, 2000. arXiv

[Kaibel and Ziegler 2003] V. Kaibel and G. M. Ziegler, "Counting lattice triangulations", pp. 277–307 in *Surveys in combinatorics* (Bangor, Wales, 2003), edited by C. D. Wensley, London Math. Soc. Lecture Note Ser. **307**, Cambridge Univ. Press, 2003. MR Zbl arXiv

[Katz et al. 1997] S. Katz, A. Klemm, and C. Vafa, "Geometric engineering of quantum field theories", *Nuclear Phys. B* **497**:1-2 (1997), 173–195. MR Zbl

[Lau 2014] S.-C. Lau, "Open Gromov–Witten invariants and SYZ under local conifold transitions", *J. Lond. Math. Soc.* (2) **90**:2 (2014), 413–435. MR Zbl

[Lau 2018] S. C. Lau, "Affine elliptic surfaces with type-A singularities and orbi-conifolds", preprint, 2018. arXiv

[Lau et al. 2012] S.-C. Lau, N. C. Leung, and B. Wu, "Mirror maps equal SYZ maps for toric Calabi–Yau surfaces", *Bull. Lond. Math. Soc.* **44**:2 (2012), 255–270. MR Zbl

[Leung 2005] N. C. Leung, "Mirror symmetry without corrections", *Comm. Anal. Geom.* **13**:2 (2005), 287–331. MR Zbl

[Miemiec 2000] A. Miemiec, "Branes between geometry and gauge theory", *Fortschr. Phys.* **48**:12 (2000), 1143–1227. MR Zbl

[Morrison 1999] D. R. Morrison, "Through the looking glass", pp. 263–277 in *Mirror symmetry, III* (Montreal, 1995), edited by D. H. Phong et al., AMS/IP Stud. Adv. Math. **10**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl arXiv

[Rossi 2006] M. Rossi, "Geometric transitions", *J. Geom. Phys.* **56**:9 (2006), 1940–1983. MR Zbl

[Schoen 1988] C. Schoen, "On fiber products of rational elliptic surfaces with section", *Math. Z.* **197**:2 (1988), 177–199. MR Zbl

[Smith et al. 2002] I. Smith, R. P. Thomas, and S.-T. Yau, "Symplectic conifold transitions", *J. Differential Geom.* **62**:2 (2002), 209–242. MR Zbl

[Strominger et al. 1996] A. Strominger, S.-T. Yau, and E. Zaslow, "Mirror symmetry is T-duality", *Nuclear Phys. B* **479**:1-2 (1996), 243–259. MR Zbl

[Szendrői 2004] B. Szendrői, "Contractions and monodromy in homological mirror symmetry", pp. 301–314 in *Strings and geometry* (Cambridge, 2002), edited by M. Douglas et al., Clay Math. Proc. **3**, Amer. Math. Soc., Providence, RI, 2004. MR

[Tian 1992] G. Tian, "Smoothing 3-folds with trivial canonical bundle and ordinary double points", pp. 458–479 in *Essays on mirror manifolds*, edited by S.-T. Yau, Internat. Press, 1992. MR Zbl

Received March 12, 2018. Revised November 1, 2018.

ATSUSHI KANAZAWA
DEPARTMENT OF MATHEMATICS
KYOTO UNIVERSITY
JAPAN
akanazawa@math.kyoto-u.ac.jp

SIU-CHEONG LAU
DEPARTMENT OF MATHEMATICS AND STATISTICS
BOSTON UNIVERSITY
UNITED STATES
lau@math.bu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.519

SELF-DUAL EINSTEIN ACH METRICS AND CR GJMS OPERATORS IN DIMENSION THREE

TAIJI MARUGAME

By refining Matsumoto's construction of Einstein ACH metrics, we construct a one-parameter family of ACH metrics which solve the Einstein equation to infinite order and have a given three-dimensional CR structure at infinity. When the parameter is 0, the metric is self-dual to infinite order. As an application, we give another proof of the fact that three-dimensional CR manifolds admit CR invariant powers of the sublaplacian (CR GJMS operators) of all orders, which has been proved by Gover and Graham. We also prove the convergence of the formal solutions when the CR structure is real analytic.

1. Introduction

The GJMS operator $\mathcal{P}_{2k}$ on a conformal manifold of dimension N is an invariant linear differential operator acting on conformal densities of weight $k - N/2$ whose principal part is the power Δ^k of the laplacian [Graham et al. 1992]. It plays an important role in geometric analysis on conformal manifolds, and is also related to a fundamental curvature quantity, called the Q-curvature, whose integral gives a global conformal invariant [Fefferman and Graham 2002; Fefferman and Hirachi 2003; Graham and Zworski 2003]. The GJMS operator is constructed via the (Fefferman–Graham) ambient metric [2012] or equivalently via the Poincaré metric whose boundary at infinity is the given conformal manifold [Fefferman and Graham 2002; Graham and Zworski 2003]. The ambient metric is a formal solution to the Ricci flat equation, which corresponds to the Einstein equation for the Poincaré metric. When the dimension N is odd, the equation can be solved to infinite order and $\mathcal{P}_{2k}$ is defined for all $k \geq 1$. On the other hand, when N is even, an obstruction to the existence of a formal solution appears, and $\mathcal{P}_{2k}$ can only be defined for $1 \leq k \leq N/2$ due to the ambiguity of the ambient metric at higher orders. Moreover, it is known that this result of the existence of $\mathcal{P}_{2k}$ is sharp [Gover and Hirachi 2004].

MSC2010: primary 32V05; secondary 53A55.
Keywords: ACH metrics, the Einstein equation, self-duality, CR manifolds, CR invariant differential operators.

The CR counterpart of these operators are CR invariant powers of the sublaplacian

$$P_{2k} : \mathcal{E}\Big(\frac{k-n-1}{2}, \frac{k-n-1}{2}\Big) \longrightarrow \mathcal{E}\Big(\frac{-k-n-1}{2}, \frac{-k-n-1}{2}\Big)$$

on a $(2n+1)$-dimensional CR manifold M, which are called the CR GJMS operators or the Gover–Graham operators [Gover and Graham 2005; Hislop et al. 2008]. Here, $\mathcal{E}(w, w')$ is a complex line bundle over M called the CR density of weight (w, w'); see Section 2A for the definition. One can associate a conformal structure to a circle bundle over M, called the Fefferman conformal structure [1976], and apply the GJMS construction to produce P_{2k} for $1 \le k \le n+1$. Gover and Graham [2005] gave more operators by using techniques of CR tractor calculus; they proved that for each (w, w') such that $k = w + w' + n + 1 \in \mathbb{N}_+$ and $(w, w') \notin \mathbb{N} \times \mathbb{N}$, there exists a CR invariant linear differential operator $P_{w,w'} : \mathcal{E}(w, w') \to \mathcal{E}(w-k, w'-k)$ whose principal part is Δ_b^k. In cases where $w = w'$, these operators provide CR invariant modifications of Δ_b^k for all k with $k \equiv n \mod 2$. When $n = 1$, even more operators can be constructed: CR structure is a Cartan geometry modeled on the CR sphere $S^{2n+1} = \mathrm{SU}(n+1, 1)/P$, where P is the isotropy subgroup of a point in S^{2n+1}, and three-dimensional CR structure has a special feature from this viewpoint in that P is a Borel subgroup. Then the BGG machinery developed in [Čap et al. 2001] gives operators $P_{w,w'}$ for $(w, w') \in \mathbb{N} \times \mathbb{N}$ when $n = 1$. Thus one has:

Theorem 1.1 [Gover and Graham 2005, Theorem 1.3]. *Suppose M is a three-dimensional strictly pseudoconvex CR manifold. For each (w, w') such that $k = w + w' + 2 \in \mathbb{N}_+$, there exists a CR invariant linear differential operator $P_{w,w'} : \mathcal{E}(w, w') \to \mathcal{E}(w-k, w'-k)$ on M, whose principal part is Δ_b^k.*

In this paper, we provide a unified proof of Theorem 1.1 for cases in which $w = w'$. To this end, we construct an ACH (asymptotically complex hyperbolic) metric on a manifold with boundary M whose Taylor expansion along M is completely determined by local data of M. Our ACH metric is a refinement of the ACH Einstein metric which Matsumoto [2013; 2014] constructed for partially integrable CR manifolds. To state the results, let us recall some basic notions related to ACH metrics. Let M be a $(2n+1)$-dimensional strictly pseudoconvex partially integrable CR manifold. Namely, M has a contact distribution $H \subset TM$ together with an almost complex structure $J \in \mathrm{End}(H)$, and the eigenspace $T^{1,0}M \subset \mathbb{C}H$ with the eigenvalue i satisfies the partial integrability: $[\Gamma(T^{1,0}M), \Gamma(T^{1,0}M)] \subset \Gamma(\mathbb{C}H)$. A Θ-structure on a manifold $\overline{X}$ with boundary M is a conformal class $[\Theta]$ of sections $\Theta \in \Gamma(M, T^*\overline{X})$ such that $\Theta|_{TM}$ is a contact form on M. A diffeomorphism which preserves a Θ-structure is called a Θ-diffeomorphism. On the product $M \times [0, \infty)_\rho$, we define the standard Θ-structure by extending each contact form θ on M to Θ so that $\Theta(\partial/\partial\rho) = 0$. Fix a contact form θ on M and let $\{T, Z_\alpha\}$ be an admissible frame. We take the local frame $\{\boldsymbol{Z}_\infty = \rho\partial_\rho, \boldsymbol{Z}_0 = \rho^2 T, \boldsymbol{Z}_\alpha = \rho Z_\alpha, \boldsymbol{Z}_{\overline{\alpha}} = \rho Z_{\overline{\alpha}}\}$ and its dual coframe $\{\boldsymbol{\theta}^\infty, \boldsymbol{\theta}^0, \boldsymbol{\theta}^\alpha, \boldsymbol{\theta}^{\overline{\alpha}}\}$ on $M \times (0, \infty)_\rho$. Then for any ACH metric g on X, there exists a Θ-diffeomorphism $\Phi : M \times [0, \infty)_\rho \to \overline{X}$ which is defined near M and restricts to the identity on M, such that $\Phi^* g = g_{IJ}\boldsymbol{\theta}^I\boldsymbol{\theta}^J$ satisfies

$$\begin{aligned} g_{\infty\infty} &= 4, & g_{\infty 0} &= g_{\infty\alpha} = 0, & g_{00} &= 1 + O(\rho), \\ g_{0\alpha} &= O(\rho), & g_{\alpha\beta} &= O(\rho), & g_{\alpha\overline{\beta}} &= h_{\alpha\overline{\beta}} + O(\rho), \end{aligned}$$

where $h_{\alpha\overline{\beta}}$ is the Levi form on M. The CR manifold M is called the *CR structure at infinity* of g. Matsumoto [2013; 2014] proved that for any partially integrable CR manifold M, there exists an ACH metric g on $M \times [0, \infty)_\rho$ which satisfies

$$\begin{aligned} E_{IJ} &:= \mathrm{Ric}_{IJ} + \frac{n+2}{2} g_{IJ} = O(\rho^{2n+2}), \\ \mathrm{Scal} &= -(n+1)(n+2) + O(\rho^{2n+3}), \end{aligned}$$

where Ric is the Ricci tensor and Scal is the scalar curvature. Up to the pull-back by a Θ-diffeomorphism which fixes M, such a metric is unique modulo tensors which have $O(\rho^{2n+2})$ coefficients and $O(\rho^{2n+3})$ trace in the frame $\{\boldsymbol{Z}_I\}$. The order $O(\rho^{2n+2})$ in the above equation is optimal in general since $(\rho^{-2n-2}E_{\alpha\beta})|_M$ is independent of the choice of a solution g and defines a CR invariant tensor $\mathcal{O}_{\alpha\beta} \in \mathcal{E}_{\alpha\beta}(-n, -n)$, called the *CR obstruction tensor*. Matsumoto [2016] generalized the CR GJMS operators P_{2k} to the partially integrable case by using Dirichlet-to-Neumann type operators for the eigenvalue equations of the laplacian of g, but the order is again restricted to $1 \le k \le n+1$ due to the presence of the obstruction.

If we confine ourselves to the case where M is an integrable CR manifold, there is a possibility to refine the construction of ACH metrics. In fact, the CR obstruction tensor vanishes for integrable CR manifolds, in particular for three-dimensional CR manifolds since the CR structure is always integrable in this dimension. However, we need an additional normalization condition on the metric to ensure the uniqueness since the Einstein equation does not determine the $O(\rho^{2n+2})$-term of the metric. A possible normalization is the Kähler condition; Fefferman [1976] constructed an approximate solution to the complex Monge–Ampère equation on a strictly pseudoconvex domain Ω with boundary M and defined a Kähler metric which satisfies $E_{IJ} = O(\rho^{2n+4})$ as an ACH metric on the “square root” of Ω. However, this construction also has an obstruction $\mathcal{O} \in \mathcal{E}(-n-2, -n-2)$, called the *CR obstruction density*, and the metric is only determined modulo $O(\rho^{2n+4})$.

In this paper, we show that the self-dual equation $W^- = 0$ works as a better normalization when M is three-dimensional. The anti-self-dual part W^- of the Weyl curvature is connected to the Ricci tensor by the Bianchi identity

$$\nabla^I W^-_{IJKL} = C^-_{JKL}, \tag{1-1}$$

where C^-_{IJK} is the anti self-dual part of the Cotton tensor C_{IJK}, which is defined by

$C_{IJK} := \nabla_K P_{IJ} - \nabla_J P_{IK}$ with the Schouten tensor $P_{IJ} = \frac{1}{2}\operatorname{Ric}_{IJ} - \frac{1}{12}\operatorname{Scal} g_{IJ}$. It follows from (1-1) that the equation $E_{IJ} = O(\rho^4)$ implies $W^-_{IJKL} = O(\rho^4)$, and it turns out that the further normalization $W^-_{IJKL} = O(\rho^5)$ determines g_{IJ} modulo $O(\rho^5)$. We can then solve $E_{IJ} = O(\rho^6)$, which implies $W^-_{IJKL} = O(\rho^6)$. In the next step, besides the Einstein equation, we have freedom to prescribe the value of

$$\eta := \left(\rho^{-6} W^-_{\infty 0 \infty 0}\right)\big|_M .$$

If the Taylor coefficients of g_{IJ} along M have universal expressions in terms of pseudohermitian structure, η defines a CR invariant of weight $(-3, -3)$ (see Lemma 4.1). Thus, we should prescribe η to be a CR invariant in order to obtain a CR invariant normalization condition. It is known that a CR invariant in $\mathcal{E}(-3, -3)$ on a three-dimensional CR manifold is unique up to a constant multiple [Graham 1987], so there is no choice but to set $\eta = \lambda\mathcal{O}$ with a constant $\lambda \in \mathbb{R}$. After this step, the Einstein equation determines g_{IJ} to infinite order, and in the case $\lambda = 0$, the self-duality follows automatically from (1-1). Thus our main theorem reads as follows:

Theorem 1.2. *Let M be a three-dimensional strictly pseudoconvex CR manifold, and let $\lambda \in \mathbb{R}$. Then there exists an ACH metric g^λ_{IJ} on $M \times [0, \infty)_\rho$ which has M as the CR structure at infinity and satisfies*

$$\operatorname{Ric}_{IJ} + \tfrac{3}{2} g^\lambda_{IJ} = O(\rho^\infty), \quad W^-_{IJKL} = O(\rho^6), \quad \eta = \lambda\mathcal{O},$$

where η is the density defined by (4-1). *The metric g^λ_{IJ} is unique modulo $O(\rho^\infty)$ up to the pull-back by a Θ-diffeomorphism which fixes M. Moreover, g^0_{IJ} satisfies $W^-_{IJKL} = O(\rho^\infty)$.*

The Taylor coefficients of g^λ_{IJ} along the boundary have universal expressions in terms of the pseudohermitian structure for a fixed contact form.

By applying the construction of the CR GJMS operators via the ACH metric [Matsumoto 2016], we obtain the following theorem, which is a special case of Theorem 1.1:

Theorem 1.3. *Let M be a three-dimensional strictly pseudoconvex CR manifold, and let $\lambda \in \mathbb{R}$. Then, there exists a CR invariant linear differential operator*

$$P^\lambda_{2k} : \mathcal{E}(k/2 - 1,\ k/2 - 1) \to \mathcal{E}(-k/2 - 1,\ -k/2 - 1)$$

which is a polynomial in λ of degree $\leq k/3$, and has the principal part Δ_b^k.

Let us mention a similar construction in conformal geometry. Fefferman and Graham [2012] constructed a formal solution to the self-dual Einstein equation for the Poincaré metric with a given three-dimensional conformal manifold $\mathcal{M}$ as its conformal infinity. Thus our result is a CR analogue of their construction. When $\mathcal{M}$ is real analytic, LeBrun [1982] showed by twistor methods that there exists a real analytic self-dual Einstein metric on $\mathcal{M} \times (0, \epsilon)$ with the conformal infinity $\mathcal{M}$. The metric of Fefferman and Graham gives the Taylor expansion of LeBrun's metric. In the CR case, Biquard [2007] showed the existence of a self-dual Einstein ACH metric with a given real analytic CR three-manifold as its infinity by using twistor methods. Thus our formal solution g^0_{IJ} gives the Taylor expansion of Biquard's metric. In this paper, we prove the convergence of g^λ_{IJ} by applying the results of Baouendi and Goulaouic [1976] on singular nonlinear Cauchy problems.

Theorem 1.4. *Suppose M is a real analytic strictly pseudoconvex CR manifold of dimension three. Then the formal solution g^λ_{IJ} in Theorem 1.2 converges to a real analytic ACH metric near M.*

This paper is organized as follows. In Section 2, we review pseudohermitian geometry on a CR manifold and basic notions on ACH metrics. By following Matsumoto [2013], we describe the Levi-Civita connection of an ACH metric in terms of the extended Tanaka–Webster connection. In Section 3, we clarify the relationship between the Einstein equation and the self-dual equation, and compute the variation of curvature quantities under a perturbation of the metric. Section 4 is devoted to the proof of Theorem 1.2; we construct a one-parameter family of formal solutions to the Einstein equation and examine the dependence on the parameter. Then, in Section 5 we use these metrics to construct the CR GJMS operators and prove Theorem 1.3. Finally, in Section 6 we show the convergence of the formal solutions in the case where M is a real analytic CR manifold.

2. CR structure and ACH metric

2A. ***Pseudohermitian geometry.*** Let M be a $(2n+1)$-dimensional C^∞ manifold. A pair (H, J) is called a *CR structure* on M if H is a rank $2n$ subbundle of TM and J is an almost complex structure on H which satisfies the (formal) integrability condition

$$[\Gamma(T^{1,0}M), \Gamma(T^{1,0}M)] \subset \Gamma(T^{1,0}M),$$

where $T^{1,0}M \subset \mathbb{C}H$ is the eigenspace of J with the eigenvalue i. We note that the integrability condition automatically holds when M is three-dimensional. For any real 1-form θ such that $\operatorname{Ker}\theta = H$, we define the Levi form h_θ by

$$h_\theta(Z, \overline{W}) = -\,\mathrm{i}\,d\theta(Z, \overline{W})$$

for $Z, W \in T^{1,0}M$. We say the CR structure is *strictly pseudoconvex* if h_θ is positive definite for some θ. Since $h_{f\theta} = f h_\theta$ for any function f, such θ is determined up to a multiple by a positive function. When M is strictly pseudoconvex, H defines a contact structure, so we call θ a contact form. The *Reeb vector field* is the real vector field T uniquely determined by the conditions

$$\theta(T)=1, \quad T \lrcorner d\theta = 0.$$

Let $\{Z_\alpha\}$ be a local frame for $T^{1,0}M$. If we put $Z_{\overline{\alpha}} := \bar{Z}_\alpha$, then $\{T, Z_\alpha, Z_{\overline{\alpha}}\}$ gives a local frame for $\mathbb{C}TM$, which we call an *admissible frame*. The dual coframe $\{\theta, \theta^\alpha, \theta^{\overline{\alpha}}\}$ is called an *admissible coframe* and satisfies

$$d\theta = i h_{\alpha\overline{\beta}} \theta^\alpha \wedge \theta^{\overline{\beta}},$$

where $h_{\alpha\overline{\beta}} = h_\theta(Z_\alpha, Z_{\overline{\beta}})$.

The *CR canonical bundle* is defined by $K_M := \wedge^{n+1}(T^{0,1}M)^\perp \subset \wedge^{n+1}\mathbb{C}T^*M$, where $T^{0,1}M := \overline{T^{1,0}M}$. When K_M^{-1} admits an $(n+2)$-nd root $\mathcal{E}(1,0)$, the *CR density bundle* is defined by

$$\mathcal{E}(w, w') = \mathcal{E}(1,0)^{\otimes w} \otimes \overline{\mathcal{E}(1,0)}^{\otimes w'} \tag{2-1}$$

for each $(w, w') \in \mathbb{C}^2$ with $w - w' \in \mathbb{Z}$. In this paper, we restrict ourselves to the cases $w = w'$. In these cases, the definition (2-1) is independent of the choice of $\mathcal{E}(1,0)$ so we can define $\mathcal{E}(w,w)$ without assuming the global existence of $\mathcal{E}(1,0)$. We also denote the space of sections of these bundles by the same symbols, and call them *CR densities*.

For any contact form θ, there exists a local nonvanishing section ζ of K_M, unique up to a multiple of a $U(1)$-valued function, which satisfies

$$\theta \wedge (d\theta)^n = i^{n^2} n! \theta \wedge (T \lrcorner \zeta) \wedge (T \lrcorner \overline{\zeta}).$$

Then, the weighted contact form $\boldsymbol{\theta} := \theta \otimes |\zeta|^{-2/(n+2)} \in \Gamma(T^*M \otimes \mathcal{E}(1,1))$ is defined globally and independent of the choice of θ. Thus, there is a one-to-one correspondence between the set of contact forms and the set of positive sections $\tau \in \mathcal{E}(1,1)$, called *CR scales*. We define the CR invariant weighted Levi form $\boldsymbol{h}_{\alpha\overline{\beta}} := \tau h_{\alpha\overline{\beta}}$ by putting a weight to h_θ with the CR scale τ corresponding to θ. We raise and lower the indices of tensors on $\mathbb{C}H$ by $\boldsymbol{h}_{\alpha\overline{\beta}}$ and its inverse $\boldsymbol{h}^{\alpha\overline{\beta}}$, which has weight $(-1,-1)$.

For a fixed contact form θ, we can define a canonical linear connection ∇ on TM, called the *Tanaka–Webster connection*. It preserves $T^{1,0}M$ and satisfies $\nabla T = 0$, $\nabla h_\theta = 0$. In an admissible frame $\{T, Z_\alpha, Z_{\overline{\alpha}}\}$, the connection 1-forms $\omega_\beta{}^\alpha$ satisfy the structure equation

$$d\theta^\alpha = \theta^\beta \wedge \omega_\beta{}^\alpha + A^\alpha{}_{\overline{\beta}} \boldsymbol{\theta} \wedge \theta^{\overline{\beta}}.$$

The tensor $A_{\alpha\beta} := \bar{A}_{\overline{\alpha}\overline{\beta}}$ satisfies $A_{\alpha\beta} = A_{\beta\alpha}$ and is called the *Tanaka–Webster torsion tensor*. We use the index 0 for the direction of T, and we denote the components of covariant derivatives of a tensor by indices preceded by a comma, e.g., $A_{\alpha\gamma,\overline{\beta}} = \nabla_{\overline{\beta}} A_{\alpha\gamma}$. We omit the comma for covariant derivatives of a function. The curvature form $\Omega_\alpha{}^\beta = d\omega_\alpha{}^\beta - \omega_\alpha{}^\gamma \wedge \omega_\gamma{}^\beta$ is given by

$$\begin{aligned}(2\text{-}2)\quad \Omega_\alpha{}^\beta = R_\alpha{}^\beta{}_{\gamma\overline{\mu}}\theta^\gamma \wedge \theta^{\overline{\mu}} + A_{\alpha\gamma,}{}^\beta\theta^\gamma \wedge \boldsymbol{\theta} - A^\beta{}_{\overline{\gamma},\alpha}\theta^{\overline{\gamma}} \wedge \boldsymbol{\theta}\\ -iA_{\alpha\gamma}\theta^\gamma \wedge \theta^\beta + i\boldsymbol{h}_{\alpha\overline{\gamma}}A^\beta{}_{\overline{\mu}}\theta^{\overline{\gamma}} \wedge \theta^{\overline{\mu}}.\end{aligned}$$

The tensor $R_\alpha{}^\beta{}_{\gamma\overline{\mu}}$ is called the *Tanaka–Webster curvature tensor*. Taking traces with the weighted Levi form, we define the Tanaka–Webster Ricci tensor $\mathrm{Ric}_{\alpha\overline{\beta}} := R_\gamma{}^\gamma{}_{\alpha\overline{\beta}}$ and the Tanaka–Webster scalar curvature $\mathrm{Scal} := \mathrm{Ric}_\alpha{}^\alpha$. The *sublaplacian* is the differential operator $\Delta_b : \mathcal{E}(w, w') \to \mathcal{E}(w-1, w'-1)$ defined by

$$\Delta_b f = -\boldsymbol{h}^{\alpha\overline{\beta}}(\nabla_\alpha\nabla_{\overline{\beta}} + \nabla_{\overline{\beta}}\nabla_\alpha)f.$$

If we rescale the contact form as $\widehat{\theta} = e^{\Upsilon}\theta$, the Tanaka–Webster connection and its curvature quantities satisfy transformation formulas involving the derivatives of the scaling factor Υ; see, e.g., [Lee 1988]. We note that in dimension three the rank of $T^{1,0}M$ is 1 and the curvature form (2-2) is reduced to

$$\Omega_1{}^1 = \mathrm{Scal}\ \boldsymbol{h}_{1\overline{1}}\theta^1 \wedge \theta^{\overline{1}} + A_{11,}{}^1\theta^1 \wedge \boldsymbol{\theta} - A^1{}_{\overline{1},1}\theta^{\overline{1}} \wedge \boldsymbol{\theta}.$$

Also, in this dimension, M is locally CR diffeomorphic to the standard sphere S^3 if and only if the *Cartan tensor*

$$Q_{11} := \frac{1}{6}\,\mathrm{Scal}_{11} + \frac{i}{2}\,\mathrm{Scal}\ A_{11} - A_{11,0} - \frac{2i}{3}A_{11,\overline{1}}{}^{\overline{1}}$$

vanishes identically. The Cartan tensor is a CR invariant tensor of weight $(-1, -1)$. We also have a CR invariant density defined by

$$(2\text{-}3)\qquad \mathcal{O} := (\nabla^1\nabla^1 - iA^{11})\, Q_{11} \in \mathcal{E}(-3, -3),$$

called the *obstruction density*. It follows from the Bianchi identity for the Cartan tensor that $\mathcal{O}$ is a real density [Cheng and Lee 1990]. There is also a CR invariant density, called the obstruction density, on higher-dimensional CR manifolds and it appears as the logarithmic coefficient in the asymptotic expansion of the solution to the complex Monge–Ampère equation on a strictly pseudoconvex domain [Lee and Melrose 1982]. In dimension three, a CR invariant of weight $(-3, -3)$ is unique up to a constant multiple [Graham 1987], so it is necessarily a multiple of $\mathcal{O}$.

2B. ***ACH metrics.*** The ACH metric was introduced by Epstein, Melrose and Mendoza [Epstein et al. 1991] as a generalization of the complex hyperbolic metric on the ball. In this paper, we define it by using the characterization via the normal form.

Let X be the interior of a $(2n+2)$-dimensional C^∞ manifold whose boundary M is equipped with a strictly pseudoconvex CR structure (H, J). A conformal class $[\Theta]$

in $\Gamma(M, T^*\overline{X})$ is called a Θ-*structure* if $\Theta|_{TM}$ gives a contact form on M for each $\Theta \in [\Theta]$. We call $(\overline{X}, [\Theta])$ a Θ-*manifold*. Let $(\overline{X}', [\Theta'])$ be another Θ-manifold with the same boundary M. Then, a diffeomorphism Φ from a neighborhood of M in $\overline{X}$ to a neighborhood of M in $\overline{X}'$ is called a Θ-*diffeomorphism* if it fixes M and satisfies $[\Phi^*\Theta'] = [\Theta]$. We take a boundary-defining function $\rho \in C^\infty(\overline{X})$ which is positive on X. A vector field V on $\overline{X}$ is called a Θ-*vector field* if it satisfies

$$V|_M = 0, \quad \widetilde{\Theta}(V) = O(\rho^2),$$

where $\widetilde{\Theta}$ is an arbitrary extension of a $\Theta \in [\Theta]$. Note that the definition is independent of the choice of Θ and $\widetilde{\Theta}$. We extend $\{d\rho, \widetilde{\Theta}\}$ to a local coframe $\{d\rho, \widetilde{\Theta}, \alpha^1, \dots, \alpha^{2n}\}$ for $T^*\overline{X}$ near M. Let $\{N, T, Y_1, \dots, Y_{2n}\}$ be the dual frame. Then, any Θ-vector field V can be written as

$$V = V^\infty(\rho N) + V^0(\rho^2 T) + V^i(\rho Y_i), \quad V^\infty, V^0, V^i \in C^\infty(\overline{X}).$$

If we take another local coframe $\{d\rho', \widetilde{\Theta}', \alpha'^i\}$ and its dual $\{N', T', Y_i'\}$, then the transition function between $\{\rho N, \rho^2 T, \rho Y_i\}$ and $\{\rho' N', \rho'^2 T', \rho' Y_i'\}$ is smooth and nondegenerate up to M, so there exists a vector bundle ${}^\Theta T\overline{X}$ over $\overline{X}$ for which $\{\rho N, \rho^2 T, \rho Y_i\}$ gives a local frame. A Θ-vector field is identified with a section of this bundle and we call ${}^\Theta T\overline{X}$ the Θ-*tangent bundle*. A fiber metric on ${}^\Theta T\overline{X}$ is called a Θ-*metric*. Since the restriction ${}^\Theta T\overline{X}|_X$ is canonically isomorphic to TX, a Θ-metric defines a Riemannian metric on X. A local frame $\{\boldsymbol{Z}_I\}$ of ${}^\Theta T\overline{X}$ is called a Θ-*frame*. We also consider the dual ${}^\Theta T^*\overline{X}$ of the Θ-tangent bundle and various tensor bundles, whose sections are called Θ-*tensors*. A Θ-tensor is said to be $O(\rho^m)$ if each component in a Θ-frame is $O(\rho^m)$. The Θ-vector fields are closed under the Lie bracket, and those which vanish at a fixed point $p \in M$ form an ideal. Thus the fiber ${}^\Theta T_p\overline{X}$ becomes a Lie algebra, which we call the *tangent algebra*.

The product $M \times [0, \infty)_\rho$ has a canonical Θ-structure, called the *standard* Θ-*structure*, which is defined by extending each contact form θ on M to $\Theta \in \Gamma(M, T^*\overline{X})$ with $\Theta(\partial/\partial\rho) = 0$. Let θ be a contact form and $\{T, Z_\alpha, Z_{\overline{\alpha}}\}$ an admissible frame for $\mathbb{C}TM$. We extend $\{T, Z_\alpha, Z_{\overline{\alpha}}\}$ to $M \times [0, \infty)_\rho$ in the trivial way, and define a (complexified) Θ-frame $\{\boldsymbol{Z}_I\}$ by

$$\boldsymbol{Z}_\infty = \rho\partial_\rho, \quad \boldsymbol{Z}_0 = \rho^2 T, \quad \boldsymbol{Z}_\alpha = \rho Z_\alpha, \quad \boldsymbol{Z}_{\overline{\alpha}} = \rho Z_{\overline{\alpha}},$$

where $\partial_\rho = \partial/\partial\rho$. A Θ-metric g on $M \times [0, \infty)_\rho$ is called a *normal form ACH metric* if the components $g_{IJ} = g(\boldsymbol{Z}_I, \boldsymbol{Z}_J)$ satisfy

$$\begin{aligned} &g_{\infty\infty} = 4, && g_{\infty 0} = g_{\infty\alpha} = 0, && g_{00} = 1 + O(\rho), \\ &g_{0\alpha} = O(\rho), && g_{\alpha\beta} = O(\rho), && g_{\alpha\overline{\beta}} = h_{\alpha\overline{\beta}} + O(\rho), \end{aligned} \tag{2-4}$$

where $h_{\alpha\overline{\beta}} = h_\theta(Z_\alpha, Z_{\overline{\beta}})$. On a general Θ-manifold $(\overline{X}, [\Theta])$, the ACH metric is defined as follows:

Definition 2.1. A Θ-metric g on $\overline{X}$ is called an ACH metric if for any contact form θ on M, there exist a neighborhood $U \subset \overline{X}$ of M and a Θ-diffeomorphism $\Phi_\theta : M \times [0, \infty)_\rho \to U$ such that $\Phi_\theta^* g$ is a normal form ACH metric.

We remark that there is an alternative definition of the ACH metric which involves only the boundary value of g; see [Matsumoto 2013, Definition 4.6].

The germ of Φ_θ along M is unique, and we call $\rho \circ \Phi_\theta^{-1}$ the *model defining function* for θ. We identify a neighborhood of M in $\overline{X}$ with $M \times [0, \epsilon)_\rho$ through Φ_θ and regard $\{\boldsymbol{Z}_I\}$ as a Θ-frame on $\overline{X}$. The following proposition will be used in the proof of Lemma 4.1.

Proposition 2.2. *The boundary values $\mathbf{Z}_\infty|_M, \mathbf{Z}_0|_M$ are independent of θ and determined only by the ACH metric g.*

Proof. By strict pseudoconvexity of (H, J), the derived Lie algebras of the tangent algebra ${}^\Theta T_p\overline{X}$ at a point $p \in M$ are given by

$$\begin{aligned}\mathcal{D}^1 &:= [{}^\Theta T_p\overline{X}, {}^\Theta T_p\overline{X}] = \operatorname{span}\{(\mathbf{Z}_0)_p, (\mathbf{Z}_\alpha)_p, (\mathbf{Z}_{\overline{\alpha}})_p\},\\ \mathcal{D}^2 &:= [\mathcal{D}^1, \mathcal{D}^1] = \operatorname{span}\{(\mathbf{Z}_0)_p\}.\end{aligned}$$

Thus, $(\mathbf{Z}_\infty)_p$ and $(\mathbf{Z}_0)_p$ are oriented bases of $(\mathcal{D}^1)^\perp$ and $\mathcal{D}^2$ respectively. Since they are normalized by $|(\mathbf{Z}_\infty)_p|_g^2 = 4$ and $|(\mathbf{Z}_0)_p|_g^2 = 1$, they are independent of θ. $\square$

Let $\theta, \widehat{\theta} = e^\Upsilon\theta$ be contact forms on M and ρ and $\widehat{\rho}$ be the corresponding model-defining functions. Then there exists a positive function f on $\overline{X}$ such that $\widehat{\rho} = f\rho$. Since the Reeb vector fields are related as $\widehat{T} = e^{-\Upsilon}(T - ih^{\alpha\overline{\gamma}}\Upsilon_{\overline{\gamma}}Z_\alpha + ih^{\gamma\overline{\alpha}}\Upsilon_\gamma Z_{\overline{\alpha}})$, we have

$$\widehat{\mathbf{Z}}_0 = \widehat{\rho}^2\widehat{T} = e^{-\Upsilon} f^2 \mathbf{Z}_0 + O(\rho)$$

as a Θ-vector field, where we regard Υ as a function on a neighborhood of M. It follows from $\widehat{\mathbf{Z}}_0|_M = \mathbf{Z}_0|_M$ that $f|_M = e^{\Upsilon/2}$. Thus we have

$$\widehat{\rho} = e^{\Upsilon/2}\rho + O(\rho^2). \tag{2-5}$$

In particular, a contact form is recovered from the 1-jet of the corresponding model-defining function along the boundary.

2C. *The Levi-Civita connection.* Let g be an ACH metric on a Θ-manifold $(\overline{X}, [\Theta])$ with boundary M. Here and after, we assume that M is three-dimensional. We lower and raise the indices of Θ-tensors by g_{IJ} and its inverse g^{IJ}. In order to describe the Levi-Civita connection of g, we introduce an extension of the Tanaka–Webster connection by following Matsumoto. We refer the reader to [Matsumoto 2013, §6.2] or [Matsumoto 2014, §4] for a more detailed exposition.

Let θ be a contact form on M. We identify a neighborhood of M in $\overline{X}$ with $M \times [0, \epsilon)_\rho$ by the Θ-diffeomorphism determined by θ. We take an admissible frame $\{T, Z_1, Z_{\overline{1}}\}$ and define the *extended Tanaka–Webster connection* $\overline{\nabla}$ on $T\overline{X}$ by

$$\overline{\nabla}\partial_\rho = 0, \qquad \overline{\nabla}_{\partial_\rho} T = \overline{\nabla}_{\partial_\rho} Z_1 = 0,$$
$$\overline{\nabla}_T Z_1 = \nabla^{\mathrm{TW}}_T Z_1, \qquad \overline{\nabla}_{Z_1} Z_1 = \nabla^{\mathrm{TW}}_{Z_1} Z_1, \qquad \overline{\nabla}_{Z_{\overline{1}}} Z_1 = \nabla^{\mathrm{TW}}_{Z_{\overline{1}}} Z_1,$$

where ∇^{TW} denotes the Tanaka–Webster connection associated with θ. Then, $\overline{\nabla}$ is a Θ-*connection* in the sense that if V, W are Θ-vector fields, so is the covariant derivative $\overline{\nabla}_V W$. We take the Θ-frame $\{\boldsymbol{Z}_I\} = \{\rho\partial_\rho, \rho^2 T, \rho Z_1, \rho Z_{\overline{1}}\}$ and define the Christoffel symbols $\overline{\Gamma}_{IJ}{}^K$ by $\overline{\nabla}_{\boldsymbol{Z}_I} \boldsymbol{Z}_J = \overline{\Gamma}_{IJ}{}^K \boldsymbol{Z}_K$. A simple calculation shows that

$$\begin{aligned} &\overline{\Gamma}_{\infty\infty}{}^{\infty} = 1, && \overline{\Gamma}_{\infty 0}{}^{0} = 2, && \overline{\Gamma}_{\infty 1}{}^{1} = 1, \\ &\overline{\Gamma}_{01}{}^{1} = \rho^2 \Gamma_{01}{}^{1}, && \overline{\Gamma}_{11}{}^{1} = \rho \Gamma_{11}{}^{1}, && \overline{\Gamma}_{\overline{1}1}{}^{1} = \rho \Gamma_{\overline{1}1}{}^{1}, \end{aligned} \tag{2-6}$$

where $\Gamma_{ij}{}^k$ are the Christoffel symbols of ∇^{TW} with respect to $\{T, Z_1, Z_{\overline{1}}\}$; the components which cannot be obtained by taking complex conjugates of (2-6) are 0. It follows from (2-6) that the components of the covariant derivative of a Θ-tensor $S_{I_1\cdots I_p}{}^{J_1\cdots J_q}$ are computed as

$$\begin{aligned} \overline{\nabla}_\infty S_{I_1\cdots I_p}{}^{J_1\cdots J_q} &= \big(\rho\partial_\rho - \#(I_1\cdots I_p) + \#(J_1\cdots J_q)\big) S_{I_1\cdots I_p}{}^{J_1\cdots J_q}, \\ \overline{\nabla}_0 S_{I_1\cdots I_p}{}^{J_1\cdots J_q} &= \rho^2 \nabla^{\mathrm{TW}}_0 S_{I_1\cdots I_p}{}^{J_1\cdots J_q}, \\ \overline{\nabla}_1 S_{I_1\cdots I_p}{}^{J_1\cdots J_q} &= \rho \nabla^{\mathrm{TW}}_1 S_{I_1\cdots I_p}{}^{J_1\cdots J_q}, \end{aligned} \tag{2-7}$$

where $\#(I_1\cdots I_p) := p +$ (the number of 0s) and we regard S as a tensor on $\mathbb{C}H$ when we apply ∇^{TW} to it [Matsumoto 2013, Lemma 6.2; 2014, (4.9)].

The torsion tensor $\overline{T}_{IJ}{}^K$ and the curvature tensor $\overline{R}_I{}^J{}_{KL}$ of $\overline{\nabla}$ are defined by

$$(\overline{\nabla}_V W - \overline{\nabla}_W V - [V, W])^K = \overline{T}_{IJ}{}^K V^I W^J,$$
$$(\overline{\nabla}_V \overline{\nabla}_W Y - \overline{\nabla}_W \overline{\nabla}_V Y - \overline{\nabla}_{[V,W]} Y)^J = \overline{R}_I{}^J{}_{KL} Y^I V^K W^L,$$

respectively. In the Θ-frame $\{\boldsymbol{Z}_I\}$, the components are given by

$$\overline{T}_{1\overline{1}}{}^{0} = i h_{1\overline{1}}, \quad \overline{T}_{01}{}^{\overline{1}} = \rho^2 A_1{}^{\overline{1}}, \tag{2-8}$$

and

$$\overline{R}_1{}^1{}_{1\overline{1}} = \rho^2\, \mathrm{Scal}^{\mathrm{TW}}\, h_{1\overline{1}}, \quad \overline{R}_1{}^1{}_{01} = -\rho^3 A_{11,}{}^{1}, \quad \overline{R}_1{}^1{}_{0\overline{1}} = \rho^3 A^1{}_{\overline{1},1}, \tag{2-9}$$

where $\mathrm{Scal}^{\mathrm{TW}}$ denotes the Tanaka–Webster scalar curvature, and we have removed the CR weights in the Tanaka–Webster tensors by the CR scale corresponding to θ. The components which cannot be obtained from (2-8) and (2-9) by the symmetries of $\overline{T}$ and $\overline{R}$, or by taking the complex conjugates, are all 0. The nonzero components of the Ricci tensor $\bar{R}_{IJ} = \bar{R}_I{}^K{}_{KJ}$ are given by

$$\bar{R}_{1\bar{1}} = \rho^2 \operatorname{Scal}^{\mathrm{TW}} h_{1\bar{1}}, \quad \bar{R}_{10} = \rho^2 A_{11,}{}^1.$$

Let ∇ be the Levi-Civita connection of g, which is also a Θ-connection ([Matsumoto 2013, Proposition 4.4]). We define the *difference* Θ*-tensor* $D_{IJ}{}^K$ by

$$\nabla_I V^K = \overline{\nabla}_I V^K + D_{IJ}{}^K V^J.$$

Since ∇ is torsion-free, we have

$$D_{IJ}{}^K = D_{JI}{}^K + \overline{T}_{JI}{}^K. \tag{2-10}$$

Using this relation and the fact $\nabla g = 0$, we obtain

$$2D_{IJK} = \overline{\nabla}_I g_{JK} + \overline{\nabla}_J g_{KI} - \overline{\nabla}_K g_{IJ} - \overline{T}_{IJK} + \overline{T}_{JKI} - \overline{T}_{KIJ}. \tag{2-11}$$

We will compute $D_{IJ}{}^K$ by these formulas. Since the components g_{IJ} satisfy (2-4), g is described by ρ-dependent tensors φ_{ij} on M defined by

$$g_{00} = 1 + \varphi_{00}, \quad g_{01} = \varphi_{01}, \quad g_{11} = \varphi_{11}, \quad g_{1\bar{1}} = h_{1\bar{1}} + \varphi_{1\bar{1}}.$$

In the construction of a formal solution to the self-dual Einstein equation, we need to examine the effect of a perturbation

$$\varphi_{ij} \mapsto \varphi_{ij} + \psi_{ij}, \quad \psi_{ij} = O(\rho^m) \tag{2-12}$$

on the curvature quantities of g. It is useful in the computation to ignore irrelevant terms on which the perturbation causes only changes in higher orders. Such terms are of the form

$$O(\rho) \cdot (\rho\partial_\rho)^l \mathcal{D}\varphi_{ij}, \tag{2-13}$$

where $\mathcal{D}$ is a ρ-dependent differential operator on M. These are called *negligible terms*. In fact, a negligible term changes by $O(\rho^{m+1})$ under the perturbation (2-12). Thus, it suffices to compute $D_{IJ}{}^K$ modulo negligible terms. For simplicity, we assume that the admissible frame $\{Z_1\}$ is unitary with respect to the Levi form; namely $h_{1\bar{1}} = 1$. Noting that $\varphi_{ij} = O(\rho)$, we have

$$g^{\infty\infty} = \tfrac{1}{4}, \quad g^{\infty 0} = g^{\infty 1} = 0, \quad g^{00} \equiv 1 - \varphi_{00}, \quad g^{1\bar{1}} \equiv -\varphi_{1\bar{1}}, \quad g^{11} \equiv -\varphi_{\bar{1}\bar{1}} \tag{2-14}$$

modulo negligible terms. By computing with (2-7), (2-8), (2-10), (2-11), (2-14) we obtain the following result:

Lemma 2.3 [Matsumoto 2013, Lemma 6.4; 2014, Table 1]. *Let $\{T, Z_1, Z_{\bar{1}}\}$ be a unitary admissible frame and $\{\mathbf{Z}_I\} = \{\rho\partial_\rho, \rho^2 T, \rho Z_1, \rho Z_{\bar{1}}\}$ the associated Θ-frame.*

Then, modulo negligible terms, the components $D_{IJ}{}^K$ are given by

$$
\begin{aligned}
&D_{\infty\infty}{}^{\infty} \equiv -1, && D_{\infty 0}{}^{\infty} \equiv D_{\infty 1}{}^{\infty} \equiv 0,\\
&D_{00}{}^{\infty} \equiv \tfrac12 - \tfrac18(\rho\partial_\rho - 4)\varphi_{00}, && D_{01}{}^{\infty} \equiv -\tfrac18(\rho\partial_\rho - 3)\varphi_{01},\\
&D_{1\bar 1}{}^{\infty} \equiv \tfrac14 - \tfrac18(\rho\partial_\rho - 2)\varphi_{1\bar 1}, && D_{11}{}^{\infty} \equiv -\tfrac18(\rho\partial_\rho - 2)\varphi_{11},
\end{aligned}
$$

$$
\begin{aligned}
&D_{\infty\infty}{}^{1} \equiv D_{00}{}^{1} \equiv D_{\bar 1\bar 1}{}^{1} \equiv 0, && D_{\infty\bar 1}{}^{1} \equiv \tfrac12\rho\partial_\rho\varphi_{\bar 1\bar 1}, && D_{0\bar 1}{}^{1} \equiv \tfrac{i}{2}\varphi_{\bar 1\bar 1},\\
&D_{\infty 0}{}^{1} \equiv \tfrac12(\rho\partial_\rho + 1)\varphi_{0\bar 1}, && D_{01}{}^{1} \equiv \tfrac{i}{2}(1 + \varphi_{00} - \varphi_{1\bar 1}), && D_{1\bar 1}{}^{1} \equiv \tfrac{i}{2}\varphi_{0\bar 1},\\
&D_{\infty 1}{}^{1} \equiv -1 + \tfrac12\rho\partial_\rho\varphi_{1\bar 1}, && D_{\bar 1 0}{}^{1} \equiv \tfrac{i}{2}\varphi_{\bar 1\bar 1} + \rho^2 A_{\bar 1\bar 1}, && D_{11}{}^{1} \equiv i\varphi_{01},\\
&D_{\infty\infty}{}^{0} \equiv D_{00}{}^{0} \equiv 0, && D_{1\bar 1}{}^{0} \equiv -\tfrac{i}{2}, && D_{11}{}^{0} \equiv -\rho^2 A_{11},\\
&D_{\infty 0}{}^{0} \equiv -2 + \tfrac12\rho\partial_\rho\varphi_{00}, && D_{\infty 1}{}^{0} \equiv \tfrac12(\rho\partial_\rho - 1)\varphi_{01}, && D_{01}{}^{0} \equiv -\tfrac{i}{2}\varphi_{01}.
\end{aligned}
$$

The components which are not displayed are obtained by taking the complex conjugates or using the relation (2-10).

Remark 2.4. We have modified a typographical error in [Matsumoto 2013, Table 6.2; 2014, Table 1]; the value of $D_{01}{}^1$ above differs by $-\frac{i}{2}\varphi_{1\bar 1}$ from that in [Matsumoto 2013; 2014]. (Note that $D_{IJ}{}^K$ is denoted by $D^K{}_{IJ}$ in [Matsumoto 2013] and by $D_J{}^K{}_I$ in [Matsumoto 2014].) The correct value is used in the other computations in [Matsumoto 2013; 2014].

3. The self-dual Einstein equation

Let g be an ACH metric on a four-dimensional Θ-manifold $(\overline{X}, [\Theta])$ which has a strictly pseudoconvex CR manifold M as its boundary. We fix a contact form θ on M and identify a neighborhood of M as $M \times [0, \epsilon)_\rho$, where ρ is the model-defining function for θ. We take a unitary admissible frame $\{T, Z_1, Z_{\bar 1}\}$ on M and work in the associated Θ-frame $\{\boldsymbol{Z}_I\} = \{\rho\partial_\rho, \rho^2 T, \rho Z_1, \rho Z_{\bar 1}\}$.

3A. ***The Einstein equation.*** We will recall from [Matsumoto 2013; 2014] the computation of the Einstein tensor modulo negligible terms which is needed in the construction of the Einstein ACH metric. We set

$$E_{IJ} := \mathrm{Ric}_{IJ} + \tfrac32 g_{IJ}.$$

In terms of the extended Tanaka–Webster connection and the difference Θ-tensor, the curvature tensor of g is expressed as

$$
\begin{aligned}
\text{(3-1)}\quad R_I{}^J{}_{KL} = \overline{R}_I{}^J{}_{KL} &+ \overline{\nabla}_K D_{LI}{}^J\\
&- \overline{\nabla}_L D_{KI}{}^J + D_{KM}{}^J D_{LI}{}^M - D_{LM}{}^J D_{KI}{}^M + \overline{T}_{KL}{}^M D_{MI}{}^J.
\end{aligned}
$$

Hence, the Ricci tensor is given by

$$
\begin{aligned}
\text{(3-2)}\quad \mathrm{Ric}_{IJ} &= R_J{}^K{}_{KI} \\
&= \bar{R}_{JI} + \overline{\nabla}_K D_{IJ}{}^K - \overline{\nabla}_I D_{KJ}{}^K \\
&\qquad\qquad + D_{KM}{}^K D_{IJ}{}^M - D_{IM}{}^K D_{KJ}{}^M + \bar{T}_{KI}{}^M D_{MJ}{}^K \\
&= \bar{R}_{JI} + \overline{\nabla}_K D_{IJ}{}^K - \overline{\nabla}_I D_{KJ}{}^K + D_{KM}{}^K D_{IJ}{}^M - D_{MI}{}^K D_{KJ}{}^M.
\end{aligned}
$$

In the last equality, we have used (2-10). With this formula and Lemma 2.3, we can compute E_{IJ} modulo negligible terms:

Lemma 3.1 [Matsumoto 2013, Lemma 6.5; 2014, Lemma 4.2]. *Let $\{T, Z_1, Z_{\bar{1}}\}$ be a unitary admissible frame and $\{\boldsymbol{Z}_I\} = \{\rho\partial_\rho, \rho^2 T, \rho Z_1, \rho Z_{\bar{1}}\}$ the associated Θ-frame. Then, modulo negligible terms, the components of the Einstein tensor E_{IJ} are given by*

$$
\begin{aligned}
E_{\infty\infty} &\equiv -\tfrac{1}{2}\rho\partial_\rho(\rho\partial_\rho - 4)\varphi_{00} - \rho\partial_\rho(\rho\partial_\rho - 2)\varphi_{1\bar{1}}, \\
E_{\infty 0} &\equiv 0, \\
E_{\infty 1} &\equiv -\tfrac{i}{2}(\rho\partial_\rho + 1)\varphi_{01}, \\
E_{00} &\equiv -2\rho^4|A|^2 - \tfrac{1}{8}\big((\rho\partial_\rho)^2 - 6\rho\partial_\rho - 4\big)\varphi_{00} + \tfrac{1}{2}(\rho\partial_\rho - 2)\varphi_{1\bar{1}}, \\
E_{01} &\equiv \rho^3 A_{11,}{}^1 - \tfrac{1}{8}(\rho\partial_\rho + 1)(\rho\partial_\rho - 5)\varphi_{01}, \\
E_{1\bar{1}} &\equiv \rho^2\,\mathrm{Scal}^{\mathrm{TW}} - \tfrac{1}{8}\big((\rho\partial_\rho)^2 - 6\rho\partial_\rho - 8\big)\varphi_{1\bar{1}} + \tfrac{1}{8}(\rho\partial_\rho - 4)\varphi_{00}, \\
E_{11} &\equiv i\rho^2 A_{11} - \rho^4 A_{11,0} - \tfrac{1}{8}\rho\partial_\rho(\rho\partial_\rho - 4)\varphi_{11}.
\end{aligned}
$$

The components which are not displayed are obtained by the symmetry or by taking the complex conjugates.

Remark 3.2. We have corrected the value of E_{00} in [Matsumoto 2013, Lemma 6.5; 2014, Lemma 4.2], where the term $-2\rho^4|A|^2$ is missed, though this modification has no significant effect on the construction of the Einstein ACH metric.

3B. *The self-dual equation.* Let $\{\boldsymbol{\theta}^I\}$ be the dual Θ-coframe of $\{\boldsymbol{Z}_I\}$. We take the orientation of $\overline{X}$ such that $\theta \wedge d\theta \wedge d\rho > 0$, and define a skew symmetric Θ-tensor ε_{IJKL} by

$$\mathrm{vol}_g = \frac{1}{4!}\varepsilon_{IJKL}\boldsymbol{\theta}^I \wedge \boldsymbol{\theta}^J \wedge \boldsymbol{\theta}^K \wedge \boldsymbol{\theta}^L,$$

where vol_g is the volume form of g. Since $\det(g_{IJ}) \equiv -4(1 + \varphi_{00} + 2\varphi_{1\bar{1}})$ modulo negligible terms, we have

$$
\begin{aligned}
\mathrm{vol}_g &= |\det(g_{IJ})|^{1/2}\, i\boldsymbol{\theta}^0 \wedge \boldsymbol{\theta}^1 \wedge \boldsymbol{\theta}^{\bar{1}} \wedge \boldsymbol{\theta}^\infty \\
&\equiv (2i + i\varphi_{00} + 2i\varphi_{1\bar{1}})\, \boldsymbol{\theta}^0 \wedge \boldsymbol{\theta}^1 \wedge \boldsymbol{\theta}^{\bar{1}} \wedge \boldsymbol{\theta}^\infty,
\end{aligned}
$$

and hence

$$\text{(3-3)}\qquad \varepsilon_{01\bar{1}\infty} \equiv 2i + i\varphi_{00} + 2i\varphi_{1\bar{1}}.$$

Let $P_{IJ} = \frac{1}{2}\operatorname{Ric}_{IJ} - \frac{1}{12}\operatorname{Scal} g_{IJ}$ be the Schouten tensor, and let

$$W_{IJKL} = R_{IJKL} + g_{IK}P_{JL} - g_{JK}P_{IL} + g_{JL}P_{IK} - g_{IL}P_{JK}$$

be the Weyl curvature. Since $\overline{X}$ is four-dimensional, we can define the anti-self-dual part of the Weyl curvature, which is given by

$$W^{-}_{IJKL} = \tfrac{1}{2}\big(W_{IJKL} - \tfrac{1}{2}\varepsilon_{KL}{}^{PQ}W_{IJPQ}\big).$$

Note that W^{-}_{IJKL} has the same symmetry as the Weyl curvature and satisfies

$$\tfrac{1}{2}\varepsilon_{KL}{}^{PQ}W^{-}_{IJPQ} = -W^{-}_{IJKL}.$$

Thus, by (2-14) and (3-3), we have

$$W^{-}_{\infty 0\infty 0} \equiv -W^{-}_{\infty 1\infty\bar{1}} - W^{-}_{\infty\bar{1}\infty 1} = -2W^{-}_{\infty 1\infty\bar{1}}, \tag{3-4}$$

$$W^{-}_{IJ01} = -\varepsilon_{01\bar{1}\infty}W^{-}_{IJ}{}^{\bar{1}\infty} \equiv -\frac{i}{2}W^{-}_{IJ1\infty}, \tag{3-5}$$

$$W^{-}_{IJ1\bar{1}} = -\varepsilon_{1\bar{1}0\infty}W^{-}_{IJ}{}^{0\infty} \equiv -\frac{i}{2}W^{-}_{IJ0\infty} \tag{3-6}$$

modulo $O(\rho)\cdot W^{-}_{IJKL}$. Since $W^{-}_{IJKL} = W^{-}_{KLIJ}$, we also have

$$W^{-}_{01KL} \equiv -\frac{i}{2}W^{-}_{1\infty KL},$$
$$W^{-}_{1\bar{1}KL} \equiv -\frac{i}{2}W^{-}_{0\infty KL}$$

modulo $O(\rho)\cdot W^{-}_{IJKL}$. As a consequence, we have the following lemma:

Lemma 3.3. *Let m be a positive integer. If $W^{-}_{\infty 1\infty 1}, W^{-}_{\infty 0\infty 1}, W^{-}_{\infty 0\infty 0} = O(\rho^m)$, then $W^{-}_{IJKL} = O(\rho^m)$.*

Thus, in order to solve the self-dual equation $W^{-}_{IJKL} = O(\rho^\infty)$, we only have to deal with the three components indicated above.

Next, we consider the Bianchi identity which relates the self-dual equation to the Einstein equation. Let $C_{IJK} := \nabla_K P_{IJ} - \nabla_J P_{IK}$ be the Cotton tensor of g and define the anti-self-dual part C^{-}_{IJK} by

$$C^{-}_{IJK} = \tfrac{1}{2}\big(C_{IJK} - \tfrac{1}{2}\varepsilon_{JK}{}^{PQ}C_{IPQ}\big).$$

Then, since $\nabla_I\varepsilon_{JKLM} = 0$, the Bianchi identity $\nabla^I W_{IJKL} = C_{JKL}$ yields

$$\nabla^I W^{-}_{IJKL} = C^{-}_{JKL}.$$

If g satisfies $E_{IJ} = O(\rho^m)$ for some $m \geq 1$, then we have $P_{IJ} = -\frac{1}{4}g_{IJ} + O(\rho^m)$ and hence $C^{-}_{IJK} = O(\rho^m)$ since the covariant differentiation does not decrease the vanishing order of a Θ-tensor. Therefore, it holds that

$$E_{IJ} = O(\rho^m) \Rightarrow \nabla^I W^{-}_{IJKL} = O(\rho^m).$$

To derive the consequence of the latter equation, we will compute

$$\nabla^I W^-_{IJKL} = \overline{\nabla}^I W^-_{IJKL} - W^-_{MJKL} D^I{}_I{}^M - W^-_{IMKL} D^I{}_J{}^M - W^-_{IJML} D^I{}_K{}^M - W^-_{IJKM} D^I{}_L{}^M \tag{3-7}$$

modulo $O(\rho)\cdot \mathcal{D}W^-_{IJKL}$, where $\mathcal{D}$ is a ρ-dependent differential operator on M. By computations similar to (3-4), (3-5), (3-6), we have

$$\nabla^I W^-_{I0\infty 0} \equiv -2\nabla^I W^-_{I1\infty\bar{1}}, \tag{3-8}$$

$$\nabla^I W^-_{IJ01} \equiv -\frac{i}{2}\nabla^I W^-_{IJ1\infty}, \tag{3-9}$$

$$\nabla^I W^-_{IJ1\bar{1}} \equiv -\frac{i}{2}\nabla^I W^-_{IJ0\infty} \tag{3-10}$$

modulo $O(\rho)\cdot\mathcal{D}W^-_{IJKL}$. By (3-9) and (3-10), it suffices to consider the cases where $K=\infty$. Then, taking complex conjugates we may assume that $L=0$ or 1, and the case $(J, K, L)=(\bar{1}, \infty, 0)$ is reduced to the case $(J, K, L)=(1, \infty, 0)$. Moreover, by (3-8) the case $(J, K, L)=(\bar{1}, \infty, 1)$ is reduced to the case $(J, K, L)=(0, \infty, 0)$. Thus, it suffices to compute (3-7) for

$$(J, K, L) = (1, \infty, 1), (0, \infty, 0), (0, \infty, 1), (1, \infty, 0), (\infty, \infty, 1), (\infty, \infty, 0).$$

By (2-7), we have

$$\begin{aligned}\overline{\nabla}^I W^-_{IJKL} &= \overline{\nabla}^\infty W^-_{\infty JKL} + \overline{\nabla}^1 W^-_{1JKL} + \overline{\nabla}^{\bar{1}} W^-_{\bar{1}JKL} + \overline{\nabla}^0 W^-_{0JKL}\\ &\equiv \tfrac{1}{4}\big(\rho\partial_\rho - \#(\infty JKL)\big) W^-_{\infty JKL}.\end{aligned}$$

The other terms in the right-hand side of (3-7) can be computed by Lemma 2.3. The final results are:

$$\begin{aligned}&\nabla^I W^-_{I1\infty 1} \equiv \frac{1}{4}(\rho\partial_\rho - 4) W^-_{\infty 1\infty 1}, &&\nabla^I W^-_{I0\infty 0} \equiv \frac{1}{4}(\rho\partial_\rho - 6) W^-_{\infty 0\infty 0},\\ &\nabla^I W^-_{I0\infty 1} \equiv \frac{1}{4}(\rho\partial_\rho - 6) W^-_{\infty 0\infty 1}, &&\nabla^I W^-_{I1\infty 0} \equiv \frac{1}{4}(\rho\partial_\rho - 5) W^-_{\infty 1\infty 0},\\ &\nabla^I W^-_{I\infty\infty 1} \equiv \frac{i}{2} W^-_{\infty 0\infty 1}, &&\nabla^I W^-_{I\infty\infty 0} \equiv 0.\end{aligned} \tag{3-11}$$

Consequently, by an inductive argument, we have the following implication:

$$E_{IJ} = O(\rho^4) \Rightarrow W^-_{IJKL} = O(\rho^4).$$

Moreover, if $E_{IJ} = O(\rho^5)$ then $W^-_{\infty 0\infty 0}$ and $W^-_{\infty 0\infty 1} = O(\rho^5)$, but we cannot conclude that $W^-_{\infty 1\infty 1} = O(\rho^5)$. Thus, we may use the equation $W^-_{\infty 1\infty 1} = O(\rho^5)$ as a normalization on the metric which is independent of the Einstein equation. We will also use a normalization on the ρ^6-term in $W^-_{\infty 0\infty 0}$ whose vanishing is not imposed by the Einstein equation. To make sure that such normalizations in fact work, we must calculate the variations of $W^-_{\infty 1\infty 1}$ and $W^-_{\infty 0\infty 0}$ under the perturbation (2-12).

First, we calculate the relevant components of the curvature tensor modulo negligible terms. Since the curvature tensor is given by (2-9) and $R_{IJ\infty K} = -4R_K{}^{\infty}{}_{IJ}$, we obtain the following result by a straightforward computation using (2-7) and Lemma 2.3:

$$
\begin{aligned}
&R_{\infty 0\infty 0} \equiv 4+\frac{1}{2}\big((\rho\partial_\rho)^2-4\rho\partial_\rho+8\big)\varphi_{00}, && R_{01\infty 0} \equiv \frac{i}{4}(\rho\partial_\rho+1)\varphi_{01},\\
&R_{\infty 1\infty 1} \equiv \frac{1}{2}\big((\rho\partial_\rho)^2-2\rho\partial_\rho+2\big)\varphi_{11}, && R_{01\infty 1} \equiv \rho^2 A_{11}-\frac{i}{4}\rho\partial_\rho\varphi_{11},\\
&R_{0\bar{1}\infty 1} \equiv -\frac{i}{2}+\frac{i}{4}(\rho\partial_\rho-2)\varphi_{00}-\frac{i}{4}\rho\partial_\rho\varphi_{1\bar{1}}, && R_{1\bar{1}\infty 1} \equiv \frac{3i}{4}(\rho\partial_\rho-1)\varphi_{01},\\
&R_{1\bar{1}\infty 0} \equiv -i+\frac{i}{2}(\rho\partial_\rho-2)\varphi_{00}-\frac{i}{2}\rho\partial_\rho\varphi_{1\bar{1}}.
\end{aligned}
\tag{3-12}
$$

These equations enable us to compute the variations of the curvature components under the perturbation (2-12), which we denote by putting "δ" to each component. For example, by the first equation in (3-12), we have

$$\delta R_{\infty 0\infty 0} = \tfrac{1}{2}(m^2-4m+8)\psi_{00} + O(\rho^{m+1}).$$

Next, we calculate the variation of the Schouten tensor

$$P_{IJ} = \tfrac{1}{2}E_{IJ} - \tfrac{1}{12}(E_K{}^K+3)g_{IJ}.$$

Since $E_{IJ} = O(\rho)$ by Lemma 3.1, we have

$$\delta P_{IJ} = \tfrac{1}{2}\delta E_{IJ} - \tfrac{1}{12}g^{KL}(\delta E_{KL})g_{IJ} - \tfrac{1}{4}\delta g_{IJ} + O(\rho^{m+1}),$$

which yields

$$
\begin{aligned}
\delta P_{\infty\infty} &= -\frac{1}{6}(m^2-3m-1)\psi_{00} - \frac{1}{6}(2m^2-m+2)\psi_{1\bar{1}} + O(\rho^{m+1}),\\
\delta P_{\infty 0} &= O(\rho^{m+1}),\\
\delta P_{\infty 1} &= -\frac{i}{4}(m+1)\psi_{01} + O(\rho^{m+1}),\\
\delta P_{00} &= -\frac{1}{24}(m^2-6m-1)\psi_{00} + \frac{1}{24}(m^2+m-14)\psi_{1\bar{1}} + O(\rho^{m+1}),\\
\delta P_{01} &= -\frac{1}{16}(m^2-4m-1)\psi_{01} + O(\rho^{m+1}),\\
\delta P_{1\bar{1}} &= \frac{1}{48}(m^2-3m-10)\psi_{00} - \frac{1}{48}(m^2-8m-8)\psi_{1\bar{1}} + O(\rho^{m+1}),\\
\delta P_{11} &= -\frac{1}{16}(m^2-4m+4)\psi_{11} + O(\rho^{m+1}).
\end{aligned}
\tag{3-13}
$$

From (2-14), (3-12), and (3-13), we have

$$
\begin{aligned}
W_{\infty 1}{}^{0\bar{1}} &= O(\rho),\\
W_{\infty 0}{}^{\bar{1}1} &= -i + O(\rho),\\
\delta W_{\infty 1}{}^{0\bar{1}} &= -\frac{i}{4}(m-2)\psi_{11} + O(\rho^{m+1}),\\
\delta W_{\infty 0}{}^{\bar{1}1} &= \frac{i}{2}(m-2)\psi_{00} - \frac{i}{2}(m-4)\psi_{1\bar{1}} + O(\rho^{m+1}),\\
\delta W_{\infty 1\infty 1} &= \frac{1}{4}(m^2-4)\psi_{11} + O(\rho^{m+1}),\\
\delta W_{\infty 0\infty 0} &= \frac{1}{6}(m^2-3m+20)\psi_{00} - \frac{1}{6}(m^2-2m+16)\psi_{1\bar{1}} + O(\rho^{m+1}).
\end{aligned}
\tag{3-14}
$$

Finally, by (3-3) and (3-14), we obtain

$$
\begin{aligned}
\delta W^{-}_{\infty 1\infty 1} &= \tfrac{1}{2}(\delta W_{\infty 1\infty 1} - \delta\varepsilon_{\infty 10\bar{1}} \cdot W_{\infty 1}{}^{0\bar{1}} - \varepsilon_{\infty 10\bar{1}} \cdot \delta W_{\infty 1}{}^{0\bar{1}}) + O(\rho^{m+1})\\
&= \tfrac{1}{8}(m^2-2m)\psi_{11} + O(\rho^{m+1}),
\end{aligned}
\tag{3-15}
$$

$$
\begin{aligned}
\delta W^{-}_{\infty 0\infty 0} &= \tfrac{1}{2}(\delta W_{\infty 0\infty 0} - \delta\varepsilon_{\infty 0\bar{1}1} \cdot W_{\infty 0}{}^{\bar{1}1} - \varepsilon_{\infty 0\bar{1}1} \cdot \delta W_{\infty 0}{}^{\bar{1}1}) + O(\rho^{m+1})\\
&= \tfrac{1}{12}(m^2+3m+2)\psi_{00} - \tfrac{1}{12}(m^2+4m+4)\psi_{1\bar{1}} + O(\rho^{m+1}).
\end{aligned}
\tag{3-16}
$$

3C. ***Bianchi identities.*** Since the Einstein equation is an overdetermined system, we need some relations which are satisfied by the components of the Einstein tensor in order to construct a formal solution to the Einstein equation. Some of them are given by the Bianchi identity $g^{IJ}\nabla_K E_{IJ} = 2g^{IJ}\nabla_I E_{JK}$:

Lemma 3.4 [Matsumoto 2013, Lemma 6.6; 2014, Lemma 6.1]. *Suppose g satisfies $E_{IJ} = O(\rho^m)$ for an integer $m \geq 1$. Then, we have*

$$(m-8)E_{\infty\infty} - 4(m-4)E_{00} - 8(m-2)E_{1\bar{1}} = O(\rho^{m+1}), \tag{3-17}$$

$$(m-6)E_{\infty 0} = O(\rho^{m+1}), \tag{3-18}$$

$$(m-5)E_{\infty 1} - 4iE_{01} = O(\rho^{m+1}). \tag{3-19}$$

We will also use some equations obtained from the Bianchi identity $\nabla^I W^{-}_{IJKL} = C^{-}_{JKL}$ in the construction of g. Since the Cotton tensor is given by

$$C_{IJK} = \tfrac{1}{2}(\nabla_K E_{IJ} - \nabla_J E_{IK}) - \tfrac{1}{12}\big((\nabla_K E_L{}^L)g_{IJ} - (\nabla_J E_L{}^L)g_{IK}\big),$$

we can compute the components C^{-}_{IJK} in terms of E_{IJ} by using (2-7), (2-14), (3-3), and Lemma 2.3. As a result, we have the following lemma:

Lemma 3.5. *Suppose g satisfies $E_{IJ} = O(\rho^m)$ for an integer $m \geq 1$. Then,*

$$C^-_{1\infty 1} = -\tfrac{1}{4}(m-2)E_{11} + O(\rho^{m+1}), \tag{3-20}$$

$$C^-_{0\infty 0} = -\tfrac{5}{24}mE_{00} + \tfrac{1}{96}(m-12)E_{\infty\infty} + \tfrac{1}{12}(m+6)E_{1\bar{1}} + O(\rho^{m+1}), \tag{3-21}$$

$$C^-_{\infty\infty 0} = -\tfrac{1}{4}(m-2)E_{\infty 0} + O(\rho^{m+1}). \tag{3-22}$$

4. Construction of the metric

4A. ***The formal solution to the self-dual Einstein equation.*** Let M be a three-dimensional strictly pseudoconvex CR manifold. We fix a contact form θ and construct a one-parameter family of ACH metrics g^λ on $\overline{X} = M \times [0, \infty)_\rho$ which are in normal form with respect to θ and satisfy the Einstein equation to infinite order. The parameter $\lambda \in \mathbb{R}$ is involved in the normalization on the ρ^6-term in g^λ, and if $\lambda = 0$ the metric is self-dual to infinite order. As in the previous section, we take the Θ-frame $\{\boldsymbol{Z}_I\} = \{\rho\partial_\rho, \rho^2 T, \rho Z_1, \rho Z_{\bar{1}}\}$ associated with a unitary admissible frame $\{T, Z_1, Z_{\bar{1}}\}$ on M. We suppress the superscript λ in the following.

First we show a lemma which assures that our normalization condition is independent of the choice of θ.

Lemma 4.1. *Suppose that an ACH metric g on $\overline{X}$ satisfies $W^-_{IJKL} = O(\rho^6)$, and let ρ_θ be the model-defining function associated with a contact form θ. Then,*

$$\eta_\theta := \left(\rho_\theta^{-6} W^-_{\infty 0\infty 0}\right)\big|_M \tag{4-1}$$

satisfies $\eta_{\widehat{\theta}} = e^{-3\Upsilon}\eta_\theta$ for the rescaling $\widehat{\theta} = e^\Upsilon\theta$.

Proof. By Proposition 2.2, $\mathbf{Z}_\infty|_M$ and $\mathbf{Z}_0|_M$ are determined by g and independent of θ. Thus, we have

$$\widehat{W}^-_{\infty 0\infty 0} = W^-_{\infty 0\infty 0} + O(\rho^7).$$

Since $\rho_{\widehat{\theta}} = e^{\Upsilon/2}\rho_\theta + O(\rho^2)$ by (2-5), we obtain $\eta_{\widehat{\theta}} = e^{-3\Upsilon}\eta_\theta$. □

This lemma implies that if η_θ has a universal expression in terms of the Tanaka–Webster connection, then it defines a CR invariant $\eta \in \mathcal{E}(-3, -3)$. Since such a CR invariant is necessarily a multiple of the obstruction density [Graham 1987], we are led to the CR invariant normalization $\eta = \lambda\mathcal{O}$.

Now we construct the metric and prove Theorem 1.2. We start with an arbitrary normal form ACH metric $g^{(1)}_{IJ}$, which automatically satisfies $E_{IJ} = O(\rho)$ by Lemma 3.1. Supposing that we have a normal form ACH metric $g^{(m)}_{IJ}$ such that $E_{IJ} = O(\rho^m)$, we consider a perturbed metric

$$g^{(m+1)}_{IJ} = g^{(m)}_{IJ} + \psi_{IJ}, \quad \psi_{\infty J} = 0, \;\; \psi_{IJ} = O(\rho^m)$$

and try to solve $E_{IJ} = O(\rho^{m+1})$. We also take W^-_{IJKL} into consideration in each inductive step by using the following equations modulo $O(\rho)\cdot\mathcal{D}W^-_{IJKL}$ from (3-11):

$$\nabla^I W^-_{I1\infty 1} \equiv \frac{1}{4}(\rho\partial_\rho - 4)W^-_{\infty 1\infty 1}, \tag{4-2}$$

$$\nabla^I W^-_{I0\infty 0} \equiv \frac{1}{4}(\rho\partial_\rho - 6)W^-_{\infty 0\infty 0}, \tag{4-3}$$

$$\nabla^I W^-_{I\infty\infty 1} \equiv \frac{i}{2}W^-_{\infty 0\infty 1}, \tag{4-4}$$

$$\nabla^I W^-_{I\infty\infty 0} \equiv 0. \tag{4-5}$$

By Lemma 3.1, the variation of E_{IJ} is given by

$$\delta E_{\infty\infty} = -\frac{1}{2}m(m-4)\psi_{00} - m(m-2)\psi_{1\bar{1}} + O(\rho^{m+1}), \tag{4-6}$$

$$\delta E_{\infty 0} = O(\rho^{m+1}), \tag{4-7}$$

$$\delta E_{\infty 1} = -\frac{i}{2}(m+1)\psi_{01} + O(\rho^{m+1}), \tag{4-8}$$

$$\delta E_{00} = -\frac{1}{8}(m^2 - 6m - 4)\psi_{00} + \frac{1}{2}(m-2)\psi_{1\bar{1}} + O(\rho^{m+1}), \tag{4-9}$$

$$\delta E_{01} = -\frac{1}{8}(m+1)(m-5)\psi_{01} + O(\rho^{m+1}), \tag{4-10}$$

$$\delta E_{1\bar{1}} = -\frac{1}{8}(m^2 - 6m - 8)\psi_{1\bar{1}} + \frac{1}{8}(m-4)\psi_{00} + O(\rho^{m+1}), \tag{4-11}$$

$$\delta E_{11} = -\frac{1}{8}m(m-4)\psi_{11} + O(\rho^{m+1}). \tag{4-12}$$

The determinant of the coefficients of (4-9) and (4-11) as a system of linear equations for ψ_{00} and $\psi_{1\bar{1}}$ is

$$\det\begin{pmatrix} -\frac{1}{8}(m^2-6m-4) & \frac{1}{2}(m-2) \\ \frac{1}{8}(m-4) & -\frac{1}{8}(m^2-6m-8) \end{pmatrix} = \tfrac{1}{64}m(m+2)(m-6)(m-8).$$

First we consider the case of $m \le 5$, where the determinant is nonzero. We determine ψ_{00} and $\psi_{1\bar{1}}$ (modulo $O(\rho^{m+1})$) by (4-9) and (4-11) so that $E_{00}, E_{1\bar{1}} = O(\rho^{m+1})$ holds. Then, by the Bianchi identities (3-17) and (3-18), we have $E_{\infty\infty}$ and $E_{\infty 0} = O(\rho^{m+1})$. We determine ψ_{01} by (4-8) to obtain $E_{\infty 1} = O(\rho^{m+1})$. Then, (3-19) gives $E_{01} = O(\rho^{m+1})$. When $m \le 3$, (4-12) determines ψ_{11} so that $E_{11} = O(\rho^{m+1})$, thus we have $E_{IJ} = O(\rho^{m+1})$. Moreover, by (4-2)–(4-4) and Lemma 3.3, we also have $W^-_{IJKL} = O(\rho^{m+1})$. When $m = 4$, we cannot use (4-12) to obtain $E_{11} = O(\rho^5)$. However, since $W^-_{IJKL} = O(\rho^4)$, it follows from (4-2) that

$$C^-_{1\infty 1} = \nabla^I W^-_{I1\infty 1} = \tfrac{1}{4}(4-4)W^-_{\infty 1\infty 1} + O(\rho^5) = O(\rho^5),$$

so we have $E_{11}=O(\rho^5)$ by (3-20). (This also follows from the fact that the CR obstruction tensor $\mathcal{O}_{11}=(\rho^{-4}E_{11})|_M$ vanishes in three dimensions; see [Matsumoto 2013; 2014].) Thus, we have $E_{IJ}=O(\rho^5)$ and by (4-2)–(4-4), it holds that

$$W^-_{\infty1\infty1}=O(\rho^4),\quad W^-_{\infty0\infty0},W^-_{\infty0\infty1}=O(\rho^5).$$

We can choose ψ_{11} so that $W^-_{\infty1\infty1}=O(\rho^5)$ holds since

$$\delta W^-_{\infty1\infty1}=\psi_{11}+O(\rho^5)$$

by (3-15). Thus we obtain unique $g^{(5)}_{IJ}$ modulo $O(\rho^5)$ with $E_{IJ},W^-_{IJKL}=O(\rho^5)$. When $m=5$, we can construct $g^{(6)}_{IJ}$ with $E_{IJ}=O(\rho^6)$ in the same way as for $m\le 3$ and we also have $W^-_{IJKL}=O(\rho^6)$ by (4-2)–(4-4).

Next we consider the case of $m=6$, where the equations (4-6), (4-9), (4-11) are not pairwise independent. We determine ψ_{01} by (4-8) so that $E_{\infty1}=O(\rho^7)$. Then we also have $E_{01}=O(\rho^7)$ by (3-19). We determine ψ_{11} by (4-12) and obtain $E_{11}=O(\rho^7)$. By (3-16), we have

$$\delta W^-_{\infty0\infty0}=\tfrac{14}{3}\psi_{00}-\tfrac{16}{3}\psi_{1\bar1}+O(\rho^7). \tag{4-13}$$

We use this equation and (4-6) to determine ψ_{00}, $\psi_{1\bar1}$ so that

$$E_{\infty\infty}=O(\rho^7),\quad \eta=\lambda\mathcal{O}$$

holds. Thus we have determined $g^{(7)}_{IJ}$ and we must check that it also satisfies $E_{00},E_{1\bar1},E_{\infty0}=O(\rho^7)$. Since $W^-_{IJKL}=O(\rho^6)$, by (4-3) we have

$$C^-_{0\infty0}=\tfrac14(6-6)W^-_{\infty0\infty0}+O(\rho^7)=O(\rho^7).$$

Then it follows from (3-21) that

$$-\tfrac54E_{00}+E_{1\bar1}=O(\rho^7).$$

Also, (3-17) gives

$$E_{00}+4E_{1\bar1}=O(\rho^7).$$

Therefore, we have $E_{00},E_{1\bar1}=O(\rho^7)$. Moreover, by $W^-_{IJKL}=O(\rho^6)$ and (4-5), it holds that $C^-_{\infty\infty0}=O(\rho^7)$, which implies $E_{\infty0}=O(\rho^7)$ by (3-22). Thus, $g^{(7)}_{IJ}$ satisfies $E_{IJ}=O(\rho^7)$, $W^-_{IJKL}=O(\rho^6)$, and $\eta=\lambda\mathcal{O}$. We note that it satisfies $W^-_{IJKL}=O(\rho^7)$ when $\lambda=0$.

When $m=7$, we can determine $g^{(8)}_{IJ}$ so that it satisfies $E_{IJ}=O(\rho^8)$ in the same way as in $m\le 3$. If $\lambda=0$, it also satisfies $W^-_{IJKL}=O(\rho^8)$ by (4-2)–(4-4).

Let us consider the case of $m=8$. In this case, (4-9) and (4-11) are not independent. We use (4-6) and (4-9) to determine ψ_{00} and $\psi_{1\bar1}$ so that $E_{\infty\infty},E_{00}=O(\rho^9)$. Then (3-17) gives $E_{1\bar1}=O(\rho^9)$. We determine ψ_{01} and ψ_{11} by (4-8) and (4-12) respectively and obtain $E_{\infty1},E_{11}=O(\rho^9)$. By (3-18), (3-19), we have

$E_{\infty 0}$, $E_{01} = O(\rho^9)$. Thus we have constructed $g_{IJ}^{(9)}$ with $E_{IJ} = O(\rho^9)$, which satisfies $W^-_{IJKL} = O(\rho^9)$ when $\lambda = 0$ by (4-2)–(4-4).

Finally, let $m \geq 9$. In this case, the equation $E_{IJ} = O(\rho^{m+1})$ determines $g_{IJ}^{(m+1)}$ in the same way as in $m \leq 3$, and it satisfies $W^-_{IJKL} = O(\rho^{m+1})$ by (4-2)–(4-4) when $\lambda = 0$.

Consequently, we can construct all $g_{IJ}^{(m+1)}$ inductively, and by Borel's lemma we obtain a solution g^λ_{IJ} to

$$E_{IJ} = O(\rho^\infty), \quad W^-_{IJKL} = O(\rho^6), \quad \eta = \lambda\mathcal{O},$$

which is unique modulo $O(\rho^\infty)$. By the construction, g^0_{IJ} satisfies $W^-_{IJKL} = O(\rho^\infty)$. Thus we complete the proof of Theorem 1.2.

4B. *Dependence on* λ. We can read off the dependence of g^λ_{IJ} on the parameter λ from the construction.

Proposition 4.2. *The metric g^λ_{IJ} admits the following asymptotic expansion*:

$$g^\lambda_{IJ} \sim g^0_{IJ} + \sum_{k=1}^{\infty} \lambda^k \rho^{6k} \phi^{(k)}_{IJ}(\rho), \quad \phi^{(k)}_{I\infty} = \phi^{(k)}_{01} = \phi^{(k)}_{11} = 0.$$

Here, $\phi^{(k)}_{IJ}(\rho)$ is a formal power series in ρ.

Proof. We write the Taylor expansion of $g^\lambda_{IJ} - g^0_{IJ}$ as

$$g^\lambda_{IJ} - g^0_{IJ} \sim \sum_{k=0}^{\infty} \rho^k \Phi^{\lambda,k}_{IJ}, \quad \Phi^{\lambda,k}_{I\infty} = 0.$$

Then, it suffices to show that $\Phi^{\lambda,k}_{01} = \Phi^{\lambda,k}_{11} = 0$ and each $\Phi^{\lambda,k}_{IJ}$ is a polynomial in λ of degree $\leq k/6$. First, we note that $g^\lambda_{IJ} - g^0_{IJ} = O(\rho^6)$, so $\Phi^{\lambda,k}_{IJ} = 0$ for $k \leq 5$. Since both g^λ_{IJ} and g^0_{IJ} satisfy $E_{\infty 1}$, $E_{11} = O(\rho^\infty)$, we also have that $\Phi^{\lambda,k}_{01} = \Phi^{\lambda,k}_{11} = 0$ for $k \geq 6$, by (4-8), (4-12). From (4-6), (4-13), we see $\Phi^{\lambda,6}_{00}$ and $\Phi^{\lambda,6}_{1\bar{1}}$ are determined by

$$-6\Phi^{\lambda,6}_{00} - 24\Phi^{\lambda,6}_{1\bar{1}} = 0,$$
$$\tfrac{14}{3}\Phi^{\lambda,6}_{00} - \tfrac{16}{3}\Phi^{\lambda,6}_{1\bar{1}} = \lambda\mathcal{O}.$$

Thus we have $\deg \Phi^{\lambda,6}_{00} = \deg \Phi^{\lambda,6}_{1\bar{1}} = 1$. Now we prove $\deg \Phi^{\lambda,k}_{00}$, $\deg \Phi^{\lambda,k}_{1\bar{1}} \leq k/6$ by the induction on k. When $k \geq 7$, $\Phi^{\lambda,k}_{00}$ and $\Phi^{\lambda,k}_{1\bar{1}}$ are determined by the condition $\partial^k_\rho E_{00}|_{\rho=0} = \partial^k_\rho E_{1\bar{1}}|_{\rho=0} = 0$ for $k \neq 8$ and $\partial^k_\rho E_{\infty\infty}|_{\rho=0} = \partial^k_\rho E_{00}|_{\rho=0} = 0$ for $k = 8$. These conditions can be regarded as a system of linear equations for $\Phi^{\lambda,k}_{00}$ and $\Phi^{\lambda,k}_{1\bar{1}}$, and in view of (2-7), (2-11), (3-2), the terms involving the other components are linear combinations of

$$\mathcal{D}_1 \Phi^{\lambda,l_1}_{I_1 J_1} \cdots \mathcal{D}_p \Phi^{\lambda,l_p}_{I_p J_p}, \quad (l_1 + \cdots + l_p \leq k,\ l_j < k),$$

where $\mathcal{D}_j$ is a differential operator on M. Then, by the induction hypothesis,

$$\deg \Phi^{\lambda,k}_{IJ} \leq \frac{l_1+\cdots+l_p}{6} \leq \frac{k}{6}.$$

Thus, we complete the proof. $\square$

4C. *Evenness.* Let g be a normal form ACH metric on $M \times [0,\infty)_\rho$. Then it can be written in the form

$$g = \frac{h_\rho + 4d\rho^2}{\rho^2}, \tag{4-14}$$

where h_ρ is a family of Riemannian metrics on M. We say g is *even* when h_ρ has even Laurent expansion at $\rho = 0$. In other words, g is even if and only if the components g_{00}, g_{11}, $g_{1\bar{1}}$ are even in ρ, and g_{01} is odd in ρ. An ACH metric is said to be even if its normal form is even for any choice of θ.

Proposition 4.3. *The ACH metric g^λ is even.*

Proof. Fix a contact form θ and suppose g^λ is in the normal form as (4-14). By using the Laurent expansion of h_ρ, we can regard the right-hand side of (4-14) as an ACH metric g^λ_- defined on $M \times (-\infty, 0]_\rho$. Then, g^λ_- also satisfies

$$E_{IJ} = O(\rho^\infty), \quad W^-_{IJKL} = O(\rho^6), \quad \eta = \lambda\mathcal{O},$$

with respect to the orientation satisfying

$$i\boldsymbol{\theta}^0 \wedge \boldsymbol{\theta}^1 \wedge \boldsymbol{\theta}^{\bar{1}} \wedge \boldsymbol{\theta}^\infty = i\rho^{-5}\theta \wedge \theta^1 \wedge \theta^{\bar{1}} \wedge d\rho > 0.$$

We consider the ACH metric $\iota^* g^\lambda_-$ on $M \times [0,\infty)_\rho$, where $\iota(x,\rho) := (x, -\rho)$. Since ι preserves the orientation, $\iota^* g^\lambda_-$ satisfies

$$E_{IJ} = O(\rho^\infty), \quad W^-_{IJKL} = O(\rho^6).$$

Noting that $\iota_* \boldsymbol{Z}_\infty = \boldsymbol{Z}_\infty$ and $\iota_* \boldsymbol{Z}_0 = \boldsymbol{Z}_0$, we have

$$\begin{aligned}
\rho^{-6} W^-[\iota^* g^\lambda_-]_{\infty 0 \infty 0} &= (\iota^*\rho)^{-6} (\iota^* W^-[g^\lambda_-])(\boldsymbol{Z}_\infty, \boldsymbol{Z}_0, \boldsymbol{Z}_\infty, \boldsymbol{Z}_0) \\
&= \iota^*\big(\rho^{-6} W^-[g^\lambda_-](\iota_*\boldsymbol{Z}_\infty, \iota_*\boldsymbol{Z}_0, \iota_*\boldsymbol{Z}_\infty, \iota_*\boldsymbol{Z}_0)\big) \\
&= \iota^*(\rho^{-6} W^-[g^\lambda_-]_{\infty 0 \infty 0}).
\end{aligned}$$

Thus, $\iota^* g^\lambda_-$ also satisfies $\eta = \lambda\mathcal{O}$. Therefore, by the uniqueness we obtain

$$\iota^* g^\lambda_- = g^\lambda + O(\rho^\infty),$$

which implies that g^λ is even. $\square$

5. CR GJMS operators

Matsumoto [2016] generalized the CR GJMS operators to partially integrable CR manifolds via Dirichlet-to-Neumann type operators associated with eigenvalue equations for the laplacian of the ACH metric. In dimension three, it is stated as follows:

Theorem 5.1 [Matsumoto 2016, Theorem 3.3]. *Let M be a three-dimensional strictly pseudoconvex CR manifold and g an ACH metric on a Θ-manifold $\overline{X}$ with the boundary M. Let θ be a contact form on M and let ρ be the model-defining function associated with θ. Then, for any $k \in \mathbb{N}_+$ and $f \in C^\infty(M)$, there exist $F, G \in C^\infty(\overline{X})$ with $F|_M = f$ such that the function $u := \rho^{-k+2}F + (\rho^{k+2}\log\rho)G$ satisfies*

$$\Big(\Delta + \frac{k^2}{4} - 1\Big)u = O(\rho^\infty),$$

where $\Delta = -g^{IJ}\nabla_I\nabla_J$ is the laplacian of g. The function G is unique modulo $O(\rho^\infty)$ and $P_{2k} f := (-1)^{k+1}k!(k-1)!/2 \cdot G|_M$ defines a formally self-adjoint linear differential operator $\mathcal{E}(k/2-1, k/2-1) \to \mathcal{E}(-k/2-1, -k/2-1)$ which is independent of the choice of θ and has the principal part Δ_b^k.

We apply this theorem to our metric g^λ. Since g^λ is determined to infinite order and the Taylor expansion has a universal expression in terms of the pseudohermitian structure, the operator P_{2k}^λ has a universal expression in terms of the Tanaka–Webster connection. Thus, we obtain the CR GJMS operators P_{2k}^λ for all $k \geq 1$.

In order to prove that P_{2k}^λ is a polynomial in λ of degree $\leq k/3$, we will review the details of its construction. A linear differential operator on $\overline{X}$ is called a Θ-*differential operator* if it is the sum of linear differential operators of the form $aY_1 \cdots Y_N$, where $a \in C^\infty(\overline{X})$ and $Y_j \in \Gamma({}^\Theta T\overline{X})$. Note that a Θ-differential operator preserves the subspace $\rho^m C^\infty(\overline{X}) \subset C^\infty(\overline{X})$ for each $m \geq 1$. We fix a contact form θ and denote the associated Tanaka–Webster connection by ∇^{TW}. Suppose that g^λ is of the normal form

$$g^\lambda = k_\rho + 4\frac{d\rho^2}{\rho^2}$$

for θ, where k_ρ is a family of Riemannian metrics on M. Then, the laplacian Δ of g^λ is written as

$$\Delta = -\tfrac{1}{4}(\rho\partial_\rho)^2 + \rho\partial_\rho + \rho^2\Delta_b - \rho^4 T^2 + \rho\Psi \tag{5-1}$$

with the Θ-differential operator Ψ defined by

$$\begin{aligned}\Psi f = &-\tfrac{1}{8}(\partial_\rho \log\det k_\rho)\rho\partial_\rho f - \rho^{-1}\big((k_\rho^{-1})^{ij} - (k_0^{-1})^{ij}\big)\nabla_i^{\mathrm{TW}}\nabla_j^{\mathrm{TW}} f\\ &+ \tfrac{1}{2}(k_\rho^{-1})^{ij}(k_\rho^{-1})^{kl}\rho^{-1}\big(\nabla_i^{\mathrm{TW}}(k_\rho)_{jk} + \nabla_j^{\mathrm{TW}}(k_\rho)_{ik} - \nabla_k^{\mathrm{TW}}(k_\rho)_{ij}\big)\nabla_l^{\mathrm{TW}} f.\end{aligned}$$

Here, the components are with respect to a Θ-frame $\{\boldsymbol{Z}_I\}$, and we note that $(k_\rho^{-1})^{ij} - (k_0^{-1})^{ij}$ and $\nabla_i^{\mathrm{TW}}(k_\rho)_{jk} + \nabla_j^{\mathrm{TW}}(k_\rho)_{ik} - \nabla_k^{\mathrm{TW}}(k_\rho)_{ij}$ are $O(\rho)$ by (2-6). In particular, Ψ involves $\partial_\rho g^\lambda_{IJ}$ but not higher order derivatives.

Given a function $f \in C^\infty(M)$, we try to solve the equation

$$\Big(\Delta + \frac{k^2}{4} - 1\Big)(\rho^{-k+2}F) = 0$$

for $F \in C^\infty(\overline{X})$ with $F|_M = f$. Let $F \sim \sum_{j=0}^\infty f^{(j)}\rho^j$, $(f^{(j)} \in C^\infty(M))$ be the Taylor expansion of F along M. By (5-1), we have

$$\text{(5-2)}\quad \Big(\Delta + \frac{k^2}{4} - 1\Big)(\rho^{-k+2+j} f^{(j)}) = \rho^{-k+2+j}\Big(-\frac{1}{4}j(j-2k)f^{(j)} + \rho\mathcal{D}_j f^{(j)}\Big),$$

where $\mathcal{D}_j$ is a ρ-dependent linear differential operator on M. Starting with $f^{(0)} = f$, we inductively define $f^{(j)}$ so that F satisfies

$$\Big(\Delta + \frac{k^2}{4} - 1\Big)(\rho^{-k+2}F) = O(\rho^{-k+3+j}).$$

Let $\mathcal{D}_j \sim \sum_{l=0}^\infty \mathcal{D}_j^{(l)}\rho^l$ be the Taylor expansion of $\mathcal{D}_j$. Then, by (5-2), $f^{(j)}$ is determined for $j \le 2k-1$ as

$$f^{(j)} = \frac{4}{j(j-2k)}\sum_{l=0}^{j-1}\mathcal{D}_l^{(j-1-l)} f^{(l)}.$$

We cannot define $f^{(2k)}$ due to the vanishing of the coefficient of $f^{(2k)}$ in (5-2), and we need to introduce the logarithmic term $(\rho^{k+2}\log\rho)G$ in which the coefficient $G|_M$ is a multiple of

$$\Big(\rho^{-k-2}\Big(\Delta + \frac{k^2}{4} - 1\Big)(\rho^{-k+2}F)\Big)\Big|_M.$$

Therefore, up to a constant multiple, $P_{2k}^\lambda f$ is given by

$$\sum_{j=0}^{2k-1}\mathcal{D}_j^{(2k-1-j)} f^{(j)}.$$

Since Ψ involves only g^λ_{IJ} and their first order derivatives in ρ, $\mathcal{D}_j^{(l)}$ involves $\partial_\rho^m g^\lambda_{IJ}$ for $m \le l+1$. Consequently, P_{2k}^λ is written in terms of $\partial_\rho^m g^\lambda_{IJ}$ $(m \le 2k)$, and by Proposition 4.2 it is a polynomial in λ of degree $\le k/3$. Thus we complete the proof of Theorem 1.3.

6. Convergence of the formal solutions

We will prove Theorem 1.4, which asserts that the formal solution g^λ converges to a real analytic ACH metric near M when M is a real analytic CR manifold. In the case of $\lambda = 0$, this recovers the result of Biquard [2007]. The key tool is the result of Baouendi and Goulaouic [1976] on the unique existence of the solution to a singular nonlinear Cauchy problem. Let us state their theorem in a form which fits to our setting.

We regard local coordinates (x, ρ) of $M \times [0, \infty)_\rho$ as complex variables and consider an equation for a $\mathbb{C}^N$-valued holomorphic function $v(x, \rho)$ of the form

$$(6\text{-}1)\quad (\rho\partial_\rho)^m v + A_{m-1}(\rho\partial_\rho)^{m-1} v + \cdots + A_0 v = F(x, \rho, \{(\rho\partial_\rho)^l \partial_x^\alpha (\rho v)\}_{l+|\alpha|\le m,\ l<m}),$$

where A_j is an $N \times N$ matrix and $F(x, \rho, \{y_{l,\alpha}\}_{l+|\alpha|\le m,\ l<m})$ is a holomorphic function near 0. For each $k \in \mathbb{N}$, we set

$$\mathcal{P}(k) := k^m I + k^{m-1} A_{m-1} + \cdots + A_0,$$

where I is the identity matrix of size N. Then, by [Baouendi and Goulaouic 1976, Theorem 3.1] we have the following theorem:

Theorem 6.1. *If* $\det \mathcal{P}(k) \neq 0$ *for all* $k \in \mathbb{N}$, (6-1) *has a unique holomorphic solution* $v(x, \rho)$ *near* $(0, 0)$.

In the original statement of [Baouendi and Goulaouic 1976, Theorem 3.1], the right-hand side of the (6-1) is replaced by $G(\rho, \{(\rho\partial_\rho)^l \partial_x^\alpha (\rho v)\}_{l+|\alpha|\le m,\ l<m})$ with G a C^∞-map

$$G : \mathbb{C} \times B^{N'} \to B,$$

where B is the Banach space of $\mathbb{C}^N$-valued bounded holomorphic functions of x on a fixed polydisc, and N' is the number of multiindices (l, α) such that $l + |\alpha| \le m$, $l < m$. Also, the solution v is given as a C^∞-function of ρ valued in B. In our (6-1), G is given by $G(\rho, \{y_{l,\alpha}\}) := F(x, \rho, \{y_{l,\alpha}(x)\})$. Since this is analytic in ρ, it follows from [Baouendi and Goulaouic 1976, Remark 2.2] and the proof of [Baouendi and Goulaouic 1976, Theorem 3.1] that the solution $v(x, \rho)$ is C^∞ and $v(x, \rho^m)$ is holomorphic, which implies that $v(x, \rho)$ itself is holomorphic. Thus we obtain Theorem 6.1 as a special case of their theorem.

Now we apply this theorem to our case. We assume that M is a real analytic CR manifold. Let g^λ_{IJ} be the components of the formal solution g^λ in a Θ-frame $\{\boldsymbol{Z}_I\}$, and let

$$g^{(k)}_{IJ} := \frac{1}{k!} \partial_\rho^k g^\lambda_{IJ}\big|_M$$

be the Taylor coefficients, which are analytic functions on M. We consider an ACH metric of the form

$$\widetilde{g}^\lambda_{IJ} = \sum_{k=0}^{8} \rho^k g^{(k)}_{IJ} + \rho^9 \widetilde{\varphi}_{IJ},$$

which automatically satisfies $E_{IJ} = O(\rho^9)$. Then, we consider the equation

$$-8\rho^{-9}(E_{00}, E_{1\bar{1}}, E_{01}, E_{11}) = 0 \tag{6-2}$$

for $v = (\widetilde{\varphi}_{00}, \widetilde{\varphi}_{1\bar{1}}, \widetilde{\varphi}_{01}, \widetilde{\varphi}_{11})$. We shall show that this equation is written in the form (6-1) for $m = 2$ and satisfies the assumption of Theorem 6.1; then we can conclude that g^{λ}_{IJ} converges since it gives the Taylor expansion of the solution v.

We see that in Lemma 3.1 the negligible term which we ignored in the computation of E_{IJ} is an analytic function in

$$x,\ \rho,\ \rho(\rho\partial_\rho)^l\partial_x^\alpha(\rho^9\widetilde{\varphi}) \quad \text{for } l + |\alpha| \le 2,\ l < 2.$$

Thus, it can be written in the form

$$f^{(1)}_{IJ}(x, \rho) + \rho^9 f^{(2)}_{IJ}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2})$$

with analytic functions $f^{(1)}_{IJ}, f^{(2)}_{IJ}$. Then, by Lemma 3.1, we have

$$\begin{aligned}-8E_{00} = {} & I_1(\rho\partial_\rho)\varphi_{00} + I_2(\rho\partial_\rho)\varphi_{1\bar{1}} \\ & + 16\rho^4|A|^2 + f^{(1)}_{00}(x, \rho) + \rho^9 f^{(2)}_{00}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2}),\end{aligned}$$

where

$$\varphi_{IJ} = \sum_{k=1}^{8} \rho^k g^{(k)}_{IJ} + \rho^9\widetilde{\varphi}_{IJ}$$

and

$$I_1(t) = t^2 - 6t - 4, \quad I_2(t) = -4(t-2).$$

Since $E_{00} = O(\rho^9)$, we have

$$I_1(\rho\partial_\rho)\Big(\sum_{k=1}^{8}\rho^k g^{(k)}_{00}\Big) + I_2(\rho\partial_\rho)\Big(\sum_{k=1}^{8}\rho^k g^{(k)}_{1\bar{1}}\Big) + 16\rho^4|A|^2 + f^{(1)}_{00}(x, \rho) = \rho^9 f^{(0)}_{00}(x, \rho)$$

with some analytic function $f^{(0)}_{00}$. Therefore, the equation $-8\rho^{-9}E_{00} = 0$ is written

$$I_1(\rho\partial_\rho + 9)\widetilde{\varphi}_{00} + I_2(\rho\partial_\rho + 9)\widetilde{\varphi}_{1\bar{1}} + F_{00}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2}) = 0$$

with an analytic function F_{00}.

Similarly, the equations $-8\rho^{-9}E_{IJ} = 0$ for $(I, J) = (1, \bar{1}), (0, 1), (1, 1)$ are respectively written as

$$\begin{aligned}I_3(\rho\partial_\rho + 9)\widetilde{\varphi}_{00} + I_4(\rho\partial_\rho + 9)\widetilde{\varphi}_{1\bar{1}} + F_{1\bar{1}}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2}) &= 0,\\ I_5(\rho\partial_\rho + 9)\widetilde{\varphi}_{01} + F_{01}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2}) &= 0,\\ I_6(\rho\partial_\rho + 9)\widetilde{\varphi}_{11} + F_{11}(x, \rho, \{(\rho\partial_\rho)^l\partial_x^\alpha(\rho\widetilde{\varphi})\}_{l+|\alpha|\le 2,\ l<2}) &= 0,\end{aligned}$$

where $F_{1\bar{1}}$, F_{01}, F_{11} are analytic functions and

$$I_3(t) = -t+4, \quad I_4(t) = t^2 - 6t - 8, \quad I_5(t) = (t+1)(t-5), \quad I_6(t) = t(t-4).$$

Hence (6-2) is of the form (6-1), and we have

$$\begin{aligned}\det \mathcal{P}(k) &= \det \begin{pmatrix} I_1(k+9) & I_2(k+9) & & \\ I_3(k+9) & I_4(k+9) & & \\ & & I_5(k+9) & \\ & & & I_6(k+9) \end{pmatrix} \\ &= (k+1)(k+3)(k+4)(k+5)(k+9)^2(k+10)(k+11) \\ &\neq 0\end{aligned}$$

for any $k \in \mathbb{N}$. Thus, by Theorem 6.1, (6-2) has a unique holomorphic solution and we complete the proof of Theorem 1.4.

Acknowledgements

The author is grateful to Kengo Hirachi and Yoshihiko Matsumoto for invaluable comments on the results. He also thanks Matsumoto for checking the author's suggestion of some minor corrections to the computation in his thesis [Matsumoto 2013].

References

[Baouendi and Goulaouic 1976] M. S. Baouendi and C. Goulaouic, "Singular nonlinear Cauchy problems", *J. Differential Equations* **22**:2 (1976), 268–291. MR Zbl

[Biquard 2007] O. Biquard, "Sur les variétés CR de dimension 3 et les twisteurs", *Ann. Inst. Fourier (Grenoble)* **57**:4 (2007), 1161–1180. MR Zbl

[Čap et al. 2001] A. Čap, J. Slovák, and V. Souček, "Bernstein–Gelfand–Gelfand sequences", *Ann. of Math.* (2) **154**:1 (2001), 97–113. MR Zbl

[Cheng and Lee 1990] J. H. Cheng and J. M. Lee, "The Burns–Epstein invariant and deformation of CR structures", *Duke Math. J.* **60**:1 (1990), 221–254. MR Zbl

[Epstein et al. 1991] C. L. Epstein, R. B. Melrose, and G. A. Mendoza, "Resolvent of the Laplacian on strictly pseudoconvex domains", *Acta Math.* **167**:1-2 (1991), 1–106. MR Zbl

[Fefferman 1976] C. L. Fefferman, "Monge–Ampère equations, the Bergman kernel, and geometry of pseudoconvex domains", *Ann. of Math.* (2) **103**:2 (1976), 395–416. Correction in **104**:3 (1976), 393–394. MR Zbl

[Fefferman and Graham 2002] C. Fefferman and C. R. Graham, "Q-curvature and Poincaré metrics", *Math. Res. Lett.* **9**:2-3 (2002), 139–151. MR Zbl

[Fefferman and Graham 2012] C. Fefferman and C. R. Graham, *The ambient metric*, Ann. of Math. Studies **178**, Princeton Univ. Press, 2012. MR Zbl

[Fefferman and Hirachi 2003] C. Fefferman and K. Hirachi, "Ambient metric construction of Q-curvature in conformal and CR geometries", *Math. Res. Lett.* **10**:5-6 (2003), 819–831. MR Zbl

[Gover and Graham 2005] A. R. Gover and C. R. Graham, “CR invariant powers of the sub-Laplacian”, *J. Reine Angew. Math.* **583** (2005), 1–27. MR Zbl

[Gover and Hirachi 2004] A. R. Gover and K. Hirachi, “Conformally invariant powers of the Laplacian: a complete nonexistence theorem”, *J. Amer. Math. Soc.* **17**:2 (2004), 389–405. MR Zbl

[Graham 1987] C. R. Graham, “Scalar boundary invariants and the Bergman kernel”, pp. 108–135 in *Complex analysis, II* (College Park, MD, 1985/86), edited by C. A. Berenstein, Lect. Notes in Math. **1276**, Springer, 1987. MR Zbl

[Graham and Zworski 2003] C. R. Graham and M. Zworski, “Scattering matrix in conformal geometry”, *Invent. Math.* **152**:1 (2003), 89–118. MR Zbl

[Graham et al. 1992] C. R. Graham, R. Jenne, L. J. Mason, and G. A. J. Sparling, “Conformally invariant powers of the Laplacian, I: Existence”, *J. London Math. Soc.* (2) **46**:3 (1992), 557–565. MR Zbl

[Hislop et al. 2008] P. D. Hislop, P. A. Perry, and S.-H. Tang, “CR-invariants and the scattering operator for complex manifolds with boundary”, *Anal. PDE* **1**:2 (2008), 197–227. MR Zbl

[LeBrun 1982] C. R. LeBrun, “$\mathscr{H}$-space with a cosmological constant”, *Proc. Roy. Soc. London Ser. A* **380**:1778 (1982), 171–185. MR Zbl

[Lee 1988] J. M. Lee, “Pseudo-Einstein structures on CR manifolds”, *Amer. J. Math.* **110**:1 (1988), 157–178. MR Zbl

[Lee and Melrose 1982] J. M. Lee and R. Melrose, “Boundary behaviour of the complex Monge–Ampère equation”, *Acta Math.* **148** (1982), 159–192. MR Zbl

[Matsumoto 2013] Y. Matsumoto, *Asymptotically complex hyperbolic Einstein metrics and CR geometry*, Ph.D. thesis, University of Tokyo, 2013, Available at https://tinyurl.com/matsthes.

[Matsumoto 2014] Y. Matsumoto, “Asymptotics of ACH-Einstein metrics”, *J. Geom. Anal.* **24**:4 (2014), 2135–2185. MR Zbl

[Matsumoto 2016] Y. Matsumoto, “GJMS operators, Q-curvature, and obstruction tensor of partially integrable CR manifolds”, *Differential Geom. Appl.* **45** (2016), 78–114. MR Zbl

Received February 11, 2018. Revised November 25, 2018.

TAIJI MARUGAME
INSTITUTE OF MATHEMATICS, ACADEMIA SINICA
TAIPEI
TAIWAN

marugame@gate.sinica.edu.tw

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.547

DOUBLE GRAPH COMPLEX AND CHARACTERISTIC CLASSES OF FIBRATIONS

TAKAHIRO MATSUYUKI

In this paper, we construct a double chain complex generated by certain graphs and a chain map from that to the Chevalley–Eilenberg double complex of the differential graded Lie algebra (dgl) of symplectic derivations on a free dgl. It is known that the target of the map is related to characteristic classes of fibrations. We can describe some characteristic classes of fibrations whose fiber is a 1-punctured even-dimensional manifold by linear combinations of graphs though the cohomology of the dgl of derivations.

1. Introduction

The Chevalley–Eilenberg complex of the limit of the Lie algebra of symplectic derivations on (graded) free Lie algebras is isomorphic to the graph complex defined by the cyclic Lie operad (details in [Kontsevich 1994; 1993; Conant and Vogtmann 2003; Hamilton 2006]). In this paper, we introduce an extension of (the dual of) the construction to a Lie algebra of symplectic derivations on free differential graded Lie algebras (dgls). Let (W, ω) be a finite-dimensional graded vector space with symmetric inner product of even degree N and δ a differential of degree -1 on the completed free Lie algebra $\hat{L}W$ satisfying the symplectic condition $\delta\omega = 0$. An important example is the case that $(\hat{L}W, \delta)$ is a Chen dgl model of an even-dimensional manifold and ω is its intersection form. We construct a W-labeled graph complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ and a chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\hat{L}W))$$

to the Chevalley–Eilenberg (double) complex $C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\hat{L}W))$ of the differential graded Lie algebra $(\mathrm{Der}_\omega(\hat{L}W), \mathrm{ad}(\delta))$ of symplectic derivations on $\hat{L}W$. Furthermore, from the nonlabeled part $C^{\bullet;\bullet}_{\mathrm{com}}(N, Z)$ of the graph complex, which depends on only the integer N and the set Z of degrees of W, we can obtain a chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(N, Z) \subset C^{\bullet;\bullet}_{\mathrm{com}}(W)^{\mathrm{Sp}(W,\delta)} \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\hat{L}W))^{\mathrm{Sp}(W,\delta)},$$

I would like to thank Y. Terashima and H. Kajiura for many helpful comments. This work was supported by Grant-in-Aid for Japan Society for the Promotion of Science Research Fellow (17J01757).
MSC2010: primary 55R40; secondary 55P62.
Keywords: graph complex, characteristic class, derivation, fiber bundle.

where $\mathrm{Sp}(W,\delta)$ is the group of graded linear isomorphisms of W preserving ω and δ. In the case of $N=0$ and $Z=\{0\}$, the map corresponds to the Kontsevich one [1994; 1993].

The construction above gives characteristic classes of fibrations. It is known that characteristic classes of simply connected fibrations are related to Lie algebras of derivations [Schlessinger and Stasheff 2012; Tanré 1983]. In nonsimply connected cases, we got relations between characteristic classes and Lie algebras of derivations as in [Matsuyuki and Terashima 2016; Kajiura et al. 2016]. In this paper, we consider the case that the boundary of a fiber is a sphere. For a simply connected compact manifold X with $\partial X = S^{n-1}$, let $\mathrm{aut}_\partial(X)$ be the monoid of self-homotopy equivalences of X fixing the boundary pointwisely and $\mathrm{aut}_{\partial,0}(X)$ its connected component containing id_X. According to [Berglund and Madsen 2014], the isomorphism

$$H^\bullet(B\,\mathrm{aut}_{\partial,0}(X);\mathbb{Q}) \simeq H^\bullet_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))$$

is obtained. Here L_X is a cofibrant dgl model of X. The underlying Lie algebra of L_X is generated by the linear dual W of the suspension of the reduced cohomology of X. So the graph complex above gives the invariant part of the cohomology $H^\bullet_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))$ with respect to the action of the group $\mathrm{Sp}(W,\delta)$ of automorphisms of W with intersection form preserving the differential δ of L_X. Using the Serre spectral sequence for the fibration

$$B\,\mathrm{aut}_{\partial,0}(X) \to B\,\mathrm{aut}_\partial(X) \to B\pi_0(\mathrm{aut}_\partial(X)),$$

the image of the natural map $H^\bullet(B\,\mathrm{aut}_\partial(X);\mathbb{Q}) \to H^\bullet(B\,\mathrm{aut}_{\partial,0}(X);\mathbb{Q})$ is included in the invariant part. We give a chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(N,Z)_+ \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))^{\mathrm{Sp}(W,\delta)}$$

using a positive truncated version $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ of $C^{\bullet;\bullet}_{\mathrm{com}}(W)$. Considering W-labeled graphs, we can also obtain a W-labeled version $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ and a chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(W)_+ \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X)).$$

2. Preliminary

In this paper, all vector spaces are over a field K whose characteristic is zero. A field K is regarded as a $\mathbb{Z}$-graded vector space all of whose elements have degree 0.

For a finite set U, the number of elements in U is denoted by $\#U$.

All tensor products of linear maps between $\mathbb{Z}$-graded vector spaces contain their signs: for homogeneous linear maps $f : A \to V$ and $g : B \to W$ between $\mathbb{Z}$-graded vector spaces, we set

$$(f \otimes g)(a, b) := (-1)^{ga} f(a) \otimes g(b)$$

for $a \in A$ and $b \in B$. (We often denote by $|a|$ the degree of an element a. But we omit the symbol $|\cdot|$ of the degree when it appears in a power of -1. For example, $(-1)^{ga}$ means $(-1)^{|g||a|}$.)

Let V be a $\mathbb{Z}$-graded vector space. We denote by V^i the subspace of elements of V of *cohomological degree* i and $V_i = V^{-i}$ the subspace of elements of *homological degree* i. Note that the *linear dual* $V^* = \mathrm{Hom}(V, \mathbb{R})$ of V is graded by $(V^*)^i = \mathrm{Hom}(V_i, \mathbb{R})$.

The *p-fold suspension* $V[p]$ of V for an integer p is defined by

$$V[p]^i := V^{i+p},$$

and elements of $V[p]^i$ are presented by $x\sigma$ for $x \in V^{i+p}$ using the symbol σ of cohomological degree $-p$. The p-suspension map $V \to V[p]$ is also denoted by σ. In this paper, the N-suspension σ for an even number N often appears. It is used for adjusting degrees of elements though we can ignore it when calculating signs.

Let V be a $\mathbb{Z}$-graded vector space and $\alpha : V \otimes V \to K$ be a nondegenerate bilinear map of (cohomological) degree n. Out of the two conditions

(i) $\alpha(x, y) = (-1)^{xy}\alpha(y, x)$ for homogeneous elements $x, y \in V$, and

(ii) $\alpha(x, y) = -(-1)^{xy}\alpha(y, x)$ for homogeneous elements $x, y \in V$,

the pair (V, α) is called a *symmetric vector space* with degree n if satisfying (i), and a *symplectic vector space* with degree n if satisfying (ii).

2A. *Algebras and signs.* Let V be a finite-dimensional $\mathbb{Z}$-graded vector space.

Definition 2.1. We define the following quotient algebras of the tensor algebra TV generated by V.

- The *symmetric algebra* SV generated by V is the $\mathbb{Z}$-graded commutative algebra which is the quotient algebra obtained from the $\mathbb{Z}$-graded tensor algebra TV by introducing the relation

  $$xy = (-1)^{xy} yx$$

  for $x, y \in V$. The image of $V^{\otimes k}$ for an integer k in SV is denoted by $S^k V$.
- The *exterior algebra* ΛV generated by V is the $\mathbb{Z}$-graded anticommutative algebra which is the quotient algebra obtained from the $\mathbb{Z}$-graded tensor algebra TV by introducing the relation

  $$xy = -(-1)^{xy} yx$$

  for $x, y \in V$. The image of $V^{\otimes k}$ for an integer k in ΛV is denoted by $\Lambda^k V$.

Definition 2.2. For distinct elements $v_1, \dots, v_k \in V$ and a permutation $\pi \in \mathfrak{S}_k$, the sign ϵ defined by the equation on $S^k V$

$$v_1 \cdots v_k = \epsilon \cdot v_{\pi(1)} \cdots v_{\pi(k)}$$

is called the *Koszul sign* of $(v_1, \dots, v_k) \mapsto (v_{\pi(1)}, \dots, v_{\pi(k)})$. Similarly the sign $\bar{\epsilon}$ defined by the same equation in $\Lambda^k V$ is called the *anti-Koszul sign*. Note the equation $\bar{\epsilon} = \operatorname{sgn} \pi \cdot \epsilon$.

2B. ***Derivations.*** Let W be a finite-dimensional $\mathbb{Z}$-graded vector space.

Completed tensor algebras. We denote the *completed tensor algebra* by

$$\widehat{T}W := \prod_{r=0}^{\infty} W^{\otimes r}.$$

Its product μ and coproduct Δ are defined by

$$\mu(x_1 \otimes \cdots \otimes x_s, x_{s+1} \otimes \cdots \otimes x_r) = x_1 \otimes \cdots \otimes x_r,$$

$$\Delta(x_1 \otimes \cdots \otimes x_r) = \sum_{s=0}^{r} \sum_{\tau \in \mathrm{Ush}(s, r-s)} \epsilon \cdot (x_{\tau(1)} \otimes \cdots \otimes x_{\tau(s)}) \otimes (x_{\tau(s+1)} \otimes \cdots \otimes x_{\tau(r)})$$

for homogeneous elements $x_1, \dots, x_r \in W$, where $\mathrm{Ush}(s, r-s)$ is the set of $(r, s-r)$-unshuffles and ϵ is the Koszul sign of the permutation $(x_1, \dots, x_r) \mapsto (x_{\tau(1)}, \dots, x_{\tau(r)})$ (Definition 2.2). The primitive part of $\widehat{T}W$ is the completed free Lie algebra $\hat{L}W$. These algebras have the gradings defined by the grading of W.

Derivations on a completed tensor algebra. Let $\mathrm{Der}(\hat{L}W)$ be the Lie algebra of (continuous) derivations on the completed algebra $\widehat{T}W$. Given a symplectic inner product ω of degree N on W, we define the Lie algebra of *symplectic derivations* on $\widehat{T}W$

$$\mathrm{Der}_\omega(\widehat{T}W) := \{D \in \mathrm{Der}(\widehat{T}W);\ D(\omega) = 0\}.$$

Here ω is identified with the element of $\hat{L}W$ described by

$$\sum_{i<j} \omega_{ij} [x^i, x^j],$$

where $\{x^i\}$ is a basis of W and matrix $(\omega_{ij})_{i,j}$ is the inverse matrix of $(\omega(x^i, x^j))_{i,j}$.

Since derivations on $\widehat{T}W$ are determined by the values on the generating space W, we get the isomorphism as graded vector space

$$\Phi_\omega : \mathrm{Der}(\widehat{T}W) \simeq \mathrm{Hom}(W, \widehat{T}W) \simeq \widehat{T}W \otimes W[-N] = \prod_{r=1} W^{\otimes r}[-N],$$

where the second isomorphism is induced by the isomorphism $\mathrm{Hom}(W,\mathbb{R}) \simeq W[-N]$ derived from nondegeneracy of ω. Furthermore, we also have the identification by Φ_ω

$$\begin{aligned}\mathrm{Der}^r(\widehat{T}W) &:= \{D \in \mathrm{Der}(\widehat{T}W);\ D(W) \subset W^{\otimes(r+1)}\}\\ &\simeq \mathrm{Hom}(W, W^{\otimes(r+1)}) \simeq W^{\otimes(r+2)}[-N].\end{aligned}$$

Fixing a homogeneous basis $x^1,\dots,x^m$ of W, the derivations $x^{i_1}\cdots x^{i_k}\,\partial/\partial x^i$ ($1 \le i_1,\dots,i_k, i \le m$), which are these elements corresponding to the linear map $x^i \mapsto x^{i_1}\cdots x^{i_k}$, comprise a basis of $\mathrm{Der}^{k+1}(\widehat{T}W)$. On the basis, Φ_ω is described by

$$\Phi_\omega\left(x^{i_1}\cdots x^{i_k}\frac{\partial}{\partial x^i}\right) = \sum_j \omega_{ij} x^{i_1}\cdots x^{i_k} x^j \sigma^{-1},$$

where σ^{-1} is a symbol of the $(-N)$-suspension which has homological degree $-N$.

By the identification Φ_ω, the space of symplectic derivations is described by

$$\mathrm{Der}_\omega(\widehat{T}W) \overset{\Phi_\omega}{\simeq} \prod_{r=1}^{\infty}(W^{\otimes r})^{\mathbb{Z}/r\mathbb{Z}}[-N] = \prod_{r=1}^{\infty} W^{(r)}_{\mathrm{cyc}}[-N].$$

Here $W^{(r)}_{\mathrm{cyc}} := (W^{\otimes r})^{\mathbb{Z}/r\mathbb{Z}}$ is the space of invariant tensors by cyclic permutations of tensor factors, which is also defined in Definition 3.2.

Therefore, the Lie algebra $\mathrm{Der}_\omega(\hat{L}W)$ of symplectic derivations on $\hat{L}W$ is described by

$$\begin{aligned}\mathrm{Der}_\omega(\hat{L}W) &:= \mathrm{Der}(\hat{L}W) \cap \mathrm{Der}_\omega(\widehat{T}W) \overset{\Phi_\omega}{\simeq} \prod_{r=2}^{\infty} W(r)[-N],\\ \mathrm{Der}^{r+2}_\omega(\hat{L}W) &:= \mathrm{Der}^{r+2}(\hat{L}W) \cap \mathrm{Der}_\omega(\widehat{T}W) \overset{\Phi_\omega}{\simeq} W(r)[-N],\end{aligned}$$

where $W(r) := (LW \otimes W) \cap W^{(r)}_{\mathrm{cyc}}$.

Through the isomorphism Φ_ω, the Lie algebra structure of $\mathrm{Der}(\widehat{T}W)$ is described as follows:

Lemma 2.3. *Let* $[\,\cdot\,,\cdot\,]$ *be the Lie bracket of* $\mathrm{Der}(\widehat{T}W)$. *Then the linear map* $[\,\cdot\,,\cdot\,]_\omega := \sigma\Phi_\omega \circ [\,\cdot\,,\cdot\,] \circ (\Phi_\omega^{-1}\sigma^{-1})^{\otimes 2}$ *is equal to*

$$\sum_{d_1+d_2=N} (\mathrm{id} \otimes \omega_{(d_1,d_2)})\left(\sum_{1\le t<r_2} \pi^{r_1,r_2}_{1;t} + \sum_{1\le s<r_1} \pi^{r_1,r_2}_{2;s}\right) : W^{\otimes r_1} \otimes W^{\otimes r_2} \to W^{\otimes r_1+r_2-2}$$

where $\omega_{(d_1,d_2)} : W \otimes W \to \mathbb{R}$ *for integers* d_1, d_2 *is the composition of the projection* $W \otimes W \to W_{d_1} \otimes W_{d_2}$ *and the restriction of* ω *to* $W_{d_1} \otimes W_{d_2}$, *and* $\pi^{r_1,r_2}_{1;j}, \pi^{r_1,r_2}_{2;i}$:

$W^{\otimes r_1} \otimes W^{\otimes r_2} \to W^{\otimes r_1+r_2}$ for $1 \le i \le r_1$ and $1 \le j \le r_2$ is defined by

$$\pi^{r_1,r_2}_{1;j}(a^{(1)}_1 \cdots a^{(1)}_{r_1} \otimes a^{(2)}_1 \cdots a^{(2)}_{r_2}) = \epsilon \cdot a^{(2)}_1 \cdots a^{(2)}_{j-1} a^{(1)}_1 \cdots a^{(1)}_{r_1-1} a^{(2)}_{j+1} \cdots a^{(2)}_{r_2} a^{(1)}_{r_1} a^{(2)}_j,$$
$$\pi^{r_1,r_2}_{2;i}(a^{(1)}_1 \cdots a^{(1)}_{r_1} \otimes a^{(2)}_1 \cdots a^{(2)}_{r_2}) = \epsilon \cdot a^{(1)}_1 \cdots a^{(1)}_{i-1} a^{(2)}_1 \cdots a^{(2)}_{r_2-1} a^{(2)}_{j+1} \cdots a^{(2)}_{r_2-1} a^{(1)}_i a^{(2)}_{r_2}$$

for homogeneous elements $a^{(1)}_1, \dots, a^{(1)}_{r_1}, a^{(2)}_1, \dots, a^{(2)}_{r_2}$. Here ϵ is the Koszul sign of the corresponding permutations.

Proof. Let $x^1, \dots, x^m$ be a homogeneous basis of W. The Lie bracket for the basis is described by

$$\left[x^{i_1} \cdots x^{i_k} \frac{\partial}{\partial x^i}, x^{j_1} \cdots x^{j_l} \frac{\partial}{\partial x^j}\right] = \sum_t \epsilon \delta^{j_t}_i x^{j_1} \cdots x^{j_{t-1}} x^{i_1} \cdots x^{i_k} x^{j_{t+1}} \cdots x^{j_l} \frac{\partial}{\partial x^j} - \sum_s \epsilon' \delta^{i_s}_j x^{i_1} \cdots x^{i_{s-1}} x^{j_1} \cdots x^{j_l} x^{i_{s+1}} \cdots x^{i_k} \frac{\partial}{\partial x^i}$$

where $\epsilon = (-1)^{(x^{i_1}+\cdots+x^{i_k}-x^i)(x^{j_1}+\cdots+x^{j_{t-1}})}$, $\epsilon' = (-1)^{(x^{j_1}+\cdots+x^{j_l}-x^j)(x^{i_{s+1}}+\cdots+x^{i_k}-x^i)}$, and δ^i_j is the Kronecker's delta.

Then, for $A = x^{i_1} \cdots x^{i_{r_1}}$ and $B = x^{j_1} \cdots x^{j_{r_2}}$, we obtain

$$\begin{aligned}[A, B]_\omega &= \sum_t \epsilon x^{j_1} \cdots x^{i_1} \cdots x^{i_{r_1-1}} \cdots x^{j_{r_2}} \omega^{i_{r_1} j_t} + \sum_s \epsilon' x^{i_1} \cdots x^{j_1} \cdots x^{j_{r_2-1}} \cdots x^{i_{r_1}} \omega^{i_s j_{r_2}} \\ &= \sum_{d_1+d_2=N} (\mathrm{id} \otimes \omega_{(d_1,d_2)}) \Bigg(\sum_{1 \le t < r_2} \pi^{r_1,r_2}_{1;t} + \sum_{1 \le s < r_1} \pi^{r_1,r_2}_{2;s} \Bigg)(A \otimes B),\end{aligned}$$

where ϵ and ϵ' are the Koszul signs of

$$(x^{i_1}, \dots, x^{i_{r_1}}, x^{j_1}, \dots, x^{j_{r_2}}) \mapsto (x^{j_1}, \dots, x^{i_1}, \dots, x^{i_{r_1-1}}, \dots, x^{j_{r_2}}, x^{i_{r_1}}, x^{j_t}),$$
$$(x^{i_1}, \dots, x^{i_{r_1}}, x^{j_1}, \dots, x^{j_{r_2}}) \mapsto (x^{i_1}, \dots, x^{j_1}, \dots, x^{j_{r_2-1}}, \dots, x^{i_{r_1}}, x^{i_s}, x^{j_{r_2}}),$$

respectively. In the calculus above, note that we use the assumption that N is even. □

The lemma above is needed to prove Theorem 3.9.

Derivations on a dgl. Let δ be an element in $\mathrm{Der}_\omega(\hat{L}W)$ of homological degree -1 such that $\delta^2 = 0$. Then $\mathrm{ad}(\delta)$ is a differential operator on $\mathrm{Der}_\omega(\hat{L}W)$.

In the case that W_0 is positively graded, i.e., $W_i = 0$ for $i \le 0$, we can regard that $\delta \in \mathrm{Der}_\omega(LW)$ since δ is described by only finite sums. Then we often consider the positive truncation $(\mathrm{Der}^+_\omega(LW), \mathrm{ad}(\delta))$ of the chain complex $(\mathrm{Der}_\omega(LW), \mathrm{ad}(\delta))$ defined by

$$\mathrm{Der}^+_\omega(LW)_i := \begin{cases} \mathrm{Der}_\omega(LW)_i & (i > 2), \\ \mathrm{Ker}(\mathrm{ad}(\delta))_1 & (i = 1), \\ 0 & (\text{otherwise}). \end{cases}$$

Definition 2.4 (Chevalley–Eilenberg complex). Let (L, δ) be a dgl. We define the Chevalley–Eilenberg complex as

$$C^{p,q}_{\mathrm{CE}}(L) := (\Lambda^p L^*)^q,$$

where $\Lambda^\bullet L^*$ is the exterior algebra generated by the graded vector space L^*. The first differential d_{CE} is defined by the formula for $c \in C^{p,q}_{\mathrm{CE}}(L)$ and $D_1, \ldots, D_{p+1} \in L$,

$$(d_{\mathrm{CE}}c)(D_1, \ldots, D_{p+1}) = \sum_{i<j} \bar{\epsilon} \cdot c([D_i, D_j], D_1, \ldots, \widehat{D_i}, \ldots, \widehat{D}_j, \ldots, D_{p+1}),$$

where $\bar{\epsilon} = (-1)^{D_i(D_1+\cdots+D_{i-1})+D_j(D_1+\cdots+D_{j-1})+D_iD_j+i+j-1}$, and the second differential L_δ derived from δ is defined by

$$L_\delta = i_\delta d_{\mathrm{CE}} - d_{\mathrm{CE}} i_\delta,$$

using the *interior product* defined by

$$(i_\delta c)(D_1, \ldots, D_p) = c(\delta, D_1, \ldots, D_p),$$

for $c \in C^{p+1,q}_{\mathrm{CE}}(L)$ and $D_1, \ldots, D_p \in L$. Then the triple $(C^{\bullet;\bullet}_{\mathrm{CE}}(L), d_{\mathrm{CE}}, L_\delta)$ is a double complex.

We will consider the Chevalley–Eilenberg complexes of dgls $(\mathrm{Der}_\omega(\hat{L}W), \mathrm{ad}(\delta))$ and $(\mathrm{Der}^+_\omega(LW), \mathrm{ad}(\delta))$, and the invariant space $C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(LW))^{\mathrm{Sp}(W,\delta)}$, where $\mathrm{Sp}(W, \delta)$ is the group of symplectic linear isomorphisms $W \to W$ preserving δ.

2C. ***A dgl model with symplectic form of manifolds.*** In this subsection, we review a Chen dgl model of a manifold. Let X be a smooth manifold. Put $A = A^\bullet(X)$ and $H = H^\bullet_{DR}(X)$. Fix a homotopy transfer diagram

$$\circlearrowright A \rightleftarrows H;$$

e.g., in the case that X is a closed manifold, it is obtained by using the Hodge decomposition of the de Rham complex A. Since A is a commutative differential graded algebra (dga) with symmetric form (intersection form), H has the structure of a minimal cyclic C_∞-algebra by the diagram (details in [Kontsevich and Soibelman 2001; Merkulov 1999; Kadeishvili 2009; Markl et al. 2002; Hamilton and Lazarev 2004] for instance).

Let I be the intersection form on H, m the cyclic C_∞-algebra structure on H obtained by the homotopy transfer diagram, and $s : H \to H[1]$ the suspension map. We denote $V = H[1]^*$. Defining the suspension of m_i by $\overline{m}_i := s \circ m_i \circ (s^{-1})^{\otimes i}$ for all $i \geq 1$ and of I by $\omega := I \circ (s^{-1})^{\otimes 2}$, then the duals of these define the symplectic inner product ω on $H[1]^*$ of degree $N = n - 2$ and the linear map $\bar{\delta}_i : V \to V^{\otimes n}$

of homological degree -1. Thus, extending the unique derivation $\bar{\delta}_i : \hat{L}V \to \hat{L}V$ by the Leibniz rule, then we have the derivation of homological degree -1

$$\bar{\delta} := \sum_{i=1}^{\infty} \bar{\delta}_i \in \mathrm{Der}_{\omega}(\hat{L}V).$$

Furthermore we can prove that $\bar{\delta}$ is a differential since m satisfies the A_∞-relations and *quadratic*, i.e., $\bar{\delta}(V) \subset \prod_{i\geq 2} V^{\otimes i}$, since (H, I, m) is minimal.

The Chen dgl model is a reduced version of the construction. Suppose X is connected, and put

$$W := H[1]^*_{\geq 0} = H_+(X; \mathbb{R})[-1].$$

Then we have the restriction $\delta : \hat{L}W \to \hat{L}W$ of $\bar{\delta}$ and $\omega : W^{\otimes 2} \to \mathbb{R}$. If X is simply connected, we can restrict the differential δ on the free Lie algebra $LW \subset \hat{L}W$ since $\delta(w)$ for $w \in W$ has only finitely many nontrivial terms.

Theorem 2.5 [Chen 1977]. *For a simply connected closed manifold X with base point $*$, the dgl (LW, δ) is a Quillen model of X, i.e., there is a Lie algebra isomorphism*

$$H_\bullet(LW, \delta) \simeq \pi_\bullet(\Omega X) \otimes \mathbb{Q}.$$

3. Graph complex

3A. ***Orientation and ordering of graded sets.*** The set of *orderings* on a set U is defined by

$$\mathrm{Ord}(U) := \{(u_1, \ldots, u_k) \in U^{\times k};\ U = \{u_1, \ldots, u_k\}\},$$

where $k := \#U$.

Definition 3.1. Let U be a $\mathbb{Z}$-graded set, i.e., a finite set U given a map $|\cdot| : U \to \mathbb{Z}$.

- The graded vector space generated by U is denoted by $\mathbb{R}U$.
- The symmetric algebra generated by U is denoted by $SU := S(\mathbb{R}U)$.
- The exterior algebra generated by U is denoted by $\Lambda U := \Lambda(\mathbb{R}U)$.

For an element $(u_1, \ldots, u_k) \in \mathrm{Ord}(U)$, we denote the image of $u_1 \otimes \cdots \otimes u_k$ in ΛU by $[u_1, \ldots, u_k]$. The 1-dimensional vector space generated by this element is written by

$$O(U) := \langle [u_1, \ldots, u_k] \rangle \subset \Lambda U.$$

Definition 3.2. Let V be a $\mathbb{Z}$-graded vector space. We define the subspace $V^{(k)}_{\mathrm{cyc}}$ of *cyclic tensors* in $V^{\otimes k}$ by the image of the map $[\,\cdot\,, \ldots, \cdot\,]_{\mathrm{cyc}} : V^{\otimes k} \to V^{\otimes k}$ obtained by

$$x_1 \otimes \cdots \otimes x_k \mapsto \sum_{\tau \in \mathbb{Z}/k\mathbb{Z}} \epsilon \cdot x_{\tau(1)} \otimes \cdots \otimes x_{\tau(k)},$$

where $\mathbb{Z}/k\mathbb{Z}$ is identified with the group of cyclic permutations and ϵ is the Koszul sign of $(x_1, \dots, x_k) \mapsto (x_{\tau(1)}, \dots, x_{\tau(k)})$. For a $\mathbb{Z}$-graded set U, we denote

$$\mathrm{Cyc}(U) := \langle [u_1, \dots, u_k]_{\mathrm{cyc}};\ (u_1, \dots, u_k) \in \mathrm{Ord}(U) \rangle \subset (\mathbb{R}U)^{(k)}_{\mathrm{cyc}}.$$

3B. *Definition of graph complex.* Let W be a finite-dimensional symplectic vector space with form ω of degree N and suppose N is even and $Z := \{a \in \mathbb{Z};\ W_a \neq 0\} \subset \{0, \dots, N\}$. Our labeled graph complex depends on (W, ω).

Definition of graphs.

Definition 3.3. An *N-graded graph* Γ consists of the following information:

- The set $H(\Gamma)$ of *half-edges*.
- The set $V(\Gamma)$ of *vertices*. It is a partition of the set $H(\Gamma)$; i.e.,

  $$H(\Gamma) = \coprod_{v \in V(\Gamma)} v, \quad v \neq \varnothing \quad (v \in V(\Gamma)).$$

  The number $\#v$ of elements of any $v \in V(\Gamma)$ is called the *valency* of v. A vertex with valency > 1 is called an *internal vertex*, and one with valency 1 is called an *external vertex*. The sets of internal and external vertices are denoted by $V_i(\Gamma)$ and $V_e(\Gamma)$, respectively.
- The set $E(\Gamma)$ of *edges*. It is a partition of the set $H(\Gamma)$ such that the number of elements of any $e \in E(\Gamma)$ is two, i.e.,

  $$H(\Gamma) = \coprod_{e \in E(\Gamma)} e, \quad \#e = 2 \quad (e \in E(\Gamma)).$$

- The cohomological *degree of half-edges*. It is a map $|\cdot| : H(\Gamma) \to Z$ such that $|h_1| + |h_2| = N$ for an edge $e = \{h_1, h_2\} \in E(\Gamma)$. Then the cohomological degrees of vertices and edges are defined by

  $$|v| := |h_1| + \cdots + |h_r| - N, \quad |e| := N,$$

  for $v = \{h_1, \dots, h_r\} \in V(\Gamma)$ and $e \in E(\Gamma)$.
- The division of the set $V_i(\Gamma)$ of internal vertices to two disjoint sets

  $$V_i(\Gamma) = V_n(\Gamma) \amalg V_s(\Gamma)$$

  such that all elements in $V_s(\Gamma)$ have cohomological degree -1 and valency ≥ 3. An element of $V_n(\Gamma)$ is called a *normal vertex*, and one of $V_s(\Gamma)$ is called a *special vertex*.

The set of isomorphism classes of such graphs is denoted by $\mathcal{G}(N)$. Here an isomorphism between N-graded graphs is a bijection between the sets of half-edges preserving all information of N-graded graphs.

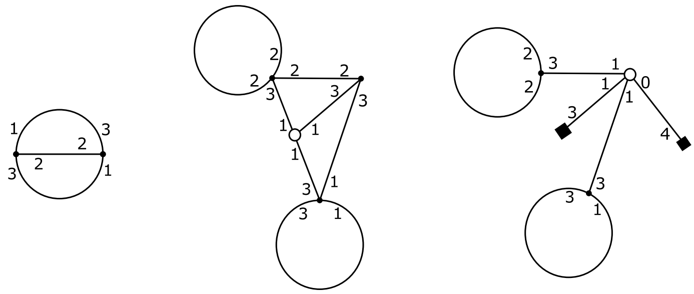


Figure 1. Examples of 4-graded graphs.

Example 3.4. In the case of $N = 4$ and $Z = \{0, 1, 2, 3, 4\}$, we can give examples of 4-graded graphs in Figure 1. In these figures,

- a black vertex $\bullet$ means a normal vertex, a white vertex $\circ$ a special vertex, and a square vertex $\blacksquare$ a univalent vertex, and
- a number drawn beside a half-edge is its degrees.

Decoration on vertices. We shall give the relation equivalent to the dual of vertices defined by the cyclic Lie operad as in [Conant and Vogtmann 2003; Hamilton 2006; Markl 1999].

Definition 3.5. Let Γ be an N-graded graph.

- We introduce to $\mathrm{Cyc}(v)[N]$ for $v \in V_i(\Gamma)$ the *commutativity relation*
$$S_{v,h_r;s}(o) := \sum_{\tau \in \mathrm{Sh}(s,r-s-1)} o^{\tau^{(v,h_r)}} = 0,$$
$$o^{\tau^{(v,h_r)}} := \epsilon[h_{\tau(1)}, \dots, h_{\tau(r-1)}, h_r]_{\mathrm{cyc}}\sigma,$$
for $r-1 > s > 0$ and $o = [h_1, \dots, h_r]_{\mathrm{cyc}}\sigma \in \mathrm{Cyc}(v)[N]$, where $\mathrm{Sh}(p, q)$ is the set of (p, q)-shuffles, σ is the symbol of the N-fold suspension, and ϵ is the Koszul sign. Then we denote by $C(v) = \mathrm{Cyc}(v)[N]/(\text{com. rel.})$ the obtained space. This relation for $r = 3, 4$ is described in Figure 2. (In the case of $r = 3$, it is the AS-relation for Jacobi diagrams.)

Decoration on N-graded graphs. Set
$$\widetilde{O}_{\mathrm{com}}(W, \Gamma) := \bigodot_{e \in E(\Gamma)} O(e) \otimes \bigodot_{u \in V_e(\Gamma)} W[-N]_{|u|} \otimes \bigwedge_{v^s \in V_s(\Gamma)} C(v^s) \otimes \bigwedge_{v \in V_n(\Gamma)} C(v),$$
where
$$\bigodot_{u \in U} V(u) := \left\{ v_{u_1} \cdots v_{u_k} \in S^k\Bigl(\bigoplus_{u \in U} V(u)\Bigr);\, v_{u_i} \in V(u_i),\, (u_1, \dots, u_k) \in \mathrm{Ord}(U) \right\},$$
$$\bigwedge_{u \in U} V(u) := \left\{ v_{u_1} \cdots v_{u_k} \in \Lambda^k\Bigl(\bigoplus_{u \in U} V(u)\Bigr);\, v_{u_i} \in V(u_i),\, (u_1, \dots, u_k) \in \mathrm{Ord}(U) \right\}$$

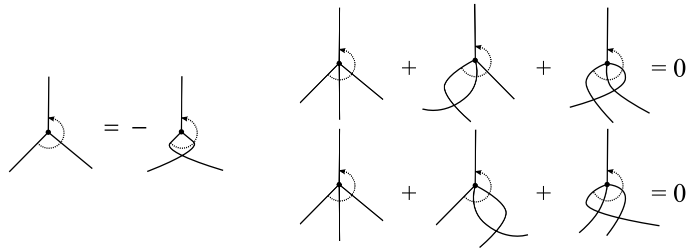


Figure 2. Commutativity ($r = 3, 4$). (Koszul signs are omitted in figures.)

for a family $(V(u))_{u\in U}$ of $\mathbb{Z}$-graded vector spaces indexed by a finite set U. This tensor product consists of four factors: the first factor means directions of edges of Γ, the second factor W-labels of external vertices of Γ, the third factor (equivalence classes of) cyclic orderings on special vertices of Γ, and the fourth factor the same on normal vertices of Γ. Note that $W[-N]_{|u|} = W_{|h|}[-N]$ for an external vertex $u = \{h\}$.

We need to identify elements of $\widetilde{O}_{\mathrm{com}}(W, \Gamma)$ by the symmetry of Γ. An automorphism α of an N-graded graph $\Gamma \in \mathcal{G}(N)$ induces the linear isomorphism $C(v) \to C(\alpha(v))$ for $v \in V_i(\Gamma)$ described by

$$[h_1, \dots, h_k]_{\mathrm{cyc}} \mapsto [\alpha(h_1), \dots, \alpha(h_k)]_{\mathrm{cyc}},$$

and the identity map $W[-N]_{|u|} \to W[-N]_{|\alpha(u)|} = W[-N]_{|u|}$ for $u \in V_e(\Gamma)$. Therefore, the automorphism group of Γ acts on the vector space $\widetilde{O}_{\mathrm{com}}(W, \Gamma)$ by the induced permutation of half-edges. Then the coinvariant vector space of $\widetilde{O}_{\mathrm{com}}(W, \Gamma)$ by this action is denoted by $O_{\mathrm{com}}(W, \Gamma)$. We often consider an element o of $O_{\mathrm{com}}(W, \Gamma)$ described by the form

$$\begin{aligned} o &= [o_1, \dots, o_l; w_1, \dots, w_{k_e}; c^s_1, \dots, c^s_{k_s}; c_1, \dots, c_{k_n}] \\ &:= (o_1 \cdots o_l) \otimes (w_1 \cdots w_{k_e}) \otimes (c^s_1 \cdots c^s_{k_s}) \otimes (c_1 \cdots c_{k_n}) \end{aligned}$$

where $w_i \in W[-N]_{|u_i|}$ and

$$o_i = [\hat{o}_i], \qquad c^s_i = [\hat{c}^s_i]_{\mathrm{cyc}}\sigma, \qquad c_i = [\hat{c}_i]_{\mathrm{cyc}}\sigma,$$

for $\hat{o}_i \in \mathrm{Ord}(e_i)$, $\hat{c}_i \in \mathrm{Ord}(v_i)$, and $\hat{c}^s_i \in \mathrm{Ord}(v^s_i)$. Such element o is called an *orientation* of Γ, a pair (Γ, o) is an *oriented graph*, and the information

$$\hat{o} = (\hat{o}_1, \dots, \hat{o}_l; w_1, \dots, w_{k_e}; \hat{c}^s_1, \dots, \hat{c}^s_{k_s}; \hat{c}_1, \dots, \hat{c}_{k_n})$$

is called a *lift* of an orientation $o = [\hat{o}]$ on Γ. The vector space $O_{\mathrm{com}}(W, \Gamma)$ is generated by orientations.

Example 3.6. In the case of $N = 4$ and $Z = \{0, 1, 2, 3, 4\}$, we can give examples of decorated 4-graded graphs in Figure 3. In these figures,

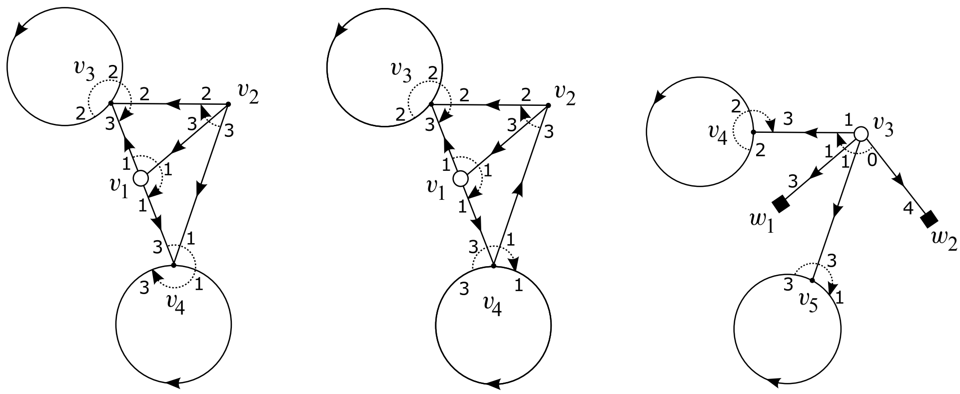


Figure 3. Left and center: nonlabeled examples (Γ, o_1) and (Γ, o_2). The ordering of vertices is $v_1v_2v_3v_4$ and $v_1v_2v_4v_3$. Right: a labeled example. The ordering of vertices and labels is $(w_1\omega^{-1})(w_2\omega^{-1})v_3v_4v_5$.

- an arrow on an edge means a direction, and
- an arc drawn around a vertex is an ordering of half-edges incident to this vertex.

In Figure 3, left and center, the degrees of vertices are $v_1 = -1$, $v_2 = 4$, $v_3 = 5$, and $v_4 = 4$. In the space $O(\Gamma)$, we have

$$o_1 = (-1)^{5\cdot 4+1}(-1)^{3\cdot 1+1}(-1)^{3\cdot(3+1+1)}o_2 = o_2,$$

where the signs $(-1)^{5\cdot 4+1}$, $(-1)^{3\cdot 1+1}$, $(-1)^{3\cdot(3+1+1)}$ are coming from changes of the ordering of vertices, the direction of the edge between v_2 and v_4, and the ordering of half-edges incident to v_4, respectively.

In Figure 3, right, elements $w_1 \in W_3$ and $w_2 \in W_4$ are labels of univalent vertices v_1, v_2 (their names v_1, v_2 of vertices are omitted in the figure). Note their degrees $|v_1| = |w_1\sigma^{-1}| = -1$ and $|v_2| = |w_2\sigma^{-1}| = 0$.

Definition of the bigraded vector space $\widehat{C}^{\bullet,\bullet}_{\mathrm{com}}(W)$. The cohomological bidegree $(p, q) \in \mathbb{Z} \times \mathbb{Z}$ of $\Gamma \in \mathcal{G}(N)$ is defined by

$$p = \#V_n(\Gamma), \qquad q = \sum_{v \in V_n(\Gamma)} |v| = \#V_s(\Gamma) + N(\#E(\Gamma) - \#V(\Gamma)) - \sum_{u \in V_e(\Gamma)} |u|,$$

and the bidegree of elements in $O_{\mathrm{com}}(W, \Gamma)$ is defined by that of Γ. We define *the space of N-graded ribbon graphs* by

$$\widehat{C}^{\bullet,\bullet}_{\mathrm{com}}(W) := \bigoplus_{\Gamma \in \mathcal{G}(N)} O_{\mathrm{com}}(W, \Gamma), \qquad \widehat{C}^{p,q}_{\mathrm{com}}(W) := \bigoplus_{\Gamma \in \mathcal{G}^{p,q}(W)} O_{\mathrm{com}}(W, \Gamma),$$

where $\mathcal{G}^{p,q}(W)$ is the subset of $\mathcal{G}(N)$ consisting of N-graded graphs of degree (p, q). Then $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ can be regarded as a bigraded vector space. We often denote an element in $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ corresponding to $o \in O_{\mathrm{com}}(W, \Gamma)$ for $\Gamma \in \mathcal{G}(N)$ by (Γ, o).

Definition of the first differential d. We define a linear map $d^{a,b}_{v;h^1,h^2} : O_{\mathrm{com}}(W, \Gamma) \to \widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ for an N-graded graph $\Gamma \in \mathcal{G}(N)$, a normal vertex $v \in V_n(\Gamma)$, two distinct half-edges h^1, h^2 incident to v, and $a, b \in Z$ satisfying $a + b = N$. For an order $h_1, \ldots, h_r$ of half-edges incident to v such that $h^1 = h_r$ and $h^2 = h_i$, put

$$\begin{aligned} &d^{a,b}_{v;h^1,h^2}(\Gamma, [\,\cdot\,;\cdot\,;\cdot\,;[h_1, \ldots, h_r]\sigma, \cdot\,]) \\ &\qquad = (\Gamma^{a,b}_{v;h^1,h^2}, [\,\cdot\,, [h', h'']; \cdot\,; \cdot\,; [h_1, \ldots, h_i, h']\sigma, [h'', h_{i+1}, \ldots, h_r]\sigma, \cdot\,]). \end{aligned}$$

Here σ is the N-fold suspension, and the N-graded graph $\Gamma^{a,b}_{v;h^1,h^2}$ is defined by

$$\begin{aligned} H(\Gamma^{a,b}_{v;h^1,h^2}) &= H(\Gamma) \amalg \{h', h''\}, & V(\Gamma^{a,b}_{v;h^1,h^2}) &= (V(\Gamma) \setminus \{v\}) \amalg \{v', v''\}, \\ V_s(\Gamma^{a,b}_{v;h^1,h^2}) &= V_s(\Gamma), & E(\Gamma^{a,b}_{v;h^1,h^2}) &= E(\Gamma) \amalg \{e_0\}, \end{aligned}$$

where $v' = \{h_1, \ldots, h_i, h'\}$, $v'' = \{h'', h_{i+1}, \ldots, h_r\}$, $e_0 = \{h', h''\}$, $|h'| = a$, and $|h''| = b$. Note that the equation above is enough to define the operator $d^{a,b}_{v;h^1,h^2}$ and the operator is well-defined. A picture of the map $d^{a,b}_{v;h^1,h^2}$ is described in Figure 4.

Then we obtain the linear map $d : \widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W) \to \widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ by

$$d_v(\Gamma, o) := \frac{1}{2} \sum_{a+b=N} \sum_{h^1 \neq h^2 \in v} d^{a,b}_{v;h^1,h^2}(\Gamma, o), \qquad d(\Gamma, o) := \sum_{v \in V_n(\Gamma)} d_v(\Gamma, o).$$

The map d can be also described by

$$d_v(\Gamma, o) = \sum_{a+b=N} \sum_{0 \leq s < t < r} d^{a,b}_{v;h_s,h_t}(\Gamma, o),$$

where $o = [\,\cdot\,;\cdot\,;\cdot\,;[h_1, \ldots, h_r]\sigma, \cdot\,]$ and $v = \{h_1, \ldots, h_r\}$. Note the relation

$$d^{a,b}_{v;h^1,h^2}(\Gamma, o) = d^{b,a}_{v;h^2,h^1}(\Gamma, o)$$

for half-edges $h^1 \neq h^2 \in v$. Here well-definedness of d is proved by the relation with the commutativity relation:

Proposition 3.7. *Using the notations above, $d_v S_{v,h_r;i}(\Gamma, o)$ is equal to zero under the commutativity relation.*

Proof. For integers p, q, we define the linear ordered set $[p, q]$ by $\{p < p+1 < \cdots < q-1 < q\}$. If $p > q$, put $[p, q] = \varnothing$. For partially ordered sets P_1, P_2, we denote their direct sum by $P_1 + P_2$ (in the category of posets), and their ordinal sum by $P_1 \oplus P_2$. Then a (p, q)-shuffle is equivalent to the inverse of an order-preserving bijection $[1, p] + [p+1, p+q] \to [1, p+q]$.

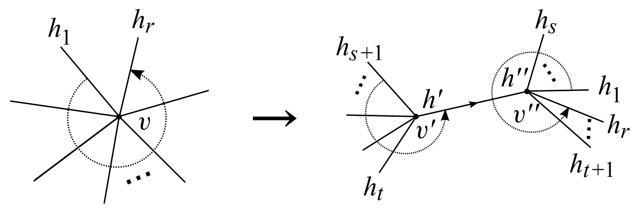


Figure 4. The operator $d^{a,b}_{v;h_s,h_t}$.

Let $\tau^{-1} : [1, i] + [i + 1, r - 1] \to [1, r - 1]$ be an $(i, r - i - 1)$-shuffle and $0 \le s < t < r$ integers. Put $L = \tau([s + 1, t])$ and $l = t - s$.

If $\tau(s + 1), \dots, \tau(t)$ are $\le i$, then we have $\tau(s + m) = \tau(s + 1) + (m - 1)$ for $1 \le m \le t - s$ since $[1, i] \to \tau^{-1}([1, i])$ is an isomorphism between posets. Put $a = \tau(s + 1) - 1$. Then we obtain the shuffle τ_2 by τ:

$$\begin{tikzcd}
{[1, i-l+1] + [i-l+2, r-l]} \arrow[r, "\tau_2^{-1}"] & {[1, r-l]} \\
{[1, a] \oplus \{*\} \oplus [a+l, i] + [i+1, r-1]} \arrow[r, "\text{bij.}"] \arrow[u, "\text{canonical isom.}"] & {[1, s] \oplus \{*\} \oplus [t+1, r-1]} \arrow[u, "\text{canonical isom.}"'] \\
{[1, i] + [i+1, r-1]} \arrow[r, "\tau^{-1}"] \arrow[u, two heads] & {[1, r-1]} \arrow[u, two heads]
\end{tikzcd}$$

The shuffle τ can be recovered from a pair (a, l, τ_2), where $\{a+1, \dots, a+l\} \subset [1, i]$ and τ_2 is an $(i - l + 1, r - i - 1)$-shuffle.

Similarly, if $\tau(s+1), \dots, \tau(t)$ are $\ge i+1$, we can obtain a triple (a, l, τ_2), where $\{a + 1, \dots, a + l\} \subset [i + 1, r - 1]$ and τ_2 is an $(i - l + 1, r - i - 1)$-shuffle.

Otherwise, put $p = \#(L \cap [1, i])$. Then we obtain the shuffle τ_1 by restricting τ:

$$\begin{tikzcd}
{[1, p] + [p+1, l]} \arrow[r, "\tau_1^{-1}"] \arrow[d, "\text{canonical isom.}"'] & {[1, l]} \arrow[d, "\text{canonical isom.}"] \\
L \arrow[r, "\tau^{-1}"] \arrow[d, hook] & \tau^{-1}(L) \arrow[d, hook] \\
{[1, i] + [i+1, r-1]} \arrow[r, "\tau^{-1}"] & {[1, r-1]}
\end{tikzcd}$$

We consider $\bar{L} = ([1, i] + [i + 1, r - 1]) \setminus L$ and the order-preserving bijection $\rho^{-1} : \bar{L} \to [1, s] \oplus [t + 1, r - 1]$ defined by the restriction of τ^{-1}. The shuffle τ is recovered from a pair (ρ, τ_1), where $\rho^{-1} : \bar{L} \to [1, s] \oplus [t + 1, r - 1]$ is an order-preserving bijection and τ_1 is a $(p, l - p)$-shuffle.

Thus, we have

$$\begin{aligned} d_v S_{v,h_r;i}([h_1,\dots,h_r]\sigma) &= \sum_{l=1}^{r-1}\Bigl(\sum_{p=1}^{l-1}\sum_{\rho}\sum_{\tau_1} o_\rho^{\tau_1^{(v',h')}} + \sum_a\sum_{\tau_2} o_{a,l}^{\tau_2^{(v'',h_r)}}\Bigr) \\ &= \sum_{l=1}^{r-1}\Bigl(\sum_{p=1}^{l-1}\sum_{\rho} S_{v',h';p}(o_\rho) + \sum_a S_{v'',h_r;i-l+1}(o_{a,l})\Bigr), \end{aligned}$$

where $L = \{1,\dots,r-1\}\setminus \bar{L} = \{u_1 < \cdots < u_p \text{ as integers}\}$,

$$\begin{aligned} o_\rho &= \epsilon\bigl[[h_{u_1},\dots,h_{u_p},h']\sigma, [h_{\rho(1)},\dots,h_{\rho(s)},h'',h_{\rho(t+1)},\dots,h_{\rho(r-1)},h_r]\sigma\bigr], \\ o_{a,l} &= \epsilon'\bigl[[h_{a+1},\dots,h_{a+l},h']\sigma, [h_1,\dots,h_a,h'',h_{a+l+1},\dots,h_r]\sigma\bigr], \end{aligned}$$

and ϵ, ϵ' are appropriate Koszul signs. (In these equations, the subscripts cyc are omitted.) $\square$

Definition of the second differential L. For $\Gamma \in \mathcal{G}(N)$, let $i_v(\Gamma)$ be the N-graded graph obtained by converting a normal vertex v of degree -1 to a special vertex. We define the linear map $i_v : O_{\text{com}}(W,\Gamma) \to O_{\text{com}}(W, i_v(\Gamma))$ for $o \in O_{\text{com}}(W,\Gamma)$ such that

$$i_v(\Gamma, [\,\cdot\,;\cdot\,;\cdot\,;c,\cdot\,]) = (i_v(\Gamma), [\,\cdot\,;\cdot\,;\cdot\,,c;\cdot\,])$$

for $c \in C(v)$ if v has degree -1 and valency ≥ 3, and $i_v(\Gamma, o) = 0$ if v does not. Since the relation

$$i_{v_1} S_{v_2,h_r;k}(\Gamma, o) = S_{v_2,h_r;k} i_{v_1}(\Gamma, o)$$

for $v_1, v_2 \in V_i(\Gamma)$ holds clearly, the map i_v is well-defined. Then the linear map $L : \widehat{C}^{\bullet;\bullet}_{\text{com}}(W) \to \widehat{C}^{\bullet;\bullet}_{\text{com}}(W)$ is defined by

$$L := id - di,$$

where the linear map $i : \widehat{C}^{\bullet;\bullet}_{\text{com}}(W) \to \widehat{C}^{\bullet;\bullet}_{\text{com}}(W)$ is obtained by

$$i(\Gamma, o) := \sum_{v \in V_n(\Gamma)} i_v(\Gamma, o).$$

The map L is also described by

$$L(\Gamma, o) = \sum_{v \in V_n(\Gamma)} (i_{v'} + i_{v''}) d_v(\Gamma, o)$$

since $i_u d_v = d_v i_u$ for normal vertices $u \neq v$.

Then d, i, and L have (cohomological) bidegree $(1,0)$, $(-1,1)$, and $(0,1)$, respectively.

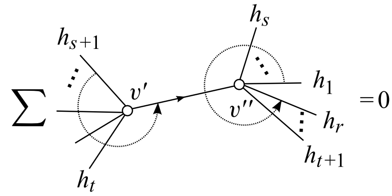


Figure 5. A_∞-relation.

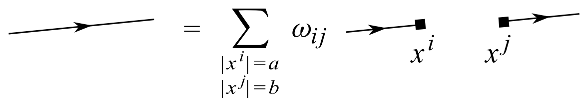


Figure 6. Cut-off relation.

Definition of the underlying bigraded vector space $C^{\bullet;\bullet}_{\mathrm{com}}(W)$. The space $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ is the quotient space of $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ by

- (*A_∞-relation*) $\qquad R_v(\Gamma, o) := i_{v'} i_{v''} d_v(\Gamma, o) = 0$

  for $\Gamma \in \mathscr{G}(N)$ and a normal vertex v (of degree -2). This relation is described in Figure 5.

- (*Cut-off relation*) For $\Gamma \in \mathscr{G}(N)$ and $e = \{h_1, h_2\} \in E(\Gamma)$, we define the N-graded graph Γ_e as

$$\begin{aligned} H(\Gamma_e) &= H(\Gamma) \amalg \{\bar{h}_1, \bar{h}_2\}, \\ E(\Gamma_e) &= (E(\Gamma) \setminus \{e\}) \amalg \{\{h_1, \bar{h}_1\}, \{h_2, \bar{h}_2\}\}, \\ V(\Gamma_e) &= V(\Gamma) \amalg \{\{\bar{h}_1\}, \{\bar{h}_2\}\}, \\ |\bar{h}_1| &= N - |h_1| =: a, \qquad |\bar{h}_2| = N - |h_2| =: b. \end{aligned}$$

  Then

$$(\Gamma, [[h_1, h_2], \cdot\,; \cdot\,; \cdot\,; \cdot\,]) = \sum_{\substack{|x^i|=a \\ |x^j|=b}} \omega_{ij}(\Gamma_e, [[h_1, \bar{h}_1], [\bar{h}_2, h_2], \cdot\,; x^i\sigma^{-1}, x^j\sigma^{-1}, \cdot\,; \cdot\,; \cdot\,]),$$

  where $\{x^i\}$ is a homogeneous basis of W and (ω_{ij}) is the inverse matrix of $(\omega(x^i, x^j))$. This relation is described in Figure 6.

Note that $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ is generated by W-labeled graphs with only one internal vertex by cut-off relation.

On well-definedness of three operators d, i, L on $C^{\bullet;\bullet}_{\mathrm{com}}(W)$. The endomorphisms d, i, and L of $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$ induce endomorphisms of $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ by the equations

$$dR_v(\Gamma, o) = \sum_{u \neq v} R_v d_u(\Gamma, o), \qquad iR_v(\Gamma, o) = \sum_{u \neq v} R_v i_u(\Gamma, o)$$

for a normal vertex v of an N-graded graph Γ.

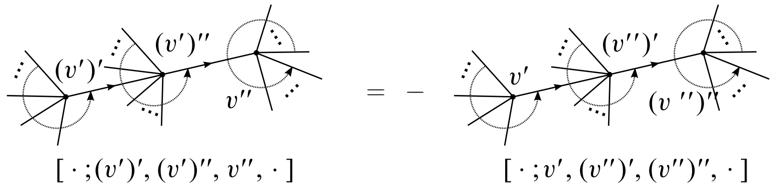


Figure 7. $d_{v'}d_v(\Gamma, o) = -d_{v''}d_v(\Gamma, o)$.

On two differentials d, L on $C^{\bullet;\bullet}_{\mathrm{com}}(W)$.

Proposition 3.8. *The bigraded vector space $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ is a double complex with respect to differentials d and L. We call $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ a* double graph complex.

Proof. First, we show the equation $d^2 = 0$. It is proved in the same way as Kontsevich's original graph complex. For a normal vertex v of an N-graded graph (Γ, o), let v', v'' be new vertices obtained by splitting at v. Then

$$d_{v'}d_v(\Gamma, o) = -d_{v''}d_v(\Gamma, o), \qquad d_u d_v(\Gamma, o) = -d_v d_u(\Gamma, o)$$

for $u \neq v$. The first equation is shown by Figure 7. In the figure, v' and v'' are defined such that the direction of the new edge is from v' to v'', and $(v')'$, $(v')''$, $(v'')'$, $(v'')''$ are also defined in the same way. So we obtain $d^2(\Gamma, o) = 0$ by cancellation.

Next, we show $L^2 = 0$. From the equation in $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$

$$\begin{aligned}(iL - Li)(\Gamma, o) &= \Bigl(\sum_u i_u(i_{v'} + i_{v''})d_v - \sum_{u\neq v}(i_{v'} + i_{v''})d_v i_u\Bigr)(\Gamma, o)\\ &= \sum_v (i_{v''}i_{v'} + i_{v'}i_{v''})d_v(\Gamma, o)\\ &= 2\sum_v R_v(\Gamma, o),\end{aligned}$$

we obtain the relation $iL - Li = 0$ in $C^{\bullet;\bullet}_{\mathrm{com}}(W)$. So the equations

$$\begin{aligned}L^2 &= (id - di)L = idL - diL = idL - dLi = \quad idid - didi,\\ L^2 &= L(id - di) = Lid - Ldi = iLd - Ldi = -idid + didi\end{aligned}$$

hold. Then we obtain $L^2 = 0$. Since $Ld + dL = -did + did = 0$ holds by definition of L, we get the proposition. □

3C. ***Construction of the map to Chevalley–Eilenberg complexes.*** Let (W, ω) and Z be as in Section 3B and δ be a symplectic and quadratic differential of homological degree -1 on $\hat{L}W$. In this section, the Lie algebra $\mathrm{Der}_\omega(\hat{L}W)$ of symplectic derivations is denoted by $\mathfrak{D}$. We construct a double chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{D})$$

from the graph complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ to the Chevalley–Eilenberg complex of the dgl $(\mathfrak{D}, \mathrm{ad}(\delta))$.

Let (Γ, o) be an oriented graph and $\hat{o}$ be a lift of o. Put

$$k = \#V(\Gamma), \qquad k_e = \#V_e(\Gamma), \qquad k_s = \#V_s(\Gamma), \qquad k_n = \#V_n(\Gamma),$$
$$\begin{aligned}(r_1, \dots, r_k) &:= (\underbrace{1, \dots, 1}_{k_e}, a_1, \dots, a_{k_s+k_n}) \\ &:= (\underbrace{1, \dots, 1}_{k_e}, \#v^s_1, \dots, \#v^s_{k_s}, \#v_1, \dots, \#v_{k_n}).\end{aligned}$$

We denote by $\tau(\hat{o})$ the linear isomorphism (the permutation of factors of the tensor product)

$$W^{\otimes r_1} \otimes \cdots \otimes W^{\otimes r_k} \to W^{\otimes 2} \otimes \cdots \otimes W^{\otimes 2} = (W^{\otimes 2})^{\otimes l}$$

corresponding to the permutation of half-edges

$$(h_1, \dots, h_{k_e}, \hat{c}^s_1, \dots, \hat{c}^s_{k_s}, \hat{c}_1, \dots, \hat{c}_{k_n}) \mapsto (\hat{o}_1, \dots, \hat{o}_l).$$

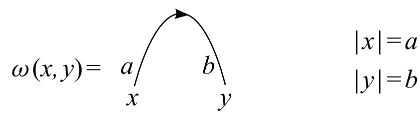
Then we define the linear map $\alpha(\Gamma, \hat{o})$ of cohomological degree $(l-k)N$ by composing these maps

$$\begin{aligned}\alpha(\Gamma, \hat{o}) : W[-N]^{\otimes k_e} \otimes \mathrm{Der}_\omega(\hat{L}W)^{\otimes(k_s+k_n)} &\xrightarrow{\text{proj.}} W[-N]^{\otimes k_e} \otimes \bigotimes_{i=1}^{k_s+k_n} \mathrm{Der}^{a_i+2}_\omega(\hat{L}W) \\ &\overset{\Phi}{\simeq} W[-N]^{\otimes k_e} \otimes \bigotimes_{i=1}^{k_s+k_n} W(a_i)[-N] \subset \bigotimes_{i=1}^{k}(W^{\otimes r_i}[-N]) \\ &\xrightarrow{\sigma^{\otimes k}} \bigotimes_{i=1}^{k} W^{\otimes r_i} \xrightarrow{\tau(\hat{o})} (W^{\otimes 2})^{\otimes l} \xrightarrow{\omega_E} \mathbb{R},\end{aligned}$$

where $\Phi := \mathrm{id}^{\otimes k_e}_{W[-N]} \otimes \Phi^{\otimes(k_s+k_n)}_\omega$, $\omega_E := \omega_{e_1} \otimes \cdots \otimes \omega_{e_l}$, and $\omega_{e_j} := \omega_{(|h^{e_j}_1|, |h^{e_j}_2|)}$ if $e_j = \{h^{e_j}_1, h^{e_j}_2\}$. Here we denote by $\omega_{(d_1,d_2)}$ for integers d_1, d_2 the composition of the projection $W \otimes W \to W_{d_1} \otimes W_{d_2}$ and the restriction of ω to $W_{d_1} \otimes W_{d_2}$. The map $\alpha(\Gamma, \hat{o})$ is independent of a choice of linear orders of half-edges representing cyclic orders, and compatible with the commutativity relation.

We define the map $\hat{\psi}(\Gamma, \hat{o}) : \mathfrak{D}^{\otimes k_n} \to \mathbb{R}$ by

$$\hat{\psi}(\Gamma, \hat{o})(D_1, \dots, D_{k_n}) := \alpha(\Gamma, \hat{o})(w_1, \dots, w_{k_e}, \underbrace{\delta, \dots, \delta}_{k_s}, D_1, \dots, D_{k_n})$$



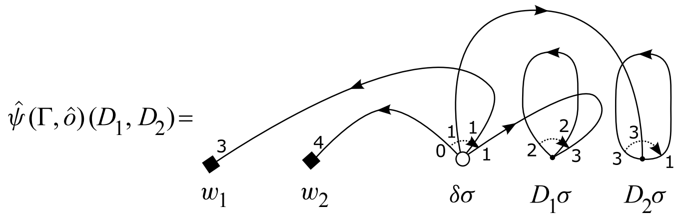


Figure 8. An example of $\hat{\psi}(\Gamma, \hat{o})(D_1, D_2)$. ($\Gamma$ is the decorated graph in Figure 3, right.)

for $D_i \in \mathfrak{D}$. (An example of $\hat{\psi}(\Gamma, \hat{o})$ is described in Figure 8.) Restricting the map[1] on the exterior algebra, we can get the map

$$\psi(\Gamma, o) = \hat{\psi}(\Gamma, \hat{o}) \circ \mathrm{Alt}_{k_n} : \Lambda^{k_n}\mathfrak{D} \to \mathbb{R}.$$

The map is independent of a representation $\hat{o}$ of o by the definition of an orientation. So we obtain the map $\psi : C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{D})$.

Well-definedness of ψ is proved by the correspondence through ψ between relations in the graph complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ corresponding to properties of derivations:

graph complex	derivations
cyclicity	symplectic derivation
commutativity	Lie derivation
A_∞-relation	$\delta^2 = 0$
cut-off	symplectic form

By definition, it is clear except for the A_∞-relation. The correspondence for the A_∞-relation is proved in the end of the proof of the following theorem.

Theorem 3.9. *The map* $\psi : C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{D})$ *is a double chain map.*

Proof. First, we shall show that $d_{\mathrm{CE}}\psi = \psi d$ on $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$. To prove this, we need Lemma 2.3.

[1]For a graded vector space V, the injective map $\mathrm{Alt}_n : \Lambda^n V \to V^{\otimes n}$ is defined by

$$\mathrm{Alt}_n(v_1 \cdots v_n) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \bar{\epsilon}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

for $v_1, \dots, v_n \in V$, where $\bar{\epsilon}(\sigma)$ is the corresponding anti-Koszul sign.

For an oriented graph (Γ, o), we define the two lifts $\hat{o}^1$, $\hat{o}^2$ on $\Gamma^{s,t}_{v_i;h_\nu,h_\mu}$ as

$$\begin{aligned}
\hat{o}^1 &= ((h', h''), \cdot\,; \cdot\,; \cdot\,; v_1, \ldots, v_i', v_i'', \ldots, v_p),\\
\hat{o}^2 &= ((h'', h'), \cdot\,; \cdot\,; \cdot\,; v_1, \ldots, v_i'', v_i', \ldots, v_p),\\
v_i' &= (h^i_{\nu+1}, \ldots, h^i_\mu, h'), \qquad v_i'' = (h^i_1, \ldots, h^i_\nu, h'', h^i_{\mu+1}, \ldots, h^i_{r_i}),
\end{aligned}$$

where $r_i = \# v_i$. The signs ϵ_i are defined by the equations

$$o^1 := \epsilon_1[\hat{o}^1], \qquad o^2 := \epsilon_2[\hat{o}^2], \qquad d^{s,t}_{v_i,h_\nu,h_\mu} o = (-1)^{i-1} o^1 = (-1)^{i-1} o^2.$$

So we obtain

$$\begin{aligned}
d(\Gamma, o) &= \sum_{i=1}^{k} \sum_{\nu<\mu} \sum_{a+b=N} (-1)^{i-1} (\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, o^1)\\
&= \sum_{i=1}^{k} \sum_{\nu<\mu} \sum_{a+b=N} (-1)^{i-1} (\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, o^2).
\end{aligned}$$

Note that

$$d_{\mathrm{CE}}(\chi \circ \mathrm{Alt}_p) = \frac{1}{2} \sum_{s=1}^{p} (-1)^{s-1} \chi \circ (1^{\otimes s-1} \otimes [\,\cdot\,,\cdot\,] \otimes 1^{\otimes p-s}) \circ \mathrm{Alt}_{p+1}$$

for a linear map $\chi : W^{\otimes r_1}[-N] \otimes \cdots \otimes W^{\otimes r_p}[-N] \to \mathbb{R}$ and the antisymmetrization Alt_p for p-components. So we should prove

$$\begin{aligned}
&\hat{\psi}(\Gamma, \hat{o}) \circ (1^{\otimes i-1} \otimes [\,\cdot\,,\cdot\,] \otimes 1^{\otimes p-i-1})\\
&\qquad = \sum_{\nu<\mu} \sum_{a+b=N} (\epsilon_1 \hat{\psi}(\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, \hat{o}^1) + \epsilon_2 \hat{\psi}(\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, \hat{o}^2) \circ \tau),
\end{aligned}$$

where the map τ means the permutation

$$\begin{aligned}
&X_1 \otimes \cdots \otimes (x_{\nu+1} \cdots x_\mu x') \otimes (x_1 \cdots x_\nu x'' x_{\mu+1} \cdots x_{r_i}) \otimes \cdots \otimes X_p\\
&\qquad \mapsto \epsilon \cdot X_1 \otimes \cdots \otimes (x_1 \cdots x_\nu x' x_{\mu+1} \cdots x_{r_i}) \otimes (x_{\nu+1} \cdots x_\mu x'') \otimes \cdots \otimes X_p
\end{aligned}$$

and ϵ is the Koszul sign. It follows from the equations

$$\begin{aligned}
\hat{\psi}(\Gamma, \hat{o}) \circ (1^{\otimes i-1} \otimes \sigma^{-1}(1 \otimes \omega_{(a,b)}) \pi^{r',r''}_{1;t} \sigma^{\otimes 2} \otimes 1^{\otimes p-i-1}) &= \epsilon_1 \hat{\psi}(\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, \hat{o}^1),\\
\hat{\psi}(\Gamma, \hat{o}) \circ (1^{\otimes i-1} \otimes \sigma^{-1}(1 \otimes \omega_{(a,b)}) \pi^{r',r''}_{2;t} \sigma^{\otimes 2} \otimes 1^{\otimes p-i-1}) &= \epsilon_2 \hat{\psi}(\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, \hat{o}^2) \circ \tau,
\end{aligned}$$

for $r' = \mu - \nu + 1$, $r'' = r - \mu + \nu + 1$, and $t = \nu + 1$. The first equation is verified as follows: we have by the definition of $\hat{\psi}$

$$\omega(x', x'') \hat{\psi}(\Gamma, \hat{o})(X_1, \ldots, X_p) = \epsilon_1 \hat{\psi}(\Gamma^{a,b}_{v_i;h_\nu,h_\mu}, \hat{o}^1)(X_1, \ldots, X_i', X_i'', \ldots, X_p)$$

for $X_s \in W^{\otimes r_s}$, $x' \in W_a$, and $x'' \in W_b$. Here we put $X_i' = x_{\nu+1} \cdots x_\mu x' \sigma^{-1}$ and $X_i'' = x_1 \cdots x_\nu x'' x_{\mu+1} \cdots x_r \sigma^{-1}$ for $X_i = x_1 \cdots x_r \sigma^{-1}$. So we obtain the first equation from

$$\epsilon_1 X_i \omega(x', x'') = \sigma^{-1}(1 \otimes \omega)\pi_{1;t}^{r',r''} \otimes \sigma^{\otimes 2}(X_i' \otimes X_i'').$$

The second is also verified in the same way.

Next, we shall prove $i_\delta \psi = \psi i$ on $\widehat{C}^{\bullet;\bullet}_{\mathrm{com}}(W)$. The ordering

$$\hat{o}_i := (\,\cdot\,;\,\cdot\,;\,\cdot\,;\, v_i, v_1, \ldots, \hat{v}_i, \ldots, v_p)$$

is a lift of $\bar{\epsilon}_i \cdot o$, where $\bar{\epsilon}_i$ is the anti-Koszul sign of the permutation

$$(v_1, \ldots, v_p) \mapsto (v_i, v_1, \ldots, \hat{v}_i, \ldots, v_p).$$

So we have

$$\begin{aligned}
&\psi i(\Gamma, o)(X_1, \ldots, X_{p-1})\\
&\quad = \sum_{s=1}^{j} \bar{\epsilon}_i \cdot \alpha(i_{v_i}(\Gamma), \hat{o}_i)(w_1, \ldots, w_{k_e}, \underbrace{\delta, \ldots, \delta}_{k_s+1}, \mathrm{Alt}_{p-1}(X_1, \ldots, X_{p-1}))\\
&\quad = \sum_{s=1}^{j} \sum_{\pi \in \mathfrak{S}_{p-1}} \bar{\epsilon} \cdot \alpha(\Gamma, \hat{o})(w_1, \ldots, w_{k_e}, \underbrace{\delta, \ldots, \delta}_{k_s}, X_{\pi(1)}, \ldots, \delta, \ldots, X_{\pi(p-1)})\\
&\quad = \alpha(\Gamma, \hat{o})(w_1, \ldots, w_{k_e}, \underbrace{\delta, \ldots, \delta}_{k_s}, \mathrm{Alt}_p(\delta, X_1, \ldots, X_{p-1}))\\
&\quad = i_\delta \psi(\Gamma, o)(X_1, \ldots, X_{p-1})
\end{aligned}$$

where $\bar{\epsilon}$ is the anti-Koszul sign of

$$(\delta, X_1, \ldots, X_{p-1}) \mapsto (X_{\pi(1)}, \ldots, \delta, \ldots, X_{\pi(p-1)}).$$

From the discussion above, the relation $\psi(R_v(\Gamma, o)) = 0$ follows from

$$\psi(R_v(\Gamma, o)) = \psi(i_{v'} i_{v''} d_v(\Gamma, o)) = \psi(\Gamma, o)([\delta, \delta], \,\cdot\,) = 0.$$

Thus, ψ induces the map $\psi : C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{D})$. Furthermore, since ψ is commutative with d and i, so is L. So we complete the proof. □

The group $\mathrm{Sp}(W, \delta)$ acts on $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ by the action on their labels. Then, the chain map $\psi : C^{\bullet;\bullet}_{\mathrm{com}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{D})$ is $\mathrm{Sp}(W, \delta)$-equivariant clearly. In particular, we can consider the $\mathrm{Sp}(W, \delta)$-invariant part $C^{\bullet;\bullet}_{\mathrm{com}}(W)^{\mathrm{Sp}(W,\delta)}$ of the complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)$. It has the double subcomplex $C^{\bullet;\bullet}_{\mathrm{com}}(N, Z)$ consisting of N-graded graphs which have no external vertex. This complex $C^{\bullet;\bullet}_{\mathrm{com}}(N, Z)$ does not depend on the symplectic vector space W. It depends only a range Z of degrees and a degree N of a symplectic inner product.

Remark 3.10. We can define the associative version of $C^{\bullet;\bullet}_{\mathrm{com}}(W)$ as follows. Set

$$\widetilde{O}_{\mathrm{ass}}(W,\Gamma) := \bigodot_{e\in E(\Gamma)} O(e) \otimes \bigodot_{u\in V_e(\Gamma)} W[-N]_{|u|} \otimes \bigwedge_{v^s\in V_s(\Gamma)} \mathrm{Cyc}(v^s)[N] \otimes \bigwedge_{v\in V_n(\Gamma)} \mathrm{Cyc}(v)[N],$$

$$C^{\bullet;\bullet}_{\mathrm{ass}}(W) := \bigoplus_{\Gamma\in\mathcal{G}(N)} O_{\mathrm{ass}}(W,\Gamma), \qquad O_{\mathrm{ass}}(W,\Gamma) := \widetilde{O}_{\mathrm{ass}}(W,\Gamma)_{\mathrm{Aut}(\Gamma)}.$$

Then $(C^{\bullet;\bullet}_{\mathrm{ass}}(W), d, L)$ is also a double $\mathrm{Sp}(W,\delta)$-chain complex, and the chain map

$$C^{\bullet;\bullet}_{\mathrm{ass}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\widehat{T}W))$$

can be defined in the same way. In this case, we can also consider the double subcomplex $C^{\bullet;\bullet}_{\mathrm{ass}}(N,Z)$ which consists of N-graded graphs without external vertices.

4. Applications and examples

Examples of relations between our chain map and a known notion are given in the following two examples.

Example 4.1. For a cyclic minimal A_∞-algebra (H, I, m) with even degree, putting $W := H^*[-1]$, we have the map $C^{\bullet;\bullet}_{\mathrm{ass}}(W) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\widehat{T}W))$. Here $\widehat{T}W$ is the dual of the bar construction of (H, I, m). The map induced by the chain map

$$C^{0,\bullet}_{\mathrm{ass}}(N,Z) \to C^{0,\bullet}_{\mathrm{CE}}(\mathrm{Der}_\omega(\widehat{T}W)) = \mathbb{R}$$

is known as the Kontsevich cocycle [Kontsevich 1994; Penkava and Schwarz 1995; Hamilton and Lazarev 2008] of a cyclic A_∞-algebra (H, I, m).

Example 4.2. In the case of $Z = \{0\}$ and $\delta = 0$, the chain map

$$C^{\bullet,0}_{\mathrm{ass}}(0,\{0\}) \to C^{\bullet,0}_{\mathrm{CE}}(\mathrm{Der}_\omega(\widehat{T}W))^{\mathrm{Sp}(W)}$$

is equal to Kontsevich's chain map [1994; 1993].

In the case that W is positively graded, we define a chain complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ by

$$C^{\bullet;\bullet}_{\mathrm{com}}(W)_+ = C^{\bullet;\bullet}_{\mathrm{com}}(W)/(\mathrm{positivity}),$$

where the positivity relation is as follows:

- (*Positivity*) (i) a graph which has a normal vertex v satisfying $|v| \le 0$ is zero, and (ii) $(i_{v'} + i_{v''})d_v(\Gamma, o) = 0$ for an oriented graph (Γ, o) and a normal vertex v of degree 0.

The differentials d, L are also defined on $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$, while i is not.

Proposition 4.3. *The operators d, L induce the differentials on $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$.*

Proof. It is clear that these operators are compatible with the former condition (i) of the positivity relation. Note that, to prove compatibility with L for a graph including a vertex with degree 0, we need to use (ii).

We shall prove they are compatible with (ii). First, we shall calculate the image of (ii) by the operator d. For $\Gamma \in \mathcal{G}(N)$ and a normal vertex v of degree 0, we have

$$\begin{aligned} d(i_{v'}+i_{v''})d_v &= d_{v''}i_{v'}d_v + d_{v'}i_{v''}d_v + \sum_{u\neq v',v''} d_u(i_{v'}+i_{v''})d_v \\ &= d_{v''}i_{v'}d_v + d_{v'}i_{v''}d_v - \sum_{u\neq v}(i_{v'}+i_{v''})d_v d_u. \end{aligned}$$

Here we used the equations in the proof of Proposition 3.8. For a splitting of v such that $|v'|=-1$, $d_{v''}i_{v'}d_v$ must have a nonpositive vertex since $|v''|=1$. In the same way, $d_{v'}i_{v''}d_v$ must also have a nonpositive vertex. So $d(i_{v'}+i_{v''})d_v$ is equal to zero under the positivity relation.

Next, we shall calculate the image of (ii) by the operator L:

$$\begin{aligned} L(i_{v'}+i_{v''})d_v &= \sum_u (i_{u'}+i_{u''})d_u(i_{v'}+i_{v''})d_v \\ &= (i_{(v'')'}+i_{(v'')''})d_{v''}i_{v'}d_v + (i_{(v')'}+i_{(v')''})d_{v'}i_{v''}d_v \\ &\qquad - \sum_{u\neq v}(i_{v'}+i_{v''})d_v(i_{u'}+i_{u''})d_u \\ &= (i_{(v'')'}+i_{(v'')''})i_{v'}d_{v''}d_v + (i_{(v')'}+i_{(v')''})i_{v''}d_{v'}d_v \\ &\qquad - \sum_{u\neq v}(i_{v'}+i_{v''})d_v(i_{u'}+i_{u''})d_u. \end{aligned}$$

By changing names of vertices like in the proof of Proposition 3.8, we get

$$(i_{(v'')'}+i_{(v'')''})i_{v'}d_{v''}d_v = -(i_{(v')''}+i_{v''})i_{(v')'}d_{v'}d_v = -R_{v'}d_{v'}d_v - i_{(v')'}i_{v''}d_{v'}d_v,$$

and

$$\begin{aligned} (i_{(v'')'}+i_{(v'')''})i_{v'}d_{v''}d_v + (i_{(v')'}+i_{(v')''})i_{v''}d_{v'}d_v &= -R_{v'}d_{v'}d_v + i_{(v')''}i_{v''}d_{v'}d_v \\ &= -R_{v'}d_{v'}d_v - i_{(v'')'}i_{(v'')''}d_{v''}d_v \\ &= -R_{v'}d_{v'}d_v - R_{v''}d_{v''}d_v. \end{aligned}$$

Using the A_∞-relation, $L(i_{v'}+i_{v''})d_v$ is equal to zero under the positivity relation. □

Then we can also get the chain map

$$\psi_+ : C^{\bullet;\bullet}_{\mathrm{com}}(W)_+ \to C^\bullet(\mathrm{Der}^+_\omega(LW))$$

induced by ψ.

Example 4.4. Suppose $X = \#_g(S^n \times S^n) \setminus \operatorname{Int} D^{2n}$. Its Quillen model is described by

$$L_X = L(u_1, v_1, \dots, u_g, v_g) \quad (\deg u_i = \deg v_i = n-1), \quad \delta = 0,$$
$$\omega(u_i, v_j) = \delta_{ij}, \qquad \omega(u_i, u_j) = \omega(v_i, v_j) = 0.$$

It means $N = 2n-2$, $W = \langle u_1, v_1, \dots, u_g, v_g \rangle$, and $Z = \{n-1\}$. Then the dgl $(\mathrm{Der}^+_\omega(L_X), 0)$ is a Quillen model of $B\operatorname{aut}_{\partial,0}(X)$ (which is proved in [Berglund and Madsen 2014]). In this case, we can forget all special vertices in the graph complex since $\delta = 0$. So we have the chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(2n-2, \{n-1\})_+/(\text{special vertices}) \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))^{\mathrm{Sp}(W)}.$$

This map is constructed by [Berglund and Madsen 2014], and it is proved that the map is an isomorphism under the limit $g \to \infty$.

Example 4.5. Suppose $X = \mathbb{C}P^3 \setminus \operatorname{Int} D^6$. Its Quillen model is described by

$$L_X = L(u_1, u_2) \quad (\deg u_i = 2i-1), \qquad \delta = \tfrac{1}{2}[u_1, u_1]\frac{\partial}{\partial u_2},$$
$$\omega(u_1, u_2) = \omega(u_2, u_1) = 1.$$

It means $N = 4$, $W = \langle u_1, u_2 \rangle$, and $Z = \{1, 3\}$. Then the dgl $(\mathrm{Der}^+_\omega(L_X), \delta)$ is a Quillen model of $B\operatorname{aut}_{\partial,0}(X)$. Since $\mathrm{Sp}(W, \delta) = 1$, we have the chain map

$$C^{\bullet;\bullet}_{\mathrm{com}}(W)_+ \to C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X)) = C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))^{\mathrm{Sp}(W,\delta)}.$$

We shall define a certain sub dgl $\mathfrak{d}$ of $\mathrm{Der}_\omega(L_X)$. Put

$$A_1 = \tfrac{1}{2}[u_1, u_1]\frac{\partial}{\partial u_2}, \qquad A_2 = \tfrac{1}{2}[u_2, u_2]\frac{\partial}{\partial u_1}$$
$$B_1 = \tfrac{1}{2}[u_1, u_1]\frac{\partial}{\partial u_1} + [u_1, u_2]\frac{\partial}{\partial u_2}, \qquad B_2 = [u_1, u_2]\frac{\partial}{\partial u_1} + \tfrac{1}{2}[u_2, u_2]\frac{\partial}{\partial u_2}.$$

Then we have

$$\delta(A_1) = \delta(B_1) = \delta(B_2) = 0,$$
$$\delta(A_2) = \tfrac{1}{2}[[u_1, u_1], u_2]\frac{\partial}{\partial u_1} + \tfrac{1}{2}[[u_2, u_2], u_1]\frac{\partial}{\partial u_2} = [A_1, A_2] = -[B_1, B_2] =: C,$$
$$[A_i, B_j] = [A_i, A_i] = [B_j, B_j] = 0 \quad (i, j = 1, 2),$$
$$\deg A_1 = -1, \qquad \deg A_2 = 5, \qquad \deg B_1 = 1, \qquad \deg B_2 = 3, \qquad \deg C = 4.$$

Here we put $\delta(Z) := [\delta, Z]$ for simplicity. By the relation above,

$$\mathfrak{d} := \langle A_1, A_2, B_1, B_2, C \rangle = \mathrm{Der}^1_\omega(L_X) \oplus \mathrm{Der}^2_\omega(L_X)$$

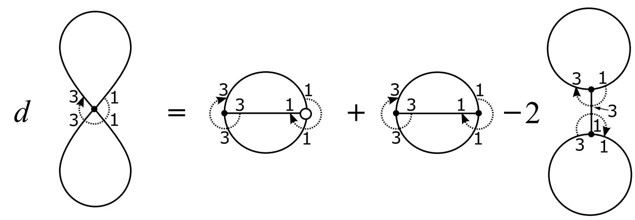


Figure 9. The relation of graphs (the orientations are omitted).

is a sub dgl of $\mathrm{Der}_\omega(L_X)$. Its positive truncation $\mathfrak{d}^+$ is described by

$$\mathfrak{d}^+ = \langle A_2, B_1, B_2, C\rangle,$$
$$\delta(A_2) = -[B_1, B_2] = C, \qquad \delta(B_1) = \delta(B_2) = \delta(C) = 0,$$
$$[A_2, B_i] = [A_2, A_2] = [B_i, B_i] = [A_2, C] = [B_i, C] = 0 \quad (i = 1, 2).$$

Let x, y_1, y_2, z be the suspension of the dual basis of A_2, B_1, B_2, C. Then the Chevalley–Eilenberg complex of the dgl $\mathfrak{d}^+$ is written by

$$C^{\bullet;\bullet}_{\mathrm{CE}}(\mathfrak{d}^+) = \Lambda(x, y_1, y_2, z) \quad (\deg x = 6,\ \deg y_1 = 2,\ \deg y_2 = 4,\ \deg z = 5),$$
$$dx = dy_1 = dy_2 = 0, \qquad dz = x - y_1 y_2$$

and its total cohomology

$$H^\bullet_{\mathrm{CE}}(\mathfrak{d}^+) = \Lambda(x, y_1, y_2)/(x - y_1 y_2).$$

Since $\mathfrak{d}^+$ is the rank ≤ 2 part of $\mathrm{Der}^+_\omega(L_X)$, the map $H^\bullet_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X)) \to H^\bullet_{\mathrm{CE}}(\mathfrak{d}^+)$ induced by the inclusion has a section. So nontrivial classes in $H^\bullet_{\mathrm{CE}}(\mathfrak{d}^+)$ give nontrivial classes in $H^\bullet_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))$.

The relation $dz = x - y_1 y_2$ in the Chevalley–Eilenberg complex is corresponding to the relation in the graph complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ described in Figure 9. Here the classes x and $y_1 y_2$ correspond to the first term and the sum of the second and third terms in the figure. Note that y_1 and y_2 do not correspond to graphs without external vertices. According to the positivity relation, all the trivalent graphs appearing in the right-hand side are cycles since the degrees of two half-edges incident to a permitted bivalent vertex in $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ must be 3.

Example 4.6. Suppose $X = \mathbb{C}P^4 \setminus \mathrm{Int}\, D^8$. Its Quillen model is described by

$$L_X = L(u_1, u_2, u_3) \quad (\deg u_i = 2i - 1), \qquad \delta = \tfrac{1}{2}[u_1, u_1]\frac{\partial}{\partial u_2} + [u_1, u_2]\frac{\partial}{\partial u_3},$$
$$\omega(u_2, u_2) = \omega(u_1, u_3) = 1.$$

It means $N=6$, $W=\langle u_1, u_2, u_3\rangle$, and $Z=\{1,3,5\}$. Then the dgl $(\mathrm{Der}^+_\omega(L_X), \delta)$ is a Quillen model of $B\,\mathrm{aut}_{\partial,0}(X)$. Defining the linear transformation τ by $\tau(u_1)=-u_1$, $\tau(u_2)=u_2$, and $\tau(u_3)=-u_3$, we have $\mathrm{Sp}(W,\delta)=\{1,\tau\}$. So $C^{\bullet;\bullet}_{\mathrm{com}}(W)^{\mathrm{Sp}(W,\delta)}$ is generated by graphs labeled by u_1, u_2, u_3 satisfying $\#\{u_1, u_3\text{-labeled vertex}\}$ even. For simplicity, we put

$$[u_{i_1}\cdots u_{i_k}] := [u_{i_1},\ldots,u_{i_k}]_{\mathrm{cyc}} = \sum_{s=1}^{k}(-1)^{s(k-s)} u_{i_{s+1}}\cdots u_{i_k} u_{i_1}\cdots u_{i_s} \in W^k_{\mathrm{cyc}}.$$

Using notations in Section 3C, we can take a basis of $W(2)$

$$[u_i u_j] \quad (\{i<j\}\subset\{1,2,3\}),$$

a basis of $W(3)$

$$\tfrac{1}{3}[u_iu_iu_i],\ [u_iu_ju_j],\ [u_iu_iu_j] \quad (\{i<j\}\subset\{1,2,3\}), \qquad [u_1u_2u_3]+[u_1u_3u_2],$$

and a basis of $W(4)$

$$[u_iu_iu_ju_j] \quad (i<j),$$
$$[u_1u_1u_2u_3]+[u_1u_1u_3u_2],\ [u_1u_2u_2u_3]-[u_1u_3u_2u_2],\ [u_1u_2u_3u_3]-[u_1u_3u_3u_2].$$

We put the corresponding rank-0, rank-1, and rank-2 bases of $\mathrm{Der}_\omega(L_X)$

$$P_{ij},\ A_{iii},\ A_{ijj},\ A_{iij},\ A_{123},\ B_{iijj},\ B_{1123},\ B_{1223},\ B_{1233},$$

and these dual bases p_{ij}, x_{ijk}, and y_{ijkl} of P_{ij}, A_{ijk}, and B_{ijkl}. Then by direct calculation we have the equations in $C^{\bullet;\bullet}_{\mathrm{CE}}(\mathrm{Der}^+_\omega(L_X))$

$$\begin{aligned}
dy_{1122} &= x_{222} - 2x_{123} + x_{122}x_{113} - x_{122}x_{122},\\
dy_{2233} &= x_{333}x_{122} + x_{233}x_{222} - x_{223}x_{223} - 2x_{123}x_{233} + x_{133}x_{223} + 2p_{23}y_{1233},\\
dy_{1133} &= x_{233} - x_{133}x_{113} - x_{123}x_{123} - 2p_{23}y_{1123},\\
dy_{1123} &= x_{223} - x_{133} - p_{23}y_{1122},\\
dy_{1223} &= x_{233} + x_{223}x_{122} + x_{123}x_{123} - x_{223}x_{113} - x_{123}x_{222} - x_{133}x_{122} + p_{23}y_{1123},\\
dy_{1233} &= x_{333} + x_{233}x_{122} - x_{123}x_{223} - x_{233}x_{113} - x_{123}x_{122} + p_{23}y_{1133}.
\end{aligned}$$

Here all terms appearing in the right-hand side of the equations are cocycles. For example, the fifth relation is corresponding to the relation in the graph complex $C^{\bullet;\bullet}_{\mathrm{com}}(W)_+$ described in Figure 10. In Figure 10, the image by ψ_+ of each graph appearing the last term of the right-hand side is zero.

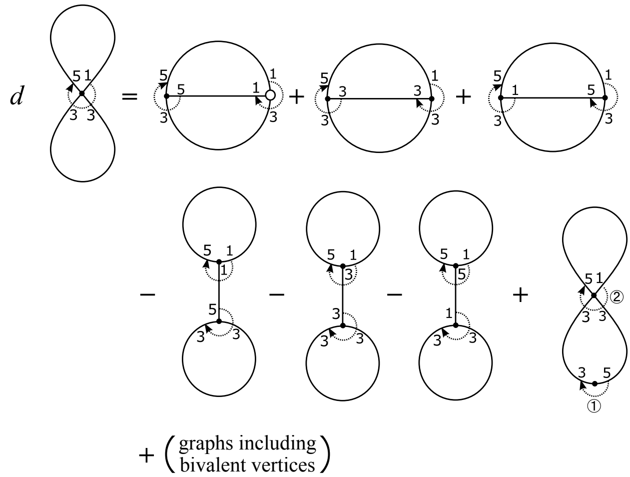


Figure 10. The relation of graphs (① and ② mean the orientation of vertices and the other orientations are omitted).

References

[Berglund and Madsen 2014] A. Berglund and I. Madsen, "Rational homotopy theory of automorphisms of manifolds", preprint, 2014. arXiv

[Chen 1977] K. T. Chen, "Iterated path integrals", *Bull. Amer. Math. Soc.* **83**:5 (1977), 831–879. MR Zbl

[Conant and Vogtmann 2003] J. Conant and K. Vogtmann, "On a theorem of Kontsevich", *Algebr. Geom. Topol.* **3** (2003), 1167–1224. MR Zbl

[Hamilton 2006] A. Hamilton, "A super-analogue of Kontsevich's theorem on graph homology", *Lett. Math. Phys.* **76**:1 (2006), 37–55. MR Zbl

[Hamilton and Lazarev 2004] A. Hamilton and A. Lazarev, "Homotopy algebras and noncommutative geometry", preprint, 2004. arXiv

[Hamilton and Lazarev 2008] A. Hamilton and A. Lazarev, "Characteristic classes of A_∞-algebras", *J. Homotopy Relat. Struct.* **3**:1 (2008), 65–111. MR Zbl

[Kadeishvili 2009] T. Kadeishvili, "Cohomology C_∞-algebra and rational homotopy type", pp. 225–240 in *Algebraic topology: old and new* (Będlewo, Poland, 2007), edited by M. Golasiński et al., Banach Center Publ. **85**, Polish Acad. Sci. Inst. Math., Warsaw, 2009. MR Zbl

[Kajiura et al. 2016] H. Kajiura, T. Matsuyuki, and Y. Terashima, "Homotopy theory of A_∞-algebras and characteristic classes of fiber bundles", preprint, 2016. arXiv

[Kontsevich 1993] M. Kontsevich, "Formal (non)commutative symplectic geometry", pp. 173–187 in *The Gelfand Mathematical Seminars, 1990–1992*, edited by L. Corwin et al., Birkhäuser, Boston, 1993. MR Zbl

[Kontsevich 1994] M. Kontsevich, "Feynman diagrams and low-dimensional topology", pp. 97–121 in *First European Congress of Mathematics, II* (Paris, 1992), Progr. Math. **120**, Birkhäuser, Basel, 1994. MR Zbl

[Kontsevich and Soibelman 2001] M. Kontsevich and Y. Soibelman, "Homological mirror symmetry and torus fibrations", pp. 203–263 in *Symplectic geometry and mirror symmetry* (Seoul, 2000), edited by K. Fukaya et al., World Sci., River Edge, NJ, 2001. MR Zbl

[Markl 1999] M. Markl, "Cyclic operads and homology of graph complexes", *Rend. Circ. Mat. Palermo* (2) *Suppl.* **59** (1999), 161–170. MR Zbl

[Markl et al. 2002] M. Markl, S. Shnider, and J. Stasheff, *Operads in algebra, topology and physics*, Math. Surv. Monographs **96**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Matsuyuki and Terashima 2016] T. Matsuyuki and Y. Terashima, "Characteristic classes of fiber bundles", *Algebr. Geom. Topol.* **16**:5 (2016), 3029–3050. MR Zbl

[Merkulov 1999] S. A. Merkulov, "Strong homotopy algebras of a Kähler manifold", *Int. Math. Res. Not.* **1999**:3 (1999), 153–164. MR Zbl

[Penkava and Schwarz 1995] M. Penkava and A. Schwarz, "A_∞ algebras and the cohomology of moduli spaces", pp. 91–107 in *Lie groups and Lie algebras: E. B. Dynkin's Seminar*, edited by S. G. Gindikin and E. B. Vinberg, Amer. Math. Soc. Transl. (2) **169**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Schlessinger and Stasheff 2012] M. Schlessinger and J. Stasheff, "Deformation theory and rational homotopy type", preprint, 2012. arXiv

[Tanré 1983] D. Tanré, *Homotopie rationnelle: modèles de Chen, Quillen, Sullivan*, Lecture Notes in Math. **1025**, Springer, 1983. MR Zbl

Received June 17, 2018. Revised January 1, 2019.

TAKAHIRO MATSUYUKI
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF TOKYO
MEGURO
TOKYO
JAPAN

matsuyuk@ms.u-tokyo.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.575

INTEGRATION OF MODULES I: STABILITY

DMITRIY RUMYNIN AND MATTHEW WESTAWAY

We explore the integration of representations from a Lie algebra to its algebraic group in positive characteristic. An integrable module is stable under the twists by group elements. Our aim is to investigate cohomological obstructions for passing from stability to an algebraic group action. As an application, we prove integrability of bricks for a semisimple algebraic group.

Over a field of positive characteristic, an algebraic group G acts on its Lie algebra $\mathfrak{g}$ and the restricted enveloping algebra $U_1(\mathfrak{g})$ by automorphisms. This yields twists: an element $x \in G$ twists a $\mathfrak{g}$-module (V, θ) into $(V, \theta)^x := (V, \theta \circ \mathrm{Ad}(x))$. A $\mathfrak{g}$-module is G*-stable* if it is isomorphic to all its twists. A $\mathfrak{g}$-module coming from a G-module is necessarily G-stable but the converse is not true. An important question in the modular representation theory of Lie algebras and algebraic groups is to determine for which modules the converse is true. We investigate this question in this paper.

Our method is subtly different from the known approach. Not only Cline [1972] and Donkin [1982] but also Parshall and Scott in their modern exposition [2013] pursue a certain unipotent extension G^* of the group G that acts on a G-stable $\mathfrak{g}$-module (V, θ). We, instead, contemplate projective actions of G on (V, θ). In particular, we completely avoid the theory of Schreier systems.

Our approach instead has similarities to the work of Dade [1981] and Thévenaz [1983] on a related question for abstract groups. They study whether a G-stable representation (V, θ) of a normal subgroup L of an abstract group G can be extended to a representation of the entire group G. They show that when the automorphism group $\mathrm{Aut}_L(V)$ of V is abelian the extension is controlled by $H^2(G/L, \mathrm{Aut}_L(V))$. Furthermore, the uniqueness of such extensions is controlled by $H^1(G/L, \mathrm{Aut}_L(V))$.

By introducing the terminology of (L, H)-morphs — a type of function which is partway to being a homomorphism — we are able to reinterpret the results of Dade and Thévenaz in a more general context (Theorem 5). When we apply these results

The research was partially supported by the Russian Academic Excellence Project "5–100" and by the Max Planck Society.
MSC2010: primary 20G05; secondary 17B45.
Keywords: Frobenius kernel, representation, cohomology, obstructions.

to the question of module extensions we repeat Corollary 1.8 and Proposition 2.1 in [Thévenaz 1983], however our formalism allows us to generalise this to the case where $\mathrm{Aut}_L(V)$ is only soluble rather than abelian via an inductive process (Theorem 7).

The other key difference between our results and those of Dade and Thévenaz is that we work with a slightly different relative cochain complex, denoted $(C^\bullet(G, L; A), d)$, while they work with the more standard complex $(C^\bullet(G/L; A), d)$. Whilst the cohomology of these complexes differs from the second cohomology group on, Theorem 5 in fact works in either case. However, in order to apply similar methods to the case of algebraic groups, the study of this new cochain complex becomes necessary. These considerations are explained in more detail in Sections 1.5 and 2.5.

Now we reveal the detailed content of the present paper, emphasising the main results. In Section 1, we devise all the machinery to discuss G-stable modules in the setting of abstract groups: a group G, its normal subgroup L and a G-stable L-module (V, θ). We introduce weak (L, H)-morphs and the relative cochain complex $C^\bullet(G, L; A)$ in Section 1.2, where A is an abelian group with a G-action. They feature in a key exact sequence (see Theorem 5) that controls both uniqueness and existence of G-actions for a large class of G-stable L-modules.

The main result of this chapter is Theorem 7, a somewhat algorithmic result pinpointing completely uniqueness and existence of a G-module structure on a G-stable L-module. Notice that it has been established by Xanthopoulos [1992] that $H^1(G/L; A)$ controls uniqueness. Since $H^1(G/L; A) = H^1(G, L; A)$, our results about uniqueness are known. However, $H^2(G/L; A) \neq H^2(G, L; A)$ (and the latter controls existence), hence our results on existence are new, even in the setting of abstract groups. Our approach is useful because it fuses uniqueness and existence into a single process controlled by the relative cohomology.

In Section 2 we extend our Section 1 results from abstract groups to algebraic groups. We face some technical challenges. An important case for applications is when L is a Frobenius kernel of G. Hence, we must assume that L is a closed subgroup scheme, not just a closed algebraic subgroup. The second challenge is poles: we need to distinguish rational and algebraic cohomology, since we encounter rational cocycles $\mu : G^n \to A$ that are not necessarily algebraic. We deal with technicalities in Sections 2.1 and 2.2.

We exhibit a key exact sequence for rational cohomology (Theorem 27 — an analogue of Theorem 5) in Section 2.3. Again, this sequence controls both uniqueness and existence of G-actions. Immediately we put it to good use: a G-stable $\mathfrak{g}$-brick (a module with trivial endomorphisms) is a G-module (Theorem 28).

A greater generality than $\mathfrak{g}$-bricks is $\mathfrak{g}$-modules with a soluble group of automorphisms. These are our assumptions in Section 2.4. Our main result in this section is Theorem 29, an analogue of Theorem 7 for algebraic groups. Again, this theorem pinpoints completely uniqueness and existence of a G-module structure on a G-stable $\mathfrak{g}$-module.

It is interesting to see whether our results could be applied to two old conjectures in the area: Humphreys–Verma Conjecture [Donkin 1982; Parshall and Scott 2013; Jantzen 1987, Chapter 11] and Verma Conjecture [Donkin 1980; Xanthopoulos 1992]

1. G-stable modules for abstract groups

In this chapter we study $\mathbb{A}G$-modules, where G is a group, $\mathbb{A}$ is an associative ring.

1.1. ***Automorphisms of indecomposable modules.*** Let $\mathbb{B}$ be a finite-dimensional algebra over a field $\mathbb{K}$ (of any characteristic), M a finite-dimensional $\mathbb{B}$-module, $\mathbb{E} = \mathrm{End}(M)$ its endomorphism ring, $J = J(\mathbb{E})$ its Jacobson radical, and $H = \mathrm{Aut}(M)$ its automorphism group. We start with the following useful observation:

Proposition 1. (1) *The quotient algebra $\mathbb{E}/J$ is a division algebra if and only if M is indecomposable.*

(2) *If M is indecomposable and $\mathbb{E}/J$ is separable*, *then $H \cong \mathrm{GL}_1(\mathbb{D}) \ltimes U$*, *where $\mathbb{D} = \mathbb{E}/J$ is a division algebra and $U = 1 + J$ is a connected unipotent group.*

(3) *Further to the conditions of* (2), *if $\mathbb{D} = \mathbb{K}$, then $H = \mathrm{GL}_1(\mathbb{K}) \times U$.*

Proof. (1) It is a standard fact that a finite length module is indecomposable if and only if its endomorphism ring is local. Since $\mathbb{E}$ is finite-dimensional, this is equivalent to $\mathbb{E}/J$ being a division ring.

(2) By (1), $\mathbb{D} = \mathbb{E}/J$. Since $\mathbb{D}$ is separable, we can use the Malcev–Wedderburn theorem to split off the radical, i.e., to realize $\mathbb{D}$ as a subalgebra of $\mathbb{E}$ such that $\mathbb{E} = \mathbb{D} \oplus J$.

Clearly, $H = \mathrm{GL}_1(\mathbb{E})$. Consider an element $x = d + j$, $d \in \mathbb{D}$, $j \in J$. Since $x^n = d^n + j'$ for some $j' \in J$, x is nilpotent if and only if $d = 0$. By the fitting lemma, $x \in H$ if and only if $d \neq 0$. The key isomorphism is given by the multiplication map:

$$\mathrm{GL}_1(\mathbb{D}) \ltimes U \xrightarrow{\cong} H = \mathrm{GL}_1(\mathbb{E}), \quad (d, 1 + j) \mapsto d + dj,$$
$$H = \mathrm{GL}_1(\mathbb{E}) \xrightarrow{\cong} \mathrm{GL}_1(\mathbb{D}) \ltimes U, \quad d + j \mapsto (d, 1 + d^{-1} j).$$

It remains to observe that $U = 1 + J$ is a connected unipotent algebraic group. It is connected because it is isomorphic to J as a variety. It is unipotent because each of its elements is unipotent in $\mathrm{GL}(M)$.

(3) The Malcev–Wedderburn decomposition turns J into a $\mathbb{D}$-$\mathbb{D}$-bimodule. Our condition forces $\mathbb{D} \otimes_{\mathbb{K}} \mathbb{D}^{op} = \mathbb{K} \otimes_{\mathbb{K}} \mathbb{K}^{op} = \mathbb{K}$ so that the bimodule structure is just the $\mathbb{K}$-vector space structure. Hence, $\mathrm{GL}_1(\mathbb{D}) = \mathrm{GL}_1(\mathbb{K})$ and U commute. $\square$

1.2. (L, H)-morphs. Let $G \geq L$, $K \geq H$ be two group-subgroup pairs. Let $N = N_K(H)$ and $C_K(H)$ be the normaliser and the centraliser of H in K. By *an (L, H)-morph from G to K* we understand a function $f : G \to K$ satisfying the following four conditions:

(1) $f\,|_L$ is a group homomorphism.

(2) $f(G) \subset N_K(H)$.

(3) $f(x)f(y) \in f(xy)H$ for all $x, y \in G$.

(4) $f(L) \subset C_K(H)$.

By a *weak (L, H)-morph from G to K* we understand a function $f : G \to K$ satisfying only the first three conditions.

One can observe that a weak (L, H)-morph is just a homomorphism $G \to N/H$ with a choice of lifting to N satisfying an additional condition. For instance, weak $(G, 1)$-morphs are the same as homomorphisms $G \to K$ and weak $(1, K)$-morphs are just functions $G \to K$ which preserve the identity. Furthermore, the same statements also hold if we replace weak morphs with morphs in the previous sentence.

Commonly (L, H)-morphs originate from K-G-sets $X = {}_K X_G$, i.e., G acts on the right, K on the left and the actions commute. Let $\theta \in X$ such that its G-orbit is inside its K-orbit. Let H be the stabiliser of θ in K. Choose a section $K/H \to K$ which sends the coset H to 1_K. The composition of the section with the G-orbit map of θ is a function

$$f : G \to K \ \text{ characterised by } \ {}^{f(x)}\theta = \theta^x \text{ for all } x \in G.$$

Lemma 2. *The map f defined above is a $(1, H)$-morph.*

Proof. By definition, ${}^{f(xy)}\theta = \theta^{xy}$. On the other hand, $\theta^{xy} = (\theta^x)^y = ({}^{f(x)}\theta)^y = {}^{f(x)f(y)}\theta$. Hence, $\theta = {}^{f(xy)^{-1}f(xy)}\theta = {}^{f(xy)^{-1}f(x)f(y)}\theta$ and $f(xy)^{-1}f(x)f(y) \in H$.

Now pick $h \in H$. Then ${}^{f(x)^{-1}hf(x)}\theta = {}^{f(x)^{-1}h}\theta^x = {}^{f(x)^{-1}}\theta^x = {}^{f(x)^{-1}f(x)}\theta = \theta$ so that $f(x)^{-1}hf(x) \in H$. □

We would like to identify weak (L, H)-morphs that define the same homomorphisms $G \to N/H$. More precisely, we say that two weak (L, H)-morphs f and f' are equivalent if $f'(x) \in f(x)H$ for all $x \in G$. We denote the set of equivalence classes of weak (L, H)-morphs by $[LH]\mathrm{mo}(G, K)$. Furthermore, given a fixed homomorphism $\theta : L \to K$ we denote by $[LH]^\theta\mathrm{mo}(G, K)$ the set of equivalence classes of those weak (L, H)-morphs that restrict to θ on L.

Let A be an additive abelian group with a G-action (a $\mathbb{Z}G$-module). We consider a subcomplex $(\widetilde{C}^\bullet(G, L; A), d)$ of the standard complex $(C^\bullet(G; A), d)$ that consists of such cochains μ_n that are trivial on L^n, i.e., $\mu_n\,|_{L\times\ldots\times L} \equiv 0_A$.

We see that this cochain complex fits into an exact sequence of cochain complexes

$$0 \to \widetilde{C}^\bullet(G, L; A) \to C^\bullet(G; A) \to C^\bullet(L; A) \to 0.$$

This then allows us to form a long exact sequence of cohomology

$$\dots \to H^{n-1}(G; A) \to H^{n-1}(L; A) \to \widetilde{H}^n(G, L; A) \to H^n(G; A) \to H^n(L; A) \to \dots$$

For our purposes, we have to modify this subcomplex slightly. We consider a subcomplex $(C^\bullet(G, L; A), d)$ of the standard complex $(C^\bullet(G; A), d)$ which is obtained from $(\widetilde{C}^\bullet(G, L; A), d)$ in the following way: for $n > 0$, $C^n(G, L; A) = \widetilde{C}^n(G, L; A)$, whilst $C^0(G, L; A) = A^L$. We can furthermore replace the complex $C^\bullet(L; A)$ with the complex $\widetilde{C}^\bullet(L; A)$, which is defined by $\widetilde{C}^n(L; A) = \mathrm{Coker}(C^n(G, L; A) \to C^n(G; A))$ for all $n \geq 0$. In particular, we observe that $\widetilde{C}^n(L; A) = C^n(L; A)$ for all $n \geq 1$. This then recovers an exact sequence of cochain complexes:

$$0 \to C^\bullet(G, L; A) \to C^\bullet(G; A) \to \widetilde{C}^\bullet(L; A) \to 0.$$

In particular, noting that for the cochain complex $\widetilde{C}^\bullet(L; A)$ we have $\widetilde{H}^0(L; A) = 0$ and $\widetilde{H}^n(L; A) = H^n(L; A)$ for $n \geq 1$, we can form the long exact sequence of cohomology

$$\begin{aligned} 0 \to H^1(G, L; A) \to \dots \to H^{n-1}(L; A) \\ \to H^n(G, L; A) \to H^n(G; A) \to H^n(L; A) \to \dots \end{aligned}$$

What can we say about the natural map $f_n : H^n(G, L; A) \to H^n(G; A)$? From this long exact sequence, the following proposition is clear.

Proposition 3. (1) *For $n > 0$, $H^n(L; A) = 0$ if and only if f_n is surjective and f_{n+1} is injective.*

(2) *For $n > 1$, f_n is injective if and only if the restriction map $Z^{n-1}(G; A) \to Z^{n-1}(L; A)$ is surjective.*

Proof. (1) This follows from the exact sequence.

(2) Suppose $Z^{n-1}(G; A) \to Z^{n-1}(L; A)$ is surjective. Pick $\mu \in Z^n(G, L; A)$ such that $[\mu] \in \ker(f_n)$. Then $\mu \in B^n(G; A)$ and $\mu = d\eta$ for some $\eta \in C^{n-1}(G; A)$. Moreover, $d(\eta|_L) = \mu|_L \equiv 0$ so that $\eta|_L \in Z^{n-1}(L; A)$. Our assumption gives $\zeta \in Z^{n-1}(G; A)$ such that $\zeta|_L = \eta|_L$. Hence, $\eta - \zeta \in C^{n-1}(G, L; A)$ and $\mu = d(\eta - \zeta) \in B^n(G, L; A)$.

Now suppose that f_n is injective. Pick $\mu \in Z^{n-1}(L; A)$, and extend it to $\chi \in C^{n-1}(G; A)$. Hence $d\chi \in Z^n(G, L; A)$ and $[d\chi] \in \ker(f_n)$. So $d\chi = d\zeta$ for some $\zeta \in C^{n-1}(G, L; A)$. Now $\chi - \zeta \in Z^{n-1}(G; A)$ and $(\chi - \zeta)|_L = \mu$. □

Corollary 4. *For $n > 1$, $H^n(G, L; A) = 0$ if and only if $H^{n-1}(G; A) \to H^{n-1}(L; A)$ is surjective and $H^n(G; A) \to H^n(L; A)$ is injective. Furthermore, $H^1(G, L; A) = 0$ if and only if $H^1(G; A) \to H^1(L; A)$ is injective.*

The next theorem clarifies the origin of this new complex. Let us fix a homomorphism $\theta = f|_L : L \to N$ and choose a subgroup $\widetilde{H} \leq H$, normal in $N = N_K(H)$ such that $A := H/\widetilde{H}$ is abelian. Notice that the conjugation ${}^{gH}h\widetilde{H} := ghg^{-1}\widetilde{H}$ defines a structure of an N/H-module (and a G-module via any weak (L, H)-morph) on A. Informally, we should think of the next theorem as "an exact sequence"

(1) $\quad H^1(G, L; A) \dashrightarrow [L\widetilde{H}]^\theta\mathrm{mo}(G, N) \to [LH]^\theta\mathrm{mo}(G, N) \to H^2(G, L; A)$

keeping in mind that the second and the third terms are sets (not even pointed sets) and the first arrow is an "action" rather than a map. Let us make it more precise: a weak (L, H)-morph defines a G-module structure ρ on A. For each particular ρ (not just its isomorphism class) we define

$$[L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho \subseteq [L\widetilde{H}]^\theta\mathrm{mo}(G, N), \quad [LH]^\theta\mathrm{mo}(G, N)_\rho \subseteq [LH]^\theta\mathrm{mo}(G, N)$$

as subsets of those weak (L, H)-morphs that define this particular G-action ρ. These subsets could be empty, in which case we consider the following theorem true for trivial reasons. The reader should consider this theorem and its proof as a generalisation of the results in Sections 1 and 2 in [Thévenaz 1983] to the situation of weak (L, H)-morphs.

Theorem 5. *We retain the notation preceding this theorem. For each G-action ρ on A the following statements hold*:

(1) *There is a restriction map*

$$\mathrm{Res} : [L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho \to [LH]^\theta\mathrm{mo}(G, N)_\rho, \quad \mathrm{Res}(\langle f\rangle) = [f],$$

where $\langle f\rangle$ and $[f]$ are the equivalence classes in $[L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho$ and $[LH]^\theta\mathrm{mo}(G, N)_\rho$.

(2) *The abelian group $Z^1(G, L; (A, \rho))$ acts freely on the set $[L\widetilde{H}]^\theta\,\mathrm{mo}(G, N)_\rho$ by*

$$\gamma \cdot \langle f\rangle := \langle \dot{\gamma} f\rangle, \text{ where } \dot{\gamma} f(x) = \dot{\gamma}(x) f(x) \text{ for all } x \in G$$

and $\dot{\gamma} : G \xrightarrow{\gamma} A \to H$ is a lift of γ to a map $G \to H$ with $\dot{\gamma}(1) = 1$.

(3) *The corestricted restriction map* $\mathrm{Res} : [L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho \to \mathrm{Im}(\mathrm{Res})$ *is a quotient map by the $Z^1(G, L; (A, \rho))$-action.*

(4) *Two classes $\langle f\rangle, \langle g\rangle \in [L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho$ lie in the same $B^1(G, L; (A, \rho))$-orbit if and only if there exist $h \in H$, $f' \in \langle f\rangle$, $g' \in \langle g\rangle$ such that $[f(L), h] \subset \widetilde{H}$ and $f'(x) = hg'(x)h^{-1}$ for all $x \in G$.*

(5) *There is an obstruction map*

$$\mathrm{Obs} : [LH]^\theta\mathrm{mo}(G, N)_\rho \to H^2(G, L; (A, \rho)), \quad \mathrm{Obs}([f]) = [f^\sharp]$$

where the cocycle $f^\sharp$ is defined by $f^\sharp(x, y) = f(x)f(y)f(xy)^{-1}\widetilde{H}$.

(6) *The sequence* (1) *is exact, i.e., the image of* Res *is equal to* $\mathrm{Obs}^{-1}([0])$.

Proof. Suppose $\langle f\rangle = \langle g\rangle$. This gives a function $\alpha : G \to \widetilde{H}$ such that $\alpha|_L \equiv 1$ and $f(x) = \alpha(x)g(x)$ for all $x \in G$. Since $H \supseteq \widetilde{H}$, we conclude that $[f] = [g]$ and the map Res is well-defined. This proves (1).

Suppose $\mathrm{Res}(\langle f\rangle) = \mathrm{Res}(\langle g\rangle)$. Then $[f] = [g]$ gives a function $\alpha : G \to H$ such that $\alpha|_L \equiv 1$ and $f(x) = \alpha(x)g(x)$ for all $x \in G$. We can also obtain such a function from a cochain $\gamma \in C^1(G, L; (A, \rho))$ by lifting $\alpha = \dot{\gamma}$. Let us compute in the group $N/\widetilde{H}$ denoting $a\widetilde{H}$ by $\bar{a}$. The weak (L, H)-morph condition for f is equivalent to the following equality:

$$\begin{aligned}\overline{\alpha(xy)}\,\overline{g(xy)} = \overline{f(xy)} = \overline{f(x)}\,\overline{f(y)} &= \overline{\alpha(x)g(x)}\,\overline{\alpha(y)g(y)}\\ &= \overline{\alpha(x)g(x)\alpha(y)g(x)^{-1}}\,\overline{g(x)g(y)}.\end{aligned}$$

Now notice that

$$\overline{g(xy)} = \overline{g(x)g(y)} = \overline{g(x)}\,\overline{g(y)}$$

is the weak (L, H)-morph condition for g, while

$$\overline{\alpha(xy)} = \overline{\alpha(x)g(x)\alpha(y)g(x)^{-1}} = \overline{\alpha(x)}\,\overline{g(x)\alpha(y)g(x)^{-1}} = \overline{\alpha(x)}\,[\rho(x)(\bar{\alpha})](y)$$

is the cocycle condition for $\bar{\alpha} = \alpha\widetilde{H}$. Any two of these three conditions imply the third one, which proves both (2) and (3), except the action freeness.

Suppose $\langle f\rangle = \gamma \cdot \langle f\rangle = \langle \dot{\gamma} f\rangle$. This gives a function $\alpha : G \to \widetilde{H}$ such that $\alpha|_L \equiv 1$ and $\dot{\gamma}(x)f(x) = \alpha(x)f(x)$ for all $x \in G$. Hence, $\dot{\gamma} = \alpha$ and $\gamma = \bar{\alpha} \equiv 1$. Thus, the action is free.

Let us examine $da \cdot \langle f\rangle = \langle \dot{da}\, f\rangle$ for some $a \in A^L$. Since $da(x) = -a + \rho(x)(a)$ and $\rho(x)$ can be computed by conjugating with $f(x)$, we immediately conclude that

$$[\dot{da}\, f](x) = \dot{a}^{-1} f(x)\dot{a} f(x)^{-1} f(x) = \dot{a}^{-1} f(x)\dot{a}.$$

It is easy to see that $[f(L), \dot{a}] \subset \widetilde{H}$. The argument we have just given is reversible, i.e., if $f(x) = hg(x)h^{-1}$ then $\langle g\rangle = d\bar{h} \cdot \langle f\rangle$ and $\bar{h} \in A^L$. This proves (4).

Suppose $[f] = [g]$. This gives a function $\alpha : G \to H$ such that $\alpha|_L \equiv 1$ and $f(x) = \alpha(x)g(x)$ for all $x \in G$. Let us compute the cocycles in $N/\widetilde{H}$, keeping in mind that $H/\widetilde{H}$ is abelian:

$$\begin{aligned} f^\sharp(x, y) = \overline{f(x)f(y)}\,\overline{f(xy)^{-1}} &= \overline{\alpha(x)}\,\overline{g(x)}\,\overline{\alpha(y)}\,\overline{g(y)}\,\overline{g(xy)}^{-1}\overline{\alpha(xy)}^{-1}\\ &= \big(\overline{\alpha(xy)}^{-1}\overline{\alpha(x)}\,\overline{g(x)\alpha(y)}\,\overline{g(x)^{-1}}\big)\overline{g(x)g(y)}\,\overline{g(xy)^{-1}}\\ &= d\,\bar{\alpha}(x, y) + g^\sharp(x, y).\end{aligned}$$

Thus $[f^\sharp] = [g^\sharp]$, proving (5).

It is clear that $f^\sharp \equiv 1$ for $f \in [L\widetilde{H}]^\theta\mathrm{mo}(G, N)_\rho$. Hence, $\mathrm{Obs}(\mathrm{Res}(\langle f\rangle)) = [0]$. Suppose now that $\mathrm{Obs}([f]) = [0]$. This gives a function $\alpha : G \to H$ such that $\alpha|_L \equiv 1$

and $d\bar{\alpha} = f^\sharp$. Consider $g : G \to N$ defined by $g(x) = \alpha(x)^{-1} f(x)$ for all $x \in G$. Then $[g] = [f]$ and we can verify that $g \in [L\widetilde{H}]^\theta \mathrm{mo}(G, N)_\rho$ by checking $g^\sharp \equiv 1$ in $N/\widetilde{H}$:

$$\begin{aligned} g^\sharp(x, y) = \overline{\alpha(x)}^{-1}\overline{f(x)}\,\overline{\alpha(y)}^{-1}\overline{f(y)}\,\overline{f(xy)}^{-1}\overline{\alpha(xy)} \\ \sim \overline{\alpha(xy)}\,\overline{\alpha(x)}^{-1}\big(\overline{f(x)}\,\overline{\alpha(y)}\,\overline{f(x)}^{-1}\big)^{-1} f^\sharp(x, y) \\ = (d\,\bar{\alpha}(x, y))^{-1} f^\sharp(x, y) \equiv 1. \end{aligned}$$

This proves (6). □

Let us quickly reexamine how the last section works for (L, H)-morphs. All of its results including Theorem 5 clearly work, although the objects that appear have additional properties. Most crucially, since $f(L) \subseteq C_K(H)$, the L-action on the abelian group A is trivial. If L is normal in G, this just means that A is a $\mathbb{Z}G/L$-module.

An important feature is that $Z^1(L; A)$ consists of homomorphisms $L \to A$ in this case. This means that Proposition 3 yields the following corollary:

Corollary 6. *If the group L is perfect, then $f_1 : H^1(G, L; A) \to H^1(G; A)$ is surjective and $f_2 : H^2(G, L; A) \to H^2(G; A)$ is injective.*

1.3. ***Module extensions.*** We now assume that L is a normal subgroup of G. Let $\mathbb{A}$ be an associative ring, (V, θ) an $\mathbb{A}L$-module, $K = \mathrm{Aut}_{\mathbb{A}} V$ and $H = \mathrm{Aut}_{\mathbb{A}L} V$ its automorphism groups. We can think of θ as an element of the set of $\mathbb{A}L$-structures $X = \hom(L, K)$. Then H is the centraliser in K of $\theta(L)$. By N, as before, we denote the normaliser of H in K.

Naturally, X is a K-G-set: G acts by conjugation on L twisting the $\mathbb{A}L$-module structure, K acts by conjugations on the target, while $H = \mathrm{Stab}_K(\theta)$. The module V is called *G-stable* if $(V, \theta) \cong (V, \theta^g)$ for all $g \in G$. This is equivalent to the orbit inclusion $\theta^G \subseteq {}^K\theta$. By Lemma 2 this gives a $(1, H)$-morph $f : G \to K$.

If $g \in L$, the isomorphism $f(g) : (V, \theta) \to (V, \theta^g)$ can be chosen to be $\theta(g)$. Indeed,

$$\theta(g)(\theta(h)v) = \theta(gh)(v) = \theta(ghg^{-1})(\theta(g)(v)) = \theta^g(h)(\theta(g)(v))$$

for all $g, h \in L$. Then, without loss of generality $f|_L = \theta$, and f is an (L, H)-morph in $[LH]^\theta \mathrm{mo}(G, N)$.

Suppose that the group $H = \mathrm{Aut}_{\mathbb{A}L} V$ is soluble. We can always find a subnormal series $H = H_0 \rhd H_1 \rhd \ldots \rhd H_k = \{1\}$ with abelian quotients $A_j = H_{j-1}/H_j$ such that each H_j is normal in N. For instance, we can use the commutator series $H_j = H^{(j)}$. In this case, every abelian group A_j becomes an N-module.

If $\mathbb{A}$ is finite-dimensional over the field $\mathbb{K}$ and V is a finite-dimensional indecomposable $\mathbb{A}L$-module, we can use Proposition 1 to derive useful information about its automorphisms. In particular, if $\mathbb{D} = \mathrm{End}_{\mathbb{A}L}(V)/J$ is a separable field extension of $\mathbb{K}$, then $H = \mathrm{GL}_1(\mathbb{D}) \ltimes (1 + J)$ is soluble. It admits another standard N-stable subnormal series:

$$H_m = 1 + J^m, \ m \geq 1, \quad A_m = (1 + J^m)/(1 + J^{m+1}).$$

As groups, we have $A_m = ((1+J^m)/(1+J^{m+1}), \cdot) \cong (J^m/J^{m+1}, +)$. The following theorem is the direct application of Theorem 5. It determines the uniqueness and existence of a G-module structure on a G-stable L-module. The proof is obvious.

Theorem 7. *Let $V = (V, \theta)$ be a G-stable $\mathbb{A}L$-module with a soluble automorphism group H, where $\mathbb{A}$ is an associative ring. Let $H = H_0 \rhd H_1 \rhd \ldots \rhd H_k = \{1\}$ be a subnormal N-stable series with abelian factors $A_j = H_{j-1}/H_j$.*

Any $\mathbb{A}G$-module structure Θ on (V, θ) compatible with its $\mathbb{A}L$-structure (i.e., $\Theta|_{\mathbb{A}L} = \theta$) can be discovered by the following recursive process in k steps. One initialises the process with an (L, H_0)-morph $f_0 = f$ coming from the G-stability. The step m is the following.

(1) *The (L, H_{m-1})-morph $f_{m-1} : G \to N$ such that $f_{m-1}|_L = \theta$ determines a G-module structure ρ_m on A_m.*
(2) *If* $\mathrm{Obs}([f_{m-1}]) \neq 0 \in H^2(G, L; (A_m, \rho_m))$, *then this branch of the process terminates.*
(3) *If* $\mathrm{Obs}([f_{m-1}]) = 0 \in H^2(G, L; (A_m, \rho_m))$, *then we choose an (L, H_m)-morph $f_m : G \to N$ such that* $\mathrm{Res}([f_m]) = [f_{m-1}]$.
(4) *For each element of $H^1(G, L; (A_m, \rho_m))$ we choose a different f_m branching the process. (The choices different by an element of $B^1(G, L; (A_m, \rho_m))$ are equivalent, not requiring the branching.)*
(5) *We change m to $m + 1$ and go to step* (1).

An $\mathbb{A}G$-module structure Θ on (V, θ) compatible with its $\mathbb{A}L$-structure is equivalent to f_k for one of the nonterminated branches. Distinct nonterminated branches produce (as f_k) nonequivalent compatible $\mathbb{A}G$-module structures.

This process is subtle as ρ_m is revealed only when f_{m-1} is computed. It would be useful to have stability, i.e., the fact the G-modules (A_m, ρ_m) are the same (isomorphic) for different branches. The actions ρ_m on $A_m = H_{m-1}/H_m$ on different branches differ by conjugation via a function $G \to H_{m-2}$. Thus, one needs all two-step quotients H_{m-1}/H_{m+1} to be abelian to ensure stability. Having said that, we can still have some easy criteria for existence, uniqueness and nonuniqueness.

Corollary 8 (existence test). *Suppose $H^2(G, L; (A_m, \rho_m)) = 0$ for all m for one of the branches. Then this branch does not terminate and an $\mathbb{A}G$-module structure exists.*

Corollary 9 (uniqueness test). *Suppose $H^1(G, L; (A_m, \rho_m)) = 0$ for all m for one of the nonterminating branches. Then this branch is the only branch. Moreover, if an $\mathbb{A}G$-module structure exists, it is unique up to an isomorphism.*

Corollary 10 (nonuniqueness test). *Suppose $H^1(G, L; (A_k, \rho_k)) \neq 0$ for one of the nonterminating branches. Then there exist nonequivalent $\mathbb{A}G$-module structures.*

1.4. ***Extension from not necessarily normal subgroups.*** In Section 1.3 we restrict our attention to the case of L being a normal subgroup of G. Let us take a moment to examine how Section 1.3 works if L is not normal.

Set $P := \bigcap_{g\in G} L^g$, where $L^g := g^{-1}Lg$. Let $\mathbb{A}$ be an associative ring, (V, θ) an $\mathbb{A}L$-module. Note that (V, θ) is also an $\mathbb{A}P$-module under restriction, so we can view θ as an element of the set $X = \hom(P, K)$. Let $K = \mathrm{Aut}_{\mathbb{A}} V$ and $H = \mathrm{Aut}_{\mathbb{A}P} V$ be its automorphism groups, so H is the centraliser in K of $\theta(P)$. By N, as before, we denote the normaliser of H in K.

As in Section 1.3, X is a K-G-set. The $\mathbb{A}L$-module V is called *G-stable-by-conjugation* if $(V, \theta) \cong (V, \theta^g)$ as $\mathbb{A}[L \cap L^g]$-modules for all $g \in G$. Note that this condition guarantees that V is G-stable as an $\mathbb{A}P$-module. This is equivalent to the orbit inclusion $\theta^G \subseteq {}^K\theta$. By Lemma 2 this gives a $(1, H)$-morph $f : G \to K$.

If $g \in L$, the $\mathbb{A}[L \cap L^g]$-isomorphism $f(g) : (V, \theta) \to (V, \theta^g)$ can be chosen to be $\theta(g)$. Indeed, $\theta(g)(\theta(h)v) = \theta(gh)(v) = \theta(ghg^{-1})(\theta(g)(v)) = \theta^g(h)(\theta(g)(v))$ for $g \in L$, $h \in L \cap L^g$. Then, without loss of generality $f|_L = \theta$, and f is an (L, H)-morph in $[LH]^\theta\mathrm{mo}(G, N)$.

This then allows us to proceed with the inductive process of Theorem 7 as before, when $H = \mathrm{Aut}_{\mathbb{A}P} V$ is soluble.

1.5. ***Comparison with $C^\bullet(G/L; A)$.*** When studying the question of extending representations from a normal subgroup, Dade and Thévenaz use the cohomology of the cochain complex $(C^\bullet(G/L; A), d)$ to control existence and uniqueness of such extensions. In this paper, however, we use the cohomology complex $(C^\bullet(G, L; A), d)$ instead. It is worth taking a moment to compare the cohomology of these two complexes, and see where the difference in approaches arises. We use the notation of Section 1.2, assuming that cochains are normalised since this does not affect the cohomology groups.

In order for the action of G/L on A to make sense, we need to make the assumption that L acts on A trivially. The reader can observe that this assumption holds in the case considered in Section 1.3, and, in fact, holds whenever one obtains the G-action on A from an (L, H)-morph as opposed to a weak (L, H)-morph. With this assumption, we have the following proposition.

Proposition 11. *Under the aforementioned conditions we have isomorphisms of groups $H^0(G, L; A) \cong H^0(G/L; A)$ and $H^1(G, L; A) \cong H^1(G/L; A)$.*

Proof. It is easy to see that $H^0(G, L; A) = A^G = H^0(G/L; A)$. The natural map from the group of normalised cochains

$$\inf : \widehat{C}^1(G/L; A) \to C^1(G, L; A), \quad \inf(\mu)(g) = \mu(gL),$$

defines a map $\mathrm{Inf} := [\inf] : H^1(G/L; A) \to H^1(G, L; A)$ of cohomology groups. It is injective because $\mathrm{Inf}([\mu]) = 0$ means that $\inf(\mu) = da$ for some $a \in A$. Then $\mu = da$ and $[\mu] = 0$.

It is surjective because for $\eta \in Z^1(G, L; A)$ we have $d\eta = 0$ that translates as

$$\eta(gh) = {}^g(\eta(h)) + \eta(g) \quad \text{for all } g, h \in G.$$

If one chooses $h \in L$, then it tells us that $\eta(gh) = \eta(g)$, i.e., that η is constant on L-cosets. Thus, the cocycle

$$\mu \in \widehat{Z}^1(G/L; A), \quad \mu(gL) := \eta(g)$$

is well-defined. By definition $\inf(\mu) = \eta$. □

Considering the second cohomology of these complexes, it is still possible to construct the inflation map $\mathrm{Inf} : H^2(G/L; A) \to H^2(G, L; A)$ in the natural way, but this map is no longer an isomorphism in general. We can still view $H^2(G/L; A)$ as a subgroup of $H^2(G, L; A)$:

Proposition 12. *The map* $\mathrm{Inf} : H^2(G/L; A) \to H^2(G, L; A)$ *is injective.*

Proof. If $\mathrm{Inf}([\eta]) = 0 \in H^2(G, L; A)$ then there exists $\mu \in C^1(G, L; A)$ such that $d\mu = \inf(\eta)$. Note that $\inf(\eta)$ is constant on $L \times L$-cosets by construction. In particular, for $g \in G$ and $h \in L$, we have

$$\begin{aligned} \mu(g) - \mu(gh) &= {}^g(\mu(h)) + \mu(g) - \mu(gh) = \inf(\eta)(g, h) \\ &= \inf(\eta)(g, 1) = \inf(\eta)(1, 1) = 0, \end{aligned}$$

using the cocycle condition in the penultimate equality. Hence, μ is constant on cosets of L in G. In particular, if we define $\widetilde{\mu} \in \widehat{C}^1(G/L; A)$ by $\widetilde{\mu}(gL) = \mu(g)$ then we obtain that $\eta = d\widetilde{\mu}$ and so $[\eta] = 0 \in H^2(G/L; A)$. □

In the context of Theorem 5, we can see that $H^2(G/L; A)$ and $H^2(G, L; A)$ can be made to play the same role in certain key cases. To that end, we say that an (L, H)-morph f is *normalised* if $f(gh) = f(g)f(h)$ whenever $g \in G$ and $h \in L$. Note that this definition is independent of the subgroup H.

Lemma 13. *In the context of Theorem 7, the (L, H_i)-morphs f_i can be assumed to be normalised for each i. Furthermore, with this assumption, the cocycles $f_i^\sharp \in Z^2(G, L; A_{i+1})$ are constant on cosets of $L \times L$ in $G \times G$.*

Proof. These results follow from Lemmas 9.2 and 9.4(i) in [Karpilovsky 1989]. □

For the rest of this section we assume all morphs are normalised. The second statement of Lemma 13 immediately yields that, given an (L, H)-morph f, $\mathrm{Obs}([f])$ lies in the image of the natural homomorphism $\mathrm{Inf}: H^2(G/L; A) \to H^2(G, L; A)$. The discussion in this section yields the following result:

Corollary 14. *Let f be a normalised (L, H)-morph. There exists $\eta \in Z^2(G/L; A)$ with* $\mathrm{Inf}([\eta]) = \mathrm{Obs}([f])$. *Furthermore*, $\mathrm{Obs}([f]) = 0 \in H^2(G, L; A)$ *if and only if* $[\eta] = 0 \in H^2(G/L; A)$.

Combining Proposition 11 and Corollary 14, we observe that Sections 1.2 and 1.3 could be interpreted using the cochain complex $C^\bullet(G/L; A)$ at all points instead of the complex $C^\bullet(G, L; A)$ (although doing so would force us to work exclusively with normalised morphs instead of not-necessarily-normalised weak morphs). Indeed, this is the approach taken by Dade and Thévenaz in the contexts they consider. Our reasons for not taking this approach are threefold. Firstly, our new complex fits nicely into an exact sequence as described in Section 1.2. Secondly, this complex is more natural to work with — Dade and Thévenaz essentially move from the complex $C^\bullet(G/L; A)$ to the complex $C^\bullet(G, L; A)$ as described in this section, and then proceed as we do. Finally, our main motivation in studying the case for abstract groups is to gain insight into the question for algebraic groups, where the procedures described in this section do not work smoothly (cf. Section 2.5).

In particular, note that if H is abelian then the corollaries at the end of Section 1.3 give precisely Corollary 1.8 and Proposition 2.1 in [Thévenaz 1983].

2. *G*-stable modules for algebraic groups

In this chapter we consider algebraic groups over an algebraically closed field $\mathbb{K}$ of positive characteristic p. Algebraic groups are affine and reduced, groups schemes are affine and not necessarily reduced.

2.1. ***Rational and algebraic G-modules.*** We distinguish algebraic and rational maps of algebraic varieties. In particular, we can talk about algebraic and rational homomorphisms of algebraic groups $f : G \to H$. The latter are defined on an open dense subset $U = \mathrm{dom}(f)$ of G containing 1 and satisfy $f(x)f(y) = f(xy)$ whenever $x, y, xy \in U$.

A rational automorphic G-action on a commutative algebraic group H is a rational map $G \times H \to H$, defined on an open set $U \times H$ containing $1 \times H$, with the usual action conditions and also such that for each $g \in U$ the map $x \mapsto {}^g x$ is a group automorphism of H. An algebraic G-action on H is the same, but where the map $G \times H \to H$ is algebraic.

In an important case, the distinction between rational and algebraic maps can be essentially forgotten, as observed by Rosenlicht [1956].

Lemma 15 [Rosenlicht 1956, Theorem 3]. *Let G and H be algebraic groups with G connected. Suppose $f : G \to H$ is a rational homomorphism. Then f extends uniquely to an algebraic group homomorphism $G \to H$.*

When H is commutative, this lemma is a special case of the next lemma. Indeed, if one takes the G-action on H to be trivial, then the condition in the following lemma is precisely the condition for a map to be a homomorphism.

Lemma 16. *Suppose that G is a connected algebraic group and $(H, +)$ is a commutative algebraic group with an algebraic automorphic G-action ρ. Let $f : G \to H$ be a rational map such that $f(xy) = f(x) + {}^{x}f(y)$ for all $x, y, xy \in \operatorname{dom}(f)$ (where ${}^{x}f(y) := \rho(x)(f(y))$). Then f extends to an algebraic map satisfying $f(xy) = f(x) + {}^{x}f(y)$ for all $x, y \in G$.*

Proof. Since f is rational and G is connected, $\operatorname{dom}(f) = U \subset G$ is a dense open subset. Set $V = U \cap U^{-1}$.

Fix $x \in V$. Consider the rational map

$$f_x : G \to H, \qquad f_x(y) := f(yx) + {}^{yx}f(x^{-1}).$$

This map is rational since it is defined on the dense open set Vx^{-1}. Observe that on $V \cap Vx^{-1}$ we have that $f_x = f$ by the assumption on f. Now, let $x, z \in V$ and define the rational map

$$f_{x,z} : G \to H, \qquad f_{x,z}(y) := f_x(y) - f_z(y).$$

Then $f_{x,z}$ is defined on $Vx^{-1} \cap Vz^{-1}$. If the set $f_{x,z}^{-1}(H \setminus \{0\})$ is nonempty, it is open dense. Hence, it has nonempty intersection with $V \cap Vx^{-1} \cap Vz^{-1}$. However, since on $V \cap Vx^{-1} \cap Vz^{-1}$ we have $f = f_x = f_z$, this is impossible. Thus, we must have $f_{x,z} \equiv 0$ on $Vx^{-1} \cap Vz^{-1}$. In particular, if $y \in Vx^{-1} \cap Vz^{-1}$ then $f_x(y) = f_z(y)$.

Therefore, the following map is a well-defined locally-algebraic, and hence algebraic, map:

$$\widehat{f} : G \to H, \qquad \widehat{f}(y) := f_w(y), \text{ where } w \in y^{-1}V.$$

This map clearly restricts to f on V. Furthermore, it satisfies the condition from the lemma:

Let $a, b \in G$. Choose $w \in b^{-1}a^{-1}V \cap b^{-1}V$ — this exists since both these sets are open dense in G. We then have $abw \in V$ and $bw \in V$. The condition on f tells us that $0 = f(1) = f(bw) + {}^{bw}f(w^{-1}b^{-1})$. Hence, we have the equations

$$\begin{aligned}
\widehat{f}(ab) &= f_w(ab) = f(abw) + {}^{abw}f(w^{-1}),\\
\widehat{f}(a) &= f_{bw}(a) = f(abw) + {}^{abw}f(w^{-1}b^{-1}),\\
{}^{a}\widehat{f}(b) &= {}^{a}f_w(b) = {}^{a}f(bw) + {}^{abw}f(w^{-1}).
\end{aligned}$$

This then gives us that $\widehat{f}(ab) = \widehat{f}(a) + {}^{a}\widehat{f}(b)$, as required. □

Recall that a rational[1] representation of an algebraic group G is a vector space V, equipped with an algebraic homomorphism $\theta : G \to \mathrm{GL}(V)$. An immediate consequence of Lemma 15 is that if G is connected, then θ is uniquely determined by any of its restrictions to an open subset and any rational homomorphism of algebraic groups $G \to GL(V)$ determines a representation.

Similar to the case of abstract groups, we have the following proposition:

Proposition 17 [Xanthopoulos 1992, Section 4.3]. (cf. Proposition 1) *Suppose that V is a finite-dimensional indecomposable $\mathfrak{g}$-module, where $\mathfrak{g}$ is the Lie algebra of the algebraic group G over $\mathbb{K}$. Then as algebraic groups we have*

$$\mathrm{Aut}_{\mathfrak{g}}(V) = \mathbb{K}^{\times} \times (1 + J),$$

where J is the Jacobson radical of $\mathrm{End}_{\mathfrak{g}}(V)$. *Furthermore*, $1 + J$ *is a connected unipotent algebraic subgroup of* $\mathrm{Aut}_{\mathfrak{g}}(V)$.

2.2. *Rational and algebraic cohomologies.* Let H be an affine group scheme acting on an additive algebraic group $(A, +)$ algebraically by automorphisms. The following easy lemma shall be useful in what follows.

Lemma 18. *Let H be an irreducible affine group scheme. Then H is primary, i.e., every zero-divisor in $\mathbb{K}[H]$ lies inside the nilradical.*

Proof. The affinity of H tells us that $\mathbb{K}[H] = \mathbb{K}[y_1, \dots, y_n]/I$ for some $n \geq 1$ and some Hopf ideal I. In particular, I has a primary decomposition $I = Q_0 \cap \cdots \cap Q_r$ (which we assume to be normal) with associated primes $P_0 = \sqrt{I}, P_1, \dots, P_r$. From the perspective of group schemes, this uniquely endows H with a finite collection $p_0, p_1, \dots, p_r$ of embedded points of H, where p_i is a generic point of the irreducible closed subscheme given by Q_i. Furthermore, for $i > 0$ each p_i is of codimension at least one. If x is a closed point in H, then the set $xp_0, xp_1, \dots, xp_r$ corresponds to the associated primes of another primary decomposition of I. Hence, by uniqueness, x acts on the set $p_0, p_1, \dots, p_r$ by permutation. Thus,

$$H_{\mathrm{red}} = \bigcup_{i=1}^{r} \Big(\bigcup_{x \text{ closed point}} x p_i \Big)_{\mathrm{red}} = \bigcup_{i=1}^{r} (p_i)_{\mathrm{red}}.$$

However, over an algebraically closed field, H_{red} cannot be a finite union of proper subvarieties. Hence, $r = 0$ and the result follows. $\square$

Define the cochain complex $(C^n_{\mathrm{Rat}}(H; A), d)$ to consist of the rational maps $H^n \to A$ defined at $(1, 1, \dots, 1)$ with the standard differentials of group cohomology.

A rational function f on H^n is defined on an open dense subset $U \subseteq H^n$, thus, U has a nonempty intersection $U_\alpha = U \cap H^n_\alpha$ with each irreducible component H^n_α

[1] It is a standard terminology, which slightly disagrees with our usage of the adjective *rational*.

of H^n. Since H^n is a group scheme, its irreducible components are connected components that yields the direct sum decomposition of functions:

$$\mathbb{K}[H^n] = \oplus_\alpha \mathbb{K}[H^n_\alpha].$$

Note that each H_α is isomorphic to an irreducible affine group scheme, so we can apply Lemma 18. Thus, U_α is of the form $U(s_\alpha)$ for a non-zero-divisor $s_\alpha \in \mathbb{K}[H^n_\alpha]$ and $f = hs^{-1}$ for some $h \in \mathbb{K}[H^n]$ and a non-zero-divisor $s := (s_\alpha) \in \mathbb{K}[H^n]$. Thus, $f \in \mathbb{K}[H^n]_S$, the localised ring of functions obtained by inverting the set S of all non-zero-divisors.

Writing functions on the algebraic group A as $\mathbb{K}[A] = \mathbb{K}[x_1, \dots x_m]/I$, a rational n-cochain μ is uniquely determined by an m-tuple of rational functions $(\mu_i) \in \mathbb{K}[H^n]_S^m$ satisfying the relations of I. In particular, if each component of H is infinitesimal,

$$\mathbb{K}[H^n]_S = \mathbb{K}[H^n] \quad \text{and} \quad C^n_{\text{Rat}}(H; A) = C^n_{\text{Alg}}(H; A),$$

where, in general, $(C^n_{\text{Alg}}(H; A), d)$ is the cochain subcomplex of $(C^n_{\text{Rat}}(H; A), d)$ that consists of those rational maps $H^n \to A$ which are, in fact, algebraic.

Let us now concentrate on a connected algebraic group G and its connected subgroup scheme L. There is another subcomplex of $(C^n_{\text{Rat}}(G; A), d)$ which we are interested in: we define $(\widetilde{C}^\bullet_{\text{Rat}}(G, L; A), d)$ to consist of rational maps $G^n \to A$ that are trivial on L^n (i.e., everywhere 0 on L^n). As in the case of abstract groups, we define $(C^\bullet_{\text{Rat}}(G, L; A), d)$ by

$$C^n_{\text{Rat}}(G, L; A) = \begin{cases} \widetilde{C}^n_{\text{Rat}}(G, L; A), & \text{if } n > 0, \\ A^L, & \text{if } n = 0. \end{cases}$$

There is a natural inclusion of cochain complexes $C^\bullet_{\text{Rat}}(G, L; A) \to C^\bullet_{\text{Rat}}(G; A)$. We can hence define the cochain complex $\widetilde{C}^\bullet_{\text{Rat}}(L; A)$ such that $\widetilde{C}^n_{\text{Rat}}(L; A) := \text{Coker}(C^n_{\text{Rat}}(G, L; A) \to C^n_{\text{Rat}}(G; A))$ for all $n \geq 0$.

In particular, this gives us the short exact sequence of cochain complexes

$$0 \to C^\bullet_{\text{Rat}}(G, L; A) \to C^\bullet_{\text{Rat}}(G; A) \to \widetilde{C}^\bullet_{\text{Rat}}(L; A) \to 0.$$

We define the algebraic complexes $C^\bullet_{\text{Alg}}(G, L; A)$ and $\widetilde{C}^\bullet_{\text{Alg}}(L; A)$ in the expected way, and once again get a short exact sequence of cochain complexes. In either case, this allows us to construct the long exact sequence in cohomology (suppressing the "Rat" and "Alg"):

$$\begin{aligned} (2) \qquad 0 \to H^1(G, L; A) \to \dots \to \widetilde{H}^{n-1}(L; A) \\ \to H^n(G, L; A) \to H^n(G; A) \to \widetilde{H}^n(L; A) \to \dots \end{aligned}$$

Note that $\widetilde{H}^0_{\text{Rat}}(L; A) = \widetilde{H}^0_{\text{Alg}}(L; A) = 0$, hence this exact sequence starts in degree one.

These long exact sequences can be connected, using the maps induced by the inclusions $C^n_{\mathrm{Alg}}(G, L; A) \hookrightarrow C^n_{\mathrm{Rat}}(G, L; A)$ and $C^n_{\mathrm{Alg}}(G; A) \hookrightarrow C^n_{\mathrm{Rat}}(G; A)$:

$$\begin{array}{ccccccccccc}
\dots & \to & H^n_{\mathrm{Alg}}(G, L; A) & \to & H^n_{\mathrm{Alg}}(G; A) & \to & \widetilde{H}^n_{\mathrm{Alg}}(L; A) & \to & H^{n+1}_{\mathrm{Alg}}(G, L; A) & \to & \dots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
\dots & \to & H^n_{\mathrm{Rat}}(G, L; A) & \to & H^n_{\mathrm{Rat}}(G; A) & \to & \widetilde{H}^n_{\mathrm{Rat}}(L; A) & \to & H^{n+1}_{\mathrm{Rat}}(G, L; A) & \to & \dots
\end{array}$$

Since we identify $C^0_{\mathrm{Alg}}(G; A)$ with algebraic maps from the trivial algebraic group to A (and similarly in the other complexes), there is no distinction between rational and algebraic maps. Hence,

$$H^0_{\mathrm{Rat}}(G; A) = H^0_{\mathrm{Alg}}(G; A) = H^0_{\mathrm{Rat}}(G, L; A) = H^0_{\mathrm{Alg}}(G, L; A) = A^G.$$

The cocycle condition on $f \in C^1_{\mathrm{Rat}}(G; A)$ is precisely the condition considered in Lemma 16 for a rational map $f : G \to A$. Since G is connected, Lemma 16 tells us the map extends to an algebraic map. Hence, in this case

$$H^1_{\mathrm{Rat}}(G; A) = H^1_{\mathrm{Alg}}(G; A) \quad \text{and} \quad H^1_{\mathrm{Rat}}(G, L; A) = H^1_{\mathrm{Alg}}(G, L; A).$$

This leads to the following proposition. The first part of it follows from the exact sequence. The second part has a similar proof to Proposition 3.

Proposition 19 (cf. Proposition 3). (1) *If* $\widetilde{H}^1_{\mathrm{Rat}}(L; A) = 0$, *then* $H^1_{\mathrm{Rat}}(G, L; A) = H^1_{\mathrm{Rat}}(G; A)$.

(2) *For* $n > 0$, *if the natural map* $Z^{n-1}_{\mathrm{Rat}}(G; A) \to \widetilde{Z}^{n-1}_{\mathrm{Rat}}(L; A)$ *is surjective*, *then the natural map* $H^n_{\mathrm{Rat}}(G, L; A) \to H^n_{\mathrm{Rat}}(G; A)$ *is injective.*

The appropriate long exact sequence yields the following.

Corollary 20. $H^2_{\mathrm{Rat}}(G, L; A) = 0$ *if and only if* $H^1_{\mathrm{Rat}}(G; A) \to \widetilde{H}^1_{\mathrm{Rat}}(L; A)$ *is surjective and* $H^2_{\mathrm{Rat}}(G; A) \to \widetilde{H}^2_{\mathrm{Rat}}(L; A)$ *is injective.*

When the action is trivial, we can learn more about what these cohomology groups are.

Lemma 21. *If* G *acts trivially on* A *and* $\mathrm{Hom}(L, A) = 0$, *then* $\widetilde{Z}^1_{\mathrm{Rat}}(L; A) = 0$.

Proof. Let $\mu + C^1_{\mathrm{Rat}}(G, L; A) \in \widetilde{Z}^1_{\mathrm{Rat}}(L; A)$, so $d\mu \in C^2_{\mathrm{Rat}}(G, L; A)$. In particular, $d\mu|_{L^2} = 0$. However, since the action is trivial, $d\mu|_{L^2} = 0$ if and only if $\mu|_L$ is a rational homomorphism $L \to A$ if and only if $\mu|_L$ is a homomorphism $L \to A$ (since L is connected, by assumption). Since $\mathrm{Hom}(L, A) = 0$, we conclude that $\mu + C^1_{\mathrm{Rat}}(G, L; A) = 0 + C^1_{\mathrm{Rat}}(G, L; A)$. Hence, $\widetilde{Z}^1_{\mathrm{Rat}}(L; A) = 0$. □

Lemma 22. *Let* G *be a connected algebraic group which acts trivially on a commutative algebraic group* A. *Let* $L \leq G$ *be a closed connected subgroup scheme. Then* $H^1_{\mathrm{Rat}}(G; A) = \mathrm{Hom}(G, A)$ *and* $H^1_{\mathrm{Rat}}(G, L; A) = \{\mu \in \mathrm{Hom}(G, A) \mid \mu|_L \equiv 0\}$.

Proof. Since the G-action on A is trivial, the coboundary map $C^0_{\mathrm{Rat}}(G; A) \to C^1_{\mathrm{Rat}}(G; A)$ is just the trivial map. Hence, we get that $H^1_{\mathrm{Rat}}(G; A) = Z^1_{\mathrm{Rat}}(G; A)$, the rational 1-cocycles of G. However, as the action is trivial, rational 1-cocycles of G on A are the same as homomorphisms of algebraic groups $G \to A$. Hence, $H^1_{\mathrm{Rat}}(G; A) = \mathrm{Hom}(G, A)$.

A similar argument gives $H^1_{\mathrm{Rat}}(G, L; A) = \{\mu \in \mathrm{Hom}(G, A) \mid \mu|_L \equiv 0\}$. □

Combining Lemmas 21, 22 and Proposition 19 (2), we get the following corollary:

Corollary 23. *Let G be a connected algebraic group acting algebraically* (*not necessarily trivially*) *by automorphisms on a commutative algebraic group A. Let $L \leq G$ be a connected closed subgroup scheme of G such that the action of L on A is trivial*, *and* $\mathrm{Hom}(L, A) = 0$. *Then $H^1_{\mathrm{Rat}}(G, L; A) = H^1_{\mathrm{Alg}}(G; A)$ and $H^2_{\mathrm{Rat}}(G, L; A) \to H^2_{\mathrm{Rat}}(G; A)$ is injective.*

The following lemma by van der Kallen [1973, Proposition 2.2] will be useful:

Lemma 24. *Let G be a semisimple*, *simply-connected algebraic group. Suppose further that*, *if $p = 2$*, *the Lie algebra $\mathfrak{g}$ of G does not contain A_1*, *B_2 or C_l $(l \geq 3)$ as a direct summand. Then $\mathfrak{g}$ is perfect*, *i.e.*, $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

Proof. It is enough to prove this result for G simple and simply-connected, with irreducible root system Φ. It is well known that $\mathfrak{g}$ is simple and nonabelian (and so $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$) in the following cases: $p \nmid l + 1$ in type A_l, $p \neq 2$ in types B_l, C_l, D_l, $p \neq 2, 3$ in types E_6, E_7, F_4, G_2, and $p \neq 2, 3, 5$ in type E_8. It is further known [Capdeboscq et al. 2017] that $\mathfrak{g}$ is simple and nonabelian in the following cases: $p = 2$ in types E_6, G_2, $p = 3$ in types E_7, F_4, and $p = 2, 3, 5$ in type E_8.

Furthermore, it is known from Table 1 in [Hogeweij 1982] that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ in all the remaining cases except for $p = 2$ in types A_1, B_2, C_l $(l \geq 3)$. □

Lemma 25. *Let G be a semisimple*, *simply-connected algebraic group over an algebraically closed field $\mathbb{K}$ of characteristic p which acts trivially on a commutative algebraic group A. Suppose further that*, *if $p = 2$*, *the Lie algebra $\mathfrak{g}$ of G does not contain A_1*, *B_2 or C_l $(l \geq 3)$ as a direct summand. Let G_1 be the first Frobenius kernel of G. Then $H^2_{\mathrm{Rat}}(G, G_1; A) = 0$.*

Proof. Let us first show that $H^2_{\mathrm{Rat}}(G; A) = 0$. Let $\mu : G \times G \to A$ be a rational cocycle defined on the open set $U \times U$ with $U^{-1} = U$. We can define a local group structure on the set $A \times G$ by setting

$$(a, g)(b, h) = (a + b + \mu(g, h), gh) \quad \text{and} \quad (a, g)^{-1} = (-a - \mu(g, g^{-1}), g^{-1}).$$

In the language of Weil [1955], $A \times U$ is a group-chunk in the pre-group $A \times G$. By Weil's theorem [1955], there exists an algebraic group H birationally equivalent to $A \times U$ with $\Phi : A \times U \to \Phi(A \times U)$ an isomorphism of algebraic group-chunks and $\Phi(A \times U)$ a dense open set in H.

Since H is connected it is generated by $\Phi(A\times U)$. Let $f:A\to H$ be the natural algebraic group homomorphism coming from $A\to A\times U$. This is clearly injective and, since A commutes with each element of $A\times U$, we have that $f(A)\subset Z(H)$. Furthermore, the natural projection $A\times U\to G$ extends to a rational (and so algebraic) homomorphism $\pi:H\to G$, which is surjective as U generates G (since G connected). Finally, it is clear that $f(A)=\ker\pi\cap\Phi(A\times U)$. Hence, π descends to a homomorphism $\bar{\pi}:H/f(A)\to G$, whose kernel is discrete (since $\Phi(A\times U)$ is dense in H) and, hence, central (as G connected).

In other words, we have a central extension $1\to A\to H\to G\to 1$ of algebraic groups, which corresponds to an algebraic cocycle $\widetilde{\mu}:G\times G\to A$. It is straightforward to see that $\widetilde{\mu}|_{U\times U}=\mu|_{U\times U}$, and hence $[\mu]$ lies in the image of the natural map $H^2_{\mathrm{Alg}}(G;A)\to H^2_{\mathrm{Rat}}(G;A)$. Thus the map $H^2_{\mathrm{Alg}}(G;A)\to H^2_{\mathrm{Rat}}(G;A)$ is surjective.

It suffices to prove that $H^2_{\mathrm{Alg}}(G;A)=0$ when A is $\mathbb{G}_a$ or $\mathbb{G}_m$ or a finite group: the long exact sequence in cohomology reduces the case of arbitrary A to one of these cases. It is known that $H^2_{\mathrm{Alg}}(G;\mathbb{G}_a)=H^2(G;\mathbb{K}_{\mathrm{triv}})=0$ [Jantzen 1987, II.4.11].

Consider a nontrivial cohomology class in $H^2_{\mathrm{Alg}}(G;A)$ when A is $\mathbb{G}_m$ or a nontrivial finite group. It yields a nonsplit central extension $1\to A\to\widetilde{G}\to G\to 1$. Pick a nontrivial character $\chi:A\to\mathbb{G}_m$. There exists an irreducible representation of $\widetilde{G}$ with a central character χ. It is an irreducible projective representation of G. By the original version of Steinberg's tensor product theorem [1963] it is linear. Hence, χ is trivial. This contradiction proves that $H^2_{\mathrm{Alg}}(G;A)=0$ for these two particular A. We have finished the proof that $H^2_{\mathrm{Rat}}(G;A)=0$ for an arbitrary A.

Since G_1 is a height 1 group scheme, rational homomorphisms of schemes $G_1\to A$ are fully controlled by the corresponding restricted homomorphisms of Lie algebras $\mathfrak{g}\to\mathrm{Lie}(A)$. By Lemma 24, $\mathfrak{g}=[\mathfrak{g},\mathfrak{g}]$ and thus all such homomorphism of Lie algebras are trivial. Hence, we can apply Corollary 23 to get that $H^2_{\mathrm{Rat}}(G,G_1;A)\to H^2_{\mathrm{Rat}}(G;A)$ is injective, and so

$$H^2_{\mathrm{Rat}}(G,G_1;A)=0. \qquad \square$$

2.3. ***G-stable bricks.*** In Section 1, we have introduced the notions of weak (L,H)-morphs and (L,H)-morphs for abstract groups. In this section, we discuss how these notions apply to algebraic groups and see how they can be used to shed some light on the lifting of $\mathfrak{g}$-modules to G-modules.

Suppose that G, K are algebraic groups over $\mathbb{K}$, where G is connected, and that L, H are closed subgroup schemes of G, K respectively. We say that a rational map $f:G\to K$ is *a (weak) (L,H)-morph of algebraic groups* if it satisfies the conditions for a (weak) (L,H)-morph of abstract groups, where condition (3) is interpreted for only those $x, y, xy\in\mathrm{dom}(f)$.

In analogy with the case of abstract groups, a weak (L,H)-morph of algebraic groups is a homomorphism $G\to N/H$ with a rational lifting $N/H\to N$ which satisfies an additional condition. It is clear that if H is normal in K then condition (2) is trivially satisfied. We again have that weak $(L, 1)$-morphs are just homomorphisms $G \to K$, and that weak $(1, K)$-morphs are rational maps $G \to K$ which preserve the identity.

We say that two weak (L, H)-morphs of algebraic groups, f and g, are equivalent if $f(x)g(x)^{-1} \in H$ for all $x \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$. Given a homomorphism of algebraic groups $\theta : L \to K$, we denote by $[LH]^\theta\mathrm{mo}(G, K)$ the quotient by this equivalence relation of the set of weak (L, H)-morphs of algebraic groups from G to K which restrict to θ on L.

Suppose that X is a separated algebraic scheme on which G acts rationally on the right (i.e., the action $X \times G \to X$ is a rational map), K acts algebraically on the left, and the actions commute. Suppose further that $\theta \in X(\mathbb{K})$ is such that $\theta^G \subset {}^K\theta$, and that there exists a rational section $K/H \to K$ where $H = \mathrm{Stab}_K(\theta)$ is the scheme-theoretic stabiliser of θ.

As in the case for abstract groups, this gives us a rational map

$$f : G \to K \;\; \text{characterised by} \;\; {}^{f(x)}\theta = \theta^x \text{ for all } x \in U \overset{\text{open}}{\subset} G.$$

Lemma 26. *The map f defined above is a $(1, H)$-morph of algebraic groups.*

Proof. We can think of f as the composition of the rational maps

$$G \hookrightarrow \{\theta\} \times G \to {}^K\theta \to K/H \to K.$$

Note that ${}^K\theta \to K/H$ is an algebraic map by [Demazure and Gabriel 1970, Proposition 3.2.1]. We then have that the composition is rational since each domain of definition intersects the previous map's image.

The proof that $f(x)f(y) \in f(xy)H$ for $x, y \in G$ with $f(x)$, $f(y)$ and $f(xy)$ defined is the same as in the abstract case, as is the proof that $f(G) \subset N_K(H)$. $\square$

Now we fix algebraic (group, subgroup scheme) pairs (G, L) and (K, H) with H soluble and G connected. Denote by m_G, m_K the corresponding multiplication maps, Δ_G, Δ_K the diagonal embeddings, and $\mathrm{inv}_G, \mathrm{inv}_K$ the inverse maps. Let $\theta : L \to K$ be a homomorphism of algebraic group schemes. Furthermore, choose $\widetilde{H}$ to be an algebraic subgroup of H, characteristic in $N = N_K(H)$ such that $A := H/\widetilde{H}$ is commutative. We denote the quotient map $H \to A$ by π.

We can define an N-action on H by conjugation. Note that since $\widetilde{H}$ is characteristic in N, so preserved by conjugation, this passes to an algebraic N-action on A. Hence, we have an algebraic action of N on A which is trivial on H (since A is commutative). This gives us an algebraic N/H-action on A. For an element $f \in [LH]^\theta\mathrm{mo}(G, K)$, we get a rational homomorphism $G \to N/H$ which is, in fact, algebraic by Lemma 15. Thus, every element of $[LH]^\theta\mathrm{mo}(G, K)$ induces an algebraic G-action on A. This G-action respects the multiplication operation of A, i.e., it is an algebraic automorphic G-action.

As in the case for abstract groups, we can form something resembling an exact sequence. Let ρ be a rational G-action on A, and define

$$[L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho \subset [L\widetilde{H}]^\theta\mathrm{mo}(G,N), \quad [LH]^\theta\mathrm{mo}(G,N)_\rho \subset [LH]^\theta\mathrm{mo}(G,N)$$

as the subsets of weak morphs which induce the action ρ.

We get the following theorem:

Theorem 27 (cf. Theorem 5). *For a rational G-action ρ on A the following statements hold*:

(1) *There is a restriction map*

$$\mathrm{Res} : [L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho \to [LH]^\theta\mathrm{mo}(G,N)_\rho, \quad \mathrm{Res}(\langle f\rangle) = [f],$$

where $\langle f\rangle$ and $[f]$ are the equivalence classes in $[L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho$ and $[LH]^\theta\mathrm{mo}(G,N)_\rho$.

(2) *The abelian group $Z^1_{\mathrm{Rat}}(G,L;(A,\rho))$ acts freely on the set $[L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho$ by*

$$\gamma\cdot\langle f\rangle := \langle\dot\gamma f\rangle, \quad \textit{where } \dot\gamma f = m_K\circ(\dot\gamma\times f)\circ\Delta_G$$

and $\dot\gamma : G \xrightarrow{\gamma} A \to H$ comes from a rational Rosenlicht section $A \to H$ (cf. [Rosenlicht 1956, Theorem 10]) with $\dot\gamma(1) = 1$.

(3) *The corestricted restriction map* $\mathrm{Res} : [L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho \to \mathrm{Im}(\mathrm{Res})$ *is a quotient map by the $Z^1_{\mathrm{Rat}}(G,L;(A,\rho))$-action.*

(4) *If H, $\widetilde{H}$ and A are reduced, two classes $\langle f\rangle, \langle g\rangle \in [L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho$ lie in the same $B^1_{\mathrm{Rat}}(G,L;(A,\rho))$-orbit if and only if there exist $h\in H$, $f'\in\langle f\rangle$, $g'\in\langle g\rangle$ such that $[f(L),h]\subset\widetilde{H}$ and $f'(x) = hg'(x)h^{-1}$ for all $x\in G$.*

(5) *There is an obstruction map*

$$\mathrm{Obs} : [LH]^\theta\mathrm{mo}(G,N)_\rho \to H^2_{\mathrm{Rat}}(G,L;(A,\rho)), \quad \mathrm{Obs}([f]) = [f^\sharp],$$

where the cocycle $f^\sharp$ is defined by

$$G\times G \xrightarrow{(p_1,p_2,m_K)} G\times G\times G \xrightarrow{(f,f,\mathrm{inv}_K f)} K\times K\times K \xrightarrow{m_K} H \xrightarrow{\pi} A.$$

Here, p_1 and p_2 denote projection to the first and second coordinate respectively.

(6) *The sequence* (*cf. sequence* (1))

$$[L\widetilde{H}]^\theta\mathrm{mo}(G,N)_\rho \to [LH]^\theta\mathrm{mo}(G,N)_\rho \to H^2_{\mathrm{Rat}}(G,L;(A,\rho))$$

is exact, i.e., the image of Res *is equal to* $\mathrm{Obs}^{-1}([0])$.

Proof. If $\langle f \rangle = \langle g \rangle$ then the map

$$\alpha : G \xrightarrow{(f, \mathrm{inv}_K\, g)} K \times K \xrightarrow{m} K$$

has image in $\widetilde{H}$ and is trivial on L. It is rational as it is a composition of rational maps, and the identity is in the domain of definition and image of each map.

We also observe that given an analogous $\alpha : G \to H$ (i.e., corresponding to $[f] = [g]$) we get $\pi\alpha : G \to A$. Denoting the Rosenlicht section [1956, Theorem 10] $A \to H$ by τ, we see that $\tau\pi\alpha = \alpha$ and thus $(\dot{\pi\alpha}) = \alpha$. Note that we may assume the Rosenlicht section is defined at 0_A by composing with a translation if necessary. All the maps here are rational. In particular, $\pi\alpha \in C^1_{\mathrm{Rat}}(G, L; (A, \rho))$.

With these observations in mind, the remainder of the proof follows in the same way as in the proof of Theorem 5 does for abstract groups, doing everything diagrammatically. □

Before going any further, let's consider the following case where we can use this exact sequence directly. A restricted $\mathfrak{g}$-module (V, θ) satisfying the condition that $\mathrm{Aut}_{\mathfrak{g}}(V) = \mathbb{K}^\times$ is called a *brick*. A brick is necessarily an indecomposable $\mathfrak{g}$-module.

Theorem 28. *Suppose G is a semisimple, simply-connected algebraic group over an algebraically closed field $\mathbb{K}$ of characteristic $p > 0$, with Lie algebra $\mathfrak{g}$. Suppose further that, if $p = 2$, $\mathfrak{g}$ does not contain A_1, B_2 or C_l $(l \geq 3)$ as a direct summand. Let (V, θ) be a finite-dimensional G-stable brick. Then there exists a unique G-module structure Θ on V with $\Theta|_{G_1} = \theta$.*

Proof. We use Theorem 27 in the following situation:

- $L = G_1$, the first Frobenius kernel of G,
- $K = \mathrm{GL}(V)$,
- $H = \mathrm{Aut}_{\mathfrak{g}}(V) = \mathbb{K}^\times$,
- $N = N_K(H)$,
- $X = \mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, \mathfrak{gl}(V))$, a separated algebraic scheme with $\theta \in X(\mathbb{K})$.

Observe that G acts on X on the right via the adjoint map on the domain and $\mathrm{GL}(V)$ acts on X on the left via conjugation on the image. Furthermore, the actions commute, and the G-stability of V gives us that $\theta^G \subset {}^{\mathrm{GL(V)}}\theta$.

Hence, Lemma 26 gives a $(1, H)$-morph of algebraic groups, say $f : G \to \mathrm{GL}(V)$. In particular, it gives a homomorphism of algebraic groups $f : G \to \mathrm{PGL}(V)$, together with a rational lifting $\eta : \mathrm{PGL}(V) \to \mathrm{GL}(V)$. This rational lifting can be defined as follows: fix a basis of V and let U be the open subset of $\mathrm{PGL}(V)$ consisting of all cosets which can be represented by a (unique) matrix $A = (a_{ij}) \in \mathrm{GL}(V)$

with $a_{11} = 1$. Then define the map $\eta : U \to \mathrm{GL}(V)$ by assigning to each coset this representative.

Currently f and θ give the same maps from G_1 to N/H, since

$$^{\theta(x)}\theta(a)(v) = \theta(x)\theta(a)\theta(x^{-1})(v) = \theta(xax^{-1})(v) = \theta^x(a)(v)$$

for $x, a \in G_1(\mathbb{S})$, $v \in V(\mathbb{S})$ for any commutative $\mathbb{K}$-algebra $\mathbb{S}$. Note, however, that the maps $G_1 \to K$ do not necessarily agree.

To fix this potential disagreement, we define a rational map $R : G_1 \to H = \mathbb{K}^\times$ by $R(g) = f(g)^{-1}\theta(g)$ for $g \in G_1(\mathbb{S})$. There exists a rational map $\widetilde{R} : G \to H = \mathbb{K}^\times$ which restricts to R on G_1. Indeed, we have $R \in \mathbb{K}[G_1]$ (as G_1 is infinitesimal), so we can lift it to $\widetilde{R} \in \mathbb{K}[G]$ (since $\mathbb{K}[L]$ is a quotient of $\mathbb{K}[G]$). Let $U = G \setminus \widetilde{f}^{-1}(0)$. This is open in G, and on U we have that the image of $\widetilde{R}$ lies inside $\mathbb{K}^\times$, so $\widetilde{R}$ is a rational map $G \to \mathbb{K}^\times$. If now we define $\widetilde{f} : G \to \mathrm{GL}(V)$ by $\widetilde{f}(g) = f(g)\widetilde{R}(g)$, we get that $\widetilde{f}$ is a (G_1, H)-morph which restricts to θ on G_1, fixing the disagreement.

Observe that with $\widetilde{H} := 1$, we get (in the notation of the Theorem 27) $A = H$ and G acting on A trivially. Hence, the "exact sequence" from Theorem 27 is

$$H^1_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times) \dashrightarrow [G_1 1]^\theta \mathrm{mo}(G, N)_1 \to [G_1 H]^\theta \mathrm{mo}(G, N)_1 \to H^2_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times).$$

By Lemma 25, $H^2_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times) = 0$. Hence $[\widetilde{f}] \in [G_1 H]^\theta \mathrm{mo}(G, N)_1$ can be lifted to $\widehat{f} \in [G_1 1]^\theta \mathrm{mo}(G, N)_1$. This means that $\Theta := \widehat{f} : G \to \mathrm{GL}(V)$ is a homomorphism of algebraic groups which restricts to θ on G_1. Furthermore, this representation is unique (up to equivalence) if $H^1_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times) = 0$.

By Lemma 22,

$$H^1_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times) = \{\mu \in \mathrm{Hom}(G; \mathbb{K}^\times) \mid \mu|_{G_1} \equiv 1\}.$$

Since G is perfect, $H^1_{\mathrm{Rat}}(G, G_1; \mathbb{K}^\times) = 0$ and the extension is unique. □

2.4. ***G-stable modules with soluble automorphisms.*** We return to the general situation, where (G, L), (K, H) are algebraic (group, subgroup scheme) pairs with H soluble, G connected, and H reduced. However, from now on we suppose that L is a normal subgroup scheme of G. We also fix a homomorphism of algebraic groups $\theta : L \to K$, where the image commutes with H, so we are now dealing with (L, H)-morphs. Everything in the previous section can be reformulated in terms of (L, H)-morphs without difficulty — the key difference is that the G-action on A is now trivial on L. Since H is soluble, we can find a subnormal series $H = H_0 \rhd H_1 \rhd \ldots \rhd H_k = \{1\}$ with commutative quotients $A_j = H_{j-1}/H_j$ and each H_j characteristic in $N = N_K(H)$ and reduced.

Suppose that f is an (L, H)-morph of algebraic groups such that $f|_L = \theta$. As in the case of abstract groups, we get the following theorem — it generalises the procedure which we have used for bricks in the previous section.

Theorem 29 (cf. Theorem 7). *Given an (L, H)-morph of algebraic groups $f = f_0$ with $f|_L = \theta$, we obtain any $(L, 1)$-morph extending θ by applying the following procedure. Step m is the following*:

(1) *The (L, H_{m-1})-morph $f_{m-1} : G \to N$ such that $f_{m-1}|_L = \theta$ determines a rational G-action ρ_m on A_m.*

(2) *If* $\operatorname{Obs}([f_{m-1}]) \neq 0 \in H^2_{\mathrm{Rat}}(G, L; (A_m, \rho_m))$, *then this branch of the process terminates.*

(3) *If* $\operatorname{Obs}([f_{m-1}]) = 0 \in H^2_{\mathrm{Rat}}(G, L; (A_m, \rho_m))$, *then we choose an (L, H_m)-morph $f_m : G \to N$ such that* $\operatorname{Res}([f_m]) = [f_{m-1}]$.

(4) *For each element of $H^1_{\mathrm{Rat}}(G, L; (A_m, \rho_m))$ we choose a different f_m branching the process.* (*The choices different by an element of $B^1_{\mathrm{Rat}}(G, L; (A_m, \rho_m))$ are conjugate by an element of H.*)

(5) *We change m to $m + 1$ and go to step* (1).

An $(L, 1)$-morph which restricts to θ on L is equivalent to f_k for one of the nonterminated branches. Two $(L, 1)$-morphs f, g come from different branches if and only if there is no $h \in H$ such that $f(x) = hg(x)h^{-1}$ for all $x \in G$.

We get the following corollaries, similarly to Section 1.3:

Corollary 30. *Suppose $H^2_{\mathrm{Rat}}(G, L; (A_m, \rho_m)) = 0$ for all m for one of the branches. Then this branch does not terminate and there is a homomorphism $f : G \to K$ which restricts to θ on L.*

Corollary 31. *Suppose $H^1_{\mathrm{Rat}}(G, L; (A_m, \rho_m)) = 0$ for all m for one of the nonterminating branches. Then this branch is the only branch. Moreover*, *if a homomorphism of algebraic groups $f : G \to K$ restricting to θ exists*, *then it is unique up to conjugation by an element of H.*

Corollary 32. *Suppose $H^1_{\mathrm{Rat}}(G, L; (A_k, \rho_k)) \neq 0$ for one of the nonterminating branches. Then there exist algebraic homomorphisms $G \to K$ which are not conjugate by an element of H.*

We apply this theorem (and these corollaries) in the following case — a generalisation of the case from the previous section:

- G is a connected algebraic group over $\mathbb{K}$ with Lie algebra $\mathfrak{g}$,
- $L = G_1$,
- $K = \mathrm{GL}(V)$, where (V, θ) is a finite-dimensional G-stable indecomposable $\mathfrak{g}$-module,
- $H = \mathrm{Aut}_{\mathfrak{g}}(V)$,
- $X = \mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, \mathfrak{gl}(V))$, a separated algebraic scheme with $\theta \in X(\mathbb{K})$.

Applying exactly the same argument as in Theorem 28, we only start to encounter problems when trying to extend the rational map $R : G_1 \to H$ to a rational map on the whole of G. This can be fixed without much difficulty.

As a variety, we have that

$$H = \mathbb{K}^\times \times \mathbb{K}^n \subset \mathbb{K}^{n+1}$$

for some n (Proposition 17). Hence, we get $R = (R_0, R_1, \ldots, R_n)$, where $R_i \in \mathbb{K}[G_1]$ for $i = 0, 1, \ldots, n$. We can then lift each of these to elements of $\mathbb{K}[G]$, so we obtain $\widetilde{R} = (\widetilde{R_0}, \widetilde{R_1} \ldots, \widetilde{R_n}) : G \to \mathbb{K}^{n+1}$. We would like the image to lie in H. Thus, we define $U = G \setminus R_0^{-1}(0)$. This is an open set in G, so we can view $\widetilde{R}$ as a rational map from G to $\mathbb{K}^\times \times \mathbb{K}^n = H$ which is defined on U, and restricts to R on G_1.

Now we can define $\widetilde{f} : G \to \mathrm{GL}(V)$ as $\widetilde{f}(g) = f(g)\widetilde{R}(g)$. This is a (G_1, H)-morph of algebraic groups, which restricts to θ on G_1. Hence, we are in the situation of Theorem 29. Observe that $\theta : G_1 \to \mathrm{GL}(V)$ extends to a homomorphism of algebraic groups $\Theta : G \to \mathrm{GL}(V)$ if and only if there exists a $(G_1, 1)$-morph of algebraic groups extending θ. In particular, the corollaries to Theorem 29 can be used to determine the existence and uniqueness of a G-module structure on V.

Corollary 33 (existence test). *Suppose G is a connected algebraic group over $\mathbb{K}$ with Lie algebra $\mathfrak{g}$, and suppose V is an indecomposable G-stable finite-dimensional $\mathfrak{g}$-module. Then there exists a G-action on V which respects the $\mathfrak{g}$-module structure if and only if there is a branch* (*in the terminology of Theorem 29*) *which does not terminate*; *for instance, a branch such that $H^2_{\mathrm{Rat}}(G, G_1; (A_m, \rho_m)) = 0$ for all (A_m, ρ_m) on that branch.*

Corollary 34 (uniqueness test). *Suppose that G is a connected algebraic group over $\mathbb{K}$ with Lie algebra $\mathfrak{g}$, and that V is an indecomposable G-stable finite-dimensional $\mathfrak{g}$-module. Suppose further that there exists a G-action on V which extends the $\mathfrak{g}$-module structure. This G-action is unique* (*up to isomorphism*) *if and only if there is a branch* (*in the terminology of Theorem 29*) *such that $H^1_{\mathrm{Rat}}(G, G_1; (A_m, \rho_m)) = 0$ for all (A_m, ρ_m) on that branch.*

Observe that combining Corollary 34 with Corollary 23 for the N-stable subnormal series $H_m = 1 + J^m$, $m \geq 1$, we get a similar result to Proposition 4.3.1 in [Xanthopoulos 1992].

2.5. ***Comparison with $C^\bullet_{\mathrm{Rat}}(G/L; A)$.*** Let us now mimic the approach we took in Section 1.5 and examine how our cochain complex $(C^\bullet_{\mathrm{Rat}}(G, L; A), d)$ compares with the complex $(C^\bullet_{\mathrm{Rat}}(G/L; A), d)$ on the level of cohomology. We use the notation of Section 2.3. As with our discussion in Section 1.5 we have to assume that L acts trivially on A for this discussion to be meaningful — a condition which holds in the examples considered.

Similar to the case for abstract groups, we have the following proposition:

Proposition 35. *Under the aforementioned conditions we have isomorphisms of groups* $H^0_{\mathrm{Alg}}(G, L; A) \cong H^0_{\mathrm{Alg}}(G/L; A)$ *and* $H^1_{\mathrm{Alg}}(G, L; A) \cong H^1_{\mathrm{Alg}}(G/L; A)$.

Proof. Making use of the universal property of the quotient for algebraic groups, the proof follows word-for-word as in Proposition 11. □

Recalling the observation that there is no distinction between H^i_{Alg} and H^i_{Rat} for $i = 0, 1$, this tells us that $H^0_{\mathrm{Rat}}(G, L; A) \cong H^0_{\mathrm{Alg}}(G/L; A)$ and $H^1_{\mathrm{Rat}}(G, L; A) \cong H^1_{\mathrm{Alg}}(G/L; A)$ in these circumstances.

The universal property of the quotient for algebraic groups further yields an analogue of Proposition 12.

Proposition 36. *The maps* $\mathrm{Inf}_{\mathrm{Alg}} : H^2_{\mathrm{Alg}}(G/L; A) \to H^2_{\mathrm{Alg}}(G, L; A)$ *and* $\mathrm{Inf}_{\mathrm{Rat}} : H^2_{\mathrm{Rat}}(G/L; A) \to H^2_{\mathrm{Rat}}(G, L; A)$ *are injective.*

Proof. The proof follows as in Proposition 12. □

In the case of abstract groups, Section 1.5 shows that by making careful choices of (L, H)-morphs in Theorem 7 the image of the obstruction maps

$$\mathrm{Obs} : [LH]^{\theta}\mathrm{mo}(G, N)_{\rho_i} \to H^2(G, L; (A_i, \rho_i))$$

always lies inside $H^2(G/L; (A_i, \rho_i)) \hookrightarrow H^2(G, L; (A_i, \rho_i))$. As such, it is possible to reinterpret Theorem 7 using the complex $(C^\bullet(G/L; A), d)$ instead of $(C^\bullet(G, L; A), d)$ at all points. This conclusion for abstract groups, however, relies on the observation that it is always possible to assume that the (L, H)-morphs being considered are normalised. When translating the results to the case of algebraic groups it is far from clear that the analogues of Lemma 13 and Corollary 14 hold.

Question. Can the (L, H)-morphs considered in Sections 2.3 and 2.4 be chosen to be normalised?

Acknowledgements

We would like to thank Ami Braun, Simon Goodwin and Jim Humphreys for valuable discussions. We are indebted to Stephen Donkin for encouragement, interest in our work and sharing Xanthopoulos' thesis.

References

[Capdeboscq et al. 2017] I. Capdeboscq, K. Kirkina, and D. Rumynin, "Subgroup growth in some profinite Chevalley groups", *Proc. Amer. Math. Soc.* **145**:10 (2017), 4187–4200. MR Zbl

[Cline 1972] E. Cline, "Stable Clifford theory", *J. Algebra* **22** (1972), 350–364. MR Zbl

[Dade 1981] E. C. Dade, "Extending irreducible modules", *J. Algebra* **72**:2 (1981), 374–403. MR Zbl

[Demazure and Gabriel 1970] M. Demazure and P. Gabriel, *Groupes algébriques, I: Géométrie algébrique, généralités, groupes commutatifs*, Masson, Paris, 1970. MR Zbl

[Donkin 1980] S. Donkin, "On a question of Verma", *J. London Math. Soc.* (2) **21**:3 (1980), 445–455. MR Zbl

[Donkin 1982] S. Donkin, "A note on the characters of the projective modules for the infinitesimal subgroups of a semisimple algebraic group", *Math. Scand.* **51**:1 (1982), 142–150. MR Zbl

[Hogeweij 1982] G. M. D. Hogeweij, "Almost-classical Lie algebras, I", *Indag. Math.* **44**:4 (1982), 441–452. MR Zbl

[Jantzen 1987] J. C. Jantzen, *Representations of algebraic groups*, Pure and Applied Mathematics **131**, Academic Press, Boston, 1987. MR Zbl

[van der Kallen 1973] W. L. J. van der Kallen, *Infinitesimally central extensions of Chevalley groups*, Lecture Notes in Mathematics **356**, Springer, 1973. MR Zbl

[Karpilovsky 1989] G. Karpilovsky, *Clifford theory for group representations*, North-Holland Mathematics Studies **156**, North-Holland, Amsterdam, 1989. MR Zbl

[Parshall and Scott 2013] B. J. Parshall and L. L. Scott, "Variations on a theme of Cline and Donkin", *Algebr. Represent. Theory* **16**:3 (2013), 793–817. MR Zbl

[Rosenlicht 1956] M. Rosenlicht, "Some basic theorems on algebraic groups", *Amer. J. Math.* **78** (1956), 401–443. MR Zbl

[Steinberg 1963] R. Steinberg, "Representations of algebraic groups", *Nagoya Math. J.* **22** (1963), 33–56. MR Zbl

[Thévenaz 1983] J. Thévenaz, "Extensions of group representations from a normal subgroup", *Comm. Algebra* **11**:4 (1983), 391–425. MR Zbl

[Weil 1955] A. Weil, "On algebraic groups of transformations", *Amer. J. Math.* **77** (1955), 355–391. MR Zbl

[Xanthopoulos 1992] S. Xanthopoulos, *On a question of Verma about indecomposable representations of algebraic groups and of their Lie algebras*, Ph.D. thesis, Queen Mary University of London, 1992.

Received May 7, 2018. Revised December 7, 2018.

DMITRIY RUMYNIN
MATHEMATICS INSTITUTE
ZEEMAN BUILDING
UNIVERSITY OF WARWICK
COVENTRY
UNITED KINGDOM
d.rumynin@warwick.ac.uk

MATTHEW WESTAWAY
MATHEMATICS INSTITUTE
ZEEMAN BUILDING
UNIVERSITY OF WARWICK
COVENTRY
UNITED KINGDOM
m.p.westaway@warwick.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.601

UNIFORM BOUNDS OF THE PILTZ DIVISOR PROBLEM OVER NUMBER FIELDS

WATARU TAKEDA

We consider the upper bound of the Piltz divisor problem over number fields. The Piltz divisor problem is known as a generalization of the Dirichlet divisor problem. We deal with this problem over number fields and improve the error term of this function for many cases. Our proof uses the estimate of exponential sums. We also show uniform results for the ideal counting function and relatively r-prime lattice points as one of its applications.

1. Introduction

The behavior of arithmetic functions has long been studied, and it is one of the most important areas of research in analytic number theory. But many arithmetic functions $f(n)$ fluctuate as n increases, and it becomes difficult to deal with them. Thus, many authors study partial sums $\sum_{n\le x} f(n)$ to obtain some information about arithmetic functions $f(n)$. In this paper we consider the Piltz divisor function $I_K^m(x)$ over a number field. Let K be a number field with extension degree $[K:\mathbb{Q}]=n$, and let $\mathcal{O}_K$ be its ring of integers. Let D_K be the absolute value of the discriminant of K. Then the Piltz divisor function $I_K^m(x)$ counts the number of m-tuples of ideals $(\mathfrak{a}_1,\mathfrak{a}_2,\dots,\mathfrak{a}_m)$ such that the product of their ideal norms $\mathfrak{N}\mathfrak{a}_1\cdots\mathfrak{N}\mathfrak{a}_m\le x$. It is known that

$$I_K^m(x)\sim \operatorname*{Res}_{s=1}\left(\zeta_K(s)^m\frac{x^s}{s}\right). \tag{1-1}$$

We denote by $\Delta_K^m(x)$ the error term of $I_K^m(x)$, that is, $I_K^m(x)-\operatorname{Res}_{s=1}\bigl(\zeta_K(s)^m\frac{x^s}{s}\bigr)$.

In the case $m=1$, this function is the ordinary ideal counting function over K. For simplicity we substitute $I_K(x)$ and $\Delta_K(x)$ for $I_K^1(x)$ and $\Delta_K^1(x)$, respectively. There are many results about $I_K(x)$ from the 1900s. In the case $K=\mathbb{Q}$, integer ideals of $\mathbb{Z}$ and positive integers are in one-to-one correspondence, so $I_{\mathbb{Q}}(x)=[x]$, where $[\cdot]$ is the Gauss symbol. For the general case, the best estimate of $\Delta_K(x)$ hitherto is the following theorem.

MSC2010: primary 11N45; secondary 11R42, 11H06, 11P21.
Keywords: ideal counting function, exponential sum, Piltz divisor problem.

Theorem 1-2. *The following estimates hold for all $\varepsilon > 0$:*

$n = [K : \mathbb{Q}]$	$\Delta_K(x)$	
2	$O\big(x^{\frac{131}{416}}(\log x)^{\frac{18627}{8320}}\big)$	[Huxley 2002]
3	$O\big(x^{\frac{43}{96}+\varepsilon}\big)$	[Müller 1988]
4	$O\big(x^{\frac{41}{72}+\varepsilon}\big)$	[Bordellès 2015]
$5 \le n \le 10$	$O\big(x^{1-\frac{4}{2n+1}+\varepsilon}\big)$	[Bordellès 2015]
$11 \le n$	$O\big(x^{1-\frac{3}{n+6}+\varepsilon}\big)$	[Lao 2010]

There are also many results about $I_{\mathbb{Q}}^m$ from the 1800s. In 1849 Dirichlet showed that

$$I_{\mathbb{Q}}^2(x) = x\log x + (2\gamma - 1)x + O\big(x^{\frac{1}{2}}\big),$$

where γ is the Euler constant, defined by

$$\gamma = \lim_{n\to\infty}\Bigg(\sum_{k=1}^{n}\frac{1}{k} - \log n\Bigg).$$

The O-term has been improved by many researchers many times; the best estimate hitherto is $x^{\frac{517}{1648}+\varepsilon}$ [Bourgain and Watt 2017].

As we have mentioned above, there exist many results about other divisor problems, but it seems that there are not many results about the Piltz divisor problem over number fields.

Theorem 1-3 [Nowak 1993]. *When $n = [K : \mathbb{Q}] \ge 2$, then we get*

$$\Delta_K^m(x) = \begin{cases} O_K\big(x^{1-\frac{2}{mn}+\frac{8}{mn(5mn+2)}}(\log x)^{m-1-\frac{10(m-2)}{5n+2}}\big) & \textit{for } 3 \le mn \le 6,\\ O_K\big(x^{1-\frac{2}{mn}+\frac{3}{2m^2n^2}}(\log x)^{m-1-\frac{2(m-2)}{mn}}\big) & \textit{for } mn \ge 7.\end{cases}$$

For the estimate of the lower bound, Girstmair, Kühleitner, Müller, and Nowak obtain the following Ω-results:

Theorem 1-4 [Girstmair et al. 2005]. *For any fixed number field K with $n = [K : \mathbb{Q}] \ge 2$,*

$$\Delta_K^m(x) = \Omega\big(x^{\frac{1}{2}-\frac{1}{2mn}}(\log x)^{\frac{1}{2}-\frac{1}{2mn}}(\log\log x)^{\kappa}(\log\log\log x)^{-\lambda}\big), \tag{1-5}$$

where κ and λ are constants depending on K. To be more precise, let K^{gal} be the Galois closure of $K/\mathbb{Q}$, $G = Gal(K^{\mathrm{gal}}/\mathbb{Q})$ its Galois group, and $H = Gal(K^{\mathrm{gal}}/K)$ the subgroup of G corresponding to K. Then

$$\kappa = \frac{mn+1}{2mn}\Bigg(\sum_{\nu=1}^{n}\delta_\nu \nu^{\frac{2mn}{mn+1}} - 1\Bigg) \quad \textit{and} \quad \lambda = \frac{mn+1}{4mn}R + \frac{mn-1}{2mn},$$

where

$$\delta_\nu = \frac{\bigl|\bigl\{\tau \in G \bigm| |\{\sigma \in G \mid \tau \in \sigma H \sigma^{-1}\}| = \nu |H|\bigr\}\bigr|}{|G|}$$

and R is the number of $1 \le \nu \le n$ *with* $\delta_\nu > 0$.

We know the following conditional result. If we assume the Lindelöf hypothesis for the Dedekind zeta function, it holds that for all $\varepsilon > 0$, for all K, and for all m,

$$\Delta_K^m(x) = O_\varepsilon\bigl(x^{\frac{1}{2}+\varepsilon} D_K^{\varepsilon}\bigr). \tag{1-6}$$

In this paper we estimate the error term of $\Delta_K^m(x)$ by using exponential sums. In [Nowak 1993; Girstmair et al. 2005], they use other approaches, so we expect new development for the Piltz divisor problem over number fields. As a results, we improve the estimate of upper bound of $\Delta_K^m(x)$ for many K and many m.

In Section 2, we show some auxiliary theorems to consider the upper bound of the error term $\Delta_K^m(x)$. First we give a review of the convexity bound for the Dedekind zeta function and generalized Atkinson lemma [1941]. Next we show Proposition 2-6, which reduces an ideal counting problem to a problem of exponential sums. This proposition plays a crucial role in our computing $\Delta_K^m(x)$.

In Section 3, we prove the following theorem about the error term $\Delta_K^m(x)$ by using estimates of exponential sums.

Theorem 1-7. *For every $\varepsilon > 0$ the following estimate holds. When $mn \ge 4$, then*

$$\Delta_K^m(x) = O_{n,m,\varepsilon}\bigl(x^{\frac{2mn-3}{2mn+1}+\varepsilon} D_K^{\frac{2m}{2mn+1}+\varepsilon}\bigr).$$

This theorem gives improvement of the upper bound of $\Delta_K^m(x)$ for $mn \ge 4$.

In Section 4, we give some applications. First we give a uniform estimate for the ideal counting function over number fields. Second we show a good uniform upper bound of the distribution of relatively r-prime lattice points over number fields as a corollary of the first application.

In Section 5, we consider a conjecture about estimates for the Piltz divisor functions over number fields. It is proposed that for all number fields K and for all m the best upper bound of the error term is better than that on the assumption of the Lindelöf hypothesis (1-6).

2. Auxiliary theorem

In this section, we show some important lemmas for our argument. Let $s = \sigma + it$ and $n = [K : \mathbb{Q}]$. We use the convexity bound of the Dedekind zeta function to obtain an upper bound of the error term of the Piltz divisor function $\Delta_K^m(x)$.

It is well known that the Dedekind zeta function satisfies the functional equation

$$\zeta_K(1-s) = D_K^{s-\frac{1}{2}} 2^{n(1-s)} \pi^{-ns} \Gamma(s)^n \Big(\cos\frac{\pi s}{2}\Big)^{r_1+r_2} \Big(\sin\frac{\pi s}{2}\Big)^{r_2} \zeta_K(s), \tag{2-1}$$

where r_1 is the number of real embeddings of K and r_2 is the number of pairs of complex embeddings.

The Phragmen–Lindelöf principle and (2-1) give the well known convexity bound of the Dedekind zeta function [Rademacher 1959]: for any $\varepsilon > 0$ and $n = [K : \mathbb{Q}]$,

$$\zeta_K(\sigma+it) = \begin{cases} O_{n,\varepsilon}\big(|t|^{\frac{n}{2}-n\sigma+\varepsilon} D_K^{\frac{1}{2}-\sigma+\varepsilon}\big) & \text{if } \sigma \le 0, \\ O_{n,\varepsilon}\big(|t|^{\frac{n(1-\sigma)}{2}+\varepsilon} D_K^{\frac{1-\sigma}{2}+\varepsilon}\big) & \text{if } 0 \le \sigma \le 1, \\ O_{n,\varepsilon}(|t|^{\varepsilon} D_K^{\varepsilon}) & \text{if } 1 \le \sigma \end{cases} \tag{2-2}$$

as $|t|^n D_K \to \infty$, where K runs through number fields with $[K : \mathbb{Q}] = n$. In the previous papers, we also use this convexity bound (2-2) to estimate the distribution of ideals. In the following sections, we show some estimates for $\Delta_K^m(x)$ in the similar way to our previous papers.

Lemma 2-3 states the growth of the product of the Gamma function and trigonometric functions in the functional equation (2-1) of the Dedekind zeta function.

Lemma 2-3. *Let $\tau \in \{\cos, \sin\}$ and n be a positive integer. Then*

$$\frac{\Gamma(s)^n}{1-s}\Big(\cos\frac{\pi s}{2}\Big)^{r_1+r_2}\Big(\sin\frac{\pi s}{2}\Big)^{r_2} = Cn^{-ns}\Gamma\Big(ns - \frac{n+1}{2}\Big)\tau\Big(\frac{n\pi s}{2}\Big) + O_n\big(|t|^{-2+n\sigma-\frac{n}{2}}\big),$$

where C is a constant and $s = \sigma + it$.

Proof. This lemma is shown from the Stirling formula and estimate for trigonometric functions. □

Next we introduce the generalized Atkinson lemma. This lemma is quite useful for calculating integrals of the Dedekind zeta function.

Lemma 2-4 [Atkinson 1941]. *Let $y > 0$, $1 < A \le B$, and $\tau \in \{\cos, \sin\}$, and define*

$$I = \frac{1}{2\pi i}\int_{A-iB}^{A+iB} \Gamma(s)\tau\Big(\frac{\pi s}{2}\Big) y^{-s}\, ds.$$

If $y \le B$, then

$$I = \tau(y) + O\Big(y^{-\frac{1}{2}}\min\Big(\Big(\log\frac{B}{y}\Big)^{-1}, B^{\frac{1}{2}}\Big) + y^{-A}B^{A-\frac{1}{2}} + y^{-\frac{1}{2}}\Big).$$

If $y > B$, then

$$I = O\Big(y^{-A}\Big(B^{A-\frac{1}{2}}\min\Big(\Big(\log\frac{y}{B}\Big)^{-1}, B^{\frac{1}{2}}\Big) + A^{A-\frac{1}{2}}\Big)\Big).$$

Finally we introduce the following lemma to reduce the ideal counting problem to an exponential sum problem.

Lemma 2-5 [Bordellès 2015]. *Let $1 \le L \le R$ be a real number and f be an arithmetical function satisfying $f(m) = O(m^{\varepsilon})$, and let $\boldsymbol{e}(x) = \exp(2\pi i x)$ and $F = f * \mu$, where $*$ is the Dirichlet product symbol. For $a \in \mathbb{R} - \{1\}$ and $b, x \in \mathbb{R}$ and for every $\varepsilon > 0$ the following estimate holds*:

$$\sum_{m \le R} \frac{f(m)}{m^a} \tau(2\pi x m^b) = O_{n,\varepsilon}\biggl(L^{1-a} + R^{\varepsilon} \max_{L < S \le R} S^{-a} \times \max_{S < S_1 \le 2S} \max_{\substack{M,N \le S_1 \\ MN \asymp S}} \max_{\substack{M \le M_1 \le 2M \\ N \le N_1 \le 2N}} \biggl| \sum_{M < m \le M_1} F(m) \sum_{N < n \le N_1} \boldsymbol{e}(x(mn)^b) \biggr| \biggr).$$

The next proposition plays a crucial role in our computing $I_K^m(x)$. We consider the distribution of ideals of $\mathcal{O}_K$, where K runs through extensions with $[K : \mathbb{Q}] = n$ and some conditions. The detail of the conditions will be determined later, but they state the relation of the principal term and the error term.

Proposition 2-6. *Let $F_K = I_K^m * \mu$. For every $\varepsilon > 0$ the following estimate holds*:

$$\Delta_K^m(x) = O_{n,m,\varepsilon}\biggl(L^{1-\alpha} + x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}} R^{\varepsilon} \max_{L \le S \le R} S^{-\frac{mn+1}{2mn}} \times \max_{S < S_1 \le 2S} \max_{\substack{M,N \le S_1 \\ MN \asymp S}} \max_{\substack{M \le M_1 \le 2M \\ N \le N_1 \le 2N}} \biggl| \sum_{M < l \le M_1} F_K(m) \sum_{N < k \le N_1} \boldsymbol{e}\biggl(mn \Bigl(\frac{xlk}{D_K} \Bigr)^{\frac{1}{mn}} \biggr) \biggr| + x^{\frac{mn-2}{2mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{\frac{mn-2}{2mn}+\varepsilon} + x^{\frac{mn-1}{mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{-\frac{1}{mn}+\varepsilon} \biggr),$$

where K runs through number fields with $[K : \mathbb{Q}] = n$ and some conditions.

Proof. Let $d_K^m(l)$ be the number of m-tuples of ideals $(\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_m)$ such that the product of their ideal norms $\mathfrak{N}\mathfrak{a}_1 \cdots \mathfrak{N}\mathfrak{a}_m = l$. Then one can easily check that

$$\zeta_K(s)^m = \sum_{l=1}^{\infty} \frac{d_K^m(l)}{l^s} \quad \text{for } \Re s > 1 \tag{2-7}$$

and

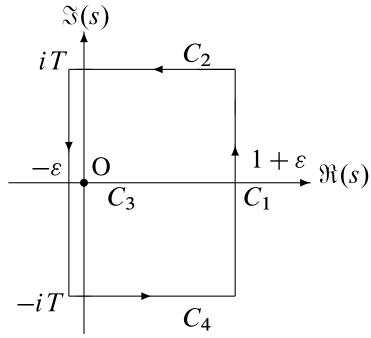
$$I_K^m(x) = \sum_{l \le x} d_K^m(l).$$

Thus, Perron's formula plays a crucial role in this proof.

We consider the integral

$$\frac{1}{2\pi i} \int_C \zeta_K(s)^m \frac{x^s}{s}\, ds,$$

where C is the contour $C_1 \cup C_2 \cup C_3 \cup C_4$ shown below:



In a way similar to the well known proof of Perron's formula, we estimate

$$\frac{1}{2\pi i}\int_{C_1}\zeta_K(s)^m\frac{x^s}{s}\,ds = I_K^m(x)+O_\varepsilon\biggl(\frac{x^{1+\varepsilon}}{T}\biggr). \tag{2-8}$$

We can select the large T, so that the O-term in the right-hand side is sufficiently small. For estimating the left-hand side by using estimate (2-2), we divide it into the integrals over C_2, C_3, and C_4.

First we consider the integrals over C_2 and C_4 as

$$\begin{aligned}&\left|\frac{1}{2\pi i}\int_{C_2\cup C_4}\zeta_K(s)^m\frac{x^s}{s}\,ds\right|\\ &\qquad\le\frac{1}{2\pi}\int_{-\varepsilon}^{1+\varepsilon}|\zeta_K(\sigma+iT)|^m\frac{x^\sigma}{T}\,d\sigma+\frac{1}{2\pi}\int_{-\varepsilon}^{1+\varepsilon}|\zeta_K(\sigma-iT)|^m\frac{x^\sigma}{T}\,d\sigma.\end{aligned}$$

It holds by the convexity bound of the Dedekind zeta function (2-2) that their sum is estimated as

$$\begin{aligned}\left|\frac{1}{2\pi i}\int_{C_2\cup C_4}\zeta_K(s)^m\frac{x^s}{s}\,ds\right| &= O_{n,m,\varepsilon}\biggl(\int_{-\varepsilon}^{1+\varepsilon}(T^{mn}D_K^m)^{\frac{1-\sigma}{2}+\varepsilon}\frac{x^\sigma}{T}\,d\sigma\biggr)\\ &= O_{n,m,\varepsilon}\biggl(\frac{x^{1+\varepsilon}D_K^\varepsilon}{T^{1-\varepsilon}}+T^{\frac{mn}{2}-1+\varepsilon}D_K^{\frac{m}{2}+\varepsilon}x^{-\varepsilon}\biggr).\end{aligned} \tag{2-9}$$

By the Cauchy residue theorem, (2-8), and (2-9) we obtain

$$\Delta_K^m(x)=\int_{C_3}\zeta_K(s)^m\frac{x^s}{s}\,ds+O_{n,m,\varepsilon}\biggl(\frac{x^{1+\varepsilon}D_K^\varepsilon}{T^{1-\varepsilon}}+T^{\frac{mn}{2}-1+\varepsilon}D_K^{\frac{m}{2}+\varepsilon}x^{-\varepsilon}\biggr). \tag{2-10}$$

Thus, it suffices to consider the integral over C_3 as

$$\frac{1}{2\pi i}\int_{C_3}\zeta_K(s)^m\frac{x^s}{s}\,ds=\frac{1}{2\pi i}\int_{-\varepsilon-iT}^{-\varepsilon+iT}\zeta_K(s)^m\frac{x^s}{s}\,ds.$$

Changing the variable s to $1-s$, we have

$$\frac{1}{2\pi i}\int_{C_3}\zeta_K(s)^m\frac{x^s}{s}\,ds=\frac{1}{2\pi i}\int_{1+\varepsilon-iT}^{1+\varepsilon+iT}\zeta_K(1-s)^m\frac{x^{1-s}}{1-s}\,ds.$$

From functional equation (2-1), it holds that

$$\frac{1}{2\pi i}\int_{C_3}\zeta_K(s)^m\frac{x^s}{s}\,ds=\frac{1}{2\pi i}\int_{1+\varepsilon-iT}^{1+\varepsilon+iT}\Bigl(D_K^{s-\frac12}2^{n(1-s)}\pi^{-ns}\Gamma(s)^n\Bigl(\cos\frac{\pi s}{2}\Bigr)^{r_1+r_2}\times\Bigl(\sin\frac{\pi s}{2}\Bigr)^{r_2}\zeta_K(s)\Bigr)^m\frac{x^{1-s}}{1-s}\,ds.$$

By Lemma 2-3 the integral over C_3 can be expressed as

$$\begin{aligned}&\frac{1}{2\pi i}\int_{C_3}\zeta_K(s)\frac{x^s}{s}\,ds\\&=\frac{Cx}{2\pi i}\int_{1+\varepsilon-iT}^{1+\varepsilon+iT}D_K^{-\frac{m}{2}}\Bigl(\frac{(2n)^{mn}\pi^{mn}x}{D_K^m}\Bigr)^{-s}\Gamma\Bigl(mns-\frac{mn+1}{2}\Bigr)\tau\Bigl(\frac{mn\pi s}{2}\Bigr)\zeta_K(s)\,ds\\&\qquad+O_{n,m,\varepsilon}\bigl(D_K^{\frac{m}{2}+\varepsilon}T^{\frac{mn}{2}-1+\varepsilon}x^{-\varepsilon}\bigr).\end{aligned}$$

Changing the variable $mns-\frac{mn+1}{2}$ to s, we have

$$\begin{aligned}\frac{1}{2\pi i}\int_{C_3}\zeta_K(s)\frac{x^s}{s}\,ds=\frac{Cx^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}}{2\pi i}\int_{\frac{mn-1}{2}+mn\varepsilon-mniT}^{\frac{mn-1}{2}+mn\varepsilon+mniT}\Bigl(2mn\pi\Bigl(\frac{x}{D_K^m}\Bigr)^{\frac{1}{mn}}\Bigr)^{-s}\\\times\Gamma(s)\tau\Bigl(\frac{\pi s}{2}+\frac{(mn+1)\pi}{4}\Bigr)\zeta_K\Bigl(\frac{s}{mn}+\frac{mn+1}{2mn}\Bigr)\,ds\\+O_{n,m,\varepsilon}\bigl(D_K^{\frac{m}{2}+\varepsilon}T^{\frac{mn}{2}-1+\varepsilon}x^{-\varepsilon}\bigr).\end{aligned}$$

From (2-7) the function $\zeta_K(s)^m$ can be expressed as a Dirichlet series. It is absolutely and uniformly convergent on compact subsets on $\Re(s)>1$. Therefore, we can interchange the order of summation and integral. Thus, we obtain

$$\begin{aligned}&\int\Bigl(2mn\pi\Bigl(\frac{x}{D_K^m}\Bigr)^{\frac{1}{mn}}\Bigr)^{-s}\Gamma(s)\tau\Bigl(\frac{\pi s}{2}+\frac{(mn+1)\pi}{4}\Bigr)\zeta_K\Bigl(\frac{s}{mn}+\frac{mn+1}{2mn}\Bigr)\,ds\\&\qquad=\sum_{l=1}^{\infty}\frac{d_K^m(l)}{l^{\frac{mn+1}{2mn}}}\int\Bigl(2mn\pi\Bigl(\frac{lx}{D_K^m}\Bigr)^{\frac{1}{mn}}\Bigr)^{-s}\Gamma(s)\tau\Bigl(\frac{\pi s}{2}+\frac{(mn+1)\pi}{4}\Bigr)\,ds,\end{aligned}$$

where the integration is on the vertical line from $\frac{mn-1}{2}+mn\varepsilon-mniT$ to $\frac{mn-1}{2}+mn\varepsilon+mniT$. Properties of trigonometric functions lead to

$$\tau\Bigl(\frac{\pi s}{2}+\frac{(mn+1)\pi}{4}\Bigr)=\pm\begin{cases}\tau\bigl(\frac{\pi s}{2}\bigr) & \text{if } mn \text{ is odd,}\\ \frac{1}{\sqrt2}\bigl(\tau\bigl(\frac{\pi s}{2}\bigr)\pm\tau_1\bigl(\frac{\pi s}{2}\bigr)\bigr) & \text{if } mn \text{ is even,}\end{cases}$$

where $\{\tau, \tau_1\} = \{\sin, \cos\}$. Hence, it holds that

$$\frac{1}{2\pi i}\int_{C_3} \zeta_K(s)^m \frac{x^s}{s}\,ds = \frac{C x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}}}{2\pi i}\sum_{l=1}^{\infty}\frac{d_K^m(l)}{l^{\frac{mn+1}{2mn}}} \times \int_{\frac{mn-1}{2}+mn\varepsilon-mniT}^{\frac{mn-1}{2}+mn\varepsilon+mniT}\left(2mn\pi\left(\frac{lx}{D_K^m}\right)^{\frac{1}{mn}}\right)^{-s}\Gamma(s)\tau\left(\frac{\pi s}{2}\right)ds + O_{n,m,\varepsilon}\big(D_K^{\frac{m}{2}+\varepsilon}T^{\frac{mn}{2}-1+\varepsilon}x^{-\varepsilon}\big).$$

Applying Lemma 2-4 to this integral with $y = 2mn\pi\big(\frac{lx}{D_K^m}\big)^{\frac{1}{mn}}$, $A = \frac{mn-1}{2} + mn\varepsilon$, $B = mnT$, and $T = 2\pi\big(\frac{xR}{D_K^m}\big)^{\frac{1}{mn}}$, this becomes

$$\begin{aligned}\frac{1}{2\pi i}\int_{C_3} \zeta_K(s)^m \frac{x^s}{s}\,ds &= \frac{C x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}}}{2\pi i}\sum_{l\le R}\frac{d_K^m(l)}{l^{\frac{mn+1}{2mn}}}\tau\left(2mn\pi\left(\frac{lx}{D_K^m}\right)^{\frac{1}{mn}}\right)\\
&+ O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{l\le R}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\min\left\{\left(\log\frac{R}{l}\right)^{-1},\left(\frac{Rx}{D_K^m}\right)^{\frac{1}{2mn}}\right\}\right)\\
&+ O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{l\le R}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\left(\left(\frac{R}{l}\right)^{\frac{mn-2}{2mn}}+1\right)\right)\\
&+ O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}R^{\frac{mn-2}{2mn}+\varepsilon}\sum_{l>R}\frac{d_K^m(l)}{l^{1+\varepsilon}}\min\left\{\left(\log\frac{l}{R}\right)^{-1},\left(\frac{Rx}{D_K^m}\right)^{\frac{1}{2mn}}\right\}\right)\\
&+ O_{n,m,\varepsilon}\big(x^{\frac{mn-2}{2mn}+\varepsilon}D_K^{\frac{1}{n}+\varepsilon}R^{\frac{mn-2}{2mn}+\varepsilon}\big).\end{aligned}$$

We evaluate three O-terms as follows. First we consider the first O-term. One can estimate $\big(\log\frac{R}{l}\big)^{-1} = O\big(\frac{R}{R-l}\big)$, so we obtain

$$\begin{aligned}&O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{l\le R}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\min\left\{\left(\log\frac{R}{l}\right)^{-1},\left(\frac{Rx}{D_K^m}\right)^{\frac{1}{2mn}}\right\}\right)\\
&= O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{l\le[R]-1}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\left(\log\frac{R}{l}\right)^{-1} + x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{[R]\le l\le R}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\left(\frac{Rx}{D_K^m}\right)^{\frac{1}{2mn}}\right)\\
&= O_{n,m,\varepsilon}\left(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}\sum_{l\le[R]-1}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\frac{R}{R-l} + x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}R^{\frac{1}{2mn}}\sum_{[R]\le l\le R}\frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\right)\\
&= O_{n,m,\varepsilon}\big(x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}R^{\frac{mn-2}{2mn}+\varepsilon} + x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}R^{-\frac{mn+1}{2mn}}\big).\end{aligned}$$

Next we calculate the second O-term

$$O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} \sum_{l\le R} \frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\biggl(\Bigl(\frac{R}{l}\Bigr)^{\frac{mn-2}{2mn}}+1\biggr)\biggr) = O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}} \sum_{l\le R} \frac{d_K^m(l)}{l}\biggr).$$

Since it is well known that $d_K^m(l) = O(l^{\varepsilon})$, we get

$$\begin{aligned} O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} \sum_{l\le R} \frac{d_K^m(l)}{l^{\frac{mn+2}{2mn}}}\biggl(\Bigl(\frac{R}{l}\Bigr)^{\frac{mn-2}{2mn}}+1\biggr)\biggr) &= O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}} \int_1^R \frac{t^{\varepsilon}}{t}\,dt\biggr) \\ &= O_{n,m,\varepsilon}\bigl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon}\bigr). \end{aligned}$$

Finally we estimate the third O-term in a similar way to calculate the first O-term. One can estimate $\bigl(\log \frac{l}{R}\bigr)^{-1} = O\bigl(\frac{R}{l-R}\bigr)$, so we obtain

$$\begin{aligned} &O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon} \sum_{l>R} \frac{d_K^m(l)}{l^{1+\varepsilon}} \min\biggl\{\Bigl(\log\frac{l}{R}\Bigr)^{-1}, \Bigl(\frac{Rx}{D_K^m}\Bigr)^{\frac{1}{2mn}}\biggr\}\biggr) \\ &= O_{n,m,\varepsilon}\biggl(x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon}\biggl(\sum_{R<l\le[R]+1} \frac{d_K^m(l)}{l^{1+\varepsilon}}\Bigl(\frac{Rx}{D_K^m}\Bigr)^{\frac{1}{2mn}} + \sum_{[R]+2\le l} \frac{d_K^m(l)}{l^{1+\varepsilon}}\Bigl(\log\frac{l}{R}\Bigr)^{-1}\biggr)\biggr) \\ &= O_{n,m,\varepsilon}\biggl(x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}} R^{\frac{mn-1}{2mn}+\varepsilon} \sum_{R<l\le[R]+1} \frac{d_K^m(l)}{l^{1+\varepsilon}} + x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon} \sum_{[R]+2\le l} \frac{d_K^m(l)}{l^{1+\varepsilon}} \frac{R}{l-R}\biggr) \\ &= O_{n,m,\varepsilon}\bigl(x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}} R^{-\frac{mn+1}{2mn}+\varepsilon} + x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon}\bigr). \end{aligned}$$

From above results, we obtain

$$\frac{1}{2\pi i}\int_{C_3} \zeta_K(s)^m \frac{x^s}{s}\,ds = \frac{Cx^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}}}{2\pi i} \sum_{l\le R} \frac{d_K^m(l)}{l^{\frac{mn+1}{2mn}}}\tau\biggl(2n\pi\Bigl(\frac{lx}{D_K^m}\Bigr)^{\frac{1}{mn}}\biggr) + O_{n,m,\varepsilon}\bigl(x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}} R^{-\frac{mn+1}{2mn}+\varepsilon} + x^{\frac{mn-2}{2mn}} D_K^{\frac{1}{n}} R^{\frac{mn-2}{2mn}+\varepsilon}\bigr). \tag{2-11}$$

From estimates (2-10) and (2-11), it is obtained that

$$\Delta_K^m(x) = \frac{Cx^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}}}{2\pi i} \sum_{l \le R} \frac{d_K^m(l)}{l^{\frac{mn+1}{2mn}}} \tau\left(2mn\pi\left(\frac{lx}{D_K^m}\right)^{\frac{1}{mn}}\right) + O_{n,m,\varepsilon}\big(x^{\frac{mn-2}{2mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{\frac{mn-2}{2mn}+\varepsilon} + x^{\frac{mn-1}{mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{-\frac{1}{mn}+\varepsilon}\big).$$

Next we consider the above sum. Let $F_K = d_K^m * \mu$, where $*$ is the Dirichlet product symbol. From Lemma 2-5 this becomes

$$\begin{aligned}\Delta_K^m(x) = O_{n,m,\varepsilon}\Bigg(&L^{1-\alpha} + x^{\frac{mn-1}{2mn}} D_K^{\frac{1}{2n}} R^{\varepsilon} \max_{L \le S \le R} S^{-\frac{mn+1}{2mn}} \\ &\times \max_{S < S_1 \le 2S} \max_{\substack{M,N \le S_1 \\ MN \asymp S}} \max_{\substack{M \le M_1 \le 2M \\ N \le N_1 \le 2N}} \left| \sum_{M < l \le M_1} F_K(l) \sum_{N < k \le N_1} \boldsymbol{e}\left(mn\left(\frac{xlk}{D_K^m}\right)^{\frac{1}{mn}}\right)\right| \\ &+ x^{\frac{mn-2}{2mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{\frac{mn-2}{2mn}+\varepsilon} + x^{\frac{mn-1}{mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} R^{-\frac{1}{mn}+\varepsilon}\Bigg). \qquad \square\end{aligned}$$

Let $\mathcal{S}_K(x, S)$ be the sum in the O-term, that is,

$$S^{-\frac{mn+1}{2mn}} \max_{S < S_1 \le 2S} \max_{\substack{M,N \le S_1 \\ MN \asymp S}} \max_{\substack{M \le M_1 \le 2M \\ N \le N_1 \le 2N}} \left| \sum_{M < l \le M_1} F_K(l) \sum_{N < k \le N_1} \boldsymbol{e}\left(mn\left(\frac{xlk}{D_K^m}\right)^{\frac{1}{mn}}\right)\right|.$$

This proposition reduces the initial problem to an exponential sums problem. There are many results to estimate an exponential sums. In the next section, we estimate the Piltz divisor function by using some results for exponential sums established by many authors.

3. Estimate of counting function

In the last section, we showed that the error term of the Piltz divisor function $\Delta_K^m(x)$ can be expressed as an exponential sum. Let $X > 1$ be a real number, $1 \le M < M_1 \le 2M$ and $1 \le N < N_1 \le 2N$ be integers, and $(a_m), (b_n) \subset \mathbb{C}$ be sequences of complex numbers, and let $\alpha, \beta \in \mathbb{R}$. Then we define

$$\mathcal{S} = \sum_{M < m \le M_1} a_m \sum_{N < n \le N_1} b_n \boldsymbol{e}\left(X\left(\frac{m}{M}\right)^{\alpha}\left(\frac{n}{N}\right)^{\beta}\right). \tag{3-1}$$

Lemma 3-2 [Wu 1998]. *Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha\beta(\alpha-1)(\beta-1) \neq 0$, and $|a_m| \le 1$ and $|b_n| \le 1$ and $\mathcal{L} = \log(XMN+2)$. Then*

$$\mathcal{L}^{-2}\mathcal{S} = O\big((XM^3N^4)^{\frac{1}{5}} + (X^4M^{10}N^{11})^{\frac{1}{16}} + (XM^7N^{10})^{\frac{1}{11}} + MN^{\frac{1}{2}} + (X^{-1}M^{14}N^{23})^{\frac{1}{22}} + X^{-\frac{1}{2}}MN\big).$$

Next Bordellès also shows this lemma by using estimates for triple exponential sums by Robert and Sargos.

Lemma 3-3 [Bordellès 2015]. *Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha\beta(\alpha-1)(\beta-1) \neq 0$, and $|a_m| \leq 1$ and $|b_n| \leq 1$. If $X = O(M)$, then*

$$(MN)^{-\varepsilon}\mathcal{S} = O\big((XM^5N^7)^{\frac{1}{8}} + N(X^{-2}M^{11})^{\frac{1}{12}} + (X^{-3}M^{21}N^{23})^{\frac{1}{24}} + M^{\frac{3}{4}}N + X^{-\frac{1}{4}}MN\big).$$

The following Srinivasan result is important for our estimating $\Delta_K^m(x)$.

Lemma 3-4 [Srinivasan 1962]. *Let N and P be positive integers and $u_n \geq 0$, $v_p > 0$, A_n, and B_p denote constants for $1 \leq n \leq N$ and $1 \leq p \leq P$. Then there exists q with properties*

$$Q_1 \leq q \leq Q_2$$

and

$$\sum_{n=1}^{N} A_n q^{u_n} + \sum_{p=1}^{P} B_p q^{-v_p} = O\bigg(\sum_{n=1}^{N}\sum_{p=1}^{P} \sqrt[u_n+v_p]{A_n^{v_p} B_p^{u_n}} + \sum_{n=1}^{N} A_n Q_1^{u_n} + \sum_{p=1}^{P} B_p Q_2^{-v_p}\bigg).$$

The constant involved in the O-symbol is less than $N + P$.

Srinivasan [1962] remarks that the inequality in Lemma 3-4 corresponds to the "best possible" choice of q in the range $Q_1 \leq q \leq Q_2$. We apply Lemma 3-4 to improve the error term $\Delta_K^m(x)$.

Theorem 3-5. *For every $\varepsilon > 0$ the following estimate holds. When $mn \geq 4$, then*

$$\Delta_K^m(x) = O_{n,m,\varepsilon}\big(x^{\frac{2mn-3}{2mn+1}+\varepsilon} D_K^{\frac{2m}{2mn+1}+\varepsilon}\big)$$

as x tends to infinity.

Proof. We note that

$$\bigg|\sum_{M<l\leq M_1} F_K(l) \sum_{N<k\leq N_1} \boldsymbol{e}\bigg(mn\bigg(\frac{xlk}{D_K^m}\bigg)^{\frac{1}{mn}}\bigg)\bigg| = \bigg|\sum_{M<l\leq M_1} F_K(l) \sum_{N<k\leq N_1} \boldsymbol{e}\bigg(mn\bigg(\frac{xMN}{D_K^m}\bigg)^{\frac{1}{mn}}\bigg(\frac{l}{M}\bigg)^{\frac{1}{mn}}\bigg(\frac{k}{N}\bigg)^{\frac{1}{mn}}\bigg)\bigg|.$$

We use the above lemmas with $X=mn\big(\frac{xMN}{D_K^m}\big)^{\frac{1}{mn}}>0$. Let $0\le\alpha\le\frac{1}{3}$; we consider four cases:

$$\begin{aligned}
&\text{case 1,} && S^{\alpha}\ll N\ll S^{\frac{1}{2}},\\
&\text{case 2,} && S^{\frac{1}{2}}\ll N\ll S^{1-\alpha},\\
&\text{case 3,} && S^{1-\alpha}\ll N,\\
&\text{case 4,} && N\ll S^{\alpha}.
\end{aligned}$$

When $S^{\alpha}\ll N\ll S^{\frac{1}{2}}$, we apply Lemma 3-2 and this gives

$$\begin{aligned}
\text{(3-6)}\quad & S^{-\varepsilon}x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}\mathcal{S}_K(x,S)\\
&=O_{n,m,\varepsilon}\big(x^{\frac{5mn-3}{10mn}}D_K^{\frac{3}{10n}}R^{\frac{2mn-3}{10mn}}+x^{\frac{2mn-1}{4mn}}D_K^{\frac{1}{4n}}R^{\frac{5mn-8}{32mn}}+x^{\frac{11mn-9}{22mn}}D_K^{\frac{9}{22n}}R^{\frac{6mn-9}{22mn}}\\
&\qquad+x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}R^{\frac{mn-1}{2mn}-\frac{1}{2}\alpha}+x^{\frac{11mn-12}{22mn}}D_K^{\frac{6}{11n}}R^{\frac{15mn-24}{44mn}}+x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}R^{\frac{mn-2}{2mn}}\big).
\end{aligned}$$

When $S^{\frac{1}{2}}\ll N\ll S^{1-\alpha}$ we use Lemma 3-2 again reversing the role of M and N. We obtain the same estimate for the case that $S^{\alpha}\ll N\ll S^{\frac{1}{2}}$. For case 3, we use Lemma 3-3:

$$\begin{aligned}
\text{(3-7)}\quad & S^{-\varepsilon}x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}\mathcal{S}_K(x,S)\\
&=O_{n,m,\varepsilon}\big(x^{\frac{4mn-3}{8mn}}D_K^{\frac{3}{8n}}R^{\frac{mn-3}{8mn}+\frac{1}{4}\alpha}+x^{\frac{3mn-4}{6mn}}D_K^{\frac{2}{3n}}R^{\frac{5mn-8}{12mn}+\frac{1}{12}\alpha}\\
&\quad+x^{\frac{4mn-5}{8mn}}D_K^{\frac{5}{8n}}R^{\frac{3mn-5}{8mn}-\frac{1}{12}\alpha}+x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}R^{\frac{mn-2}{4mn}+\frac{1}{4}\alpha}+x^{\frac{2mn-3}{4mn}}D_K^{\frac{3}{4n}}R^{\frac{2mn-3}{4mn}}\big).
\end{aligned}$$

If $x^{\frac{1}{mn(1-\alpha)-1}}D_K^{-\frac{m}{mn(1-\alpha)-1}}\ll S$, the condition of Lemma 3-3, $X=O(N)$, is satisfied. Therefore, it suffices to choose $L=x^{\frac{1}{mn(1-\alpha)-1}}D_K^{-\frac{m}{mn(1-\alpha)-1}}$. For case 4, we use Lemma 3-3 again reversing the role of M and N. We obtain the same estimate for the case that $N\ll S^{\alpha}$. Combining (3-6) and (3-7) with Proposition 2-6, we obtain

$$\begin{aligned}
\text{(3-8)}\quad \Delta_K^m(x)=O_{n,m,\varepsilon}\big(&x^{\frac{5mn-3}{10mn}}D_K^{\frac{3}{10n}}R^{\frac{2mn-3}{10mn}+\varepsilon}+x^{\frac{2mn-1}{4mn}}D_K^{\frac{1}{4n}}R^{\frac{5mn-8}{32mn}+\varepsilon}\\
&+x^{\frac{11mn-9}{22mn}}D_K^{\frac{9}{22n}}R^{\frac{6mn-9}{22mn}+\varepsilon}+x^{\frac{mn-1}{2mn}}D_K^{\frac{1}{2n}}R^{\frac{mn-1}{2mn}-\frac{1}{2}\alpha+\varepsilon}\\
&+x^{\frac{11mn-12}{22mn}}D_K^{\frac{6}{11n}}R^{\frac{15mn-24}{44mn}+\varepsilon}+x^{\frac{mn-2}{2mn}}D_K^{\frac{1}{n}}R^{\frac{mn-2}{2mn}+\varepsilon}\\
&+x^{\frac{4mn-3}{8mn}}D_K^{\frac{3}{8n}}R^{\frac{mn-3}{8mn}+\frac{1}{4}\alpha+\varepsilon}+x^{\frac{3mn-4}{6mn}}D_K^{\frac{2}{3n}}R^{\frac{5mn-8}{12mn}+\frac{1}{12}\alpha+\varepsilon}\\
&+x^{\frac{4mn-5}{8mn}}D_K^{\frac{5}{8n}}R^{\frac{3mn-5}{8mn}+\frac{1}{12}\alpha+\varepsilon}+x^{\frac{2mn-3}{4mn}}D_K^{\frac{3}{4n}}R^{\frac{2mn-3}{4mn}+\varepsilon}\\
&+x^{\frac{mn-1}{mn}+\varepsilon}D_K^{\frac{1}{n}+\varepsilon}R^{-\frac{1}{mn}+\varepsilon}+x^{\frac{1-\alpha}{mn(1-\alpha)-1}}D_K^{-\frac{m(1-\alpha)}{mn(1-\alpha)-1}}\big).
\end{aligned}$$

By Lemma 3-4 with $x^{\frac{1}{mn(1-\alpha)-1}} D_K^{-\frac{m}{mn(1-\alpha)-1}} \le R \le xD$ there exists R such that the error term of estimate (3-8) is much less than

$$
\begin{aligned}
&x^{\frac{2mn}{2mn+7}+\varepsilon} D_K^{\frac{2m}{2mn+7}+\varepsilon} + x^{\frac{5mn+3}{5mn+24}+\varepsilon} D_K^{\frac{5m}{5mn+24}+\varepsilon} + x^{\frac{6mn-4}{6mn+13}+\varepsilon} D_K^{\frac{6m}{6mn+13}+\varepsilon} \\
&+ x^{\frac{(1-\alpha)mn+\alpha-1}{(1-\alpha)mn+1}+\varepsilon} D_K^{\frac{(1-\alpha)m}{(1-\alpha)mn+1}+\varepsilon} + x^{\frac{15mn-17}{15mn+20}+\varepsilon} D_K^{\frac{3m}{3mn+4}+\varepsilon} + x^{\frac{mn-2}{mn}+\varepsilon} D_K^{\frac{1}{n}+\varepsilon} \\
&+ x^{\frac{(2\alpha+1)mn-2\alpha}{(2\alpha+1)mn+5}+\varepsilon} D_K^{\frac{(2\alpha+1)m}{(2\alpha+1)mn+5}+\varepsilon} + x^{\frac{(\alpha+5)mn-\alpha-7}{(\alpha+5)mn+4}+\varepsilon} D_K^{\frac{(\alpha+5)m}{(\alpha+5)mn+4}+\varepsilon} \\
&+ x^{\frac{(2\alpha+9)mn-2\alpha-12}{(2\alpha+9)mn+9}+\varepsilon} D_K^{\frac{(2\alpha+9)m}{(2\alpha+9)mn+9}+\varepsilon} + x^{\frac{2mn-3}{2mn+1}+\varepsilon} D_K^{\frac{2m}{2mn+1}+\varepsilon} \\
&+ x^{\frac{5mn(1-\alpha)-6+3\alpha}{10mn(1-\alpha)-10}+\varepsilon} D_K^{\frac{m-3m\alpha}{10mn(1-\alpha)-10}+\varepsilon} + x^{\frac{16mn(1-\alpha)-19+8\alpha}{32mn(1-\alpha)-32}+\varepsilon} D_K^{\frac{3m-8m\alpha}{32mn(1-\alpha)-32}+\varepsilon} \\
&+ x^{\frac{11mn(1-\alpha)-14+9\alpha}{22mn(1-\alpha)-22}+\varepsilon} D_K^{\frac{3m-9m\alpha}{22mn(1-\alpha)-22}+\varepsilon} + x^{\frac{1}{2}+\varepsilon} \\
&+ x^{\frac{22mn(1-\alpha)-31+24\alpha}{44mn(1-\alpha)-44}+\varepsilon} D_K^{\frac{9m-24m\alpha}{44mn(1-\alpha)-44}+\varepsilon} + x^{\frac{mn(1-\alpha)-2+2\alpha}{2mn(1-\alpha)-2}+\varepsilon} D_K^{\frac{m-2m\alpha}{2mn(1-\alpha)-2}+\varepsilon} \\
&+ x^{\frac{4mn(1-\alpha)-6+5\alpha}{8mn(1-\alpha)-8}+\varepsilon} D_K^{\frac{2m-5m\alpha}{8mn(1-\alpha)-8}+\varepsilon} + x^{\frac{6mn(1-\alpha)-9+9\alpha}{12mn(1-\alpha)-12}+\varepsilon} D_K^{\frac{3m-9m\alpha}{12mn(1-\alpha)-12}+\varepsilon} \\
&+ x^{\frac{12mn(1-\alpha)-18+17\alpha}{24mn(1-\alpha)-24}+\varepsilon} D_K^{\frac{6m-17m\alpha}{24mn(1-\alpha)-24}+\varepsilon} + x^{\frac{2mn(1-\alpha)-3+3\alpha}{4mn(1-\alpha)-4}+\varepsilon} D_K^{\frac{m-3m\alpha}{4mn(1-\alpha)-4}+\varepsilon} \\
&+ x^{\frac{1-\alpha}{mn(1-\alpha)-1}} D_K^{-\frac{m(1-\alpha)}{mn(1-\alpha)-1}}.
\end{aligned}
$$

When $mn \ge 4$ and $\alpha = \frac{mn+3}{7mn-5}$, then we have

$$\Delta_K^m(x) = O_{n,m,\varepsilon}\big(x^{\frac{2mn-3}{2mn+1}+\varepsilon} D_K^{\frac{2m}{2mn+1}+\varepsilon}\big). \qquad \square$$

For $mn \ge 4$ this theorem gives new results for the Piltz divisor problem over number fields. In particular, if we fix K with $[K:\mathbb{Q}] = 4$, then we improve the estimate for $\Delta_K(x)$ as follows:

Corollary 3-9. *For any number field K with $[K:\mathbb{Q}] = 4$,*

$$\Delta_K(x) = O_{K,\varepsilon}\big(x^{\frac{5}{9}+\varepsilon}\big).$$

This result is better than Bordellès' result.

4. Application

In this section we introduce some applications of our theorems. First we obtain a uniform estimate for the ideal counting function $I_K(x)$. From the proof of Theorem 3-5, we obtain the following theorem.

Theorem 4-1. *For all $\varepsilon > 0$ for any fixed $0 \le \beta \le \frac{8}{2n+5} - \varepsilon$ and $C > 0$ the following holds. If K runs through number fields with $[K:\mathbb{Q}] \le n$ and $D_K \le Cx^\beta$, then*

$$\Delta_K(x) = O_{C,n,\varepsilon}\bigl(x^{\frac{2n-3+2\beta}{2n+1}+\varepsilon}\bigr).$$

The condition $D_K \le Cx^\beta$ is caused by the relation between the principal term and the error term. It is well known that $I_K(x)$ is very important to estimate the distribution of relatively r-prime lattice points. We regard an ℓ-tuple of ideals $(\mathfrak{a}_1, \mathfrak{a}_2, \dots, \mathfrak{a}_\ell)$ of $\mathbb{O}_K$ as a lattice point in K^ℓ. We say that a lattice point $(\mathfrak{a}_1, \mathfrak{a}_2, \dots, \mathfrak{a}_\ell)$ is *relatively r-prime* for a positive integer r if there exists no prime ideal $\mathfrak{p}$ such that $\mathfrak{a}_1, \mathfrak{a}_2, \dots, \mathfrak{a}_\ell \subset \mathfrak{p}^r$. Let $V_\ell^r(x, K)$ denote the number of relatively r-prime lattice points $(\mathfrak{a}_1, \mathfrak{a}_2, \dots, \mathfrak{a}_\ell)$ such that their ideal norm $\mathfrak{N}\mathfrak{a}_i \le x$.

B. D. Sittinger [2010] shows that

$$V_\ell^r(x, K) \sim \frac{\rho_K^\ell}{\zeta_K(r\ell)} x^\ell,$$

where ρ_K is the residue of ζ_K at $s = 1$. It is well known that

$$\rho_K = \frac{2^{r_1}(2\pi)^{r_2} h_K R_K}{w_K\sqrt{D_K}}, \tag{4-2}$$

where h_K is the class number of K, R_K is the regulator of K, and w_K is the number of roots of unity in $\mathbb{O}_K^*$.

After that we show some results for the error term:

$$E_\ell^r(x, K) = V_\ell^r(x, K) - \frac{\rho_K^\ell}{\zeta_K(r\ell)} x^\ell.$$

In [Takeda 2017; Takeda and Koyama 2018] we consider the relation between the relatively r-prime problem and other mathematical problems. If we assume the Lindelöf hypothesis for $\zeta_K(s)$, then it holds that for all $\varepsilon > 0$

$$E_\ell^r(x, K) = \begin{cases} O_\varepsilon\bigl(x^{\frac{1}{r}(\frac{3}{2}+\varepsilon)}\bigr) & \text{if } r\ell = 2, \\ O_\varepsilon\bigl(x^{\ell-\frac{1}{2}+\varepsilon}\bigr) & \text{otherwise.} \end{cases} \tag{4-3}$$

From easy calculation, we obtain the following corollary.

Corollary 4-4. *For all $\varepsilon > 0$ and for any fixed $0 \le \beta \le \frac{8}{2n+5} - \varepsilon$ and $C > 0$ the following holds. If K runs through number fields with $[K:\mathbb{Q}] \le n$ and $D_K \le Cx^\beta$, then*

$$E_\ell^r(x, K) = \begin{cases} O_{C,n,\varepsilon}\bigl(x^{\frac{4n-2}{r(2n+1)}+\frac{4}{2n+1}\beta+\varepsilon}\bigr) & \textit{if } r\ell = 2, \\ O_{C,n,\varepsilon}\bigl(x^{\ell-\frac{4}{2n+1}+\frac{2n+5-(2n+1)\ell}{2(2n+1)}\beta+\varepsilon}\bigr) & \textit{otherwise.} \end{cases}$$

For the proof of this corollary, please see the proof of Theorem 4.1 of [Takeda and Koyama 2018].

5. Conjecture

Theorem 4-1 states good uniform upper bounds. It is proposed that for all number fields K the best uniform upper bound of the error term is better than that on the assumption of the Lindelöf hypothesis (1-6).

Conjecture 5-1. *If K runs through number fields with $D_K < x$, then*

$$\Delta_K^m(x) = o\big(x^{\frac{1}{2}}\big).$$

From estimate (1-5), this conjecture may give the best estimate for uniform upper bound of $\Delta_K^m(x)$. As we remarked above (Theorem 1-2) this conjecture is very difficult even when K is fixed and $m = 1$.

References

[Atkinson 1941] F. V. Atkinson, "A divisor problem", *Quart. J. Math.* **12** (1941), 193–200. MR Zbl

[Bordellès 2015] O. Bordellès, "On the ideal theorem for number fields", *Funct. Approx. Comment. Math.* **53**:1 (2015), 31–45. MR Zbl

[Bourgain and Watt 2017] J. Bourgain and N. Watt, "Mean square of zeta function, circle problem and divisor problem revisited", preprint, 2017. arXiv

[Girstmair et al. 2005] K. Girstmair, M. Kühleitner, W. Müller, and W. G. Nowak, "The Piltz divisor problem in number fields: an improved lower bound by Soundararajan's method", *Acta Arith.* **117**:2 (2005), 187–206. MR Zbl

[Huxley 2002] M. N. Huxley, "Integer points, exponential sums and the Riemann zeta function", pp. 275–290 in *Number theory for the millennium, II* (Urbana, IL, 2000), edited by M. A. Bennett et al., Peters, Natick, MA, 2002. MR Zbl

[Lao 2010] H. Lao, "On the distribution of integral ideals and Hecke Grössencharacters", *Chin. Ann. Math. Ser. B* **31**:3 (2010), 385–392. MR Zbl

[Müller 1988] W. Müller, "On the distribution of ideals in cubic number fields", *Monatsh. Math.* **106**:3 (1988), 211–219. MR Zbl

[Nowak 1993] W. G. Nowak, "On the distribution of integer ideals in algebraic number fields", *Math. Nachr.* **161** (1993), 59–74. MR Zbl

[Rademacher 1959] H. Rademacher, "On the Phragmén–Lindelöf theorem and some applications", *Math. Z.* **72** (1959), 192–204. MR Zbl

[Sittinger 2010] B. D. Sittinger, "The probability that random algebraic integers are relatively r-prime", *J. Number Theory* **130**:1 (2010), 164–171. MR Zbl

[Srinivasan 1962] B. R. Srinivasan, "On van der Corput's and Nieland's results on the Dirichlet's divisor problem and the circle problem", *Proc. Natl. Inst. Sci. India A* **28**:5 (1962), 732–742. Zbl

[Takeda 2017] W. Takeda, "Visible lattice points and the extended Lindelöf hypothesis", *J. Number Theory* **180** (2017), 297–309. MR Zbl

[Takeda and Koyama 2018] W. Takeda and S.-y. Koyama, "Estimates of lattice points in the discriminant aspect over abelian extension fields", *Forum Math.* **30**:3 (2018), 767–773. MR Zbl

[Wu 1998] J. Wu, "On the average number of unitary factors of finite abelian groups", *Acta Arith.* **84**:1 (1998), 17–29. MR Zbl

Received July 23, 2018. Revised August 29, 2018.

WATARU TAKEDA
DEPARTMENT OF MATHEMATICS
NAGOYA UNIVERSITY
CHIKUSA-KU
NAGOYA
JAPAN
d18002r@math.nagoya-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.617

EXPLICIT WHITTAKER DATA FOR ESSENTIALLY TAME SUPERCUSPIDAL REPRESENTATIONS

GEO KAM-FAI TAM

Based on the ideas of Bushnell and Henniart, and of Paskunas and Stevens, we construct explicit Whittaker data for an essentially tame supercuspidal representation of $\mathbf{GL}_n(F)$, where F is a non-Archimedean local field.

1. Introduction

Let F be a non-Archimedean local field, V be an n-dimensional F-vector space, and G be the group $\mathrm{Aut}_F(V)$ of F-linear automorphisms of V, usually regarded as $\mathrm{GL}_n(F)$ by choosing a basis of V. Let π be a supercuspidal representation of G. As a classical result in [Gel'fand and Kajdan 1975], we know that π admits a unique Whittaker model. More precisely, take a tuple of Whittaker data (N, ψ) consisting of a maximal unipotent subgroup N of G and a nondegenerate character ψ of N, in the sense that its restriction to every simple root subgroup of N is nontrivial, then we have

$$\mathrm{Hom}_G(\pi, \mathrm{Ind}_N^G \psi) = 1.$$

As another classical result in [Bushnell and Kutzko 1993], we know that π is isomorphic to a compactly induced representation from a finite-dimensional representation Λ of an open compact-mod-center subgroup $\mathbf{J}$ of G; that is,

$$\pi \cong \mathrm{cInd}_{\mathbf{J}}^G \Lambda.$$

Using Frobenius reciprocity and Mackey's formula [Kutzko 1977], the existence and uniqueness of a Whittaker model is equivalent to the existence of a pair (N, ψ) as above such that

$$\mathrm{Hom}_{N\cap\mathbf{J}}(\psi, \Lambda) \neq 0, \tag{1-1}$$

and the pair is unique up to conjugation by $\mathbf{J}$ [Bushnell and Henniart 1998].

The above is the starting point of [Bushnell and Henniart 1998] in describing an explicit Whittaker function for a supercuspidal representation. This description, together with the result in [Paskunas and Stevens 2008], turn out to be useful in computing the epsilon factor for a certain pair of supercuspidal representations which

MSC2010: primary 22E50; secondary 20G05.
Keywords: Whittaker data, essentially tame supercuspidal representations.

differ only at the "tame level" (see [Paskunas and Stevens 2008, Section 7] and [Kim 2014, Theorem 4.4.3]). However, as pointed out in the introduction of [Paskunas and Stevens 2008], the proof of [Bushnell and Henniart 1998, Lemma 2.10] contains a gap, so they have to bypass the problem using a "black box" case (explained below) in [Paskunas and Stevens 2008].

The purpose of this paper is to construct explicit Whittaker data (N, ψ) for an essentially tame supercuspidal representation π. The essential tameness condition means that, by the definition in [Bushnell and Henniart 2005], if the group

$$\{\chi : \text{ unramified character of } F^\times \text{ such that } (\chi \circ \det) \otimes \pi \cong \pi\}$$

has order f, which is necessarily a divisor of n, then the residual characteristic p of F does not divide n/f. We will explain the advantage of restricting to the essentially tame case at the end of this introduction.

We summarize briefly the method of constructing our Whittaker data, mostly following Theorem 3.3 and Section 4.2 of [Paskunas and Stevens 2008]. Let θ be the simple character of a compact subgroup H^1 of G, in the sense of [Bushnell and Kutzko 1993, Section 3.2], underlying a chosen inducing type Λ of π. Associated to θ is an element $\beta \in A = \mathrm{End}_F(V)$ such that $E_0 = F[\beta]$ is a subfield of A and is tamely ramified over F in the essentially tame case. We will construct a maximal unipotent subgroup N satisfying

$$\theta|_{H^1 \cap N} = \psi_\beta|_{H^1 \cap N},$$

where $\psi_\beta : A \to \mathbb{C},\ x \mapsto \psi_F \circ \mathrm{tr}_{A/F}(\beta(x-1))$ with ψ_F being an additive character of F trivial on $\mathfrak{p}_F$ but not on $\mathfrak{o}_F$, together with other conditions in [Paskunas and Stevens 2008, Theorem 3.3].

The above unipotent subgroup N is defined by a particular ordered basis $\mathfrak{b}$ of V, given in (4-4). To describe it briefly, associated to the element β is a set $\{\beta_i\}_{i=0}^t$ of approximation elements such that $\sum_{i=0}^t \beta_i = \beta$ and, in the essentially tame case, that $E_j = F\big[\sum_{j \le i} \beta_i\big]$ form a tower of intermediate (tamely ramified) extensions between E_0 and F. When $t = 0$, which is known as the minimal case in [Bushnell and Kutzko 1993, (1.4.15)], we define $\mathfrak{b}$ cyclicly using the element β, similar to the one defined in [Bushnell and Henniart 1998, 2.1 Proposition]. In the presence of multiple approximation elements (i.e., $t > 0$), we define $\mathfrak{b}$ cyclicly also, but in an inductive way along these elements. Our $\mathfrak{b}$ is different from the one defined in [Bushnell and Henniart 1998, 2.1 Proposition] at the level of the complexity of approximations.

It is unknown whether ψ_β can be extended to a character of N; however, using the matrix presentation of β with respect to $\mathfrak{b}$, we construct an analogous element $\alpha \in A$ such that $\psi = \psi_\alpha$ is a nondegenerate character of N. As the main result in Theorem 5.1, we will show that ψ satisfies

$$\psi|_{H^1 \cap N} = \psi_\beta|_{H^1 \cap N},$$

and that the conditions in [Paskunas and Stevens 2008, Theorem 3.3] are satisfied as well, which is enough to imply that (1-1) holds for our (N, ψ). As a note, the case when $[E_0 : F] = n$, i.e., E_0 is a maximal subfield in A, is the "black box" case deemed by [Paskunas and Stevens 2008]. Hence when E_0/F is moreover tamely ramified, our (N, ψ) serve as a "black box" character for the arguments in [Paskunas and Stevens 2008, Section 3].

The description of our basis $\mathfrak{b}$, combined with the results in [Paskunas and Stevens 2008; Kim 2014], provides a direct formula of the conductor of the epsilon factor for a certain pair of supercuspidal representations mentioned above and which are, as in our present paper, essentially tame. Such a formula can also be deduced from the general conductor formula in [Bushnell et al. 1998], obtained using the theory of intertwining operators of Shahidi, with an inductive calculation of a certain discriminant of β [Bushnell and Henniart 2003]. We will explain it briefly in the last section.

Note that, using the rational canonical form of β, we can of course extend ψ_β to a character of the maximal unipotent subgroup N_β defined using the cyclic basis generated by β (as in [Bushnell and Henniart 1998, 2.1 Proposition]). However, (N_β, ψ_β) may not be good Whittaker data; in particular, they do not give the correct conductor formula for the epsilon factors of pairs in Paskunas and Stevens' case.

Finally, we remark that the whole development of our main result requires E_0/F to be tamely ramified. As we will see, the tower of intermediate extensions $\{E_j\}_{j=0}^t$ between E_0 and F allows us to define inductively the basis $\mathfrak{b}$ in (4-4), and consequently decompose the simple character θ and the compact subgroup $H^1 \cap N$ inductively to derive our main result. The author believes that a more complex technique is required for the general case beyond the essentially tame case, and hopes to deal with it in his future work.

1A. ***Notations.*** Let F be a non-Archimedean local field with an algebraic closure denoted by $\bar{F}$. Denote by $\mathfrak{o}_F$ the ring of integers of F and by $\mathfrak{p}_F$ the maximal ideal of $\mathfrak{o}_F$. The residue field $\mathbf{k}_F = \mathfrak{o}_F/\mathfrak{p}_F$ of F is a finite extension of $\mathbb{F}_p$. Denote by $v_F : F \to \mathbb{Z} \cup \{\infty\}$ the valuation of F.

If $r \in \mathbb{R}$, we denote by $r+$ the smallest integer strictly greater than r.

2. Tamely ramified extensions

The main purpose of this section is to gather some known facts concerning tamely ramified extensions. More importantly, we consider minimal elements in a tamely ramified extension, and study how they form bases with nice properties on lattice filtrations.

2A. ***Complementary subgroup.*** Let E/F be a tamely ramified extension of degree $n = n(E/F)$ and ramification degree $e = e(E/F)$. Put $f = f(E/F) = n/e$.

Throughout we fix a chosen uniformizer ϖ_F of F, and let $\boldsymbol{\mu}_F$ be the group of roots of unity with order coprime to p. We also fix ϖ_E and let $\boldsymbol{\mu}_E$ similarly for the field E, and assume that $\varpi_E^e \varpi_F^{-1} \in \boldsymbol{\mu}_E$. We define the complementary subgroup C_E of $E^\times$ to be the subgroup generated by ϖ_E and $\boldsymbol{\mu}_E$. It can be shown that C_E depends only on the choice of ϖ_F. Moreover, if K/F is an intermediate field extension in E, then $C_K \subseteq C_E$. We denote by C_F^{tame} the union of all C_E, with E ranges over all tamely ramified extensions of F.

If r is a positive integer, let U_F^r be the r-th unit group $1+\mathfrak{p}_F^r$. In general for $r \in \mathbb{R}$, we write $U_F^r = U_F^{\lceil r \rceil}$ where $\lceil r \rceil$ is the smallest integer $\geq r$, and write $U_F^{r+} = U_F^{\lceil r \rceil_+}$ where $\lceil r \rceil_+$ is the smallest integer $> r$. We define U_E^r similarly. Any element $b \in E^\times$ can be uniquely decomposed as cu where $c \in C_E$ and $u \in U_E^1$. We call c the *first term* of b.

2B. ***Minimality.*** At the beginning of this subsection, we only require E/F to be a finite separable extension of degree n. Later we will require E/F to be moreover tamely ramified.

Let $E = F[\alpha]$ for some $\alpha \in E$. Denote $e = e(E/F)$ and $v = v_E(\alpha)$. From [Kutzko and Manderscheid 1988, Proposition 1.5], we say that α is *minimal* over F if it satisfies

(I) $\gcd(v, e) = 1$, and

(II) any one of the following conditions:

(a) $\mathfrak{o}_F[\beta] = \mathfrak{o}_K$, where K/F is the maximal unramified extension in E/F and $\beta = N_{E/K}(\alpha)/\varpi_F^v$.

(b) The elements $\{x_j\}_{j=1}^n$, where $x_j = \alpha^j/\varpi_F^{\lfloor jv/e \rfloor}$, form an $\mathfrak{o}_F$-basis of $\mathfrak{o}_E$. In particular we have $\mathfrak{o}_E = \oplus_{j=0}^{n-1} \mathfrak{o}_F x_j$.

(c) $\mathbf{k}_E = \mathbf{k}_F[\gamma + \mathfrak{p}_E]$, where $\gamma = x_e = \alpha^e/\varpi_F^v$ ([Bushnell and Kutzko 1993, (1.4.15)]).

By [Kutzko and Manderscheid 1988, Proposition 1.5], given (I), the three conditions in (II) are equivalent.

To incorporate the construction of simple characters, we recall another equivalent minimality condition from [Bushnell and Kutzko 1993]. Let V be a finite-dimensional E-vector space. We first regard V as an F-vector space and denote $A = \text{End}_F(V)$. Let $\mathfrak{A}$ be an hereditary $\mathfrak{o}_F$-order in A, with Jacobson radical $\mathfrak{P}$ and normalized by $E^\times$. Let $v_{\mathfrak{A}}$ be the valuation on A associated with $\mathfrak{A}$. Let B be the centralizer of E in A, and denote $\mathfrak{B} = \mathfrak{A} \cap B$. Recall from [Bushnell and Kutzko 1993, 1.4] the $\mathfrak{o}_F$-lattice

$$\mathfrak{N}_k(\alpha, \mathfrak{A}) = \{x \in \mathfrak{A} : \alpha x - x\alpha \in \mathfrak{P}^k\}$$

and the critical exponent

$$k_0(\alpha, \mathfrak{A}) = \max\{k \in \mathbb{Z} : \mathfrak{N}_k(\alpha, \mathfrak{A}) \not\subset \mathfrak{B} + \mathfrak{P}\}.$$

By convention, if $\alpha \in F$, we put $k_0(\alpha, \mathfrak{A}) = \infty$. In fact, by [Bushnell and Kutzko 1993, (1.4.13)(ii)], the definition is independent of the vector space V. One important property is [Bushnell and Kutzko 1993, (1.4.15)]

$$v_{\mathfrak{A}}(\alpha) \leq k_0(\alpha, \mathfrak{A}), \tag{2-1}$$

with equality if and only if α is minimal over F.

Proposition 2.1. *An element $\alpha \in \bar{F}^\times$ with finite order modulo $F^\times$ coprime to p is minimal over F. In particular, any element in C_F^{tame} is minimal over F.*

Proof. The second statement follows directly from the first, so we focus on proving the first statement. The idea of the proof comes from [Reimann 1991, Lemma 2.8]. Let α have finite order modulo $F^\times$, and assume that $\alpha \notin F$ (otherwise the result is trivial). We will use condition (2-1) and, since the condition does not depend on the choice of the vector space V, we can assume V to be $E = F[\alpha]$ as an F-vector space, and denote A, $\mathfrak{A}$, and $\mathfrak{P}$ as above, so that $B = E$ and $\mathfrak{B} = \mathfrak{o}_E$. The statement can be proved if we show that

$$\mathfrak{N}_k(\alpha, \mathfrak{A}) = \mathfrak{o}_E + \mathfrak{P}^{k - v_{\mathfrak{A}}(\alpha)}$$

for all $k \in \mathbb{Z}$. Let $\tau(x) = \alpha x \alpha^{-1}$ for all $x \in A$, which is an F-algebra automorphism of A. We hence take $x \in \mathfrak{A}$ such that $\tau(x) - x \in \mathfrak{P}^{k - v_{\mathfrak{A}}(\alpha)}$. If m is the order of α in $\bar{F}^\times / F^\times$, we define

$$s : A \to A, \qquad X \mapsto \frac{1}{m} \sum_{i=0}^{m-1} \tau^i(X),$$

which is an F-linear projection onto E, and so $s(x) \in \mathfrak{o}_E$ as $m \in \mathfrak{o}_F^\times$. The relation

$$x = s(x) - \sum_{i=0}^{m-1} \sum_{j=1}^{i} \tau^{j-1}(\tau(x) - x)$$

implies that $x \in \mathfrak{o}_E + \mathfrak{P}^{k - v_{\mathfrak{A}}(\alpha)}$. The converse inclusion is straightforward. $\square$

Corollary 2.2. *Suppose that the field $E = F[\alpha]$, for some $\alpha \in E$, is tamely ramified over F. Then α is minimal if and only if the first term of α also generates E over F.*

Proof. We write $\alpha = au$ for some $a \in C_E$ and $u \in U_E^1$. It is straightforward to see that α satisfies minimality conditions (I) and (c) if and only if a does the same for field E. $\square$

We provide a useful calculation of the critical exponent of an element generating a tamely ramified field extension.

Proposition 2.3. *Suppose that $\beta \in A$ such that $E = F[\beta]$ is a tamely ramified extension of F, and $\mathfrak{A}$ is an $\mathfrak{o}_F$-hereditary order normalized by $E^\times$. Take $c \in C_E$ and denote $\gamma = \beta - c$. If $k_0(\gamma, \mathfrak{A}) < \upsilon_{\mathfrak{A}}(c)$, then*

$$k_0(\beta, \mathfrak{A}) = \begin{cases} k_0(\gamma, \mathfrak{A}) & \text{if } c \in F[\gamma], \\ \upsilon_{\mathfrak{A}}(c) & \text{otherwise.} \end{cases}$$

Proof. This can be derived from [Bushnell and Kutzko 1993, (2.2.8)]. □

2C. *A special property.* Suppose that V is an n-dimensional F-vector space containing an $\mathfrak{o}_F$-lattice chain $\mathcal{L}$. We call an F-basis $\{x_j\}_{j=1}^n$ of V an $\mathfrak{o}_F$-basis of $\mathcal{L}$, in the sense of [Bushnell and Kutzko 1993, (1.1.7)], if

(A) it is an $\mathfrak{o}_F$-basis of $\mathcal{L}(r)$ for some $r \in \mathbb{Z}$, and

(B) there exist $a(j,r) \in \mathbb{Z}$, for all $j = 1, \dots, n$ and $r \in \mathbb{Z}$, such that

$$\mathcal{L}(r) = \bigoplus_j \mathfrak{p}_F^{a(j,r)} x_j.$$

We may arrange the integers such that $a(j,r) \le a(j+1,r)$.

For example, if V is a field extension $E = F[\alpha]$ as in the last section, then the set $\{x_j\}_{j=1}^n$ in the minimality condition (b) is an $\mathfrak{o}_F$-basis of $\{\mathfrak{p}_E^r\}_{r\in\mathbb{Z}}$. Indeed, suppose that $\{y_j\}$ is an ordered set equal to $\{x_j\}$ as a set but with the order rearranged such that $\upsilon_E(y_j) = i$ if $j = fi + k$ with $0 \le i < e$ and $1 \le k \le f$, then we have

$$\mathfrak{p}_E^r = \bigoplus_{i=t}^{e-1} \bigoplus_{k=1}^{f} \mathfrak{p}_F^s y_{fi+k} \oplus \bigoplus_{i=0}^{t-1} \bigoplus_{k=1}^{f} \mathfrak{p}_F^{s+1} y_{fi+k} \tag{2-2}$$

if $r = se+t$ for all $s \in \mathbb{Z}$ and $t = 0, \dots, e-1$. Indeed we always have the inclusion $\supseteq$ for all $r \in \mathbb{Z}$, and we just have to show the equality for $r = 0, \dots, e-1$ by periodicity. We of course have the equality for $r = 0$ and $r = e$. We then obtain the equality for other r by counting the $\mathbf{k}_F$-dimensions of successive quotients on both sides of (2-2).

For constructing Whittaker data, we require a special property. Denote by $\upsilon_{\mathcal{L}} : V \to \mathbb{Z} \cup \{\infty\}$ the associated valuation of $\mathcal{L}$. Let $\{u_j\}_{j=1}^n$ be an $\mathfrak{o}_F$-basis of $\mathcal{L}$ satisfying the following condition.

(*) For every $u = \sum_j a_j u_j \in V$ with $a_j \in F$, we have $\upsilon_{\mathcal{L}}(a_j u_j) \ge \upsilon_{\mathcal{L}}(u)$ for all j.

This condition leads to the following simple useful result:

Proposition 2.4. *Suppose further that $\upsilon_{\mathcal{L}}(u_i) \ge \upsilon_{\mathcal{L}}(u_j)$ if $i \le j$. For every $u = \sum_i a_i u_i \in V$, if $\upsilon_{\mathcal{L}}(u) > \upsilon_{\mathcal{L}}(u_i)$ for some i, then $a_j \in \mathfrak{p}_F$ for all $j \ge i$.*

Proof. This is because $\upsilon_{\mathcal{L}}(a_j u_j) \ge \upsilon_{\mathcal{L}}(u) > \upsilon_{\mathcal{L}}(u_i) \ge \upsilon_{\mathcal{L}}(u_j)$, where the first inequality comes from condition (*) above. □

For example, the basis of E in the minimality condition (b), and hence the cyclic basis $\{\alpha^i\}_{i=0}^{[E:F]-1}$, satisfy the condition (*), by observing from (2-2). If moreover $v_E(\alpha) \leq 0$, then the cyclic basis also satisfies the conclusion in Proposition 2.4.

3. Essentially tame supercuspidal representations

In this section, we recall the construction of essentially tame supercuspidal representations of G using admissible characters.

3A. ***Structure of admissible characters.*** Given a character ξ of $F^\times$, the level of ξ is the smallest integer $r = r_F(\xi) \geq 0$ such that $\xi|_{U_F^{r+1}}$ is trivial. We call ξ tamely ramified if $r = 0$.

Suppose that E/F is a tamely ramified extension and ξ is an admissible character of $E^\times$ over F in the sense of [Howe 1977], which means that for some intermediate subfield K between E and F,

- if ξ factors through $N_{E/K}$, then $E = K$, and
- if $\xi|_{U_E^1}$ factors through $N_{E/K}$, then E/K is unramified.

From [Howe 1977, Corollary of Lemma 11] we know that an admissible character ξ admits a factorization

$$\xi = \xi_{-1}(\xi_0 \circ N_{E/E_0}) \cdots (\xi_t \circ N_{E/E_t})(\xi_{t+1} \circ N_{E/F}), \tag{3-1}$$

with notations specified as follows.

- We have a tower of field extensions

$$E = E_{-1} \supseteq E_0 \supsetneq E_1 \cdots \supsetneq E_t \supsetneq E_{t+1} = F, \tag{3-2}$$

  and each ξ_i is a character of $E_i^\times$. This tower is uniquely determined by ξ.
- Let r_i be the level of $\xi_i \circ N_{E/E_i}$, then $r = r_{t+1}$ is the level of ξ. We assume that ξ_{t+1} is trivial if $r_{t+1} = r_t$. We call the increasing sequence of integers $r_0 < \cdots < r_t$ the *jumps* of ξ, which are uniquely determined by ξ. For later computation, we put $r_{-1} = 0$.
- If $E_0 = E$, then we replace $(\xi_0 \circ N_{E/E_0})\xi_{-1}$ by ξ_0. If $E_0 \subsetneq E$, then we assume that ξ_{-1} is tamely ramified and E/E_0 is unramified.

We put $\xi\xi_{-1}^{-1} = \Xi_0 \circ N_{E/E_0}$, where $\Xi_0 = \xi_0(\xi_1 \circ N_{E_0/E_1}) \cdots (\xi_t \circ N_{E_0/E_t})(\xi_{t+1} \circ N_{E_0/F})$. Note that the jumps $\{r_i\}_{i=0}^t$ depend only on $\Xi_0 \circ N_{E/E_0}|_{U_E^1}$, and are invariant under the Galois action on ξ.

We fix an additive character ψ of F, which is assumed to be trivial on $\mathfrak{p}_F$ but not on $\mathfrak{o}_F$. For any tamely ramified extension K/F, we write $\psi_K = \psi_F \circ \mathrm{tr}_{K/F}$.

We recall several results from [Moy 1986, Section 2.2]. For $i = 0, \dots, t+1$, suppose that s_i is the level of ξ_i, which means that $s_i e(E/E_i) = r_i$; then there is $\beta_i \in \mathfrak{p}_{E_i}^{-s_i} - \mathfrak{p}_{E_i}^{-s_i+}$ such that

(3-3) $$\xi_i(x) = \psi_{E_i}(\beta_i(x-1)), \qquad \text{for all } x \in U_{E_i}^{s_i/2+}.$$

This β_i, depending on the choice of ξ_i, can be regarded as in $\mathfrak{p}_E^{-r_i} - \mathfrak{p}_E^{-r_i+}$ and chosen mod $\mathfrak{p}_E^{(-r_i/2)+}$. Let $c_i \in C_{E_i}$ be the first term of ξ_i. For $i = 0, \dots, t$, each character ξ_i is generic over E_{i+1}, in the sense that

(3-4) $$E_{i+1}[c_i] = E_i,$$

which implies that

(3-5) $$\gcd(s_i, e(E_i/E_{i+1})) = 1.$$

We write

(3-6) $$\beta = \beta(\xi) = \beta_0 + \cdots + \beta_{t+1}.$$

Note that $v_E(\beta) = -r$, the level of ξ. When $r = 0$, i.e., ξ is tamely ramified, then all ξ_i, with $i = 0, \dots, t+1$, are trivial, and we take $\beta = 0$.

Proposition 3.1. *For $i = 0, \dots, t$:*

(i) $E_i = F[\beta_{t+1} + \cdots + \beta_i]$.

(ii) *Each $\beta_i \in E_i$ is minimal over E_{i+1}.*

Proof. We know (i) is true because we have a decreasing sequence (3-2) of field extensions, while (ii) follows from (3-4) and Corollary 2.2. □

3B. ***Construction of simple characters.*** We identify E as an n-dimensional vector space V and hence obtain an embedding $E \hookrightarrow A$. We define an hereditary order

$$\mathfrak{A} = \{X \in A : X\mathfrak{p}_E^k \subseteq \mathfrak{p}_E^k \text{ for all } k \in \mathbb{Z}\}$$

and its j-th radical

$$\mathfrak{P}_{\mathfrak{A}}^j = \{X \in A : X\mathfrak{p}_E^k \subseteq \mathfrak{p}_E^{k+j} \text{ for all } k \in \mathbb{Z}\}, \text{ for } j \in \mathbb{Z}.$$

We also extend the definition such that $\mathfrak{P}_{\mathfrak{A}}^r = \mathfrak{P}_{\mathfrak{A}}^{\lceil r \rceil}$ and $\mathfrak{P}_{\mathfrak{A}}^{r+} = \mathfrak{P}_{\mathfrak{A}}^{\lceil r \rceil +}$ for $r \in \mathbb{R}$. We then define the following subgroups in G,

$$U_{\mathfrak{A}} = U_{\mathfrak{A}}^0 = \mathfrak{A}^\times \quad \text{and} \quad U_{\mathfrak{A}}^j = 1 + \mathfrak{P}_{\mathfrak{A}}^j, \qquad \text{for all } j \in \mathbb{Z}_{>0},$$

and define $U_{\mathfrak{A}}^r$ and $U_{\mathfrak{A}}^{r+}$ similarly for $r \in \mathbb{R}_{\geq 0}$. Finally, we define $\mathfrak{B}_i$, $\mathfrak{P}_{\mathfrak{B}_i}^r$, and $\mathfrak{P}_{\mathfrak{B}_i}^{r+}$ as the centralizers of E_i in $\mathfrak{A}$, $\mathfrak{P}_{\mathfrak{A}}^r$, and $\mathfrak{P}_{\mathfrak{A}}^{r+}$ respectively, and define the subgroups $U_{\mathfrak{B}_i}$, $U_{\mathfrak{B}_i}^r$, and $U_{\mathfrak{B}_i}^{r+}$ in $U_{\mathfrak{A}}$ as the centralizers of $E_i^\times$ in $U_{\mathfrak{A}}$, $U_{\mathfrak{A}}^r$, and $U_{\mathfrak{A}}^{r+}$ respectively.

Given an element $\alpha \in A$, we denote a map

$$\psi_\alpha : A \to \mathbb{C}, \qquad x \mapsto \psi_F \circ \mathrm{tr}_{A/F}(\alpha(x-1)).$$

If $v = v_{\mathfrak{A}}(\alpha) < 0$, then the restriction of ψ_α on $U_{\mathfrak{A}}^{-(v/2)+}$ defines a character, which is trivial on $U_{\mathfrak{A}}^{-v+}$.

Given an admissible character ξ of $E^\times$, we recall the construction of a simple character $\theta = \theta_\xi$, in the sense of [Bushnell and Kutzko 1993, Section 3.2], on the compact subgroup

$$H^1 = H^1_\xi := U^1_{\mathfrak{B}_0} U^{r_0/2+}_{\mathfrak{B}_1} \cdots U^{r_{t-1}/2+}_{\mathfrak{B}_t} U^{r_t/2+}_{\mathfrak{B}_{t+1}}$$

(note that $\mathfrak{B}_{t+1} = \mathfrak{A}$) and whose restriction onto U^1_E coincides with $\xi|_{U^1_E}$. Like ξ, this simple character also admits a factorization

$$\theta = \theta_0\theta_1 \cdots \theta_{t+1}$$

such that

$$\theta_i|_{U^1_{\mathfrak{B}_0} U^{r_0/2+}_{\mathfrak{B}_1} \cdots U^{r_{i-1}/2+}_{\mathfrak{B}_i}} = \xi_i \circ \det{}_{B_i/E_i} \qquad \text{and} \qquad \theta_i|_{U^{r_i/2+}_{\mathfrak{B}_{i+1}} \cdots U^{r_t/2+}_{\mathfrak{B}_{t+1}}} = \psi_{\beta_i}.$$

It is well-defined since on the intersection $U^{r_i/2+}_{\mathfrak{B}_i}$ the characters are equal, by (3-3). Note that when $r = 0$, we take $\mathfrak{A} = \mathrm{M}_n(\mathfrak{o}_F)$ with $H^1 = U^1_{\mathfrak{A}}$, and θ is the trivial character of H^1.

Proposition 3.2. *The assignment* $\xi|_{U^1_E} \mapsto \theta$ *is well-defined*, *i.e.*, *it is independent of the factorization* (3-1).

Proof. The verifying arguments are quite routine, so we only provide a brief idea as follows. Before we begin, in order to reduce the load of notations, we denote the restriction of any character ϕ of some $E^\times$ to U^1_E just by ϕ instead of $\phi|_{U^1_E}$, and similarly if we replace E by other fields.

First of all, remember that the jumps $\{r_i\}$ and the intermediate subfields $\{E_i\}$ in (3-2) are uniquely determined by ξ. Suppose we have another factorization of ξ whose factors are $\{\xi'_i\}_{i=-1}^{t+1}$, then we can inductively deduce that, for $i = 0, \dots, t+1$,

$$\xi_i^{-1}\xi'_i\phi_{i-1} = \phi_i \circ N_{E_i/E_{i+1}} \tag{3-7}$$

for some characters ϕ_i of $U^1_{E_{i+1}}$, each of whose level t_i is less than $s_i = r_i/e(E/E_i)$ because of (3-5). We remark that here we take ϕ_{-1} and ϕ_{t+1} to be trivial. In the additive level, suppose that

$$\phi_i(x) = \psi_{E_{i+1}}(\gamma_i(x-1)) \quad \text{for all } x \in U^{t_i/2+}_{E_{i+1}}, \tag{3-8}$$

then (3-7) becomes, for $i = 0, \dots, t+1$,

$$\beta'_i + \gamma_{i-1} - \beta_i = \gamma_i \tag{3-9}$$

for some element $\gamma_i \in E_{i+1}$, and we take $\gamma_{-1} = \gamma_{t+1} = 0$.

Now we consider the restriction of θ to $U^{r_i/2+}_{\mathfrak{B}_{i+1}}$, on which each factor θ_j is equal to

$$\psi_{\beta_j} \text{ if } j \leq i, \quad \text{and} \quad \xi_j \circ \det_{B_j/E_j} = (\xi_j \circ N_{E_i/E_j}) \circ \det_{B_i/E_i} \text{ if } j > i.$$

Similar results apply to each factor θ'_j of θ'. We then apply (3-7) and (3-9) to obtain

$$\theta(\theta')^{-1}|_{U^{r_i/2+}_{\mathfrak{B}_{i+1}}} = \phi_i \circ \det_{B_{i+1}/E_{i+1}} \cdot \psi^{-1}_{\gamma_i},$$

which is just trivial because of (3-8). Therefore, we have $\theta = \theta'$. □

Given ξ with β as in (3-6), we associate a stratum $[\mathfrak{A}, r, 0, \beta]$, in the sense of [Bushnell and Kutzko 1993, (1.5)], to ξ, where $r = -v_E(\beta)$. Note that we have taken $\beta = 0$ when ξ is tamely ramified, in which case the associated stratum is null $[\mathrm{M}_n(\mathfrak{o}_F), 0, 0, 0]$.

Proposition 3.3. (*i*) *If the level r of ξ is positive, then the stratum $[\mathfrak{A}, r, 0, \beta]$ is simple, with a sequence of approximation strata $[\mathfrak{A}, r, r_i, \gamma_i]$, where*

$$\gamma_i = \sum_{j=i}^{t+1} \beta_j,$$

and each with a derived stratum $[\mathfrak{B}_i, r_i, r_i - 1, c_i]$, all in the sense of [Bushnell and Kutzko 1993, (2.4.2)].

(*ii*) $\theta \in \mathcal{C}(\mathfrak{A}, 0, \beta)$, *the set of simple characters in the sense of [Bushnell and Kutzko 1993,* (3.2.3)].

Proof. We first prove (i), which is to show that the sequence $[\mathfrak{A}, n, r_i, \gamma_i]$ satisfies the conditions in [Bushnell and Kutzko 1993, (2.4.1)]. In fact, many of the arguments are routine, mostly following from constructions. One technical part is [Bushnell and Kutzko 1993, (2.4.1)(iv)], where we have to show that

$$k_0(\gamma_i, \mathfrak{A}) = -r_i \quad \text{for each } i = 0, \dots, t.$$

We first decompose β term by term as $\sum_{i=1}^{r} a_i$ with $a_i \in C_E$ and $v_E(a_i) = -i$. Hence $\beta_i = \sum_{j=r_{i-1}+1}^{r_i} a_j$ and $a_{r_i} = c_i$. We now apply induction, assuming that $k_0(\gamma_{i+1}, \mathfrak{A}) = -r_{i+1}$, which is less than $v_E(c_i)$. By the second case of Proposition 2.3, we have $k_0(c_i + \gamma_{i+1}, \mathfrak{A}) = -r_i$. Now notice that each a_k with $k = r_{i-1}+1, \dots, r_i$, lies in $E_i = F\bigl[\gamma_{i+1} + \sum_{l=k}^{r_i} a_l\bigr]$. In particular $\gamma_{i+1} + \sum_{l=r_{i-1}+1}^{r_i} a_l = \gamma_i$, and so by the first case of Proposition 2.3, $k_0(\gamma_i, \mathfrak{A}) = -r_i$.

Once (i) is established, (ii) can be checked just by the definition in [Bushnell and Kutzko 1993, (3.2.3)]. The case for θ being trivial (when ξ is tamely ramified) is just by convention, so we move on to the positive level case. By induction along the approximation sequence in (i), it suffices to show that for each $i = 0, \dots, t+1$, we have

$$\Theta_i := \theta_i \cdots \theta_{t+1} \in \mathcal{C}(\mathfrak{A}, r_{i-1}/2+, \gamma_i).$$

Now the subgroup $H^{r_{i-1}/2+}$ is $U_{\mathfrak{B}_i}^{r_{i-1}/2+} \cdots U_{\mathfrak{B}_{t+1}}^{r_t/2+}$. For each $j \geq i$, the factor $\theta_j|_{H^{r_{i-1}/2+}}$ is equal to

$$\xi_j \circ \det_{B_j/E_j}|_{U_{\mathfrak{B}_i}^{r_{i-1}/2+} \cdots U_{\mathfrak{B}_j}^{r_{j-1}/2+}} \cdot \psi_{\beta_j}|_{U_{\mathfrak{B}_{j+1}}^{r_j/2+} \cdots U_{\mathfrak{B}_{t+1}}^{r_t/2+}}.$$

We hence check the conditions in [Bushnell and Kutzko 1993, (3.2.3)] for the character Θ_i.

(a) We have $\Theta_i|_{U_{\mathfrak{B}_i}^{r_{i-1}/2+}} = (\xi_i(\xi_{i+1} \circ N_{E_i/E_{i+1}}) \cdots (\xi_{t+1} \circ N_{E_i/F})) \circ \det_{B_i/E_i}$.

(b) The compact subgroup $H^{r_{i-1}/2+}$ is clearly normalized by

$$\mathfrak{K}(\mathfrak{B}_i) = \{x \in B_i^\times : x^{-1}\mathfrak{B}_i x = \mathfrak{B}_i\},$$

and so are the characters $\xi_j \circ \det_{B_j/E_j}$ and ψ_{β_j} for $j \geq i$.

(c) We have $H^{r_i/2+} = U_{\mathfrak{B}_{i+1}}^{r_i/2+} \cdots U_{\mathfrak{B}_{t+1}}^{r_t/2+}$, on which the factor θ_i is equal to ψ_{β_j}, and $\Theta_{i+1} \in \mathcal{C}(\mathfrak{A}, r_i/2+, \gamma_{i+1})$ by induction assumption. $\square$

We show very briefly that our θ agrees with the one in [Bushnell and Henniart 2005, Section 2.3]. We will not go into detail as it incurs heavy definitions and notations from transfers [Bushnell and Kutzko 1993], endo-classes [Bushnell and Henniart 1996, Section 7], and tame liftings [Bushnell and Henniart 1996, Section 9], but only refer to the references as given.

Suppose that ξ is an admissible character of $E^\times$, with an associated stratum $[\mathfrak{A}, r, 0, \beta]$ as constructed in the previous section, and $\theta \in \mathcal{C}(\mathfrak{A}, 0, \beta)$ is a simple character of a compact subgroup H^1 of G. Recall from [Bushnell and Henniart 2005, Section 2.3] that, if we write $\xi|_{U_E^1} = \Xi_0 \circ N_{E/E_0}$ for some character Ξ_0 of $U_{E_0}^1$, and denote the endo-classes of θ and Ξ_0 by $\mathcal{E}_F(\theta)$ and $\mathcal{E}_{E_0}(\Xi_0)$ respectively, then a specific simple character θ_0 is characterized by the condition that

$$\mathcal{E}_{E_0}(\Xi_0) \text{ is a } E_0/F\text{-lift of } \mathcal{E}_F(\theta_0).$$

Our simple character θ constructed above satisfies this condition, because of the relation

$$\theta|_{U_E^1} = \xi|_{U_E^1}$$

which is exactly the relation in [Bushnell and Henniart 1996, (9.2)] that defines the tame lifting of a simple character. Hence $\theta_0 = \theta$.

We continue to follow [Bushnell and Henniart 2005, Section 2.3]. On the compact mod-center subgroup

$$\mathbf{J} = \mathbf{J}_\xi := E^\times U_{\mathfrak{B}_0}^1 U_{\mathfrak{B}_1}^{r_0/2} \cdots U_{\mathfrak{B}_t}^{r_{t-1}/2} U_{\mathfrak{A}}^{r_t/2},$$

we define an extended maximal simple type $\Lambda = \Lambda_\xi$, which is a finite-dimensional irreducible representation, depending on ξ and whose restriction onto H^1 is a direct sum of $\theta = \theta_\xi$. We then put $\pi = \pi_\xi := \mathrm{cInd}_{\mathbf{J}}^G \Lambda$.

Proposition 3.4. (i) *The representation π is irreducible, supercuspidal, and essentially tame. Moreover, any such representation arises from the above construction.*

(ii) *We have $f(\pi) = n/e(E_0/F)$.*

(iii) *The isomorphism class of π depends only on the orbit of $(E/F, \xi)$ under Galois conjugation.* (*This orbit is called an admissible pair in* [*Bushnell and Henniart 2005*].)

Proof. All statements can be deduced from Proposition 2.3 and Theorem 2.3 of [Bushnell and Henniart 2005]. □

4. A special choice of ordered basis

We continue from the last section. In Section 4A, we provide the desired properties of our F-basis of E for constructing our explicit Whittaker data. In Section 4B, we construct such an ordered basis, and express the element β and the compact subgroup H^1 with respect to this basis. Finally, in Section 4C we provide a factorization of $H^1 \cap N$, where N is the maximal unipotent subgroup defined by this ordered basis, according to the one defined by θ.

4A. ***An inductively subordinate condition.*** We first consider a general situation. Suppose that E_0/F is a finite extension with a tower of subextensions $\{E_i\}_{i=0}^{t+1}$ similar to (3-2), except that we do not require E_0/F to be tamely ramified. Let V be an E_0-vector space with an $\mathfrak{o}_{E_0}$-lattice chain $\mathcal{L}$ in V. Suppose that, for each $i = 0, \dots, t+1$, there is an ordered E_i-decomposition of V as

$$V = \bigoplus_{j \in I^i} W_j^i \tag{4-1}$$

for an ordered set I^i of indices, such that the following conditions hold:

(I) There is a decomposition of ordered sets $I^{i+1} = \bigsqcup_{j \in I^i} I_j^{i+1}$ such that

$$W_j^i = \bigoplus_{k \in I_j^{i+1}} W_k^{i+1}. \tag{4-2}$$

(II) For each $r \in \mathbb{Z}$, we have

$$\mathcal{L}(r) = \bigoplus_{j \in I^i} \mathcal{L}(r) \cap W_j^i,$$

which means that the decomposition (4-1) conforms with $\mathcal{L}$ over E_i, in the sense of [Bushnell and Kutzko 1993, (7.1.1(i))] or [Bushnell and Henniart 1996, (10.5)].

We call the refining decompositions in (4-1) *inductively subordinate* to $\mathcal{L}$. Let $\mathcal{F}_{E_i}$ be the associated E_i-flag. By regarding all flags $\mathcal{F}_{E_i}$ as F-flags by restriction of scalars, we have a successive refinement of F-flags

$$\mathcal{F}_{E_0} \subset \cdots \subset \mathcal{F}_{E_t} \subseteq \mathcal{F}, \tag{4-3}$$

which gives rise to a tower of unipotent subgroups

$$N_{\mathcal{F}_{E_0}} \subset \cdots \subset N_{\mathcal{F}_{E_t}} \subseteq N = N_{\mathcal{F}}.$$

4B. ***The ordered basis.*** Let ξ be an admissible character of $E^\times$ over F, with $\{\beta_i\}_{i=0}^t$ the set of approximation elements as in Section 3A. For future computation, we define an extra element β_{-1} to be a primitive root of unity in $\boldsymbol{\mu}_E$ when $E \neq E_0$, and put $\beta_{-1} = 0$ when $E = E_0$.

We choose the following ordered F-basis $\mathfrak{b} = \{x_j\}_{j=1}^n$ of $V = E$,

$$x_1 = 1 \quad \text{and} \quad x_{j+1} = \beta_i x_j \tag{4-4}$$

for $i = -1, \ldots, t$, if j is a multiple of $[E_{i+1} : F]$ but not a multiple of $[E_i : F]$. Note that:

- $v_E(x_j) \leq v_E(x_k)$ for all $j > k$, with equality if and only if $E \neq E_0$ and k is a multiple of $[E_0 : F]$ with $j = k + 1$, in which case $x_{k+1} = \beta_{-1} x_k$.
- If $\beta \notin F$, then $v_E(x_j) < 0$ for all $j > 1$.

We can also define this basis inductively as follows. Let

$$\mathfrak{b}^{-1} = \{1, \beta_{-1}, \beta_{-1}^2, \ldots, \beta_{-1}^{[E:E_0]-1}\}.$$

This is an ordered cyclic E_0-basis of E. For $i = 0, \ldots, t+1$, we define

$$\mathfrak{b}_{E_{i+1}}(\beta_i) = \{1, \beta_i, \beta_i^2, \ldots, \beta_i^{[E_i:E_{i+1}]-1}\}$$

and, if $\mathfrak{b}^{i-1}$ is ordered as $\{z_1, \ldots, z_{[E:E_i]}\}$, define

$$\mathfrak{b}_j^i = z_j \beta_i^{(j-1)([E_i:E_{i+1}]-1)} \mathfrak{b}_{E_{i+1}}(\beta_i) \tag{4-5}$$

for $j = 1, \ldots, [E : E_i]$. Each $\mathfrak{b}_j^i$ is an E_{i+1}-basis of an E_i-vector space, and

$$\mathfrak{b}^i = \mathfrak{b}_1^i \sqcup \cdots \sqcup \mathfrak{b}_{[E:E_i]}^i,$$

is an E_{i+1}-basis of E. Finally, we have $\mathfrak{b} = \mathfrak{b}^t$.

We hence define, for $i = 0, \ldots, t+1$ and $j = 1, \ldots, [E : E_t]$,

$$W_j^i = \operatorname{span}_F \mathfrak{b}_j^{i-1}, \tag{4-6}$$

which is an E_i-vector space of dimension 1. Condition (I) in the previous section is clearly satisfied.

Proposition 4.1. (i) *Let $\mathcal{L}(r) = \mathfrak{p}_E^r$ in $V = E$; then the refining decompositions defined by* (4-6) *are inductively subordinate to $\mathcal{L}$.*

(ii) *In particular, the condition of Proposition 2.4 is satisfied by the basis $\mathfrak{b}$.*

Proof. For (i), we claim that condition (II) is satisfied by (4-6). If we show that, for each $r \in \mathbb{Z}$,

$$\mathfrak{p}_E^r = \bigoplus_{j=1}^{[E:E_i]} \mathfrak{p}_{E_i}^{a_i(j,r)} z_j \quad (\text{here } \mathfrak{b}^{i-1} = \{z_1, \dots, z_{[E:E_i]}\}) \tag{4-7}$$

for suitable integers $a_i(j,r)$, for $i = 0, \dots, t+1$, then the claim is implied by the last i. The minimality of β_{-1} implies that (4-7) holds for $i = 0$, by the remark after Proposition 2.4 (indeed in this case $a_0(j,r) = r$ for all j). If (4-7) holds for some i, then by the same remark again (substituting E, F and α by E_i, E_{i+1} and β_i, respectively) and the minimality in Proposition 3.1, we show that, for $s \in \mathbb{Z}$,

$$\mathfrak{p}_{E_i}^{s-v} = \beta_i^{-(j-1)([E_i:E_{i+1}]-1)} \mathfrak{p}_{E_i}^s = \bigoplus_{k=0}^{[E_i:E_{i+1}]-1} \mathfrak{p}_{E_{i+1}}^{b_{i+1}^j(k,s)} \beta_i^k$$

for suitable integers $b_{i+1}^j(s,k)$ and where $v = (j-1)([E_i : E_{i+1}] - 1) v_{E_i}(\beta_i)$. We obtain

$$\mathfrak{p}_E^r = \bigoplus_{j,k} \mathfrak{p}_{E_{i+1}}^{b_{i+1}^j(k, a_i(r,j))} w_{j,k}$$

with $w_{j,k} = z_j \beta_i^{(j-1)([E_i:E_{i+1}]-1)+k}$ forming the basis $\mathfrak{b}^i$ by (4-5). Hence (4-7) holds for $i+1$.

For (ii), it is enough to show that (*) is satisfied. We again apply induction on i. Condition (*) is satisfied by the cyclic basis $\mathfrak{b}^{-1}$, and suppose it is satisfied by $\mathfrak{b}^{i-1}$, so that if $u = \sum_{z_j \in \mathfrak{b}^{i-1}} a_j z_j$ for $a_j \in E_i$ then $v_E(a_j z_j) \geq v_E(u)$. Write

$$a_j = \sum_{k=0}^{[E_i:E_{i+1}]-1} b_{j,k} y_{j,k}$$

for some $b_{j,k} \in E_{i+1}$ and $y_{j,k} = \beta_i^{(j-1)([E_i:E_{i+1}]-1)+k}$, then by applying (*) on the E_{i+1}-basis $\{y_{j,k}\}_k$ for E_i we obtain $v_{E_i}(b_{j,k} y_{j,k}) \geq v_{E_i}(a_j)$. Now

$$u = \sum_{w_{j,k} \in \mathfrak{b}^i} b_{j,k} w_{j,k},$$

with $w_{j,k}$ as above forming the basis $\mathfrak{b}^i$, and $v_E(b_{j,k} w_{j,k}) \geq v_E(a_j z_j) \geq v_E(u)$. Hence (*) is satisfied by $\mathfrak{b}^i$. □

We now provide some properties of the matrix presentations of the elements β and β_{-1}, and also the compact subgroup H^1, with respect to the ordered basis $\mathfrak{b}$.

Proposition 4.2. (i) *The matrix presentation $\beta_{j,k}$, where $j, k = 1, \dots, n$, of β with respect to $\mathfrak{b}$ takes the form*

$$\beta_{j,k} \in \begin{cases} 1+\mathfrak{p}_F & \text{if } j-k=1 \text{ and } k \text{ is not a multiple of } [E_0:F], \\ \mathfrak{p}_F & \text{if } j-k>1 \text{ or if } j-k=1 \text{ and } k \text{ is a multiple of } [E_0:F]. \end{cases}$$

(ii) *When $E \neq E_0$, the matrix presentation $(\beta_{-1})_{j,k}$ of β_{-1} with respect to $\mathfrak{b}$ takes the form*

$$(\beta_{-1})_{j,k} \in \begin{cases} 1+\mathfrak{p}_F & \text{if } j-k=1 \text{ and } k \text{ is a multiple of } [E_0:F], \\ \mathfrak{p}_F & \text{if } j-k>1 \text{ or if } j-k=1 \text{ and } k \text{ is not a multiple of } [E_0:F]. \end{cases}$$

(iii) *In the matrix presentation of H^1 with respect to $\mathfrak{b}$, the entries in the strictly upper triangle belong to $\mathfrak{o}_F$.*

(We remark that, in cases (i) and (ii), we do not need to study the (j,k)-entries with $j \leq k$.)

Proof. To prove (i), for each $k = 1, \dots, n$, we will determine where the entries of the k-th column of β belong with respect to $\mathfrak{b}$. Let $i = i(k)$ be the index such that k is a multiple of $[E_{i+1}:F]$ but not a multiple of $[E_i:F]$; then $x_{k+1} = \beta_i x_k$ by construction. If $E = E_0$, we want to show that the product

$$\beta x_k = \sum_{i=0}^{t+1} \beta_i x_k$$

lies in

$$\bigoplus_{l=1}^{k} F x_l + x_{k+1} + \bigoplus_{l=k+1}^{n} \mathfrak{p}_F x_l.$$

First of all, we have

$$\beta_{t+1} x_k + \cdots + \beta_{i+1} x_k \in \bigoplus_{l=1}^{k} F x_l, \tag{4-8}$$

because if we write $x_k = \beta_t^{m_t} \cdots \beta_{j+1}^{m_{j+1}} \beta_j^{m_j}$ for some integers $m_{t+1}, \dots, m_j > 0$, then $i \geq j-1$, and we see that $E_{i+1} x_k \in \bigoplus_{l=1}^{k} F x_l$; in particular (4-8) holds. We then show that

$$\beta_{i-1} x_k + \cdots + \beta_0 x_k \in \bigoplus_{l=k+1}^{n} \mathfrak{p}_F x_l. \tag{4-9}$$

For all $j < i$,

$$v_E(\beta_j x_k) > v_E(\beta_i x_k) = v_E(x_{k+1}).$$

By Proposition 2.4, the coefficients of x_l, for $l \geq k+1$, of all $\beta_{i-1}x_k, \dots, \beta_0 x_k$, lie in $\mathfrak{p}_F$, and (4-9) holds. When $E \neq E_0$, the proof is similar, except that when k is a multiple of $[E_0 : F]$, we have $x_{k+1} = \beta_{-1}x_k$, and so

$$\beta x_k \in \bigoplus_{l=1}^{k} F x_l \oplus \bigoplus_{l=k+1}^{n} \mathfrak{p}_F x_l.$$

For (ii), the arguments are similar to above. If k is a multiple of $[E_0 : F]$, then $\beta_{-1}x_k = x_{k+1}$. Otherwise, we have $v_E(\beta_{-1}x_k) = v(x_k) > v(x_{k+1})$, and so $\beta_{-1}x_k \in \bigoplus_{l=1}^{k} F x_l \oplus \bigoplus_{l=k+1}^{n} \mathfrak{p}_F x_l$.

For (iii), notice that $\{x_j/\varpi_F^{\lfloor v_E(x_j)/e \rfloor}\}_{j=1}^n$ is an $\mathfrak{o}_F$-basis for the lattice chain $\mathcal{L}(r) = \mathfrak{p}_E^r$ in $V = E$. With this basis, the entries of $U_{\mathfrak{B}_i}$ for all i, hence those of H^1, belong to $\mathfrak{o}_F$. If we use the basis $\mathfrak{b} = \{x_j\}_{j=1}^n$ instead, then the (j,k)-entry is multiplied by $\varpi_F^{\lfloor v_E(x_j)/e \rfloor - \lfloor v_E(x_k)/e \rfloor}$. In the upper triangle consisting of (j,k)-entries where $j < k$, we have $v_E(x_j) > v_E(x_k)$, and so the (j,k)-entry with respect to $\mathfrak{b}$ is still in $\mathfrak{o}_F$. □

Corollary 4.3. $\psi_{\beta+\beta_{-1}}|_{N \cap H^1}(x) = \psi_F\big(\sum_{j=1}^{n-1} x_{j,j+1}\big)$, *where* $x_{j,k}$ *is the* (j,k)*-entry of the matrix presentation of* $x \in A$ *with respect to* $\mathfrak{b}$.

Proof. With respect to the basis $\mathfrak{b}$, it is easy to see that the entries of $\beta + \beta_{-1}$ in the lower sub-diagonal belong to $1 + \mathfrak{p}_F$, and those underneath belong to $\mathfrak{p}_F$. Also, N is defined by this ordered basis, and the entries of $N \cap H^1$ in the strictly upper triangle belong to $\mathfrak{o}_F$. Since ψ_F is trivial on $\mathfrak{p}_F$ but not on $\mathfrak{o}_F$, we have the desired result. □

As a remark, for a fixed β, there are other bases which also serve our purpose. For instance, we can take the basis constructed in the same way as $\mathfrak{b}$ but with all β_i replaced by their first terms c_i. One can prove, almost verbatim, that β takes the same form as in the proposition. Also, another factorization of ξ yields another set of elements $\{\beta_i\}$, and so another β, but the matrix presentation of that β takes the same form.

We end this subsection with a few examples.

Example 4.4. Let $[\mathfrak{A}, r, 0, \beta]$ be a minimal stratum, and let m be the degree of $E_0 = F[\beta]$ over F. For a positive integer d, let E be the unramified extension of E_0 of degree d, and take a primitive root of unity in $\boldsymbol{\mu}_E$. We construct the basis

$$\mathfrak{b} = \{1, \beta, \dots, \beta^{m-1}, \zeta\beta^{m-1}, \zeta\beta^m, \dots, \zeta\beta^{2m-2}, \zeta^2\beta^{2m-2} \dots, \zeta^{d-1}\beta^{d(m-1)}\}.$$

We consider the matrix of β relative to $\mathfrak{b}$. On the j-th column where j is not a multiple of m, the entries are all 0 except the $(j+1)$-th entry, which is 1. For $k = 1, \dots, d$, if $\beta^m = \phi(\beta)$ for some F-polynomial ϕ of degree $m-1$, then we have $\beta \cdot x_{km} = \beta \cdot \zeta^{k-1}\beta^{k(m-1)} = \zeta^{k-1}\beta^{(k-1)(m-1)}\phi(\beta)$, which lies in the F-span of x_j for $(k-1)m < j \le km$. Therefore, on the (km)-th column, the j-th entries for $j > km$ are all 0.

We then consider the matrix of ζ relative to $\mathfrak{b}$. On the j-th column where j is a multiple of m, the entries are all 0 except the $(j{+}1)$-th entry, which is 1. For $s = 1, \dots, m-1$, if $\beta^{-s} = \phi_s(\beta)$ for some F-polynomial ϕ_s of degree $m-1$, then minimality implies that its coefficients must lie in $\mathfrak{p}_F$. Now if $j = (k-1)m + l$, where $k = 1, \dots, d$ and $l = 1, \dots, m-1$, then

$$\zeta \cdot x_j = \zeta \cdot \zeta^{k-1} \beta^{(k-1)(m-1)+(l-1)} = \zeta^k \beta^{k(m-1)} \phi_{m-l}(\beta),$$

which lies in the $\mathfrak{p}_F$-span of x_i where $km < i \le (k+1)m$, in particular $i > j+1$. Therefore, on the j-th column, the i-th entry for $i \le j+1$ is 0, and lies in $\mathfrak{p}_F$ for $i > j+1$.

We hence see that the element $\zeta + \beta$ has the desired form as in Corollary 4.3. Note that we did not assume that E_0/F is tamely ramified in the minimal case: all we need to know is that the valuation of β is negative.

Example 4.5. We provide one more example for small n which exhibits the situation when multiple jumps are present. Let's take $n = 4$. As the minimal case is covered in the previous example, we assume that our simple stratum $[\mathfrak{A}, r, 0, \beta]$ has two jumps. Consider a tower of the form $E \supset K \supset F$ where $[E : K] = [K : F] = 2$. For simplicity, we only consider two extreme cases.

(i) Suppose that E/F is totally ramified, and so $p \ne 2$. We fix a uniformizer ϖ_F and choose ϖ_K and ϖ_E such that $\varpi_K^2 = a\varpi_F$ and $\varpi_E^2 = b\varpi_K$ for some $a, b \in \boldsymbol{\mu}_F$. Consider the element

$$\beta = \varpi_F^{-r} + \varpi_K^{-s} + \varpi_E^{-t},$$

where $4r > 2s > t > 0$ and both s and t are odd. The basis constructed by $\beta_0 = \varpi_E^{-t}$ and $\beta_1 = \varpi_K^{-s}$ is

$$\{1,\ \varpi_K^{-s},\ \varpi_E^{-t}\varpi_K^{-s},\ \varpi_E^{-t}\varpi_K^{-2s}\}.$$

The matrix of β takes the form

$$\begin{bmatrix} \varpi_F^{-r} & (a\varpi_F)^{-s} & b^{-t}(a\varpi_F)^{-(t+s)/2} & * \\ 1 & \varpi_F^{-r} & 0 & * \\ 0 & 1 & \varpi_F^{-r} & * \\ (a\varpi_F)^s & 0 & 1 & * \end{bmatrix}$$

(the last column is unimportant for our purposes).

(ii) Suppose now that E/F is unramified. Let $K = F[\zeta]$ and $E = K[\eta]$, where $\zeta, \eta \in \boldsymbol{\mu}_E$ and satisfy the equations $\zeta^2 = a\zeta + b$ and $\eta^2 = (c\zeta + d)\eta + (e\zeta + f)$ with all $a, \dots, f \in \mathfrak{o}_F$. Write $\varpi = \varpi_F$ and consider for example the element

$$\beta = \varpi^{-r} + \zeta\varpi^{-s} + \eta\varpi^{-t},$$

where $r > s > t > 0$. With the basis

$$\{1,\ \zeta\varpi^{-s},\ \zeta\eta\varpi^{-(t+s)},\ \zeta^2\eta\varpi^{-(t+2s)}\},$$

the matrix of β takes the form

$$\begin{bmatrix} \varpi^{-r} & b\varpi^{-2s} & b(ae+f)\varpi^{-(s+2t)} & * \\ 1 & \varpi^{-r}+a\varpi^{-s} & (ae+af+be)\varpi^{-2t} & * \\ (-a/b)\varpi^{s} & 1 & \varpi^{-r}+bc\varpi^{-t} & * \\ (1/b)\varpi^{2s} & 0 & 1+(ac+d)\varpi^{s-t} & * \end{bmatrix}.$$

4C. ***A factorization for maximal unipotent subgroups.*** We first work on a general situation. Let E/F be a finite extension and V be a finite-dimensional E-vector space, also regarded as an F-vector space. Denote $A = \mathrm{End}_F(V)$ and let $\mathfrak{A}$ be an hereditary $\mathfrak{o}_F$-order in A defined by an $\mathfrak{o}_E$-lattice chain $\mathcal{L}$ in V. Let B be the centralizer of E in A, and denote $\mathfrak{B} = \mathfrak{A} \cap B$.

We suppose that V admits an ordered decomposition $\bigoplus_j W_j$ into a direct sum of E-subspaces, and $\mathcal{F}_E$ is the associated E-flag. We further suppose that each W_j, viewed as an F-vector space, admits an ordered decomposition into a direct sum $\bigoplus_i W_j^i$ of F-subspaces, altogether forming an F-flag $\mathcal{F}$ in V. Let $M_{\mathcal{F}}$ be the subgroup of G stabilizing all W_j^i, let $N_{\mathcal{F}}$ be the unipotent subgroup in G defined by the flag $\mathcal{F}$, and let $N_{\mathcal{F}}^-$ be its opposite. Also, define $M_{\mathcal{F}_E}$, $N_{\mathcal{F}_E}$, and $N_{\mathcal{F}_E}^-$ similarly, using the flag $\mathcal{F}_E$.

Lastly, we suppose that the refining decompositions above are both subordinate to $\mathcal{L}$, in the sense of the conditions in Section 4A.

Proposition 4.6. (i) *For each positive integer k, the subgroups $U_{\mathfrak{A}}^k$ and $U_{\mathfrak{B}}^k$ admit an Iwahori decomposition*

$$U_{\mathfrak{A}}^k = (U_{\mathfrak{A}}^k \cap N_{\mathcal{F}_E})(U_{\mathfrak{A}}^k \cap M_{\mathcal{F}_E})(U_{\mathfrak{A}}^k \cap N_{\mathcal{F}_E}^-)$$

and similarly for $U_{\mathfrak{B}}^k$.

(ii) *For positive integers $k_1 < k_2$, we have*

$$(U_{\mathfrak{B}}^{k_1} U_{\mathfrak{A}}^{k_2}) \cap N_{\mathcal{F}} = (U_{\mathfrak{B}}^{k_1} \cap N_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}}).$$

Proof. Part (i) is given by [Bushnell and Henniart 1996, (10.4)] and noting that, if the decomposition conforms with $\mathcal{L}$ over E, it also conforms with $\mathcal{L}$ over F. For part (ii), we first prove the "maximal" case, i.e., when V is 1-dimensional over E, in which case $N_{\mathcal{F}_E}$ is trivial, and the right-hand side is $U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}}$. This is equal to the left-hand side by [Blondel and Stevens 2009, Lemma A.5 Appendix]. If E is not maximal in A, we can follow the idea in [Blondel and Stevens 2009, Corollary A.6 Appendix]. By (i), we have

$$(U_{\mathfrak{B}}^{k_1} U_{\mathfrak{A}}^{k_2}) \cap P_{\mathcal{F}_E} = (U_{\mathfrak{B}}^{k_1} \cap N_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}_E})(U_{\mathfrak{B}}^{k_1} \cap M_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap M_{\mathcal{F}_E}),$$

and note that $N_{\mathcal{F}} \subset P_{\mathcal{F}_E}$, so

$$U_{\mathfrak{B}}^{k_1} U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}} = (U_{\mathfrak{B}}^{k_1} \cap N_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}_E})((U_{\mathfrak{B}}^{k_1} \cap M_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap M_{\mathcal{F}_E}) \cap N_{\mathcal{F}}).$$

The last bracket lies in the maximal case for the Levi subgroup $M_{\mathcal{F}_E}$, and so is equal to $U_{\mathfrak{A}}^{k_2} \cap M_{\mathcal{F}_E} \cap N_{\mathcal{F}}$. Since $(U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}_E})(U_{\mathfrak{A}}^{k_2} \cap M_{\mathcal{F}_E} \cap N_{\mathcal{F}}) = U_{\mathfrak{A}}^{k_2} \cap N_{\mathcal{F}}$, we have the desired result. $\square$

We further assume that E/F is tamely ramified, with a tower of intermediate subfields (3-2) coming from an admissible character. Let H^1 be the subgroup defined in Section 3B, and $N_{\mathcal{F}_{E_i}}$, for $i = 0, \ldots, t+1$, be the maximal E_i-flags defined by the ordered decompositions in (4-6), which are inductively subordinate by Proposition 4.1.

Corollary 4.7. *Given a sequence of flags as in* (4-3), *then*

$$H^1 \cap N_{\mathcal{F}} = (U_{\mathfrak{B}_0}^1 \cap N_{\mathcal{F}_{E_0}}) \cdots (U_{\mathfrak{B}_t}^{r_{t-1}/2+} \cap N_{\mathcal{F}_{E_t}})(U_{\mathfrak{A}}^{r_t/2+} \cap N_{\mathcal{F}}).$$

Proof. By the inductive subordination, we apply Proposition 4.6 (ii) inductively. First regard $U_{\mathfrak{B}_0}^1 \cdots U_{\mathfrak{B}_t}^{r_{t-1}/2+}$ as a subgroup of $U_{\mathfrak{B}_t}^1$ and so

$$H^1 \cap N_{\mathcal{F}} = (U_{\mathfrak{B}_0}^1 \cdots U_{\mathfrak{B}_t}^{r_{t-1}/2+} \cap N_{\mathcal{F}_{E_t}})(U_{\mathfrak{A}}^{r_t/2+} \cap N_{\mathcal{F}}).$$

We can therefore apply induction on $U_{\mathfrak{B}_0}^1 \cdots U_{\mathfrak{B}_t}^{r_{t-1}/2+} \cap N_{\mathcal{F}_{E_t}}$ and obtain the desired result. $\square$

Proposition 4.8. $\theta|_{N \cap H^1} = \psi_\beta|_{N \cap H^1}$.

Proof. For each i, we already know that θ_i is equal to ψ_{β_i} on

$$(U_{\mathfrak{B}_{i+1}}^{r_i/2+} \cap N_{\mathcal{F}_{E_t}}) \cdots (U_{\mathfrak{A}}^{r_t/2+} \cap N_{\mathcal{F}})$$

from its construction. It suffices to show that θ_i on

$$(U_{\mathfrak{B}_0}^1 \cap N_{\mathcal{F}_{E_0}}) \cdots (U_{\mathfrak{B}_i}^{r_{i-1}/2+} \cap N_{\mathcal{F}_{E_i}}),$$

which is $\xi_i \circ \det_{B_i/E_i}$, is also equal to ψ_{β_i}. Indeed, on all $N_{\mathcal{F}_{E_j}}$ for $j \leq i$, the character $\det_{B_i/E_i}$ is trivial, while ψ_{β_i} is also trivial since $\beta_i \in M_{\mathcal{F}_{E_i}} \subset M_{\mathcal{F}_{E_j}}$. $\square$

5. The main result

Let π be an essentially tame supercuspidal representation compactly induced by an extended maximal type $(\mathbf{J}, \Lambda)$ which contains a simple character $\theta \in \mathcal{C}(\mathfrak{A}, 0, \beta)$ associated to an admissible character ξ, and $N = N_{\mathcal{F}}$ be the maximal unipotent subgroup defined by the F-flag

$$\mathcal{F} = \{V_j\}_{j=1}^n, \qquad \text{where } V_j = \bigoplus_{k=1}^{j} F x_k,$$

and $\{x_j\}_{j=1}^n$ is the ordered basis constructed in (4-4).

Let:

- α_0 be an element in $\mathrm{Mat}_n(F)$ whose matrix representation $(\alpha_0)_{j,k}$ with respect to $\mathfrak{b}$ is 1 if $j-k=1$ but is 0 if k is a multiple of $[E_0:F]$, and is 0 if $j-k>1$ (and can be anything if $j\le k$).
- α_{-1} be an element in $\mathrm{Mat}_n(F)$ whose matrix representation $(\alpha_{-1})_{j,k}$ with respect to $\mathfrak{b}$ is 0 if $j-k=1$ but is 1 if k is a multiple of $[E_0:F]$, and is 0 if $j-k>1$ (and can be anything if $j\le k$).
- $\alpha=\alpha_{-1}+\alpha_0$.

Hence, with the notation defined in [Paskunas and Stevens 2008], we have

$$\alpha\in X_{\mathcal{F}}^{+}:=\{x\in A: xV_i\subset V_{i+1}\text{ and }xV_i\not\subset V_i\text{ for all }i\},$$

and so by [Paskunas and Stevens 2008, Lemma 1.2] ψ_α defines a non-degenerate character of N. (Note that, in contrast, ψ_β may not extend to a character of the whole N.)

Theorem 5.1. $\mathrm{Hom}_{N\cap\mathbf{J}}(\psi_\alpha,\Lambda)\neq 0$.

Proof. We show that the condition at the beginning of [Paskunas and Stevens 2008, Section 4.2] is satisfied, then our result is implied by Corollary 4.13 of the same work. Hence it suffices to show that $(\mathcal{F},\psi_{\alpha_0},\alpha_{-1})$ satisfies the conditions (i)-(iv) in Theorem 3.3 of the same work.

(i) This condition is just $\mathcal{F}_{E_0}\subset\mathcal{F}$ in our notation, which is true by construction.

(ii) In Proposition 4.8, we showed that $\theta|_{N\cap H^1}=\psi_\beta|_{N\cap H^1}$. Now with the matrix representation of β and elements in H^1 in Proposition 4.2(i), we know that $\psi_\beta|_{N\cap H^1}=\psi_{\alpha_0}|_{N\cap H^1}$.

(iii) If $E=E_0$, then $N_{\mathcal{F}_{E_0}}$ is trivial and the result is clearly satisfied. If $E\neq E_0$, then ψ_{α_0} on $N_{\mathcal{F}_{E_0}}$ is trivial since the matrix entry $(\alpha_0)_{k,k+1}$ with respect to $\mathfrak{b}$ is 0 when k is a multiple of $[E_0:F]$.

(iv) The maximal unipotent subgroup $N_{\mathcal{F}_{E_0}}\cap U_{\mathfrak{B}_0}/U^1_{\mathfrak{B}_0}$ of $U_{\mathfrak{B}_0}/U^1_{\mathfrak{B}_0}$ is defined by the cyclic basis

$$\{\bar{1},\bar{\beta}_{-1},\bar{\beta}_{-1}^2,\dots,\bar{\beta}_{-1}^{[E:E_0]-1}\},$$

where each $\bar{x}$ is $x+\mathfrak{P}_{\mathfrak{B}_0}\in U_{\mathfrak{B}_0}/U^1_{\mathfrak{B}_0}$ for $x\in U_{\mathfrak{B}_0}$. The character $\psi_{\beta_{-1}}$ clearly defines a nondegenerate character, by arguments similar to [Bushnell and Henniart 1998, 2.1]. This character is equal to $\psi_{\alpha_{-1}}$ by Proposition 4.2 (ii). □

5A. ***A formula for the Artin conductor.*** Suppose that (π_1,π_2) is a pair of essentially tame supercuspidal representations of $\mathrm{GL}_{n_i}(F)$, for $i=1,2$, such that their extended maximal simple types contain the same simple character, hence the same associated simple or null stratum. Recall the conductor of the epsilon factor for the pair (π_1, π_2) computed in [Paskunas and Stevens 2008; Kim 2014], which is

$$\mathfrak{f}(\pi_1 \times \pi_2) = \mathfrak{f}(\tau_1 \times \tau_2) + \frac{n_1 n_2}{e(E_0/F)[E_0 : F]} v_E(x_{[E_0:F]}/x_1),$$

where τ_i is a supercuspidal representation of $\mathrm{GL}_{n_i/[E_0:F]}(E_0)$, compactly induced from the "level-zero" component of the extended maximal simple type of π_i (see [Paskunas and Stevens 2008, Section 7]).

Let's compare this result with the calculation in [Bushnell et al. 1998, Theorem 6.5]. The conductor formula implies that

$$\mathfrak{f}(\pi_1 \times \pi_2) = \mathfrak{f}(\tau_1 \times \tau_2) + n_1 n_2 \frac{\mathfrak{c}(\beta)}{[E_0 : F]^2}.$$

Here $\mathfrak{c}(\beta)$ is a certain kind of "discriminant", whose value can be inductively computed from [Bushnell and Henniart 2003, 3.1] as

$$\frac{\mathfrak{c}(\beta_i)}{[E_i : F]^2} = \frac{\mathfrak{c}(\beta_{i+1})}{[E_{i+1} : F]^2} + \frac{k_0(\beta_i, \mathfrak{A})}{e(E_0/F)} \left(\frac{1}{[E_{i+1} : F]} - \frac{1}{[E_i : F]} \right).$$

We can rewrite it into a direct formula as

$$\mathfrak{c}(\beta) = \frac{[E_0 : F]}{e(E_0/F)} \sum_{i=0}^{t} ([E : E_{i+1}] - [E : E_i]) k_0(\beta_i, \mathfrak{A}).$$

In the essentially tame case, our result implies that

$$v_E(x_{[E_0:F]}/x_1) = \sum_{i=0}^{t} ([E : E_{i+1}] - [E : E_i]) v_E(\beta_i). \tag{5-1}$$

We can use Proposition 2.3 to see that our result (5-1) agrees with the calculation in the above literatures.

Acknowledgements

This article was written when the author was visiting the University of British Columbia. He would like to thank Julia Gordon for her support. During the review process, he was supported by the Max Planck Institute for Mathematics. He would also like to thank Vincent Sécherre for providing a note on tame ramifications and his comments on the first draft of the article, Shaun Stevens for his careful reading and useful remarks on the first draft, and Ju-Lee Kim for providing an electronic version of her paper [Kim 2014]. Finally, he would like to thank the two referees for their careful readings and suggestions.

References

[Blondel and Stevens 2009] C. Blondel and S. Stevens, "Genericity of supercuspidal representations of p-adic Sp_4", *Compos. Math.* **145**:1 (2009), 213–246. MR Zbl

[Bushnell and Henniart 1996] C. J. Bushnell and G. Henniart, "Local tame lifting for $\mathrm{GL}(N)$, I: Simple characters", *Inst. Hautes Études Sci. Publ. Math.* **83** (1996), 105–233. MR Zbl

[Bushnell and Henniart 1998] C. J. Bushnell and G. Henniart, "Supercuspidal representations of GL_n: explicit Whittaker functions", *J. Algebra* **209**:1 (1998), 270–287. MR Zbl

[Bushnell and Henniart 2003] C. J. Bushnell and G. Henniart, "Local tame lifting for $\mathrm{GL}(n)$, IV: Simple characters and base change", *Proc. London Math. Soc.* (3) **87**:2 (2003), 337–362. MR Zbl

[Bushnell and Henniart 2005] C. J. Bushnell and G. Henniart, "The essentially tame local Langlands correspondence, I", *J. Amer. Math. Soc.* **18**:3 (2005), 685–710. MR Zbl

[Bushnell and Kutzko 1993] C. J. Bushnell and P. C. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Annals of Mathematics Studies **129**, Princeton University Press, 1993. MR Zbl

[Bushnell et al. 1998] C. J. Bushnell, G. M. Henniart, and P. C. Kutzko, "Local Rankin–Selberg convolutions for GL_n: explicit conductor formula", *J. Amer. Math. Soc.* **11**:3 (1998), 703–730. MR Zbl

[Gel'fand and Kajdan 1975] I. M. Gel'fand and D. A. Kajdan, "Representations of the group $\mathrm{GL}(n, K)$ where K is a local field", pp. 95–118 in *Lie groups and their representations* (Budapest, 1971), edited by I. M. Gel'fand, Halsted, New York, 1975. MR Zbl

[Howe 1977] R. E. Howe, "Tamely ramified supercuspidal representations of Gl_n", *Pacific J. Math.* **73**:2 (1977), 437–460. MR Zbl

[Kim 2014] J.-L. Kim, "An inductive formula for ϵ-factors", pp. 243–260 in *Automorphic forms and related geometry: assessing the legacy of I. I. Piatetski-Shapiro*, edited by J. W. Cogdell et al., Contemp. Math. **614**, Amer. Math. Soc., Providence, RI, 2014. MR Zbl

[Kutzko 1977] P. C. Kutzko, "Mackey's theorem for nonunitary representations", *Proc. Amer. Math. Soc.* **64**:1 (1977), 173–175. MR Zbl

[Kutzko and Manderscheid 1988] P. Kutzko and D. Manderscheid, "On intertwining operators for $\mathrm{GL}_N(F)$, F a non-Archimedean local field", *Duke Math. J.* **57**:1 (1988), 275–293. MR Zbl

[Moy 1986] A. Moy, "Local constants and the tame Langlands correspondence", *Amer. J. Math.* **108**:4 (1986), 863–930. MR Zbl

[Paskunas and Stevens 2008] V. Paskunas and S. Stevens, "On the realization of maximal simple types and epsilon factors of pairs", *Amer. J. Math.* **130**:5 (2008), 1211–1261. MR Zbl

[Reimann 1991] H. Reimann, "Representations of tamely ramified p-adic division and matrix algebras", *J. Number Theory* **38**:1 (1991), 58–105. MR Zbl

Received October 8, 2017. Revised October 16, 2018.

GEO KAM-FAI TAM
INSTITUTE FOR MATHEMATICS, ASTROPHYSICS AND PARTICLE PHYSICS
RADBOUD UNIVERSITY
NIJMEGEN
NETHERLANDS
k.tam@math.ru.nl

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.639

K-THEORY OF AFFINE ACTIONS

JAMES WALDRON

For a Lie group G and a vector bundle E we study those actions of the Lie group TG on E for which the action map $TG \times E \to E$ is a morphism of vector bundles, and call those *affine actions*. We prove that the category $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ of such actions over a fixed G-manifold X is equivalent to a certain slice category $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$. We show that there is a monadic adjunction relating $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ to $\mathrm{Vect}_G(X)$, and the right adjoint of this adjunction induces an isomorphism of Grothendieck groups $K^{\mathrm{aff}}_{TG}(X) \cong KO_G(X)$. Complexification produces analogous results involving $T_{\mathbb{C}}G$ and $K_G(X)$.

1. Introduction

1A. Let G be a Lie group. The tangent bundle TG carries two structures: it is a vector bundle over G, and a Lie group, with multiplication given by the derivative of the multiplication of G. These structures are compatible, in the sense that the multiplication $TG \times TG \to TG$ is a morphism of vector bundles, so that TG is a *group object* in the category of vector bundles. It is therefore natural to study actions of TG on vector bundles, such that the action map $TG \times E \to E$ is a morphism of vector bundles (see Definition 3.1 below). We refer to such actions as *affine actions*, as each element of TG necessarily acts by an affine linear transformation between fibres of E (see Remark 3.3 below). A basic example of an affine action is the following:

Example 1.1 (tangent bundles). For any action $t : G \times X \to X$ of a Lie group G on a smooth manifold X, the derivative defines an affine action $t_* : TG \times TX \to TX$ of TG on TX. Note that restricting the action t_* to G defines the natural action $G \times TX \to TX$ of G on TX, whilst restricting t_* to the Lie algebra $\mathfrak{g} = T_eG$ allows one to define a map $\mathfrak{g} \to \Gamma(TX)$ which is exactly the infinitesimal action associated to t. These maps are compatible in the sense that $\mathfrak{g} \to \Gamma(TX)$ is G-equivariant.

Example 1.1 suggests the question of whether the action $t_* : TG \times TX \to TX$, or more generally any affine action $\mu : TG \times E \to E$, can be reconstructed from its restrictions to G and $\mathfrak{g}$.

MSC2010: 19E99.
Keywords: vector bundles, equivariant K-theory, differential geometry.

One motivation for studying equivariant vector bundles is their use in defining the (real) equivariant K-theory $KO_G(X)$ of a G-manifold X (see [Segal 1968]). Recall that (at least if G and X are compact) $KO_G(X)$ is the Grothendieck group of the commutative monoid of isomorphism classes of G-equivariant real vector bundles over X. A natural question to ask is whether one can emulate this construction in the case of affine actions to define an abelian group $K^{\mathrm{aff}}_{TG}(X)$. If so, how is this group related to $KO_G(X)$?

A different motivation for studying affine actions comes from the theory of *Lie algebroids*. Recall that a Lie algebroid $A \to X$ is a vector bundle over X equipped with an $\mathbb{R}$-linear Lie bracket on $\Gamma(A)$ and a map $A \to TX$, such that a certain Leibniz rule is satisfied. (See [Mackenzie 2005] for more details.) There exists a notion of equivariant Lie algebroid (called a *Harish-Chandra Lie algebroid* in [Beilinson and Bernstein 1993]) which involves both a G-action $G \times A \to A$ and a linear map $\mathfrak{g} \to \Gamma(A)$ satisfying certain conditions. Variants of this notion have appeared in [Alekseev and Meinrenken 2009; Bruzzo et al. 2009; Marrero et al. 2012; Ginzburg 1999]. It was shown in [Marrero et al. 2012] that equivariant Lie algebroids give rise to examples of affine actions (see Example 4.5 below). One motivation for our results is therefore to generalise the notion of affine action and study this concept at the level of vector bundles.

1B. *Main results.* Throughout the paper G is a real Lie group and X is a G-manifold. See Section 2 for our notation and conventions and Section 3A for the definition of affine actions and their morphisms. We use $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ (respectively $\mathrm{Vect}_G(X)$) to denote the category of affine actions (respectively the category of real G-equivariant vector bundles) over X. We use $\mathfrak{g}_X$ to denote the G-equivariant vector bundle associated to the adjoint representation of G. See Section 3C for the definition of the slice category $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$.

Theorem A. *The following three categories are isomorphic*:

(1) *The category* $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ *of affine actions of* TG *over* X.

(2) *The category of pairs* $(E, \rho_{\mathfrak{g}})$, *defined as follows*:

- *The objects are pairs* $(E, \rho_{\mathfrak{g}})$, *where* E *is a* G*-equivariant vector bundle over* X *and* $\rho_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(E)$ *is a* G*-equivariant linear map.*
- *The morphisms* $(E, \rho_{\mathfrak{g}}) \to (E', \rho_{\mathfrak{g}})$ *are morphisms* $\psi : E \to E'$ *of* G*-equivariant vector bundles over* X *such that* $\Gamma(\psi) \circ \rho_{\mathfrak{g}} = \rho'_{\mathfrak{g}}$.
- *Composition is given by composition of morphisms of vector bundles over* X.

(3) *The slice category* $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$.

Our second main result compares affine actions to equivariant vector bundles. There is a canonical forgetful functor

$$\mathrm{U} : \mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \to \mathrm{Vect}_G(X).$$

Via the isomorphism $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \cong \mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ of Theorem A, the functor U is equal to the canonical forgetful functor $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X) \to \mathrm{Vect}_G(X)$ which maps an object $\mathfrak{g}_X \xrightarrow{\phi} E$ to E. We also define a pair of functors $\mathrm{F}, \sigma : \mathrm{Vect}_G(X) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$. In terms of $\mathfrak{g}\backslash \mathrm{Vect}_G(X)$, they are defined on objects by $\sigma : E \mapsto (E, 0)$ and $\mathrm{F} : E \mapsto (\mathfrak{g}_X \oplus E, i_{\mathfrak{g}_X})$. See Section 5A for the precise definitions.

Theorem B. *The following statements hold*:

(1) F *is left adjoint to* U.

(2) *The adjunction* $\mathrm{F} \dashv \mathrm{U}$ *is monadic.*

(3) σ *is the unique section of* U.

Our third main result concerns the Grothendieck group of $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$. We denote by

$$KO_G(-) : G\text{-Man} \to \mathrm{Ab}$$

the functor from the category G-Man of G-manifolds to the category Ab of abelian groups, which maps a G-manifold X to the Grothendieck group of G-equivariant real vector bundles over X. (This agrees with real G-equivariant topological K-theory as defined in [Segal 1968] if both G and X are compact.) Although the category $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ is not additive we show in Section 6B that it does have finite products. This allows us to define the Grothendieck group $K^{\mathrm{aff}}_{TG}(X)$ (see Section 2C and Definition 6.13). This construction extends to a contravariant functor

$$K^{\mathrm{aff}}_{TG}(-) : G\text{-Man} \to \mathrm{Ab}.$$

Our third main result shows that $K^{\mathrm{aff}}_{TG}(-)$ agrees with $KO_G(-)$:

Theorem C. *For X a G-manifold the functor* U *induces a group isomorphism*

$$K(\mathrm{U}) : K^{\mathrm{aff}}_{TG}(X) \to KO_G(X).$$

Its inverse is

$$K(\sigma) : KO_G(X) \to K^{\mathrm{aff}}_{TG}(X).$$

These isomorphisms are natural in X, and thus define an isomorphism of functors

$$K^{\mathrm{aff}}_{TG}(-) \xrightarrow{\cong} KO_G(-).$$

1C. ***The complex case.*** It is possible to reformulate the notion of affine action in the complex setting by replacing TG by the complexified tangent bundle $T_{\mathbb{C}}G$ and considering actions $T_{\mathbb{C}}G \times E \to E$ on complex vector bundles E. The analogues of Theorems A, B and C hold with essentially the same proofs. See Section 7 for the precise statements.

1D. ***The proofs.*** Theorem A is proved using the facts that $TG \cong G \times \mathfrak{g}$ as a vector bundle, and $TG \cong \mathfrak{g} \rtimes G$ as a Lie group, where the semidirect product is defined via the adjoint representation of G. This allows one to decompose an action μ of TG into an action μ_G of G and a linear map $\rho_{\mathfrak{g}}$ with domain $\mathfrak{g}$. Parts of the proof are similar to that of Theorem 3.5 in [Marrero et al. 2012], which deals with the particular case where E is a Lie algebroid, and constructs, at the level of objects, one direction of the isomorphism of Theorem A.

Using the isomorphism $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \cong \mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ of Theorem A, Theorems B and C are proved using the following two category-theoretic lemmas regarding over slice categories:

Lemma D. *Let $\mathscr{C}$ be a category, m an object in $\mathscr{C}$, and $\mathcal{U} : m\backslash\mathscr{C} \to \mathscr{C}$ the standard forgetful functor. If the coproduct $m \amalg a$ exists in $\mathscr{C}$ for all objects a in $\mathscr{C}$ then the functor $\mathcal{F} : a \mapsto (m \amalg a, i_m)$ is left adjoint to $\mathcal{U}$, and this adjunction is monadic.*

Lemma E. *Let $\mathscr{C}$ be an additive category and m an object in $\mathscr{C}$. Let $\mathcal{U} : m\backslash\mathscr{C} \to \mathscr{C}$ be the standard forgetful functor and $\mathcal{S} : \mathscr{C} \to m\backslash\mathscr{C}$ the section $a \mapsto (a, 0)$. Then the group homomorphism*

$$K(\mathcal{U}) : K(m\backslash\mathscr{C}) \to K(\mathscr{C})$$

is an isomorphism. Its inverse is

$$K(\mathcal{S}) : K(\mathscr{C}) \to K(m\backslash\mathscr{C}).$$

Here, $K(\mathcal{U})$ denotes the homomorphism of Grothendieck groups associated to the product preserving functor $\mathcal{U}$, and similarly for $K(\mathcal{S})$, see Section 2C. We expect that Lemma D is well known to experts (in particular it is stated without proof in [nLab 2009–]), but we are unaware of a complete reference and so have provided a proof.

1E. ***Outline.*** In Section 2 we fix notation and conventions. In Section 3 we define affine actions and morphisms between them. The main result of Section 3 is Theorem A, the proof of which is broken into Lemma 3.7 and Propositions 3.8, 3.10 and 3.11. In Section 4 we describe a number of examples of affine actions, and describe the category $\mathrm{Vect}_{TG}(X)$ for certain classes of groups G and G-manifolds X. In Section 5 we define several functors between $\mathrm{Vect}_{TG}(X)$ and $\mathrm{Vect}_G(X)$. We then prove Lemma D, which is used to prove Theorem B. In Section 6 we define pullback functors for affine actions. We then prove that the category $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ has finite products and use this to define the abelian group $K^{\mathrm{aff}}_{TG}(X)$ and the functor $K^{\mathrm{aff}}_{TG}(-) : G\text{-Man} \to \mathrm{Ab}$. We then prove Lemma E from which we prove Theorem C.

2. Preliminaries

2A. ***Notation and conventions.*** By "manifold" we shall always mean smooth finite-dimensional real manifold. Maps between manifolds are assumed to be smooth. Unless stated otherwise, by "vector bundle" we mean finite-dimensional real vector bundle. We will usually denote manifolds by X or Y, vector bundles by E or F, vector fields by v or w, and sections of vector bundles by ξ or ν. For a vector bundle E over X we use $\pi_E : E \to X$ to denote the bundle projection, and $0_E : X \to E$ to denote the zero section. We allow morphisms of vector bundles over different bases. If E and F are vector bundles over X, then by "morphism of vector bundles over X" we mean a vector bundle morphism $\phi : E \to F$ which satisfies $\pi_F \circ \phi = \pi_E$. If ϕ is a morphism of this type, then $\Gamma(\phi) : \Gamma(E) \to (F)$ denotes the associated linear map.

For $E \to X$ a vector bundle, $x \in X$ and $\xi \in \Gamma(E)$, we use ξ_x to denote ξ evaluated at x. We use e_x to denote an element of E_x and v_x to denote an element of $T_x X$. We denote the zero element of E_x by 0_x. For a morphism of vector bundles $\phi : E \to F$ over X and $x \in X$ we denote by $\phi_x : E_x \to F_x$ the restriction of ϕ.

If $E \to X$ and $F \to Y$ are vector bundles then $E \times F$ is a vector bundle over $X \times Y$ in a natural way, with fibre over (x, y) canonically isomorphic to $E_x \oplus F_y$.

We denote hom-sets in a category $\mathscr{C}$ by $\mathrm{Hom}_{\mathscr{C}}(-,-)$. We reserve the unadorned Hom for morphisms of real vector spaces (i.e., linear maps). If G is a Lie group then we use Hom_G for morphisms of representations of G (i.e., G-equivariant linear maps). We denote identity morphisms by 1_a or id_a. If $\mathscr{C}$ is an additive category then we denote any zero-morphism by 0. If $a \times b$ is a product in a category $\mathscr{C}$ then we denote the associated projections by $\mathrm{pr}_a : a \times b \to a$ and $\mathrm{pr}_b : a \times b \to b$. Similarly, if $a \amalg b$ is a coproduct then we denote the associated inclusions by $i_a : a \to a \amalg b$ and $i_b : b \to a \amalg b$. If $a \amalg b$ is a coproduct and $f : a \to c$ and $g : b \to c$ are morphisms, then we denote by $(f, g) : a \amalg b \to c$ the associated morphism. We use a similar notation for morphisms into products.

If G is a Lie group then by a "G-manifold" we shall mean a smooth manifold X equipped with a smooth left action $t : G \times X \to X$. We denote by $t_g = t(g, -) : X \to X$ the diffeomorphism associated to $g \in G$, which we also denote by $x \mapsto g \cdot x$. If X and Y are G-manifolds then by a G-map $f : X \to Y$ we mean a G-equivariant smooth map.

For a G-manifold X, by "G-equivariant vector bundle", or just "G-vector bundle", we mean a vector bundle $E \to X$ equipped with a left action $\mu_G : G \times E \to E$ which covers $t : G \times X \to X$ and which is fibrewise linear. A morphism $\phi : E \to F$

of G-equivariant vector bundles is a morphism of vector bundles over X which is also G-equivariant. We denote by $\mathrm{Vect}_G(X)$ (respectively $\mathrm{Vect}_G^{\mathbb{C}}(X)$) the category of real (respectively complex) G-equivariant vector bundles over X. (See [Segal 1968] for generalities on equivariant vector bundles.) $\mathrm{Vect}_G(X)$ and $\mathrm{Vect}_G^{\mathbb{C}}(X)$ are both additive categories. In particular, for E and F G-vector bundles the G-vector bundle $E \oplus F$ is both the product and the coproduct of E and F.

We denote the category of finite-dimensional real (respectively complex) representations of G by $\mathrm{Rep}(G)$ (respectively $\mathrm{Rep}_{\mathbb{C}}(G)$). If V is such a representation then we denote by V_X the associated G-vector bundle, i.e., the trivial bundle $X \times V$ with G-action $g \cdot (x, v) = (g \cdot x, g \cdot v)$. If E is a G-equivariant vector bundle then we consider $\Gamma(E)$ as a G-representation with G-action $(g \cdot \xi)_x = g \cdot \xi_{g^{-1} \cdot x}$. If V and W are representations of G, and E is a G-vector bundle, then there exist bijections

$$\mathrm{Hom}_{\mathrm{Vect}_G(X)}(V_X, W_X) \xrightarrow{\cong} C^\infty(X, \mathrm{Hom}(V, W))^G \qquad \phi \mapsto (x \mapsto \phi_x), \tag{1}$$

$$\mathrm{Hom}_{\mathrm{Vect}_G(X)}(V_X, E) \xrightarrow{\cong} \mathrm{Hom}_G(V, \Gamma(E)) \qquad \phi \mapsto (v \mapsto (x \mapsto \phi(x, v))). \tag{2}$$

2B. ***Tangent groups.*** Let G be a Lie group and $\mathfrak{g} = T_e G$ its Lie algebra. We will usually denote elements of G by g or h, and elements of $\mathfrak{g}$ by α or β. The tangent bundle TG of G carries a natural Lie group structure with multiplication defined by the composite map $TG \times TG \xrightarrow{\cong} T(G \times G) \xrightarrow{m_*} TG$, where $m : G \times G \to G$ is the multiplication of G. We will denote the multiplication in TG by $\bullet$. If $v_g \in T_g G$ and $w_h \in T_h G$ then it follows from the chain rule that

$$v_g \bullet w_h = (L_g)_* w_h + (R_h)_* v_g.$$

In particular, $0_g \bullet (-) = (L_g)_*$, $(-) \bullet 0_h = (R_h)_*$, $0_g \bullet 0_h = 0_{gh}$, and if $\alpha, \beta \in \mathfrak{g}$ then $\alpha \bullet \beta = \alpha + \beta$. If one considers $\mathfrak{g}$ as an abelian Lie group upon which G acts via the adjoint representation then the associated semidirect product $\mathfrak{g} \rtimes G$ has multiplication

$$(\alpha, g) \bullet (\beta, h) = (\alpha + \mathrm{Ad}_g \beta, gh).$$

There is a Lie group isomorphism $\mathfrak{g} \rtimes G \xrightarrow{\cong} TG$ given by $(\alpha, g) \mapsto (R_g)_* \alpha$. Under this isomorphism, the inclusion $\mathfrak{g} \hookrightarrow \mathfrak{g} \rtimes G$ corresponds to $\mathfrak{g} = T_e G \hookrightarrow TG$, the inclusion $G \hookrightarrow \mathfrak{g} \rtimes G$ corresponds to $0_{TG} : G \to TG$, and the projection $\mathfrak{g} \rtimes G \to G$ corresponds to $\pi_{TG} : TG \to G$.

2C. ***Grothendieck groups.*** Let $\mathscr{C}$ be an essentially small category (one where the collection of isomorphism classes of objects in $\mathscr{C}$ is a set) with finite products. The set $\mathscr{C}/\cong$ of isomorphism classes of objects in $\mathscr{C}$ forms a commutative monoid under the operation $[E]+[E'] \equiv [E \times E']$. We denote by $K(\mathscr{C})$ the associated abelian group defined by the Grothendieck construction. If $\mathscr{C}$ is an additive category then this agrees with the standard notion of the "split" Grothendieck group of $\mathscr{C}$, i.e., the abelian group generated by isomorphism classes of objects and relations $[A]+[B]=[A \oplus B]$.

If $\mathcal{F}:\mathscr{C} \to \mathscr{C}'$ is a product preserving functor between categories satisfying the above assumptions then there is a group homomorphism $K(\mathcal{F}):K(\mathscr{C}) \to K(\mathscr{C}')$ defined by $[E]-[E'] \mapsto [\mathcal{F}(E)]-[\mathcal{F}(E')]$. The group homomorphism $K(\mathcal{F})$ depends functorially on $\mathcal{F}$. If $\mathcal{F},\mathcal{F}':\mathscr{C} \to \mathscr{C}'$ are naturally isomorphic functors then $K(\mathcal{F})=K(\mathcal{F}')$.

We write $KO_G(X)$ for $K(\mathrm{Vect}_G(X))$ and $K_G(X)$ for $K(\mathrm{Vect}_G^{\mathbb{C}}(X))$. This agrees with G-equivariant topological K-theory as defined by Segal [1968] if both G and X are compact.

3. Affine actions

Throughout the paper, G denotes a Lie group and X denotes a G-manifold with action $t: G \times X \to X$.

3A. *Affine actions.*

Definition 3.1. An *affine action* of TG on a real vector bundle $E \to X$ is a left action $\mu: TG \times E \to E$ of the Lie group TG on the total space of E such that μ is a vector bundle morphism covering $t: G \times X \to X$.

Example 3.2. The derivative of t defines an affine action $t_*: TG \times TX \to TX$ of TG on TX.

Remark 3.3. The condition that μ is a morphism of vector bundles covering t is equivalent to requiring that:

(i) The following diagram commutes:

$$\begin{array}{ccc} TG \times E & \xrightarrow{\mu} & E \\ \downarrow & & \downarrow \\ G \times X & \xrightarrow[t]{} & X \end{array} \tag{3}$$

(ii) For all $(g,x) \in G \times X$ the restriction of μ to the fibre $T_gG \oplus E_x$ over (g,x) is a linear map $T_gG \oplus E_x \to E_{g\cdot x}$.

The second of these conditions implies that $v_g \in T_gG$ acts on $e_x \in E_x$ by the composite map $e_x \mapsto (v_g, e_x) \mapsto \mu(v_g, e_x)$. This is an affine linear map from E_x to $E_{g\cdot x}$, which justifies the name *affine action*.

Definition 3.4. A *morphism* from an affine action $\mu : TG \times E \to E$ to an affine action $\mu' : TG \times E' \to E'$ is a morphism $\psi : E \to E'$ of vector bundles over X which is TG equivariant, i.e., the following diagrams commute:

$$\begin{tikzcd} E \arrow[r, "\psi"] \arrow[dr, "\pi_E"'] & E' \arrow[d, "\pi_{E'}"] \\ & X \end{tikzcd} \qquad \begin{tikzcd} TG \times E \arrow[r, "\mu"] \arrow[d, "\mathrm{id}\times\psi"'] & E \arrow[d, "\psi"] \\ TG \times E' \arrow[r, "\mu'"'] & E' \end{tikzcd}$$

Definition 3.5. Affine actions of TG over X form a category $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$.

Remark 3.6. If one considers TG as a group object in the category of vector bundles, then affine actions coincide with the notion of actions of group objects — see, for example [MacLane 1971].

3B. *The structure of affine actions.* Lemma 3.7 and Propositions 3.8, 3.10 and 3.11 below describe how an affine action $TG \times E \to E$ can be decomposed into a G-action $G \times E \to E$ and a linear map $\mathfrak{g} \to \Gamma(E)$, and how morphisms between affine actions can be described in terms of this decomposition. These propositions will be used in Section 3C to prove our first main result, Theorem A. The first proposition shows that an affine action can be recovered from its restriction to G and $\mathfrak{g}$, as motivated by Example 1.1 in the introduction.

Lemma 3.7. *If $\mu : TG \times E \to E$ is an affine action on a vector bundle $E \to X$ then*

$$\mu(v_g, e_x) = \mu(0_g, e_x) + \mu((R_{g^{-1}})_* v_g, 0_{g\cdot x}).$$

Proof. Using the fact that μ is fibrewise linear we have

$$\mu(v_g, e_x) = \mu(v_g + 0_g, e_x + 0_x) = \mu((0_g, e_x) + (v_g, 0_x)) = \mu(0_g, e_x) + \mu(v_g, 0_x).$$

Using the fact that μ is a left action of TG we have

$$\begin{aligned} \mu(v_g, 0_x) &= \mu(v_g \bullet 0_{g^{-1}} \bullet 0_g, 0_x) = \mu((R_{g^{-1}})_* v_g, \mu(0_g, 0_x)) \\ &= \mu((R_{g^{-1}})_* v_g, 0_{g\cdot x}). \end{aligned}$$ □

Proposition 3.8. *Let $E \to X$ be a vector bundle. There is a bijection*

(4) $$\{\textit{affine actions } \mu : TG \times E \to E\} \xrightarrow{\cong} \{\textit{pairs } (\mu_G, \rho_{\mathfrak{g}}) \textit{ satisfying } (\star), (\star\star)\},$$

where:

($\star$) *$\mu_G : G \times E \to E$ is a left action of G on E making E into a G-equivariant vector bundle over X.*

($\star\star$) *$\rho_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(E)$ is a G-equivariant linear map, where the G-action on $\Gamma(E)$ is induced from μ_G.*

The bijection (4) *maps an affine action* μ *to the pair* $(\mu_G, \rho_{\mathfrak{g}})$ *defined by*

$$\mu_G(g, e_x) = \mu(0_g, e_x), \tag{5}$$

$$\rho_{\mathfrak{g}}(\alpha) = (x \mapsto \mu(\alpha, 0_x)). \tag{6}$$

The inverse of (4) *maps a pair* $(\mu_G, \rho_{\mathfrak{g}})$ *satisfying* $(\star)$, $(\star\star)$ *to the affine action* μ *defined by*

$$\mu(v_g, e_x) = \mu_G(g, e_x) + \rho_{\mathfrak{g}}((R_{g^{-1}})_* v_g)_{g\cdot x} \tag{7}$$

for $v_g \in T_g G$ *and* $e_x \in E_x$.

Note that the addition on the right-hand side of (7) is defined as the assumptions on μ_G and $\rho_{\mathfrak{g}}$ imply that both terms are elements of $E_{g\cdot x}$.

Proof. We first show that if $\mu : TG \times E \to E$ is an affine action then the pair $(\mu_G, \rho_{\mathfrak{g}})$ defined by (5) and (6) satisfies $(\star)$ and $(\star\star)$, which shows that the map (4) is well defined. That μ_G is a left action of G follows from the facts that μ is equal to the composite map

$$G \times E \xrightarrow{0_{TG}\times \mathrm{id}_E} TG \times E \xrightarrow{\mu} E$$

and that $0_{TG} : G \to TG$ is a Lie group homomorphism. The fact that μ is a vector bundle morphism covering $t : G \times X \to X$ implies that for fixed $g \in G$ the map $\mu_G(g, -) : E \to E$ is a vector bundle morphism covering $t(g, -) : X \to X$. This shows that μ_G satisfies $(\star)$. We will sometimes denote this G action by $g \cdot e_x = \mu_G(g, e_x)$.

If $\alpha \in \mathfrak{g} = T_e G$ then the commutativity of (3) and the fact that $(\alpha, 0_x) \in T_e G \oplus E_x$ implies that $\mu(\alpha, 0_x) \in E_x$, so that the map $\rho_{\mathfrak{g}}(\alpha) = (x \mapsto \mu(\alpha, 0_x))$ is a smooth section of E. The fibrewise linearity of μ implies that the map $\alpha \mapsto \rho_{\mathfrak{g}}(\alpha) \in \Gamma(E)$ is linear. If $g \in G$, $\alpha \in \mathfrak{g}$ and $x \in X$ then

$$\begin{aligned}
(\rho_{\mathfrak{g}}(\mathrm{Ad}_g \alpha))_x &= \mu(0_g \bullet \alpha \bullet 0_{g^{-1}}, 0_x) \\
&= 0_g \cdot \mu(\alpha, 0_{g^{-1}} \cdot 0_x) \\
&= 0_g \cdot \mu(\alpha, 0_{g^{-1}\cdot x}) \\
&= g \cdot \rho_{\mathfrak{g}}(\alpha)_{g^{-1}\cdot x} \\
&= (g \cdot \rho_{\mathfrak{g}}(\alpha))_x.
\end{aligned}$$

This shows that $\rho_{\mathfrak{g}}$ satisfies $(\star\star)$.

Now let $(\mu_G, \rho_{\mathfrak{g}})$ be a pair satisfying $(\star)$ and $(\star\star)$, and let us show that μ as defined in (7) defines a fibrewise linear left action of TG. That μ is fibrewise linear follows from the fibrewise linearity of μ_G and the linearity of $\rho_{\mathfrak{g}}$. Transporting the action from TG to $\mathfrak{g} \rtimes G$ via the Lie group isomorphism $v_g \mapsto ((R_{g^{-1}})_* v_g, g)$, we have

$$(\alpha, g) \cdot e_x = g \cdot e_x + \rho_{\mathfrak{g}}(\alpha)_{g \cdot x}.$$

If $(\alpha, g), (\beta, h) \in G \ltimes \mathfrak{g}$ then using $(\star)$ and $(\star\star)$ we have

$$\begin{aligned}(\alpha, g) \cdot ((\beta, h) \cdot e_x) &= (\alpha, g) \cdot (h \cdot e_x + \rho_{\mathfrak{g}}(\beta)_{h \cdot x}) \\ &= g \cdot (h \cdot e_x) + g \cdot \rho_{\mathfrak{g}}(\beta)_{h \cdot x} + \rho_{\mathfrak{g}}(\alpha)_{gh \cdot x} \\ &= gh \cdot e_x + \rho_{\mathfrak{g}}(\mathrm{Ad}_g \beta)_{gh \cdot x} + \rho_{\mathfrak{g}}(\alpha)_{gh \cdot x} \\ &= gh \cdot (e_x) + \rho_{\mathfrak{g}}(\alpha + \mathrm{Ad}_g \beta)_{gh \cdot x} \\ &= (\alpha + \mathrm{Ad}_g \beta, gh) \cdot e_x \\ &= ((\alpha, g) \bullet (\beta, h)) \cdot e_x.\end{aligned}$$

This shows that μ defines a left action of TG.

It remains to show that (4) is a bijection with inverse defined by (7). It follows from Lemma 3.7 that for an affine action μ mapped to $(\mu_G, \rho_{\mathfrak{g}})$ by (4) we have

$$\begin{aligned}\mu_G(g, e_x) + \rho_{\mathfrak{g}}((R_{g^{-1}})_* v_g)_{g \cdot x} &= \mu(0_g, e_x) + \mu((R_{g^{-1}})_* v_g, 0_{g \cdot x}) \\ &= \mu(v_g, e_x).\end{aligned}$$

Conversely, if $(\mu_G, \rho_{\mathfrak{g}})$ is a pair satisfying $(\star)$, $(\star\star)$, and μ is the affine action defined by (7), then

$$\begin{aligned}\mu(0_g, e_x) &= \mu_G(g, e_x) + \rho_{\mathfrak{g}}((R_{g^{-1}})_* 0_g)_{g \cdot x} \\ &= \mu_G(g, e_x) + 0_{g \cdot x} \\ &= \mu_G(g, e_x),\end{aligned}$$

and

$$\begin{aligned}\mu(\alpha, 0_x) &= \mu_G(e, 0_x) + \rho_{\mathfrak{g}}((R_e)_* \alpha)_{e \cdot x} \\ &= 0_x + \rho_{\mathfrak{g}}(\alpha)_x \\ &= \rho_{\mathfrak{g}}(\alpha)_x.\end{aligned}$$

□

Remark 3.9. It follows from Lemma 3.7 and the proof of Proposition 3.8 that if one transports an affine action $\mu : TG \times E \to E$ from TG to $\mathfrak{g} \rtimes G$ via the Lie group isomorphism $v_g \mapsto (g, (R_{g^{-1}})_* v_g)$ then

$$(\alpha, g) \cdot e_x = g \cdot e_x + \rho_{\mathfrak{g}}(\alpha)_{g \cdot x} \tag{8}$$

$$= \mu(0_g, e_x) + \mu(\alpha, 0_{g \cdot x}). \tag{9}$$

In particular, $g \cdot e_x = \mu(0_g, e_x)$ and $\alpha \cdot e_x = e_x + \mu(\alpha, 0_x)$, so that elements of $\mathfrak{g}$ act as fibrewise affine linear transformations. This motivates the name *affine action*.

Proposition 3.10. *Let $E \to X$ be a G-equivariant vector bundle with G-action $\mu_G : G \times E \to E$. There is a bijection*

$$
\begin{aligned}
(10)\quad & \{\rho_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(E) \mid \rho_{\mathfrak{g}} \text{ is a } G\text{-equivariant linear map}\} \\
& \xrightarrow{\cong} \{\phi : \mathfrak{g}_X \to E \mid \phi \text{ is a morphism of } G\text{-equivariant vector bundles}\}.
\end{aligned}
$$

The bijection (10) *maps $\rho_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(E)$ to the morphism $\phi : \mathfrak{g}_X \to E$ defined by*

$$\phi(x, \alpha) = \rho_{\mathfrak{g}}(\alpha)_x.$$

Proof. This follows from the bijection (2) applied to the G-representation $\mathfrak{g}$ and the G-equivariant vector bundle E. ☐

Proposition 3.11. *Let $\mu : TG \times E \to E$ and $\mu' : TG \times E' \to E'$ be affine actions corresponding to pairs $(\mu_G, \rho_{\mathfrak{g}})$, (μ_G, ϕ) and $(\mu'_G, \rho'_{\mathfrak{g}})$, (μ'_G, ϕ') under the bijections of Propositions 3.8 and 3.10. Let $\psi : E \to E'$ be a morphism of vector bundles over X. The following are equivalent*:

(1) *ψ is TG-equivariant.*

(2) *ψ is G-equivariant and $\Gamma(\psi) \circ \rho_{\mathfrak{g}} = \rho'_{\mathfrak{g}}$.*

(3) *ψ is G-equivariant and $\psi \circ \phi = \phi'$.*

Proof. ($1 \Leftrightarrow 2$) Via the isomorphism $TG \cong \mathfrak{g} \rtimes G$, a map $\psi : E \to E'$ is TG-equivariant if and only if it is $\mathfrak{g} \rtimes G$-equivariant for the action (8). If $(\alpha, g) \in \mathfrak{g} \rtimes G$ then

$$
\begin{aligned}
\psi((\alpha, g) \cdot e_x) &= \psi(g \cdot e_x + \rho_{\mathfrak{g}}(\alpha)_{g \cdot x}) \\
&= \psi(g \cdot e_x) + \psi(\rho_{\mathfrak{g}}(\alpha)_{g \cdot x})
\end{aligned}
$$

and

$$(\alpha, g) \cdot \psi(e_x) = g \cdot \psi(e_x) + \rho'_{\mathfrak{g}}(\alpha)_{g \cdot x}.$$

Therefore, ψ is TG-equivariant if and only if

$$\psi(g \cdot e_x) + \psi(\rho_{\mathfrak{g}}(\alpha)_{g \cdot x}) = g \cdot \psi(e_x) + \rho'_{\mathfrak{g}}(\alpha)_{g \cdot x}. \tag{11}$$

Setting first $\alpha = 0$, and then $g = e$ and $e_x = 0$, one sees that (11) is equivalent to the two equations

$$\psi(g \cdot e_x) = g \cdot \psi(e_x), \tag{12}$$

$$\psi(\rho_{\mathfrak{g}}(\alpha)) = \rho'_{\mathfrak{g}}(\alpha). \tag{13}$$

Equation (12) is the condition that ψ is G-equivariant, and (13) is the condition that $\Gamma(\psi) \circ \rho_{\mathfrak{g}} = \rho'_{\mathfrak{g}}$.

$(2 \Leftrightarrow 3)$ If $\alpha \in \mathfrak{g}$ and $x \in X$ then

$$(\Gamma(\psi) \circ \rho_{\mathfrak{g}}(\alpha))_x = \psi(\rho_{\mathfrak{g}}(\alpha)_x) = \psi(\phi(x, \alpha))$$

and

$$\rho'_{\mathfrak{g}}(\alpha)_x = \phi'(x, \alpha).$$

Therefore, $\Gamma(\psi) \circ \rho_{\mathfrak{g}} = \rho'_{\mathfrak{g}}$ if and only if $\psi \circ \phi = \phi'$. □

3C. ***The category of affine actions.*** Recall that for a fixed object m in a category $\mathscr{C}$, the *over-slice category* $m \backslash \mathscr{C}$ is defined as follows:

- The objects are pairs (a, ϕ), where a is an object in $\mathscr{C}$ and $\phi : m \to a$ is a morphism in $\mathscr{C}$.
- The morphisms $(a, f) \to (a', f')$ are morphisms $\chi : a \to a'$ in $\mathscr{C}$ such that $\chi \circ f = f'$:

$$\begin{tikzcd} m \arrow[r, "f"] \arrow[dr, "f'"'] & a \arrow[d, "\chi"] \\ & a' \end{tikzcd}$$

- The composition of morphisms is induced from that of $\mathscr{C}$.

There is a canonical faithful forgetful functor $m \backslash \mathscr{C} \to \mathscr{C}$ defined by $(a, f) \mapsto a$ and $(\chi : (a, f) \to (a', f')) \mapsto (\chi : a \to a')$.

Theorem A. *The following three categories are isomorphic*:

(1) *The category* $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ *of affine actions of* TG *over* X.

(2) *The category of pairs* $(E, \rho_{\mathfrak{g}})$, *defined as follows*:

- *The objects are pairs* $(E, \rho_{\mathfrak{g}})$, *where* E *is a* G*-equivariant vector bundle over* X *and* $\rho_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(E)$ *is a* G*-equivariant linear map.*
- *The morphisms* $(E, \rho_{\mathfrak{g}}) \to (E', \rho_{\mathfrak{g}})$ *are morphisms* $\psi : E \to E'$ *of* G*-equivariant vector bundles over* X *such that* $\Gamma(\psi) \circ \rho_{\mathfrak{g}} = \rho'_{\mathfrak{g}}$.
- *Composition is given by composition of morphisms of vector bundles over* X.

(3) *The slice category* $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$.

Proof. The bijections of Propositions 3.8 and 3.10 provide bijections between the classes of objects of each of these three categories. The bijections of Proposition 3.11 provide bijections between hom-sets. These bijections are functorial because the composition of morphisms in all three categories is given by the composition of morphisms of vector bundles over X. □

Remark 3.12. For the remainder of the paper we shall use the isomorphisms of Theorem A implicitly. We reserve the notation $(E, \rho_{\mathfrak{g}})$ and (E, ϕ) for objects in the second and third categories in Theorem A respectively, and refer to either as a *pair*. In particular, Theorem A shows that associated to every affine action $TG \times E \to E$ is an "underlying" G-vector bundle E, equal to the E in either of the associated pairs. Explicitly, the underlying G-action μ_G is defined by the formula (5) and is equal to the restriction of μ along the Lie group homomorphism $0_{TG} : G \to TG$.

In Section 4 we give a number of examples, some phrased in terms of affine actions and some in terms of pairs. In Sections 5 and 6 we work mostly with the category $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$, but where it is easy to do so describe the corresponding statements in terms of affine actions — see Remarks 5.5, 5.7 and 6.2, and Corollaries 6.7 and 6.9.

4. Examples and special cases

4A. *Examples.*

Example 4.1 (the tangent bundle). The composite map

$$TG \times TX \xrightarrow{\cong} T(G \times X) \xrightarrow{t_*} TX$$

is an affine action on TX. The corresponding pair $(\mu_G, i_{\mathfrak{g}})$ is given by $g \cdot w = (t_g)_*(w)$ and $i_{\mathfrak{g}}(\alpha) = \alpha^{\#}$, where $\alpha^{\sharp} \in \Gamma(TX)$ is the induced vector field associated to α, defined by

$$\alpha^{\sharp}_x = t_*(\alpha, 0_x) = \tfrac{d}{dt}|_{t=0}(\exp(t\alpha) \cdot x).$$

In terms of $\mathfrak{g} \rtimes G$ we have

$$(\alpha, g) \cdot w_x = (t_g)_* w_x + \alpha^{\#}_{g \cdot x}.$$

Example 4.2 (equivariant vector bundles). If E is a G-equivariant vector bundle with G-action $\mu_G : G \times E \to E$ then the composite map

$$TG \times E \xrightarrow{\pi_G \times \mathrm{id}} G \times E \xrightarrow{\mu_G} E$$

is an affine action. The corresponding pair is $(\mu_G, 0)$. In terms of $\mathfrak{g} \rtimes G$ we have

$$(\alpha, g) \cdot e_x = g \cdot e_x.$$

Example 4.3 (the G-bundle $\mathfrak{g}_X$). The G-vector bundle $\mathfrak{g}_X$ defines a canonical object $(\mathfrak{g}_X, \mathrm{id})$ in the category $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$. The corresponding affine action is given by

$$v_g \cdot (x, \beta) = (g \cdot x, \mathrm{Ad}_g\beta + (R_{g^{-1}})_* v_g).$$

In terms of $\mathfrak{g} \rtimes G$ we have

$$(\alpha, g) \cdot (x, \beta) = (g \cdot x, \mathrm{Ad}_g\beta + \alpha).$$

Example 4.4 (G-modules). As a generalization of Example 4.3, suppose that M is a finite-dimensional real G-representation and $\bar{\phi} : \mathfrak{g} \to M$ is a G-map. Then $\bar{\phi}$ extends to a constant morphism of G-vector bundles $\phi : \mathfrak{g}_X \to M_X$ and (M_X, ϕ) is an object in $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$. The corresponding affine action is given by

$$v_g \cdot (x, m) = (g \cdot x, g \cdot m + \bar{\phi}((R_{g^{-1}})_* v_g)).$$

In terms of $\mathfrak{g} \rtimes G$ we have

$$(\alpha, g) \cdot (x, m) = (g \cdot x, g \cdot m + \bar{\phi}(\alpha)).$$

This construction extends to a functor

$$\mathfrak{g}\backslash\mathrm{Rep}(G) \to \mathfrak{g}_X\backslash\mathrm{Vect}_G(X) \cong \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$$

which coincides with the pullback functor $\mathrm{Vect}^{\mathrm{aff}}_{TG}(*) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ associated to the G-map $X \to *$ (see Definition 6.1 in Section 6A below).

Example 4.5 (equivariant Lie algebroids). Recall that a *Lie algebroid* is a vector bundle $A \to X$ equipped with an $\mathbb{R}$-linear Lie bracket on $\Gamma(A)$ and a vector bundle morphism $\rho : A \to TX$, such that the Leibniz rule $[\xi, f\xi'] = \rho(\xi)(f)\xi' + f[\xi, \xi']$ is satisfied for all $\xi, \xi' \in \Gamma(A)$ and $f \in C^\infty(X)$. See [Mackenzie 2005] for more details.

A *G-equivariant Lie algebroid* (called a *Harish-Chandra Lie algebroid* in [Beilinson and Bernstein 1993]) is a Lie algebroid $A \to X$ equipped with an action $G \times A \to A$ of G on A by Lie algebroid automorphisms and a G-equivariant Lie algebra morphism $i_{\mathfrak{g}} : \mathfrak{g} \to \Gamma(A)$ such that $\alpha \cdot \xi = [i_{\mathfrak{g}}(\alpha), \xi]$ for all $\alpha \in \mathfrak{g}$ and $\xi \in \Gamma(A)$, where $\xi \mapsto \alpha \cdot \xi$ is the action of $\mathfrak{g}$ on $\Gamma(A)$ given by the derivative of the action of G. Variants of this notion have appeared in [Alekseev and Meinrenken 2009; Bruzzo et al. 2009; Marrero et al. 2012; Ginzburg 1999].

By forgetting the Lie brackets, every G-equivariant Lie algebroid gives rise to a pair $(A, i_{\mathfrak{g}})$ and therefore to an affine action of TG. Affine actions of this type were considered in [Marrero et al. 2012].

Example 4.6 (affine actions on $\mathbb{R}_X$). Let $\mathbb{R}_X$ be the G-vector bundle associated to the trivial representation of G. There is a bijection between affine actions with underlying G-vector bundle $\mathbb{R}_X$, and the space $C^\infty(X, \mathfrak{g}^*)^G$. This follows from the bijection (2) and the isomorphism $\mathrm{Hom}(\mathfrak{g}, \mathbb{R}) \cong \mathfrak{g}^*$ of G-representations:

$$\mathrm{Hom}_{\mathrm{Vect}_G(X)}(\mathfrak{g}_X, \mathbb{R}_X) \cong C^\infty(X, \mathrm{Hom}(\mathfrak{g}, \mathbb{R}))^G \cong C^\infty(X, \mathfrak{g}^*)^G.$$

Given a function $f \in C^\infty(X, \mathfrak{g}^*)^G$, the corresponding affine action is given by

$$v_g \cdot (x, \lambda) = (g \cdot x, \lambda + f_x((R_{g^{-1}})_* v_g)).$$

In terms of $\mathfrak{g} \rtimes G$ we have

$$(\alpha, g) \cdot (x, \lambda) = (g \cdot x, \lambda + f_x(\alpha)).$$

Example 4.7 (products). If (E, ϕ) and (E', ϕ') are objects in $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$ then so is $(E \oplus E', (\phi, \phi'))$. In fact,

$$(E \oplus E', (\phi, \phi')) = (E, \phi) \times (E', \phi'),$$

where the right-hand side is the product of (E, ϕ) and (E', ϕ') in the category $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$ (see Section 6B). The corresponding affine action is the diagonal action

$$v_g \cdot (e_x, e'_x) = (v_g \cdot e_x, v_g \cdot e'_x).$$

In terms of $\mathfrak{g} \rtimes G$ we have

$$v_g \cdot (e_x, e'_x) = (g \cdot e_x + \rho_{\mathfrak{g}}(\alpha)_x, g \cdot e'_x + \rho'_{\mathfrak{g}}(\alpha)_x).$$

4B. *Special cases.*

Example 4.8 (discrete groups). If G is a discrete group then $TG \cong G$, and affine actions are the same as equivariant vector bundles.

Example 4.9 (tori). Suppose that G is one-dimensional and abelian. In this case the adjoint representation is one-dimensional and trivial. This implies that for every G-equivariant vector bundle E there is an isomorphism $\mathrm{Hom}_G(\mathfrak{g}, \Gamma(E)) \cong \Gamma(E)^G$. Affine actions are therefore equivalent to pairs (E, ξ), where E is a G-equivariant vector bundle and ξ is a G-invariant section of E.

Example 4.10 (points). If $X = *$ is a point then $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X) = \mathfrak{g}\backslash\mathrm{Rep}(G)$.

Assuming that G is simple and compact, one can use the description of products of Example 4.7 to describe the objects in $\mathfrak{g}\backslash\mathrm{Rep}(G)$ explicitly:

Proposition 4.11. *If G is a simple compact Lie group and (V, ϕ) is an object in $\mathfrak{g}\backslash\mathrm{Rep}(G)$ then either*

(1) $(V, \phi) = (V, 0)$, *or*

(2) $(V, \phi) \cong (\mathfrak{g}, \mathrm{id})^n \times (W, 0)$,

where $n \in \mathbb{N}$ and W is a finite-dimensional G-representation with no summand isomorphic to $\mathfrak{g}$.

Remark 4.12. Note that in the setting of Proposition 4.11, if $(V, \phi) = (V, 0)$ then $(V, \phi) \cong (\mathfrak{g}, 0)^n \times (W, 0)$ for some $n \in \mathbb{Z}^{\geq 0}$ and W a finite-dimensional G-representation with no summand isomorphic to $\mathfrak{g}$.

Proof. The proof is an application of complete reducibility (G is compact) and Schur's Lemma ($\mathfrak{g}$ is irreducible as a G-representation because G is simple). Note that as G is simple and compact the adjoint representation $\mathfrak{g}$ of G is absolutely irreducible ([Onishchik 2004]). In particular, $\mathrm{End}_G(\mathfrak{g}) \cong \mathbb{R}$.

Let (V, ϕ) be an object in $\mathfrak{g}\backslash\mathrm{Rep}(G)$. As a G-representation $V \cong \mathfrak{g}^{\oplus n} \oplus W$ for some nonnegative integer n and G-representation W with no summand isomorphic to $\mathfrak{g}$. It follows from Schur's Lemma that there is an isomorphism

$$\mathbb{R}^n \xrightarrow{\cong} \mathrm{Hom}_G(\mathfrak{g}, \mathfrak{g}^{\oplus n} \oplus W), \quad \underline{\lambda} \mapsto (\alpha \mapsto (\lambda_1\alpha, \dots, \lambda_n\alpha, 0)). \tag{14}$$

Suppose that $(V', \phi') \cong \mathfrak{g}^{\oplus n'} \oplus W'$ is a second object in $\mathfrak{g}\backslash\mathrm{Rep}(G)$ and that under the bijection (14) (V, ϕ) (respectively (V', ϕ')) corresponds to $(\mathfrak{g}^{\oplus n} \oplus W, \underline{\lambda})$ (respectively $(\mathfrak{g}^{\oplus n'} \oplus W', \underline{\lambda}')$). If $(V, \phi) \cong (V', \phi')$ then we necessarily have $n' = n$ and $W' \cong W$. Using Schur's Lemma again, we have a bijection

$$\mathrm{Aut}_G(\mathfrak{g}^{\oplus n} \oplus W) \cong GL_n(\mathbb{R}) \times \mathrm{Aut}_G(W),$$

from which it follows that isomorphisms in $\mathfrak{g}\backslash\mathrm{Rep}(G)$ from (V, ϕ) to (V', ϕ') correspond to diagrams

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\underline{\lambda}"] \arrow[dr, "\underline{\lambda}'"'] & \mathfrak{g}^{\oplus n} \oplus W \arrow[d, "(A,\psi)"] \\ & \mathfrak{g}^{\oplus n} \oplus W \end{tikzcd}$$

in which $(A, \psi) \in GL_n(\mathbb{R}) \times \mathrm{Aut}_G(W)$ and $A\underline{\lambda} = \underline{\lambda}'$. The proposition now follows from the fact that there are exactly two $GL_n(\mathbb{R})$ orbits in $\mathbb{R}^n$: the zero orbit $\{\underline{0}\}$, and its complement $\mathbb{R}^n\backslash\{\underline{0}\}$. Under the bijection (14) the zero vector $\underline{0}$ corresponds to the zero morphism $\mathfrak{g} \to \mathfrak{g}^{\oplus n} \oplus W$ and therefore to the object $(\mathfrak{g}, 0)^n \times (W, 0) \cong (V, 0)$, and the vector $(1, \dots, 1) \in \mathbb{R}^n\backslash\{\underline{0}\}$ corresponds to the map $(\mathrm{diag}, 0) : \mathfrak{g} \to \mathfrak{g}^{\oplus n} \oplus W$ and therefore to the object $(\mathfrak{g}, \mathrm{id})^n \times (W, 0)$. $\square$

Remark 4.13. In the Grothendieck group $K^{\mathrm{aff}}_{TG}(*)$ (see Definition 6.13) the objects $[(\mathfrak{g}, \mathrm{id})]$ and $[(\mathfrak{g}, 0)]$ are, surprisingly, in fact equal. See Remark 6.16 below for a further discussion of this point.

Example 4.14 (trivial G-spaces). As a generalization of Example 4.10 consider the case where X is a trivial G-space. In this case an affine action $TG \times E \to E$ over X is equivalent to a smoothly varying family $TG \times E_x \to E_x$ of affine actions parametrised by $x \in X$. In particular, associated to each point $x \in X$ there is an object (E_x, ϕ_x) in $\mathfrak{g}\backslash\mathrm{Rep}(G)$.

Proposition 4.15 (homogeneous G-spaces). *Let G/H be a homogeneous G-space. Then there is an equivalence of categories*

$$\mathrm{Vect}^{\mathrm{aff}}_{TG}(G/H) \simeq \mathfrak{g}\backslash\mathrm{Rep}(H), \tag{15}$$

where $\mathfrak{g}$ is considered as an H-module by restriction of the adjoint representation of G.

Proof. Under the equivalence $\mathrm{Vect}_G(G/H) \to \mathrm{Rep}(H)$ the G-vector bundle $\mathfrak{g}_X$ is mapped to the H-representation $\mathfrak{g}$. Therefore

$$\mathrm{Vect}^{\mathrm{aff}}_{TG}(G/H) \simeq \mathfrak{g}_X\backslash\mathrm{Vect}_G(X) \simeq \mathfrak{g}\backslash\mathrm{Rep}(H). \qquad \square$$

Example 4.16 (the tangent bundle of G/H). Under the equivalence (15) the tangent bundle $T(G/H)$ corresponds to the quotient map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$, considered as an object in the category $\mathfrak{g}\backslash\mathrm{Rep}(H)$.

Proposition 4.17 (free G-spaces). *Suppose that P is a free, proper G-space so that P/G is a manifold and $P \to P/G$ is a principal G-bundle. Then*

$$\mathrm{Vect}^{\mathrm{aff}}_{TG}(P) \simeq \mathrm{ad}(P)\backslash\mathrm{Vect}(P/G),$$

where $\mathrm{ad}(P) = P \times_G \mathfrak{g}$ *is the adjoint bundle of P.*

Proof. The quotient construction $E \mapsto E/G$ yields an equivalence of categories $\mathrm{Vect}_G(P) \to \mathrm{Vect}(P/G)$ under which the G-vector bundle $\mathfrak{g}_P$ is mapped to the adjoint bundle $\mathrm{ad}(P) = P \times_G \mathfrak{g}$. Therefore

$$\mathrm{Vect}^{\mathrm{aff}}_{TG}(P) \simeq \mathfrak{g}_P\backslash\mathrm{Vect}_G(P) \simeq \mathrm{ad}(P)\backslash\mathrm{Vect}(P/G). \tag{16} \qquad \square$$

Example 4.18 (Atiyah algebroids). If $\pi : P \to X$ is a principal G-bundle then the *Atiyah algebroid* of P is the Lie algebroid (see Example 4.5) TP/G over X (see [Mackenzie 2005]). The Atiyah algebroid fits into a short exact sequence

$$0 \to \mathrm{ad}(P) \to TP/G \to TX \to 0$$

which arises from an application of the quotient construction to the short exact sequence

$$0 \to \mathrm{Ker}\,\pi_* \to TP \xrightarrow{\pi_*} \pi^*TX \to 0$$

and the isomorphism of G-equivariant vector bundles $\mathrm{Ker}\,\pi_* \cong \mathfrak{g}_P$. Under (16), the tangent bundle TP corresponds to the object $(TP/G, \mathrm{ad}(P) \to TP/G)$ in $\mathrm{ad}(P)\backslash\mathrm{Vect}(X)$.

5. $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ and $\mathrm{Vect}_G(X)$

In this section we relate affine actions to equivariant vector bundles and in Section 5C prove our second main result Theorem B. We mostly describe the statements in terms of the category $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$, but will indicate the corresponding results for $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ — see Remarks 5.5 and 5.7.

5A. ***The functors* U, F *and* σ.** We will define several functors between the categories $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$ and $\mathrm{Vect}_G(X)$. In Remark 5.5 we explain what are the corresponding functors between $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ and $\mathrm{Vect}_G(X)$.

Definition 5.1. We denote by U the forgetful functor

$$\mathrm{U} : \mathfrak{g}_X\backslash\mathrm{Vect}_G(X) \to \mathrm{Vect}_G(X)$$

which maps a pair (E,ϕ) to the G-vector bundle E and maps a morphism $\psi : (E,\phi) \to (E',\phi')$ to the morphism $\psi : E \to E'$.

Definition 5.2. We denote by σ the functor

$$\sigma : \mathrm{Vect}_G(X) \to \mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$$

defined on objects by $\sigma(E) = (E,0)$ and on morphisms by $\sigma(\psi : E \to E') = (\psi : (E,0) \to (E',0))$.

Definition 5.3. We denote by F the functor

$$\mathrm{F} : \mathrm{Vect}_G(X) \to \mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$$

which maps a G-vector bundle E to the object $(\mathfrak{g}_X \oplus E, i_{\mathfrak{g}_X})$, where $i_{\mathfrak{g}_X} : \mathfrak{g}_X \to \mathfrak{g}_X \oplus E$ is the natural inclusion $(x,\alpha) \mapsto ((x,\alpha),0_x)$, and maps a morphism $\psi : E \to E'$ to the morphism $\mathrm{id}_{\mathfrak{g}} \oplus \psi : (\mathfrak{g}_X \oplus E, i_{\mathfrak{g}_X}) \to (\mathfrak{g}_X \oplus E', i_{\mathfrak{g}_X})$.

Remark 5.4. Note that F can be written in terms of coproducts in $\mathrm{Vect}_G(X)$ or products in $\mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$ (see Example 4.7 and Proposition 6.8):

$$\mathrm{F} = (\mathfrak{g}_X \oplus -, i_{\mathfrak{g}_X}) = ((\mathfrak{g}_X, \mathrm{id}) \times -) \circ \sigma.$$

Remark 5.5. Via the isomorphism of Theorem A, the functors U, σ and F correspond to functors between $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ and $\mathrm{Vect}_G(X)$, which we denote by the same symbols. The functor

$$\mathrm{U} : \mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \to \mathrm{Vect}_G(X)$$

is given by the restriction of actions along the Lie group morphism $0_{TG} : G \to TG$. It maps an affine action $\mu : TG \times E \to E$ to the G-vector bundle E with action $g \cdot e_x = 0_g \cdot e_x$. The functor

$$\sigma : \mathrm{Vect}_G(X) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$$

is given by the restriction of actions along $\pi_{TG} : TG \to G$. It maps a G-vector bundle E to the affine action defined by $v_g \cdot e_x = g \cdot e_x$. See also Example 4.2. The functor

$$\mathrm{F} : \mathrm{Vect}_G(X) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$$

maps a G-equivariant vector bundle E to the affine action $TG \times (\mathfrak{g}_X \oplus E) \to \mathfrak{g}_X \oplus E$ given by

$$v_g \cdot ((x,\alpha), e_x) = ((g \cdot x, \mathrm{Ad}_g\alpha + (R_{g^{-1}})_* v_g), g \cdot e_x).$$

5B. ***Lemma on slice categories.*** This subsection contains the proof of the following category-theoretic result. Lemma D will only be used to prove Theorem C and will not be referred to elsewhere, so the reader uninterested in category-theoretic abstractions may wish to skip to Section 5C.

Recall that if $\mathcal{U} : \mathscr{D} \rightleftarrows \mathscr{C} : \mathcal{F}$ are a pair of adjoint functors, with $\mathcal{F}$ left adjoint to $\mathcal{U}$, then there is an associated *monad* acting on the category $\mathscr{C}$. This monad consists of the endofunctor $\mathbb{T} = \mathcal{U}\mathcal{F} : \mathscr{C} \to \mathscr{C}$, together with a certain pair χ, η of natural transformations which are constructed from the unit and counit of the adjunction. There is a canonical *comparison functor* $\mathcal{K} : \mathscr{D} \to \mathscr{C}^{\langle \mathbb{T}, \chi, \eta \rangle}$, from $\mathscr{D}$ to the category of $\langle \mathbb{T}, \chi, \eta \rangle$*-algebras*, and the adjunction is called *monadic* if $\mathcal{K}$ is an isomorphism. One can therefore understand the statement that an adjunction is monadic as saying that the category $\mathscr{D}$ can be reconstructed from $\mathscr{C}$ and the monad $\langle \mathbb{T}, \chi, \eta \rangle$ in a canonical way, which shows the importance of the notion. See [MacLane 1971] for further details.

Lemma D. *Let $\mathscr{C}$ be a category, m an object in $\mathscr{C}$, and $\mathcal{U} : m \backslash \mathscr{C} \to \mathscr{C}$ the standard forgetful functor. If the coproduct $m \amalg a$ exists in $\mathscr{C}$ for all objects a in $\mathscr{C}$ then the functor $\mathcal{F} : a \mapsto (m \amalg a, i_m)$ is left adjoint to $\mathcal{U}$, and this adjunction is monadic.*

Proof. Let (a', ϕ') be an object in $m \backslash \mathscr{C}$. Via the universal property of the coproduct $m \amalg a$ the natural bijection

$$\operatorname{Hom}_{\mathscr{C}}(m, a') \times \operatorname{Hom}_{\mathscr{C}}(a, a') \xrightarrow[\Sigma]{\cong} \operatorname{Hom}_{\mathscr{C}}(m \amalg a, a')$$

restricts to a natural bijection

$$\begin{aligned} \operatorname{Hom}_{\mathscr{C}}(a, a') &\xrightarrow{\cong} \operatorname{Hom}_{m\backslash\mathscr{C}}((m \amalg a, i_m), (a', \phi')) \\ f &\mapsto \Sigma((\phi', f)). \end{aligned} \tag{17}$$

The bijection (17) can be rewritten as

$$\operatorname{Hom}_{\mathscr{C}}(a, \mathcal{U}((a', \phi'))) \xrightarrow{\cong} \operatorname{Hom}_{m\backslash\mathscr{C}}(\mathcal{F}(a), (a', \phi')). \tag{18}$$

This shows that $\mathcal{F}$ is left adjoint to $\mathcal{U}$.

It remains to show that the adjunction $\mathcal{F} \dashv \mathcal{U}$ defined by (18) is monadic. We must show that the canonical comparison functor $\mathcal{K} : m \backslash \mathscr{C} \to \mathscr{C}^{\langle \mathbb{T}, \chi, \eta \rangle}$ from $m \backslash \mathscr{C}$ to the category of $\langle \mathbb{T}, \chi, \eta \rangle$-algebras is an isomorphism, where $\langle \mathbb{T}, \chi, \eta \rangle$ is the monad associated to the adjunction, and $\mathbb{T}$, χ, η and $\mathcal{K}$ are defined below (see [MacLane 1971] for the general case).

Straightforward calculations show that the unit $\eta : 1_{\mathscr{C}} \Rightarrow \mathcal{U}\mathcal{F}$ and the counit $\varepsilon : \mathcal{F}\mathcal{U} \Rightarrow 1_{m\backslash\mathscr{C}}$ of the adjunction have components

$$\begin{aligned} \eta_a &= i_a : a \to m \amalg a, \\ \varepsilon_{(a', \phi')} &= (\phi', 1_{a'}) : (m \amalg a', i_m) \to (a', \phi'). \end{aligned}$$

We denote the associated monad by $\langle \mathbb{T}, \chi, \eta\rangle$, where $\mathbb{T} = \mathcal{U}\mathcal{F} : \mathscr{C} \to \mathscr{C}$, $\chi = \mathcal{U}\varepsilon\mathcal{F} : \mathbb{T}^2 \Rightarrow \mathbb{T}$, and $\eta : 1_{\mathscr{C}} \Rightarrow \mathbb{T}$. The functor $\mathbb{T}$ is the functor $a \mapsto m \amalg a$. A further calculation shows that χ has components

$$\chi_a = (i_m, 1_{m\amalg a}) : m \amalg (m \amalg a) \to m \amalg a.$$

We can now describe the category $\mathscr{C}^{\langle \mathbb{T},\chi,\eta\rangle}$ of $\langle \mathbb{T}, \chi, \eta\rangle$-algebras, or $\mathbb{T}$-algebras for short. Let $\langle a, h\rangle$ be a $\mathbb{T}$-algebra. Then h is a morphism $h : \mathbb{T}a = m \amalg a \to a$ in $\mathscr{C}$ such that $h \circ \eta_a = 1_a$ as morphisms $a \to a$, and $h \circ \mathbb{T}h = h \circ \chi_a$ as morphisms $\mathbb{T}^2 a \to a$. Under the natural bijection

$$\mathrm{Hom}_{\mathscr{C}}(m \amalg a, a) \xrightarrow{\cong} \mathrm{Hom}_{\mathscr{C}}(m, a) \times \mathrm{Hom}_{\mathscr{C}}(a, a),$$

the first condition on h is equivalent to the condition that h is of the form $(\bar{h}, 1_a)$. A calculation then shows that the second condition on h holds automatically. It follows that $\mathbb{T}$-algebras can be identified with objects of $m\backslash\mathscr{C}$.

Let $f : \langle a, h\rangle \to \langle a', h'\rangle$ be a morphism of $\mathbb{T}$-algebras. Then $f : a \to a'$ is a morphism in $\mathscr{C}$ such that $h' \circ \mathbb{T} f = f \circ h$ as morphisms $\mathbb{T}a \to a'$. Identifying $h = (\bar{h}, 1_a)$ and $h' = (\bar{h}', 1_{a'})$ as above, the condition on f is

$$(\bar{h}', 1_a) \circ (1_m \amalg f) = f \circ (\bar{h}, 1_a)$$

which is equivalent to

$$(\bar{h}', f) = (f\bar{h}, f)$$

and therefore to $\bar{h}' = f\bar{h}$. It follows that morphisms between $\mathbb{T}$-algebras can be identified with morphisms in $m\backslash\mathscr{C}$. The discussion above shows that there is an isomorphism of categories $m\backslash\mathscr{C} \xrightarrow{\cong} \mathscr{C}^{\langle \mathbb{T},\chi,\eta\rangle}$ defined on objects by

$$(a, \bar{h}) \mapsto \langle a, (\bar{h}, 1_a)\rangle = \langle \mathcal{U}(a, \bar{h}), \mathcal{U}(\varepsilon_{(a,\bar{h})})\rangle$$

and on morphisms by

$$\begin{aligned}(f : (a, \bar{h}) \to (a', \bar{h}')) &\mapsto (f : \langle a, (\bar{h}, 1_a)\rangle \to \langle a', (\bar{h}', 1_{a'})\rangle) \\ &= \big(\mathcal{U}f : \langle \mathcal{U}(a, \bar{h}), \mathcal{U}(\varepsilon_{(a,\bar{h})})\rangle \to \langle \mathcal{U}(a', \bar{h}'), \mathcal{U}(\varepsilon_{(a',\bar{h}')})\rangle\big)\end{aligned}$$

This functor is exactly the comparison functor $\mathcal{K} : m\backslash\mathscr{C} \to \mathscr{C}^{\langle \mathbb{T},\chi,\eta\rangle}$ (see [MacLane 1971]), which implies that the adjunction $\mathcal{F} \dashv \mathcal{U}$ is monadic. □

5C. *Theorem B.*

Theorem B. *The following statements hold*:

(1) F *is left adjoint to* U.

(2) *The adjunction* F ⊣ U *is monadic.*

(3) σ *is the unique section of* U.

Proof. The category $\mathrm{Vect}_G(X)$ is additive and so, in particular, has finite coproducts. If E is an object in $\mathrm{Vect}_G(X)$ then the coproduct $\mathfrak{g}_X \amalg E$ is given by the G-vector bundle $\mathfrak{g}_X \oplus E$ together with the natural inclusions $i_{\mathfrak{g}_X} : \mathfrak{g}_X \to \mathfrak{g}_X \oplus E$ and $i_E : E \to \mathfrak{g}_X \oplus E$. It follows that F corresponds to the functor $\mathcal{F} : a \mapsto (c \amalg a, i_c)$ in the statement of Lemma D, and U to the functor $\mathcal{U}$. Therefore (1) and (2) follow from Lemma D.

We now prove (3). If E is an object in $\mathrm{Vect}_G(X)$ then $\mathrm{U}(\sigma(E)) = \mathrm{U}(E, 0) = E$, and if $\psi : E \to E'$ is a morphism then $\mathrm{U}(\sigma(\psi)) = \mathrm{U}(\psi : (E, 0) \to (E', 0)) = \psi : E \to E'$, so that σ is a section of U. It remains to show that it is unique. Let σ' be a section of U. If E is an object in $\mathrm{Vect}_G(X)$ then $\sigma'(E) = (E, \phi)$ for some morphism $\phi : \mathfrak{g}_X \to E$. Applying σ to the zero-morphism $0 : E \to E$ produces a morphism $0 : (E, \phi) \to (E, \phi)$ in $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ which implies that $\phi = 0 \circ \phi = 0$. Therefore $\sigma'(E) = (E, 0) = \sigma(E)$. If $\psi : E \to E'$ is a morphism in $\mathrm{Vect}_G(X)$ then the fact that σ' is a section of U implies that $\sigma'(\psi) = \psi : (E, 0) \to (E', 0)$. Therefore $\sigma' = \sigma$. □

Remark 5.6. It follows from the details of the proof of Lemma D, and in particular (18), that the adjunction F ⊣ U of Theorem B is given in terms of hom-sets by:

$$\mathrm{Hom}_{\mathrm{Vect}_G(X)}(E, \mathrm{U}((E', \phi'))) \xrightarrow{\cong} \mathrm{Hom}_{\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)}(\mathrm{F}(E), (E', \phi'))$$

$$\begin{tikzcd} E \arrow[d, "f"] & \longmapsto & \mathfrak{g}_X \arrow[r, "i_{\mathfrak{g}_X}"] \arrow[dr, "\phi'"'] & \mathfrak{g}_X \oplus E \arrow[d, "(\phi', f)"] \\ E' & & & E' \end{tikzcd}$$

Remark 5.7. It follows from the isomorphism $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \cong \mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ of Theorem A, that the analogue of Theorem B involving the functors

$$\mathrm{U} : \mathrm{Vect}^{\mathrm{aff}}_{TG}(X) \to \mathrm{Vect}_G(X) \quad \text{and} \quad \mathrm{F}, \sigma : \mathrm{Vect}_G(X) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$$

defined in Remark 5.5 also holds.

6. Pullbacks, products and K-theory

In this section we first show that a G-map $f : X \to Y$ determines a pullback functor $\mathrm{Vect}^{\mathrm{aff}}_{TG}(Y) \to \mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$. Unlike $\mathrm{Vect}_G(X)$, $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ does not carry any obvious additive structure. Nonetheless, $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ does have initial and terminal objects, and finite products. This will allow us to define the Grothendieck group $K^{\mathrm{aff}}_{TG}(X)$, and to show that that the above-mentioned pullback functors determine group homomorphisms between these groups. Our third main result, Theorem C, shows that the functors U and σ defined in Section 5 determine a natural isomorphism between $K^{\mathrm{aff}}_{TG}(X)$ and the equivariant K-theory $KO_G(X)$.

6A. ***Pullback functors.*** If $f_1 : X_1 \to X_2$ is a G-equivariant smooth map then the pullback of G-vector bundles determines a functor $f_1^* : \mathrm{Vect}_G(X_2) \to \mathrm{Vect}_G(X_1)$. If $f_2 : X_2 \to X_3$ is a composable G-equivariant smooth map then there is a natural isomorphism of functors $(f_2 f_1)^* \cong f_1^* f_2^*$. Note that there is a canonical isomorphism of G-vector bundles $f_1^* \mathfrak{g}_{X_2} \cong \mathfrak{g}_{X_1}$ given by $(x, (f_1(x), \xi)) \mapsto (x, \xi)$.

Definition 6.1. If $f : X \to Y$ is a G-equivariant smooth map then we define the functor
$$\tilde{f}^* : \mathfrak{g}_Y \backslash \mathrm{Vect}_G(Y) \to \mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$$
to be the composition
$$\mathfrak{g}_X \backslash \mathrm{Vect}_G(Y) \to (f^* \mathfrak{g}_Y) \backslash \mathrm{Vect}_G(X) \xrightarrow{\cong} \mathfrak{g}_X \backslash \mathrm{Vect}_G(X), \tag{19}$$
where the first functor in (19) is the functor between slice categories determined by the pullback of G-vector bundles, and the second is determined by the canonical isomorphism $f^* \mathfrak{g}_Y \cong \mathfrak{g}_X$.

Remark 6.2. In terms of affine actions the action of TG on $f^* E$ is given by $v_g \cdot (x, e) = (g \cdot x, v_g \cdot e)$.

The following proposition follows immediately from the corresponding result for G-vector bundles:

Proposition 6.3. *If f_1, f_2 are composable G-equivariant maps and $\tilde{f}_1^*, \tilde{f}_2^*$ are the pullback functors defined in Definition 6.1, then there is a natural isomorphism $\widetilde{(f_2 f_1)}^* \cong \tilde{f}_1^* \tilde{f}_2^*$.*

Remark 6.4. The categories $\mathrm{Vect}_{TG}^{\mathrm{aff}}(X)$ for varying X, functors of Definition 6.1, and natural isomorphisms of Proposition 6.3 in fact constitute a *pseudo-functor* G-Man $\to$ Cat from the category of G-manifolds to the 2-category of (essentially small) categories, but we shall not make use of this fact.

The pullback functors defined above are compatible with the functors U, σ and F defined in Section 5A in the following sense. Let us temporarilly record the dependence on X by U_X, σ_X and F_X. The following proposition then follows from the definitions of these functors:

Proposition 6.5. *Let $f : X \to Y$ be a G-map. Then there are natural isomorphisms*
$$\mathrm{U}_X \tilde{f}^* \cong f^* \mathrm{U}_Y, \quad \sigma_X f^* \cong \tilde{f}^* \sigma_Y, \quad \mathrm{F}_X f^* \cong \tilde{f}^* \mathrm{F}_Y.$$

6B. ***Products.***

Proposition 6.6. $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ *has a terminal object, which is given by* $(X \times \{0\}, 0)$, *and an initial object, given by* $(\mathfrak{g}_X, \mathrm{id})$.

Proof. This follows from the definition of the slice category $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ and the fact that $X \times \{0\}$ is a terminal object in $\mathrm{Vect}_G(X)$. $\square$

Corollary 6.7. *The category* $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ *has a terminal object given by the zero-vector bundle* $X \times \{0\}$ *with action* $v_g \cdot (x, 0) = (g \cdot x, 0)$, *and an initial object given by the vector bundle* $\mathfrak{g}_X$ *with action* $v_g \cdot (x, \alpha) = (g \cdot x, \mathrm{Ad}_g \alpha + (R_{g^{-1}})_* v_g)$.

Proof. This follows from the isomorphism of Theorem A. □

Proposition 6.8. *The category* $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ *has finite products. The product of* (E, ϕ) *and* (E', ϕ') *is given by the object* $(E \oplus E', (\phi, \phi'))$ *and the canonical projections to* (E, ϕ) *and* (E', ϕ').

Proof. The result follows from general facts about limits in slice categories. Explicitly, let (F, ψ) be an object in $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ and $f : (F, \psi) \to (E, \phi)$ and $f' : (F, \psi) \to (E', \phi')$ be morphisms. In particular, $f \psi = \phi$ and $f' \psi = \phi'$. As $E \oplus E'$ is a product in $\mathrm{Vect}_G(X)$ there is a unique morphism $(f, f') : F \to E \oplus E'$ satisfying $\mathrm{pr}_E \circ (f, f') = f$ and $\mathrm{pr}_{E'} \circ (f, f') = f'$. We have $(f, f') \circ \psi = (f\psi, f'\psi) = (\phi, \phi')$ and therefore (f, f') defines a morphism $(f, f') : (F, \psi) \to (E \oplus E', (\phi, \phi'))$ in $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$. Finally, the identities above involving pr_E and $\mathrm{pr}_{E'}$ imply that the same identities hold when these morphisms are considered as morphisms in $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$. This shows that $(E \oplus E', (\phi, \phi'))$ and the canonical projections to (E, ϕ) and (E', ϕ') satisfy the required universal property. □

Corollary 6.9. *The category* $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$ *has finite products. If* E *and* E' *are vector bundles equipped with affine actions then their product is the vector bundle* $E \oplus E'$ *with affine action* $v_g \cdot (e, e') = (v_g \cdot e, v_g \cdot e')$.

Proof. This follows from the isomorphism of Theorem A. □

Proposition 6.10. *The functors* U *and* σ *preserve products.*

Proof. This follows immediately from the definitions of the functors U and σ and the description of products in $\mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$ given in Proposition 6.8. □

Remark 6.11. Note that if G is not discrete then the left adjoint functor F of Definition 5.3 does not preserve products:

$$\begin{aligned} \mathrm{F}(E) \times \mathrm{F}(E') &= ((\mathfrak{g}_X \oplus E) \oplus (\mathfrak{g}_X \oplus E'), (i_{\mathfrak{g}_X}, i_{\mathfrak{g}_X})) \\ &\ncong (\mathfrak{g}_X \oplus (E \oplus E'), i_{\mathfrak{g}_X}) \\ &= \mathrm{F}(E \times E'). \end{aligned}$$

Proposition 6.12. *If* $f : X \to Y$ *is a* G*-map then the functor*

$$\tilde{f}^* : \mathfrak{g}_Y \backslash \mathrm{Vect}_G(Y) \to \mathfrak{g}_X \backslash \mathrm{Vect}_G(X)$$

of Definition 6.1 preserves products.

Proof. This follows from the definition of the functor $\tilde{f}^*$ and the fact that the functor $f^* : \mathrm{Vect}_G(Y) \to \mathrm{Vect}_G(X)$ preserves products. □

6C. *K-theory.*

Definition 6.13. Let $K^{\mathrm{aff}}_{TG}(X)$ be the Grothendieck group of the category $\mathrm{Vect}^{\mathrm{aff}}_{TG}(X)$.

It follows from Proposition 6.12 and the functoriality of the construction of the Grothendieck group (see Section 2C) that the functor

$$\tilde{f}^* : \mathfrak{g}_Y\backslash\mathrm{Vect}_G(Y) \to \mathfrak{g}_X\backslash\mathrm{Vect}_G(X)$$

associated to a G-map $f : X \to Y$ induces a group homomorphism $K(\tilde{f}^*) : K^{\mathrm{aff}}_{TG}(Y) \to K^{\mathrm{aff}}_{TG}(X)$.

Definition 6.14. We denote by

$$K^{\mathrm{aff}}_{TG}(-) : G\text{-Man} \to \mathrm{Ab}$$

the functor which maps a G-manifold X to the abelian group $K^{\mathrm{aff}}_{TG}(X)$ and maps a G-map $f : X \to Y$ to the group homomorphism $K(\tilde{f}^*)$.

Our last main result is that the functor $K^{\mathrm{aff}}_{TG}(-)$ is isomorphic to $K\,O_G(-)$ (recall from Section 2C that $K\,O_G(X)$ denotes the Grothendieck group of $\mathrm{Vect}_G(X)$). The proof is based on the following Lemma, which states the category-theoretic reasons for this result.

Lemma E. *Let $\mathscr{C}$ be an additive category and m an object in $\mathscr{C}$. Let $\mathcal{U} : m\backslash\mathscr{C} \to \mathscr{C}$ be the standard forgetful functor and $\mathcal{S} : \mathscr{C} \to m\backslash\mathscr{C}$ the section $a \mapsto (a, 0)$. Then the group homomorphism*

$$K(\mathcal{U}) : K(m\backslash\mathscr{C}) \to K(\mathscr{C})$$

is an isomorphism. Its inverse is

$$K(\mathcal{S}) : K(\mathscr{C}) \to K(m\backslash\mathscr{C}).$$

Proof. As $\mathcal{U} \circ \mathcal{S} = 1_{\mathscr{C}}$, the functoriality of the Grothendieck group implies that the homomorphism $K(\mathcal{U})$ is surjective, $K(\mathcal{S})$ is injective, and $K(\mathcal{U}) \circ K(\mathcal{S}) = 1_{K(\mathscr{C})}$. The lemma will therefore follow from the fact that $K(\mathcal{S})$ is surjective, which we will prove.

Let (a, f) be an object in $m\backslash\mathscr{C}$. Consider the diagram

$$\begin{tikzcd} m \arrow[r, "{(f,f)}"] \arrow[dr, "{(f,0)}"'] & a \oplus a \arrow[d, "h"] \\ & a \oplus a \end{tikzcd} \tag{20}$$

where $h = (\mathrm{pr}_1, \mathrm{pr}_2 - \mathrm{pr}_1)$, or in terms of elements $h(e_1, e_2) = (e_1, e_2 - e_1)$. It is clear that (20) commutes, and h is an isomorphism with inverse $(\mathrm{pr}_1, \mathrm{pr}_2 + \mathrm{pr}_1)$.

Therefore, h determines an isomorphism

$$(a, f) \times (a, f) \xrightarrow{\cong} (a, f) \times (a, 0)$$

in $m\backslash\mathscr{C}$. It follows that in the Grothendieck group $K(m\backslash\mathscr{C})$ we have

$$[(a, f)] = [(a, 0)] = K(\mathcal{S})([a]).$$

Therefore, if (a_1, f_1) and (a_2, f_2) are objects in $m\backslash\mathscr{C}$, then

$$[(a_1, f_1)] - [(a_2, f_2)] = K(\mathcal{S})([a_1] - [a_2]),$$

which shows that $K(\mathcal{S})$ is surjective. □

Remark 6.15. There exist similar results describing the Grothendieck groups of various categories associated to an additive category $\mathscr{C}$. For example, the Main Theorem in [Almkvist 1974] describes the group $K(\operatorname{end} P(A))$, where $\operatorname{end} P(A)$ is the additive category of endomorphisms of finitely generated projected modules over a commutative ring A.

Theorem C. *If X is a G-manifold then the functor* U *induces a group isomorphism*

$$K(\mathrm{U}) : K^{\mathrm{aff}}_{TG}(X) \to K O_G(X).$$

Its inverse is

$$K(\sigma) : K O_G(X) \to K^{\mathrm{aff}}_{TG}(X).$$

These isomorphisms are natural in X, and thus define an isomorphism of functors

$$K^{\mathrm{aff}}_{TG}(-) \xrightarrow{\cong} K O_G(-).$$

Proof. For fixed X, the fact that $K(\mathrm{U})$ and $K(\sigma)$ are mutually inverse isomorphisms follows from Lemma E. The fact that these isomorphisms are natural in X follows from the first two natural isomorphisms in Proposition 6.5 and the functoriality of the Grothendieck group. □

Remark 6.16. Note that Theorem C does not hold at the level of monoids. The functor U induces a morphism of commutative monoids

$$(\mathfrak{g}_X\backslash\mathrm{Vect}_G(X))/\cong \;\to (\mathrm{Vect}_G(X))/\cong$$

which is an isomorphism if and only if G is discrete. In particular, $[(\mathfrak{g}_X, 0)]$ and $[(\mathfrak{g}_X, \mathrm{id})]$ are both mapped to $[\mathfrak{g}_X]$. This can be seen most explicitly in the situation of Proposition 4.11 (see also Remark 4.12). In this case the equality $[(\mathfrak{g}, \mathrm{id})] = [(\mathfrak{g}, 0)]$ in $K^{\mathrm{aff}}_{TG}(*)$ implies the equality

$$[(\mathfrak{g}, \mathrm{id})^n \times (W, 0)] = [(\mathfrak{g}, 0)^n \times (W, 0)]$$

in $K^{\mathrm{aff}}_{TG}(*)$ between the two classes of elements appearing in the classification of objects in $\mathfrak{g}\backslash\mathrm{Rep}(G)$ given in Proposition 4.11.

7. The complex case

It is possible to reformulate the notion of affine actions in the complex setting. The complexified tangent bundle $T_{\mathbb{C}}G$ is both a complex vector bundle and a *real* Lie group. We define an affine action of $T_{\mathbb{C}}G$ on a complex vector bundle E to be an action for which the action map $T_{\mathbb{C}}G \times E \to E$ is a morphism of *complex* vector bundles. For example, if X is a G-manifold with action $G \times X \to X$ then the derivative defines a complex affine action $T_{\mathbb{C}}G \times T_{\mathbb{C}}X \to T_{\mathbb{C}}X$ of $T_{\mathbb{C}}G$ on the complexified tangent bundle $T_{\mathbb{C}}X$ of X.

We use $\mathrm{Vect}^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X)$ (respectively $\mathrm{Vect}^{\mathbb{C}}_{G}(X)$) to denote the category of complex affine actions (respectively G-equivariant complex vector bundles) over X, and $K^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X)$ (respectively $K_G(X)$) to denote the Grothendieck group of $\mathrm{Vect}^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X)$ (respectively $\mathrm{Vect}^{\mathbb{C}}_{G}(X)$). As standard, $\mathfrak{g}_{\mathbb{C}}$ denotes the complexified Lie algebra of G equipped with the complexified adjoint representation.

The functors U, σ and F defined in Section 5A have analogues in the complex case, which we denote by

$$\mathrm{U}^{\mathbb{C}} : \mathrm{Vect}^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X) \to \mathrm{Vect}^{\mathbb{C}}_{G}(X)$$

and

$$\mathrm{F}^{\mathbb{C}}, \sigma^{\mathbb{C}} : \mathrm{Vect}^{\mathbb{C}}_{G}(X) \to \mathrm{Vect}^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X).$$

7A. ***Complex theorems.*** The following analogues of Theorems A, B and C hold with essentially the same proofs.

Theorem A′. *The following three categories are isomorphic*:

(1) *The category* $\mathrm{Vect}^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X)$ *of complex affine actions of* $T_{\mathbb{C}}G$ *over* X.

(2) *The category of pairs* $(E, \rho_{\mathfrak{g}})$, *defined as follows*:
- *The objects are pairs* $(E, \rho_{\mathfrak{g}})$, *where* E *is a* G-*equivariant complex vector bundle over* X *and* $\rho_{\mathfrak{g}_{\mathbb{C}}} : \mathfrak{g}_{\mathbb{C}} \to \Gamma(E)$ *is a* G-*equivariant complex linear map.*
- *The morphisms* $(E, \rho_{\mathfrak{g}}) \to (E', \rho_{\mathfrak{g}})$ *are morphisms* $\psi : E \to E'$ *of* G-*equivariant complex vector bundles over* X *such that* $\Gamma(\psi) \circ \rho_{\mathfrak{g}_{\mathbb{C}}} = \rho'_{\mathfrak{g}_{\mathbb{C}}}$.
- *Composition is given by composition of morphisms of complex vector bundles over* X.

(3) *The slice category* $(\mathfrak{g}_{\mathbb{C}})_X \backslash \mathrm{Vect}^{\mathbb{C}}_{G}(X)$.

Theorem B′. *The following statements hold*:

(1) $\mathrm{F}^{\mathbb{C}}$ *is left adjoint to* $\mathrm{U}^{\mathbb{C}}$.

(2) *The adjunction* $\mathrm{F}^{\mathbb{C}} \dashv \mathrm{U}^{\mathbb{C}}$ *is monadic.*

(3) $\sigma^{\mathbb{C}}$ *is the unique section of* $\mathrm{U}^{\mathbb{C}}$.

Theorem C′. *If X is a G-manifold then the functor* $\mathrm{U}^{\mathbb{C}}$ *induces a group isomorphism*

$$K(\mathrm{U}^{\mathbb{C}}) : K^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X) \to K_G(X).$$

Its inverse is

$$K(\sigma^{\mathbb{C}}) : K_G(X) \to K^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(X).$$

These isomorphisms are natural in X, and thus define an isomorphism of functors

$$K^{\mathrm{aff},\mathbb{C}}_{T_{\mathbb{C}}G}(-) \xrightarrow{\cong} K_G(-).$$

Acknowledgements

The author would like to thank Martina Balagović for useful comments on the first draft of this paper, and Peter Jørgensen for his continued guidance during the completion of this work. The author would also like to thank the Department of Mathematics at the University of Zagreb for their hospitality during part of the time in which this work was completed.

References

[Alekseev and Meinrenken 2009] A. Alekseev and E. Meinrenken, “The Atiyah algebroid of the path fibration over a Lie group”, *Lett. Math. Phys.* **90**:1-3 (2009), 23–58. MR Zbl

[Almkvist 1974] G. Almkvist, “The Grothendieck ring of the category of endomorphisms”, *J. Algebra* **28** (1974), 375–388. MR Zbl

[Beilinson and Bernstein 1993] A. Beilinson and J. Bernstein, “A proof of Jantzen conjectures”, pp. 1–50 in *I. M. Gel’fand Seminar*, edited by S. Gel’fand and S. Gindikin, Adv. Soviet Math. **16**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

[Bruzzo et al. 2009] U. Bruzzo, L. Chirio, P. Rossi, and V. N. Rubtsov, “Equivariant cohomology and localization for Lie algebroids”, *Funktsional. Anal. i Prilozhen.* **43**:1 (2009), 22–36. In Russian; translated in *Funct. Anal. Appl.* **43**:1 (2009), 18–29. MR Zbl

[Ginzburg 1999] V. L. Ginzburg, “Equivariant Poisson cohomology and a spectral sequence associated with a moment map”, *Int. J. Math.* **10**:8 (1999), 977–1010. MR Zbl

[Mackenzie 2005] K. C. H. Mackenzie, *General theory of Lie groupoids and Lie algebroids*, London Math. Soc. Lecture Note Series **213**, Cambridge Univ. Press, 2005. MR Zbl

[MacLane 1971] S. MacLane, *Categories for the working mathematician*, Graduate Texts in Math. **5**, Springer, 1971. MR Zbl

[Marrero et al. 2012] J. C. Marrero, E. Padrón, and M. Rodríguez-Olmos, “Reduction of a symplectic-like Lie algebroid with momentum map and its application to fiberwise linear Poisson structures”, *J. Phys. A* **45**:16 (2012), art. id. 165201. MR Zbl

[nLab 2009–] nLab authors, “Over category”, wiki article, 2009–, available at https://ncatlab.org/nlab/show/over+category.

[Onishchik 2004] A. L. Onishchik, *Lectures on real semisimple Lie algebras and their representations*, Eur. Math. Soc., Zürich, 2004. MR Zbl

[Segal 1968] G. Segal, “Equivariant K-theory”, *Inst. Hautes Études Sci. Publ. Math.* **34** (1968), 129–151. MR Zbl

Received October 23, 2017.

JAMES WALDRON
SCHOOL OF MATHEMATICS, STATISTICS AND PHYSICS
NEWCASTLE UNIVERSITY
NEWCASTLE UPON TYNE
UNITED KINGDOM
james.waldron@newcastle.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.667

OPTIMAL DECAY ESTIMATE OF STRONG SOLUTIONS FOR THE 3D INCOMPRESSIBLE OLDROYD-B MODEL WITHOUT DAMPING

RENHUI WAN

We obtain the decay estimate of the global solution for the 3D incompressible Oldroyd-B model with only dissipation. The decay rate is optimal in the sense that this rate coincides with that of the linear system, which improves upon work by Zhu (*J. Funct. Anal.* 274:7 (2018) 2039–2060.)

1. Introduction

In this paper, we consider the Cauchy problem for the three-dimensional (3D) incompressible Oldroyd-B model given by

$$\begin{cases} \partial_t u + u\cdot\nabla u - \nu\Delta u + \nabla p = \mu_1 \operatorname{div}\tau, \ (t,x)\in\mathbb{R}^+\times\mathbb{R}^3 \\ \partial_t \tau + u\cdot\nabla\tau + a\tau + Q(\tau,\nabla u) = \mu_2 D(u), \\ \operatorname{div} u = 0, \\ u|_{t=0} = u_0(x), \quad \tau|_{t=0} = \tau_0(x), \end{cases} \tag{1-1}$$

where $u = (u_1, u_2, u_3)$ stands for the 3D velocity field, p the pressure and τ the non-Newtonian part of the stress tensor which can be seen as a symmetric matrix here. The values ν, a, μ_1 and μ_2 are nonnegative parameters, where we call μ_1 and μ_2 the coupling parameters. $D(u)$ and $W(u)$ are the deformation tensor and vorticity tensor, which can be given by

$$D(u) \triangleq \frac{\nabla u + (\nabla u)^\top}{2}, \quad W(u) \triangleq \frac{\nabla u - (\nabla u)^\top}{2},$$

and

$$Q(\tau,\nabla u) = \tau W(u) - W(u)\tau - \alpha(D(u)\tau + \tau D(u)), \ \alpha\in[-1,1].$$

The initial data are (u_0,τ_0) satisfying $\operatorname{div} u_0 = 0$ and $(\tau_0)_{ij} = (\tau_0)_{ji}$.

We refer to [Bird et al. 1977; Chemin and Masmoudi 2001; Oldroyd 1958] for the details about the derivation of (1-1). Guillopé and Saut [1990a; 1990b] obtained

MSC2010: primary 76A05, 76D03; secondary 42A38, 42B25.
Keywords: Oldroyd-B model, global small solution, integral representation.

local existence and uniqueness, when the initial data is in Sobolev space, which was extended to the Lebesgue space in [Fernández Cara et al. 1994]. Lions and Masmoudi [2000] obtained global existence of the weak solution for the case $\alpha = 0$. However, to the best of our knowledge, whether the case $\alpha \neq 0$ can yield a global weak solution is an open question. Considering the initial data in critical Besov space, Chemin and Masmoudi [2001] showed local well-posedness of the solution and global well-posedness with small initial data, where they required the small coupling parameter for the global result. We refer to [Chen and Miao 2008; Zi et al. 2014] for the results in generalized space. By using some techniques developed in the study of compressible Navier–Stokes, Zi et al. [2014] removed the condition concerning the small coupling parameter. Recently, Fang and Zi [2016] proved global well-posedness with a new class of large initial data admitting large initial vertical velocity. By using a decomposition technique, we [Wan 2019] improved the initial condition in [Fang and Zi 2016] to the more generalized initial condition. In particular, for the 2D case, [Wan 2019] proved global well-posedness with large initial velocity, which improved upon the corresponding work in [Fang and Zi 2016].

We point out that all the above works are based on $a > 0$. If $a = 0$, it seems difficult to proved the global well-posedness even under small initial data. By using a new method, i.e., constructing the time-weighted energies, Zhu [2018] first proved global well-posedness in 3D under small initial data, which can be showed as follows:

Theorem 1.1 [Zhu 2018]. *Let $\nu = \mu_1 = \mu_2 = 1$ and $a = 0$. Let $N \geq 7$ be a big enough constant. Assume we have the initial data $(|D|^{-1}u_0, |D|^{-1}\tau_0) \in H^N(\mathbb{R}^3) \times H^N(\mathbb{R}^3)$. There exists a small enough constant ϵ such that if*

$$\||D|^{-1}u_0\|_{H^N} + \||D|^{-1}\tau_0\|_{H^N} < \epsilon, \tag{1-2}$$

the system (1-1) *has a unique global solution (u, τ) satisfying*

$$\|u(t)\|_{H^{N-1}} + \|\tau(t)\|_{H^{N-1}} \lesssim \epsilon \quad \text{for all } t > 0.$$

Remark 1.2. Zhu [2018] only assumed $N = 3$; here the assumption of $N \geq 7$ in Theorem 1.1 is used to get the faster decay rate of the solution in the following context. For instance, we need high regularity of the solution when showing the decay estimate of $\|\nabla u(t)\|_{L^\infty}$; see Section 5 for the details.

Let us remark that the approach in [Zhu 2018] seems difficult to use for the 2D case, since $\|u(t)\|_{L^\infty(\mathbb{R}^2)}$ may yield a weakened decay, which is not integrable in time and may bring the growth of the solutions. Recently, based on a time-space approach, a new system concerning $u - 2\mathbb{P}\,\mathrm{div}\tau$ and the estimate in Besov space, [Wan 2017b] proved the global small solution for the 2D case.

It is known that long-time behavior of the global solution is an interesting and important issue in the studies of many fluids. Notice that Zhu [2018] obtained the decay estimate in the 3D case by constructing some time-weighted estimates such as

$$\|u(t)\|_{L^2} \lesssim (1+t)^{-\frac{1}{2}}, \qquad \|\nabla u(t)\|_{L^2} \lesssim (1+t)^{-1}. \tag{1-3}$$

So a natural question is whether the decay estimate (1-3) is optimal. Motivated by this, in this paper, we will focus on the long-time behavior of the global solution for (1-1) with $a = 0$. Let us first introduce some notation:

$$X(t) \triangleq \sup_{0\le\tau\le t} \big\{\langle\tau\rangle^{\frac{3}{4}}\|u(\tau)\|_{L^2} + \langle\tau\rangle^{\frac{3}{2}}\|\hat{u}(\tau)\|_{L^1}\big\},$$

$$Y(t) \triangleq \sup_{0\le\tau\le t} \big\{\langle\tau\rangle^{\frac{5}{4}}(\|\nabla u(\tau)\|_{L^2} + \|\mathcal{A}(\tau)\|_{L^2}) + \langle\tau\rangle^{2}\||\xi|\hat{u}(\tau)\|_{L^1}\big\},$$

where $\mathcal{A} = \mathbb{P}\,\mathrm{div}\tau$. Now, we state our main result.

Theorem 1.3. *Let* (u,τ) *be the solution of* (1-1) *obtained in Theorem 1.1 with the initial data satisfying* (1-2) *and* $\|(u_0,\tau_0)\|_{L^1} < \epsilon$. *Then we have*

$$X(t) + Y(t) \lesssim \epsilon.$$

In particular, *for all* $t > 0$, *the following decay estimates hold*:

$$\|u(t)\|_{L^2} \lesssim \epsilon\langle t\rangle^{-\frac{3}{4}}, \qquad \|\nabla u(t)\|_{L^2} + \|\mathcal{A}(t)\|_{L^2} \lesssim \epsilon\langle t\rangle^{-\frac{5}{4}}$$

and

$$\|u(t)\|_{L^\infty} \lesssim \epsilon\langle t\rangle^{-\frac{3}{2}}, \qquad \|\nabla u(t)\|_{L^\infty} \lesssim \epsilon\langle t\rangle^{-2}.$$

Remark 1.4. One can also get $\|\mathcal{A}(t)\|_{L^\infty} \lesssim \epsilon\langle t\rangle^{-2}$ by following the same idea. Since this decay estimate does not play an essential role in the estimate of $X(t)$ and $Y(t)$, we omit it.

Remark 1.5. One can easily find the obtained decay rate is faster than (1-3). In addition, this faster decay rate yields many better time-weighted estimates. For instance, by using our decay estimate we get that

$$\int_0^t \langle t'\rangle^a \|\nabla u(t')\|_{L^2}\,dt' \lesssim \epsilon \quad \text{for all } a \in \big(0, \tfrac{1}{4}\big),$$

while the associated result in [Zhu 2018, page 7] can only yield

$$\int_0^t \langle t'\rangle^a \|\nabla u(t')\|_{L^2}\,dt' \lesssim \epsilon \quad \text{for all } a < 0.$$

Remark 1.6. By using interpolation inequality, we can obtain the explicit decay rate in L^p $(2 \le p \le \infty)$ space, which is optimal since it is consistent with the linear part of system (1-1).

The paper is structured as follows. In Section 2, we provide some definitions of spaces and several lemmas. Section 3 is devoted to giving the integral representation. Section 4 bounds the estimate of $\|u(t)\|_{H^1}$. Section 5 provides the estimate of $\|\hat{u}(t)\|_{L^1}$, $\||\xi|\hat{u}(t)\|_{L^1}$ and $\|\mathcal{A}(t)\|_{L^2}$. In the last section, we prove Theorem 1.3. In the Appendix, we provide the proof of Lemma 4.1 and (5-3).

Let us complete this section by describing the notation we shall use in this paper.

Notation. For two operators A and B, we denote by $[A, B] = AB - BA$ the commutator between A and B, $\langle t\rangle$ means $1 + |t|$, and $A \lesssim B$ means that there exists a constant C such that $A \le CB$.

2. Preliminaries

In this section, we give some necessary definitions, propositions and lemmas in d dimensions.

The fractional Laplacian operator $|D|^\alpha = (-\Delta)^{\frac{\alpha}{2}}$ is defined through the Fourier transform, namely,

$$\widehat{|D|^\alpha f}(\xi) \triangleq |\xi|^\alpha \hat{f}(\xi),$$

where the Fourier transform is given by

$$\hat{f}(\xi) \triangleq \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx.$$

Let $\mathfrak{C} = \left\{\xi \in \mathbb{R}^d,\ \frac{3}{4} \le |\xi| \le \frac{8}{3}\right\}$. Choose a smooth radial function φ supported on $\mathfrak{C}$ such that

$$\sum_{j\in\mathbb{Z}} \varphi(2^{-j}\xi) = 1, \quad \xi \in \mathbb{R}^d \setminus \{0\}.$$

We denote $\varphi_j = \varphi(2^{-j}\xi)$, $h = \mathfrak{F}^{-1}\varphi$, where $\mathfrak{F}^{-1}$ stands for the inverse Fourier transform. Then the dyadic blocks Δ_j and S_j can be defined as

$$\Delta_j f = \varphi(2^{-j}D) f = 2^{jd} \int_{\mathbb{R}^d} h(2^j y) f(x-y)\, dy, \quad S_j f = \sum_{k\le j-1} \Delta_k f.$$

We can easily verify that with our choice of φ

$$\Delta_j \Delta_k f = 0 \text{ if } |j-k| \ge 2 \quad \text{and} \quad \Delta_j(S_{k-1} f \Delta_k f) = 0 \text{ if } |j-k| \ge 5.$$

Let us recall the definition of the Besov space.

Definition 2.1. Let $s \in \mathbb{R}$ and $(p, q) \in [1, \infty]^2$. The homogeneous Besov space $\dot{B}^s_{p,q}(\mathbb{R}^d)$ is defined by

$$\dot{B}^s_{p,q}(\mathbb{R}^d) = \{ f \in \mathfrak{S}'(\mathbb{R}^d);\ \|f\|_{\dot{B}^s_{p,q}(\mathbb{R}^d)} < \infty \},$$

where

$$\|f\|_{\dot{B}^s_{p,q}(\mathbb{R}^d)} = \begin{cases} \left(\sum_{j \in \mathbb{Z}} 2^{sqj} \|\Delta_j f\|^q_{L^p(\mathbb{R}^d)} \right)^{\frac{1}{q}} & \text{for } 1 \le q < \infty, \\ \sup_{j \in \mathbb{Z}} 2^{sj} \|\Delta_j f\|_{L^p(\mathbb{R}^d)} & \text{for } q = \infty, \end{cases}$$

and $\mathfrak{S}'(\mathbb{R}^d)$ denotes the dual space of

$$\mathfrak{S}(\mathbb{R}^d) = \{ f \in \mathcal{S}(\mathbb{R}^d);\ \partial^\alpha \hat{f}(0) = 0, \text{ for all } \alpha \in \mathbb{N}^d \text{ multi-index} \}$$

and can be identified by the quotient space of $\mathcal{S}'/\mathcal{P}$ with the polynomial space $\mathcal{P}$.

Thanks to the definition of Δ_j, we have

$$\|[\Delta_j, f]g\|_{L^p} \lesssim 2^{-j} \|\nabla f\|_{L^{p_1}} \|g\|_{L^{p_2}}, \tag{2-1}$$

where $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. Equation (2-1) will be used in the Appendix.

The following proposition provides Bernstein type inequalities. For more details about Besov space such as some useful embedding relations, see, e.g., [Bahouri et al. 2011; Stein 1970].

Proposition 2.2. *Let $1 \le p \le q \le \infty$. Then for any $\beta, \gamma \in (\mathbb{N} \cup \{0\})^d$, there exists a constant C independent of f, j such that*:

(1) *If f satisfies*

$$\operatorname{supp} \hat{f} \subset \{ \xi \in \mathbb{R}^d : |\xi| \le \mathcal{K} 2^j \},$$

then

$$\|\partial^\gamma f\|_{L^q(\mathbb{R}^d)} \le C 2^{j|\gamma| + jd\left(\frac{1}{p} - \frac{1}{q}\right)} \|f\|_{L^p(\mathbb{R}^d)}.$$

(2) *If f satisfies*

$$\operatorname{supp} \hat{f} \subset \{ \xi \in \mathbb{R}^d : \mathcal{K}_1 2^j \le |\xi| \le \mathcal{K}_2 2^j \}$$

then

$$\|f\|_{L^p(\mathbb{R}^d)} \le C 2^{-j|\gamma|} \sup_{|\beta| = |\gamma|} \|\partial^\beta f\|_{L^p(\mathbb{R}^d)}.$$

For the special case $p = q = 2$, we have

$$\|f\|_{\dot{H}^s(\mathbb{R}^d)} \approx \|f\|_{\dot{B}^s_{2,2}(\mathbb{R}^d)}.$$

The $\dot{H}^s(\mathbb{R}^d)$ and $H^s(\mathbb{R}^d)$ norm of f can be also defined as follows:

$$\|f\|_{\dot{H}^s(\mathbb{R}^d)} \triangleq \||D|^s f\|_{L^2(\mathbb{R}^d)}, \quad s \in \mathbb{R},$$
$$\|f\|_{H^s(\mathbb{R}^d)} \triangleq \|f\|_{L^2(\mathbb{R}^d)} + \||D|^s f\|_{L^2(\mathbb{R}^d)}, \quad s > 0.$$

Let us introduce the homogeneous Bony's decomposition:

$$uv = T_u v + T_v u + R(u, v),$$

where

$$T_u v = \sum_{j\in\mathbb{Z}} S_{j-1} u \Delta_j v, \quad T_v u = \sum_{j\in\mathbb{Z}} \Delta_j u S_{j-1} v, \quad R(u, v) = \sum_{j\in\mathbb{Z}} \Delta_j u \tilde{\Delta}_j v;$$

here $\tilde{\Delta}_j = \Delta_{j-1} + \Delta_j + \Delta_{j+1}$.

Lemma 2.3. [Kato and Ponce 1988] (i) *Let $s > 0$ and $1 \le p, r \le \infty$, then*

$$\|fg\|_{\dot{B}^s_{p,r}(\mathbb{R}^d)} \le C\big\{\|f\|_{L^{p_1}(\mathbb{R}^d)}\|g\|_{\dot{B}^s_{p_2,r}(\mathbb{R}^d)} + \|g\|_{L^{r_1}(\mathbb{R}^d)}\|f\|_{\dot{B}^s_{r_2,r}(\mathbb{R}^d)}\big\}, \tag{2-2}$$

where $1 \le p_1, r_1 \le \infty$ such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{r_1} + \frac{1}{r_2}$.

[Kenig et al. 1991] (ii) *Let $s > 0$, and $1 < p < \infty$, then*

$$\begin{aligned}(2\text{-}3)\quad &\|[|D|^s, f]g\|_{L^p(\mathbb{R}^d)} \\ &\le C\big\{\|\nabla f\|_{L^{p_1}(\mathbb{R}^d)}\||D|^{s-1} g\|_{L^{p_2}(\mathbb{R}^d)} + \||D|^s f\|_{L^{p_3}(\mathbb{R}^d)}\|g\|_{L^{p_4}(\mathbb{R}^d)}\big\},\end{aligned}$$

where $1 < p_2, p_3 < \infty$ such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p_3} + \frac{1}{p_4}$.

The inequalities below are used frequently in the proof.

Lemma 2.4 [Wan 2017a]. *If $0 < s_1 \le s_2$, then*

$$\int_0^t \langle t-\tau\rangle^{-s_1}\langle\tau\rangle^{-s_2}\,d\tau \le \begin{cases} C\langle t\rangle^{-s_1} & \text{if } s_2 > 1, \\ C\langle t\rangle^{-s_1}\ln(1+t) & \text{if } s_2 = 1, \\ C\langle t\rangle^{1-s_1-s_2} & \text{if } s_2 < 1. \end{cases}$$

Remark 2.5. For the case $s_1 \ge s_2$, since one can get a similar result by using the change of variable, we omit the details.

Let us introduce a generalized estimate of the solution to the heat equation.

Lemma 2.6. *If $f \in L^1(\mathbb{R}^d)$, we have*

$$\|e^{t\Delta} f\|_{L^2(\mathbb{R}^d)} \lesssim t^{-\frac{d}{4}}\|f\|_{\dot{B}^0_{1,\infty}(\mathbb{R}^d)}.$$

Proof. Thanks to the interpolation inequality

$$\|g\|_{L^2} \lesssim \|g\|^{\frac{1}{2}}_{\dot{B}^{-1}_{2,\infty}}\|g\|^{\frac{1}{2}}_{\dot{B}^{1}_{2,\infty}},$$

we have

$$\|e^{t\Delta} f\|_{L^2} \lesssim \|e^{t\Delta} f\|^{\frac{1}{2}}_{\dot{B}^{-1}_{2,\infty}} \|e^{t\Delta} f\|^{\frac{1}{2}}_{\dot{B}^{1}_{2,\infty}}.$$

So the aim reduces to the estimate of $\|e^{t\Delta} f\|_{\dot{B}^{a}_{2,\infty}}$, $a = -1, 1$. By Bernstein's inequality, for all $a \geq -\frac{3}{2}$, we have

$$\begin{aligned} \|e^{t\Delta} f\|_{\dot{B}^{a}_{2,\infty}} &\lesssim \sup_{j\in\mathbb{Z}} 2^{aj} \|\Delta_j e^{t\Delta} f\|_{L^2} \lesssim \sup_{j\in\mathbb{Z}} 2^{aj} e^{-2^{2j-1}t} \|\Delta_j f\|_{L^2} \\ &\leq \sup_{j\in\mathbb{Z}} 2^{\left(a+\frac{d}{2}\right)j} e^{-2^{2j-1}t} \|\Delta_j f\|_{L^1} \lesssim t^{-\left(\frac{a}{2}+\frac{d}{4}\right)} \|f\|_{\dot{B}^{0}_{1,\infty}}. \end{aligned} \tag{2-4}$$

Setting $a = -1$ and $a = 1$ in (2-4), respectively, we can conclude the proof. □

3. Spectral analysis

In this section, we give the integral representation of the solution to (1-1). In fact, we repeat the procedure in Section 3 of [Wan 2017b]. We use v_i to stand for the i-th component of the vector v in the following context. Denote

$$\mathcal{A} \triangleq \mathbb{P}\,\mathrm{div}\tau.$$

Let us investigate the spectrum properties of the following system:

$$\begin{cases} \partial_t u_i - \Delta u_i = \mathcal{A}_i + G_i, \\ \partial_t \mathcal{A}_i = \frac{1}{2}\Delta u_i + F_i + H_i, \\ G = -\mathbb{P}(u\cdot\nabla u), \quad F = -\mathbb{P}\,\mathrm{div}(u\cdot\nabla\tau), \\ H = -\mathbb{P}\,\mathrm{div}(Q(\tau,\nabla u)), \quad i = 1,2,3, \end{cases} \tag{3-1}$$

where $\mathbb{P}$ is the Leray operator. Denote

$$A \triangleq \begin{pmatrix} -|\xi|^2 & 1 \\ -\frac{|\xi|^2}{2} & 0 \end{pmatrix},$$

then the eigenvalues of the matrix A can be given as

$$\lambda_{\pm} = \begin{cases} \frac{1}{2}(-|\xi|^2 \pm i|\xi|\sqrt{2-|\xi|^2}) & \text{when } |\xi| < \sqrt{2}, \\ \frac{1}{2}(-|\xi|^2 \pm |\xi|\sqrt{|\xi|^2-2}) & \text{when } |\xi| \geq \sqrt{2}, \end{cases}$$

where $i = \sqrt{-1}$. After Fourier transform, (3-1) reduces to

$$\partial_t \begin{pmatrix} \widehat{u_i} \\ \widehat{\mathcal{A}_i} \end{pmatrix}(\xi) = A \begin{pmatrix} \widehat{u_i} \\ \widehat{\mathcal{A}_i} \end{pmatrix}(\xi) + \begin{pmatrix} \widehat{G_i} \\ \widehat{F_i} + \widehat{H_i} \end{pmatrix}(\xi), \tag{3-2}$$

By using the standard method of diagonalization via the eigenvalues and eigenvectors, we can get from (3-2) that

$$
\begin{aligned}
u(t,x) &= M_{11}(\partial,t)u_0(x) + M_{12}(\partial,t)\mathcal{A}_0(x) \\
&\qquad + \int_0^t M_{11}(\partial,t-s)G\,ds + \int_0^t M_{12}(\partial,t-s)(F+H)\,ds \\
\mathcal{A}(t,x) &= M_{21}(\partial,t)u_0(x) + M_{22}(\partial,t)\mathcal{A}_0(x) \\
&\qquad + \int_0^t M_{21}(\partial,t-s)G\,ds + \int_0^t M_{22}(\partial,t-s)(F+H)\,ds,
\end{aligned}
\tag{3-3}
$$

where

$$\widehat{M_{ij}f}(\xi,t) \triangleq \widehat{M_{ij}}(\xi,t)\widehat{f}(\xi), \quad (i,j) \in \{1,2\}^2,$$

and

$$\begin{pmatrix} \widehat{M_{11}}(\xi,t) & \widehat{M_{12}}(\xi,t) \\ \widehat{M_{21}}(\xi,t) & \widehat{M_{22}}(\xi,t) \end{pmatrix} \triangleq \begin{pmatrix} \frac{\lambda_+ e^{\lambda_+ t} - \lambda_- e^{\lambda_- t}}{\lambda_+ - \lambda_-} & \frac{e^{\lambda_+ t} - e^{\lambda_- t}}{\lambda_+ - \lambda_-} \\ \frac{|\xi|^2}{2}\frac{e^{\lambda_+ t} - e^{\lambda_- t}}{\lambda_- - \lambda_+} & \frac{\lambda_- e^{\lambda_+ t} - \lambda_+ e^{\lambda_- t}}{\lambda_- - \lambda_+} \end{pmatrix}.$$

To bound $M_{ij}(\partial,t)$, we split the whole space $\mathbb{R}^3$ into the following four regions:

$$
\begin{aligned}
D_1 &\triangleq \{\xi \in \mathbb{R}^3 : |\xi| < 1\}, \\
D_2 &\triangleq \{\xi \in \mathbb{R}^3 : 1 \le |\xi| < \sqrt{2}\}, \\
D_3 &\triangleq \{\xi \in \mathbb{R}^3 : \sqrt{2} \le |\xi| < 2\}, \\
D_4 &\triangleq \{\xi \in \mathbb{R}^3 : |\xi| \ge 2\}.
\end{aligned}
$$

Let us keep the fact that $|\xi| \approx 1$ when $\xi \in D_2 \cup D_3$ in mind. Next, a proposition devoted to the estimates of $\widehat{M_{ij}}(\xi,t)$ is given as follows.

Proposition 3.1 [Wan 2017b]. *For all $(i,j) \in \{1,2\}^2$, $\widehat{M_{ij}}(\xi,t)$ satisfies the following estimates*:

(1) *When $\xi \in D_1$*,

$$
\begin{aligned}
|\widehat{M_{11}}(\xi,t)| &\lesssim e^{-\frac{|\xi|^2}{2}t}, & |\widehat{M_{12}}(\xi,t)| &\lesssim |\xi|^{-1}e^{-\frac{|\xi|^2}{2}t}, \\
|\widehat{M_{21}}(\xi,t)| &\lesssim |\xi|e^{-\frac{|\xi|^2}{2}t}, & |\widehat{M_{22}}(\xi,t)| &\lesssim e^{-\frac{|\xi|^2}{2}t}.
\end{aligned}
\tag{3-4}
$$

(2) *When $\xi \in D_2$*,

$$
\begin{aligned}
|\widehat{M_{11}}(\xi,t)| &\lesssim e^{-\frac{|\xi|^2}{4}t}, & |\widehat{M_{12}}(\xi,t)| &\lesssim |\xi|^{-1}e^{-\frac{|\xi|^2}{4}t}, \\
|\widehat{M_{21}}(\xi,t)| &\lesssim |\xi|e^{-\frac{|\xi|^2}{4}t}, & |\widehat{M_{22}}(\xi,t)| &\lesssim e^{-\frac{|\xi|^2}{4}t}.
\end{aligned}
$$

(3) *When* $\xi \in D_3$,

$$|\widehat{M_{11}}(\xi,t)| \lesssim e^{-\frac{|\xi|^2}{16}t}, \qquad |\widehat{M_{12}}(\xi,t)| \lesssim |\xi|^{-1}e^{-\frac{|\xi|^2}{16}t},$$
$$|\widehat{M_{21}}(\xi,t)| \lesssim |\xi|e^{-\frac{|\xi|^2}{16}t}, \quad |\widehat{M_{22}}(\xi,t)| \lesssim e^{-\frac{|\xi|^2}{16}t}.$$

(4) *When* $\xi \in D_4$,

$$\begin{aligned} &|\widehat{M_{11}}(\xi,t)| \lesssim e^{-\frac{t}{2}}, \quad |\widehat{M_{12}}(\xi,t)| \lesssim |\xi|^{-2}e^{-\frac{t}{2}}, \\ &|\widehat{M_{21}}(\xi,t)| \lesssim e^{-\frac{t}{2}}, \quad |\widehat{M_{22}}(\xi,t)| \lesssim e^{-\frac{t}{2}}. \end{aligned} \tag{3-5}$$

4. The estimate of $\|u(t)\|_{H^1}$

We will give the estimate of $\|u(t)\|_{L^2}$ and $\|\nabla u(t)\|_{L^2}$ in order.

The estimate of $\|u(t)\|_{L^2}$. Thanks to $\|\hat{f}\|_{L^2} = \|f\|_{L^2}$, we have

$$\|u(t)\|_{L^2} = \|\hat{u}(t)\|_{L^2} = \sum_{i=1}^{4} \|\hat{u}(t)\|_{L^2(D_i)},$$

which is sufficient for us to estimate the four terms on the right-hand side.

4A. ***The estimate of*** $\|\hat{u}(t)\|_{L^2(\mathbb{R}^3\setminus D_4)}$. Using (3-3), we can get

$$\begin{aligned} \|\hat{u}(t)\|_{L^2(D_1)} \le \|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^2(D_1)} &+ \|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_1)} \\ &+ \int_0^t \|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^2(D_1)}\,ds + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^2(D_1)}\,ds \\ &+ \int_0^t \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^2(D_1)}\,ds \triangleq I_1 + I_2 + \cdots + I_5. \end{aligned}$$

For I_1, using (3-4) and the estimate of the solution for the heat equation given by

$$\|e^{t\Delta} f\|_{L^2} \lesssim t^{-\frac{3}{4}}\|f\|_{L^1}, \tag{4-1}$$

we get

$$I_1 \lesssim \|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^2(D_1)} \lesssim \|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^2} = \|e^{\frac{1}{2}\Delta t}u_0\|_{L^2} \lesssim t^{-\frac{3}{4}}\|u_0\|_{L^1}.$$

Together with $I_1 \lesssim \|u_0\|_{L^2}$, this yields

$$I_1 \lesssim \langle t\rangle^{-\frac{3}{4}}\|u_0\|_{L^1\cap L^2}.$$

A similar method leads to

$$I_2 \lesssim \langle t\rangle^{-\frac{3}{4}}\|\tau_0\|_{L^1\cap L^2}.$$

Before we proceed, note that

$$\|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^2(D_1)} \le \|\widehat{M_{11}}(t-s)\widehat{u\cdot\nabla u}\|_{L^2(D_1)}.$$

For I_3, using (3-4) and (4-1), we have

$$I_3 \lesssim \int_0^t (t-s)^{-\frac{3}{4}} \|u\cdot\nabla u\|_{L^1}\,ds,$$

which, with the estimate $I_3 \lesssim \int_0^t \|u\cdot\nabla u\|_{L^2}\,ds$, implies

$$\begin{aligned}
I_3 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{4}} (\|u\cdot\nabla u\|_{L^1} + \|u\cdot\nabla u\|_{L^2})\,ds \\
&\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{4}} \|u\|_{L^2}\|\nabla u\|_{L^2\cap L^\infty}\,ds \\
&\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{4}} (\langle s\rangle^{-\frac{11}{4}} + \langle s\rangle^{-2})\,ds\; X(t)Y(t) \\
&\lesssim \langle t\rangle^{-\frac{3}{4}} X(t)Y(t).
\end{aligned}$$

For I_4, using

$$\mathbb{P}\,\mathrm{div}(u\cdot\nabla\tau) = \sum_{i=1,2,3} \partial_i\,\mathbb{P}\,\mathrm{div}(u_i\tau)$$

and (3-4), we get

$$I_4 \lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}|\xi||\widehat{u\otimes\tau}|\|_{L^2(D_1)}\,ds.$$

Using the spherical coordinate system

$$\xi_1 = r\sin\varphi\cos\theta, \quad \xi_2 = r\sin\varphi\sin\theta, \quad \xi_3 = r\cos\varphi,$$

we can obtain

$$\begin{aligned}
\|e^{-\frac{|\xi|^2}{2}t}|\xi|\widehat{f}\|_{L^2(D_1)} &= \left(\int_{|\xi|\le 1} e^{-|\xi|^2 t}|\xi|^2|\widehat{f}|^2\,d\xi\right)^{\frac{1}{2}} \\
&= \left(\int_0^{2\pi} d\theta \int_0^{\pi} d\varphi \int_0^1 e^{-r^2 t} r^4 \sin\varphi |\widehat{f}(r,\theta,\varphi)|\,dr\right)^{\frac{1}{2}} \\
&\lesssim \left(\int_0^1 e^{-r^2 t} r^4 dr\right)^{\frac{1}{2}} \|\widehat{f}\|_{L^\infty} \\
&\lesssim t^{-\frac{5}{4}} \|f\|_{L^1} \left(\int_0^{\sqrt{t}} e^{-r^2} r^4\,dr\right)^{\frac{1}{2}} \\
&\lesssim t^{-\frac{5}{4}} \|f\|_{L^1}.
\end{aligned}$$

This yields

$$I_4 \lesssim \int_0^t (t-s)^{-\frac{5}{4}} \|u \otimes \tau\|_{L^1}\, ds \lesssim \int_0^t (t-s)^{-\frac{5}{4}} \|u\|_{L^2} \|\tau\|_{L^2}\, ds.$$

Direct computations lead to

$$I_4 \lesssim \int_0^t \|\widehat{u \otimes \tau}\|_{L^2(D_1)}\, ds \lesssim \int_0^t \|u \otimes \tau\|_{L^2}\, ds.$$

So

$$\begin{aligned} I_4 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} \|u\|_{L^2\cap L^\infty} \|\tau\|_{L^2}\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} (\langle s\rangle^{-\frac{3}{4}} + \langle s\rangle^{-\frac{3}{2}})\, ds (X(t)+Y(t)) \|\tau\|_{L^\infty_t(L^2)} \\ &\lesssim \langle t\rangle^{-\frac{3}{4}} (X(t)+Y(t)) \|\tau\|_{L^\infty_t(L^2)}. \end{aligned}$$

For I_5, using (3-4) and (4-1), we have

$$\begin{aligned} I_5 &\lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^2(D_1)}\, ds \lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^2}\, ds \\ &\lesssim \int_0^t \|e^{\frac{1}{2}\Delta(t-s)} Q(\tau,\nabla u)\|_{L^2}\, ds \lesssim \int_0^t (t-s)^{-\frac{3}{4}} \|Q(\tau,\nabla u)\|_{L^1}\, ds \\ &\lesssim \int_0^t (t-s)^{-\frac{3}{4}} \|\tau\|_{L^2} \|\nabla u\|_{L^2}\, ds. \end{aligned}$$

We also have

$$I_5 \lesssim \int_0^t \|Q(\tau,\nabla u)\|_{L^2}\, ds \lesssim \int_0^t \|\tau\|_{L^2} \|\nabla u\|_{L^\infty}\, ds,$$

so

$$\begin{aligned} I_5 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{4}} \|\tau\|_{L^2} \|\nabla u\|_{L^2\cap L^\infty}\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{4}} (\langle s\rangle^{-\frac{5}{4}} + \langle s\rangle^{-2})\, ds (X(t)+Y(t)) \|\tau\|_{L^\infty_t(L^2)} \\ &\lesssim \langle t\rangle^{-\frac{3}{4}} (X(t)+Y(t)) \|\tau\|_{L^\infty_t(L^2)}. \end{aligned}$$

Collecting the above estimates of I_i yields

$$\begin{aligned} \|\hat{u}\|_{L^2(D_1)} \lesssim \langle t\rangle^{-\frac{3}{4}} \|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{3}{4}} X(t)Y(t) \\ + \langle t\rangle^{-\frac{3}{4}} (X(t)+Y(t)) \|\tau\|_{L^\infty_t(L^2)}. \end{aligned} \tag{4-2}$$

Since the estimates of $\widehat{M_{ij}}$ in $D_2 \cup D_3$ are similar to that in D_1, repeating the above procedure yields

$$\sum_{i=2,3} \|\hat{u}\|_{L^2(D_i)} \lesssim \langle t\rangle^{-\frac{3}{4}} \|(u_0, \tau_0)\|_{L^1 \cap L^2} + \langle t\rangle^{-\frac{3}{4}} X(t)Y(t) + \langle t\rangle^{-\frac{3}{4}} (X(t) + Y(t)) \|\tau\|_{L^\infty_t(L^2)}. \tag{4-3}$$

4B. *The estimate of $\|\hat{u}(t)\|_{L^2(D_4)}$.* For the estimate in D_4, we will use (3-5) in the following context.

$$\begin{aligned}
\|\hat{u}(t)\|_{L^2(D_4)} &\le \|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^2(D_4)} + \|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_4)} \\
&\quad + \int_0^t \|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^2(D_4)}\, ds + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^2(D_4)}\, ds \\
&\quad + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^2(D_4)}\, ds \\
&= I_1' + I_2' + \cdots + I_5'.
\end{aligned}$$

We have

$$I_1' + I_2' \lesssim e^{-\frac{t}{2}} \|(u_0, \tau_0)\|_{L^2}.$$

For I_3', we have

$$\begin{aligned}
I_3' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|u \cdot \nabla u\|_{L^2}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|u\|_{L^2} \|\nabla u\|_{L^\infty}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-\frac{11}{4}}\, ds\; X(t)Y(t) \\
&\lesssim \langle t\rangle^{-\frac{11}{4}} X(t)Y(t).
\end{aligned}$$

For I_4', we have

$$\begin{aligned}
I_4' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|u \otimes \tau\|_{L^2}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|u\|_{L^\infty} \|\tau\|_{L^2}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-\frac{3}{2}}\, ds\; X(t) \|\tau\|_{L^\infty_t L^2} \\
&\lesssim \langle t\rangle^{-\frac{3}{2}} X(t) \|\tau\|_{L^\infty_t L^2}.
\end{aligned}$$

For I_5', we get

$$\begin{aligned}
I_5' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|Q(\tau, \nabla u)\|_{L^2}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|\nabla u\|_{L^\infty} \|\tau\|_{L^2}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\, ds\; Y(t) \|\tau\|_{L^\infty_t L^2} \\
&\lesssim \langle t\rangle^{-2} Y(t) \|\tau\|_{L^\infty_t L^2}.
\end{aligned}$$

Thus we can infer

$$\|\widehat{u}(t)\|_{L^2(D_4)} \lesssim e^{-\frac{t}{2}}\|(u_0,\tau_0)\|_{L^2}+\langle t\rangle^{-\frac{3}{2}}\big(X(t)Y(t)+(X(t)+Y(t))\|\tau\|_{L^\infty_t L^2}\big). \tag{4-4}$$

Combining (4-2), (4-3) and (4-4) implies that

$$\|u(t)\|_{L^2} \lesssim \langle t\rangle^{-\frac{3}{4}}\|(u_0,\tau_0)\|_{L^1\cap L^2}+\langle t\rangle^{-\frac{3}{4}}X(t)Y(t) +\langle t\rangle^{-\frac{3}{4}}(X(t)+Y(t))\|\tau\|_{L^\infty_t(L^2)}. \tag{4-5}$$

The estimate of $\|\nabla u(t)\|_{L^2}$. Before we begin this estimate, let us introduce a lemma dealing with a commutator estimate, which plays an important role in the following proof.

Lemma 4.1. *Let* $i=1,2,3$, *then*

$$\|[\mathbb{P}\,\mathrm{div}, u_i]\tau\|_{\dot{B}^0_{1,\infty}} \lesssim (\|\nabla u_i\|_{L^2}\|\tau\|_{L^2}+\|u_i\|_{L^2}\|\mathcal{A}\|_{L^2}).$$

The proof of this lemma will be postponed until the Appendix. As in the previous procedure, we have

$$\|\nabla u(t)\|_{L^2}=\|\widehat{\nabla u}(t)\|_{L^2}\lesssim\sum_{i=1}^{4}\||\xi|\widehat{u}\|_{L^2(D_i)}.$$

4C. ***The estimate of*** $\||\xi|\hat{u}(t)\|_{L^2(\mathbb{R}^3\setminus D_4)}$. Using (3-3), we have

$$\begin{aligned}\||\xi|\widehat{u}(t)\|_{L^2(D_1)} &\le \|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^2(D_1)}+\|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_1)}\\ &\quad+\int_0^t\|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^2(D_1)}\,ds+\int_0^t\|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^2(D_1)}\,ds\\ &\quad+\int_0^t\|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^2(D_1)}\,ds\\ &=J_1+J_2+\cdots+J_5.\end{aligned}$$

Applying (3-4) and (4-1), we get

$$J_1\lesssim\||\xi|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^2(D_1)}\lesssim t^{-\frac{1}{2}}\|e^{-\frac{|\xi|^2}{4}t}\widehat{u_0}\|_{L^2}\lesssim t^{-\frac{1}{2}}\|e^{\frac{1}{4}t\Delta}u_0\|_{L^2}\lesssim t^{-\frac{5}{4}}\|u_0\|_{L^1}.$$

With another estimate

$$J_1\lesssim\||\xi|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^2(D_1)}\lesssim\|\widehat{u_0}\|_{L^2}\lesssim\|u_0\|_{L^2},$$

we infer

$$J_1\lesssim\langle t\rangle^{-\frac{5}{4}}\|u_0\|_{L^1\cap L^2}.$$

A similar process leads to

$$J_2 \lesssim \langle t\rangle^{-\frac{5}{4}} \|\tau_0\|_{L^1\cap L^2}.$$

Getting the estimate of J_1 in the same way, we can infer that

$$\begin{aligned} J_3 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} (\|u\cdot\nabla u\|_{L^1} + \|u\cdot\nabla u\|_{L^2})\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} \|u\|_{L^2} (\|\nabla u\|_{L^2} + \|\nabla u\|_{L^\infty})\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} (\langle s\rangle^{-2} + \langle s\rangle^{-\frac{11}{4}})\, ds X(t)Y(t) \lesssim \langle t\rangle^{-\frac{5}{4}} X(t)Y(t) \end{aligned}$$

and

$$\begin{aligned} J_5 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} (\|Q(\tau,\nabla u)\|_{L^1} + \|Q(\tau,\nabla u)\|_{L^2})\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} \|\nabla u\|_{L^2} (\|\tau\|_{L^2} + \|\tau\|_{L^\infty})\, ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}} \langle s\rangle^{-\frac{5}{4}}\, ds Y(t) \|\tau\|_{L^\infty_t(L^2\cap L^\infty)} \lesssim \langle t\rangle^{-\frac{5}{4}} Y(t) \|\tau\|_{L^\infty_t(L^2\cap L^\infty)}. \end{aligned}$$

We shall use a different approach to get the estimates of J_4. Otherwise, a bad term $\|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^1}$ will appear when we use the standard estimate of the solution for the heat equation (4-1). It is known that this bad term cannot be bounded by a standard commutator estimate. Our idea uses Lemmas 2.6 and 4.1. Now, we begin the estimate of J_4. Thanks to (3-4) again, and using

$$\mathbb{P}\operatorname{div}(u\cdot\nabla\tau) = \sum_{i=1,2,3} \partial_i \mathbb{P}\operatorname{div}(u_i\tau) = \sum_{i=1,2,3} (\partial_i [\mathbb{P}\operatorname{div}, u_i]\tau + \partial_i(u_i\mathcal{A})), \tag{4-6}$$

we have

$$\begin{aligned} J_4 &\lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{\mathbb{P}\operatorname{div}(u\cdot\nabla\tau)}\|_{L^2(D_1)}\, ds \\ &\lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} |\xi| \widehat{u\otimes\mathcal{A}}\|_{L^2(D_1)}\, ds \\ &\qquad + \underbrace{\sum_{i=1,2,3} \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} |\xi_i| \widehat{[\mathbb{P}\operatorname{div}, u_i]\tau}\|_{L^2(D_1)}\, ds}_{\Upsilon_1}. \end{aligned}$$

By the previous approach, we can get

$$\int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)} |\xi| \widehat{u\otimes\mathcal{A}}\|_{L^2(D_1)}\, ds \lesssim \int_0^t (t-s)^{-\frac{5}{4}} \|u\otimes\mathcal{A}\|_{L^1}\, ds,$$

and we also have

$$\int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}|\xi|\widehat{u\otimes\mathcal{A}}\|_{L^2(D_1)}\,ds \lesssim \int_0^t \|u\otimes\mathcal{A}\|_{L^2}\,ds.$$

So we can get

$$\begin{aligned}\int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}|\xi|\widehat{u\otimes\mathcal{A}}\|_{L^2(D_1)}\,ds&\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}(\|u\otimes\mathcal{A}\|_{L^1}+\|u\otimes\mathcal{A}\|_{L^2})\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}\|\mathcal{A}\|_{L^2}(\|u\|_{L^2}+\|u\|_{L^\infty})\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}(\langle s\rangle^{-2}+\langle s\rangle^{-\frac{11}{4}})\,dsY(t)(X(t)+Y(t))\\ &\lesssim \langle t\rangle^{-\frac{5}{4}}Y(t)(X(t)+Y(t)).\end{aligned}$$

As for Υ_1, by Lemmas 2.6 and 4.1, we have

$$\begin{aligned}\Upsilon_1 &\lesssim \int_0^t (t-s)^{-\frac{1}{2}}\|e^{-\frac{|\xi|^2}{4}(t-s)}\widehat{[\mathbb{P}\operatorname{div},u_i]\tau}\|_{L^2(D_1)}\,ds\\ &\lesssim \int_0^t (t-s)^{-\frac{1}{2}}\|e^{-\frac{1}{4}(t-s)\Delta}[\mathbb{P}\operatorname{div},u_i]\tau\|_{L^2}\,ds\\ &\lesssim \int_0^t (t-s)^{-\frac{5}{4}}\|[\mathbb{P}\operatorname{div},u_i]\tau\|_{\dot{B}^0_{1,\infty}}\,ds\\ &\lesssim \int_0^t (t-s)^{-\frac{5}{4}}(\|\nabla u\|_{L^2}\|\tau\|_{L^2}+\|u\|_{L^2}\|\mathcal{A}\|_{L^2})\,ds.\end{aligned}$$

By using (2-3), we can also bound Υ_1 as follows:

$$\Upsilon_1 \lesssim \int_0^t \|[\mathbb{P}\operatorname{div},u_i]\tau\|_{L^2}\,ds \lesssim \int_0^t (\|\nabla u\|_{L^\infty}\|\tau\|_{L^2}+\|\tau\|_{L^\infty}\|\nabla u\|_{L^2})\,ds.$$

Finally, we infer

$$\begin{aligned}\Upsilon_1 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}(\|\nabla u\|_{L^2}\|\tau\|_{L^2}+\|u\|_{L^2}\|\mathcal{A}\|_{L^2}\\ &\qquad\qquad +\|\nabla u\|_{L^\infty}\|\tau\|_{L^2}+\|\tau\|_{L^\infty}\|\nabla u\|_{L^2})\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}\langle s\rangle^{-\frac{5}{4}}\,ds\; Y(t)(X(t)+\|\tau\|_{L^\infty_t(L^2\cap L^\infty)})\\ &\lesssim \langle t\rangle^{-\frac{5}{4}}Y(t)(X(t)+\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}).\end{aligned}$$

Thus

$$J_4 \lesssim \langle t\rangle^{-\frac{5}{4}}Y(t)(X(t)+Y(t)+\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}).$$

Collecting the five estimates above, we get

$$\||\xi|\widehat{u}(t)\|_{L^2(D_1)} \lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{5}{4}}Y(t) \times\{X(t)+Y(t)+\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}\}. \tag{4-7}$$

Similarly,

$$\sum_{i=2,3}\||\xi|\widehat{u}(t)\|_{L^2(D_i)} \lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{5}{4}}Y(t) \times\{X(t)+Y(t)+\|\tau\|_{L^\infty_t(H^2)}\}. \tag{4-8}$$

4D. ***The estimate of*** $\boldsymbol{\||\xi|\hat{u}(t)\|_{L^2(D_4)}}$**.** As in the previous statement, we use (3-5) for the estimate in D_4:

$$\begin{aligned}\||\xi|\widehat{u}(t)\|_{L^2(D_4)} \le \||\xi|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^2(D_4)}&+\||\xi|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_4)}\\
&+\int_0^t\||\xi|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^2(D_4)}\,ds+\int_0^t\||\xi|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^2(D_4)}\,ds\\
&+\int_0^t\||\xi|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^2(D_4)}\,ds\\
&=J_1'+J_2'+\cdots+J_5'.\end{aligned}$$

For J_1' and J_2', we have

$$J_1'+J_2' \lesssim e^{-\frac{t}{2}}(\|\nabla u_0\|_{L^2}+\|\nabla\tau_0\|_{L^2}).$$

For J_3', using product estimate (2-2), we get

$$\begin{aligned}J_3' &\lesssim \int_0^t e^{-\frac{t-s}{2}}\|\nabla(u\cdot\nabla u)\|_{L^2}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\otimes u\|_{\dot{H}^2}\,ds\\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\|_{L^\infty}\|u\|_{\dot{H}^2}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}}\langle s\rangle^{-\frac{3}{2}}\,ds\; X(t)\|u\|_{L^\infty_t\dot{H}^2}\\
&\lesssim \langle t\rangle^{-\frac{3}{2}}X(t)\|u\|_{L^\infty_t\dot{H}^2}.\end{aligned}$$

For J_4', we have

$$\begin{aligned}J_4' &\lesssim \int_0^t e^{-\frac{t-s}{2}}\||\xi|^{-2}\widehat{\mathbb{P}\,\mathrm{div}(u\cdot\nabla\tau)}\|_{L^2(D_4)}\,ds\\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}\||\xi|^{-2}\widehat{\mathbb{P}\,\mathrm{div}(u\cdot\nabla\tau)}\|_{L^2}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\otimes\tau\|_{L^2}\,ds\\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\|_{L^\infty}\|\tau\|_{L^2}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}}\langle s\rangle^{-\frac{3}{2}}\,ds X(t)\|\tau\|_{L^\infty_t L^2}\\
&\lesssim \langle t\rangle^{-\frac{3}{2}}X(t)\|\tau\|_{L^\infty_t L^2}.\end{aligned}$$

The last term J_5' can be bounded as follows:

$$\begin{aligned}J_5' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \||\xi|^{-2}\widehat{\mathbb{P}\operatorname{div} Q(\tau,\nabla u)}\|_{L^2(D_4)}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|Q(\tau,\nabla u)\|_{L^2}\,ds\\ &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\tau\|_{L^2}\|\nabla u\|_{L^\infty}\,ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\,ds Y(t)\|\tau\|_{L^\infty_t L^2}\\ &\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2}.\end{aligned}$$

Combining the five estimates above can yield

$$\begin{aligned}\||\xi|\hat{u}(t)\|_{L^2(D_4)} &\lesssim e^{-\frac{t}{2}}\|(u_0,\tau_0)\|_{H^1}\\ &\quad + \langle t\rangle^{-\frac{3}{2}} X(t)\|(u,\tau)\|_{L^\infty_t H^2} + \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2}.\end{aligned} \tag{4-9}$$

Thanks to (4-7), (4-8) and (4-9), we can get

$$\begin{aligned}\||\xi|\hat{u}(t)\|_{L^2} &\lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap H^2} + \langle t\rangle^{-\frac{5}{4}}(X(t)+Y(t))\\ &\quad \times\{X(t)+Y(t)+\|(u,\tau)\|_{L^\infty_t H^2}\}.\end{aligned} \tag{4-10}$$

5. The estimate of $\|\hat{u}(t)\|_{L^1}$, $\||\xi|\hat{u}(t)\|_{L^1}$ and $\|\mathcal{A}(t)\|_{L^2}$

We will bound $\|\hat{u}\|_{L^1}$ and $\||\xi|\hat{u}\|_{L^1}$ in order.

The estimate of $\|\hat{u}(t)\|_{L^1}$. As for the previous process, we have

$$\|\hat{u}(t)\|_{L^1} \le \sum_{i=1}^4 \|\hat{u}(t)\|_{L^1(D_i)}.$$

5A. *The estimate of* $\|\hat{u}\|_{L^1(\mathbb{R}^3\setminus D_4)}$. Thanks to (3-3), we have

$$\begin{aligned}\|\hat{u}(t)\|_{L^1(D_1)} &\le \|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^1(D_1)} + \|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^1(D_1)}\\ &\quad + \int_0^t \|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^1(D_1)}\,ds + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^1(D_1)}\,ds\\ &\quad + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\,ds\\ &= K_1+K_2+\cdots+K_5.\end{aligned}$$

We will frequently use the estimate

$$\begin{aligned}\|e^{-b|\xi|^2 t}\|_{L^2(D_1)} &= \left(\int_{|\xi|\le 1} e^{-2b|\xi|^2 t}\,d\xi\right)^{\frac{1}{2}}\\ &= t^{-\frac{3}{4}}\left(\int_{|\xi|\le\sqrt{t}} e^{-2bv^2}\,dv\right)^{\frac{1}{2}} \lesssim t^{-\frac{3}{4}} \quad \text{for all } b>0.\end{aligned} \tag{5-1}$$

For K_1, using (3-4), (5-1) and (4-1), we have

$$\begin{aligned}K_1 &\lesssim \|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^1(D_1)} \lesssim \|e^{-\frac{|\xi|^2}{4}t}\|_{L^2(D_1)}\|e^{-\frac{|\xi|^2}{4}t}\widehat{u_0}\|_{L^2(D_1)}\\ &\lesssim t^{-\frac{3}{4}}\|e^{\frac{t}{4}\Delta}u_0\|_{L^2} \lesssim t^{-\frac{3}{2}}\|u_0\|_{L^1},\end{aligned}$$

and we also have

$$K_1 \lesssim \|\widehat{u_0}\|_{L^1(D_1)} \lesssim \|u_0\|_{L^2},$$

so

$$K_1 \lesssim \langle t\rangle^{-\frac{3}{2}}\|u_0\|_{L^1\cap L^2}.$$

Similarly,

$$K_2 \lesssim \langle t\rangle^{-\frac{3}{2}}\|\tau_0\|_{L^1\cap L^2}.$$

For K_3, using (3-4), (5-1) and (4-1) again, we can obtain

$$\begin{aligned}K_3 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}(\|u\cdot\nabla u\|_{L^1}+\|u\cdot\nabla u\|_{L^2})\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}\|u\|_{L^2}(\|\nabla u\|_{L^2}+\|\nabla u\|_{L^\infty})\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}(\langle s\rangle^{-\frac{11}{4}}+\langle s\rangle^{-2})\,ds\,X(t)Y(t) \lesssim \langle t\rangle^{-\frac{3}{2}}X(t)Y(t).\end{aligned} \tag{5-2}$$

Thanks to (4-6), this implies that

$$K_4 \le \mathcal{K}_{41}+\mathcal{K}_{42},$$

where

$$\mathcal{K}_{41} = \sum_{i=1,2,3}\int_0^t \||\xi_i|\widehat{M_{12}}(t-s)\widehat{u_i\mathcal{A}}\|_{L^1(D_1)}\,ds$$

and

$$\mathcal{K}_{42} = \sum_{i=1,2,3}\int_0^t \||\xi_i|\widehat{M_{12}}(t-s)\widehat{[\mathbb{P}\operatorname{div},u_i]\tau}\|_{L^1(D_1)}\,ds.$$

In a similar way, we deduce

$$\begin{aligned}\mathcal{K}_{41} &\lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}\widehat{u_i\mathcal{A}}\|_{L^1(D_1)}\,ds \lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}(\|u\otimes\mathcal{A}\|_{L^1}+\|u\otimes\mathcal{A}\|_{L^2})\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}(\|u\|_{L^2}+\|u\|_{L^\infty})\|\mathcal{A}\|_{L^2}\,ds\\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{3}{2}}\langle s\rangle^{-2}\,ds\,X(t)Y(t)\\ &\lesssim \langle t\rangle^{-\frac{3}{2}}X(t)Y(t).\end{aligned}$$

In the following context, we shall use a new approach to bound $\mathcal{K}_{42}$ and the last term K_5, due to the fact that τ does not have the decay property like $\mathcal{A}$.

Remark 5.1. If we follow the procedure yielding the estimate of K_3 and $\mathcal{K}_{41}$, we will face an integral of the following type:

$$\int_0^t \langle t-s\rangle^{-\frac{3}{2}} \|\nabla u\|_{L^2} \|\tau\|_{L^2}\, ds,$$

which yields a weaker decay rate $\langle t\rangle^{-\frac{5}{4}}$ than $\langle t\rangle^{-\frac{3}{2}}$, since we only have

$$\|\nabla u\|_{L^2} \lesssim \langle t\rangle^{-\frac{5}{4}} Y(t).$$

Now, we begin the estimate of $\mathcal{K}_{42}$. The new idea is splitting the integral interval $[0,t]$ into $[0,\frac{t}{2})$ and $[\frac{t}{2},t]$. That is

$$\mathcal{K}_{42} = K_{41} + K_{42},$$

where

$$K_{41} = \sum_{i=1,2,3} \int_0^{\frac{t}{2}} \| |\xi| \widehat{M_{12}}(t-s) \widehat{[\mathbb{P}\operatorname{div}, u_i]\tau} \|_{L^1(D_1)}\, ds$$

and

$$K_{42} = \sum_{i=1,2,3} \int_{\frac{t}{2}}^{t} \| |\xi| \widehat{M_{12}}(t-s) \widehat{[\mathbb{P}\operatorname{div}, u_i]\tau} \|_{L^1(D_1)}\, ds.$$

For K_{41}, using (3-4), (5-1) and Lemma 2.6, we have

$$\begin{aligned}
K_{41} &\lesssim \int_0^{\frac{t}{2}} \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{[\mathbb{P}\operatorname{div}, u_i]\tau}\|_{L^1(D_1)}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} \|e^{-\frac{|\xi|^2}{4}(t-s)}\|_{L^2(D_1)} \|e^{-\frac{|\xi|^2}{4}(t-s)} \widehat{[\mathbb{P}\operatorname{div}, u_i]\tau}\|_{L^2(D_1)}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} (t-s)^{-\frac{3}{4}} \|e^{\frac{t-s}{4}\Delta} [\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} (t-s)^{-\frac{3}{2}} \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{\dot{B}^0_{1,\infty}}\, ds,
\end{aligned}$$

and we can also get

$$K_{41} \lesssim \int_0^t \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\, ds.$$

Thus, by Lemma 4.1 and the commutator estimate, we have

$$\begin{aligned}
K_{41} &\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-\frac{3}{2}} (\|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{\dot{B}^0_{1,\infty}} + \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2})\,ds \\
&\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-\frac{3}{2}} (\|\nabla u\|_{L^2\cap L^\infty}\|\tau\|_{L^2\cap L^\infty} + \|u\|_{L^2}\|\mathcal{A}\|_{L^2})\,ds \\
&\lesssim \langle t\rangle^{-\frac{3}{2}} \int_0^{\frac{t}{2}} \langle s\rangle^{-\frac{5}{4}}\,ds\; Y(t)(\|\tau\|_{L^\infty_t H^2} + X(t)) \\
&\lesssim \langle t\rangle^{-\frac{3}{2}} Y(t)(\|\tau\|_{L^\infty_t H^2} + X(t)).
\end{aligned}$$

For K_{42}, applying (5-1), we can get

$$\begin{aligned}
K_{42} \lesssim \int_{\frac{t}{2}}^{t} (t-s)^{-\frac{3}{4}} \|e^{-\frac{|\xi|^2}{4}(t-s)}\widehat{[\mathbb{P}\operatorname{div}, u_i]\tau}\|_{L^2(D_1)}\,ds \\
\lesssim \int_{\frac{t}{2}}^{t} (t-s)^{-\frac{3}{4}} \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\,ds.
\end{aligned}$$

This, together with

$$K_{42} \lesssim \int_{\frac{t}{2}}^{t} \|e^{-\frac{|\xi|^2}{2}(t-s)}\widehat{[\mathbb{P}\operatorname{div}, u_i]\tau}\|_{L^2(D_1)}\,ds \lesssim \int_{\frac{t}{2}}^{t} \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\,ds$$

leads to

$$\begin{aligned}
K_{42} &\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}} \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\,ds \\
&\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}} \||\xi|\widehat{u}\|_{L^1}\|\tau\|_{L^2}\,ds \lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}}\langle s\rangle^{-2}\,ds Y(t)\|\tau\|_{L^\infty_t L^2} \\
&\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2} \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}}\,ds \\
&\lesssim \langle t\rangle^{-\frac{7}{4}} Y(t)\|\tau\|_{L^\infty_t L^2},
\end{aligned}$$

where we have used

$$\|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2} \lesssim \||\xi|\widehat{u}\|_{L^1}\|\tau\|_{L^2}, \tag{5-3}$$

which will be proved in the Appendix, and

$$\int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}}\,ds \lesssim \int_0^{\frac{t}{2}} \langle s\rangle^{-\frac{3}{4}}\,ds \lesssim \langle t\rangle^{-\frac{1}{4}}.$$

So we have

$$K_4 \lesssim \langle t\rangle^{-\frac{3}{2}} Y(t)(\|\tau\|_{L^\infty_t H^2} + X(t)).$$

In the following, we will use the same strategy to bound K_5. We have

$$K_5 = K_{51} + K_{52},$$

where

$$K_{51} = \int_0^{\frac{t}{2}} \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\, ds$$

and

$$K_{52} = \int_{\frac{t}{2}}^{t} \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\, ds.$$

By (3-4), (5-1) and (4-1), we get

$$\begin{aligned}
K_{51} &\lesssim \int_0^{\frac{t}{2}} \|e^{-\frac{|\xi|^2}{4}(t-s)}\|_{L^2(D_1)} \|e^{-\frac{|\xi|^2}{4}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^2(D_1)}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} (t-s)^{-\frac{3}{4}} \|e^{-\frac{|\xi|^2}{4}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^2(D_1)}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} (t-s)^{-\frac{3}{4}} \|e^{\frac{1}{4}(t-s)\Delta} Q(\tau,\nabla u)\|_{L^2}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} (t-s)^{-\frac{3}{2}} \|Q(\tau,\nabla u)\|_{L^1}\, ds.
\end{aligned}$$

We also have

$$\begin{aligned}
K_{51} &\lesssim \int_0^{\frac{t}{2}} \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^1(D_1)}\, ds \lesssim \int_0^{\frac{t}{2}} \|e^{-\frac{|\xi|^2}{2}(t-s)} \widehat{Q(\tau,\nabla u)}\|_{L^2}\, ds \\
&\lesssim \int_0^{\frac{t}{2}} \|Q(\tau,\nabla u)\|_{L^2}\, ds;
\end{aligned}$$

thus

$$\begin{aligned}
K_{51} &\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-\frac{3}{2}} \|\tau\|_{L^2}(\|\nabla u\|_{L^2} + \|\nabla u\|_{L^\infty})\, ds \\
&\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-\frac{3}{2}} (\langle s\rangle^{-\frac{5}{4}} + \langle s\rangle^{-2})\, ds\, Y(t)\|\tau\|_{L^\infty_t L^2} \lesssim \langle t\rangle^{-\frac{3}{2}} Y(t)\|\tau\|_{L^\infty_t L^2}.
\end{aligned}$$

For K_{52}, by (3-4) and (5-1) again, we get

$$K_{52} \lesssim \int_{\frac{t}{2}}^{t} (t-s)^{-\frac{3}{4}} \|Q(\tau,\nabla u)\|_{L^2}\, ds,$$

and, with the estimate

$$K_{52} \lesssim \int_{\frac{t}{2}}^{t} \|Q(\tau,\nabla u)\|_{L^2}\, ds,$$

we can infer

$$
\begin{aligned}
K_{52} &\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}} \|\nabla u\|_{L^\infty} \|\tau\|_{L^2}\, ds \lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}} \||\xi|\hat{u}\|_{L^1} \|\tau\|_{L^2}\, ds \\
&\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}} \langle s\rangle^{-2} ds Y(t) \|\tau\|_{L_t^\infty L^2} \lesssim \langle t\rangle^{-2} Y(t) \|\tau\|_{L_t^\infty L^2} \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{3}{4}}\, ds \\
&\lesssim \langle t\rangle^{-\frac{7}{4}} Y(t) \|\tau\|_{L_t^\infty L^2}.
\end{aligned}
$$

Thus we can deduce that

$$
K_5 \lesssim \langle t\rangle^{-\frac{3}{2}} Y(t) \|\tau\|_{L_t^\infty L^2}.
$$

Combining this with the estimates above implies

$$
\|\hat{u}(t)\|_{L^1(D_1)} \lesssim \langle t\rangle^{-\frac{3}{2}} \|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{3}{2}} Y(t)(\|\tau\|_{L_t^\infty H^2} + X(t)). \tag{5-4}
$$

Similarly,

$$
\begin{aligned}
\sum_{i=2,3} &\|\hat{u}(t)\|_{L^1(D_i)} \\
&\lesssim \langle t\rangle^{-\frac{3}{2}} \|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{3}{2}} Y(t)(\|\tau\|_{L_t^\infty H^2} + X(t)).
\end{aligned} \tag{5-5}
$$

5B. ***The estimate of*** $\|\hat{u}\|_{L^1(D_4)}$**.** Using (3-3), we have

$$
\begin{aligned}
\|\hat{u}(t)\|_{L^1(D_4)} \le\,& \|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^1(D_4)} + \|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^1(D_4)} \\
&+ \int_0^t \|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^1(D_4)}\, ds + \int_0^t \|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^1(D_4)}\, ds \\
&+ \int_0^t \|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_4)}\, ds \\
=\,& K_1' + K_2' + \cdots + K_5'.
\end{aligned}
$$

For K_1', using (3-4), we get

$$
K_1' \lesssim e^{-\frac{t}{2}} \|\widehat{u_0}\|_{L^1} \lesssim e^{-\frac{t}{2}} \|u_0\|_{H^2}.
$$

In the same way,

$$
K_2' \lesssim e^{-\frac{t}{2}} \|\widehat{\tau_0}\|_{L^1} \lesssim e^{-\frac{t}{2}} \|\tau_0\|_{H^2}.
$$

For the three nonlinear terms' estimates, applying (3-4) and Young's inequality, we have

$$
\begin{aligned}
K_3' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\widehat{u\cdot\nabla u}\|_{L^1}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{u}\|_{L^1} \||\xi|\hat{u}\|_{L^1}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-\frac{7}{2}}\, ds X(t) Y(t) \lesssim \langle t\rangle^{-\frac{7}{2}} X(t) Y(t),
\end{aligned}
$$

$$
\begin{aligned}
K_4' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{u}\|_{L^1} \|\hat{\tau}\|_{L^1}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-\frac{3}{2}}\, ds X(t) \|\tau\|_{L_t^\infty H^2} \\
&\lesssim \langle t\rangle^{-\frac{3}{2}} X(t) \|\tau\|_{L_t^\infty H^2}
\end{aligned}
$$

$$
\begin{aligned}
K_5' &\lesssim \int_0^t e^{-\frac{t-s}{2}} \||\xi|^{-1} \widehat{Q(\tau,\nabla u)}\|_{L^1(D_4)}\, ds \lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{\tau}\|_{L^1} \||\xi|\hat{u}\|_{L^1}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\, ds Y(t) \|\tau\|_{L_t^\infty H^2} \lesssim \langle t\rangle^{-2} Y(t) \|\tau\|_{L_t^\infty H^2}.
\end{aligned}
$$

Collecting the above estimates leads to

$$
\|\hat{u}(t)\|_{L^1(D_4)} \lesssim e^{-\frac{t}{2}} \|(u_0,\tau_0)\|_{H^2} + \langle t\rangle^{-\frac{3}{2}} (X(t)+Y(t))(\|\tau\|_{L_t^\infty H^2} + X(t)). \tag{5-6}
$$

Combining the estimates (5-4), (5-5) and (5-6), we can infer that

$$
\|\hat{u}(t)\|_{L^1} \lesssim \langle t\rangle^{-\frac{3}{2}} \|(u_0,\tau_0)\|_{L^1\cap H^2} + \langle t\rangle^{-\frac{3}{2}} (X(t)+Y(t))(\|\tau\|_{L_t^\infty H^2} + X(t)). \tag{5-7}
$$

<u>*The estimate of* $\||\xi|\hat{u}(t)\|_{L^1}$</u>. Thanks to

$$
\||\xi|\hat{u}(t)\|_{L^1} \lesssim \sum_{i=1}^4 \||\xi|\hat{u}(t)\|_{L^1(D_i)},
$$

it is sufficient to bound the four terms on the right-hand side.

5C. ***The estimate of*** $\boldsymbol{\||\xi|\hat{u}(t)\|_{L^1(\mathbb{R}^3\setminus D_4)}}$**.** Using (3-3), we get

$$
\begin{aligned}
\||\xi|\hat{u}(t)\|_{L^1(D_1)} &\le \||\xi|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^1(D_1)} + \||\xi|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^1(D_1)} \\
&\quad + \int_0^t \||\xi|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^1(D_1)}\, ds + \int_0^t \||\xi|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^1(D_1)}\, ds \\
&\quad + \int_0^t \||\xi|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\, ds \\
&= L_1 + L_2 + \cdots + L_5.
\end{aligned}
$$

For L_1, by (3-4), (5-1) and (4-1), we have

$$
\begin{aligned}
L_1 &\lesssim \||\xi| e^{-\frac{|\xi|^2}{2}t} \widehat{u_0}\|_{L^1(D_1)} \lesssim t^{-\frac{1}{2}} \|e^{-\frac{|\xi|^2}{4}t} \widehat{u_0}\|_{L^1(D_1)} \\
&\lesssim t^{-\frac{1}{2}} \|e^{-\frac{|\xi|^2}{8}t}\|_{L^2(D_1)} \|e^{-\frac{|\xi|^2}{8}t} \widehat{u_0}\|_{L^2(D_1)} \lesssim t^{-\frac{5}{4}} \|e^{\frac{t}{8}\Delta} u_0\|_{L^2} \lesssim t^{-2} \|u_0\|_{L^1},
\end{aligned}
$$

which along with

$$
L_1 \le \|\widehat{u_0}\|_{L^1(D_1)} \lesssim \|\widehat{u_0}\|_{L^2(D_1)} \lesssim \|u_0\|_{L^2}
$$

yields

$$L_1 \lesssim \langle t\rangle^{-2}\|u_0\|_{L^1\cap L^2}.$$

Similarly, we can also get

$$L_2 \lesssim \langle t\rangle^{-2}\|\tau_0\|_{L^1\cap L^2},$$

and

$$\begin{aligned}
L_3 &\lesssim \int_0^t \langle t-s\rangle^{-2}(\|u\cdot\nabla u\|_{L^1}+\|u\cdot\nabla u\|_{L^2})\,ds\\
&\lesssim \int_0^t \langle t-s\rangle^{-2}(\|u\|_{L^2}\|\nabla u\|_{L^2}+\|u\|_{L^2}\|\nabla u\|_{L^\infty})\,ds\\
&\lesssim \int_0^t \langle t-s\rangle^{-2}(\langle s\rangle^{-2}+\langle s\rangle^{-\frac{11}{4}})\,ds\,X(t)Y(t) \lesssim \langle t\rangle^{-2}X(t)Y(t).
\end{aligned}$$

Next, we use the same approach as for dealing with the estimates of K_4 and K_5 to bound L_4 and L_5, respectively, using the same reasoning reason (see Remark 5.1 for the details). Indeed, we have

$$L_4 \le \mathcal{L}_{41}+\mathcal{L}_{42},$$

where

$$\mathcal{L}_{41} = \int_0^t \||\xi|^2\widehat{M_{12}}(t-s)\widehat{u\otimes\mathcal{A}}\|_{L^1(D_1)}\,ds$$

and

$$\mathcal{L}_{42} = \int_0^t \||\xi||\xi_i|\widehat{M_{12}}(t-s)\widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau}\|_{L^1(D_1)}\,ds.$$

For $\mathcal{L}_{41}$, we have

$$\begin{aligned}
\mathcal{L}_{41} &\lesssim \int_0^t \langle t-s\rangle^{-2}(\|u\otimes\mathcal{A}\|_{L^1}+\|u\otimes\mathcal{A}\|_{L^2})\,ds\\
&\lesssim \int_0^t \langle t-s\rangle^{-2}(\|u\|_{L^2}+\|u\|_{L^\infty})\|\mathcal{A}\|_{L^2}\,ds\\
&\lesssim \int_0^t \langle t-s\rangle^{-2}(\langle s\rangle^{-2}+\langle s\rangle^{-\frac{11}{4}})\,ds\,X(t)Y(t) \lesssim \langle t\rangle^{-2}X(t)Y(t).
\end{aligned}$$

For $\mathcal{L}_{42}$, like the previous procedure for $\mathcal{K}_{42}$, we have

$$\mathcal{L}_{42} = L_{41}+L_{42},$$

where

$$\begin{aligned}
L_{41} &= \int_0^{\frac{t}{2}} \||\xi||\xi_i|\widehat{M_{12}}(t-s)\widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau}\|_{L^1(D_1)}\,ds,\\
L_{42} &= \int_{\frac{t}{2}}^t \||\xi||\xi_i|\widehat{M_{12}}(t-s)\widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau}\|_{L^1(D_1)}\,ds.
\end{aligned}$$

Using the same method as for the estimate of K_{41}, we infer

$$\begin{aligned}
L_{41} &\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-2}(\|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{\dot{B}^0_{1,\infty}} + \|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2})\,ds \\
&\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-2}(\|\nabla u\|_{L^2\cap L^\infty}\|\tau\|_{L^2\cap L^\infty} + \|u\|_{L^2}\|\mathcal{A}\|_{L^2})\,ds \\
&\lesssim \langle t\rangle^{-2}\int_0^{\frac{t}{2}} \langle s\rangle^{-\frac{5}{4}}\,ds Y(t)(\|\tau\|_{L^\infty_t H^2} + X(t)) \\
&\lesssim \langle t\rangle^{-2} Y(t)(\|\tau\|_{L^\infty_t H^2} + X(t)).
\end{aligned}$$

Using the same method as for the estimate of K_{42}, we infer

$$\begin{aligned}
L_{42} &\lesssim \int_{\frac{t}{2}}^t \langle t-s\rangle^{-\frac{5}{4}}\|[\mathbb{P}\operatorname{div}, u_i]\tau\|_{L^2}\,ds \lesssim \int_{\frac{t}{2}}^t \langle t-s\rangle^{-\frac{5}{4}}\||\xi|\widehat{u}\|_{L^1}\|\tau\|_{L^2}\,ds \\
&\lesssim \int_{\frac{t}{2}}^t \langle t-s\rangle^{-\frac{5}{4}}\langle s\rangle^{-2}\,ds Y(t)\|\tau\|_{L^\infty_t L^2} \\
&\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2}\int_{\frac{t}{2}}^t \langle t-s\rangle^{-\frac{5}{4}}\,ds \\
&\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2}.
\end{aligned}$$

Following the process in the estimate of K_5, we have

$$L_5 = L_{51} + L_{52},$$

where

$$L_{51} = \int_0^{\frac{t}{2}} \||\xi|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\,ds$$

and

$$L_{52} = \int_{\frac{t}{2}}^t \||\xi|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_1)}\,ds.$$

Following the estimate of K_{51} and K_{52}, we can get

$$\begin{aligned}
L_{51} &\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-2}(\|Q(\tau,\nabla u)\|_{L^1} + \|Q(\tau,\nabla u)\|_{L^2})\,ds \\
&\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-2}\|\tau\|_{L^2}(\|\nabla u\|_{L^2} + \|\nabla u\|_{L^\infty})\,ds \\
&\lesssim \int_0^{\frac{t}{2}} \langle t-s\rangle^{-2}(\langle s\rangle^{-\frac{5}{4}} + \langle s\rangle^{-2})\,ds Y(t)\|\tau\|_{L^\infty_t L^2} \\
&\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L^\infty_t L^2},
\end{aligned}$$

and

$$\begin{aligned}
L_{52} &\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{5}{4}} \|Q(\tau, \nabla u)\|_{L^2}\, ds \\
&\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{5}{4}} \|\tau\|_{L^2} \|\nabla u\|_{L^\infty}\, ds \\
&\lesssim \int_{\frac{t}{2}}^{t} \langle t-s\rangle^{-\frac{5}{4}} \langle s\rangle^{-2}\, ds Y(t) \|\tau\|_{L_t^\infty L^2} \\
&\lesssim \langle t\rangle^{-2} Y(t) \|\tau\|_{L_t^\infty L^2}.
\end{aligned}$$

Thus we deduce

$$L_5 \lesssim \langle t\rangle^{-2} Y(t) \|\tau\|_{L_t^\infty L^2}.$$

Combining the estimates of L_i $(i = 1, 2, 3, 4, 5)$, we have

$$\||\xi|\hat{u}(t)\|_{L^1(D_1)} \lesssim \langle t\rangle^{-2} \|(u_0, \tau_0)\|_{L^1 \cap L^2} + \langle t\rangle^{-2} Y(t)(\|\tau\|_{L_t^\infty H^2} + X(t)). \tag{5-8}$$

Repeating the above procedure, we can also obtain

$$\begin{aligned}
\sum_{i=2,3} &\||\xi|\hat{u}(t)\|_{L^1(D_i)} \\
&\lesssim \langle t\rangle^{-2} \|(u_0, \tau_0)\|_{L^1 \cap L^2} + \langle t\rangle^{-2} Y(t)(\|\tau\|_{L_t^\infty H^2} + X(t)).
\end{aligned} \tag{5-9}$$

5D. ***The estimate of*** $\boldsymbol{\||\xi|\hat{u}(t)\|_{L^1(D_4)}}$**.** Using (3-3), we have

$$\begin{aligned}
&\||\xi|\hat{u}(t)\|_{L^1(D_4)} \\
&\quad \le \||\xi|\widehat{M_{11}}(t)\widehat{u_0}\|_{L^1(D_4)} + \||\xi|\widehat{M_{12}}(t)\widehat{\mathcal{A}_0}\|_{L^1(D_4)} \\
&\qquad + \int_0^t \||\xi|\widehat{M_{11}}(t-s)\widehat{G}\|_{L^1(D_4)}\, ds + \int_0^t \||\xi|\widehat{M_{12}}(t-s)\widehat{F}\|_{L^1(D_4)}\, ds \\
&\qquad + \int_0^t \||\xi|\widehat{M_{12}}(t-s)\widehat{H}\|_{L^1(D_4)}\, ds \\
&\quad = L_1' + L_2' + \cdots + L_5'.
\end{aligned}$$

Applying (3-5), we have

$$L_1' \lesssim e^{-\frac{t}{2}} \||\xi|\widehat{u_0}\|_{L^1(D_4)} \lesssim e^{-\frac{t}{2}} \|u_0\|_{H^3}$$

and

$$L_2' \lesssim e^{-\frac{t}{2}} \||\xi|\widehat{\tau_0}\|_{L^1(D_4)} \lesssim e^{-\frac{t}{2}} \|\tau_0\|_{H^3}.$$

For L'_3, by (3-5) and Young's inequality, we have

$$\begin{aligned}
L'_3 &\lesssim \int_0^t e^{-\frac{t-s}{2}} \| |\xi| \widehat{u \cdot \nabla u} \|_{L^1}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} (\| |\xi| \hat{u} \|_{L^1}^2 + \|\hat{u}\|_{L^1} \| |\xi|^2 \hat{u} \|_{L^1})\, ds \\
&\lesssim L'_{31} + L'_{32}.
\end{aligned}$$

For L'_{31}, we get

$$L'_{31} \lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s \rangle^{-4}\, ds Y(t)^2 \lesssim \langle t \rangle^{-2} Y(t)^2.$$

We cannot directly bound the estimate L'_{32} as we can for L'_{31}. Let us first consider the estimate of $\| |\xi|^2 \hat{u} \|_{L^1}$. We have

$$\begin{aligned}
\| |\xi|^2 \hat{u} \|_{L^1} &\le \| |\xi|^2 \hat{u} \|_{L^1(|\xi| \le \langle s \rangle^{\frac{3}{2}})} + \| |\xi|^2 \hat{u} \|_{L^1(|\xi| > \langle s \rangle^{\frac{3}{2}})} \\
&\le \langle s \rangle^{\frac{3}{2}} \| |\xi| \hat{u} \|_{L^1} + \langle s \rangle^{-\frac{1}{2}} \| |\xi|^{\frac{7}{3}} \hat{u} \|_{L^1} \\
&\lesssim \langle s \rangle^{\frac{3}{2}} \| |\xi| \hat{u} \|_{L^1} + \langle s \rangle^{-\frac{1}{2}} \|u\|_{H^4}.
\end{aligned}$$

So we have

$$L'_{32} \lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s \rangle^{-2}\, ds\, X(t)(Y(t) + \|u\|_{L^\infty_t H^4}) \lesssim \langle t \rangle^{-2} X(t)(Y(t) + \|u\|_{L^\infty_t H^4}).$$

Thus we have

$$L'_3 \lesssim \langle t \rangle^{-2} (X(t) + Y(t))(Y(t) + \|u\|_{L^\infty_t H^4}).$$

For L'_4, we have

$$\begin{aligned}
L'_4 &\lesssim \int_0^t e^{-\frac{t-s}{2}} (\| \widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau} \|_{L^1} + \| \widehat{u \otimes \mathcal{A}} \|_{L^1})\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} (\| |\xi| \hat{u} \|_{L^1} \|\hat{\tau}\|_{L^1} + \|\hat{u}\|_{L^1} \|\widehat{\mathcal{A}}\|_{L^1})\, ds \\
&\lesssim L'_{41} + L'_{42},
\end{aligned}$$

where we have used

$$\| \widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau} \|_{L^1} \lesssim \| |\xi| \hat{u} \|_{L^1} \|\hat{\tau}\|_{L^1},$$

which can be proved by a similar procedure as that which yielded the estimate of (5-3). For L'_{41}, we have

$$L'_{41} \lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s \rangle^{-2}\, ds\, Y(t) \|\tau\|_{L^\infty_t H^2} \lesssim \langle t \rangle^{-2} Y(t) \|\tau\|_{L^\infty_t H^2}.$$

For L'_{42}, we can get

$$\begin{aligned}
L'_{42} &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{u}\|_{L^1} \|\widehat{\mathcal{A}}\|_{L^1}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{u}\|_{L^1} \big(\|\widehat{\mathcal{A}}\|_{L^1(|\xi|\le\langle s\rangle^{\frac14})} + \|\widehat{\mathcal{A}}\|_{L^1(|\xi|>\langle s\rangle^{\frac14})}\big)\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{\frac34} \|\hat{u}\|_{L^1} \|\widehat{\mathcal{A}}\|_{L^2(|\xi|\le\langle s\rangle^{\frac14})}\, ds \\
&\qquad\qquad + \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-\frac12} \|\hat{u}\|_{L^1} \||\xi|^2\widehat{\mathcal{A}}\|_{L^1(|\xi|>\langle s\rangle^{\frac14})}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\, ds X(t)Y(t) \\
&\qquad\qquad + \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\, ds X(t)\|\tau\|_{L_t^\infty H^5} \\
&\lesssim \langle t\rangle^{-2} X(t)(Y(t) + \|\tau\|_{L_t^\infty H^5}).
\end{aligned}$$

Thus we can obtain

$$L'_4 \lesssim \langle t\rangle^{-2}(X(t)+Y(t))(Y(t)+\|\tau\|_{L_t^\infty H^5}).$$

For L'_5, we have

$$\begin{aligned}
L'_5 &\lesssim \int_0^t e^{-\frac{t-s}{2}} \|\hat{\tau}\|_{L^1} \||\xi|\hat{u}\|_{L^1}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \langle s\rangle^{-2}\, ds Y(t)\|\tau\|_{L_t^\infty H^2} \\
&\lesssim \langle t\rangle^{-2} Y(t)\|\tau\|_{L_t^\infty H^2}.
\end{aligned}$$

Collecting the estimates of L'_i $(i = 1, 2, 3, 4, 5)$, we get

$$\begin{aligned}
&\||\xi|\hat{u}(t)\|_{L^1(D_4)} \\
&\qquad \lesssim e^{-\frac{t}{2}}\|(u_0,\tau_0)\|_{H^3} + \langle t\rangle^{-2}(X(t)+Y(t))(Y(t)+\|(u,\tau)\|_{L_t^\infty H^5}).
\end{aligned} \tag{5-10}$$

It follows from combining (5-8), (5-9) and (5-10) that

$$\begin{aligned}
&\||\xi|\hat{u}\|_{L^1} \\
&\qquad \lesssim \langle t\rangle^{-2}\|(u_0,\tau_0)\|_{L^1\cap H^3} + \langle t\rangle^{-2}(X(t)+Y(t))(Y(t)+\|(u,\tau)\|_{L_t^\infty H^5}).
\end{aligned} \tag{5-11}$$

The estimate of $\|\mathcal{A}(t)\|_{L^2}$. As for the previous procedure, we have

$$\|\mathcal{A}(t)\|_{L^2} = \|\widehat{\mathcal{A}}(t)\|_{L^2} \le \sum_{i=1}^{4} \|\widehat{\mathcal{A}}(t)\|_{L^2(D_i)}.$$

5E. ***The estimate of*** $\|\widehat{\mathcal{A}}(t)\|_{L^2(\mathbb{R}^2\setminus D_4)}$. Using (3-3), we have

$$\begin{aligned}\|\widehat{\mathcal{A}}(t)\|_{L^2(D_1)} &\le \|\widehat{M_{21}}(t)\widehat{u_0}\|_{L^2(D_1)} + \|\widehat{M_{22}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_1)} \\ &\quad + \int_0^t \|\widehat{M_{21}}(t-s)\widehat{G}\|_{L^2(D_1)}\,ds + \int_0^t \|\widehat{M_{22}}(t-s)\widehat{F}\|_{L^2(D_1)}\,ds \\ &\quad + \int_0^t \|\widehat{M_{22}}(t-s)\widehat{H}\|_{L^2(D_1)}\,ds \\ &= N_1 + N_2 + \cdots + N_5.\end{aligned}$$

For N_1; using (3-4) and (4-1), we have

$$N_1 \lesssim t^{-\frac{1}{2}}\|e^{\frac{t}{4}\Delta}u_0\|_{L^2} \lesssim t^{-\frac{5}{4}}\|u_0\|_{L^1},$$

which, together with

$$N_1 \lesssim \||\xi|e^{-\frac{|\xi|^2}{2}t}\widehat{u_0}\|_{L^2(D_1)} \lesssim \|u_0\|_{L^2},$$

yields

$$N_1 \lesssim \langle t\rangle^{-\frac{5}{4}}\|u_0\|_{L^1\cap L^2}.$$

Similarly, we also have

$$N_2 \lesssim \langle t\rangle^{-\frac{5}{4}}\|\tau_0\|_{L^1\cap L^2}$$

and

$$\begin{aligned}N_3 &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}(\|u\cdot\nabla u\|_{L^1} + \|u\cdot\nabla u\|_{L^2})\,ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}\|u\|_{L^2}(\|\nabla u\|_{L^2} + \|\nabla u\|_{L^\infty})\,ds \\ &\lesssim \int_0^t \langle t-s\rangle^{-\frac{5}{4}}(\langle s\rangle^{-2} + \langle s\rangle^{-\frac{11}{4}})\,ds X(t)Y(t) \\ &\lesssim \langle t\rangle^{-\frac{5}{4}}X(t)Y(t).\end{aligned}$$

For N_4, since

$$N_4 \lesssim \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}|\xi_i|\widehat{[\mathbb{P}\,\mathrm{div}, u_i]\tau}\|_{L^2}\,ds + \int_0^t \|e^{-\frac{|\xi|^2}{2}(t-s)}|\xi|\widehat{u\otimes\tau}\|_{L^2}\,ds,$$

we can get the estimate by repeating the estimate of J_4; in fact, we can infer that

$$N_4 \lesssim \langle t\rangle^{-\frac{5}{4}}Y(t)\big(X(t) + Y(t) + \|\tau\|_{L^\infty_t(L^2\cap L^\infty)}\big).$$

Like the estimate of J_5, we have

$$N_5 \lesssim \langle t\rangle^{-\frac{5}{4}}Y(t)\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}.$$

Thus we have

$$\|\widehat{\mathcal{A}}\|_{L^2(D_1)} \lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{5}{4}}Y(t)\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}. \tag{5-12}$$

Similarly, we can also get

$$\sum_{i=2,3}\|\widehat{\mathcal{A}}\|_{L^2(D_i)} \lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap L^2} + \langle t\rangle^{-\frac{5}{4}}Y(t)\|\tau\|_{L^\infty_t(L^2\cap L^\infty)}. \tag{5-13}$$

5F. ***The estimate of*** $\|\widehat{\mathcal{A}}(t)\|_{L^2(D_4)}$**.** We have

$$\begin{aligned}
\|\widehat{\mathcal{A}}(t)\|_{L^2(D_4)} &\le \|\widehat{M_{21}}(t)\widehat{u_0}\|_{L^2(D_4)} + \|\widehat{M_{22}}(t)\widehat{\mathcal{A}_0}\|_{L^2(D_4)} \\
&\qquad + \int_0^t \|\widehat{M_{21}}(t-s)\widehat{G}\|_{L^2(D_4)}\,ds + \int_0^t \|\widehat{M_{22}}(t-s)\widehat{F}\|_{L^2(D_4)}\,ds \\
&\qquad + \int_0^t \|\widehat{M_{22}}(t-s)\widehat{H}\|_{L^2(D_4)}\,ds \\
&= N_1' + N_2' + \cdots + N_5'.
\end{aligned}$$

Using (3-5), we have

$$\begin{aligned}
N_1' + N_2' &\lesssim e^{-\frac{t}{2}}(\||\xi|\widehat{u_0}\|_{L^2(D_4)} + \||\xi|\widehat{\tau_0}\|_{L^2(D_4)}) \\
&\lesssim e^{-\frac{t}{2}}\|(u_0,\tau_0)\|_{H^1}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
N_3' &\lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\cdot\nabla u\|_{L^2}\,ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}\|u\|_{L^2}\|\nabla u\|_{L^\infty}\,ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}\langle s\rangle^{-\frac{11}{4}}\,ds X(t)Y(t) \\
&\lesssim \langle t\rangle^{-\frac{11}{4}}X(t)Y(t),
\end{aligned}$$

and

$$\begin{aligned}
N_4' &\lesssim \int_0^t e^{-\frac{t-s}{2}}\|\mathbb{P}\,\mathrm{div}(u\cdot\nabla\tau)\|_{L^2}\,ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}(\|u\|_{L^\infty}\|\nabla\tau\|_{L^2} + \|\nabla u\|_{L^\infty}\|\tau\|_{L^2})\,ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}}(\langle s\rangle^{-\frac{3}{2}} + \langle s\rangle^{-2})\,ds(X(t)+Y(t))\|\tau\|_{L^\infty_t H^1} \\
&\lesssim \langle t\rangle^{-\frac{3}{2}}(X(t)+Y(t))\|\tau\|_{L^\infty_t H^1}.
\end{aligned}$$

For the last term N_5', we have

$$
\begin{aligned}
N_5' &\lesssim \int_0^t e^{-\frac{t-s}{2}} (\|\nabla\tau\|_{L^2}\|\nabla u\|_{L^\infty} + \|\tau\|_{L^2}\|\Delta u\|_{L^2})\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} (\|\nabla\tau\|_{L^2}\|\nabla u\|_{L^\infty} + \|\tau\|_{L^2}\|\Delta u\|_{L^\infty})\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} (\|\nabla\tau\|_{L^2}\|\nabla u\|_{L^\infty} + \|\tau\|_{L^2}\||\xi|^2\hat{u}\|_{L^1})\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \Big\{\|\nabla\tau\|_{L^2}\|\nabla u\|_{L^\infty} + \|\tau\|_{L^2}(\||\xi|^2\hat{u}\|_{L^1(|\xi|<\langle s\rangle^{\frac{3}{4}})} \\
&\qquad\qquad + \||\xi|^2\hat{u}\|_{L^1(|\xi|\geq\langle s\rangle^{\frac{3}{4}})})\Big\}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} \Big\{\|\nabla\tau\|_{L^2}\|\nabla u\|_{L^\infty} + \|\tau\|_{L^2}(\langle s\rangle^{\frac{3}{4}}\||\xi|\hat{u}\|_{L^1} \\
&\qquad\qquad + \langle s\rangle^{-\frac{5}{4}}\||\xi|^{\frac{11}{3}}\hat{u}\|_{L^1})\Big\}\, ds \\
&\lesssim \int_0^t e^{-\frac{t-s}{2}} (\langle s\rangle^{-2} + \langle s\rangle^{-\frac{5}{4}})\, ds (Y(t) + \|u\|_{L_t^\infty H^6})\|\tau\|_{L_t^\infty H^1} \\
&\lesssim \langle t\rangle^{-\frac{5}{4}}(Y(t) + \|u\|_{L_t^\infty H^6})\|\tau\|_{L_t^\infty H^1}.
\end{aligned}
$$

Thus we have

$$
\|\widehat{\mathcal{A}}(t)\|_{L^2(D_4)} \lesssim e^{-\frac{t}{2}}\|(u_0,\tau_0)\|_{H^1} + \langle t\rangle^{-\frac{5}{4}}(X(t)+Y(t) + \|u\|_{L_t^\infty H^6})\|\tau\|_{L_t^\infty H^1}. \tag{5-14}
$$

Combining (5-12), (5-13) and (5-14) leads to

$$
\|\widehat{\mathcal{A}}(t)\|_{L^2} \lesssim \langle t\rangle^{-\frac{5}{4}}\|(u_0,\tau_0)\|_{L^1\cap H^1} + \langle t\rangle^{-\frac{5}{4}}(X(t)+Y(t) + \|u\|_{L_t^\infty H^6})\|\tau\|_{L_t^\infty H^1}. \tag{5-15}
$$

6. Proof of Theorem 1.3

In this section, we give the proof of our main result.

Proof of Theorem 1.3. Thanks to Sections 4 and 5, we can get

$$
X(t) \lesssim \|(u_0,\tau_0)\|_{L^1\cap H^2} + (X(t)+Y(t))(\|\tau\|_{L_t^\infty H^2} + X(t)),
$$

which can be obtained by combining (4-5), (5-7), and

$$
Y(t) \lesssim \|(u_0,\tau_0)\|_{L^1\cap H^3} + \|u\|_{L_t^\infty H^6}\|\tau\|_{L_t^\infty H^1} + (X(t)+Y(t))(Y(t) + \|(u,\tau)\|_{L_t^\infty H^6}),
$$

which can be deduced by using (4-10), (5-11) and (5-15). Since (u,τ) is the solution obtained by Theorem 1.1, we have

$$\|(u,\tau)\|_{L_t^\infty H^6} \lesssim \epsilon.$$

Using the small conditions of the initial data, we can get that there exists a positive constant C_1 such that

$$X(t)+Y(t) \le \ C_1\{\epsilon+\epsilon(X(t)+Y(t))+(X(t)+Y(t))^2\}.$$

Then one can get the desired result by using continuous arguments. This completes the proof of Theorem 1.3. $\square$

Appendix

In this section, we show the proofs of Lemma 4.1 and (5-3) in order by using the Littlewood–Paley theory and Fourier analysis technique.

Proof of Lemma 4.1. Let $f=u_i$. Using

$$\Delta_j([\mathbb{P}\,\mathrm{div}, f]\tau)=[\Delta_j\mathbb{P}\,\mathrm{div}, f]\tau+f\Delta_j\mathcal{A}-\Delta_j(f\mathcal{A}),$$

we have

$$\|\Delta_j([\mathbb{P}\,\mathrm{div}, f]\tau)\|_{L^1}\le\|[\Delta_j\mathbb{P}\,\mathrm{div}, f]\tau\|_{L^1}+\|f\Delta_j\mathcal{A}-\Delta_j(f\mathcal{A})\|_{L^1}=F_1+F_2.$$

F_2 can be easily bounded as follows:

$$F_2\lesssim\|f\|_{L^2}\|\mathcal{A}\|_{L^2}.$$

Then it remains to bound F_1. Using Bony's decomposition, we get

$$\begin{aligned}
&\|[\Delta_j\mathbb{P}\,\mathrm{div}, f]\tau\|_{L^1}\\
&\quad\le\sum_{|k-j|\le4}\|[\Delta_j\mathbb{P}\,\mathrm{div}, S_{k-1}f]\Delta_k\tau\|_{L^1}+\sum_{|k-j|\le4}\|\Delta_j\mathbb{P}\,\mathrm{div}(\Delta_k f S_{k-1}\tau)\|_{L^1}\\
&\qquad+\sum_{k\ge j-3}\|\Delta_k f S_{k+2}\Delta_j\mathcal{A}\|_{L^1}+\sum_{k\ge j-3}\|\Delta_j\mathbb{P}\,\mathrm{div}(\Delta_k f\tilde{\Delta}_k\tau)\|_{L^1}\\
&\quad=\gimel_1+\gimel_2+\gimel_3+\gimel_4.
\end{aligned}$$

Notice that

$$\sum_{|k-j|\le4}\|[\Delta_j\mathbb{P}\,\mathrm{div}, S_{k-1}f]\Delta_k\tau\|_{L^1}=\sum_{|k-j|\le4}\sum_{i=1,2,3}\|[\Delta_j\mathbb{P}\,\partial_{\mathrm{i}}, S_{k-1}f]\Delta_k\tau_{i\cdot}\|_{L^1}$$

and let $g=\tau_{i\cdot}$, and return to considering the estimate of $\|[\Delta_j\mathbb{P}\,\partial_{\mathrm{i}}, S_{k-1}f]\Delta_k g\|_{L^1}$. We can see there exists a vector function $h_i(x)$, the components of which are

Schwartz functions, such that

$$\Delta_j \mathbb{P}\,\partial_i g = 2^{4j} \int h_i(2^j(x-y))g(y)\,dy.$$

Thus, letting $Q_i(x) = |x|h_i(x)$, we have

$$\begin{aligned}
&|[\Delta_j \mathbb{P}\,\partial_{\mathrm{i}}, S_{k-1}f]\Delta_k g| \\
&\quad = |2^{4j} \int h_i(2^j(x-y))\{(S_{k-1}f)(y) - (S_{k-1}f)(x)\}\Delta_k g(y)\,dy| \\
&\quad \le 2^{4j} \int_0^1 \int |h_i(2^j(x-y))||x-y||\nabla S_{k-1}f(sx+(1-s)y)||\Delta_k g(y)|\,dy\,ds \\
&\quad \le 2^{3j} \int_0^1 \int |Q_i(2^j(x-y))||\nabla S_{k-1}f(sx+(1-s)y)||\Delta_k g(y)|\,dy\,ds \\
&\quad = 2^{3j} \int_0^1 \int |Q_i(2^j z)||\nabla S_{k-1}f(x-z+sz)||\Delta_k g(x-z)|\,dz\,ds.
\end{aligned}$$

By Hölder's inequality, we have

$$\begin{aligned}
&\|[\Delta_j \mathbb{P}\,\partial_{\mathrm{i}}, S_{k-1}f]\Delta_k g\|_{L^1} \\
&\qquad \le 2^{3j} \int_0^1 \int |Q_i(2^j z)|\|\nabla S_{k-1}f(\cdot - z + sz)\|_{L^2(\cdot)}\|\Delta_k \tau(\cdot - z)\|_{L^2(\cdot)}\,dz\,ds \\
&\qquad \le 2^{3j} \int |Q_i(2^j z)|\,dz \|\nabla f\|_{L^2}\|\Delta_k g\|_{L^2} \lesssim \|\nabla f\|_{L^2}\|g\|_{L^2}.
\end{aligned}$$

Thus we can get

$$\gimel_1 \lesssim \|\nabla f\|_{L^2}\|\tau\|_{L^2}.$$

$\gimel_2$ can be easily bounded. Next, we bound $\gimel_3$ and $\gimel_4$. For $\gimel_3$, by Bernstein's inequality, we have

$$\begin{aligned}
\gimel_3 &\lesssim \sum_{k \ge j-3} \|\Delta_k f\|_{L^2}\|\Delta_j \mathcal{A}\|_{L^2} \lesssim 2^j \|\Delta_j \tau\|_{L^2} \sum_{k \ge j-3} 2^{-k}\|\nabla \Delta_k f\|_{L^2} \\
&\lesssim \|\nabla f\|_{L^2}\|\tau\|_{L^2}.
\end{aligned}$$

For $\gimel_4$, using Bernstein's inequality again, we have

$$\begin{aligned}
\gimel_4 &\lesssim 2^j \sum_{k \ge j-3} \|\Delta_k f\|_{L^2}\|\tilde{\Delta}_k \tau\|_{L^2} \lesssim 2^j \|\tau\|_{L^2} \sum_{k \ge j-3} 2^{-k}\|\Delta_k \nabla f\|_{L^2} \\
&\lesssim 2^j \|\nabla f\|_{L^2}\|\tau\|_{L^2} \sum_{k \ge j-3} 2^{-k} \lesssim \|\nabla f\|_{L^2}\|\tau\|_{L^2}.
\end{aligned}$$

Collecting the above four estimates leads to the desired results. □

Proof of (5-3). Let $f = u_i$, and using

$$[\mathbb{P}\,\mathrm{div}, f]\tau = \sum_{i=1,2,3} [\mathbb{P}\,\partial_{\mathrm{i}}, f]\tau_{i\cdot},$$

we return to the bound $\|[\mathbb{P}\,\partial_{\mathrm{i}}, f]g_i\|_{L^2}$, where $g_i = \tau_{i\cdot}$. Notice that if $\widehat{\mathbb{P}\,\partial_i} = \Phi(\xi, \xi_i)$, satisfying $|\nabla_\xi \Phi(\xi, \xi_i)| \lesssim 1$, we have

$$|\Phi(\xi, \xi_i) - \Phi(\xi - \eta, \xi_i - \eta_i)| \lesssim |\eta|.$$

Using the Fourier transform, we have

$$\begin{aligned}
&|\widehat{[\mathbb{P}\,\partial_{\mathrm{i}}, f]g_i}| \\
&\quad = \left|\Phi(\xi, \xi_i) \int \hat{f}(\eta)\widehat{g_i}(\xi - \eta)\,d\eta - \int \hat{f}(\eta)\Phi(\xi - \eta, \xi_i - \eta_i)\widehat{g_i}(\xi - \eta)\,d\eta\right| \\
&\quad \le \left|\int \hat{f}(\eta)|\Phi(\xi, \xi_i) - \Phi(\xi - \eta, \xi_i - \eta_i)|\widehat{g_i}(\xi - \eta)\,d\eta\right| \\
&\quad \lesssim \left|\int |\eta|\hat{f}(\eta)\widehat{g_i}(\xi - \eta)\,d\eta\right|,
\end{aligned}$$

which yields the desired result by using Young's inequality. □

Acknowledgements

We are grateful to the referees for the careful reading and helpful comments. This work was supported by the NSF of China (11901301), the NSF of the Jiangsu Higher Education Institutions of China (18KJB110018) and the NSF of Jiangsu Province BK20180721.

References

[Bahouri et al. 2011] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Math. Wissenschaften **343**, Springer, 2011. MR Zbl

[Bird et al. 1977] R. B. Bird, R. C. Armstrong, and O. Hassager, *Dynamics of polymeric liquids, I: Fluid mechanics*, Wiley, New York, 1977.

[Chemin and Masmoudi 2001] J.-Y. Chemin and N. Masmoudi, "About lifespan of regular solutions of equations related to viscoelastic fluids", *SIAM J. Math. Anal.* **33**:1 (2001), 84–112. MR Zbl

[Chen and Miao 2008] Q. Chen and C. Miao, "Global well-posedness of viscoelastic fluids of Oldroyd type in Besov spaces", *Nonlinear Anal.* **68**:7 (2008), 1928–1939. MR Zbl

[Fang and Zi 2016] D. Fang and R. Zi, "Global solutions to the Oldroyd-B model with a class of large initial data", *SIAM J. Math. Anal.* **48**:2 (2016), 1054–1084. MR Zbl

[Fernández Cara et al. 1994] E. Fernández Cara, F. Guillén, and R. R. Ortega, "Existence et unicité de solution forte locale en temps pour des fluides non newtoniens de type Oldroyd (version L^s–L^r)", *C. R. Acad. Sci. Paris Sér. I Math.* **319**:4 (1994), 411–416. MR Zbl

[Guillopé and Saut 1990a] C. Guillopé and J.-C. Saut, "Existence results for the flow of viscoelastic fluids with a differential constitutive law", *Nonlinear Anal.* **15**:9 (1990), 849–869. MR Zbl

[Guillopé and Saut 1990b] C. Guillopé and J.-C. Saut, "Global existence and one-dimensional nonlinear stability of shearing motions of viscoelastic fluids of Oldroyd type", *RAIRO Modél. Math. Anal. Numér.* **24**:3 (1990), 369–401. MR Zbl

[Kato and Ponce 1988] T. Kato and G. Ponce, "Commutator estimates and the Euler and Navier–Stokes equations", *Comm. Pure Appl. Math.* **41**:7 (1988), 891–907. MR Zbl

[Kenig et al. 1991] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness of the initial value problem for the Korteweg–de Vries equation", *J. Amer. Math. Soc.* **4**:2 (1991), 323–347. MR Zbl

[Lions and Masmoudi 2000] P. L. Lions and N. Masmoudi, "Global solutions for some Oldroyd models of non-Newtonian flows", *Chinese Ann. Math. Ser. B* **21**:2 (2000), 131–146. MR Zbl

[Oldroyd 1958] J. G. Oldroyd, "Non-Newtonian effects in steady motion of some idealized elastico-viscous liquids", *Proc. Roy. Soc. London. Ser. A* **245** (1958), 278–297. MR Zbl

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Math. Ser. **30**, Princeton Univ. Press, 1970. MR Zbl

[Wan 2017a] R. Wan, "Global existence and long time behavior of smooth solution for the 2D Boussinesq–Navier–Stokes equations", preprint, 2017, Available at https://tinyurl.com/renhuiwan2.

[Wan 2017b] R. Wan, "Global small solutions to the 2D incompressible Oldroyd-B model with only dissipation", preprint, 2017, Available at https://tinyurl.com/renhuiwan.

[Wan 2019] R. Wan, "Some new global results to the incompressible Oldroyd-B model", *Z. Angew. Math. Phys.* **70**:1 (2019), art. id. 28. MR

[Zhu 2018] Y. Zhu, "Global small solutions of 3D incompressible Oldroyd-B model without damping mechanism", *J. Funct. Anal.* **274**:7 (2018), 2039–2060. MR Zbl

[Zi et al. 2014] R. Zi, D. Fang, and T. Zhang, "Global solution to the incompressible Oldroyd-B model in the critical L^p framework: the case of the non-small coupling parameter", *Arch. Ration. Mech. Anal.* **213**:2 (2014), 651–687. MR Zbl

Received February 1, 2018.

RENHUI WAN
INSTITUTE OF MATHEMATICS, SCHOOL OF MATHEMATICAL SCIENCES
NANJING NORMAL UNIVERSITY
NANJING
CHINA

rhwanmath@163.com, wrh@njnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.703

TRIANGULATED CATEGORIES WITH CLUSTER TILTING SUBCATEGORIES

WUZHONG YANG, PANYUE ZHOU AND BIN ZHU

Dedicated to Professor Idun Reiten on the occasion of her 76th birthday

For a triangulated category $\mathscr{C}$ with a cluster tilting subcategory $\mathcal{T}$ which contains infinitely many indecomposable objects, the notion of weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ is introduced. We use them to study the τ-tilting theory in the module category over $\mathcal{T}$. Inspired by the work of Iyama, Jørgensen and Yang (2014), we introduce the notion of τ-tilting subcategories of mod $\mathcal{T}$, and show that there exists a bijection between weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ and support τ-tilting subcategories of mod $\mathcal{T}$. Moreover, we describe the subcategories of mod $\mathcal{T}$ which correspond to cluster tilting subcategories of $\mathscr{C}$. This generalizes and improves results by Adachi, Iyama and Reiten (2014), Beligiannis (2013), and Yang and Zhu (2019).

1. Introduction

The links between cluster tilting objects in a (2-Calabi–Yau) triangulated category and tilting modules over the cluster-tilted algebras have been studied for a relatively long time. They stemmed from the categorification of cluster algebras, see for

Wuzhong Yang was supported by the National Natural Science Foundation of China (Grant No. 11801442) and by Scientific Research Program Funded by Shaanxi Provincial Education Department (Program No. 17JK0794), Panyue Zhou was supported by the National Natural Science Foundation of China (Grant No. 11901190) and by the Hunan Provincial Natural Science Foundation of China (Grant No. 2018JJ3205) and Bin Zhu was supported by the NSF of China (Grant No. 11671221).
MSC2010: primary 16G20, 18E30; secondary 16G99, 16G70.
Keywords: ghost cluster tilting subcategories, (support) τ-tilting subcategories, tilting subcategories, cluster tilting subcategories.

examples: [Smith 2008; Fu and Liu 2009; Beaudet et al. 2014; Lasnier 2011; Beligiannis 2013]. Adachi, Iyama and Reiten [Adachi et al. 2014] established a bijection between cluster tilting objects in a 2-Calabi–Yau triangulated category and support τ-tilting modules over a cluster-tilted algebra (see also [Chang et al. 2015; Yang et al. 2017; Yang and Zhu 2019] for various versions of this bijection). They introduced the τ-tilting theory for finite-dimensional algebras. As a generalization of classical tilting theory, it completes tilting theory from the viewpoint of mutation. Nowadays the relationships between τ-tilting theory and the various aspects of the representation theory of finite-dimensional algebras have been studied.

In order to generalize the bijection in [Adachi et al. 2014] mentioned above to arbitrary triangulated categories with cluster tilting objects, two of us [Yang and Zhu 2019] introduced the notion of relative cluster tilting objects. An object M in a triangulated category $\mathscr{C}$ with a cluster tilting object T is called a $T[1]$-cluster tilting object provided that $|M| = |T|$ and $[T[1]](M, M[1]) = 0$, where $|X|$ denotes the number of the isomorphism classes of indecomposable direct summands of X and $[T[1]](X, X[1]))$ denotes the subgroup of $\mathrm{Hom}_{\mathscr{C}}(X, X[1])$ consisting of morphisms factoring through an object in $\mathsf{add}(T[1])$. It was proved that there is a bijection between the set of basic $T[1]$-cluster tilting objects and the set of basic support τ-tilting modules over the cluster tilted algebra $\mathrm{End}_{\mathscr{C}}(T)^{\mathrm{op}}$, see [Yang and Zhu 2019], which is the bijection in [Adachi et al. 2014] when $\mathscr{C}$ is 2-Calabi–Yau.

Although the (2-Calabi–Yau) triangulated categories with cluster tilting objects are the main sources for categorifying cluster algebras, the more general triangulated categories (not necessarily 2-Calabi–Yau) with cluster tilting subcategories (the number of nonisomorphic indecomposable objects in it is not finite) appear naturally, see for examples, [Jørgensen and Palu 2013; Ng 2010; Igusa and Todorov 2015a; 2015b; Holm and Jørgensen 2012; Liu and Paquette 2017; Chang et al. 2018; Gratz et al. 2019; Stovicek and van Roosmalen 2016; Jørgensen and Yakimov 2017]. It is natural to ask which classes of subcategories of $\mathscr{C}$ correspond bijectively to support τ-tilting subcategories of $\mathsf{mod}\ \mathcal{T}$ for 2-Calabi–Yau triangulated categories, higher Calabi–Yau triangulated categories or arbitrary triangulated categories, where $\mathcal{T}$ is a cluster tilting subcategory of $\mathscr{C}$. Iyama, Jørgensen and Yang [Iyama et al. 2014] gave a functor version of τ-tilting theory. They considered modules over a category and showed that for a triangulated category $\mathscr{C}$ with a silting subcategory $\mathcal{S}$, there exists a bijection between the set of silting subcategories of $\mathscr{C}$ which are in $\mathcal{S} * \mathcal{S}[1]$ and the set of support τ-tilting pairs of $\mathsf{mod}\, \mathcal{S}$.

Motivated by this question and the bijection given by Yang and Zhu [2019], we introduce the notions of $\mathcal{T}[1]$-cluster tilting subcategories (also called ghost cluster tilting subcategories) and weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ (the precise definitions of these subcategories are given in Definition 3.1), which are generalizations of cluster tilting subcategories. When $\mathscr{C}$ has a cluster tilting object T, then weak $\mathsf{add}(T[1])$-cluster tilting subcategories coincide with $\mathsf{add}(T[1])$-cluster tilting subcategories, and are also the same as $T[1]$-cluster tilting objects introduced in [Yang and Zhu 2019].

The first part of our work is to develop a basic theory of ghost cluster tilting subcategories of $\mathscr{C}$. Some intrinsic properties and results on ghost cluster tilting subcategories will be presented. Some of our results can be summarized as follows.

$$\begin{tikzcd}
\boxed{\text{cluster tilting}} \arrow[d, shift left, "\text{Proposition 3.3}"] & \\
\boxed{\text{ghost cluster tilting}} \arrow[u, shift left, "\substack{\text{F-stable}\\ \text{Theorem 3.6}}"] \arrow[r, shift left, "\text{Corollary 3.11}"] \arrow[d, shift left, "\text{Proposition 3.7}"] & \boxed{\text{weak ghost cluster tilting}} \arrow[l, shift left, "\substack{\text{contravariantly finite}\\ \text{Definition 3.1}}"] \\
\boxed{\text{maximal ghost rigid}} \arrow[u, shift left, "\substack{\text{contravariantly finite}\\ \text{Theorem 3.10}}"] &
\end{tikzcd}$$

The second part of our paper is devoted to answering the question above. We have the following main result.

Theorem 1.1 (see Proposition 4.2 and Theorem 4.3). *Let $\mathscr{C}$ be a triangulated category with a cluster-tilting subcategory $\mathcal{T}$. The functor $\mathbb{H}:\mathscr{C}\to \mathsf{Mod}\ \mathcal{T}$ induces a bijection*

$$\Phi:\mathscr{X}\longmapsto(\mathbb{H}(\mathscr{X}),\mathcal{T}\cap\mathscr{X}[-1])$$

from the first of the following sets to the second:

(I) *$\mathcal{T}[1]$-rigid subcategories of $\mathscr{C}$.*

(II) *τ-rigid pairs of* $\mathsf{mod}\ \mathcal{T}$.

It restricts to a bijection from the first to the second of the following sets:

(I) *Weak $\mathcal{T}[1]$-tilting subcategories of $\mathscr{C}$.*

(II) *Support τ-tilting subcategories of* $\mathsf{mod}\ \mathcal{T}$.

Consequently, we also describe the subcategories of $\mathsf{mod}\ \mathcal{T}$ which correspond to cluster tilting subcategories of $\mathscr{C}$ (see Theorem 4.4). This generalizes and improves several results in the literature.

Inspired by Adachi, Iyama and Reiten [Adachi et al. 2014] and by Iyama, Jørgensen and Yang [Iyama et al. 2014], we introduce the notions of τ-tilting subcategories and tilting subcategories of $\mathsf{mod}\,\mathcal{T}$. In the third part of our paper, we give some close relationships between certain ghost cluster tilting subcategories of $\mathscr{C}$ and some important subcategories of $\mathsf{mod}\,\mathcal{T}$ (see Theorems 4.8 and 4.11).

The paper is organized as follows. In Section 2, we recall some elementary definitions and facts about cluster tilting subcategories and support τ-tilting subcategories. In Section 3, we will study the basic properties of ghost cluster tilting subcategories of $\mathscr{C}$. For a triangulated category with cluster tilting object, we will show that the definition of ghost cluster tilting objects in $\mathscr{C}$ is equivalent to the definition of relative cluster tilting objects in [Yang and Zhu 2019]. In Section 4, we explore the connections between ghost cluster tilting theory and τ-tilting theory.

We conclude this section with some conventions.

Throughout this article, k is an algebraically closed field. All modules we consider in this paper are left modules. Let $\mathscr{C}$ be an additive category. When we say that $\mathscr{D}$ is a subcategory of $\mathscr{C}$, we always assume that $\mathscr{D}$ is a full subcategory which is closed under isomorphisms, direct sums and direct summands. We denote by $[\mathscr{D}]$ the ideal of $\mathscr{C}$ consisting of morphisms which factor through objects in $\mathscr{D}$. Thus we get a new category $\mathscr{C}/[\mathscr{D}]$ whose objects are objects of $\mathscr{C}$ and whose morphisms are elements of $\mathscr{C}(X, Y)/[\mathscr{D}](X, Y)$ for $X, Y \in \mathscr{C}/[\mathscr{D}]$. For any object M, we denote by $\mathsf{add}\,M$ the full subcategory of $\mathscr{C}$ consisting of direct summands of direct sum of finitely many copies of M and simply denote $\mathscr{C}/[\mathsf{add}\,M]$ by $\mathscr{C}/[M]$. Let $\mathscr{X}$ and $\mathscr{Y}$ be subcategories of $\mathscr{C}$. We denote by $\mathscr{X} \vee \mathscr{Y}$ the smallest subcategory of $\mathscr{C}$ containing $\mathscr{X}$ and $\mathscr{Y}$. For two morphisms $f : M \to N$ and $g : N \to L$, the composition of f and g is denoted by $gf : M \to L$.

Let X be an object in $\mathscr{C}$. A morphism $f : D_0 \to X$ is called a *right $\mathscr{D}$-approximation* of X if $D_0 \in \mathscr{D}$ and $\mathrm{Hom}_{\mathscr{C}}(-, f)|_{\mathscr{D}}$ is surjective. If any object in $\mathscr{C}$ has a right $\mathscr{D}$-approximation, we call $\mathscr{D}$ *contravariantly finite* in $\mathscr{C}$. Dually, a *left $\mathscr{D}$-approximation* and a *covariantly finite subcategory* are defined. We say that $\mathscr{D}$ is *functorially finite* if it is both covariantly finite and contravariantly finite. For more details, we refer to [Auslander and Reiten 1991].

For any triangulated category $\mathscr{C}$, we assume that it is k-linear, Hom-finite, and satisfies the Krull–Remak–Schmidt property [Happel 1988]. For any object M in $\mathscr{C}$, we can write $M \simeq M_1 \oplus \cdots \oplus M_n$, where the endomorphism ring of M_i is local, for any $i = 1, 2, \dots, n$. Then M is called basic if $M_i \not\simeq M_j$ for all $i \neq j$. In $\mathscr{C}$, we denote the shift functor by [1] and for objects X and Y, define $\mathrm{Ext}^i_{\mathscr{C}}(X, Y) = \mathrm{Hom}_{\mathscr{C}}(X, Y[i])$. For two subcategories $\mathscr{X}, \mathscr{Y}$ of $\mathscr{C}$, we denote by $\mathrm{Ext}^1(\mathscr{X}, \mathscr{Y}) = 0$ when $\mathrm{Ext}^1(X, Y) = 0$ for any $X \in \mathscr{X}$ and $Y \in \mathscr{Y}$. For a subcategory $\mathscr{X}$, we use $|\mathscr{X}|$ to denote the number of nonisomorphic indecomposable objects in $\mathscr{X}$. It is easy to see that $|\mathscr{X}| < \infty$ if and only if $\mathscr{X} = \mathrm{add}\, X$ for an object X. In this case, $|\mathscr{X}|$ is denoted simply by $|X|$.

2. Background and preliminary results

In this section, we give some background material and recall some results that will be used in this paper.

Cluster tilting subcategories and relative cluster tilting objects. Let $\mathscr{C}$ be a triangulated category. An important class of subcategories of $\mathscr{C}$ are the cluster tilting subcategories, which have many nice properties. We recall the definition of cluster tilting subcategories from [Buan et al. 2006; Keller and Reiten 2007; Koenig and Zhu 2008; Iyama and Yoshino 2008].

Definition 2.1. (1) A subcategory $\mathcal{T}$ of $\mathscr{C}$ is called *rigid* if $\mathrm{Hom}_{\mathscr{C}}(\mathcal{T}, \mathcal{T}[1]) = 0$.

(2) A subcategory $\mathcal{T}$ of $\mathscr{C}$ is *maximal rigid* if it is rigid and maximal with respect to this property, that is, $\mathcal{T} = \{M \in \mathscr{C} \mid \mathrm{Hom}_{\mathscr{C}}(\mathcal{T} \vee \mathsf{add}\, M, (\mathcal{T} \vee \mathsf{add}\, M)[1]) = 0\}$.

(3) A functorially finite subcategory $\mathcal{T}$ of $\mathscr{C}$ is called *cluster tilting* if

$$\mathcal{T} = \{M \in \mathscr{C} \mid \mathrm{Hom}_{\mathscr{C}}(\mathcal{T}, M[1]) = 0\} = \{M \in \mathscr{C} \mid \mathrm{Hom}_{\mathscr{C}}(M, \mathcal{T}[1]) = 0\}.$$

(4) An object T in $\mathscr{C}$ is *cluster tilting* if $\mathsf{add}\, T$ is a cluster tilting subcategory of $\mathscr{C}$.

Remark 2.2. It is easy to see that a subcategory $\mathcal{T}$ of $\mathscr{C}$ is cluster tilting if and only if it is contravariantly finite in $\mathscr{C}$ and $\mathcal{T} = \{M \in \mathscr{C} \mid \mathrm{Hom}_{\mathscr{C}}(\mathcal{T}, M[1]) = 0\}$, see for example [Koenig and Zhu 2008].

For two subcategories $\mathscr{X}$ and $\mathscr{Y}$ of $\mathscr{C}$, we denote by $\mathscr{X} * \mathscr{Y}$ the collection of objects in $\mathscr{C}$ consisting of all such $M \in \mathscr{C}$ with triangles

$$X \longrightarrow M \longrightarrow Y \longrightarrow X[1],$$

where $X \in \mathscr{X}$ and $Y \in \mathscr{Y}$.

Recall from [Bondal and Kapranov 1989] that $\mathscr{C}$ has a Serre functor $\mathbb{S}$ provided $\mathbb{S} : \mathscr{C} \to \mathscr{C}$ is an equivalence and there exists a functorial isomorphism

$$\mathrm{Hom}_{\mathscr{C}}(A, B) \simeq D\,\mathrm{Hom}_{\mathscr{C}}(B, \mathbb{S}A)$$

for any $A, B \in \mathscr{C}$, where D is the duality over k. Thus $\mathscr{C}$ has the Auslander–Reiten translation $\tau \simeq \mathbb{S}[-1]$, see [Reiten and Van den Bergh 2002]. Define an equivalence $F = \tau^{-1} \circ [1]$. An object M in $\mathscr{C}$ is called *F-stable* if $F(M) \simeq M$ and a subcategory $\mathcal{M}$ of $\mathscr{C}$ is called *F-stable* if $F(\mathcal{M}) = \mathcal{M}$. We say that $\mathscr{C}$ is *2-Calabi–Yau* if $\mathbb{S} \simeq [2]$. Note that for a 2-Calabi–Yau category $\mathscr{C}$, $F = \mathrm{id}_{\mathscr{C}}$.

We have the following result [Keller and Reiten 2007; Koenig and Zhu 2008; Iyama and Yoshino 2008], which will be used frequently in this paper.

Proposition 2.3. *Let $\mathcal{T}$ be a cluster-tilting subcategory of $\mathscr{C}$ and C be an arbitrary object in $\mathscr{C}$. Then*:

(a) $\mathscr{C} = \mathcal{T} * \mathcal{T}[1]$.

(b) *$F\mathcal{T} = \mathcal{T}$ if $\mathscr{C}$ has a Serre functor.*

(c) *Let $C \to T_0$ be a left $\mathsf{add}\, T$-approximation of C. Let $C \to T_0 \to Y \to C[1]$ be a completed triangle. Then Y is in $\mathsf{add}\, T$.*

When $\mathscr{C}$ is a 2-Calabi–Yau triangulated category, cluster tilting objects have a very important property: if we remove some direct summand T_i from a cluster tilting object $T = T_1 \oplus T_2 \oplus \cdots \oplus T_n$ to get $T/T_i = \oplus_{j \neq i} T_j$ (which is called an almost complete cluster tilting object), then there is exactly one indecomposable object T_i^* such that $T_i^* \ncong T_i$ and $T/T_i \oplus T_i^*$ is a cluster-tilting object, which is called the mutation of T at T_i, see [Buan et al. 2006; Iyama and Yoshino 2008]. But the mutation of cluster tilting objects in triangulated categories which are not 2-Calabi–Yau is not always possible, see for example Section II1 in [Buan et al. 2009]. In order to generalize it in a more general triangulated category, Yang and Zhu [2019] introduced the notion of relative cluster tilting objects in triangulated categories as follows.

Definition 2.4 [Yang and Zhu 2019, Definition 3.1]. Let $\mathscr{C}$ be a triangulated category with a cluster tilting object.

- An object X in $\mathscr{C}$ is called *relative rigid* if there exists a cluster tilting object T such that $[T[1]](X, X[1]) = 0$. In this case, X is also called $T[1]$-*rigid*.
- An object X in $\mathscr{C}$ is called *relative cluster tilting* if there exists a cluster tilting object T such that X is $T[1]$-rigid and $|X| = |T|$. In this case, X is also called $T[1]$-*cluster tilting*.

Throughout this paper, we denote by $T[1]$-rigid $\mathscr{C}$ (respectively, $T[1]$-tilt $\mathscr{C}$) the set of isomorphism classes of basic $T[1]$-rigid (respectively, basic $T[1]$-cluster tilting) objects in $\mathscr{C}$.

Support τ-tilting modules and support τ-tilting subcategories. Let Λ be a finite-dimensional k-algebra and τ the Auslander–Reiten translation. We denote by $\operatorname{proj}\Lambda$ the subcategory of $\operatorname{mod}\Lambda$ consisting of projective Λ-modules. Support τ-tilting modules were introduced by Adachi, Iyama and Reiten [Adachi et al. 2014], they can be regarded as a generalization of tilting modules.

Definition 2.5. Let (X, P) be a pair with $X \in \operatorname{mod}\Lambda$ and $P \in \operatorname{proj}\Lambda$.

1. X is called τ-*rigid* if $\operatorname{Hom}_\Lambda(X, \tau X) = 0$.
2. X is called τ-*tilting* if X is τ-rigid and $|X| = |\Lambda|$.
3. (X, P) is called a τ-*rigid pair* if X is τ-rigid and $\operatorname{Hom}_\Lambda(P, X) = 0$.
4. (X, P) is a *support* τ-*tilting pair* if it is a τ-rigid pair and $|X| + |P| = |\Lambda|$. In this case, X is called a *support* τ-*tilting module*.

Throughout this paper, we denote by τ-rigid Λ the set of isomorphism classes of basic τ-rigid pairs of Λ, and by sτ-tiltΛ the set of isomorphism classes of basic support τ-tilting Λ-modules.

The following proposition gives a criterion for a τ-rigid Λ-module to be a support τ-tilting Λ-module.

Proposition 2.6 [Jasso 2015, Proposition 2.14]. *Let Λ be a finite-dimensional algebra and M a τ-rigid Λ-module. Then M is a support τ-tilting Λ-module if and only if there exists an exact sequence*

$$\Lambda \xrightarrow{f} M' \xrightarrow{g} M'' \to 0,$$

with $M', M'' \in \mathsf{add}\, M$ and f a left $(\mathsf{add}\, M)$-approximation of Λ.

Iyama, Jørgensen and Yang [Iyama et al. 2014, Definition 1.3] defined a functor version of $\tau-$tilting modules, and they extended the notion of support τ-tilting modules for finite dimensional algebras to that for essentially small additive categories. Let $\mathcal{T}$ be an additive category. We write $\mathsf{Mod}\ \mathcal{T}$ for the abelian category of contravariant additive functors from $\mathcal{T}$ to the category of abelian groups and $\mathsf{mod}\ \mathcal{T}$ for the full subcategory of finitely presented functors, see [Auslander 1974].

Definition 2.7 [Iyama et al. 2014, Definition 1.3]. Let $\mathcal{T}$ be an essentially small additive category.

(i) Let $\mathcal{M}$ be a subcategory of $\mathsf{mod}\ \mathcal{T}$. A class $\{\, P_1 \xrightarrow{\pi^M} P_0 \to M \to 0 \mid M \in \mathcal{M}\,\}$ of projective presentations in $\mathsf{mod}\ \mathcal{T}$ is said to have *Property (S)* if

$$\mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\pi^M, M') : \mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(P_0, M') \to \mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(P_1, M')$$

is surjective for any $M, M' \in \mathcal{M}$.

(ii) A subcategory $\mathcal{M}$ of $\mathsf{mod}\ \mathcal{T}$ is said to be *τ-rigid* if there is a class of projective presentations $\{P_1 \to P_0 \to M \to 0 \mid M \in \mathcal{M}\}$ which has Property (S).

(iii) A *τ-rigid pair* of $\mathsf{mod}\ \mathcal{T}$ is a pair $(\mathcal{M}, \mathcal{E})$, where $\mathcal{M}$ is a τ-rigid subcategory of $\mathsf{mod}\ \mathcal{T}$ and $\mathcal{E} \subseteq \mathcal{T}$ is a subcategory with $\mathcal{M}(\mathcal{E}) = 0$, that is, $M(E) = 0$ for each $M \in \mathcal{M}$ and $E \in \mathcal{E}$.

(iv) A τ-rigid pair $(\mathcal{M}, \mathcal{E})$ is *support τ-tilting* if $\mathcal{E} = \mathrm{Ker}\,(\mathcal{M})$ and for each $T \in \mathcal{T}$ there exists an exact sequence $\mathcal{T}(-, T) \xrightarrow{f} M^0 \to M^1 \to 0$ with $M^0, M^1 \in \mathcal{M}$ such that f is a left $\mathcal{M}$-approximation. In this case, $\mathcal{M}$ is called a *support τ-tilting subcategory* of $\mathsf{mod}\, \mathcal{T}$.

From triangulated categories to abelian categories. In this subsection, we assume that $\mathcal{T}$ is a cluster tilting subcategory of a triangulated category $\mathscr{C}$. A $\mathcal{T}$-module is a contravariant k-linear functor $F : \mathcal{T} \to \mathsf{Mod}\, k$. Then $\mathcal{T}$-modules form an abelian category $\mathsf{Mod}\, \mathcal{T}$. We denote by $\mathsf{mod}\, \mathcal{T}$ the subcategory of $\mathsf{Mod}\, \mathcal{T}$ consisting of finitely presented $\mathcal{T}$-modules. It is easy to see that $\mathsf{mod}\, \mathcal{T}$ is an abelian category. Moreover the restricted Yoneda functor

$$\mathbb{H} : \mathscr{C} \to \mathsf{Mod}\ \mathcal{T}, \quad M \mapsto \mathrm{Hom}_{\mathscr{C}}(-, M)\,|_{\mathcal{T}}$$

is homological and induces an equivalence

$$\mathcal{T} \xrightarrow{\sim} \mathsf{proj}(\mathsf{mod}\ \mathcal{T}).$$

The following results are crucial in this paper.

Theorem 2.8. (i) $\mathbb{H}(\mathscr{C})$ *is a subcategory of* $\mathsf{mod}\,\mathcal{T}$.

(ii) [Auslander 1974] *For* $N \in \mathsf{Mod}\ \mathcal{T}$ *and* $T \in \mathcal{T}$, *there exists a natural isomorphism*

$$\mathrm{Hom}_{\mathsf{Mod}\ \mathcal{T}}(\mathbb{H}(T), N) \xrightarrow{\sim} N(T).$$

More explicitly, if we have a map $f : T \to T'$, *where* $T' \in \mathcal{T}$, *then we have the commutative diagram*

$$\begin{tikzcd}
\mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(T'), N) \arrow[r, "\circ\mathbb{H}(f)"] \arrow[d, "\simeq"'] & \mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(T), N) \arrow[d, "\simeq"] \\
N(T') \arrow[r, "N(f)"] & N(T).
\end{tikzcd}$$

(iii) [Keller and Reiten 2007; Koenig and Zhu 2008; Iyama and Yoshino 2008] *The functor* $\mathbb{H}$ *from* (i) *induces an equivalence*

$$\mathscr{C}/[\mathcal{T}[1]] \xrightarrow{\sim} \mathsf{mod}\ \mathcal{T},$$

and $\mathsf{mod}\ \mathcal{T}$ *is Gorenstein of dimension at most one.*

Proof of (i). Since $\mathcal{T}$ is cluster tilting, for any object $C \in \mathscr{C}$, there exists a triangle

$$T_0 \xrightarrow{f} T_1 \xrightarrow{g} C \xrightarrow{h} T_0[1],$$

where $T_0, T_1 \in \mathcal{T}$. Applying the functor $\mathbb{H}$ to the above triangle, we get an exact sequence

$$\mathbb{H}(T_0) \xrightarrow{f\circ} \mathbb{H}(T_1) \xrightarrow{g\circ} \mathbb{H}(C) \longrightarrow 0.$$

This shows that $\mathbb{H}(C) \in \mathsf{mod}\,\mathcal{T}$. □

If there exists an object $T \in \mathscr{C}$ such that $\mathcal{T} = \mathsf{add}\,T$, we obtain the following.

Corollary 2.9. *Let* T *be a cluster tilting object in* $\mathscr{C}$ *and* $\Lambda = \mathrm{End}^{\mathrm{op}}_{\mathscr{C}}(T)$. *Then the functor*

$$\mathrm{Hom}_{\mathscr{C}}(T, -) : \mathscr{C} \to \mathsf{mod}\ \Lambda \tag{2-1}$$

induces an equivalence

$$\overline{(-)} : \mathscr{C}/[T[1]] \xrightarrow{\sim} \mathsf{mod}\ \Lambda. \tag{2-2}$$

This equivalence gives a close relationship between the relative cluster tilting objects in $\mathscr{C}$ and support τ-tilting Λ-modules.

Theorem 2.10 [Yang and Zhu 2019, Theorem 3.6]. *Let $\mathscr{C}$ be a triangulated category with a Serre functor $\mathbb{S}$ and a cluster tilting object T, and let $\Lambda = \operatorname{End}^{\mathrm{op}}_{\mathscr{C}}(T)$. Then the functor* (2-1) *induces the bijections*

$$T[1]\text{-rigid}\,\mathscr{C} \overset{(a)}{\longleftrightarrow} \tau\text{-rigid}\,\Lambda, \qquad T[1]\text{-tilt}\,\mathscr{C} \overset{(b)}{\longleftrightarrow} \mathrm{s}\tau\text{-tilt}\,\Lambda.$$

3. Ghost cluster tilting subcategories

In this section, our aim is to define and study ghost cluster tilting subcategories in a triangulated category with cluster tilting subcategories, in particular, to compare them to the existing notions: cluster tilting subcategories [Keller and Reiten 2007; Koenig and Zhu 2008; Iyama and Yoshino 2008] and relative cluster tilting objects [Yang and Zhu 2019].

Ghost cluster tilting subcategories. We first give the definitions and then discuss connections between them.

Definition 3.1. Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory.

(i) A subcategory $\mathscr{X}$ in $\mathscr{C}$ is called *ghost rigid* if there exists a cluster tilting subcategory $\mathcal{T}$ such that $[\mathcal{T}[1]](\mathscr{X}, \mathscr{X}[1]) = 0$. In this case, $\mathscr{X}$ is also called $\mathcal{T}[1]$*-rigid*.

(ii) A subcategory $\mathscr{X}$ in $\mathscr{C}$ is called *maximal ghost rigid* if there exists a cluster tilting subcategory $\mathcal{T}$ such that $\mathscr{X}$ is $\mathcal{T}[1]$-rigid and

$$[\mathcal{T}[1]](\mathscr{X} \vee \mathsf{add}\, M, (\mathscr{X} \vee \mathsf{add}\, M)[1]) = 0 \text{ implies } M \in \mathscr{X}.$$

In this case, $\mathscr{X}$ is also called *maximal* $\mathcal{T}[1]$*-rigid*.

(iii) A subcategory $\mathscr{X}$ in $\mathscr{C}$ is called *weak ghost cluster tilting* if there exists a cluster tilting subcategory $\mathcal{T}$ with $\mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}$ and

$$\mathscr{X} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \mathscr{X}[1]) = 0 \quad \text{and} \quad [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0\}.$$

In this case, $\mathscr{X}$ is also called *weak* $\mathcal{T}[1]$*-cluster tilting*.

(iv) A subcategory $\mathscr{X}$ in $\mathscr{C}$ is called *ghost cluster tilting* if $\mathscr{X}$ is contravariantly finite in $\mathscr{C}$ and there exists a cluster tilting subcategory $\mathcal{T}$ such that

$$\mathscr{X} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \mathscr{X}[1]) = 0 \quad \text{and} \quad [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0\}.$$

In this case, $\mathscr{X}$ is also called $\mathcal{T}[1]$*-cluster tilting*.

(v) An object X is called $\mathcal{T}[1]$-rigid, maximal $\mathcal{T}[1]$-rigid, weak $\mathcal{T}[1]$-cluster tilting, or $\mathcal{T}[1]$-cluster tilting if $\mathsf{add}\, X$ is $\mathcal{T}[1]$-rigid, maximal $\mathcal{T}[1]$-rigid, weak $\mathcal{T}[1]$-cluster tilting, or $\mathcal{T}[1]$-cluster tilting respectively.

Remark 3.2. Since ghost cluster tilting subcategories are introduced in order to generalize the notion of relative cluster tilting objects, it is natural to compare the definition of relative cluster tilting objects of Definition 2.4 (originally given in [Yang and Zhu 2019]) to the definition of ghost cluster tilting subcategories of Definition 3.1(iv). When $\mathscr{C}$ has a cluster tilting object, we will show that ghost cluster tilting objects are exactly the relative cluster tilting objects in Theorem 3.16. Therefore when $|\mathcal{T}| = \infty$, we replace the condition "$|\mathscr{X}| = |\mathcal{T}|$" by the equation

$$\text{“}\mathscr{X} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \mathscr{X}[1]) = 0 \quad \text{and} \quad [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0\}.\text{”}$$

From here until the end of the section, we prove some properties of ghost cluster tilting subcategories. We first prove a cluster tilting subcategory is a ghost cluster tilting subcategory with respect to any cluster tilting subcategory.

Proposition 3.3. *Cluster tilting subcategories are ghost cluster tilting. More precisely, let $\mathscr{X}$ be a cluster tilting subcategory. Then $\mathscr{X}$ is $\mathcal{T}[1]$-cluster tilting subcategory for any cluster tilting subcategory $\mathcal{T}$.*

Proof. Let $\mathscr{X}$ be an arbitrary cluster tilting subcategory in $\mathscr{C}$. Clearly, $\mathscr{X}$ is contravariantly finite and

$$\mathscr{X} \subseteq \{M \in \mathscr{C} \mid [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 = [\mathcal{T}[1]](M, \mathscr{X}[1])\}.$$

For any object $M \in \{M \in \mathscr{C} \mid [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 = [\mathcal{T}[1]](M, \mathscr{X}[1])\}$, we need to prove that $M \in \mathscr{X}$. Since $\mathcal{T}$ is cluster tilting, there exists a triangle

$$T_1 \xrightarrow{f} T_0 \xrightarrow{g} M \xrightarrow{h} T_1[1],$$

where $T_0, T_1 \in \mathcal{T}$. Take a left $\mathscr{X}$-approximation of T_0 and complete it to a triangle

$$T_0 \xrightarrow{u} X_1 \xrightarrow{v} X_2 \xrightarrow{w} T_0[1],$$

where $X_1 \in \mathscr{X}$. Since $\mathscr{X}$ is cluster tilting, by Proposition 2.3(c), we have that $X_2 \in \mathscr{X}$. By the octahedral axiom, we have a commutative diagram

$$\begin{tikzcd}
T_1 \arrow[r, "f"] \arrow[d, equal] & T_0 \arrow[r, "g"] \arrow[d, "u"] & M \arrow[r, "h"] \arrow[d, "a"] & T_1[1] \arrow[d, equal] \\
T_1 \arrow[r, "x=uf"] & X_1 \arrow[r, "y"] \arrow[d, "v"] & N \arrow[r, "z"] \arrow[d, "b"] & T_1[1] \\
& X_2 \arrow[r, equal] \arrow[d, "w"] & X_2 \arrow[d, "c"] & \\
& T_0[1] \arrow[r, "{g[1]}"] & M[1] &
\end{tikzcd}$$

of triangles. We claim that x is a left $\mathscr{X}$-approximation of T_1. Indeed, for any morphism $\alpha : T_1 \to X'$, where $X' \in \mathscr{X}$, since $\alpha \circ h[-1] \in [\mathcal{T}](M[-1], \mathscr{X}) = 0$, there exists a morphism $\beta : T_0 \to X'$ such that $\alpha = \beta f$. Since u is a left $\mathscr{X}$-approximation of T_0 and $X' \in \mathscr{X}$, there exists a morphism $\gamma : X_1 \to X'$ such that $\beta = \gamma u$ and then

$\alpha = \gamma(uf) = \gamma x$. This shows that x is a left $\mathscr{X}$-approximation of T_1. Note that $\mathscr{X}$ is cluster tilting. Thus $N \in \mathscr{X}$. Since $c = g[1]w \in [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0$, the triangle

$$M \xrightarrow{a} N \xrightarrow{b} X_2 \xrightarrow{c} M[1]$$

splits. It follows that M is a direct summand of N and therefore $M \in \mathscr{X}$. Thus

$$\mathscr{X} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 = [\mathcal{T}[1]](M, \mathscr{X}[1])\}$$

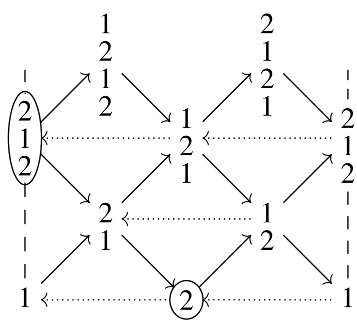
and hence $\mathscr{X}$ is $\mathcal{T}[1]$-cluster tilting. □

The following example shows that ghost cluster tilting subcategories need not be cluster tilting.

Example 3.4. Let $A = kQ/I$ be a self-injective algebra given by the quiver

$$Q:\ 1 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 2$$

and $I = \langle \alpha\beta\alpha\beta, \beta\alpha\beta\alpha \rangle$. Let $\mathscr{C}$ be the stable module category $\underline{\mathrm{mod}}\, A$ of A. This is a triangulated category whose Auslander–Reiten quiver is the following (note that projective-injective modules should be deleted):



where the leftmost and rightmost columns are identified. It is easy to see that

$$\mathcal{T} := \mathsf{add}\left(2 \oplus \begin{smallmatrix}2\\1\\2\end{smallmatrix}\right)$$

is a cluster tilting subcategory of $\mathscr{C}$. Note that $\mathscr{X} := \mathsf{add}\left(2 \oplus \begin{smallmatrix}1\\2\end{smallmatrix}\right)$ is a $\mathcal{T}[1]$-cluster tilting subcategory of $\mathscr{C}$, but not a cluster tilting subcategory of $\mathscr{C}$, since $\mathrm{Hom}\left(\begin{smallmatrix}1\\2\end{smallmatrix}, \begin{smallmatrix}1\\2\end{smallmatrix}[1]\right) = \mathrm{Hom}\left(\begin{smallmatrix}1\\2\end{smallmatrix}, \begin{smallmatrix}1\\2\end{smallmatrix}\right) \neq 0$.

As we have seen in Example 3.4, ghost cluster tilting categories need not be cluster tilting categories, however the situation is much better when we assume that the triangulated category $\mathscr{C}$ has a Serre functor $\mathbb{S}$ as we will show in Theorem 3.6.

We need the following lemma in order to prove Theorem 3.6:

Lemma 3.5. *Let $\mathscr{C}$ be a triangulated category with a Serre functor $\mathbb{S}$ and a cluster tilting subcategory $\mathcal{T}$. For two objects M and N in $\mathscr{C}$, $[\mathcal{T}[1]](M, N[1]) = 0$ and*

$[\mathcal{T}[1]](N, \tau M) = 0$ *if and only if* $\mathrm{Hom}_{\mathscr{C}}(M, N[1]) = 0$. *In particular*, *if* $\mathscr{C}$ *is* 2-*Calabi–Yau*, *then* M *is* $\mathcal{T}[1]$-*rigid if and only if* M *is rigid.*

Proof. Our argument is similar to the proof of Proposition 3.4 in [Yang and Zhu 2019]. We give the proof for the convenience of the reader.

We show the "if" part. If $\mathrm{Hom}_{\mathscr{C}}(M, N[1]) = 0$, then $[\mathcal{T}[1]](M, N[1]) = 0$. By the Serre duality, we have

$$\mathrm{Hom}_{\mathscr{C}}(N, \tau M) \simeq \mathrm{Hom}_{\mathscr{C}}(N[1], \mathbb{S}M) \simeq D\,\mathrm{Hom}_{\mathscr{C}}(M, N[1]) = 0.$$

Thus we obtain $[\mathcal{T}[1]](N, \tau M) = 0$.

We now show the "only if" part. Since $\mathcal{T}$ is a cluster tilting subcategory, by Proposition 2.3(a), we have a triangle

$$T_0 \xrightarrow{g} N \xrightarrow{f} T_1[1] \xrightarrow{h} T_0[1]$$

with $T_0, T_1 \in \mathcal{T}$. Thus we have a commutative diagram of exact sequences:

$$\begin{tikzcd}
\mathrm{Hom}_{\mathscr{C}}(T_1[1], \tau M) \arrow[r, "\sim"] \arrow[d, "\cdot f"'] & D\,\mathrm{Hom}_{\mathscr{C}}(M, T_1[2]) \arrow[d, "D(f[1]\cdot)"] \\
\mathrm{Hom}_{\mathscr{C}}(N, \tau M) \arrow[r, "\sim"] \arrow[d] & D\,\mathrm{Hom}_{\mathscr{C}}(M, N[1]) \arrow[d, "D(g[1]\cdot)"] \\
\mathrm{Hom}_{\mathscr{C}}(T_0, \tau M) \arrow[r, "\sim"] \arrow[d] & D\,\mathrm{Hom}_{\mathscr{C}}(M, T_0[1]) \arrow[d, "D(h\cdot)"] \\
\mathrm{Hom}_{\mathscr{C}}(T_1, \tau M) \arrow[r, "\sim"] & D\,\mathrm{Hom}_{\mathscr{C}}(M, T_1[1])
\end{tikzcd}$$

Since $\mathrm{Im}(\cdot f) = \{af \mid a \in \mathrm{Hom}_{\mathscr{C}}(T_1[1], \tau M)\} \subseteq [\mathcal{T}[1]](N, \tau M) = 0$, we deduce

$$\mathrm{Ker} D(g[1]\cdot) = \mathrm{Im} D(f[1]\cdot) \simeq \mathrm{Im}(\cdot f) = 0. \tag{3-1}$$

$$\begin{tikzcd}
N \arrow[r, "f"] \arrow[d] & T_1[1] \arrow[r, "h"] \arrow[dl] & T_0[1] \arrow[r, "g[1]"] & N[1] \\
\tau M & & M \arrow[u, "b"'] \arrow[ul, dotted, "\exists c"] &
\end{tikzcd}$$

Take any $b \in \mathrm{Hom}_{\mathscr{C}}(M, T_0[1])$. Since $[\mathcal{T}[1]](M, N[1]) = 0$, we have $g[1]b = 0$. Thus there exists $c : M \to T_1[1]$ such that $b = hc$, which implies that

$$(h\cdot) : \mathrm{Hom}_{\mathscr{C}}(M, T_1[1]) \longrightarrow \mathrm{Hom}_{\mathscr{C}}(M, T_0[1]), \qquad c \longmapsto hc = b$$

is surjective. Therefore, $D(h\cdot)$ is injective and

$$\operatorname{Im} D(g[1]\cdot) = \operatorname{Ker} D(h\cdot) = 0.$$

Combining this with (3-1), we deduce $D\operatorname{Hom}_{\mathscr{C}}(M, N[1]) = 0$ and $\operatorname{Hom}_{\mathscr{C}}(M, N[1])$ vanishes.

If $\mathscr{C}$ is 2-Calabi–Yau, then $\tau \simeq [1]$. The assertion is clear. □

The following result gives a characterization of cluster tilting subcategories in terms of ghost cluster tilting subcategories, which implies that in a 2-Calabi–Yau triangulated category, ghost cluster tilting subcategories coincide with cluster tilting subcategories.

Theorem 3.6. *Let $\mathscr{C}$ be a triangulated category with a Serre functor $\mathbb{S}$ and a cluster tilting subcategory. Then F-stable ghost cluster tilting subcategories of $\mathscr{C}$ are precisely cluster tilting subcategories, where $F = \tau^{-1}[1] = \mathbb{S}^{-1}[2]$.*

Proof. By Proposition 2.3, we have that cluster tilting subcategories are F-stable. By Proposition 3.3, we have that cluster tilting subcategories are ghost cluster tilting. Now we prove the other direction. Let $\mathscr{X}$ be a $\mathcal{T}[1]$-cluster tilting subcategory satisfying $F\mathscr{X} = \mathscr{X}$, where $\mathcal{T}$ is a cluster tilting subcategory. It follows that $\tau\mathscr{X} = \mathscr{X}[1]$.

(1) We show that $\mathscr{X}$ is a rigid subcategory of $\mathscr{C}$. For any two objects $M, N \in \mathscr{X}$, since $\mathscr{X}$ is $\mathcal{T}[1]$-tilting, we have

$$[\mathcal{T}[1]](M, N[1]) = 0. \tag{3-2}$$

Similarly, since $\tau\mathscr{X} = \mathscr{X}[1]$, we have $\tau M = M'[1]$, where $M' \in \mathscr{X}$. It follows that

$$[\mathcal{T}[1]](N, \tau M) = [\mathcal{T}[1]](N, M'[1]) = 0. \tag{3-3}$$

By Lemma 3.5, equalities (3-2) and (3-3) imply that $\operatorname{Hom}_{\mathscr{C}}(M, N[1]) = 0$.

(2) We show that $\mathscr{X} = \{M \in \mathscr{C} \mid \operatorname{Ext}^1_{\mathscr{C}}(\mathscr{X}, M) = 0\}$. The "$\subseteq$" part is clear. Assume that an object $M \in \mathscr{C}$ satisfies $\operatorname{Ext}^1_{\mathscr{C}}(\mathscr{X}, M) = 0$. Then

$$\begin{aligned}\operatorname{Hom}_{\mathscr{C}}(M, \mathscr{X}[1]) &\simeq D\operatorname{Hom}_{\mathscr{C}}(\mathscr{X}[1], \mathbb{S}M)\\ &\simeq D\operatorname{Hom}_{\mathscr{C}}(\tau\mathscr{X}, F\mathbb{S}M)\\ &\simeq D\operatorname{Hom}_{\mathscr{C}}(\mathscr{X}, M[1]) = 0.\end{aligned}$$

This implies that $[\mathcal{T}[1]](M, \mathscr{X}[1]) = 0$. Since $[\mathcal{T}[1]](\mathscr{X}, M[1]) = 0$ and $\mathscr{X}$ is $\mathcal{T}[1]$-cluster tilting, we obtain that $M \in \mathscr{X}$.

Note that $\mathscr{X}$ is contravariantly finite. It follows from Remark 2.2 that $\mathscr{X}$ is a cluster tilting subcategory of $\mathscr{C}$. □

Proposition 3.7. *Any ghost cluster tilting subcategory is a contravariantly finite maximal ghost rigid subcategory.*

Proof. Assume that $\mathscr{X}$ is a $\mathcal{T}[1]$-cluster tilting subcategory. If there exists an object $M \in \mathscr{C}$ such that

$$[\mathcal{T}[1]](\mathscr{X} \vee \mathsf{add}\, M, (\mathscr{X} \vee \mathsf{add}\, M)[1]) = 0,$$

then

$$[\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 \quad \text{and} \quad [\mathcal{T}[1]](M, \mathscr{X}[1]) = 0.$$

Since $\mathscr{X}$ is $\mathcal{T}[1]$-cluster tilting, we obtain $M \in \mathscr{X}$. $\square$

The converse result to Proposition 3.7 will be given in Theorem 3.10. We need the following lemma to prove this theorem.

Lemma 3.8. (a) *Let $\mathcal{T}$ be a cluster tilting subcategory and $\mathscr{X}$ a maximal $\mathcal{T}[1]$-rigid subcategory in $\mathscr{C}$. Let $T_0 \in \mathcal{T}$, let $T_0 \xrightarrow{g} X_0$ be a left $\mathscr{X}$-approximation of T_0 and consider the associated triangle*:

$$M[-1] \xrightarrow{f} T_0 \xrightarrow{g} X_0 \xrightarrow{h} M.$$

Then $M \in \mathscr{X}$.

(b) *Let $\mathcal{T}$ be a cluster tilting subcategory and $\mathscr{X}$ a maximal $\mathcal{T}[1]$-rigid subcategory in $\mathscr{C}$. Let $T_0 \in \mathcal{T}$, let $\mathscr{X}_0[-1] \xrightarrow{f} T_0$ be a left $\mathscr{X}[-1]$-approximation of T_0 and consider the associated triangle*:

$$X_0[-1] \xrightarrow{f} T_0 \xrightarrow{g} M \xrightarrow{h} X_0.$$

Then $M \in \mathscr{X}$.

Proof. We only prove (a), the proof of (b) is similar. For any $x \in [\mathcal{T}](M[-1], \mathscr{X})$, there are two morphisms $x_1 : M[-1] \to T_1$ and $x_2 : T_1 \to X_1$ such that $x = x_2 x_1$, where $T_1 \in \mathcal{T}$ and $X_1 \in \mathscr{X}$.

$$\begin{tikzcd}
X_0[-1] \arrow[r, "{h[-1]}"] & M[-1] \arrow[r, "f"] \arrow[d, "x_1"] & T_0 \arrow[r, "g"] \arrow[ddl, dotted, "a"] & X_0 \arrow[r, "h"] \arrow[ddll, dotted, "b"] & M \arrow[r, "{-f[1]}"] & T_0[1] \\
 & T_1 \arrow[d, "x_2"] & & & T_2 \arrow[u, "y_2"] \arrow[ul, dotted, "c"] & \\
 & X_1 & & & X_2[-1] \arrow[u, "y_1"] &
\end{tikzcd}$$

Since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $xh[-1] = x_2(x_1 h[-1]) = 0$. Thus, there exists $a : T_0 \to X_1$ such that $x = af$. Because g is a left $\mathscr{X}$-approximation of T_0, we deduce that there exists $b : X_0 \to X_1$ such that $a = bg$. Therefore, $x = af = b(gf) = 0$ and

$$[\mathcal{T}[1]](M, \mathscr{X}[1]) = 0. \tag{3-4}$$

For $y \in [\mathcal{T}](\mathscr{X}[-1], M)$, there are two morphisms $y_1 : X_2[-1] \to T_2$ and $y_2 : T_2 \to M$ such that $y = y_2 y_1$, where $T_2 \in \mathcal{T}$ and $X_2 \in \mathscr{X}$. Since $f[1]y_2 = 0$, there exists $c : T_2 \to X_0$ such that $y_2 = hc$. Because $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $y = y_2 y_1 = h(cy_1) = 0$. Therefore,

$$[\mathcal{T}[1]](\mathscr{X}, M[1]) = 0. \tag{3-5}$$

For any $z \in [\mathcal{T}](M[-1], M)$, there are two morphisms $z_1 : M[-1] \to T_3$ and $z_2 : T_3 \to M$ such that $z = z_2 z_1$, where $T_3 \in \mathcal{T}$. Since $f[1]z_2 = 0$, there exists $d : T_3 \to X_0$ such that $z_2 = hd$. By equality (3-4), we have $z = z_2 z_1 = h(dz_1) = 0$. Thus,

$$[\mathcal{T}[1]](M, M[1]) = 0. \tag{3-6}$$

$$\begin{tikzcd}
T_0 \arrow[r, "g"] & X_0 \arrow[r, "h"] & M \arrow[r, "-f[1]"] & T_0[1] \\
 & & T_3 \arrow[u, "z_2"] \arrow[ul, dotted, "d"] & \\
 & & M[-1] \arrow[u, "z_1"] &
\end{tikzcd}$$

Using (3-4), (3-5) and (3-6), we get $[\mathcal{T}[1]](\mathscr{X} \vee \operatorname{add} M, (\mathscr{X} \vee \operatorname{add} M)[1]) = 0$. Note that $\mathscr{X}$ is maximal $\mathcal{T}[1]$-rigid. Hence $M \in \mathscr{X}$. □

This lemma immediately yields the following important conclusion:

Corollary 3.9. *Let $\mathcal{T}$ be a cluster tilting subcategory in a triangulated category $\mathscr{C}$ and $\mathscr{X}$ be a covariantly* (*or contravariantly*) *finite maximal $\mathcal{T}[1]$-rigid subcategory. Then*

$$\mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}.$$

Now we prove that the converse of Proposition 3.7 also holds, which generalizes a result of Zhou and Zhu [2011, Theorem 2.6].

Theorem 3.10. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. Then any contravariantly finite maximal $\mathcal{T}[1]$-rigid subcategory is a $\mathcal{T}[1]$-cluster tilting subcategory.*

Proof. Assume that $\mathscr{X}$ is a contravariantly finite maximal $\mathcal{T}[1]$-rigid subcategory in $\mathscr{C}$. Clearly,

$$\mathscr{X} \subseteq \{M \in \mathscr{C} \mid [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 = [\mathcal{T}[1]](M, \mathscr{X}[1])\}.$$

For any object $M \in \{M \in \mathscr{C} \mid [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0 = [\mathcal{T}[1]](M, \mathscr{X}[1])\}$, since $\mathcal{T}$ is cluster tilting, there exists a triangle

$$T_1 \xrightarrow{f} T_0 \xrightarrow{g} M \xrightarrow{h} T_1[1],$$

where $T_0, T_1 \in \mathcal{T}$. By Corollary 3.9, there exists a triangle

$$T_0 \xrightarrow{u} X_1 \xrightarrow{v} X_2 \xrightarrow{w} T_0[1],$$

where $X_1, X_2 \in \mathscr{X}$. Since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have that u is a left $\mathscr{X}$-approximation of T_0. By the octahedral axiom, we have a commutative diagram

$$\begin{tikzcd}
T_1 \arrow[r, "f"] \arrow[d, equal] & T_0 \arrow[r, "g"] \arrow[d, "u"] & M \arrow[r, "h"] \arrow[d, "a"] & T_1[1] \arrow[d, equal] \\
T_1 \arrow[r, "x=uf"] & X_1 \arrow[r, "y"] \arrow[d, "v"] & N \arrow[r, "z"] \arrow[d, "b"] & T_1[1] \\
& X_2 \arrow[r, equal] \arrow[d, "w"] & X_2 \arrow[d, "c"] & \\
& T_0[1] \arrow[r, "{g[1]}"] & M[1] &
\end{tikzcd}$$

of triangles. Using similar arguments as in the proof of Proposition 3.3, we conclude that x is a left $\mathscr{X}$-approximation of T_1. By Lemma 3.8, we have $N \in \mathscr{X}$. Since

$$c = g[1]w \in [\mathcal{T}[1]](\mathscr{X}, M[1]) = 0.$$

This shows that the triangle

$$M \xrightarrow{a} N \xrightarrow{b} X_2 \xrightarrow{c} M[1].$$

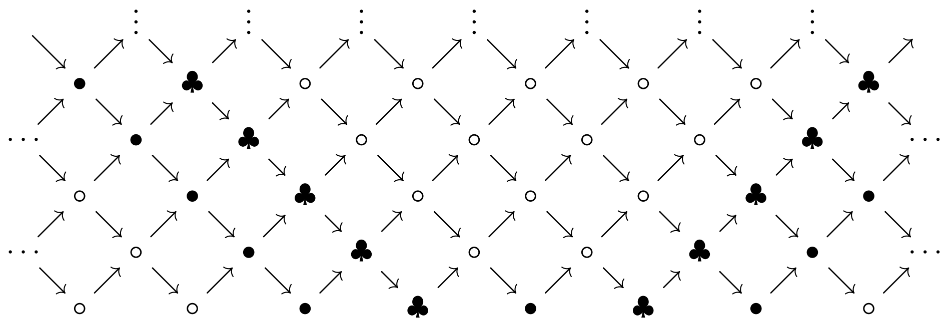
splits. It follows that M is a direct summand of N and thus $M \in \mathscr{X}$. Hence $\mathscr{X}$ is $\mathcal{T}[1]$-cluster tilting. □

Corollary 3.11. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. Then $\mathcal{T}[1]$-cluster tilting subcategories are weak $\mathcal{T}[1]$-cluster tilting subcategories.*

Proof. This follows from Theorem 3.10 and Corollary 3.9. □

The following example shows the converse is not true. More precisely, weak $\mathcal{T}[1]$-cluster tilting subcategories are not usually $\mathcal{T}[1]$-cluster tilting subcategories.

Example 3.12. The cluster category of type $\mathbb{A}_\infty$ was introduced in [Holm and Jørgensen 2012; Ng 2010]. This definition is completely analogous to the definition of the cluster category of type $\mathbb{A}_n$. Namely, it is the orbit category $D^f(\operatorname{mod}\Gamma)/\mathbb{S}[-2]$. Here Γ is a quiver of type $\mathbb{A}_\infty$ with zigzag orientation and $\mathbb{S}$ and [1] are the Serre and shift functors of the finite derived category $D^f(\operatorname{mod}\Gamma)$. Let $\mathscr{C}$ be a cluster category of type $\mathbb{A}_\infty$. The Auslander–Reiten quiver of $\mathscr{C}$ is as follows:



Set $\mathscr{X}$ to be the subcategory whose indecomposable objects are marked by bullets here, and $\mathcal{T}$ to be the subcategory whose indecomposable objects are marked by clubsuits here. It is easy to see that $\mathcal{T}$ is a cluster tilting subcategory of $\mathscr{C}$ and $\mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}$. By [Holm and Jørgensen 2012, Theorem 4.3], we have that $\mathscr{X}$ is a weak cluster tilting subcategory of $\mathscr{C}$ since the corresponding set of arcs is a maximal set of noncrossing arcs. By [Holm and Jørgensen 2012, Theorem 4.4], we obtain that $\mathscr{X}$ is not contravariantly finite in $\mathscr{C}$ since the corresponding maximal set of noncrossing arcs has no right-fountain. That is to say, $\mathscr{X}$ is a weak ghost cluster tilting subcategory in the sense of Definition 3.1, but it is not ghost cluster tilting (=cluster tilting).

As an application of Theorem 3.10, we have the following:

Corollary 3.13 [Zhou and Zhu 2011, Theorem 2.6]. *Let $\mathscr{C}$ be a 2-Calabi–Yau triangulated category with a cluster tilting subcategory $\mathcal{T}$. Then every functorially finite maximal rigid subcategory is cluster-tilting.*

Proof. This follows from Lemma 3.5 and Theorem 3.10. □

We give a characterization of weak ghost cluster tilting subcategories.

Theorem 3.14. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$, and $\mathscr{X}$ a subcategory of $\mathscr{C}$. Then $\mathscr{X}$ is a weak ghost cluster tilting subcategory if and only if $\mathscr{X}$ is a maximal ghost rigid subcategory such that $\mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}$.*

Proof. This follows from similar arguments as in the proof of Theorem 3.10. □

We conclude with a picture illustrating the relationships between ghost cluster tilting subcategories and related subcategories:

$$
\begin{tikzcd}
\boxed{\text{cluster tilting}} \arrow[d, shift left, "\text{Proposition 3.3}"] & \\
\boxed{\text{ghost cluster tilting}} \arrow[u, shift left, "\substack{\text{F-stable}\\ \text{Theorem 3.6}}"] \arrow[r, shift left, "\text{Corollary 3.11}"] \arrow[d, shift left, "\text{Proposition 3.7}"] & \boxed{\text{weak ghost cluster tilting}} \arrow[l, shift left, "\substack{\text{contravariantly finite}\\ \text{Definition 3.1}}"] \\
\boxed{\text{maximal ghost rigid}} \arrow[u, shift left, "\substack{\text{contravariantly finite}\\ \text{Theorem 3.10}}"] &
\end{tikzcd}
$$

A characterization of ghost cluster tilting objects. In this subsection, we always assume that $\mathscr{C}$ is a triangulated category with a Serre functor and a cluster tilting object T. We will prove that the $\operatorname{add}(T[1])$-cluster tilting objects are precisely the $T[1]$-cluster tilting objects introduced in [Yang and Zhu 2019], see Definition 2.4. Notice that the two objects have similar names but quite different definitions. To prove it, we need a lemma:

Lemma 3.15. (a) *Let T be a cluster tilting object and X a $T[1]$-cluster tilting object in $\mathscr{C}$. Let $T_0 \in \operatorname{add} T$, let $g : T_0 \longrightarrow X_0$ be a left $\operatorname{add} X$-approximation of T_0 and consider the associated triangle*:

$$M[-1] \xrightarrow{f} T_0 \xrightarrow{g} X_0 \xrightarrow{h} M.$$

Then $M \in \operatorname{add} X$.

(b) *Let T be a cluster tilting object and X a $T[1]$-cluster tilting object in $\mathscr{C}$. Let $T_0 \in \operatorname{add} T$, let $f : X_0[-1] \to T_0$ be a right $\operatorname{add} X[-1]$-approximation of T_0 and consider the associated triangle*:

$$X_0[-1] \xrightarrow{f} T_0 \xrightarrow{g} M \xrightarrow{h} X_0.$$

Then $M \in \operatorname{add} X$.

Proof. Using similar arguments as in the proof of Lemma 3.8 we conclude that

$$[T[1]](X \oplus M, (X \oplus M)[1]) = 0.$$

By Corollary 3.7(1) in [Yang and Zhu 2019], we know that the number of nonisomorphic indecomposable direct summands of any $T[1]$-rigid object is at most the number of nonisomorphic indecomposable direct summands of a cluster tilting object. Thus we have $|X \oplus M| \leq |T|$. Since $|X| = |T|$, we deduce that $M \in \operatorname{add} X$. $\square$

Our main result in this subsection is the following:

Theorem 3.16. *Let T be a cluster tilting object in a triangulated category $\mathscr{C}$ with a Serre functor. Let X be an object in $\mathscr{C}$. Then X is an* $\mathsf{add}(T[1])$*-cluster tilting object, that is to say,*

$$\mathsf{add}\,X = \{M \in \mathscr{C} \mid \mathsf{add}(T[1])(X, M[1]) = 0 \quad \textit{and} \quad \mathsf{add}(T[1])(M, X[1]) = 0\},$$

by Definition 3.1, if and only if X is a $T[1]$-cluster tilting object, that is to say,

$$T[1](X, X[1]) = 0 \quad \textit{and} \quad |X| = |T|,$$

by Definition 2.4.

Proof. (1) The "only if" part: Assume that X is an $\mathsf{add}\,T[1]$-cluster tilting object. Then X is $T[1]$-rigid. By [Yang and Zhu 2019, Corollary 3.7(2)], there exists an object $M \in \mathscr{C}$ such that $X \oplus M$ is a $T[1]$-cluster tilting object. That is to say, $X \oplus M$ is $T[1]$-rigid and $|X \oplus M| = |T|$. Since $X \oplus M$ is $T[1]$-rigid, we have

$$[T[1]](M, X[1]) = 0 = [T[1]](X, M[1]).$$

By Definition 3.1(iv), we have $M \in \mathsf{add}\,X$. It follows that $|X| = |X \oplus M| = |T|$. This shows that X is $T[1]$-cluster tilting.

(2) The "if" part: Assume X is a $T[1]$-cluster tilting object in Definition 2.4. Clearly,

$$\mathsf{add}\,X \subseteq \{M \in \mathscr{C} \mid [T[1]](X, M[1]) = 0 = [T[1]](M, X[1])\}.$$

Conversely, for any object $M \in \{M \in \mathscr{C} \mid [T[1]](X, M[1]) = 0 = [T[1]](M, X[1])\}$, since T is cluster tilting, we have a triangle

$$T_1 \xrightarrow{f} T_0 \xrightarrow{g} M \xrightarrow{h} T_1[1],$$

where $T_0, T_1 \in \mathsf{add}\,T$. For the object $T_0 \in \mathsf{add}\,T$, there is a left $\mathsf{add}\,X$-approximation l_1 of T_0, which can be extended to a triangle

$$X_1[-1] \xrightarrow{m} T_0 \xrightarrow{l_1} X_0 \to X_1.$$

By Lemma 3.15, we have $X_1 \in \mathsf{add}\,X$.

Let $l_2 = l_1 f$. It is easy to see that l_2 is a left $\mathsf{add}\,X$-approximation of T_1. Indeed, for any object $X' \in \mathsf{add}\,X$ and any map $a \in \mathrm{Hom}(T_1, X')$, we have that $ah[-1] \in [T](M[-1], X') = 0$. Then there exists $b : T_0 \longrightarrow X'$ such that $a = bf$. Because l_1 is a left $\mathsf{add}\,X$-approximation of T_0, there is a map $c : X_0 \longrightarrow X'$ such that $b = cl_1$. Therefore $a = bf = c(l_1 f) = cl_2$ and l_2 is a left $(\mathsf{add}\,X)$-approximation of T_1.

$$\begin{tikzcd}
M[-1] \arrow[r, "{h[-1]}"] & T_1 \arrow[r, "f"] \arrow[d, "\forall a"'] & T_0 \arrow[r, "g"] \arrow[d, "l_1"] \arrow[dl, dotted, "b"] & M \arrow[r, "h"] & T_1[1] \\
& X' & X_0 \arrow[l, dotted, "c"] & &
\end{tikzcd}$$

Using Lemma 3.15, we get a triangle

$$X_2[-1] \to T_1 \xrightarrow{l_2} X_0 \to X_2,$$

where $X_2 \in \mathsf{add}\, X$. Starting with $l_2 = l_1 f$, we get the following commutative diagram by the octahedral axiom.

$$\begin{tikzcd}
& M[-1] \arrow[r, equal] \arrow[d] & M[-1] \arrow[d] & \\
X_0[-1] \arrow[r] \arrow[d, equal] & X_2[-1] \arrow[r] \arrow[d] & T_1 \arrow[r, "l_2"] \arrow[d, "f"] & X_0 \arrow[d, equal] \\
X_0[-1] \arrow[r] & X_1[-1] \arrow[r, "m"] \arrow[d, "n"] & T_0 \arrow[r, "l_1"] \arrow[d, "g"] & X_0 \\
& M \arrow[r, equal] & M &
\end{tikzcd}$$

Since $n = gm \in [T](X_1[-1], M) = 0$, we get a split triangle and thus $M \in \mathsf{add}\, X$. This shows that X is an $\mathsf{add}\, T[1]$-cluster tilting object. $\square$

Remark 3.17. Let $\mathscr{C}$ be a triangulated category with a cluster tilting object T. One may want to define $T[1]$-cluster tilting objects in the spirit of Definition 2.1 as one of the following two possibilities:

$$\mathsf{add}\, X = \{M \in \mathscr{C} \mid [T[1]](X, M[1]) = 0\}, \tag{3-7}$$

or

$$\mathsf{add}\, X = \{M \in \mathscr{C} \mid [T[1]](X, M[1]) = 0\} = \{M \in \mathscr{C} \mid [T[1]](M, X[1]) = 0\}. \tag{3-8}$$

However, neither one of these agrees with the description in Definition 2.4 of a $T[1]$-cluster tilting object which is a $T[1]$-rigid object with the same number of non-isomorphic indecomposable direct summands as $|T|$, as one can see in Example 3.18.

Example 3.18. Let Q be the quiver $1 \xrightarrow{\alpha} 2$ and τ_Q be the Auslander–Reiten translation in $D^b(kQ)$. We consider a triangulated category, named *repetitive cluster category* in [Zhu 2011], $\mathscr{C} = D^b(kQ)/\langle \tau_Q^{-2}[2] \rangle$, whose objects are the same in $D^b(kQ)$, and whose morphisms are given by

$$\mathrm{Hom}_{D^b(kQ)/\langle \tau_Q^{-2}[2] \rangle}(X, Y) = \bigoplus_{i \in \mathbb{Z}} \mathrm{Hom}_{D^b(kQ)}(X, (\tau_Q^{-2}[2])^i Y).$$

We depict the Auslander–Reiten quiver of $\mathscr{C}$ as follows.

$$\begin{tikzcd}[column sep=small]
\cdots & \substack{1\\2} \arrow[l, dotted] \arrow[dr] & & 2[1] \arrow[ll, dotted] \arrow[dr] & & 1[1] \arrow[ll, dotted] \arrow[dr] & & \substack{1\\2}[2] \arrow[ll, dotted] \arrow[dr] & & 2[3] \arrow[ll, dotted] \arrow[dr] & & \substack{1\\2} \arrow[ll, dotted] \\
2 \arrow[ur] & & 1 \arrow[ll, dotted] \arrow[ur] & & \substack{1\\2}[1] \arrow[ll, dotted] \arrow[ur] & & 2[2] \arrow[ll, dotted] \arrow[ur] & & 1[2] \arrow[ll, dotted] \arrow[ur] & & 2 \arrow[ll, dotted] \arrow[ur]
\end{tikzcd}$$

It is easy to check that the direct sum

$$T = 1 \oplus 2[1] \oplus \substack{1\\2}\,[2] \oplus 1[2]$$

of the encircled indecomposable objects is a cluster tilting object. Thus it is also a $T[1]$-cluster tilting object. Clearly,

$$\{M \in \mathscr{C} \mid [T[1]](T, M[1]) = 0\} = \mathscr{C} \neq \mathsf{add}\, T,$$

which means that (3-7) or (3-8) does not hold.

4. Connection with τ-tilting theory

Throughout this section, we assume that $\mathscr{C}$ is a k-linear, Hom-finite triangulated category with a cluster tilting subcategory $\mathcal{T}$. It is well known that the category $\mathsf{mod}\,\mathcal{T}$ of coherent $\mathcal{T}$-modules is abelian. By Theorem 2.8, we know that the restricted Yoneda functor $\mathbb{H} : \mathscr{C} \to \mathsf{mod}\,\mathcal{T}$ induces an equivalence

$$\mathscr{C}/[\mathcal{T}[1]] \xrightarrow{\sim} \mathsf{mod}\ \mathcal{T}.$$

We will investigate this relationship between $\mathscr{C}$ and $\mathsf{mod}\ \mathcal{T}$ via $\mathbb{H}$ more closely.

On the relationship between ghost cluster tilting and support τ-tilting. In this subsection, we give a direct connection between ghost cluster tilting subcategories of $\mathscr{C}$ and support τ-tilting pairs of $\mathsf{mod}\,\mathcal{T}$. We start with the following important observation.

Lemma 4.1. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$ and $\mathscr{X}$ a subcategory of $\mathscr{C}$. For any object $X \in \mathscr{X}$, let*

$$T_1 \xrightarrow{f} T_0 \xrightarrow{g} X \xrightarrow{h} T_1[1] \tag{4-1}$$

be a triangle in $\mathscr{C}$ with $T_0, T_1 \in \mathcal{T}$. Then applying the functor $\mathbb{H}$ gives a projective presentation

$$P_1^{\mathbb{H}(X)} \xrightarrow{\pi^{\mathbb{H}(X)}} P_0^{\mathbb{H}(X)} \to \mathbb{H}(X) \to 0 \tag{4-2}$$

in $\mathsf{mod}\ \mathcal{T}$, and $\mathscr{X}$ is a $\mathcal{T}[1]$-rigid subcategory if and only if the class $\{\pi^{\mathbb{H}(X)} \mid X \in \mathscr{X}\}$ has Property (S).

Proof. Applying $\mathbb{H}$ to the triangle (4-1), we have the projective presentation (4-2). By Theorem 2.8(ii), for any object $X' \in \mathscr{X}$, we have the commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(\mathbb{H}(T_0), \mathbb{H}(X')) \arrow[r, "\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(\pi^{\mathbb{H}(X)},\, \mathbb{H}(X'))"] \arrow[d, "\simeq"'] & \mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(\mathbb{H}(T_1), \mathbb{H}(X')) \arrow[d, "\simeq"] \\
\mathrm{Hom}_{\mathscr{C}}(T_0, X') \arrow[r, "\mathrm{Hom}_{\mathscr{C}}(f,\, X')"] & \mathrm{Hom}_{\mathscr{C}}(T_1, X').
\end{tikzcd}$$

Thus the map $\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(\pi^{\mathbb{H}(X)}, \mathbb{H}(X'))$ is the same as

$$\mathrm{Hom}_{\mathscr{C}}(T_0, X') \xrightarrow{\mathrm{Hom}_{\mathscr{C}}(f,\, X')} \mathrm{Hom}_{\mathscr{C}}(T_1, X'). \tag{4-3}$$

So the class $\{\pi^{\mathbb{H}(X)} \mid X \in \mathscr{X}\}$ has Property (S) if and only if the morphism (4-3) is surjective for all $X, X' \in \mathscr{X}$.

Assume the class $\{\pi^{\mathbb{H}(X)} \mid X \in \mathscr{X}\}$ has Property (S). For any $a \in [\mathcal{T}[1]](\mathscr{X}, \mathscr{X}[1])$, we know that there exist two morphisms $a_1 : X \to T[1]$ and $a_2 : T[1] \to X'[1]$ such that $a = a_2 a_1$, where $X, X' \in \mathscr{X}$ and $T \in \mathcal{T}$. Since $\mathrm{Hom}_{\mathscr{C}}(T_0, T[1]) = 0$, there exists a morphism $b : T_1[1] \to T[1]$ such that $a_1 = bh$.

$$\begin{tikzcd}
T_1 \arrow[r, "f"] & T_0 \arrow[r, "g"] & X \arrow[r, "h"] \arrow[d, "a_1"] & T_1[1] \arrow[dl, dotted, "b"] \\
 & & T[1] \arrow[d, "a_2"] & \\
 & & X'[1] &
\end{tikzcd}$$

Since $\mathrm{Hom}_{\mathscr{C}}(f, X')$ is surjective, there exists a morphism $c : T_0 \to X'$ such that $a_2[-1] \circ b[-1] = cf$ and thus $a_2 b = c[1] \circ f[1]$. It follows that $a = a_2 a_1 = a_2 bh = c[1] \circ (f[1]h) = 0$. This shows that $[\mathcal{T}[1]](\mathscr{X}, \mathscr{X}[1]) = 0$. Hence $\mathscr{X}$ is a $\mathcal{T}[1]$-rigid subcategory.

Conversely, assume that $\mathscr{X}$ is a $\mathcal{T}[1]$-rigid subcategory. For any morphism $x : T_1 \to X'$, since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $x \circ h[-1] = 0$. So there exists a morphism $y : T_0 \to X'$ such that $x = yf$.

$$\begin{tikzcd}
X[-1] \arrow[r, "{h[-1]}"] & T_1 \arrow[r, "f"] \arrow[d, "x"] & T_0 \arrow[r, "g"] \arrow[dl, dotted, "y"] & X \arrow[r, "h"] & T_1[1] \\
 & X' & & &
\end{tikzcd}$$

This shows that $\mathrm{Hom}_{\mathscr{C}}(f, X') : \mathrm{Hom}_{\mathscr{C}}(T_0, X') \to \mathrm{Hom}_{\mathscr{C}}(T_1, X')$ is surjective. By the above discussion, we deduce that the class $\{\pi^{\mathbb{H}(X)} \mid X \in \mathscr{X}\}$ has Property (S). $\square$

The following result plays an important role in this paper:

Proposition 4.2. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. The functor $\mathbb{H} : \mathscr{C} \to \mathsf{Mod}\,\mathcal{T}$ induces a bijection between the sets of $\mathcal{T}[1]$-rigid subcategories of $\mathscr{C}$ and of τ-rigid pairs of* $\mathsf{mod}\,\mathcal{T}$, *given by*

$$\Phi : \mathscr{X} \longmapsto (\mathbb{H}(\mathscr{X}), \mathcal{T} \cap \mathscr{X}[-1]).$$

Proof. **Step 1**: The map Φ has values in τ-rigid pairs of $\mathsf{mod}\,\mathcal{T}$.

Assume that $\mathscr{X}$ is a $\mathcal{T}[1]$-rigid subcategory of $\mathscr{C}$. Since $\mathcal{T}$ is a cluster tilting subcategory, for any $X \in \mathscr{X}$, there exists a triangle in $\mathscr{C}$

$$T_1 \xrightarrow{f} T_0 \xrightarrow{g} X \xrightarrow{h} T_1[1],$$

where $T_0, T_1 \in \mathcal{T}$. By Lemma 4.1, we have that $\mathbb{H}$ sends the set of these triangles to a set of projective presentations (4-2) which has Property (S). It remains to show that for any $X \in \mathscr{X}$ and $X' \in \mathcal{T} \cap \mathscr{X}[-1]$, we have $\mathbb{H}(X)(X') = 0$. Indeed, since $\mathscr{X}$ is a $\mathcal{T}[1]$-rigid subcategory, we have $\mathbb{H}(X)(X') = \mathrm{Hom}_{\mathscr{C}}(X', X) = 0$.

$$\begin{tikzcd} X' \in \mathscr{X}[-1] \arrow[r] \arrow[dr, equal] & X \in \mathscr{X} \\ & X' \in \mathcal{T} \arrow[u] \end{tikzcd}$$

This shows that $(\mathbb{H}(\mathscr{X}), \mathcal{T} \cap \mathscr{X}[-1])$ is a τ-rigid pair of $\mathsf{mod}\,\mathcal{T}$.

Step 2: The map Φ is surjective.

Let $(\mathcal{M}, \mathcal{E})$ be a τ-rigid pair of $\mathsf{mod}\,\mathcal{T}$. For each $M \in \mathcal{M}$, take a projective presentation

$$P_1 \xrightarrow{\pi^M} P_0 \to M \to 0 \tag{4-4}$$

such that the class $\{\pi^M \mid M \in \mathcal{M}\}$ has Property (S). By Theorem 2.8(ii), there is a unique morphism $f_M : T_1 \to T_0$ in $\mathcal{T}$ such that $\mathbb{H}(f_M) = \pi^M$. Moreover, $\mathbb{H}(\mathrm{cone}(f_M)) \cong M$. Since (4-4) has Property (S), it follows from Lemma 4.1 that the category

$$\mathscr{X}_1 := \{\mathrm{cone}(f_M) \mid M \in \mathcal{M}\}$$

is a $\mathcal{T}[1]$-rigid subcategory.

Let $\mathscr{X} := \mathscr{X}_1 \vee \mathcal{E}[1]$. Now we show that $\mathscr{X}$ is a $T[1]$-rigid subcategory of $\mathscr{C}$. Let $E \in \mathcal{E} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is cluster-tilting, we have

$$[\mathcal{T}[1]](\mathrm{cone}(f_M) \oplus E[1], E[2]) = 0.$$

Applying the functor $\mathrm{Hom}_{\mathscr{C}}(E, -)$ to the triangle $T_1 \xrightarrow{f_M} T_0 \to \mathrm{cone}(f_M) \to T_1[1]$, we have an exact sequence

$$\mathrm{Hom}_{\mathscr{C}}(E, T_1) \xrightarrow{f_M \circ} \mathrm{Hom}_{\mathscr{C}}(E, T_0) \to \mathrm{Hom}_{\mathscr{C}}(E, \mathrm{cone}(f_M)) \to 0,$$

which is isomorphic to

$$P_1(E) \xrightarrow{\pi^M} P_0(E) \to M(E) \to 0.$$

The condition $\mathcal{M}(\mathcal{E}) = 0$ implies that $\mathrm{Hom}_{\mathscr{C}}(E, \mathrm{cone}(f_M)) = 0$ and therefore

$$[\mathcal{T}[1]](E[1], \mathrm{cone}(f_M)[1]) = 0.$$

Thus the assertion follows.

Now we show that $\Phi(\mathscr{X}) = (\mathcal{M}, \mathcal{E})$.

It is straightforward to check that $\mathcal{T} \cap \mathscr{X}_1[-1] = 0$. For any object $X \in \mathcal{T} \cap \mathscr{X}[-1]$, we are able to write $X = X_1[-1] \oplus E \in \mathcal{T}$, where $X_1 \in \mathscr{X}_1$ and $E \in \mathcal{E}$. Since $X_1[-1] \in \mathcal{T} \cap \mathscr{X}_1[-1] = 0$, we have $X = E \in \mathcal{E}$. Thus we have $\mathcal{T} \cap \mathscr{X}[-1] \subseteq \mathcal{E}$. By the definition of τ-rigid pair, we have $\mathcal{E} \subseteq \mathcal{T}$. Noting that $\mathcal{E} \subseteq \mathscr{X}_1[-1] \vee \mathcal{E} = \mathscr{X}[-1]$, it follows that $\mathcal{E} \subseteq \mathcal{T} \cap \mathscr{X}[-1]$. Hence $\mathcal{T} \cap \mathscr{X}[-1] = \mathcal{E}$. It remains to show that $\mathbb{H}(\mathscr{X}) = \mathcal{M}$. Indeed, since $\mathcal{E} \subseteq \mathcal{T}$, we have

$$\mathbb{H}(\mathscr{X}) = \mathrm{Hom}_{\mathscr{C}}(\mathcal{T}, \mathscr{X}) = \mathrm{Hom}_{\mathscr{C}}(\mathcal{T}, \mathscr{X}_1) = \mathbb{H}(\mathscr{X}_1) = \mathcal{M}.$$

Step 3: The map Φ is injective.

Let $\mathscr{X}$ and $\mathscr{X}'$ be two $\mathcal{T}[1]$-rigid subcategories of $\mathscr{C}$ such that $\Phi(\mathscr{X}) = \Phi(\mathscr{X}')$. Let $\mathscr{X}_1$ and $\mathscr{X}_1'$ be respectively the full subcategories of $\mathscr{X}$ and $\mathscr{X}'$ consisting of objects without direct summands in $\mathcal{T}[1]$. Then $\mathscr{X} = \mathscr{X}_1 \vee (\mathscr{X} \cap \mathcal{T}[1])$ and $\mathscr{X}' = \mathscr{X}_1' \vee (\mathscr{X}' \cap \mathcal{T}[1])$. Since $\Phi(\mathscr{X}) = \Phi(\mathscr{X}')$, it follows that $\mathbb{H}(\mathscr{X}_1) = \mathbb{H}(\mathscr{X}_1')$ and $\mathscr{X} \cap \mathcal{T}[1] = \mathscr{X}' \cap \mathcal{T}[1]$.

For any object $X_1 \in \mathscr{X}_1$, there exists $X_1' \in \mathscr{X}_1'$ such that $\mathbb{H}(X_1) = \mathbb{H}(X_1')$. By Theorem 2.8(iii), there exists an isomorphism $X_1 \oplus Y[1] \simeq X_1' \oplus Z[1]$ for some $Y, Z \in \mathcal{T}$. Since $\mathscr{C}$ is Krull–Remak–Schmidt, we have $X_1 \simeq X_1'$. This implies that $\mathscr{X}_1 \subseteq \mathscr{X}_1'$. Similarly, we obtain $\mathscr{X}_1' \subseteq \mathscr{X}_1$ and then $\mathscr{X}_1 \simeq \mathscr{X}_1'$. Therefore $\mathscr{X} = \mathscr{X}'$. This shows that Φ is injective. $\square$

Our main result in this subsection is the following:

Theorem 4.3. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. The functor $\mathbb{H}: \mathscr{C} \to \mathsf{Mod}\,\mathcal{T}$ induces a bijection between the sets of weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ and of support τ-tilting pairs of $\mathsf{mod}\,\mathcal{T}$, given by*

$$\Phi : \mathscr{X} \longmapsto (\mathbb{H}(\mathscr{X}), \mathcal{T} \cap \mathscr{X}[-1]).$$

Proof. **Step 1**: The map Φ has values in support τ-tilting pairs of $\mathsf{mod}\,\mathcal{T}$.

Assume $\mathscr{X}$ is a weak $\mathcal{T}[1]$-cluster tilting subcategory of $\mathscr{C}$. By Proposition 4.2, we get that $\Phi(\mathscr{X})$ is a τ-rigid pair of $\mathsf{mod}\,\mathcal{T}$. Therefore $\mathcal{T} \cap \mathscr{X}[-1] \subseteq \mathrm{Ker}\ \mathbb{H}(\mathscr{X})$.

Let $T \in \mathcal{T}$ be an object of $\mathrm{Ker}\ \mathbb{H}(\mathscr{X})$, that is, $\mathrm{Hom}_{\mathscr{C}}(T, X) = 0$ for each $X \in \mathscr{X}$. This implies $[\mathcal{T}[1]](X \oplus T[1], \mathscr{X}[1]) = 0$. Note $[\mathcal{T}[1]](\mathscr{X}, (X \oplus T[1])[1]) = 0$. Since $\mathscr{X}$ is a $\mathcal{T}[1]$-cluster tilting subcategory, we have $X \oplus T[1] \in \mathscr{X}$ and thus

$T \subseteq \mathscr{X}[-1]$. Therefore $T \in \mathcal{T} \cap \mathscr{X}[-1]$. This shows that Ker $\mathbb{H}(\mathscr{X}) \subseteq \mathcal{T} \cap \mathscr{X}[-1]$. Hence

$$\text{Ker } \mathbb{H}(\mathscr{X}) = \mathcal{T} \cap \mathscr{X}[-1].$$

By the definition of weak $\mathcal{T}[1]$-cluster tilting subcategories, for any $T \in \mathcal{T}$, there exists a triangle

$$T \xrightarrow{f} X_1 \xrightarrow{g} X_2 \xrightarrow{h} T[1],$$

where $X_1, X_2 \in \mathscr{X}$. Applying the functor $\mathbb{H}$ to the above triangle, we obtain an exact sequence

$$\mathbb{H}(T) \xrightarrow{\mathbb{H}(f)} \mathbb{H}(X_1) \to \mathbb{H}(X_2) \to 0.$$

For any morphism $a : T \to X$, where $X \in \mathscr{X}$, since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $ah[-1] = 0$. So there exists a morphism $b : X_1 \to X$ such that $a = bf$. This shows that $\mathrm{Hom}_{\mathscr{C}}(f, X)$ is surjective. Thus there exists the commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}_{\mathscr{C}}(X_1, X) \arrow[r, "\mathrm{Hom}_{\mathscr{C}}(f,\ X)"] \arrow[d] & \mathrm{Hom}_{\mathscr{C}}(T, X) \arrow[r] \arrow[d] & 0 \\
\mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(X_1), \mathbb{H}(X)) \arrow[r, "\circ\mathbb{H}(f)"] & \mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(T), \mathbb{H}(X)). &
\end{tikzcd}$$

Using Theorem 2.8(ii), the right vertical map is an isomorphism. It follows that $\circ\mathbb{H}(f)$ is surjective, that is, $\mathbb{H}(f)$ is a left $\mathbb{H}(\mathscr{X})$-approximation. Altogether, we have shown that $\Phi(\mathscr{X})$ is a support τ-tilting pair of $\mathsf{mod}\ \mathcal{T}$.

Step 2: The map Φ is surjective.

Let $(\mathcal{M}, \mathcal{E})$ be a support τ-tilting pair of $\mathsf{mod}\ \mathcal{T}$ and let $\mathscr{X}$ be the preimage of $(\mathcal{M}, \mathcal{E})$ under Φ constructed in Proposition 4.2. Since $\mathbb{H}(\mathscr{X}) = \mathcal{M}$ is a support τ-tilting subcategory, for each $T \in \mathcal{T}$, there is an exact sequence

$$\mathbb{H}(T) \xrightarrow{\alpha} \mathbb{H}(X_3) \to \mathbb{H}(X_4) \to 0,$$

such that $X_3, X_4 \in \mathscr{X}$ and α is a left $\mathbb{H}(\mathscr{X})$-approximation. By Yoneda's lemma, there exists a unique morphism $\beta : T \to X_3$ such that $\mathbb{H}(\beta) = \alpha$. We complete this to a triangle

$$T \xrightarrow{\beta} X_3 \xrightarrow{\gamma} Y_T \xrightarrow{\delta} T[1]. \tag{4-5}$$

Let $\widetilde{\mathscr{X}} := \mathscr{X} \vee \mathsf{add}\{\, Y_T \mid T \in \mathcal{T} \,\}$ be the additive closure of $\mathscr{X}$ and $\{\, Y_T \mid T \in \mathcal{T} \,\}$. We claim $\widetilde{\mathscr{X}}$ is a weak $\mathcal{T}[1]$-cluster tilting subcategory of $\mathscr{C}$ such that $\Phi(\widetilde{\mathscr{X}}) = (\mathcal{M}, \mathcal{E})$.

It is clear that $\mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}$. It remains to show that

$$\widetilde{\mathscr{X}} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \widetilde{\mathscr{X}}[1]) = 0 = [\mathcal{T}[1]](\widetilde{\mathscr{X}}, M[1])\}.$$

Applying the functor $\mathbb{H}$ to the triangle (4-5), we see that $\mathbb{H}(Y_T)$ and $\mathbb{H}(X_4)$ are isomorphic in $\mathsf{mod}\ \mathcal{T}$. For any object $X \in \mathscr{X}$, consider the commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}_{\mathscr{C}}(X_3, X) \arrow[r, "\mathrm{Hom}_{\mathscr{C}}(\beta,\ X)"] \arrow[d, "\mathbb{H}(-)"'] & \mathrm{Hom}_{\mathscr{C}}(T, X) \arrow[d, "\simeq"] \\
\mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(X_3), \mathbb{H}(X)) \arrow[r, "\circ\alpha"] & \mathrm{Hom}_{\mathsf{mod}\ \mathcal{T}}(\mathbb{H}(T), \mathbb{H}(X)).
\end{tikzcd}$$

By Theorem 2.8, the map $\mathbb{H}(-)$ is surjective and the right vertical map is an isomorphism. Because α is a left $\mathbb{H}(\mathscr{X})$-approximation, $\circ\alpha$ is also surjective. Therefore $\mathrm{Hom}_{\mathscr{C}}(\beta, X)$ is surjective too.

For any morphism $a \in [\mathcal{T}[1]](Y_T, X[1])$, since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $a\gamma = 0$. So there exists a morphism $b : T[1] \to X[1]$ such that $a = b\delta$.

$$\begin{tikzcd}
T \arrow[r, "\beta"] & X_3 \arrow[r, "\gamma"] & Y_T \arrow[r, "\delta"] \arrow[d, "a"] & T[1] \arrow[dl, dotted, "b"] \\
 & & X[1] &
\end{tikzcd}$$

Since $\mathrm{Hom}_{\mathscr{C}}(\beta,\ X)$ is surjective, there exists a morphism $c : X_3 \to X$ such that $c\beta = b[-1]$ and thus $b = c[1] \circ \beta[1]$. It follows that $a = b\delta = c[1] \circ (\beta[1]\delta) = 0$. This shows that

$$[\mathcal{T}[1]](Y_T, \mathscr{X}[1]) = 0. \tag{4-6}$$

For any morphism $x \in [\mathcal{T}](X[-1], Y_T)$, we know that there exist two morphisms $x_1 : X[-1] \to T_1$ and $x_2 : T_1 \to Y_T$ such that $x = x_2x_1$, where $T_1 \in \mathcal{T}$. Since $\mathcal{T}$ is cluster tilting, we have $\delta x_2 = 0$. So there exists a morphism $y : T_1 \to X_3$ such that $x_2 = \gamma y$.

$$\begin{tikzcd}
 & & X[-1] \arrow[d, "x_1"] & \\
 & & T_1 \arrow[d, "x_2"] \arrow[dl, dotted, "y"'] & \\
T \arrow[r, "\beta"] & X_3 \arrow[r, "\gamma"] & Y_T \arrow[r, "\delta"] & T[1]
\end{tikzcd}$$

Since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $x = x_2x_1 = \gamma(yx_1) = 0$. This shows that

$$[\mathcal{T}[1]](\mathscr{X}, Y_T[1]) = 0. \tag{4-7}$$

For any $T' \in \mathcal{T}$ and morphism $u \in [\mathcal{T}](Y_{T'}[-1], Y_T)$, we know that there exist two morphisms $u_1 : Y_{T'}[-1] \to T_2$ and $u_2 : T_2 \to Y_T$ such that $u = u_2u_1$, where $T_2 \in \mathcal{T}$. Since $\mathcal{T}$ is cluster tilting, we have $\delta u_2 = 0$. So there exists a morphism $v : T_2 \to X_3$ such that $u_2 = \gamma v$.

$$\begin{tikzcd}
 & & Y_{T'}[-1] \arrow[d, "u_1"] & \\
 & & T_2 \arrow[d, "u_2"] \arrow[dl, dotted, "v"'] & \\
T \arrow[r, "\beta"] & X_3 \arrow[r, "\gamma"] & Y_T \arrow[r, "\delta"] & T[1]
\end{tikzcd}$$

Since $[\mathcal{T}[1]](Y_T, \mathscr{X}[1]) = 0$, we have $vu_1 = 0$. It follows that $u = u_2u_1 = \gamma v u_1 = 0$. This shows that

$$[\mathcal{T}[1]](Y_{T'}, Y_T[1]) = 0. \tag{4-8}$$

Using equalities (4-6), (4-7) and (4-8), we deduce that $\widetilde{\mathscr{X}}$ is a $\mathcal{T}[1]$-rigid subcategory.

Now we show that $\{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \widetilde{\mathscr{X}}[1]) = 0 = [\mathcal{T}[1]](\widetilde{\mathscr{X}}, M[1])\} \subseteq \widetilde{\mathscr{X}}$.

For any object $M \in \mathscr{C}$, assume that $[\mathcal{T}[1]](M, \widetilde{\mathscr{X}}[1]) = 0 = [\mathcal{T}[1]](\widetilde{\mathscr{X}}, M[1])$. Since $\mathcal{T}$ is a cluster-tilting subcategory, there exists a triangle

$$T_5 \xrightarrow{f} T_6 \xrightarrow{g} M \xrightarrow{h} T_5[1],$$

where $T_5, T_6 \in \mathcal{T}$. By the above discussion, for an object $T_6 \in \mathcal{T}$, there exists a triangle

$$T_6 \xrightarrow{u} X_6 \xrightarrow{v} Y_{T_6} \xrightarrow{w} T_6[1],$$

where $X_6 \in \mathscr{X}$, $Y_{T_6} \in \widetilde{\mathscr{X}}$ and u is a left $\mathscr{X}$-approximation of T_6. For an object $T_5 \in \mathcal{T}$, there exists a triangle

$$T_5 \xrightarrow{u'} X_5 \xrightarrow{v'} Y_{T_5} \xrightarrow{w'} T_5[1],$$

where $X_5 \in \mathscr{X}$, $Y_{T_5} \in \widetilde{\mathscr{X}}$ and u' is a left $\mathscr{X}$-approximation of T_5. By the octahedral axiom, we have a commutative diagram

$$\begin{tikzcd}
T_5 \arrow[r, "f"] \arrow[d, equal] & T_6 \arrow[r, "g"] \arrow[d, "u"] & M \arrow[r, "h"] \arrow[d, "a"] & T_5[1] \arrow[d, equal] \\
T_5 \arrow[r, "x=uf"] & X_6 \arrow[r, "y"] \arrow[d, "v"] & N \arrow[r, "z"] \arrow[d, "b"] & T_5[1] \\
 & Y_{T_6} \arrow[r, equal] \arrow[d, "w"] & Y_{T_6} \arrow[d, "c"] & \\
 & T_6[1] \arrow[r, "{g[1]}"] & M[1] &
\end{tikzcd}$$

of triangles in $\mathscr{C}$. We claim that x is a left $\mathscr{X}$-approximation of T_5. Indeed, for any $d : T_5 \to X$, since $dh[-1] \in [\mathcal{T}](M[-1], \widetilde{\mathscr{X}}) = 0$, there exists a morphism

$e : T_6 \to X$ such that $d = ef$, where $X \in \mathscr{X}$.

$$\begin{tikzcd} M[-1] \arrow[r, "h[-1]"] & T_5 \arrow[r, "f"] \arrow[d, "d"] & T_6 \arrow[r, "g"] \arrow[dl, dotted, "e"] & M \arrow[r, "h"] & T_5[1] \\ & X & & & \end{tikzcd}$$

Since u is a left $\mathscr{X}$-approximation of T_6, there exists a morphism $k : X_6 \to X$ such that $ku = e$. It follows that $d = ef = kuf = kx$, as required.

Since x is a left $\mathscr{X}$-approximation of T_5, by Lemma 1.4.3 in [Neeman 2001], we have the commutative diagram

$$\begin{tikzcd} T_5 \arrow[r, "x"] \arrow[d, equal] & X_6 \arrow[r, "y"] \arrow[d, "\lambda"] & N \arrow[r, "z"] \arrow[d, dotted, "\varphi"] & T_5[1] \arrow[d, equal] \\ T_5 \arrow[r, "u'"] & X_5 \arrow[r, "v'"] & Y_{T_5} \arrow[r, "w'"] & T_5[1], \end{tikzcd}$$

where the middle square is homotopy cartesian and the differential $\partial = x[1] \circ w'$, that is, there exists a triangle

$$X_6 \xrightarrow{\binom{-y}{\lambda}} N \oplus X_5 \xrightarrow{(\varphi, v')} Y_{T_5} \xrightarrow{\partial} X_6[1].$$

Note that $\partial \in [\mathcal{T}[1]](\widetilde{\mathscr{X}}, \widetilde{\mathscr{X}}[1]) = 0$. Thus we have $N \oplus X_5 \simeq X_6 \oplus Y_{T_5} \in \widetilde{\mathscr{X}}$, which implies $N \in \widetilde{\mathscr{X}}$. Since $c = g[1]w \in [\mathcal{T}[1]](\widetilde{\mathscr{X}}, M[1]) = 0$, we deduce that the triangle

$$M \xrightarrow{a} N \xrightarrow{b} Y_{T_6} \xrightarrow{c} M[1]$$

splits. Hence M is a direct summand of N and thus $M \in \widetilde{\mathscr{X}}$.

This shows that

$$\widetilde{\mathscr{X}} = \{M \in \mathscr{C} \mid [\mathcal{T}[1]](M, \widetilde{\mathscr{X}}[1]) = 0 = [\mathcal{T}[1]](\widetilde{\mathscr{X}}, M[1])\}.$$

For any object $T \in \mathcal{T}$, $\mathbb{H}(Y_T) \simeq \mathbb{H}(X_4)$. Therefore

$$\mathbb{H}(\widetilde{\mathscr{X}}) \simeq \mathbb{H}(\mathscr{X}) \simeq \mathcal{M}.$$

Since $\mathcal{T} \cap \widetilde{\mathscr{X}}[-1] \supseteq \mathcal{T} \cap \mathscr{X}[-1] = \mathcal{E}$ and $\mathcal{T} \cap \widetilde{\mathscr{X}}[-1] \subseteq \operatorname{Ker} \mathbb{H}(\mathscr{X}) = \mathcal{E}$, we have

$$\mathcal{T} \cap \widetilde{\mathscr{X}}[-1] = \mathcal{E}.$$

This shows that Φ is surjective.

Step 3: The map Φ is injective.

This follows from Step 3 in Proposition 4.2. $\square$

For any support τ-tilting subcategory $\mathscr{Y}$ of $\operatorname{mod}\mathcal{T}$, by Theorem 4.3 there exists a unique weak $\mathcal{T}[1]$-cluster tilting subcategory $\mathscr{X}$ of $\mathscr{C}$ such that $\mathbb{H}(\mathscr{X}) = \mathscr{Y}$. Throughout this paper, we denote the preimage $\mathscr{X}$ by $\mathbb{H}^{-1}(\mathscr{Y})$ for simplicity. Consequently, we have the following result:

Theorem 4.4. *The bijection in Theorem 4.3 induces a bijection from the first of the following sets to the second*:

(I) $\mathcal{T}[1]$*-cluster tilting subcategories of* $\mathscr{C}$.

(II) *Support* τ*-tilting subcategories* $\mathscr{Y}$ *of* $\operatorname{mod}\mathcal{T}$ *such that* $\mathbb{H}^{-1}(\mathscr{Y})$ *is contravariantly finite in* $\mathscr{C}$.

Moreover, if $\mathscr{C}$ *admits a Serre functor* $\mathbb{S}$, *we get a bijection from the first to the second of the following sets.*

(1) *Cluster tilting subcategories of* $\mathscr{C}$.

(2) *Support* τ*-tilting subcategories* $\mathscr{Y}$ *of* $\operatorname{mod}\mathcal{T}$ *such that* $\mathbb{H}^{-1}(\mathscr{Y})$ *is contravariantly finite and* F*-stable in* $\mathscr{C}$.

Proof. The first bijection follows from the fact that $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ are precisely contravariantly finite weak $\mathcal{T}[1]$-cluster tilting subcategories, and the second bijection follows from Theorem 3.6. □

***τ*-tilting subcategories and tilting subcategories.** $\mathscr{C}$ and $\mathcal{T}$ are the same as above. By definition we know that the category $\operatorname{mod}\mathcal{T}$ is abelian and has enough projectives. Thus we can investigate the projective dimension of an object M in $\operatorname{mod}\mathcal{T}$, which we denote by $\operatorname{pd} M$. For a subcategory $\mathscr{D}$ of $\operatorname{mod}\mathcal{T}$, we say that the projective dimension of $\mathscr{D}$ is at most n, denoted by $\operatorname{pd}\mathscr{D} \leq n$, if $\operatorname{pd} M \leq n$ for any object $M \in \mathscr{D}$.

Let $X \in \mathscr{C}$, $\mathcal{I}_X(\mathcal{T}[1])$ be the ideal of $\mathcal{T}[1]$ formed by the morphisms between objects in $\mathcal{T}[1]$ factoring through the object X. For a subcategory $\mathscr{D}$ of $\mathscr{C}$, we define the *factorization ideal* of $\mathscr{D}$, denoted by $\mathcal{I}_{\mathscr{D}}(\mathcal{T}[1])$, as follows

$$\mathcal{I}_{\mathscr{D}}(\mathcal{T}[1]) := \{\mathcal{I}_X(\mathcal{T}[1]) \mid X \in \mathscr{D}\}.$$

Theorem 2.8 indicates that $\operatorname{mod}\mathcal{T}$ is Gorenstein of dimension at most one. Thus all objects in $\operatorname{mod}\mathcal{T}$ have projective dimension zero, one or infinity. The following result characterizes the objects in $\operatorname{mod}\mathcal{T}$ having finite projective dimension.

Theorem 4.5 [Beaudet et al. 2014; Lasnier 2011]. *Let* $\mathscr{C}$ *be a triangulated category with a cluster tilting subcategory* $\mathcal{T}$, *and* X *be an object in* $\mathscr{C}$ *having no direct summands in* $\mathcal{T}[1]$. *Then*

$$\operatorname{pd}\mathbb{H}(X) \leq 1 \text{ if and only if } \mathcal{I}_X(\mathcal{T}[1]) = 0.$$

In this subsection, we introduce two important classes of subcategories of $\mathsf{mod}\,\mathcal{T}$ and give a connection with ghost cluster tilting subcategories and cluster tilting subcategories of $\mathscr{C}$. We start with the following definition.

Definition 4.6. Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$.

(i) A subcategory $\mathcal{M}$ of $\mathsf{mod}\,\mathcal{T}$ is said to be *τ-tilting* if $(\mathcal{M}, 0)$ is a support τ-tilting pair of $\mathsf{mod}\,\mathcal{T}$.

(ii) [Beligiannis 2013] A subcategory $\mathcal{M}$ of $\mathsf{mod}\,\mathcal{T}$ is said to be *weak tilting* if the following three conditions are satisfied:

(T1) $\mathrm{Ext}^1_{\mathsf{mod}\,\mathcal{T}}(\mathcal{M}, \mathcal{M}) = 0$.

(T2) $\mathsf{pd}\,M \leq 1$, for any $M \in \mathcal{M}$.

(T3) For any projective object P in $\mathsf{mod}\,\mathcal{T}$, there exists a short exact sequence

$$0 \to P \to M_0 \to M_1 \to 0,$$

where $M_0, M_1 \in \mathcal{M}$.

A weak tilting subcategory $\mathcal{M}$ is called a *tilting subcategory* if it also satisfies the following additional condition:

(T4) $\mathcal{M}$ is contravariantly finite in $\mathsf{mod}\,\mathcal{T}$.

Remark 4.7. Beligiannis [2010; 2013] indicates that a contravariantly finite subcategory $\mathcal{M}$ of $\mathsf{mod}\,\mathcal{T}$ is a tilting subcategory if and only if

$$\mathsf{Fac}(\mathcal{M}) = \{X \in \mathsf{mod}\,\mathcal{T} \mid \mathrm{Ext}^1_{\mathsf{mod}\,\mathcal{T}}(\mathcal{M}, X) = 0\},$$

where $\mathsf{Fac}(\mathcal{M})$ is the full subcategory of $\mathsf{mod}\,\mathcal{T}$ consisting of all factors of objects of $\mathcal{M}$.

Immediately, we have the following result:

Theorem 4.8. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. The functor $\mathbb{H} : \mathscr{C} \to \mathsf{Mod}\ \mathcal{T}$ induces a bijection*

$$\Phi : \mathscr{X} \longmapsto \mathbb{H}(\mathscr{X})$$

from the first of the following sets to the second:

(I) *Weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$.*

(II) *τ-tilting subcategories of* $\mathsf{mod}\ \mathcal{T}$.

It restricts to a bijection from the first to the second of the following sets.

(I) *$\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$.*

(II) *τ-tilting subcategories $\mathscr{Y}$ of* $\mathsf{mod}\ \mathcal{T}$ *such that $\mathbb{H}^{-1}(\mathscr{Y})$ is contravariantly finite in $\mathscr{C}$.*

Moreover, if $\mathscr{C}$ admits a Serre functor $\mathbb{S}$, we get a bijection between the following sets.

(1) *Cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$.*

(2) *τ-tilting subcategories $\mathscr{Y}$ of $\mathsf{mod}\,\mathcal{T}$ such that $\mathbb{H}^{-1}(\mathscr{Y})$ is contravariantly finite and F-stable in $\mathscr{C}$.*

Proof. Note that objects in $\mathscr{X}$ do not have nonzero direct summands in $\mathcal{T}[1]$ if and only if $\mathcal{T} \cap \mathscr{X}[-1] = 0$. This assertion follows from Theorems 4.3, 4.4 directly. $\square$

Now we give a close relationship between τ-tilting subcategories and weak tilting subcategories.

Lemma 4.9. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. Then any weak tilting subcategory of $\mathsf{mod}\,\mathcal{T}$ is a τ-tilting subcategory.*

Proof. Let $\mathcal{M}$ be a weak tilting subcategory of $\mathsf{mod}\,\mathcal{T}$.

(1) We first show that $(\mathcal{M}, 0)$ is a τ-rigid pair of $\mathsf{mod}\,\mathcal{T}$. For any object $M \in \mathcal{M}$, since $\mathsf{pd}\,M \leq 1$, we get a short exact sequence

$$0 \to P_1 \xrightarrow{\pi^M} P_0 \to M \to 0.$$

Note that $P_1 = 0$ if $\mathsf{pd}\,M = 0$. Applying the functor $\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(-, \mathcal{M})$ to it, we get an exact sequence

$$\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(P_0, \mathcal{M}) \xrightarrow{\circ\pi^M} \mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(P_1, \mathcal{M}) \to \mathrm{Ext}^1_{\mathsf{mod}\,\mathcal{T}}(M, \mathcal{M}) = 0.$$

This means there is a class of projective presentations $\{P_1 \xrightarrow{\pi^M} P_0 \to M \to 0 \mid M \in \mathcal{M}\}$ which has Property (S). Therefore $(\mathcal{M}, 0)$ is a τ-rigid pair of $\mathsf{mod}\,\mathcal{T}$ because $\mathcal{M}(0) = 0$.

(2) We show that $(\mathcal{M}, 0)$ is a support τ-tilting pair of $\mathsf{mod}\,\mathcal{T}$. For each object $T \in \mathcal{T}$, $\mathcal{T}(-, T)$ is a projective object in $\mathsf{mod}\,\mathcal{T}$. Since $\mathcal{M}$ is weak tilting in $\mathsf{mod}\,\mathcal{T}$, there exists a short exact sequence

$$0 \to \mathcal{T}(-, T) \xrightarrow{f} M_0, \to M_1 \to 0$$

where $M_0, M_1 \in \mathcal{M}$. Applying the functor $\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(-, \mathcal{M})$ to the above exact sequence, we have the following exact sequence:

$$\mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(M_0, \mathcal{M}) \xrightarrow{\circ f} \mathrm{Hom}_{\mathsf{mod}\,\mathcal{T}}(\mathcal{T}(-, T), \mathcal{M}) \to \mathrm{Ext}^1_{\mathsf{mod}\,\mathcal{T}}(M_1, \mathcal{M}) = 0.$$

This shows that f is a left $\mathcal{M}$-approximation.

If $\mathcal{M}(E) = 0$, where $E \in \mathcal{T}$, by the above discussion, there exists an exact sequence

$$0 \to \mathcal{T}(-, E) \to M_0 \to M_1 \to 0$$

with $M^0, M^1 \in \mathcal{M}$. It follows that there exists an exact sequence

$$0 \to \mathcal{T}(E, E) \to M_0(E) \to M_1(E) \to 0$$

Since $M_0(E) = 0$, we have $\mathcal{T}(E, E) = 0$ and thus $E = 0$. Therefore $\operatorname{Ker}(\mathcal{M}) = 0$. This shows that $(\mathcal{M}, 0)$ is a support τ-tilting pair of $\operatorname{mod}\mathcal{T}$. □

The following result gives a criterion for a τ-tilting subcategory of $\operatorname{mod}\mathcal{T}$ to be a weak tilting subcategory.

Theorem 4.10. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. A τ-tilting subcategory of $\operatorname{mod}\mathcal{T}$ is a weak tilting subcategory if and only if its projective dimension is at most one.*

Proof. Let $\mathcal{M}$ be a τ-tilting subcategory of $\operatorname{mod}\mathcal{T}$ and $\operatorname{pd}\mathcal{M} \le 1$. By Theorem 4.8, there exists a weak $\mathcal{T}[1]$-tilting subcategory $\mathscr{X}$ of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$ such that $\mathbb{H}(\mathscr{X}) = \mathcal{M}$.

Step 1: We show $\operatorname{Ext}^1_{\operatorname{mod}\mathcal{T}}(\mathcal{M}, \mathcal{M}) = 0$. Namely, $\operatorname{Ext}^1_{\operatorname{mod}\mathcal{T}}(\mathbb{H}(\mathscr{X}), \mathbb{H}(\mathscr{X})) = 0$.

For any object $X_1 \in \mathscr{X}$, since $\mathcal{T}$ is cluster tilting, there exists a triangle

$$T_0 \xrightarrow{f} T_1 \xrightarrow{g} X_1 \xrightarrow{h} T_0[1], \tag{4-9}$$

where g is a minimal right $\mathcal{T}$-approximation of X_1 and $T_0, T_1 \in \mathcal{T}$. Since $\mathbb{H}(X_1) \in \mathcal{M}$, we have $\operatorname{pd}\mathbb{H}(X_1) \le 1$. Applying the functor $\mathbb{H}$ to the above triangle, we have a minimal projective presentation

$$0 \to H(T_0) \xrightarrow{f\circ} \mathbb{H}(T_1) \xrightarrow{g\circ} \mathbb{H}(X_1) \to 0$$

of $\mathbb{H}(X_1)$, since X_1 has no nonzero direct summands in $\mathcal{T}[1]$ and $\operatorname{pd}\mathbb{H}(X_1) \le 1$. Applying the functor $\operatorname{Hom}_{\operatorname{mod}\mathcal{T}}(-, \mathbb{H}(X_2))$, where $X_2 \in \mathscr{X}$, to the above exact sequence, we get an exact sequence:

$$\begin{aligned} &\operatorname{Hom}(\mathbb{H}(T_1), \mathbb{H}(X_2)) \to \operatorname{Hom}(\mathbb{H}(T_0), \mathbb{H}(X_2)) \\ &\qquad \to \operatorname{Ext}^1(\mathbb{H}(X_1), \mathbb{H}(X_2)) \to \operatorname{Ext}^1(\mathbb{H}(T_1), \mathbb{H}(X_2)) = 0, \end{aligned}$$

where the Hom and Ext groups are taken over $\operatorname{mod}\mathcal{T}$. The last item vanishes because $\mathbb{H}(T_1)$ is projective in $\operatorname{mod}\mathcal{T}$. Note that the first map is isomorphic to

$$\operatorname{Hom}_{\mathscr{C}}(T_1, X_2) \xrightarrow{\operatorname{Hom}_{\mathscr{C}}(f, X_2)} \operatorname{Hom}_{\mathscr{C}}(T_0, X_2).$$

It follows that $\operatorname{Ext}^1_{\operatorname{mod}\mathcal{T}}(\mathbb{H}(X_1), \mathbb{H}(X_2))$ is isomorphic to $\operatorname{Coker}\operatorname{Hom}_{\mathscr{C}}(f, X_2)$.

Applying the functor $\operatorname{Hom}_{\mathscr{C}}(-, X_2)$ to the triangle (4-9), we have the following exact sequence:

$$\operatorname{Hom}_{\mathscr{C}}(T_1, X_2) \xrightarrow{\operatorname{Hom}_{\mathscr{C}}(f, X_2)} \operatorname{Hom}_{\mathscr{C}}(T_0, X_2) \xrightarrow{\operatorname{Hom}_{\mathscr{C}}(h[-1], X_2)} \operatorname{Hom}_{\mathscr{C}}(X_1[-1], X_2).$$

In particular, we have the following exact sequence:

$$\mathrm{Hom}_{\mathscr{C}}(T_1, X_2) \xrightarrow{\mathrm{Hom}_{\mathscr{C}}(f, X_2)} \mathrm{Hom}_{\mathscr{C}}(T_0, X_2) \xrightarrow{\mathrm{Hom}_{\mathscr{C}}(h[-1], X_2)} \mathrm{Im}\ \mathrm{Hom}_{\mathscr{C}}(h[-1],\ X_2) \to 0.$$

We claim that $\mathrm{Im}\,\mathrm{Hom}_{\mathscr{C}}(h[-1],\ X_2) = [\mathcal{T}](X_1[-1], X_2)$. Indeed,

$$\mathrm{Im}\,\mathrm{Hom}_{\mathscr{C}}(h[-1],\ X_2) \subseteq [\mathcal{T}](X_1[-1], X_2)$$

is clear. For any morphism $x \in [\mathcal{T}](X_1[-1], X_2)$, we have two morphisms $x_1 : X_1[-1] \to T$ and $x_2 : T \to X_2$, where $T \in \mathcal{T}$ such that $x = x_2x_1$. Since $\mathrm{Hom}_{\mathscr{C}}(T_1[-1], T) = 0$, there exists a morphism $a : T_0 \to T$ such that $ah[-1] = x_1$.

$$\begin{tikzcd}
T_1[-1] \arrow[r, "{g[-1]}"] \arrow[dr, "0"'] & X_1[-1] \arrow[r, "{h[-1]}"] \arrow[d, "x_1"] & T_0 \arrow[r, "f"] \arrow[dl, dotted, "a"] & T_1 \arrow[r, "g"] & X_1 \arrow[r, "h"] & T_0[1] \\
& T & & & &
\end{tikzcd}$$

It follows that $x = x_2x_1 = (x_2a)h[-1] \in \mathrm{Im}\ \mathrm{Hom}_{\mathscr{C}}(h[-1],\ X_2)$, as required.

Since $\mathscr{X}$ is $\mathcal{T}[1]$-rigid, we have $[\mathcal{T}](X_1[-1], X_2) = 0$. Thus we obtain

$$\mathrm{Ext}^1_{\mathsf{mod}\,\mathcal{T}}(\mathbb{H}(X_1), \mathbb{H}(X_2)) \simeq \mathsf{Coker}\ \mathrm{Hom}_{\mathscr{C}}(f, X_2) = [\mathcal{T}](X_1[-1], X_2) = 0.$$

Step 2: We show that for any projective object P in $\mathsf{mod}\,\mathcal{T}$, there exists a short exact sequence

$$0 \to P \to M_0 \to M_1 \to 0,$$

where $M_0, M_1 \in \mathcal{M}$. We may assume $P = \mathcal{T}(-, T) = \mathbb{H}(T)$ in $\mathsf{mod}\,\mathcal{T}$, where $T \in \mathcal{T}$.

Since $T \in \mathcal{T} \subseteq \mathscr{X}[-1] * \mathscr{X}$, there exists a triangle

$$X_3[-1] \xrightarrow{u} T \xrightarrow{v} X_4 \xrightarrow{w} X_3,$$

where $X_3, X_4 \in \mathscr{X}$. Applying the functor $\mathbb{H}$ to the above triangle, we have the following exact sequence:

$$\mathbb{H}(X_3[-1]) \xrightarrow{\mathbb{H}(u)} \mathbb{H}(T) \to \mathbb{H}(X_4) \to \mathbb{H}(X_3) \to 0.$$

We claim that $\mathrm{Im}\,\mathbb{H}(u) = 0$. That is to say, for any morphism $y : T' \to X_3[-1]$, where $T' \in \mathcal{T}$, we have $uy = 0$. Indeed, since $\mathcal{T}$ is cluster tilting, there exists a triangle

$$T_2 \xrightarrow{\alpha} T_3 \xrightarrow{\beta} X_3 \xrightarrow{\gamma} T_2[1],$$

where β is a minimal right $\mathcal{T}$-approximation of X_3 and $T_2, T_3 \in \mathcal{T}$. Applying the functor $\mathbb{H}$ to the above triangle, we have a minimal projective presentation

$$\mathbb{H}(X_3[-1]) \xrightarrow{\mathbb{H}(\gamma[-1])} \mathbb{H}(T_2) \xrightarrow{\mathbb{H}(\alpha)} \mathbb{H}(T_3) \xrightarrow{\mathbb{H}(\beta)} \mathbb{H}(X_3) \to 0$$

of $\mathbb{H}(X_3)$, since X_1 has no nonzero direct summands in $\mathcal{T}[1]$. Since $\mathbb{H}(X_3) \in \mathcal{M}$, we have $\mathrm{pd}\,\mathbb{H}(X_3) \leq 1$. Thus we have Im $\mathbb{H}(\gamma[-1]) = 0$ and thus $\gamma[-1] \circ y = 0$. So there exists a morphism $b : T' \to T_3[-1]$ such that $y = \beta[-1] \circ b$.

$$\begin{tikzcd}
 & T' \arrow[dl, dotted, "b"'] \arrow[d, "y"] & & \\
T_3[-1] \arrow[r, "\beta[-1]"] & X_3[-1] \arrow[r, "\gamma[-1]"] & T_2 \arrow[r, "\alpha"] & T_3.
\end{tikzcd}$$

It follows that $uy = (u\beta[-1])b = 0 \circ b = 0$, as required. Hence we have the following exact sequence:

$$0 \to \mathbb{H}(T) \to \mathbb{H}(X_4) \to \mathbb{H}(X_3) \to 0,$$

where $\mathbb{H}(X_4), \mathbb{H}(X_3) \in \mathcal{M}$.

This shows that $\mathcal{M}$ is a weak tilting subcategory of $\mathrm{mod}\,\mathcal{T}$. Combining this with Lemma 4.9, the assertion follows. $\square$

Consequently, we have the following result:

Theorem 4.11. *Let $\mathscr{C}$ be a triangulated category with a cluster tilting subcategory $\mathcal{T}$. The functor $\mathbb{H} : \mathscr{C} \to \mathrm{Mod}\ \mathcal{T}$ induces a bijection*

$$\Phi : \mathscr{X} \longmapsto \mathbb{H}(\mathscr{X})$$

from the first of the following sets to the second:

(I) *Weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$ and whose factorization ideals vanish.*

(II) *Weak tilting subcategories of* $\mathrm{mod}\ \mathcal{T}$.

It restricts to a bijection from the first to the second of the following sets.

(1) *$\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$ and whose factorization ideals vanish.*

(2) *Tilting subcategories $\mathscr{Y}$ of* $\mathrm{mod}\ \mathcal{T}$ *such that $\mathbb{H}^{-1}(\mathscr{Y})$ is contravariantly finite in $\mathscr{C}$.*

Moreover, *if $\mathscr{C}$ admits a Serre functor $\mathbb{S}$, we get a bijection between the following sets.*

(1) *Cluster tilting subcategories of $\mathscr{C}$ whose objects do not have nonzero direct summands in $\mathcal{T}[1]$ and whose factorization ideals vanish.*

(2) *Tilting subcategories $\mathscr{Y}$ of* $\mathrm{mod}\ \mathcal{T}$ *such that $\mathbb{H}^{-1}(\mathscr{Y})$ is contravariantly finite and F-stable in $\mathscr{C}$.*

Proof. This follows from Theorems 4.5, 4.8 and 4.10 directly. $\square$

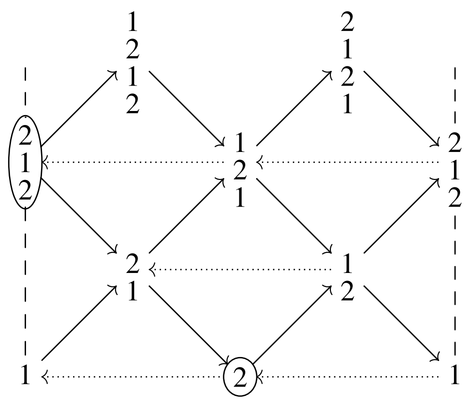
Remark 4.12. The above results generalize and improve several results in the literature. More precisely, Proposition 4.2, Theorems 4.3, 4.8 and 4.11 generalize a result in [Yang and Zhu 2019, Theorem 3.6], where analogous results were proved in the case where $\mathscr{C}$ is 2-Calabi–Yau and $\mathcal{T} = \operatorname{add} T$, see [Adachi et al. 2014, Theorem 4.1]. Theorem 4.11 generalizes a result of Beligiannis [2013, Theorem 6.6] in some cases, but we don't assume that $\operatorname{mod} \mathscr{X}$ has finite global dimension here.

We conclude this section with an example illustrating the bijections in Section 4:

Example 4.13. We revisit Example 3.4 presented in Section 3. Let $A = kQ/I$ be a self-injective algebra given by the quiver

$$Q:\ 1 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 2$$

and $I = \langle \alpha\beta\alpha\beta,\ \beta\alpha\beta\alpha \rangle$. The Auslander–Reiten quiver of $\operatorname{mod} A$ is



where the first and the last columns are identified. The stable module category

$$\mathscr{C} := \underline{\operatorname{mod}}\, A$$

is triangulated with a Serre functor. We get the Auslander–Reiten quiver of $\mathscr{C}$ by deleting the first row in above figure. By simple calculation, we obtain that

$$\mathcal{T} := \operatorname{add}\left(2 \oplus \begin{smallmatrix} 2\\1\\2 \end{smallmatrix}\right)$$

is a cluster tilting subcategory of $\mathscr{C}$. The Auslander–Reiten quiver of $\operatorname{mod} \mathcal{T}$ is

$$\begin{tikzcd}
\begin{smallmatrix} a\\b \end{smallmatrix} \arrow[dr] & & & \begin{smallmatrix} a\\b \end{smallmatrix} \\
& a \arrow[dr] & b \arrow[l, dotted] \arrow[ur] & \\
& & \begin{smallmatrix} b\\a \end{smallmatrix} \arrow[u] &
\end{tikzcd}$$

We illustrate the correspondences stated in this section as follows. In the table, weak $\mathcal{T}[1]$-cluster tilting subcategories of $\mathscr{C}$ and support τ-tilting pairs of $\mathsf{mod}\,\mathcal{T}$ are marked by ♣.

	$\mathcal{T}[1]$-rigid subcategories	τ-rigid pairs
	$\mathsf{add}\begin{pmatrix}2\\1\\2\end{pmatrix}$	$\left(\mathsf{add}\binom{a}{b}, 0\right)$
	$\mathsf{add}\begin{pmatrix}1\\2\\1\end{pmatrix}$	$\left(0, \mathsf{add}\binom{b}{a}\right)$
	$\mathsf{add}\binom{2}{1}$	$(a, 0)$
	$\mathsf{add}\binom{1}{2}$	$(b, 0)$
	$\mathsf{add}\,(2)$	$\left(\mathsf{add}\binom{b}{a}, 0\right)$
	$\mathsf{add}\,(1)$	$\left(0, \mathsf{add}\binom{a}{b}\right)$
♣	$\mathsf{add}\left(\begin{smallmatrix}2\\1\\2\end{smallmatrix} \oplus 2\right)$	$\left(\mathsf{add}\left(\begin{smallmatrix}a\\b\end{smallmatrix} \oplus \begin{smallmatrix}b\\a\end{smallmatrix}\right), 0\right)$
♣	$\mathsf{add}\left(\begin{smallmatrix}1\\2\end{smallmatrix} \oplus 2\right)$	$\left(\mathsf{add}\left(\begin{smallmatrix}b\\a\end{smallmatrix} \oplus b\right), 0\right)$
♣	$\mathsf{add}\left(\begin{smallmatrix}2\\1\\2\end{smallmatrix} \oplus \begin{smallmatrix}2\\1\end{smallmatrix}\right)$	$\left(\mathsf{add}\left(\begin{smallmatrix}a\\b\end{smallmatrix} \oplus a\right), 0\right)$
♣	$\mathsf{add}\left(\begin{smallmatrix}1\\2\end{smallmatrix} \oplus 1\right)$	$\left(b, \mathsf{add}\binom{a}{b}\right)$
♣	$\mathsf{add}\left(\begin{smallmatrix}2\\1\end{smallmatrix} \oplus \begin{smallmatrix}1\\2\\1\end{smallmatrix}\right)$	$\left(a, \mathsf{add}\binom{b}{a}\right)$
♣	$\mathsf{add}\left(1 \oplus \begin{smallmatrix}1\\2\\1\end{smallmatrix}\right)$	$\left(0, \mathsf{add}\left(\begin{smallmatrix}a\\b\end{smallmatrix} \oplus \begin{smallmatrix}b\\a\end{smallmatrix}\right)\right)$

Acknowledgements

Zhou would like to thank Huimin Chang for helpful discussions on the subject. The authors would like to thank the referee for reading the paper carefully and for many suggestions on mathematical and English expressions.

References

[Adachi et al. 2014] T. Adachi, O. Iyama, and I. Reiten, "τ-tilting theory", *Compos. Math.* **150**:3 (2014), 415–452. MR Zbl

[Auslander 1974] M. Auslander, "Representation theory of Artin algebras, I", *Comm. Algebra* **1** (1974), 177–268. MR Zbl

[Auslander and Reiten 1991] M. Auslander and I. Reiten, "Applications of contravariantly finite subcategories", *Adv. Math.* **86**:1 (1991), 111–152. MR Zbl

[Beaudet et al. 2014] L. Beaudet, T. Brüstle, and G. Todorov, "Projective dimension of modules over cluster-tilted algebras", *Algebr. Represent. Theory* **17**:6 (2014), 1797–1807. MR Zbl

[Beligiannis 2010] A. Beligiannis, "Tilting theory in abelian categories and related homotopical structures", preprint, Univ. Ioannina, 2010.

[Beligiannis 2013] A. Beligiannis, "Rigid objects, triangulated subfactors and abelian localizations", *Math. Z.* **274**:3-4 (2013), 841–883. MR Zbl

[Bondal and Kapranov 1989] A. I. Bondal and M. M. Kapranov, "Representable functors, Serre functors, and mutations", *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:6 (1989), 1183–1205. In Russian; translated in *Math. USSR-Izv.* **35**:3 (1990), 519–541. MR Zbl

[Buan et al. 2006] A. B. Buan, R. Marsh, M. Reineke, I. Reiten, and G. Todorov, "Tilting theory and cluster combinatorics", *Adv. Math.* **204**:2 (2006), 572–618. MR Zbl

[Buan et al. 2009] A. B. Buan, O. Iyama, I. Reiten, and J. Scott, "Cluster structures for 2-Calabi–Yau categories and unipotent groups", *Compos. Math.* **145**:4 (2009), 1035–1079. MR Zbl

[Chang et al. 2015] W. Chang, J. Zhang, and B. Zhu, "On support τ-tilting modules over endomorphism algebras of rigid objects", *Acta Math. Sin.* (*Engl. Ser.*) **31**:9 (2015), 1508–1516. MR Zbl

[Chang et al. 2018] H. Chang, Y. Zhou, and B. Zhu, "Cotorsion pairs in cluster categories of type A_∞^∞", *J. Combin. Theory Ser. A* **156** (2018), 119–141. MR Zbl

[Fu and Liu 2009] C. Fu and P. Liu, "Lifting to cluster-tilting objects in 2-Calabi–Yau triangulated categories", *Comm. Algebra* **37**:7 (2009), 2410–2418. MR Zbl

[Gratz et al. 2019] S. Gratz, T. Holm, and P. Jørgensen, "Cluster tilting subcategories and torsion pairs in Igusa–Todorov cluster categories of Dynkin type A_∞", *Math. Z.* **292**:1-2 (2019), 33–56. Zbl

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

[Holm and Jørgensen 2012] T. Holm and P. Jørgensen, "On a cluster category of infinite Dynkin type, and the relation to triangulations of the infinity-gon", *Math. Z.* **270**:1-2 (2012), 277–295. MR Zbl

[Igusa and Todorov 2015a] K. Igusa and G. Todorov, "Cluster categories coming from cyclic posets", *Comm. Algebra* **43**:10 (2015), 4367–4402. MR Zbl

[Igusa and Todorov 2015b] K. Igusa and G. Todorov, "Continuous cluster categories, I", *Algebr. Represent. Theory* **18**:1 (2015), 65–101. MR Zbl

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

[Iyama et al. 2014] O. Iyama, P. Jørgensen, and D. Yang, "Intermediate co-t-structures, two-term silting objects, τ-tilting modules, and torsion classes", *Algebra Number Theory* **8**:10 (2014), 2413–2431. MR Zbl

[Jasso 2015] G. Jasso, "Reduction of τ-tilting modules and torsion pairs", *Int. Math. Res. Not.* **2015**:16 (2015), 7190–7237. MR Zbl

[Jørgensen and Palu 2013] P. Jørgensen and Y. Palu, "A Caldero–Chapoton map for infinite clusters", *Trans. Amer. Math. Soc.* **365**:3 (2013), 1125–1147. MR Zbl

[Jørgensen and Yakimov 2017] P. Jørgensen and M. Yakimov, "c-vectors of 2-Calabi–Yau categories and Borel subalgebras of $\mathfrak{sl}_\infty$", preprint, 2017. arXiv

[Keller and Reiten 2007] B. Keller and I. Reiten, "Cluster-tilted algebras are Gorenstein and stably Calabi–Yau", *Adv. Math.* **211**:1 (2007), 123–151. MR Zbl

[Koenig and Zhu 2008] S. Koenig and B. Zhu, "From triangulated categories to abelian categories: cluster tilting in a general framework", *Math. Z.* **258**:1 (2008), 143–160. MR Zbl

[Lasnier 2011] A. Lasnier, "Projective dimensions in cluster-tilted categories", preprint, 2011. arXiv

[Liu and Paquette 2017] S. Liu and C. Paquette, "Cluster categories of type $\mathbb{A}_\infty^\infty$ and triangulations of the infinite strip", *Math. Z.* **286**:1-2 (2017), 197–222. MR Zbl

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

[Ng 2010] P. Ng, "A characterization of torsion theories in the cluster category of Dynkin type $\mathbb{A}_\infty$", preprint, 2010. arXiv

[Reiten and Van den Bergh 2002] I. Reiten and M. Van den Bergh, "Noetherian hereditary abelian categories satisfying Serre duality", *J. Amer. Math. Soc.* **15**:2 (2002), 295–366. MR Zbl

[Smith 2008] D. Smith, "On tilting modules over cluster-tilted algebras", *Illinois J. Math.* **52**:4 (2008), 1223–1247. MR Zbl

[Stovicek and van Roosmalen 2016] J. Stovicek and A.-C. van Roosmalen, "2-Calabi–Yau categories with a directed cluster-tilting subcategory", preprint, 2016. arXiv

[Yang and Zhu 2019] W. Yang and B. Zhu, "Relative cluster tilting objects in triangulated categories", *Trans. Amer. Math. Soc.* **371**:1 (2019), 387–412. MR Zbl

[Yang et al. 2017] W. Yang, J. Zhang, and B. Zhu, "On cluster-tilting objects in a triangulated category with Serre duality", *Comm. Algebra* **45**:1 (2017), 299–311. MR Zbl

[Zhou and Zhu 2011] Y. Zhou and B. Zhu, "Maximal rigid subcategories in 2-Calabi–Yau triangulated categories", *J. Algebra* **348** (2011), 49–60. MR Zbl

[Zhu 2011] B. Zhu, "Cluster-tilted algebras and their intermediate coverings", *Comm. Algebra* **39**:7 (2011), 2437–2448. MR Zbl

Received February 25, 2018. Revised January 16, 2019.

WUZHONG YANG
SCHOOL OF MATHEMATICS
NORTHWEST UNIVERSITY
710127, XI'AN
SHAANXI
CHINA
yangwz@nwu.edu.cn

PANYUE ZHOU
COLLEGE OF MATHEMATICS
HUNAN INSTITUTE OF SCIENCE AND TECHNOLOGY
414006, YUEYANG
HUNAN
CHINA
panyuezhou@163.com

BIN ZHU
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
100084, BEIJING
CHINA
zhu-b@mail.tsinghua.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 301, No. 2, 2019
dx.doi.org/10.2140/pjm.2019.301.741

FREE ROTA–BAXTER FAMILY ALGEBRAS AND (TRI)DENDRIFORM FAMILY ALGEBRAS

YUANYUAN ZHANG AND XING GAO

We construct free commutative Rota–Baxter family algebras, and then free noncommutative Rota–Baxter family algebras via the method of Gröbner–Shirshov bases. We introduce the concept of dendriform (resp. tridendriform) family algebras, and prove that Rota–Baxter family algebras of weight zero (resp. λ) induce dendriform (resp. tridendriform) family algebras. We also construct free commutative dendriform (resp. tridendriform) family algebras.

1. Introduction

Rota–Baxter algebras (previously called Baxter algebras) originated with the American mathematician Glen E. Baxter [1960] in the realm of probability theory. Later Baxter's work was further explored from different angles by Gian-Carlo Rota [1969a; 1969b; 1995], Pierre Cartier [1972] and Frederic V. Atkinson [1963] among others in the 1960-70s. Nowadays, Rota–Baxter algebras have many applications to a broad range of areas, such as combinatorics [Rota 1969a; 1969b; Spitzer 1956], Loday type algebras [Ebrahimi-Fard 2002; Ebrahimi-Fard and Guo 2008], pre-Lie and pre-Poisson algebras [Aguiar 2000b; An and Bai 2008], quantum field theory [Connes and Kreimer 2000; Ebrahimi-Fard et al. 2004; Guo et al. 2017], operads [Bai et al. 2013], Hopf algebras [Connes and Kreimer 1998; Zhang et al. 2016], commutative algebras [Gao et al. 2014], Loday's dendriform algebras [Ebrahimi-Fard and Guo 2008; Loday and Ronco 2004], as well as to Aguiar's associative analogue of the classical Yang–Baxter equation [Aguiar 2000a; 2001; Bai et al. 2012].

A Rota–Baxter family algebra is a generalization of a Rota–Baxter algebra. It arises naturally in renormalization of quantum field theory [Ebrahimi-Fard et al. 2007, Proposition 9.1], which plays an independent role in the physics study. See Definition 2.3 below. It is worthwhile to study the algebraic structure of Rota–Baxter family algebras. The free objects play a crucial role in the study of any algebraic structures, such as the construction of free differential algebras in terms of differential monomials, the construction of free Rota–Baxter algebras [Guo and Keigher

MSC2010: 08B20, 13P10, 16S10, 16W99.
Keywords: Rota–Baxter algebra, Rota–Baxter family algebra, dendriform family algebra, tridendriform family algebra, Gröbner–Shirshov basis.

2000b] which is more involved, the free differential Rota–Baxter algebras [Guo and Keigher 2008] composing the construction of free differential algebras followed by that of the free Rota–Baxter algebras, and the free integro-differential algebras [Gao et al. 2015; 2014] for analyzing the underlying algebraic structures of boundary problems for linear ordinary differential equations. In the present paper, we mainly construct free commutative Rota–Baxter family algebras, and free noncommutative Rota–Baxter family algebras by the method of Gröbner–Shirshov bases.

The notion of a dendriform algebra was introduced by Loday [1993] with motivation from algebraic K-theory. It has been studied extensively with connections to several areas in mathematics and physics, including operads [Loday 2004], homology [Frabetti 1998], arithmetic [Loday 2002] and quantum field theory [Ebrahimi-Fard et al. 2008] and pre-Poisson algebras [Aguiar 2000b]. In this paper, we also propose the concept of dendriform (resp. tridendriform) family algebras, and then prove that Rota–Baxter family algebras of weight zero (resp. λ) induce dendriform (resp. tridendriform) family algebras, just as the classical case where Rota–Baxter algebras of weight zero (resp. λ) induce dendriform (resp. tridendriform) algebras. The free commutative dendriform (resp. tridendriform) family algebras are also constructed.

The layout of the paper is as follows. In Section 2, after recalling the concept of Rota–Baxter family algebras, we construct the free commutative Rota–Baxter family algebra on an algebra (Theorem 2.11). In Section 3, we first recall the construction of free Ω-operated algebras $\boldsymbol{k}\mathfrak{M}(\Omega, X)$ on a set X, where Ω is a nonempty set. Then we show that the defining relation of a Rota–Baxter family algebra is a Gröbner–Shirshov basis in $\boldsymbol{k}\mathfrak{M}(\Omega, X)$ (Theorem 3.10), with respect to a known monomial order. Here Ω is a semigroup. Using the composition-diamond lemma, we obtain a $\boldsymbol{k}$-basis of the free noncommutative Rota–Baxter family algebra on a set X (Theorem 3.11). In Section 4, we introduce the concept of (tri)dendriform family algebras, which are a generalization of the classical (tri)dendriform algebras. As in the case of Rota–Baxter algebras that induce (tri)dendriform algebras, we obtain that Rota–Baxter family algebras induce (tri)dendriform family algebras (Theorem 4.4). We end the section with the construction of free commutative (tri)dendriform family algebras (Theorem 4.9).

Notation. Throughout this paper, we fix a commutative unitary ring $\boldsymbol{k}$. By an algebra we mean an associative (but not necessarily commutative) unitary $\boldsymbol{k}$-algebra, unless the contrary is specified.

2. Free commutative Rota–Baxter family algebras

This section is devoted to the construction of the free commutative Rota–Baxter family algebra on a commutative algebra.

2A. ***Free commutative Rota–Baxter algebras.*** In this subsection, we review the construction of the free commutative Rota–Baxter algebra on a commutative algebra A via mixable shuffle product [Guo 2012; Guo and Keigher 2000a].

Definition 2.1. Let λ be a given element in $\boldsymbol{k}$. A *Rota–Baxter algebra of weight* λ is a pair (R, P) consisting of an algebra R with a linear operator $P : R \to R$ that satisfies the Rota–Baxter equation

$$P(a)P(b) = P(P(a)b + aP(b) + \lambda ab) \quad \text{for } a, b \in R.$$

Then P is called a *Rota–Baxter operator of weight* λ. If, further, R is commutative, then (R, P) is called a commutative Rota–Baxter algebra of weight λ.

We often suppress the operator P on a Rota–Baxter algebra (R, P) when there is no danger of confusion. The following is the construction of the free commutative Rota–Baxter algebra.

Let A be a commutative unital algebra. Denote

$$Ш^+(A) := Ш^+_{\boldsymbol{k},\lambda}(A) = \bigoplus_{k\geq 0} A^{\otimes k} = \boldsymbol{k} \oplus A \oplus A^{\otimes 2} \oplus \cdots .$$

Define

$$Ш(A) := Ш_{\boldsymbol{k},\lambda}(A) := A \otimes Ш^+(A)$$

to be the tensor product algebra with the augmented mixable shuffle product $\diamond_\lambda$ given by

$$(a_0 \otimes \mathfrak{a}') \diamond_\lambda (b_0 \otimes \mathfrak{b}') := (a_0 b_0) \otimes (\mathfrak{a}' ш_\lambda \mathfrak{b}') \quad \text{for } a_0, b_0 \in A, \mathfrak{a}', \mathfrak{b}' \in Ш^+(A),$$

where $ш_\lambda$ is the mixable shuffle product [Guo and Keigher 2000a]. Further define

$$P := P_A : Ш(A) \to Ш(A), \quad \mathfrak{a} \mapsto 1 \otimes \mathfrak{a}, \quad \text{for } \mathfrak{a} \in Ш(A).$$

Theorem 2.2 [Guo 2012; Guo and Keigher 2000a]. *Let A be a commutative algebra and $\lambda \in \boldsymbol{k}$ be given. Then the triple* $(Ш(A), \diamond_\lambda, P)$, *together with the natural embedding $j_A : A \to Ш(A)$, is the free commutative Rota–Baxter algebra of weight λ on A.*

2B. ***The construction of free commutative Rota–Baxter family algebras.*** In this subsection, we first recall the concept of Rota–Baxter family algebras, which arise naturally in renormalization of quantum field theory [Ebrahimi-Fard et al. 2007, Proposition 9.1]. Then we proceed to construct free commutative Rota–Baxter family algebras.

Definition 2.3 [Ebrahimi-Fard et al. 2007; Guo 2009]. Let Ω be a semigroup and $\lambda \in \boldsymbol{k}$ be given. A *Rota–Baxter family* of weight λ on an algebra R is a collection of linear operators $\{P_\omega \mid \omega \in \Omega\}$ on R such that

$$(1) \quad P_\alpha(a)P_\beta(b) = P_{\alpha\beta}(P_\alpha(a)b + aP_\beta(b) + \lambda ab) \quad \text{for } a, b \in R \text{ and } \alpha, \beta \in \Omega.$$

Then the pair $(R, \{P_\omega \mid \omega \in \Omega\})$ is called a *Rota–Baxter family algebra* of weight λ. If further R is commutative, then $(R, \{P_\omega \mid \omega \in \Omega\})$ is called a commutative Rota–Baxter family algebra of weight λ.

Definition 2.4. Let $(R, \{P_\omega \mid \omega \in \Omega\})$ and $(R', \{P'_\omega \mid \omega \in \Omega\})$ be two Rota–Baxter family algebras of weight λ. A map $f : R \to R'$ is called *a Rota–Baxter family algebra morphism* if f is an algebra homomorphism and $f \circ P_\omega = P'_\omega \circ f$ for each $\omega \in \Omega$.

Remark 2.5. In a Rota–Baxter family algebra $(R, \{P_\omega \mid \omega \in \Omega\})$ of weight λ, for each $\omega \in \Omega$, it follows from (1) that

$$P_\omega(a)P_\omega(b) = P_{\omega^2}(P_\omega(a)b + aP_\omega(b) + \lambda ab) \quad \text{for } a, b \in R,$$

whence $P_\omega : R \to R$ is not necessarily a Rota–Baxter operator of weight λ. If Ω is a trivial semigroup (a semigroup with one element), then $\omega^2 = \omega$, and P_ω is a Rota–Baxter operator of weight λ.

We provide some examples.

Example 2.6. Any Rota–Baxter algebra of weight λ can be viewed as a Rota–Baxter family algebra of weight λ by taking Ω to be a trivial semigroup (a semigroup with one element).

Example 2.7. The l-jets define a Rota–Baxter family of weight -1 in perturbative quantum field theory. The reader is referred to [Ebrahimi-Fard et al. 2007, Section 6] for more details.

Remark 2.8. Let Ω be a commutative semigroup. Then (1) is Lie compatible in the sense that

$$[P_\alpha(a), P_\beta(b)] = P_{\alpha\beta}([P_\alpha(a), b] + [a, P_\beta(b)] + \lambda[a, b]) \quad \text{for } a, b \in R \text{ and } \alpha, \beta \in \Omega.$$

Here the Lie bracket is taken as the commutator. Indeed, for $a, b \in R$,

$$\begin{aligned}
&[P_\alpha(a), P_\beta(b)] \\
&= P_\alpha(a)P_\beta(b) - P_\beta(b)P_\alpha(a) \\
&= P_{\alpha\beta}(P_\alpha(a)b + aP_\beta(b) + \lambda ab) - P_{\beta\alpha}(P_\beta(b)a + bP_\alpha(a) + \lambda ba) \qquad \text{(by (1))} \\
&= P_{\alpha\beta}(P_\alpha(a)b + aP_\beta(b) + \lambda ab - P_\beta(b)a - bP_\alpha(a) - \lambda ba) \\
&\qquad\qquad\qquad \text{(by } \Omega \text{ being commutative)} \\
&= P_{\alpha\beta}([P_\alpha(a), b] + [a, P_\beta(b)] + \lambda[a, b]),
\end{aligned}$$

This Lie compatibility is the very image of a good behavior of (1).

Now we are in a position to construct free Rota–Baxter family algebras.

Definition 2.9. Let A be a commutative algebra, Ω a semigroup and $\lambda \in \boldsymbol{k}$ be given. A *free commutative Rota–Baxter family algebra* of weight λ on A is a commutative Rota–Baxter family algebra $F_{\mathrm{RBF}}(A)$ of weight λ together with an algebra homomorphism $j_A : A \to F_{\mathrm{RBF}}(A)$ that satisfies the following universal property: for any commutative Rota–Baxter family algebra $(R, \{P_{R,\omega} \mid \omega \in \Omega\})$ of weight λ and any algebra homomorphism $f : A \to R$, there is a unique Rota–Baxter family algebra morphism $\bar{f} : F_{\mathrm{RBF}}(A) \to R$ such that $f = \bar{f} \circ j_A$.

Let us first construct the underlying module of $F_{\mathrm{RBF}}(A)$. Recall that in the free commutative Rota–Baxter algebra $(Ш(A), P)$, the Rota–Baxter operator is

$$P : Ш(A) \to Ш(A), \quad \mathfrak{a} \mapsto P(\mathfrak{a}) = 1 \otimes \mathfrak{a}.$$

Roughly speaking, we can use the notation $\otimes$ to replace the Rota–Baxter operator P to obtain $Ш(A)$. Now in the case of Rota–Baxter family algebras, the Rota–Baxter family algebra operators P_ω, $\omega \in \Omega$, are indexed by a semigroup Ω. Correspondingly, we need to decorate the notation $\otimes$ with elements $\omega \in \Omega$ to obtain $\otimes_\omega$. Here the notation $\otimes_\omega$ is essentially the usual tensor product $\otimes$ over $\boldsymbol{k}$, and the index ω is just used to correspond to the operator P_ω. More precisely, for a commutative algebra A, define

$$A \otimes_\omega A := A \otimes A \ \text{ and } \ A \otimes_\Omega A := \bigoplus_{\omega \in \Omega} A \otimes_\omega A.$$

Since $\otimes_\omega$ with $\omega \in \Omega$ is the usual tensor product $\otimes$ substantially, there is no accident that the $\otimes_\Omega$ is associative. Namely,

$$\begin{aligned}(A \otimes_\Omega A) \otimes_\Omega A &= \bigoplus_{\omega_1, \omega_2 \in \Omega} (A \otimes_{\omega_1} A) \otimes_{\omega_2} A = \bigoplus_{\omega_1, \omega_2 \in \Omega} (A \otimes A) \otimes A \\ &= \bigoplus_{\omega_1, \omega_2 \in \Omega} A \otimes (A \otimes A) = A \otimes_\Omega (A \otimes_\Omega A).\end{aligned}$$

So we may denote

$$A^{\otimes_\Omega n} := A \otimes_\Omega \cdots \otimes_\Omega A = \bigoplus_{\omega_1, \dots, \omega_n \in \Omega} A \otimes_{\omega_1} \cdots \otimes_{\omega_n} A, \ \text{ for } n \geq 0,$$

with the convention that $A^{\otimes_\Omega n} = \boldsymbol{k}$ when $n = 0$. Define the $\boldsymbol{k}$-modules

$$Ш_\Omega^+(A) := \bigoplus_{n \geq 0} A^{\otimes_\Omega n} \ \text{ and } \ Ш_\Omega(A) := A \otimes_\Omega Ш_\Omega^+(A) = \bigoplus_{n \geq 1} A^{\otimes_\Omega n}.$$

Now we equip a commutative product $\diamond_\Omega$ on $Ш_\Omega(A)$, which is a generalization of the augmented mixable shuffle product $\diamond_\Omega$ [Guo 2012]. Let

(2) $$\mathfrak{a} := a_0 \otimes_{\alpha_1} \mathfrak{a}', \ \mathfrak{b} := b_0 \otimes_{\beta_1} \mathfrak{b}' \in Ш_\Omega(A) = A \otimes_\Omega Ш^+_\Omega(A),$$

where $a_0, b_0 \in A$, $\alpha_1, \beta_1 \in \Omega$ and $\mathfrak{a}', \mathfrak{b}' \in Ш^+_\Omega(A)$. Write

(3) $$\mathfrak{a}' = a_1 \otimes_{\alpha_2} a_2 \otimes_{\alpha_3} \cdots \otimes_{\alpha_m} a_m \in A^{\otimes_\Omega m} \text{ and } \mathfrak{b}' = b_1 \otimes_{\beta_2} b_2 \otimes_{\beta_3} \cdots \otimes_{\beta_n} b_n \in A^{\otimes_\Omega n}, \quad \text{with } m, n \geq 0.$$

Here we use the convention that $\mathfrak{a}' = \mathfrak{b}' = 1_{\boldsymbol{k}}$ when $m = n = 0$.

In the rest of the paper, we shall use the notation in (2) and (3), which will be employed frequently.

We now define $\mathfrak{a} \diamond_\Omega \mathfrak{b}$ inductively on $m + n \geq 0$. For the initial step of $m + n = 0$, we have $\mathfrak{a} = a_0$, $\mathfrak{b} = b_0$ and define

$$\mathfrak{a} \diamond_\Omega \mathfrak{b} := a_0 b_0.$$

For the induction step of $m + n \geq 1$, we define

(4) $$\begin{aligned} \mathfrak{a} \diamond_\Omega \mathfrak{b} &:= (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (b_0 \otimes_{\beta_1} \mathfrak{b}') \\ &:= a_0 b_0 \otimes_{\alpha_1 \beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}' + \mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}') + \lambda(\mathfrak{a}' \diamond_\Omega \mathfrak{b}')). \end{aligned}$$

Equivalently,

(5) $$\mathfrak{a} \diamond_\Omega \mathfrak{b} := \begin{cases} a_0 b_0 & \text{for } m, n = 0; \\ a_0 b_0 \otimes_{\alpha_1} a_1 \otimes_{\alpha_2} \cdots \otimes_{\alpha_m} a_m & \text{for } m \geq 1, n = 0; \\ a_0 b_0 \otimes_{\beta_1} b_1 \otimes_{\beta_2} \cdots \otimes_{\beta_n} b_n & \text{for } m = 0, n \geq 1; \\ a_0 b_0 \otimes_{\alpha_1 \beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') & \\ \quad + a_0 b_0 \otimes_{\alpha_1 \beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) + \lambda a_0 b_0 \otimes_{\alpha_1 \beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}') & \text{for } m \geq 1, n \geq 1. \end{cases}$$

Then a binary operation $\diamond_\Omega$ is given on $Ш_\Omega(A)$ after extending by biadditivity.

Remark 2.10. If Ω is a trivial semigroup with only one element, then $\diamond_\Omega$ is precisely the augmented mixable shuffle product introduced in [Guo 2012; Guo and Keigher 2000a].

Now we arrive at our main result in this section. For each $\omega \in \Omega$, define

(6) $$P_{A,\omega} : Ш_\Omega(A) \to Ш_\Omega(A), \quad \mathfrak{a} \mapsto 1 \otimes_\omega \mathfrak{a}, \quad \text{for } \mathfrak{a} \in Ш_\Omega(A).$$

Theorem 2.11. *Let A be a commutative algebra, Ω be a commutative semigroup, and let $\lambda \in \boldsymbol{k}$ be given.*

(a) *The pair $(Ш_\Omega(A), \diamond_\Omega)$ is a commutative algebra.*

(b) *The triple*

$$(\text{Ш}_\Omega(A), \diamond_\Omega, \{P_{A,\,\omega} \mid \omega \in \Omega\}),$$

together with the natural embedding $j_A : A \to \text{Ш}_\Omega(A)$, *is the free commutative Rota–Baxter family algebra of weight* λ *on* A.

Proof. We use the notation in (2) and (3).

(a) Let us first prove

$$\mathfrak{a} \diamond_\Omega \mathfrak{b} = \mathfrak{b} \diamond_\Omega \mathfrak{a}$$

by induction on $m+n \geq 0$. For the initial step of $m+n=0$, we have $m=n=0$, and by (4),

$$\mathfrak{a} \diamond_\Omega \mathfrak{b} = a_0 \diamond_\Omega b_0 = a_0 b_0 = b_0 a_0 = b_0 \diamond_\Omega a_0.$$

For the induction step of $m+n \geq 1$,

$$\begin{aligned}
\mathfrak{a} \diamond_\Omega \mathfrak{b} &= a_0 b_0 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}') + (1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}' + \lambda(\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \\
&= b_0 a_0 \otimes_{\beta_1\alpha_1} ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{a}' + \mathfrak{b}' \diamond_\Omega (1 \otimes_{\alpha_1} \mathfrak{a}') + \lambda(\mathfrak{b}' \diamond_\Omega \mathfrak{a}')) \\
&\hspace{20em} \text{(by the induction hypothesis)} \\
&= \mathfrak{b} \diamond_\Omega \mathfrak{a},
\end{aligned}$$

We next prove the associativity:

$$(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c} = \mathfrak{a} \diamond_\Omega (\mathfrak{b} \diamond_\Omega \mathfrak{c}), \tag{7}$$

where $\mathfrak{a} \in A^{\otimes_\Omega (m+1)}$, $\mathfrak{b} \in A^{\otimes_\Omega (n+1)}$, $\mathfrak{c} \in A^{\otimes_\Omega (l+1)}$, with $m, n, l \geq 0$. For this we use the induction on the sum $m+n+l \geq 0$. The initial case of $m+n+l=0$ follows from the first case of (5) and the associativity of the multiplication of A. We suppose (7) has been proved when $m+n+l=k$ for a $k \geq 0$, and consider the case when $m+n+l=k+1$.

From the second and third cases in (5), we have

$$(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c} = \mathfrak{a} \diamond_\Omega (\mathfrak{b} \diamond_\Omega \mathfrak{c}) \quad \text{for } m=0, \text{or n}=0, \text{or } l=0.$$

Now for $m \geq 1,\ n \geq 1,\ l \geq 1$, denote $\mathfrak{a} = a_0 \otimes_{\alpha_1} \mathfrak{a}'$, $\mathfrak{b} = b_0 \otimes_{\beta_1} \mathfrak{b}'$, $\mathfrak{c} = c_0 \otimes_{\gamma_1} \mathfrak{c}'$, where $a_0, b_0, c_0 \in A$, $\mathfrak{a}' \in A^{\otimes_\Omega m}$, $\mathfrak{b}' \in A^{\otimes_\Omega n}$, $\mathfrak{c}' \in A^{\otimes_\Omega l}$. For the left-hand side of (7),

$$\begin{aligned}
(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c} &= ((a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (b_0 \otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}') \\
&= (a_0 b_0 \otimes_{\alpha_1\beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') + a_0 b_0 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) \\
&\qquad + \lambda a_0 b_0 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}') \\
&= (a_0 b_0 \otimes_{\alpha_1\beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}')) \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}') \\
&\qquad + (a_0 b_0 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}'))) \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}') \\
&\qquad + (\lambda a_0 b_0 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}')
\end{aligned}$$

$$
\begin{aligned}
&= a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') \\
&\qquad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\qquad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') \diamond_\Omega \mathfrak{c}') \\
&\qquad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}'))) \diamond_\Omega \mathfrak{c}') \\
&\qquad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\qquad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') \\
&\qquad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') \\
&\qquad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega \mathfrak{b}') \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\qquad + \lambda^2 a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega \mathfrak{b}') \diamond_\Omega \mathfrak{c}').
\end{aligned}
$$

For the fifth summand on the right-hand side of the last equation, by the induction hypothesis,

$$
\begin{aligned}
&(\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}') \\
&\qquad = \mathfrak{a}' \diamond_\Omega ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\qquad = \mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}')) + \mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}'))) \\
&\qquad\qquad + \lambda \mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega \mathfrak{c}')),
\end{aligned}
$$

whence

$$
\begin{aligned}
&(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c} = \\
&\ a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} ((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') \\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (((1 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}') \diamond_\Omega \mathfrak{c}') \\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}'))) \diamond_\Omega \mathfrak{c}') \\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}'))) \\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')))) + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega \mathfrak{c}'))) \\
&\quad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega (1 \otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1 \otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}') \\
&\quad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega \mathfrak{b}') \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) + \lambda^2 a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((\mathfrak{a}' \diamond_\Omega \mathfrak{b}') \diamond_\Omega \mathfrak{c}').
\end{aligned}
$$

For the right-hand side of (7),

$$
\begin{aligned}
&\mathfrak{a} \diamond_\Omega (\mathfrak{b} \diamond_\Omega \mathfrak{c}) \\
&\ = (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega ((b_0 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega (c_0 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\ = (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \big(b_0c_0 \otimes_{\beta_1\gamma_1} ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}') + b_0c_0 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}')) \\
&\qquad\qquad + \lambda b_0c_0 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega \mathfrak{c}')\big) \\
&\ = (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (b_0c_0 \otimes_{\beta_1\gamma_1} ((1 \otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}')) \\
&\qquad + (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (b_0c_0 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega (1 \otimes_{\gamma_1} \mathfrak{c}'))) + (a_0 \otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (\lambda b_0c_0 \otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega \mathfrak{c}'))
\end{aligned}
$$

$$
\begin{aligned}
&= a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega ((1\otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}'))\\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1\otimes_{\beta_1\gamma_1} ((1\otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}')))\\
&\quad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega ((1\otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}'))\\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (\mathfrak{b}' \diamond_\Omega (1\otimes_{\gamma_1} \mathfrak{c}')))\\
&\quad + a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1\otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega (1\otimes_{\gamma_1} \mathfrak{c}'))))\\
&\quad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (\mathfrak{b}' \diamond_\Omega (1\otimes_{\gamma_1} \mathfrak{c}'))) + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} ((1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (\mathfrak{b}' \diamond_\Omega \mathfrak{c}'))\\
&\quad + \lambda a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (1\otimes_{\beta_1\gamma_1} (\mathfrak{b}' \diamond_\Omega \mathfrak{c}'))) + \lambda^2 a_0b_0c_0 \otimes_{\alpha_1\beta_1\gamma_1} (\mathfrak{a}' \diamond_\Omega (\mathfrak{b}' \diamond_\Omega \mathfrak{c})).
\end{aligned}
$$

For the first summand on the right-hand side of the last equation, by the induction hypothesis,

$$
\begin{aligned}
&(1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega ((1\otimes_{\beta_1} \mathfrak{b}') \diamond_\Omega \mathfrak{c}')\\
&\qquad = ((1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega (1\otimes_{\beta_1} \mathfrak{b}')) \diamond_\Omega \mathfrak{c}'\\
&\qquad = (1\otimes_{\alpha_1\beta_1} ((1\otimes_{\alpha_1} \mathfrak{a}') \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}' + (1\otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega (1\otimes_{\beta_1} \mathfrak{b}'))) \diamond_\Omega \mathfrak{c}'\\
&\qquad\qquad + \lambda (1\otimes_{\alpha_1\beta_1} (\mathfrak{a}' \diamond_\Omega \mathfrak{b}')) \diamond_\Omega \mathfrak{c}',
\end{aligned}
$$

which yields

$$
\begin{aligned}
&\mathfrak{a}\diamond_\Omega(\mathfrak{b}\diamond_\Omega\mathfrak{c})\\
&= a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}((1\otimes_{\alpha_1\beta_1}((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega\mathfrak{b}'))\diamond_\Omega\mathfrak{c}')\\
&\quad + a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}((1\otimes_{\alpha_1\beta_1}(\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1}\mathfrak{b}'))\diamond_\Omega\mathfrak{c}')\\
&\quad + \lambda a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}((1\otimes_{\alpha_1\beta_1}(\mathfrak{a}'\diamond_\Omega\mathfrak{b}'))\diamond_\Omega\mathfrak{c}') + a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1\gamma_1}((1\otimes_{\beta_1}\mathfrak{b}')\diamond_\Omega\mathfrak{c}')))\\
&\quad + \lambda a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega((1\otimes_{\beta_1}\mathfrak{b}')\diamond_\Omega\mathfrak{c}')) + a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega(\mathfrak{b}'\diamond_\Omega(1\otimes_{\gamma_1}\mathfrak{c}')))\\
&\quad + a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1\gamma_1}(\mathfrak{b}'\diamond_\Omega(1\otimes_{\gamma_1}\mathfrak{c}')))) + \lambda a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega(\mathfrak{b}'\diamond_\Omega(1\otimes_{\gamma_1}\mathfrak{c}')))\\
&\quad + \lambda a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega(\mathfrak{b}'\diamond_\Omega\mathfrak{c}')) + \lambda a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1\gamma_1}(\mathfrak{b}'\diamond_\Omega\mathfrak{c}')))\\
&\quad + \lambda^2 a_0b_0c_0\otimes_{\alpha_1\beta_1\gamma_1}(\mathfrak{a}'\diamond_\Omega(\mathfrak{b}'\diamond_\Omega\mathfrak{c})).
\end{aligned}
$$

Now we see that the i-th term in the expansion of $(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c}$ is equal to the $\sigma(i)$-th term in the expansion of $\mathfrak{a} \diamond_\Omega (\mathfrak{b} \diamond_\Omega \mathfrak{c})$ by the induction hypothesis, where σ is a permutation of order 11:

$$
\begin{pmatrix} i \\ \sigma(i) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & 6 & 9 & 2 & 4 & 7 & 10 & 5 & 3 & 8 & 11 \end{pmatrix}.
$$

Thus $(\mathfrak{a} \diamond_\Omega \mathfrak{b}) \diamond_\Omega \mathfrak{c} = \mathfrak{a} \diamond_\Omega (\mathfrak{b} \diamond_\Omega \mathfrak{c})$. This completes the inductive proof of (7).

(b) We divide the proof into two steps.

Step 1: We show that $\{P_{A,\omega} \mid \omega \in \Omega\}$ in (6) is a Rota–Baxter family on $(Ш_\Omega(A), \diamond_\Omega)$.

Indeed for $\alpha_1, \beta_1 \in \Omega$ and $\mathfrak{a}', \mathfrak{b}' \in Ш_\Omega(A)$, we have

$$\begin{aligned}
&P_{A,\alpha_1}(\mathfrak{a}')\diamond_\Omega P_{A,\beta_1}(\mathfrak{b}')\\
&=(1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega(1\otimes_{\beta_1}\mathfrak{b}') && \text{(by (6))}\\
&=1\otimes_{\alpha_1\beta_1}((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega\mathfrak{b}'+\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1}\mathfrak{b}')+\lambda(\mathfrak{a}'\diamond_\Omega\mathfrak{b}'))\\
&=1\otimes_{\alpha_1\beta_1}(P_{A,\alpha_1}(\mathfrak{a}')\diamond_\Omega\mathfrak{b}'+\mathfrak{a}'\diamond_\Omega P_{A,\beta_1}(\mathfrak{b}')+\lambda(\mathfrak{a}'\diamond_\Omega\mathfrak{b}')) && \text{(by (4))}\\
&=P_{A,\alpha_1\beta_1}(P_{A,\alpha_1}(\mathfrak{a}')\diamond_\Omega\mathfrak{b}')+P_{A,\alpha_1\beta_1}(\mathfrak{a}'\diamond_\Omega P_{A,\beta_1}(\mathfrak{b}'))+\lambda P_{A,\alpha_1\beta_1}(\mathfrak{a}'\diamond_\Omega\mathfrak{b}') && \text{(by (6))}
\end{aligned}$$

Step 2: We prove the universal property of $(Ш_\Omega(A), \{P_{A,\omega} \mid \omega \in \Omega\})$. Let $(R, \{P_{R,\omega} \mid \omega \in \Omega\})$ be a commutative Rota–Baxter family algebra of weight λ and let $f : A \to R$ be an algebra homomorphism.

(existence) To construct a linear map $\bar{f} : Ш_\Omega(A) \to R$, it suffices to define $\bar{f}(\mathfrak{a})$ for a pure tensor $\mathfrak{a} = a_0 \otimes_{\alpha_1} \mathfrak{a}' \in A^{\otimes_\Omega(m+1)}$ with $m \geq 0$ and $a_0 \in A$. For this we employ induction on $m \geq 0$. For the initial step of $m = 0$, we have $\mathfrak{a} = a_0$ and define

$$\bar{f}(\mathfrak{a}) := \bar{f}(a_0) := f(a_0).$$

For the induction step of $m \geq 1$, we define

$$\bar{f}(\mathfrak{a}) := \bar{f}(a_0 \otimes_{\alpha_1} \mathfrak{a}') := f(a_0)P_{R,\alpha_1}(\bar{f}(\mathfrak{a}')). \tag{8}$$

Since

$$\begin{aligned}
\bar{f} \circ P_{A,\omega}(\mathfrak{a}) &= \bar{f}(1 \otimes_\omega \mathfrak{a})\\
&= f(1)P_{R,\omega}(\bar{f}(\mathfrak{a})) = P_{R,\omega}(\bar{f}(\mathfrak{a})) = P_{R,\omega} \circ \bar{f}(\mathfrak{a}) \quad \text{for } \mathfrak{a} \in Ш_\Omega(A),
\end{aligned}$$

we have

$$\bar{f} \circ P_{A,\omega} = P_{R,\omega} \circ \bar{f} \quad \text{for } \omega \in \Omega. \tag{9}$$

We now prove the compatibility of $\bar{f}$ with the multiplication $\diamond_\Omega$:

$$\bar{f}(\mathfrak{a} \diamond_\Omega \mathfrak{b}) = \bar{f}(\mathfrak{a})\bar{f}(\mathfrak{b}) \quad \text{for } \mathfrak{a} \in A^{\otimes_\Omega(m+1)} \text{ and } \mathfrak{b} \in A^{\otimes_\Omega(n+1)} \tag{10}$$

by induction on $m + n \geq 0$. When $m = n = 0$, we have

$$\mathfrak{a} = a_0, \quad \mathfrak{b} = b_0 \in A \quad \text{and} \quad \mathfrak{a} \diamond_\Omega \mathfrak{b} = a_0 b_0,$$

and so

$$\bar{f}(\mathfrak{a} \diamond_\Omega \mathfrak{b}) = \bar{f}(a_0 b_0) = f(a_0 b_0) = f(a_0)f(b_0) = \bar{f}(\mathfrak{a})\bar{f}(\mathfrak{b}),$$

by f being an algebra homomorphism. Suppose (10) has been validated for $m+n \leq k$ with a $k \geq 0$, and consider the case of $m + n = k + 1$. Then

$$\begin{aligned}
\mathfrak{a} &= a_0 \otimes_{\alpha_1} \mathfrak{a}' = a_0 \diamond_\Omega P_{A,\alpha_1}(\mathfrak{a}') \in A^{\otimes_\Omega(m+1)},\\
\mathfrak{b} &= b_0 \otimes_{\beta_1} \mathfrak{b}' = b_0 \diamond_\Omega P_{A,\beta_1}(\mathfrak{b}') \in A^{\otimes_\Omega(n+1)},
\end{aligned}$$

and so

$$
\begin{aligned}
&\bar{f}(\mathfrak{a}\diamond_\Omega\mathfrak{b})\\
&=\bar{f}((a_0\diamond P_{A,\alpha_1}(\mathfrak{a}'))\diamond_\Omega(b_0\diamond P_{A,\beta_1}(\mathfrak{b}')))\\
&=\bar{f}((a_0b_0)\diamond_\Omega P_{A,\alpha_1}(\mathfrak{a}')\diamond_\Omega P_{A,\beta_1}(\mathfrak{b}'))\\
&=\bar{f}((a_0b_0)\diamond_\Omega P_{A,\alpha_1\beta_1}(P_{A,\alpha_1}(\mathfrak{a}')\diamond_\Omega\mathfrak{b}'+\mathfrak{a}'\diamond_\Omega P_{A,\beta_1}(\mathfrak{b}')+\lambda(\mathfrak{a}'\diamond_\Omega\mathfrak{b}'))) && \text{(by (1))}\\
&=\bar{f}(a_0b_0)(\bar{f}\circ P_{A,\alpha_1\beta_1})((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega\mathfrak{b}'+\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1}\mathfrak{b}')+\lambda(\mathfrak{a}'\diamond_\Omega\mathfrak{b}')) && \text{(by (6), (8), (9))}\\
&=f(a_0b_0)(P_{R,\alpha_1\beta_1}\circ\bar{f})((1\otimes_{\alpha_1}\mathfrak{a}')\diamond_\Omega\mathfrak{b}'+\mathfrak{a}'\diamond_\Omega(1\otimes_{\beta_1}\mathfrak{b}')+\lambda(\mathfrak{a}'\diamond_\Omega\mathfrak{b}')) && \text{(by (9))}\\
&=f(a_0b_0)P_{R,\alpha_1\beta_1}(\bar{f}(1\otimes_{\alpha_1}\mathfrak{a}')\bar{f}(\mathfrak{b}')+\bar{f}(\mathfrak{a}')\bar{f}(1\otimes_{\beta_1}\mathfrak{b}')+\lambda\bar{f}(\mathfrak{a}')\bar{f}(\mathfrak{b}'))\\
& && \text{(by the induction hypothesis)}\\
&=f(a_0b_0)P_{R,\alpha_1\beta_1}(\bar{f}(P_{A,\alpha_1}(\mathfrak{a}'))\bar{f}(\mathfrak{b}')+\bar{f}(\mathfrak{a}')\bar{f}(P_{A,\beta_1}(\mathfrak{b}'))+\lambda\bar{f}(\mathfrak{a}')\bar{f}(\mathfrak{b}')) && \text{(by (6))}\\
&=f(a_0b_0)P_{R,\alpha_1\beta_1}(P_{R,\alpha_1}(\bar{f}(\mathfrak{a}'))\bar{f}(\mathfrak{b}')+\bar{f}(\mathfrak{a}')P_{R,\beta_1}(\bar{f}(\mathfrak{b}'))+\lambda\bar{f}(\mathfrak{a}')\bar{f}(\mathfrak{b}')) && \text{(by (9))}\\
&=f(a_0)f(b_0)P_{R,\alpha_1}(\bar{f}(\mathfrak{a}'))P_{R,\beta_1}(\bar{f}(\mathfrak{b}')) && \text{(by (1))}\\
&=f(a_0)P_{R,\alpha_1}(\bar{f}(\mathfrak{a}'))f(b_0)P_{R,\beta_1}(\bar{f}(\mathfrak{b}'))\\
&=\bar{f}(\mathfrak{a})\bar{f}(\mathfrak{b}) && \text{(by (8))}
\end{aligned}
$$

(uniqueness) In fact, for $\mathfrak{a}=a_0\otimes_{\alpha_1}a_1\otimes_{\alpha_2}a_2\otimes_{\alpha_3}\cdots\otimes_{\alpha_m}a_m\in A^{\otimes(m+1)}$, we must have

$$\bar{f}(\mathfrak{a})=f(a_0)P_{R,\alpha_1}(f(a_1)P_{R,\alpha_2}(f(a_2)\cdots)).$$

So the uniqueness of $\bar{f}$ is proved. This completes the proof. □

As a direct application of Theorem 2.11, we obtain the free Rota–Baxter family algebra on a set X. Denote by $\boldsymbol{k}[X]$ the free commutative algebra on X.

Corollary 2.12. *Let X be a set and Ω a commutative semigroup, and let $\lambda\in\boldsymbol{k}$ be given. The Rota–Baxter family algebra $(Ш_\Omega(\boldsymbol{k}[X]),\{P_{\boldsymbol{k}[X],\omega}\mid\omega\in\Omega\})$, together with the natural embedding*

$$j_X : X\hookrightarrow\boldsymbol{k}[X]\hookrightarrow Ш_\Omega(\boldsymbol{k}[X]),$$

is the free commutative Rota–Baxter family algebra of weight λ on X.

Proof. This follows from Theorem 2.11 by taking $A=\boldsymbol{k}[X]$. □

3. Free noncommutative Rota–Baxter family algebras

In this section, we construct free noncommutative Rota–Baxter family algebras in terms of the method of Gröbner–Shirshov bases.

3A. ***Free Ω-operated algebras.*** The concept of algebras with (one or more) linear operators was introduced by A. G. Kurosh [1960]. Later, Guo [2009] called such algebras operated algebras and constructed the free objects. See also [Bokut and Chen 2014; Gao and Guo 2017; Guo et al. 2013].

Definition 3.1 [Guo 2009]. Let Ω be a nonempty set. An *Ω-operated algebra* is an algebra R together with a set of operators $P_\omega : R \to R,\ \omega \in \Omega$. A morphism from an Ω-operated algebra $(R, \{P_\omega \mid \omega \in \Omega\})$ to an Ω-operated algebra $(R', \{P'_\omega \mid \omega \in \Omega\})$ is an *algebra homomorphism* $f : R \to R'$ such that $f \circ P_\omega = P'_\omega \circ f$ for $\omega \in \Omega$.

The following is the construction of the free Ω-operated algebra on a set X [Guo 2009]. Denote by $M(X)$ the free monoid generated by X. For any set Y and $\omega \in \Omega$, let $\lfloor Y \rfloor_\omega$ denote the set $\{\lfloor y \rfloor_\omega \mid y \in Y\}$. So $\lfloor Y \rfloor_\omega$ is a disjoint copy of Y. Assume the sets $\lfloor Y \rfloor_\omega$ to be disjoint with each other when ω varies in Ω. We now use induction to define a direct system $\{\mathfrak{M}_{\mathrm{n}} := \mathfrak{M}_{\mathrm{n}}(\Omega, X), i_{n,n+1} : \mathfrak{M}_{\mathrm{n}} \to \mathfrak{M}_{n+1}\}_{n \geq 0}$ of free monoids. We first define

$$\mathfrak{M}_0 := M(X) \quad \text{and} \quad \mathfrak{M}_1 := M\Big(X \sqcup \Big(\bigsqcup\nolimits_{\omega \in \Omega} \lfloor \mathfrak{M}_0 \rfloor_\omega\Big)\Big),$$

with $i_{0,1}$ being the inclusion

$$i_{0,1} : \mathfrak{M}_0 = M(X) \hookrightarrow \mathfrak{M}_1 = M\Big(X \sqcup \Big(\bigsqcup\nolimits_{\omega \in \Omega} \lfloor \mathfrak{M}_0 \rfloor_\omega\Big)\Big).$$

Inductively assume that $\mathfrak{M}_{n-1}$ has been defined for $n \geq 2$, with the inclusion

$$i_{n-2,n-1} : \mathfrak{M}_{n-2} \to \mathfrak{M}_{n-1}. \tag{11}$$

We then define

$$\mathfrak{M}_n := M\Big(X \sqcup \Big(\bigsqcup\nolimits_{\omega \in \Omega} \lfloor \mathfrak{M}_{n-1} \rfloor_\omega\Big)\Big).$$

The inclusion in (11) induces the inclusion

$$\lfloor \mathfrak{M}_{n-2} \rfloor_\omega \to \lfloor \mathfrak{M}_{n-1} \rfloor_\omega \quad \text{for each } \omega \in \Omega,$$

and generates an inclusion of free monoids

$$i_{n-1,n} : \mathfrak{M}_{n-1} = M\Big(X \sqcup \Big(\bigsqcup\nolimits_{\omega \in \Omega} \lfloor \mathfrak{M}_{n-2} \rfloor_\omega\Big)\Big) \hookrightarrow M\Big(X \sqcup \Big(\bigsqcup\nolimits_{\omega \in \Omega} \lfloor \mathfrak{M}_{n-1} \rfloor_\omega\Big)\Big) = \mathfrak{M}_n.$$

This completes the inductive construction of the direct systems. Then we define the direct limit of monoids

$$\mathfrak{M}(\Omega,\ X) := \varinjlim \mathfrak{M}_n = \bigcup_{n \geq 0} \mathfrak{M}_n$$

with identity 1. Elements of $\mathfrak{M}_n \backslash \mathfrak{M}_{n-1}$ are said to have *depth n*. For each $u \in \mathfrak{M}(\Omega,\ X)$ with $u \neq 1$, we may write u as a product $u_1 \cdots u_k$ uniquely for some k with $u_i \in X \sqcup \big(\bigsqcup_{\omega \in \Omega} \lfloor \mathfrak{M}(\Omega,\ X) \rfloor_\omega\big)$ for $1 \leq i \leq k$. We call k the *breadth* of u and denote it by $|u|$. If $u = 1$, we define $|u| := 0$.

Proposition 3.2 [Guo 2009]. *Let X be a set and Ω be a nonempty set. Let $j_X : X \to \mathfrak{M}(\Omega,\ X)$ be the natural embedding. Then the pair*

$$(\boldsymbol{k}\mathfrak{M}(\Omega,\ X),\ \{\lfloor\ \rfloor_\omega \mid \omega \in \Omega\})$$

with the j_X is the free Ω-operated algebra on X.

3B. ***Composition-diamond lemma for free Ω-operated algebras.*** In this subsection, we recall the composition-diamond lemma for free Ω-operated algebras [Bokut et al. 2010; Gao and Guo 2017].

Definition 3.3. Let X be a set and Ω a nonempty set, $\star \notin X$, and $X^\star := X \sqcup \{\star\}$.

(a) By a *$\star$-bracketed word* on X, we mean any bracketed word in $\mathfrak{M}(\Omega, X)^\star := \mathfrak{M}(\Omega, X^\star)$ with exactly one occurrence of $\star$, counting multiplicities.

(b) For $q \in \mathfrak{M}(\Omega, X)^\star$ and $u \in \mathfrak{M}(\Omega, X)$, we define $q|_u$ to be the bracketed word on X obtained by replacing the symbol $\star$ in q by u.

(c) For $q \in \mathfrak{M}(\Omega, X)^\star$ and $s = \Sigma_i c_i q|_{u_i} \in \boldsymbol{k}\mathfrak{M}(\Omega, X)$, where $c_i \in \boldsymbol{k}$ and $u_i \in \mathfrak{M}(\Omega, X)$, we define

$$q|_s := \Sigma_i c_i q|_{u_i}.$$

Definition 3.4. Let X be a set and Ω a nonempty set. *A monomial order* on $\mathfrak{M}(\Omega, X)$ is a *well order* $\leq$ on $\mathfrak{M}(\Omega, X)$ such that

$$u < v \Rightarrow q\,|_u < q\,|_v \quad \text{for all } u, v \in \mathfrak{M}(\Omega, X) \text{ and all } q \in \mathfrak{M}(\Omega, X)^\star.$$

Here, as usual, we denote $u < v$ if $u \leq v$ but $u \neq v$.

Definition 3.5. Let X be a set and Ω a nonempty set. Let $\leq$ be a monomial order on $\mathfrak{M}(\Omega, X)$, and let $f \in \boldsymbol{k}\mathfrak{M}(\Omega, X)$.

(a) If $f \notin \boldsymbol{k}$, the *leading monomial* of f, denoted by $\bar{f}$, is the largest monomial appearing in f. The *leading coefficient* of f, denoted by c_f, is the coefficient of $\bar{f}$ in f.

(b) If $f \in \boldsymbol{k}$ (including the case $f = 0$), we define the *leading monomial* of f to be $1_{\boldsymbol{k}}$ and the *leading coefficient* of f to be $c_f = f$.

(c) We call f *monic* with respect to $\leq$ if $c_f = 1_{\boldsymbol{k}}$. We call a subset $S \subseteq \boldsymbol{k}\mathfrak{M}(\Omega, X)$ *monic with respect to* $\leq$ if each element of S is monic.

Definition 3.6. Let X be a set and Ω a nonempty set. Let $\leq$ be a monomial order on $\mathfrak{M}(\Omega, X)$ and let $f, g \in \boldsymbol{k}\mathfrak{M}(\Omega, X)$ be distinct and monic with respect to $\leq$. Then we define two kinds of compositions.

(a) If there exist $u, v, w \in \mathfrak{M}(\Omega, X)$ such that $w = \bar{f}u = v\bar{g}$ with $\max\{|\bar{f}|, |\bar{g}|\} < w < |\bar{f}| + |\bar{g}|$, we call

$$(f, g)_w := fu - vg$$

the *intersection composition of f and g with respect to* (u, v).

(b) If there exist $q \in \mathfrak{M}(\Omega, X)^\star$ and $w \in \mathfrak{M}(\Omega, X)$ such that $w = \bar{f} = q|_{\bar{g}}$, we call

$$(f, g)_w := f - q|_g$$

the *including composition of f and g with respect to q*.

Definition 3.7. Let X be a set and Ω a nonempty set. Let $\leq$ be a monomial order on $\mathfrak{M}(\Omega, X)$, $S \subseteq \boldsymbol{k}\mathfrak{M}(\Omega, X)$ monic and $w \in \mathfrak{M}(\Omega, X)$.

(a) For $u, v \in \boldsymbol{k}\mathfrak{M}(\Omega, X)$, we say u and v are *congruent modulo* (S, w) and denote this by
$$u \equiv v \mod (S, w)$$
if $u - v = \Sigma_i c_i q_i|_{s_i}$, where $c_i \in \boldsymbol{k}$, $q_i \in \mathfrak{M}(\Omega, X)^\star$, $s_i \in S$ and $q_i|_{\bar{s}_i} < w$.

(b) Let $f, g \in \boldsymbol{k}\mathfrak{M}(\Omega, X)$. Then the composition $(f, g)_w$ is called *trivial modulo* (S, w) if
$$(f, g)_w \equiv 0 \mod (S, w).$$

Definition 3.8. Let X be a set and Ω a nonempty set. Let $\leq$ be a monomial order on $\mathfrak{M}(\Omega, X)$ and $S \subseteq \boldsymbol{k}\mathfrak{M}(\Omega, X)$ monic. S is called a *Gröbner–Shirshov basis with respect to* $\leq$ if for all pairs $f, g \in S$ with $f \neq g$, every intersection composition $(f, g)_w$ and every including composition $(f, g)_w$ are trivial modulo (S, w).

The following result is the well-known composition-diamond lemma for Ω-operated algebras.

Theorem 3.9. *Let X be a set and Ω a nonempty set, and let $\leq$ be a monomial order on $\mathfrak{M}(\Omega, X)$. Let $S \subseteq \boldsymbol{k}\mathfrak{M}(\Omega, X)$ be monic with respect to $\leq$. Then the following statements are equivalent*:

(a) *S is a Gröbner–Shirshov basis in $\boldsymbol{k}\mathfrak{M}(\Omega, X)$.*

(b) *For every nonzero f in the operated ideal* $\mathrm{Id}(S)$, *we have $\bar{f} = q|_{\bar{s}}$ for some $q \in \mathfrak{M}(\Omega, X)^\star$ and $s \in S$.*

(c) $\boldsymbol{k}\mathfrak{M}(\Omega, X) = \boldsymbol{k}\,\mathrm{Irr}(S) \oplus \mathrm{Id}(S)$ *where*
$$\mathrm{Irr}(S) = \mathfrak{M}(\Omega, X) \backslash \{q|_{\bar{s}} \mid q \in \mathfrak{M}(\Omega, X)^\star, s \in S\},$$
and $\mathrm{Irr}(S)$ *is a $\boldsymbol{k}$-basis of* $\boldsymbol{k}\mathfrak{M}(\Omega, X)/\mathrm{Id}(S)$.

3C. ***Gröbner–Shirshov bases for free noncommutative Rota–Baxter family algebras.*** This subsection is devoted to a Gröbner–Shirshov basis for a free Rota–Baxter family algebra on a set. Then by the composition-diamond lemma for Ω-operated algebras, we obtain a linear basis of a free Rota–Baxter family algebra. We use the monomial order $\leq_{\text{lex}}$ on $\mathfrak{M}(\Omega, X)$ given in [Bokut et al. 2010].

Theorem 3.10. *Let X be a set and Ω a semigroup. With respect to the monomial order $\leq_{\text{lex}}$ on $\mathfrak{M}(\Omega, X)$, the set*

$$S = \Big\{ \lfloor x \rfloor_\alpha \lfloor y \rfloor_\beta - \lfloor \lfloor x \rfloor_\alpha y \rfloor_{\alpha\beta} - \lfloor x \lfloor y \rfloor_\beta \rfloor_{\alpha\beta} - \lambda \lfloor xy \rfloor_{\alpha\beta} \mid \alpha, \beta \in \Omega, x, y \in \mathfrak{M}(\Omega, X) \Big\}. \tag{12}$$

is a Gröbner–Shirshov basis in $\boldsymbol{k}\mathfrak{M}(\Omega, X)$.

Proof. The ambiguities of all possible composition of Ω-polynomials in S are

$$\begin{aligned} w_1 &= \lfloor x\rfloor_\alpha \lfloor y\rfloor_\beta \lfloor z\rfloor_\gamma, \\ w_2 &= \lfloor q|_{\lfloor x\rfloor_\beta \lfloor y\rfloor_\gamma}\rfloor_\alpha \lfloor z\rfloor_\delta, \\ w_3 &= \lfloor z\rfloor_\delta \lfloor q|_{\lfloor x\rfloor_\beta \lfloor y\rfloor_\gamma}\rfloor_\alpha, \end{aligned}$$

where $x, y, z \in \mathfrak{M}(\Omega,\ X),\ \alpha, \beta, \gamma, \delta \in \Omega$ and $q \in \mathfrak{M}(\Omega,\ X)^\star$. We have the following three cases to consider. Denote

$$f_{\alpha,\beta}(x,y) := \lfloor x\rfloor_\alpha \lfloor y\rfloor_\beta - \lfloor \lfloor x\rfloor_\alpha y\rfloor_{\alpha\beta} - \lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta} - \lambda\lfloor xy\rfloor_{\alpha\beta}$$
$$\text{for } \alpha,\beta \in \Omega \text{ and } x, y \in \mathfrak{M}(\Omega,\ X).$$

Case 1. $w_1 = \lfloor x\rfloor_\alpha \lfloor y\rfloor_\beta \lfloor z\rfloor_\gamma$. In this case, write

$$\begin{aligned} f_{\alpha,\beta}(x,y) &= \lfloor x\rfloor_\alpha \lfloor y\rfloor_\beta - \lfloor \lfloor x\rfloor_\alpha y\rfloor_{\alpha\beta} - \lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta} - \lambda\lfloor xy\rfloor_{\alpha\beta}, \\ f_{\beta,\gamma}(y,z) &= \lfloor y\rfloor_\beta \lfloor z\rfloor_\gamma - \lfloor \lfloor y\rfloor_\beta z\rfloor_{\beta\gamma} - \lfloor y\lfloor z\rfloor_\gamma\rfloor_{\beta\gamma} - \lambda\lfloor yz\rfloor_{\beta\gamma}. \end{aligned}$$

Then with respect to $\leq_{\text{lex}}$,

$$\overline{f_{\alpha,\beta}(x,y)} = \lfloor x\rfloor_\alpha \lfloor y\rfloor_\beta \quad \text{and} \quad \overline{f_{\beta,\gamma}(y,z)} = \lfloor y\rfloor_\beta \lfloor z\rfloor_\gamma.$$

Thus we get the intersection composition:

$$\begin{aligned} &(f_{\alpha,\beta}(x,y),\, f_{\beta,\gamma}(y,z))_w \\ &\quad= f_{\alpha,\beta}(x,y)\lfloor z\rfloor_\gamma - \lfloor x\rfloor_\alpha f_{\beta,\gamma}(y,z) \\ &\quad= -\lfloor\lfloor x\rfloor_\alpha y\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma - \lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma - \lambda\lfloor xy\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma + \lfloor x\rfloor_\alpha\lfloor\lfloor y\rfloor_\beta z\rfloor_{\beta\gamma} \\ &\qquad\quad + \lfloor x\rfloor_\alpha\lfloor y\lfloor z\rfloor_\gamma\rfloor_{\beta\gamma} + \lambda\lfloor x\rfloor_\alpha\lfloor yz\rfloor_{\beta\gamma} \\ &\quad= -f_{\alpha\beta,\gamma}(\lfloor x\rfloor_\alpha y, z) - \lfloor\lfloor\lfloor x\rfloor_\alpha y\rfloor_{\alpha\beta} z\rfloor_{\alpha\beta\gamma} - \lfloor\lfloor x\rfloor_\alpha y\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} - \lambda\lfloor\lfloor x\rfloor_\alpha yz\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad - f_{\alpha\beta,\gamma}(x\lfloor y\rfloor_\beta, z) - \lfloor\lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta} z\rfloor_{\alpha\beta\gamma} - \lfloor x\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} - \lambda\lfloor x\lfloor y\rfloor_\beta z\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad - \lambda f_{\alpha\beta,\gamma}(xy, z) - \lambda\lfloor\lfloor xy\rfloor_{\alpha\beta} z\rfloor_{\alpha\beta\gamma} - \lambda\lfloor xy\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} - \lambda^2\lfloor xyz\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad + f_{\alpha,\beta\gamma}(x, \lfloor y\rfloor_\beta z) + \lfloor\lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta z\rfloor_{\alpha\beta\gamma} + \lfloor x\lfloor\lfloor y\rfloor_\beta z\rfloor_{\beta\gamma}\rfloor_{\alpha\beta\gamma} + \lambda\lfloor x\lfloor y\rfloor_\beta z\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad + f_{\alpha,\beta\gamma}(x, y\lfloor z\rfloor_\gamma) + \lfloor\lfloor x\rfloor_\alpha y\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} + \lfloor x\lfloor y\lfloor z\rfloor_\gamma\rfloor_{\beta\gamma}\rfloor_{\alpha\beta\gamma} + \lambda\lfloor xy\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad + \lambda f_{\alpha,\beta\gamma}(x, yz) + \lambda\lfloor\lfloor x\rfloor_\alpha yz\rfloor_{\alpha\beta\gamma} + \lambda\lfloor x\lfloor yz\rfloor_{\beta\gamma}\rfloor_{\alpha\beta\gamma} + \lambda^2\lfloor xyz\rfloor_{\alpha\beta\gamma} \\ &\quad= -f_{\alpha\beta,\gamma}(\lfloor x\rfloor_\alpha y, z) - f_{\alpha\beta,\gamma}(x\lfloor y\rfloor_\beta, z) - \lambda f_{\alpha\beta,\gamma}(xy, z) - \lfloor x f_{\beta,\gamma}(y,z)\rfloor_{\alpha\beta\gamma} \\ &\qquad\quad + f_{\alpha,\beta\gamma}(x, \lfloor y\rfloor_\beta z) + f_{\alpha,\beta\gamma}(x, y\lfloor z\rfloor_\gamma) + \lambda f_{\alpha,\beta\gamma}(x, yz) + \lfloor f_{\alpha,\beta}(x,y) z\rfloor_{\alpha\beta\gamma}, \end{aligned}$$

where

$$\begin{aligned}
\overline{f_{\alpha\beta,\,\gamma}(\lfloor x\rfloor_\alpha y,\,z)} &= \lfloor\lfloor x\rfloor_\alpha y\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{f_{\alpha\beta,\,\gamma}(x\lfloor y\rfloor_\beta,\,z)} &= \lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{f_{\alpha\beta,\gamma}(xy,\,z)} &= \lfloor xy\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{f_{\alpha,\,\beta\gamma}(x,\,\lfloor y\rfloor_\beta z)} &= \lfloor x\rfloor_\alpha\lfloor\lfloor y\rfloor_\beta z\rfloor_{\beta\gamma} <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{f_{\alpha,\,\beta\gamma}(x,\,y\lfloor z\rfloor_\gamma)} &= \lfloor x\rfloor_\alpha\lfloor y\lfloor z\rfloor_\gamma\rfloor_{\beta\gamma} <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{f_{\alpha,\,\beta\gamma}(x,\,yz)} &= \lfloor x\rfloor_\alpha\lfloor yz\rfloor_{\beta\gamma} <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,
\end{aligned}$$

and

$$\begin{aligned}
\overline{\lfloor xg_{\beta,\,\gamma}(y,z)\rfloor_{\alpha\beta\gamma}} &= \lfloor x\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma\rfloor_{\alpha\beta\gamma} <_{\mathrm{lex}} \lfloor x\lfloor y\rfloor_\beta\rfloor_{\alpha\beta}\lfloor z\rfloor_\gamma <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1,\\
\overline{\lfloor f_{\alpha,\,\beta}(x,y)z\rfloor_{\alpha\beta\gamma}} &= \lfloor\lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta z\rfloor_{\alpha\beta\gamma} <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor\lfloor y\rfloor_\beta z\rfloor_{\beta\gamma} <_{\mathrm{lex}} \lfloor x\rfloor_\alpha\lfloor y\rfloor_\beta\lfloor z\rfloor_\gamma = w_1.
\end{aligned}$$

So the intersection composition of $f_{\alpha,\,\beta}(x,\,y)$ and $f_{\beta,\,\gamma}(y,z)$ is trivial modulo $(S,\,w_1)$.

Case 2. $w_2 = \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta$. In this case, denote $p := \lfloor q\rfloor_\alpha\lfloor z\rfloor_\delta$ and

$$\begin{aligned}
f_{\beta,\,\gamma}(x,\,y) &= \lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma - \lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma} - \lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma} - \lambda\lfloor xy\rfloor_{\beta\gamma},\\
f_{\alpha,\,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma},\,z) &= \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta - \lfloor\lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha z\rfloor_{\alpha\delta}\\
&\quad - \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lambda\lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}z\rfloor_{\alpha\delta}.
\end{aligned}$$

Then with respect to $\leq_{\mathrm{lex}}$,

$$\overline{f_{\alpha,\,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma},\,z)} = \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta,\quad \overline{f_{\beta,\,\gamma}(x,\,y)} = \lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma,$$

and

$$\overline{f_{\alpha,\,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma},\,z)} = p|_{\overline{f_{\beta,\,\gamma(x,y)}}} = \lfloor q|_{\overline{f_{\beta,\,\gamma}(x,y)}}\rfloor_\alpha\lfloor z\rfloor_\delta = \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta.$$

We have the including composition:

$$\begin{aligned}
&(f_{\alpha,\,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma},\,z),\,f_{\beta,\,\gamma}(x,\,y))_{\omega_2}\\
&= f_{\alpha,\,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma},\,z) - p|_{f_{\beta,\,\gamma(x,y)}}\\
&= -\lfloor\lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha z\rfloor_{\alpha\delta} - \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lambda\lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}z\rfloor_{\alpha\delta}\\
&\quad + \lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta + \lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta + \lambda\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta
\end{aligned}$$

$$
\begin{aligned}
&= -\lfloor\lfloor q|_{f_{\beta,\gamma}(x,y)}\rfloor_\alpha z\rfloor_{\alpha\delta} - \lfloor\lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta} - \lfloor\lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta} - \lambda\lfloor\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta}\\
&\quad - \lfloor q|_{f_{\beta,\gamma}(x,y)}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lambda\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta}\\
&\quad - \lambda\lfloor q|_{f_{\beta,\gamma}(x,y)} z\rfloor_{\alpha\delta} - \lambda\lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta} - \lambda\lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta} - \lambda^2\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta}\\
&\quad + f_{\alpha,\delta}(q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}, z) + \lfloor\lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta} + \lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} + \lambda\lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta}\\
&\quad + f_{\alpha,\delta}(q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}, z) + \lfloor\lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta} + \lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} + \lambda\lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta}\\
&\quad + \lambda f_{\alpha,\delta}(q|_{\lfloor xy\rfloor_{\beta\gamma}}, z) + \lambda\lfloor\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\rfloor_\alpha z\rfloor_{\alpha\delta} + \lambda\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} + \lambda^2\lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}} z\rfloor_{\alpha\delta}\\
&= -\lfloor\lfloor q|_{f_{\beta,\gamma}(x,y)}\rfloor_\alpha z\rfloor_{\alpha\delta} - \lfloor q|_{f_{\beta,\gamma}(x,y)}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} - \lambda\lfloor q|_{f_{\beta,\gamma}(x,y)} z\rfloor_{\alpha\delta} + f_{\alpha,\delta}(q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}, z)\\
&\quad + f_{\alpha,\delta}(q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}, z) + \lambda f_{\alpha,\delta}(q|_{\lfloor xy\rfloor_{\beta\gamma}}, z),
\end{aligned}
$$

where

$$
\begin{aligned}
\overline{\lfloor\lfloor q|_{f_{\beta,\gamma}(x,y)}\rfloor_\alpha z\rfloor_{\alpha\delta}} &= \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma} z\rfloor_{\alpha\delta} <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2,\\
\overline{\lfloor q|_{f_{\beta,\gamma}(x,y)}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta}} &= \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\lfloor z\rfloor_\delta\rfloor_{\alpha\delta} <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2,\\
\overline{\lfloor q|_{f_{\beta,\gamma}(x,y)} z\rfloor_{\alpha\delta}} &= \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma} z\rfloor_{\alpha\delta} <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2,\\
\overline{f_{\alpha,\delta}(q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}, z)} &= \lfloor q|_{\lfloor x\lfloor y\rfloor_\gamma\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2,\\
\overline{f_{\alpha,\delta}(q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}, z)} &= \lfloor q|_{\lfloor\lfloor x\rfloor_\beta y\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2,\\
\overline{f_{\alpha,\delta}(q|_{\lfloor xy\rfloor_{\beta\gamma}}, z)} &= \lfloor q|_{\lfloor xy\rfloor_{\beta\gamma}}\rfloor_\alpha\lfloor z\rfloor_\delta <_{\mathrm{lex}} \lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha\lfloor z\rfloor_\delta = w_2.
\end{aligned}
$$

So the including composition of $f_{\alpha,\delta}(q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}, z)$ and $f_{\beta,\gamma}(x, y)$ is trivial modulo (S, w_2).

Case 3. $w_3 = \lfloor z\rfloor_\delta\lfloor q|_{\lfloor x\rfloor_\beta\lfloor y\rfloor_\gamma}\rfloor_\alpha$. This case is similar to Case 2. □

Now we are ready for our main result in this section.

Theorem 3.11. *Let S be as in* (12) *and* $\mathrm{Id}(S)$ *be the operated ideal generated by S in $\boldsymbol{k}\mathfrak{M}(\Omega, X)$. Then $\boldsymbol{k}\mathfrak{M}(\Omega, X) = \boldsymbol{k}\,\mathrm{Irr}(S) \oplus \mathrm{Id}(S)$ where*

$$\mathrm{Irr}(S) = \mathfrak{M}(\Omega, X)\backslash\{q|_{\bar{s}} \mid q \in \mathfrak{M}(\Omega, X)^\star, s \in S\},$$

and $\mathrm{Irr}(S)$ *is a $\boldsymbol{k}$-basis of the free Rota–Baxter family algebra $\boldsymbol{k}\mathfrak{M}(\Omega, X)/\mathrm{Id}(S)$ of weight λ.*

Proof. This follows from Theorems 3.9 and 3.10. □

4. Dendriform family algebras

It is well known that Rota–Baxter algebras induce (tri)dendriform algebras [Aguiar 2000a; Ebrahimi-Fard 2002]. In this section, we generalize this result to the case of Rota–Baxter family algebras. The free commutative (tri)dendriform family algebras are also constructed.

4A. ***Dendriform family algebras and tridendriform family algebras.*** Motivated by algebraic K-theory, Loday [1993] invented the concept of a dendriform algebra. Now we propose the concept of a dendriform family algebra, with an eye toward giving a connection from a Rota–Baxter family algebra of weight zero to it.

Definition 4.1. Let Ω be a semigroup. A *dendriform family algebra* is a $\boldsymbol{k}$-module D with a family of binary operations $\{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}$ such that for $x, y, z \in D$ and $\alpha, \beta \in \Omega$,

$$(x \prec_\alpha y) \prec_\beta z = x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z), \tag{13}$$

$$(x \succ_\alpha y) \prec_\beta z = x \succ_\alpha (y \prec_\beta z), \tag{14}$$

$$x \succ_\alpha (y \succ_\beta z) = (x \prec_\beta y + x \succ_\alpha y) \succ_{\alpha\beta} z. \tag{15}$$

Some years later, Loday and Ronco [2004] introduced the concept of a tridendriform algebra (previously also called a dendriform trialgebra) in the study of polytopes and Koszul duality. Similarly, we propose the following definition:

Definition 4.2. Let Ω be a semigroup. A *tridendriform family algebra* is a $\boldsymbol{k}$-module T equipped with a family of binary operations $\{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}$ and a binary operation $\cdot$ such that for $x, y, z \in T$ and $\alpha, \beta \in \Omega$,

$$(x \prec_\alpha y) \prec_\beta z = x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z + y \cdot z), \tag{16}$$

$$(x \succ_\alpha y) \prec_\beta z = x \succ_\alpha (y \prec_\beta z), \tag{17}$$

$$x \succ_\alpha (y \succ_\beta z) = (x \prec_\beta y + x \succ_\alpha y + x \cdot y) \succ_{\alpha\beta} z, \tag{18}$$

$$(x \succ_\alpha y) \cdot z = x \succ_\alpha (y \cdot z), \tag{19}$$

$$\begin{aligned} (x \prec_\alpha y) \cdot z &= x \cdot (y \succ_\alpha z), \\ (x \cdot y) \prec_\alpha z &= x \cdot (y \prec_\alpha z), \\ (x \cdot y) \cdot z &= x \cdot (y \cdot z). \end{aligned} \tag{20}$$

Remark 4.3. When Ω is taken to be a trivial semigroup, a dendriform (resp. tridendriform) family algebra is precisely a dendriform (resp. tridendriform) algebra.

4B. ***From Rota–Baxter family algebras to (tri)dendriform family algebras.*** It is well known that a Rota–Baxter algebra of weight zero (resp. weight λ) induces a dendriform (resp. tridendriform) algebra [Aguiar 2000a; Ebrahimi-Fard 2002]. Now we generalize this result.

Theorem 4.4. *Let Ω be a semigroup.*

(a) *A Rota–Baxter family algebra $(R, \{P_\omega \mid \omega \in \Omega\})$ of weight zero induces a dendriform family algebra $(R, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\})$, where*

$$x \prec_\omega y := x P_\omega(y), \quad x \succ_\omega y := P_\omega(x) y, \quad \textit{for } x, y \in R.$$

(b) *A Rota–Baxter family algebra* $(R, \{P_\omega \mid \omega \in \Omega\})$ *of weight* λ *induces a tridendriform family algebra* $(R, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$, *where*

$$x \prec_\omega y := x P_\omega(y), \quad x \succ_\omega y := P_\omega(x) y, \quad x \cdot y := \lambda x y, \quad \textit{for } x, y \in R.$$

Proof. (a) For $x, y, z \in R$ and $\alpha, \beta \in \Omega$,

$$\begin{aligned}
(x \prec_\alpha y) \prec_\beta z &= (x P_\alpha(y)) P_\beta(z) = x(P_\alpha(y) P_\beta(z)) = x P_{\alpha\beta}(P_\alpha(y) z + y P_\beta(z)) \\
&= x \prec_{\alpha\beta} (y \succ_\alpha z + y \prec_\beta z), \\
(x \succ_\alpha y) \prec_\beta z &= (P_\alpha(x) y) P_\beta(z) = P_\alpha(x)(y P_\beta(z)) = x \succ_\alpha (y \prec_\beta z), \\
x \succ_\alpha (y \succ_\beta z) &= P_\alpha(x)(P_\beta(y) z) = (P_\alpha(x) P_\beta(y)) z = P_{\alpha\beta}(x P_\beta(y) + P_\alpha(x) y) z \\
&= (x \prec_\beta y + x \succ_\alpha y) \succ_{\alpha\beta} z.
\end{aligned}$$

(b) For $x, y, z \in R$ and $\alpha, \beta \in \Omega$,

$$\begin{aligned}
(x \prec_\alpha y) \prec_\beta z &= (x P_\alpha(y)) P_\beta(z) = x(P_\alpha(y) P_\beta(z)) \\
&= x P_{\alpha\beta}(P_\alpha(y) z + y P_\beta(z) + \lambda y z) = x \prec_{\alpha\beta} (y \succ_\alpha z + y \prec_\beta z + y \cdot z), \\
(x \succ_\alpha y) \prec_\beta z &= (P_\alpha(x) y) \prec_\beta z = P_\alpha(x)(y P_\beta(z)) = x \succ_\alpha (y \prec_\beta z), \\
x \succ_\alpha (y \succ_\beta z) &= P_\alpha(x)(P_\beta(y) z) = (P_\alpha(x) P_\beta(y)) z \\
&= P_{\alpha\beta}(x P_\beta(y) + P_\alpha(x) y + \lambda x y) z = (x \prec_\beta y + x \succ_\alpha y + x \cdot y) \succ_{\alpha\beta} z,
\end{aligned}$$

$$\begin{aligned}
&(x \succ_\alpha y) \cdot z = (P_\alpha(x) y) \cdot z = \lambda P_\alpha(x) y z = P_\alpha(x)(\lambda y z) = P_\alpha(x)(y \cdot z) = x \succ_\alpha (y \cdot z), \\
&(x \prec_\alpha y) \cdot z = (x P_\alpha(y)) \cdot z = \lambda x P_\alpha(y) z = \lambda x (P_\alpha(y) z) = x \cdot (P_\alpha(y) z) = x \cdot (y \succ_\alpha z), \\
&(x \cdot y) \prec_\alpha z = (\lambda x y) \prec_\alpha z = \lambda x (y P_\alpha(z)) = x \cdot (y \prec_\alpha z), \\
&(x \cdot y) \cdot z = (\lambda x y) \cdot z = \lambda^2 (x y z) = \lambda x (\lambda y z) = x \cdot (y \cdot z).
\end{aligned}$$

This completes the proof. □

4C. *Commutative (tri)dendriform family algebras.* In this part, we first characterize some properties of commutative (tri)dendriform family algebras. Then we proceed to construct free commutative (tri)dendriform family algebras.

Definition 4.5. Let Ω be a semigroup.

(a) A dendriform family algebra $(D, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\})$ is called *commutative* if

$$x \succ_\omega y = y \prec_\omega x \quad \text{for } x, y \in D, \omega \in \Omega.$$

(b) A tridendriform family algebra $(T, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$ is called *commutative* if

$$x \succ_\omega y = y \prec_\omega x \ \text{ and } \ x \cdot y = y \cdot x \quad \text{for } x, y \in T,\ \omega \in \Omega.$$

The following results characterize commutative (tri)dendriform family algebras.

Proposition 4.6. (a) *Let Ω be a commutative semigroup. If*

$$(D, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\})$$

is a commutative dendriform family algebra, then

$$(x \prec_\alpha y) \prec_\beta z = x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y) \quad \textit{for } x, y, z \in D,\ \alpha, \beta \in \Omega. \tag{21}$$

(b) *Conversely, given $(D, \{\prec_\omega \mid \omega \in \Omega\})$ where D is a $\boldsymbol{k}$-module and $\{\prec_\omega \mid \omega \in \Omega\}$ is a family of binary operations on D satisfying* (21), *the pair*

$$(D, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}),$$

where $x \succ_\omega y := y \prec_\omega x$, is a commutative dendriform family algebra.

Proof. (a) Suppose $(D, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\})$ is a commutative dendriform family algebra. Then

$$\begin{aligned}(x \prec_\alpha y) \prec_\beta z &= x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z) && \text{(by (13))}\\ &= x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y).\end{aligned}$$

(b) It suffices to prove (13)–(15). First, by (21),

$$(x \prec_\alpha y) \prec_\beta z = x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y) = x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z).$$

Second,

$$\begin{aligned}&(x \succ_\alpha y) \prec_\beta z\\ &\quad= (y \prec_\alpha x) \prec_\beta z = y \prec_{\alpha\beta} (x \prec_\beta z + z \prec_\alpha x) && \text{(by (21))}\\ &\quad= y \prec_{\beta\alpha} (z \prec_\alpha x + x \prec_\beta z) && \text{(by } \Omega \text{ being a commutative semigroup)}\\ &\quad= (y \prec_\beta z) \prec_\alpha x && \text{(by (21))}\\ &\quad= x \succ_\alpha (y \prec_\beta z).\end{aligned}$$

Finally,

$$\begin{aligned}&x \succ_\alpha (y \succ_\beta z)\\ &\quad= (z \prec_\beta y) \prec_\alpha x = z \prec_{\beta\alpha} (y \prec_\alpha x + x \prec_\beta y) && \text{(by (21))}\\ &\quad= z \prec_{\beta\alpha} (x \prec_\beta y + x \succ_\alpha y)\\ &\quad= z \prec_{\alpha\beta} (x \prec_\beta y + x \succ_\alpha y) && \text{(by } \Omega \text{ being a commutative semigroup)}\\ &\quad= (x \prec_\beta y + x \succ_\alpha y) \succ_{\alpha\beta} z,\end{aligned}$$

as required. □

Proposition 4.7. *Let Ω be a commutative semigroup.*

(a) *If* $(T, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$ *is a commutative tridendriform family algebra, then*

$$x \cdot y = y \cdot x, \tag{22}$$

$$(x \prec_\alpha y) \prec_\beta z = x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y + y \cdot z), \tag{23}$$

$$(x \cdot y) \prec_\alpha z = x \cdot (y \prec_\alpha z), \tag{24}$$

$$(x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \text{for } \alpha, \beta \in \Omega,\ x, y, z \in T. \tag{25}$$

(b) *On the contrary, given* $(T, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$ *where* T *is a* $\boldsymbol{k}$*-module and* $\prec_\omega$ *with* $\omega \in \Omega$, $\cdot$ *are binary operations on* T *satisfying* (22)–(25), *then the triple* $(T, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$, *where* $x \succ_\omega y := y \prec_\omega x$, *is a commutative tridendriform family algebra.*

Proof. (a) It suffices to prove (23), which follows from

$$\begin{aligned}
(x \prec_\alpha y) \prec_\beta z &= x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z + y \cdot z) && \text{(by (16))}\\
&= x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y + y \cdot z).
\end{aligned}$$

(b) We only need to prove (16)–(20). We have

$$\begin{aligned}
(x \prec_\alpha y) \prec_\beta z &= x \prec_{\alpha\beta} (y \prec_\beta z + z \prec_\alpha y + y \cdot z) && \text{(by (23))}\\
&= x \prec_{\alpha\beta} (y \prec_\beta z + y \succ_\alpha z + y \cdot z).\\
(x \succ_\alpha y) \prec_\beta z &= (y \prec_\alpha x) \prec_\beta z = y \prec_{\alpha\beta} (x \prec_\beta z + z \prec_\alpha x + x \cdot z) && \text{(by (23))}\\
&= y \prec_{\beta\alpha} (z \prec_\alpha x + x \prec_\beta z + z \cdot x)\\
& && \text{(by } \Omega \text{ being a commutative semigroup)}\\
&= (y \prec_\beta z) \prec_\alpha x && \text{(by (23))}\\
&= x \succ_\alpha (y \prec_\beta z).\\
x \succ_\alpha (y \succ_\beta z) &= (z \prec_\beta y) \prec_\alpha x = z \prec_{\beta\alpha} (y \prec_\alpha x + x \prec_\beta y + y \cdot x) && \text{(by (23))}\\
&= z \prec_{\beta\alpha} (x \prec_\beta y + x \succ_\alpha y + x \cdot y)\\
&= z \prec_{\alpha\beta} (x \prec_\beta y + x \succ_\alpha y + x \cdot y)\\
& && \text{(by } \Omega \text{ being a commutative semigroup)}\\
&= (x \prec_\beta y + x \succ_\alpha y + x \cdot y) \succ_{\alpha\beta} z.\\
(x \succ_\alpha y) \cdot z &= (y \prec_\alpha x) \cdot z = z \cdot (y \prec_\alpha x) && \text{(by (22))}\\
&= (z \cdot y) \prec_\alpha x && \text{(by (24))}\\
&= x \succ_\alpha (z \cdot y) = x \succ_\alpha (y \cdot z).\\
(x \prec_\alpha y) \cdot z &= (y \succ_\alpha x) \cdot z\\
&= y \succ_\alpha (x \cdot z) = y \succ_\alpha (z \cdot x) = (y \succ_\alpha z) \cdot x = x \cdot (y \succ_\alpha z).
\end{aligned}$$

□

4D. ***Free commutative (tri)dendriform family algebras.*** In this subsection, we are in a position to construct respectively free commutative dendriform and tridendriform family algebras.

Definition 4.8. Let A be a commutative algebra and Ω a semigroup. A free tridendriform family algebra on A is a tridendriform family algebra $(T, \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$ together with the natural algebra homomorphism $j_A : A \to T$ that satisfies the following universal property: for any tridendriform family algebra $(T', \{\prec_{T',\omega}, \succ_{T',\omega} \mid \omega \in \Omega\}, \cdot_{T'})$ and algebra homomorphism $f : A \to T'$, there is a unique tridendriform family algebra morphism $\bar{f} : T \to T'$ such that $f = \bar{f} \circ j_A$.

Let $(Ш_\Omega(A), \diamond_\Omega, \{P_{A,\omega} \mid \omega \in \Omega\})$ be the free commutative Rota–Baxter family algebra of weight 1 on an algebra A.

Theorem 4.9. *Let A be a commutative algebra and Ω a semigroup.*

(a) *Define*

$$\mathfrak{a} \prec_\omega \mathfrak{b} := \mathfrak{a} \diamond_\Omega (1 \otimes_\omega \mathfrak{b}),\ \mathfrak{a} \succ_\omega \mathfrak{b} := (1 \otimes_\omega \mathfrak{a}) \diamond_\Omega \mathfrak{b} \quad \textit{for each } \omega \in \Omega.$$

Then $(Ш_\Omega(A), \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\})$, together with the natural algebra homomorphism $i_A : A \to Ш_\Omega(A)$, is the free commutative dendriform family algebra on A.

(b) *Define*

$$\mathfrak{a} \prec_\omega \mathfrak{b} := \mathfrak{a} \diamond_\Omega (1 \otimes_\omega \mathfrak{b}),\ \mathfrak{a} \succ_\omega \mathfrak{b} := (1 \otimes_\omega \mathfrak{a}) \diamond_\Omega \mathfrak{b},\ \mathfrak{a} \cdot \mathfrak{b} := \mathfrak{a} \diamond_\Omega \mathfrak{b} \quad \textit{for each } \omega \in \Omega.$$

Then $(Ш_\Omega(A), \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$, together with the natural algebra homomorphism $i_A : A \to Ш_\Omega(A)$, is the free commutative tridendriform family algebra on A.

Proof. We only prove (b), as the proof of (a) is similar and easier.

(b) Note that in $(Ш_\Omega(A), \diamond_\Omega, \{P_{A,\omega} \mid \omega \in \Omega\})$, we have

$$\begin{aligned} \mathfrak{a} \succ_\alpha \mathfrak{b} &:= (1 \otimes_\alpha \mathfrak{a}) \diamond_\Omega \mathfrak{b} = P_{A,\alpha}(\mathfrak{a}) \diamond_\Omega \mathfrak{b}, \\ \mathfrak{a} \prec_\beta \mathfrak{b} &:= \mathfrak{a} \diamond_\Omega (1 \otimes_\beta \mathfrak{b}) = \mathfrak{a} \diamond_\Omega P_{A,\beta}(\mathfrak{b}) \end{aligned} \tag{26}$$

for $\alpha, \beta \in \Omega$ and $\mathfrak{a}, \mathfrak{b} \in Ш_\Omega(A)$. By Theorem 4.4, $(Ш_\Omega(A), \{\prec_\omega, \succ_\omega \mid \omega \in \Omega\}, \cdot)$ is a tridendriform family algebra. Further, it is commutative, as $Ш_\Omega(A)$ is commutative. We are left to show the universal property. For this, let $(T, \{\prec_{T,\omega}, \succ_{T,\omega} \mid \omega \in \Omega\}, \cdot_T)$ be a tridendriform family algebra and let $f : A \to T$ be an algebra homomorphism.

(existence) Define a linear map $\bar{f} : Ш_\Omega(A) \to T$ as follows. For $\mathfrak{a} = a_0 \otimes_\alpha \mathfrak{a}' \in A^{\otimes(m+1)}$ with $a_0 \in A$ and $\mathfrak{a}' \in A^{\otimes m}$, we define $\bar{f}(\mathfrak{a})$ by induction on $m \geq 0$. For the initial step of $m = 0$, we have $\mathfrak{a} = a_0$ and define

$$\bar{f}(a) := \bar{f}(a_0) := f(a_0).$$

For the induction step of $m \geq 1$, we define

$$\bar{f}(\mathfrak{a}) := \bar{f}(a_0 \otimes_\alpha \mathfrak{a}') := f(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}'). \tag{27}$$

We now prove that $\bar{f}$ is a morphism of commutative tridendriform family algebras:

$$\bar{f}(\mathfrak{a} \prec_{A,\beta} \mathfrak{b}) = \bar{f}(\mathfrak{a}) \prec_{T,\beta} \bar{f}(\mathfrak{b}) \ \text{ and } \ \bar{f}(\mathfrak{a}\cdot\mathfrak{b}) = \bar{f}(\mathfrak{a}) \cdot_T \bar{f}(\mathfrak{b}), \tag{28}$$

for $\mathfrak{a} = a_0 \otimes_\alpha \mathfrak{a}' \in A^{\otimes_\Omega (m+1)}$ and $\mathfrak{b} = b_0 \otimes_\beta \mathfrak{b}' \in A^{\otimes_\Omega (n+1)}$ with $m, n \geq 0$.

We proceed to prove (28) by induction on $m+n \geq 0$. When $m+n=0$, $\mathfrak{a} = a_0$ and $\mathfrak{b} = b_0$ are in A. So by (27),

$$\bar{f}(\mathfrak{a} \prec_{A,\beta} \mathfrak{b}) = \bar{f}(a_0 \otimes_\beta b_0) = f(a_0) \prec_{T,\beta} f(b_0) = \bar{f}(\mathfrak{a}) \prec_{T,\beta} \bar{f}(\mathfrak{b}),$$
$$\bar{f}(\mathfrak{a}\cdot\mathfrak{b}) = \bar{f}(a_0 b_0) = f(a_0 b_0) = f(a_0)\cdot_T f(b_0) = \bar{f}(\mathfrak{a}) \cdot_T \bar{f}(\mathfrak{b}).$$

Assume that (28) has been proved when $m+n=k$ for a $k \geq 0$, and consider the case of $m+n=k+1$. Then

$$\begin{aligned}
\bar{f}(\mathfrak{a} \prec_{A,\beta} \mathfrak{b}) &= \bar{f}(\mathfrak{a} \diamond_\Omega (1 \otimes_\beta \mathfrak{b})) \\
&= \bar{f}((a_0 \otimes_\alpha \mathfrak{a}') \diamond_\Omega (1 \otimes_\beta \mathfrak{b})) \\
&= \bar{f}(a_0 \otimes_{\alpha\beta} (\mathfrak{a}' \diamond_\Omega (1 \otimes_\beta \mathfrak{b}) + (1 \otimes_\alpha \mathfrak{a}') \diamond_\Omega \mathfrak{b} + \mathfrak{a}' \diamond_\Omega \mathfrak{b}) \\
&= \bar{f}(a_0) \prec_{T,\alpha\beta} \bar{f}(\mathfrak{a}' \prec_{A,\beta} \mathfrak{b} + \mathfrak{b} \prec_{A,\alpha} \mathfrak{a}' + \mathfrak{a}'\cdot\mathfrak{b}) && \text{(by (27))} \\
&= \bar{f}(a_0) \prec_{T,\alpha\beta} (\bar{f}(\mathfrak{a}') \prec_{T,\beta} \bar{f}(\mathfrak{b}) + \bar{f}(\mathfrak{b}) \prec_{T,\alpha} \bar{f}(\mathfrak{a}') + \bar{f}(\mathfrak{a}') \cdot_T \bar{f}(\mathfrak{b})) \\
&&& \text{(by the induction hypothesis)} \\
&= (\bar{f}(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}')) \prec_{T,\beta} \bar{f}(\mathfrak{b}) && \text{(by (23))} \\
&= \bar{f}(\mathfrak{a}) \prec_{T,\beta} \bar{f}(\mathfrak{b}) && \text{(by (27))}
\end{aligned}$$

and

$$\begin{aligned}
\bar{f}(\mathfrak{a}\cdot\mathfrak{b}) &= \bar{f}((a_0 \otimes_\alpha \mathfrak{a}') \diamond_\Omega (b_0 \otimes_\beta \mathfrak{b}')) \\
&= \bar{f}(a_0 b_0 \otimes_{\alpha\beta} (\mathfrak{a}' \diamond_\Omega (1 \otimes_\beta \mathfrak{b}') + (1 \otimes_\alpha \mathfrak{a}') \diamond_\Omega \mathfrak{b}' + \mathfrak{a}' \diamond_\Omega \mathfrak{b}')) && \text{(by (4))} \\
&= \bar{f}(a_0 b_0) \prec_{T,\alpha\beta} \bar{f}(\mathfrak{a}' \prec_{A,\beta} \mathfrak{b}' + \mathfrak{b}' \prec_{A,\alpha} \mathfrak{a}' + \mathfrak{a}' \cdot \mathfrak{b}') \\
&= f(a_0 b_0) \prec_{T,\alpha\beta} (\bar{f}(\mathfrak{a}') \prec_{T,\beta} \bar{f}(\mathfrak{b}') + \bar{f}(\mathfrak{b}') \prec_{T,\alpha} \bar{f}(\mathfrak{a}') + \bar{f}(\mathfrak{a}') \cdot_T \bar{f}(\mathfrak{b}')) \\
&&& \text{(by induction)} \\
&= (f(a_0)\cdot f(b_0)) \prec_{T,\alpha\beta} (\bar{f}(\mathfrak{a}') \prec_{T,\beta} \bar{f}(\mathfrak{b}') + \bar{f}(\mathfrak{b}') \prec_{T,\alpha} \bar{f}(\mathfrak{a}') + \bar{f}(\mathfrak{a}') \cdot_T \bar{f}(\mathfrak{b}')) \\
&= ((f(a_0)\cdot f(b_0)) \prec_{T,\alpha} \bar{f}(\mathfrak{a}')) \prec_{T,\beta} \bar{f}(\mathfrak{b}') && \text{(by (23))} \\
&= ((f(b_0)\cdot f(a_0)) \prec_{T,\alpha} \bar{f}(\mathfrak{a}')) \prec_{T,\beta} \bar{f}(\mathfrak{b}') && \text{(by (22))} \\
&= (f(b_0)\cdot (f(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}'))) \prec_{T,\beta} \bar{f}(\mathfrak{b}') && \text{(by (24))} \\
&= ((f(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}'))\cdot f(b_0)) \prec_{T,\beta} \bar{f}(\mathfrak{b}') && \text{(by (22))} \\
&= (f(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}'))\cdot (f(b_0) \prec_{T,\beta} \bar{f}(\mathfrak{b}')) && \text{(by (24))} \\
&= \bar{f}(\mathfrak{a}) \cdot_T \bar{f}(\mathfrak{b}) && \text{(by (27))}
\end{aligned}$$

(uniqueness) Suppose that $\bar{f} : Ш_\Omega(A) \to T$ is a required morphism of commutative tridendriform family algebras. Then for $\mathfrak{a} = a_0 \otimes_\alpha \mathfrak{a}' \in A^{\otimes_\Omega(m+1)}$,

$$\bar{f}(\mathfrak{a}) = \bar{f}(a_0 \otimes_\alpha \mathfrak{a}') = \bar{f}(a_0 \prec_{A,\alpha} \mathfrak{a}') = \bar{f}(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}') = f(a_0) \prec_{T,\alpha} \bar{f}(\mathfrak{a}'),$$

which determines $\bar{f}$ uniquely by recursion. This completes the proof of (b). □

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11771191 and 11861051), Fundamental Research Funds for the Central Universities (Grant No. lzujbky-2017-162), the Natural Science Foundation of Gansu Province (Grant No. 17JR5RA175).

References

[Aguiar 2000a] M. Aguiar, "Infinitesimal Hopf algebras", pp. 1–29 in *New trends in Hopf algebra theory* (La Falda, Argentina, 1999), edited by N. Andruskiewitsch et al., Contemp. Math. **267**, Amer. Math. Soc., Providence, RI, 2000. MR Zbl

[Aguiar 2000b] M. Aguiar, "Pre-Poisson algebras", *Lett. Math. Phys.* **54**:4 (2000), 263–277. MR Zbl

[Aguiar 2001] M. Aguiar, "On the associative analog of Lie bialgebras", *J. Algebra* **244**:2 (2001), 492–532. MR Zbl

[An and Bai 2008] H. An and C. Bai, "From Rota–Baxter algebras to pre-Lie algebras", *J. Phys. A* **41**:1 (2008), art. id. 015201. MR Zbl

[Atkinson 1963] F. V. Atkinson, "Some aspects of Baxter's functional equation", *J. Math. Anal. Appl.* **7** (1963), 1–30. MR Zbl

[Bai et al. 2012] C. Bai, L. Guo, and X. Ni, "$\mathcal{O}$-operators on associative algebras and associative Yang–Baxter equations", *Pacific J. Math.* **256**:2 (2012), 257–289. MR Zbl

[Bai et al. 2013] C. Bai, O. Bellier, L. Guo, and X. Ni, "Splitting of operations, Manin products, and Rota–Baxter operators", *Int. Math. Res. Not.* **2013**:3 (2013), 485–524. MR Zbl

[Baxter 1960] G. Baxter, "An analytic problem whose solution follows from a simple algebraic identity", *Pacific J. Math.* **10** (1960), 731–742. MR Zbl

[Bokut and Chen 2014] L. A. Bokut and Y. Chen, "Gröbner–Shirshov bases and their calculation", *Bull. Math. Sci.* **4**:3 (2014), 325–395. MR Zbl

[Bokut et al. 2010] L. A. Bokut, Y. Chen, and J. Qiu, "Gröbner–Shirshov bases for associative algebras with multiple operators and free Rota–Baxter algebras", *J. Pure Appl. Algebra* **214**:1 (2010), 89–100. MR Zbl

[Cartier 1972] P. Cartier, "On the structure of free Baxter algebras", *Adv. Math.* **9** (1972), 253–265. MR Zbl

[Connes and Kreimer 1998] A. Connes and D. Kreimer, "Hopf algebras, renormalization and noncommutative geometry", *Comm. Math. Phys.* **199**:1 (1998), 203–242. MR Zbl

[Connes and Kreimer 2000] A. Connes and D. Kreimer, "Renormalization in quantum field theory and the Riemann–Hilbert problem, I: The Hopf algebra structure of graphs and the main theorem", *Comm. Math. Phys.* **210**:1 (2000), 249–273. MR Zbl

[Ebrahimi-Fard 2002] K. Ebrahimi-Fard, "Loday-type algebras and the Rota–Baxter relation", *Lett. Math. Phys.* **61**:2 (2002), 139–147. MR Zbl

[Ebrahimi-Fard and Guo 2008] K. Ebrahimi-Fard and L. Guo, "Rota–Baxter algebras and dendriform algebras", *J. Pure Appl. Algebra* **212**:2 (2008), 320–339. MR Zbl

[Ebrahimi-Fard et al. 2004] K. Ebrahimi-Fard, L. Guo, and D. Kreimer, "Spitzer's identity and the algebraic Birkhoff decomposition in pQFT", *J. Phys. A* **37**:45 (2004), 11037–11052. MR Zbl

[Ebrahimi-Fard et al. 2007] K. Ebrahimi-Fard, J. M. Gracia-Bondía, and F. Patras, "A Lie theoretic approach to renormalization", *Comm. Math. Phys.* **276**:2 (2007), 519–549. MR Zbl

[Ebrahimi-Fard et al. 2008] K. Ebrahimi-Fard, D. Manchon, and F. Patras, "New identities in dendriform algebras", *J. Algebra* **320**:2 (2008), 708–727. MR

[Frabetti 1998] A. Frabetti, "Leibniz homology of dialgebras of matrices", *J. Pure Appl. Algebra* **129**:2 (1998), 123–141. MR Zbl

[Gao and Guo 2017] X. Gao and L. Guo, "Rota's classification problem, rewriting systems and Gröbner–Shirshov bases", *J. Algebra* **470** (2017), 219–253. MR Zbl

[Gao et al. 2014] X. Gao, L. Guo, and S. Zheng, "Construction of free commutative integro-differential algebras by the method of Gröbner–Shirshov bases", *J. Algebra Appl.* **13**:5 (2014), art. id. 1350160. MR Zbl

[Gao et al. 2015] X. Gao, L. Guo, and M. Rosenkranz, "Free integro-differential algebras and Gröbner–Shirshov bases", *J. Algebra* **442** (2015), 354–396. MR Zbl

[Guo 2009] L. Guo, "Operated semigroups, Motzkin paths and rooted trees", *J. Algebraic Combin.* **29**:1 (2009), 35–62. MR Zbl

[Guo 2012] L. Guo, *An introduction to Rota–Baxter algebra*, Surveys of Modern Math. **4**, Int. Press, Somerville, MA, 2012. MR Zbl

[Guo and Keigher 2000a] L. Guo and W. Keigher, "Baxter algebras and shuffle products", *Adv. Math.* **150**:1 (2000), 117–149. MR Zbl

[Guo and Keigher 2000b] L. Guo and W. Keigher, "On free Baxter algebras: completions and the internal construction", *Adv. Math.* **151**:1 (2000), 101–127. MR Zbl

[Guo and Keigher 2008] L. Guo and W. Keigher, "On differential Rota–Baxter algebras", *J. Pure Appl. Algebra* **212**:3 (2008), 522–540. MR Zbl

[Guo et al. 2013] L. Guo, W. Y. Sit, and R. Zhang, "Differential type operators and Gröbner–Shirshov bases", *J. Symbolic Comput.* **52** (2013), 97–123. MR Zbl

[Guo et al. 2017] L. Guo, S. Paycha, and B. Zhang, "Algebraic Birkhoff factorization and the Euler–Maclaurin formula on cones", *Duke Math. J.* **166**:3 (2017), 537–571. MR Zbl

[Kurosh 1960] A. G. Kurosh, "Free sums of multiple operator algebras", *Sibirsk. Mat. Z.* **1** (1960), 62–70; correction on 638. In Russian. MR Zbl

[Loday 1993] J.-L. Loday, "Une version non commutative des algèbres de Lie: les algèbres de Leibniz", *Enseign. Math.* (2) **39**:3-4 (1993), 269–293. MR Zbl

[Loday 2002] J.-L. Loday, "Arithmetree", *J. Algebra* **258**:1 (2002), 275–309. MR Zbl

[Loday 2004] J.-L. Loday, "Scindement d'associativité et algèbres de Hopf", pp. 155–172 in *Actes des Journées Mathématiques à la Mémoire de Jean Leray* (Nantes, 2002), edited by L. Guillopé and D. Robert, Sémin. Congr. **9**, Soc. Math. France, Paris, 2004. MR Zbl arXiv

[Loday and Ronco 2004] J.-L. Loday and M. Ronco, "Trialgebras and families of polytopes", pp. 369–398 in *Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic K-theory*, edited by P. Goerss and S. Priddy, Contemp. Math. **346**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Rota 1969a] G.-C. Rota, "Baxter algebras and combinatorial identities, I", *Bull. Amer. Math. Soc.* **75** (1969), 325–329. MR Zbl

[Rota 1969b] G.-C. Rota, "Baxter algebras and combinatorial identities, II", *Bull. Amer. Math. Soc.* **75** (1969), 330–334. MR Zbl

[Rota 1995] G.-C. Rota, "Baxter operators: an introduction", pp. 504–512 in *Gian-Carlo Rota on combinatorics*, edited by J. P. S. Kung, Birkhäuser, Boston, 1995. MR

[Spitzer 1956] F. Spitzer, "A combinatorial lemma and its application to probability theory", *Trans. Amer. Math. Soc.* **82** (1956), 323–339. MR Zbl

[Zhang et al. 2016] T. Zhang, X. Gao, and L. Guo, "Hopf algebras of rooted forests, cocyles, and free Rota–Baxter algebras", *J. Math. Phys.* **57**:10 (2016), art. id. 101701. MR Zbl

Received July 30, 2018. Revised October 27, 2018.

YUANYUAN ZHANG
SCHOOL OF MATHEMATICS AND STATISTICS
LANZHOU UNIVERSITY
LANZHOU
CHINA
zhangyy17@lzu.edu.cn

XING GAO
SCHOOL OF MATHEMATICS AND STATISTICS
KEY LABORATORY OF APPLIED MATHEMATICS AND COMPLEX SYSTEMS
LANZHOU UNIVERSITY
GANSU
CHINA
gaoxing@lzu.edu.cn

CONTENTS

Volume 301, no. 1 and no. 2

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 301 No. 2 August 2019

0030-8730(201908)301:2;1-G