# Pacific Journal of Mathematics

Volume 267 No. 1 January 2014

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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



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

The subscription price for 2014 is US $410/year for the electronic version, and $535/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.1

# NUMERICAL STUDY OF UNBOUNDED CAPILLARY SURFACES

YASUNORI AOKI AND HANS DE STERCK

**Unbounded capillary surfaces in domains with a sharp corner or a cusp are studied. It is shown how numerical study using a proposed computational methodology leads to two new conjectures for open problems on the asymptotic behavior of capillary surfaces in domains with a cusp. The numerical methodology contains two simple but important ingredients, a change of variable and a change of coordinates, which are inspired by known asymptotic approximations for unbounded capillary surfaces. These ingredients are combined with the finite volume element or Galerkin finite element methods. Extensive numerical tests show that the proposed computational methodology leads to a global approximation method for singular solutions of the Laplace–Young equation that recovers the proper asymptotic behavior at the singular point, is more accurate and has better convergence properties than numerical methods considered for singular capillary surfaces before. Using this computational methodology, two open problems on the asymptotic behavior of capillary surfaces in domains with a cusp are studied numerically, leading to two conjectures that may guide future analytical work on these open problems.**

## 1. Introduction

The mathematical analysis of unbounded capillary surfaces is most often done by asymptotic analysis (see [Concus and Finn 1970; 1974; 1989; Miersemann 1993; King et al. 1999; Norbury et al. 2005; Scholz 2001; 2004; Aoki 2007; Aoki and Siegel 2012]). Results for unbounded capillary surfaces in domains with sharp corners have been known for many years, and recent work of Aoki and Siegel [2012] on singular capillary surfaces in domains with a cusp fills almost all the gaps that still existed for the cusp case, though a few open problems remain. Since asymptotic analysis is a local analysis, asymptotic approximations are valid only in a sufficiently small domain near the singularity. It is also not easy to determine the precise region of validity of the asymptotic analysis results. In applications, global

*MSC2010:* 35J25, 65N30, 76B45, 35J75.

*Keywords:* singularity, asymptotic analysis, nonlinear elliptic PDE, Laplace–Young equation, finite element method.

approximations for the singular capillary surfaces that are valid in the whole domain are desirable, and such approximations cannot be provided by asymptotic analysis.

In this paper, our aim is to construct globally valid approximations of singular capillary surfaces which exhibit the proper asymptotic behavior at the singular point while also being valid away from the singularity. We do so by introducing a computational methodology for singular capillary surfaces. A second aim of this paper is to investigate two open problems on asymptotic behavior of capillary surfaces in domains with a cusp. We investigate these open problems using the proposed numerical methodology, which leads to two conjectures that may guide future analytical work on these open problems.

Our computational methodology starts from the finite volume element method (FVEM) [Bank and Rose 1987; Aoki and De Sterck 2011] or the Galerkin finite element method (FEM) [Strang and Fix 1973; Brenner and Scott 1994]. However, it is widely known (see [Grisvard 1985; Strang and Fix 1973; Aoki and De Sterck 2011]) that a lack of smoothness in the solution can spoil the accuracy of approximations of finite element type; hence it can be expected that standard finite element approximations cannot accurately approximate the unbounded singularity. There are about a half dozen published papers on numerical solutions of the Laplace–Young equation [Nigro et al. 2000; Hornung and Mittelmann 1990; Polevikov 2004; Polevikov 1999; Scott et al. 2005]. However, with the exception of the paper by Scott et al. [2005] they do not consider unbounded singular solutions. Scott et al. use the finite volume element method to approximate solutions of the Laplace–Young equation, and one of their model problems is a corner problem with unbounded singularity. Our proposed methodology enhances their approach in two important ways, leading to much more accurate and informative results, as shown in our numerical results section.

Instead of directly approximating the solution with a standard finite element expansion, our idea is to incorporate knowledge obtained from asymptotic analysis into the finite element approximation, in order to avoid inaccuracies introduced by the singularity. Roughly speaking, we first change the variable based on the known asymptotic order of the solution so that the new unknown function is bounded. (Though it is bounded, it can still be discontinuous at the location of the original singularity.) Inspired by knowledge of the leading-order term of the asymptotic series solution, we change the coordinate system so that the unknown function is smooth with respect to the new coordinate variables. We then finally approximate the smooth bounded new unknown function with respect to the new coordinate variables, using the finite volume element method or the Galerkin finite element method.

We verify the accuracy of this numerical methodology by comparing the numerical solution with known asymptotic series approximations, and by conducting numerical convergence studies. We first show that the numerical solutions we obtain have the proper singular behavior by comparing numerical solutions of the Laplace–Young equation with known asymptotic series approximations. Then we conduct numerical convergence studies to show that the numerical approximation is a globally valid approximation. In order to conduct numerical convergence studies, we need model problems with known closed-form solutions. Though there is no known unbounded closed-form solution of the Laplace–Young equation, a few closed-form solutions are known for the steep slope approximation of the Laplace–Young equation [King et al. 1999; Aoki 2007] (we shall refer to this PDE as the asymptotic Laplace–Young equation). It is known that these solutions have the same asymptotic behavior as the solution of the original problem, so we conduct the convergence study using the asymptotic Laplace–Young equation.

Using the proposed numerical methodology for computing singular solutions of the Laplace–Young equation, we investigate a few open problems of singular behavior of the Laplace–Young equation in a cusp domain. Aoki and Siegel [2012] considered the solution behavior for all possible cusp domains, attempting to generalize the results of Scholz [2004]. However, there are still a few cases that remain open. Using our computational methodology, we numerically investigate these cases and make conjectures based on the numerical approximations.

The paper proceeds as follows. In Section 2 we describe the Laplace–Young boundary value problem of interest and its asymptotic variant in domains with a sharp corner or a cusp. Section 3 describes the proposed numerical methodology for computing accurate global numerical approximations of unbounded capillary surfaces in these types of domains, and Section 4 gives extensive numerical results verifying the accuracy and convergence of the numerical methods. Section 5 presents conjectures for two open cases on asymptotic behavior of capillary surfaces in a domain with a cusp, motivated by numerical solutions for these open cases using the proposed numerical methodology. Finally, conclusions are formulated in Section 6.

## 2. The boundary value problem

In this section we first formulate the Laplace–Young boundary value problem, and describe the asymptotic behavior of its solutions in domains with a corner or a cusp and the function spaces these solutions belong to. We state some open problems on asymptotic behavior for a domain with an osculatory cusp and a cusp with infinite curvature, and define model problems that will be used in numerical tests. We then describe the asymptotic Laplace–Young equation and its known closed-form solutions on domains with a corner or a cusp, which are used to formulate additional numerical model problems.

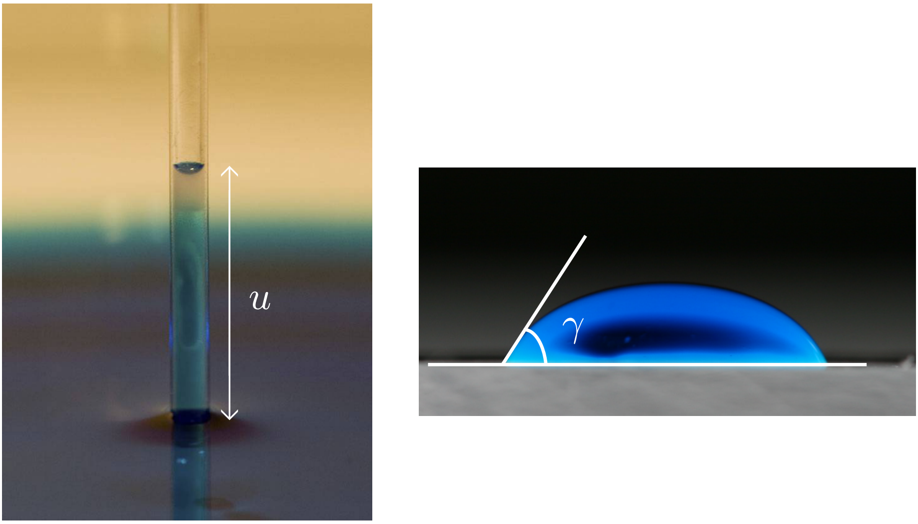



**Figure 1.** Two photos of capillarity experiments, indicating the capillary surface height $u$ and the contact angle $\gamma$.

**2.1.** ***Laplace–Young boundary value problem.*** This problem originates from observations of Laplace in 1806 and Young in 1805 that "the height of the liquid is proportional to its mean curvature" and "the angle of the contact between the solid and liquid only depends on their material." Gauss showed in 1830 that the Laplace–Young PDE is in fact the Euler–Lagrange equation of the surface energy functional. Thus the solution of the Laplace–Young boundary value problem gives the shape of the liquid surface that minimizes the surface energy, in a nonvanishing downward gravity field, and hence the Laplace–Young boundary value problem is a mathematical model for a liquid surface at equilibrium when the gravity is present. We refer the reader to Section 1.4 of a book by Finn [1986] for detailed discussion of the derivation of the Laplace–Young boundary value problem. Figure 1 shows photos of capillarity experiments, indicating the capillary surface height $u$ and the contact angle $\gamma$.

Let $\Omega$ be an unbounded open domain as in Figure 2 with boundaries $\partial\Omega_1$ and $\partial\Omega_2$ described by functions $f_1(x)$ and $f_2(x)$, and let $u \in C^2(\Omega)$ be the height of the capillary surface that satisfies the following boundary value problem (the Laplace–Young boundary value problem) on this domain:

$$\nabla \cdot T(u) = \kappa u \quad \text{in } \Omega, \tag{1}$$

$$\vec{\nu}_1 \cdot T(u) = \cos\gamma_1 \quad \text{on } \partial\Omega_1,$$

$$\vec{\nu}_2 \cdot T(u) = \cos\gamma_2 \quad \text{on } \partial\Omega_2, \tag{2}$$

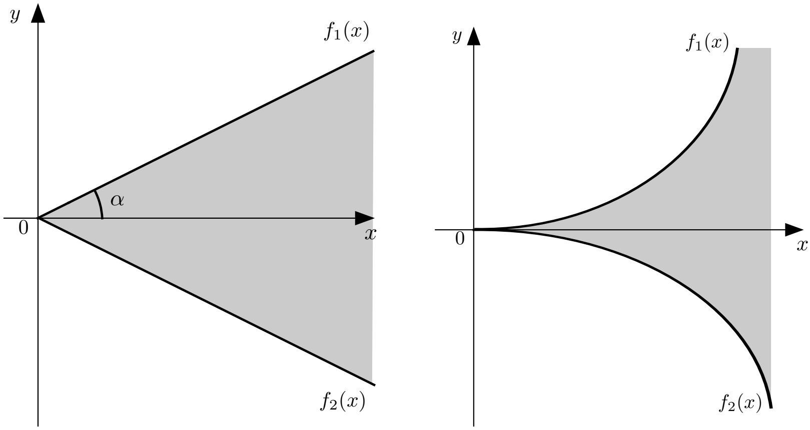



**Figure 2.** Unbounded domains. Left: a corner domain. Right: a cusp domain.

with

$$\begin{aligned} \kappa &: \text{capillarity constant (we assume } \kappa > 0\text{)}, \\ \vec{\nu}_1,\ \vec{\nu}_2 &: \text{exterior unit normal vectors on the boundaries } \partial\Omega_1 \text{ and } \partial\Omega_2, \\ \gamma_1,\ \gamma_2 &: \text{contact angles} \end{aligned}$$

and

$$T(u) = \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}.$$

Note that the capillarity constant $\kappa$ can be normalized by rescaling $x$, $y$ and $u$ when $\kappa > 0$. In the following sections we let $\kappa = 1$. The open domain $\Omega$ and boundaries $\partial\Omega_1$ and $\partial\Omega_2$ are defined more specifically as follows:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : 0 < x,\ f_2(x) < y < f_1(x)\}, \tag{3}$$

$$\partial\Omega_1 = \{(x, y) \in \mathbb{R}^2 : 0 < x,\ y = f_1(x)\}, \tag{4}$$

$$\partial\Omega_2 = \{(x, y) \in \mathbb{R}^2 : 0 < x,\ y = f_2(x)\}, \tag{5}$$

with

$$\begin{gathered} f_1(x),\ f_2(x) \in C^3(0, \infty), \\ f_1(x) > f_2(x) \quad \text{for } x > 0, \\ \lim_{x\to 0^+} f_1(x) = \lim_{x\to 0^+} f_2(x) = 0, \\ \lim_{x\to 0^+} f_1'(x) \neq \infty \neq \lim_{x\to 0^+} f_2'(x). \end{gathered} \tag{6}$$

For simplicity of discussion we focus on the two specific types of domains depicted in Figure 2: a corner domain, defined as

$$f_1(x) = x \tan\alpha \quad \text{and} \quad f_2(x) = -x \tan\alpha, \quad \text{where } 0 < \alpha < \pi/2,$$

and a cusp domain, defined as

$$\lim_{x\to 0^+} f_1'(x) = 0 \quad \text{and} \quad \lim_{x\to 0^+} f_2'(x) = 0. \tag{7}$$

**2.1.1.** *Asymptotic behavior.* It is known that the solution of the Laplace–Young boundary value problem in a corner domain is unbounded at (0, 0) if $\gamma_1+\gamma_2+2\alpha<\pi$ (see [Finn 1986]). Also, it can be shown that if $\gamma_1+\gamma_2\neq\pi$ the solution of the boundary value problem in a cusp domain is unbounded at (0, 0) (see [Scholz 2004; Aoki and Siegel 2012]). In addition, the following asymptotic behaviors are known.

Corner domain with $\gamma_1+\gamma_2+2\alpha<\pi$ (see [Concus and Finn 1970; Miersemann 1993] for a proof): If $\gamma_1=\gamma_2=\gamma$ and $\gamma+\alpha<\pi/2$, then the solution of the boundary value problem in the corner domain has the following asymptotic behavior:

$$u(r,\theta) = \frac{\cos\theta - \sqrt{k^2-\sin^2\theta}}{kr} + O(r^3) \quad \text{as } r \to 0, \tag{8}$$

where

$$(r,\theta) : \text{polar coordinate variables,}$$
$$k = \frac{\sin\alpha}{\cos\gamma}.$$

More formally, we can write that there exist constants $r_o$ and $M$ such that

$$\left| u - \frac{\cos\theta - \sqrt{k^2-\sin^2\theta}}{kr} \right| < Mr^3 \quad \text{for } 0 < r < r_o.$$

This gives the following bounds for the solution $u$:

$$\frac{\cos\theta - \sqrt{k^2-\sin^2\theta}}{kr} - Mr^3 < u < \frac{\cos\theta - \sqrt{k^2-\sin^2\theta}}{kr} + Mr^3 \qquad \text{for } 0 < r < r_0. \tag{9}$$

The proof for the asymptotic relation (8) only provides the existence of these two constants and does not give any estimate of their size. Thus, even though (8) shows that the asymptotic approximation becomes more and more accurate as we get closer to the singularity, it does not give any quantitative description of the approximation error.

Also, it is easy to show from (9) that there exist positive constants $M^+$, $M^-$, and $x_o$ such that

$$\frac{M^-}{f_1(x)-f_2(x)} < u < \frac{M^+}{f_1(x)-f_2(x)} \quad \text{for } 0 < x < x_o. \tag{10}$$

Cusp domain with $\gamma_1+\gamma_2 \neq \pi$ (see [Aoki and Siegel 2012] for a proof): An unbounded solution of the boundary value problem in a cusp domain has the asymptotic behavior

$$u(x,y) = \frac{\cos\gamma_1+\cos\gamma_2}{f_1(x)-f_2(x)} + O\left(\frac{f_1'(x)-f_2'(x)}{f_1(x)-f_2(x)}\right) \quad \text{as } x \to 0^+ \tag{11}$$

if $\gamma_1+\gamma_2 \neq \pi$ and the boundary functions $f_1(x)$ and $f_2(x)$ satisfy the asymptotic relations

$$\begin{aligned} & f_1(x)-f_2(x) = o(f_1'(x)-f_2'(x)), && (12)\\ & \frac{f_1''(x)-f_2''(x)}{f_1(x)-f_2(x)} = \alpha\frac{(f_1'(x)-f_2'(x))^2}{(f_1(x)-f_2(x))^2} + o\left(\frac{(f_1'(x)-f_2'(x))^2}{(f_1(x)-f_2(x))^2}\right), \\ & \frac{f_1'''(x)-f_2'''(x)}{f_1'(x)-f_2'(x)} = O\left(\frac{(f_1'(x)-f_2'(x))^2}{(f_1(x)-f_2(x))^2}\right), \\ & f_1'(x)+f_2'(x) = \delta(f_1'(x)-f_2'(x)) + o(f_1'(x)-f_2'(x)), && (13)\\ & f_1''(x)+f_2''(x) = O(f_1''(x)-f_2''(x)) && (14) \end{aligned}$$

as $x \to 0$, where $\alpha, \delta \in \mathbb{R}$.

Note that most boundary functions forming cusp domains satisfy the asymptotic conditions (12)–(14). One known exception is when the boundary functions forming a cusp are osculatory at the cusp. Curves are said to be osculatory if they intersect and share the tangent line and the osculating circle at the intersection point. Again it follows from (11) that there exist positive constants $M^+$, $M^-$, and $x_o$ such that

$$\frac{M^-}{f_1(x)-f_2(x)} < u < \frac{M^+}{f_1(x)-f_2(x)} \quad \text{for } 0 < x < x_o. \tag{15}$$

**2.1.2.** *Open problems.* To the authors' knowledge there are two major open problems in the solution behavior of the Laplace–Young equation in a domain with a cusp. We summarize these open problems.

Problem 1: Osculatory cusp with nonsupplementary contact angles ($\gamma_1+\gamma_2 \neq \pi$): An osculatory cusp is a cusp formed by two osculating curves. It is known that the solution is unbounded when $\gamma_1+\gamma_2 \neq \pi$, but the asymptotic expansion from the previous section does not apply in the osculatory cusp case and remains an open problem. For example, the two boundary functions

$$f_1(x) = x^2 + x^3 \quad \text{and} \quad f_2(x) = x^2 - x^3$$

form an osculatory cusp at the origin. The asymptotic orders of the sum and the difference of these boundary functions are different (i.e., $f_1(x) - f_2(x) = O(x^3)$ while $f_1(x) + f_2(x) = O(x^2)$ as $x \to 0$). Hence these choices of $f_1(x)$ and $f_2(x)$ do not satisfy the asymptotic relations (13)–(14), so that the leading-order asymptotic behavior of the solution at this cusp is unknown. The main reason why the proof for the leading-order asymptotic behavior could not be constructed for the osculatory cusp case is that the second-order term of the formal asymptotic series could not be found (see [Aoki and Siegel 2012] for details).

Problem 2: Infinite curvature cusp with supplementary contact angles ($\gamma_1 + \gamma_2 = \pi$): As was noted before, the solution of the Laplace–Young equation in a cusp domain is unbounded if $\gamma_1 + \gamma_2 \neq \pi$, but it is not necessarily true that the solution is bounded if $\gamma_1 + \gamma_2 = \pi$. Aoki and Siegel [2012] have shown that the solution is bounded if $\gamma_1 + \gamma_2 = \pi$ and the curvatures of the boundary functions are finite (i.e., $\lim_{x\to 0} f_1''(x) \neq \infty$ and $\lim_{x\to 0} f_2''(x) \neq \infty$). However, the nature of the solution for the case where the curvatures of one or both boundary functions are infinite is not known (e.g., $f_1(x) = x^{3/2}$ and $f_2(x) = -2x^{3/2}$).

**2.1.3.** *Model Problems 1 and 2.* For the numerical experiments to be reported on below we consider the following model problems (henceforth, MPs).

Consider bounded open domains $\Omega$ as depicted in Figure 3. Let $u \in C^2(\Omega)$ be the height of the capillary surface that satisfies the following boundary value problem:

$$
\begin{aligned}
\nabla \cdot T(u) &= u && \text{in } \Omega, \\
\vec{\nu}_1 \cdot T(u) &= \cos\gamma_1 && \text{on } \partial\Omega_1, \\
\vec{\nu}_2 \cdot T(u) &= \cos\gamma_2 && \text{on } \partial\Omega_2, \\
\vec{\nu}_3 \cdot T(u) &= 0 && \text{on } \partial\Omega_3,
\end{aligned}
$$

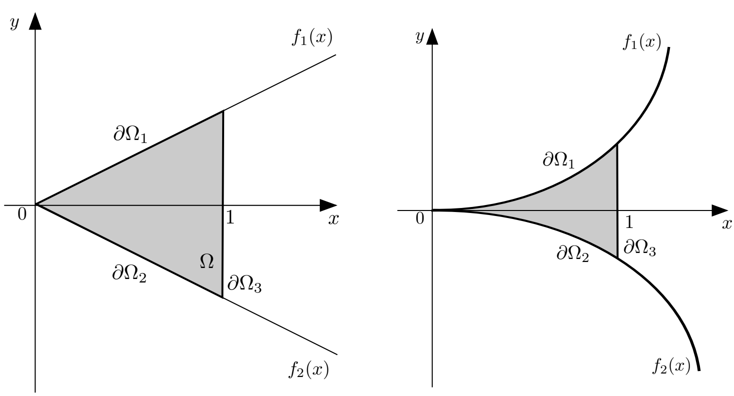



**Figure 3.** Computational domains for model problems with a corner and a cusp at (0,0). Left: problems 1 and 3. Right: problems 2 and 4.

with

$\vec{\nu}_1, \vec{\nu}_2, \vec{\nu}_3$ : exterior unit normal vectors on the boundaries $\partial\Omega_1$, $\partial\Omega_2$ and $\partial\Omega_3$,

$\gamma_1, \gamma_2$ : contact angles.

The bounded open domain $\Omega$ and boundaries $\partial\Omega_{1,2,3}$ are defined more specifically as follows:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1,\ f_2(x) < y < f_1(x)\}, \tag{16}$$

$$\partial\Omega_1 = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1,\ y = f_1(x)\},$$

$$\partial\Omega_2 = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1,\ y = f_2(x)\},$$

$$\partial\Omega_3 = \{(x, y) \in \mathbb{R}^2 : x = 1,\ f_2(1) < y < f_1(1)\}. \tag{17}$$

The boundary functions and the contact angles are chosen for each model problem as tabulated in Tables 1 and 2. Although there are cases in which the behavior of bounded solutions in a corner domain has special interest (e.g., [Lancaster 2010]), in this paper we focus our attention on unbounded solutions in a corner domain.

These model problems are chosen so that the singularity may only occur at the corner or cusp at the origin, although there are three nonsmooth points on the boundary of the domain $\Omega$. Following immediately from the regularity result of Simon [1980], this implies that the solutions $u$ of these model problems are differentiable up to the boundary except at the origin, i.e., $u \in C^1(\overline{\Omega}\backslash\{(0, 0)\})$. Also, the asymptotic behavior of the solution at the origin is known to be as stated in (8) for MP 1 ($\gamma_1 + \gamma_2 + 2\alpha < \pi$, unbounded), and as in (11) for MPs 2a ($\gamma_1 + \gamma_2 \neq \pi$, unbounded) and 2c-1 and 2c-3 ($\gamma_1 + \gamma_2 \neq \pi$, unbounded). The asymptotic expansions for MPs 2b (osculatory cusp with nonsupplementary contact angles) are an open problem (but it is known that the solutions are unbounded). The asymptotic behavior of the solution of MP 2c-2 (infinite curvature cusp with supplementary contact angles) is an open problem.

**2.1.4.** *Solution function spaces.* It is interesting to discuss the function spaces where the solutions of MP 1 and MPs 2a, 2c-1, and 2c-3 reside.

| Problem | $f_1(x)$ | $f_2(x)$ | $\gamma_1$ | $\gamma_2$ | |
|---|---|---|---|---|---|
| 1-1 | $x\tan(\pi/7)$ | $-x\tan(\pi/7)$ | $\pi/6$ | $\pi/6$ | Corner (unbounded) |
| 1-2 | $x\tan(\pi/7)$ | $-x\tan(\pi/7)$ | $\pi/4$ | $\pi/4$ | Corner (unbounded) |
| 1-3 | $x\tan(\pi/7)$ | $-x\tan(\pi/7)$ | $\pi/3$ | $\pi/3$ | Corner (unbounded) |

**Table 1.** Model Problem 1: Laplace–Young equation in a domain with a corner. All three variants have $\alpha = \pi/7$ and $\gamma_1+\gamma_2+2\alpha < \pi$, resulting in solutions that are unbounded at (0,0).

| Problem | $f_1(x)$ | $f_2(x)$ | $\gamma_1$ | $\gamma_2$ | cusp | unbd? | open? |
|---|---|---|---|---|---|---|---|
| 2a-1 | $x^2/6$ | $-x^3/8$ | $\pi/6$ | $\pi/3$ | R | yes | no |
| 2a-2 | $x^3/6$ | $-x^3/8$ | $\pi/3$ | $\pi/4$ | R | yes | no |
| 2a-3 | $x^5/6$ | $-x^4/8$ | $\pi/3$ | $\pi/4$ | R | yes | no |
| 2b-1 | $(x^2+x^3)/6$ | $(x^2-\frac{3}{4}x^3)/6$ | $\pi/3$ | $\pi/4$ | O | yes | yes |
| 2b-2 | $(3x^2+x^3)/6$ | $(3x^2-\frac{3}{4}x^3)/6$ | $\pi/3$ | $\pi/4$ | O | yes | yes |
| 2b-3 | $x^{3/2}+x^3/6$ | $x^{3/2}-\frac{3}{4}x^3/6$ | $\pi/3$ | $\pi/4$ | O | yes | yes |
| 2c-1 | $x^{3/2}/6$ | $x^{3/2}/8$ | $5\pi/6-\pi/180$ | $\pi/6$ | IC | yes | no |
| 2c-2 | $x^{3/2}/6$ | $x^{3/2}/8$ | $5\pi/6$ | $\pi/6$ | IC | ? | yes |
| 2c-3 | $x^{3/2}/6$ | $x^{3/2}/8$ | $5\pi/6+\pi/180$ | $\pi/6$ | IC | yes | no |

**Table 2.** Model Problem 2: Laplace–Young equation in a domain with a cusp. Variants 2a, 2c-1, and 2c-3 have unbounded solutions at (0,0) and known asymptotic expansions. Variants 2b also have unbounded solutions at (0,0), but asymptotic expansions are unknown and remain an open problem. The asymptotic behavior of variant 2c-2 at (0,0) is an open problem. Key for the last three columns: R = regular; O = osculatory; IC = infinite curvature; unbd = unbounded; open = open problem.

**Proposition 2.1.** *For any fixed $p$ with $1 \le p < \infty$, the solutions of MP 1, MPs 2a and MPs 2c-1 and 2c-3 are in the $L_p(\Omega)$ function space if and only if the following integral is finite for any $\epsilon$ in the interval $(0, 1]$:*

$$\int_0^{\epsilon} \frac{1}{(f_1(x)-f_2(x))^{p-1}}\,\mathrm{d}x. \tag{18}$$

*Proof.* We first note that, for the case of MP 1 and MP 2a, the comparison principle (see [Finn 1986]) gives that $u > 0$. Also recall that there exist positive constants $M^+$, $M^-$, and $x_o$ such that

$$\frac{M^-}{f_1(x)-f_2(x)} < u < \frac{M^+}{f_1(x)-f_2(x)} \quad \text{for } 0 < x < x_o.$$

We now bound the integral $\int_\Omega |u|^p\,\mathrm{d}A$ from above:

$$\begin{aligned}\int_\Omega |u|^p\,\mathrm{d}A &= \int_\Omega u^p\,\mathrm{d}A \quad (\text{since } u > 0)\\ &= \int_{x=0}^{1}\int_{f_2(x)}^{f_1(x)} u^p\,\mathrm{d}y\,\mathrm{d}x = \int_{x=0}^{x_o}\int_{f_2(x)}^{f_1(x)} u^p\,\mathrm{d}y\,\mathrm{d}x + \int_{x=x_o}^{1}\int_{f_2(x)}^{f_1(x)} u^p\,\mathrm{d}y\,\mathrm{d}x\\ &\le \int_{x=0}^{x_o}\int_{f_2(x)}^{f_1(x)} u^p\,\mathrm{d}y\,\mathrm{d}x + \int_{x=x_o}^{1}\int_{f_2(x)}^{f_1(x)} \max_{x_o<x<1} u^p\,\mathrm{d}y\,\mathrm{d}x\end{aligned}$$

$$< \int_{x=0}^{x_o} \int_{f_2(x)}^{f_1(x)} \frac{(M^+)^p}{(f_1(x)-f_2(x))^p} \, \mathrm{d}y \, \mathrm{d}x + \max_{x_o<x<1} (u)^p \int_{x=x_o}^{1} \int_{f_2(x)}^{f_1(x)} 1 \, \mathrm{d}y \, \mathrm{d}x.$$

This last sum can also be written as

$$(M^+)^p \int_0^{x_o} \frac{1}{(f_1(x) - f_2(x))^{p-1}} \, \mathrm{d}x + \max_{x_o<x<1} (u)^p \int_{x=x_o}^{1} (f_1(x) - f_2(x)) \, \mathrm{d}x. \tag{19}$$

If $p$ is chosen so that the integral (18) is finite for any $\epsilon \in (0, 1]$, then the first term of (19) is finite. Also, noting that $u$ is bounded away from the origin ($u \in C^1(\overline{\Omega}\backslash\{0\})$) and that the domains $\Omega$ for the model problems are bounded domains, the second term of (19) is also finite. Thus if $p$ is chosen so that integral (18) is finite then the solution of MPs 1 and 2a are in the $L_p(\Omega)$ function space. We now bound the integral $\int_\Omega |u|^p \, \mathrm{d}A$ from below:

$$\begin{aligned}
\int_\Omega |u|^p \, \mathrm{d}A &= \int_{x=0}^{x_o} \int_{f_2(x)}^{f_1(x)} u^p \, \mathrm{d}y \, \mathrm{d}x + \int_{x=x_o}^{1} \int_{f_2(x)}^{f_1(x)} u^p \, \mathrm{d}y \, \mathrm{d}x \\
&> \int_{x=0}^{x_o} \int_{f_2(x)}^{f_1(x)} u^p \, \mathrm{d}y \, \mathrm{d}x > \int_{x=0}^{x_o} \int_{f_2(x)}^{f_1(x)} \frac{(M^-)^p}{(f_1(x) - f_2(x))^p} \, \mathrm{d}y \, \mathrm{d}x \\
&= (M^-)^p \int_{x=0}^{x_o} \frac{1}{(f_1(x) - f_2(x))^{p-1}} \, \mathrm{d}x.
\end{aligned}$$

This gives that if $p$ is chosen so that integral (18) is not finite, then the solutions of MPs 1 and 2a are not in the $L_p$ function space.

The proof for MPs 2c-1 and 2c-3 is slightly more complicated because $u > 0$ does not hold. A sketch of the proof for these cases is as follows. Since $u \in C^1(\overline{\Omega}\backslash\{0\})$, there is a neighborhood $\Omega_s$ of the singularity where the solution is either positive or negative. Using the approach above, it can be shown that $u \in L_p(\Omega_s)$ if and only if integral (18) is finite, which is equivalent to $u \in L_p(\Omega)$ since $u$ is bounded away from the singularity. □

**Corollary 2.1.** (A) *The solution of MP* 1 *is in the* $L_{2-\delta}$ *function space where* $\delta > 0$.

(B) *The solution is in the* $L_{1+1/2-\delta}$ *function space for MP* 2*a*-1, *is in the* $L_{1+1/3-\delta}$ *function space for MP* 2*a*-2, *and is in the* $L_{1+1/4-\delta}$ *function space for MP* 2*a*-3, *where* $\delta > 0$.

(C) *The solution is in the* $L_{1+2/3-\delta}$ *function space for MPs* 2*c*-1 *and* 2*c*-3, *where* $\delta > 0$.

Note finally that all solutions of the Laplace–Young equation in a bounded domain $\Omega$ are in $L_1$, which is consistent with the physical interpretation that the volume of the fluid under the capillary surface is finite.

**2.2.** ***Asymptotic Laplace–Young equation.*** There are no closed-form solutions for the Laplace–Young equation in domains with a corner or a cusp, but closed-form solutions exist for the following simplification of the Laplace–Young PDE. These closed-form solutions will be used in Section 4 for convergence studies of the numerical methodology we propose in Section 3.

Assuming the slope of the solution of the Laplace–Young boundary value problem ((1)–(2)) is steep, i.e., $|\nabla u| \gg 1$, we can approximate the PDE and the boundary condition, by ignoring the 1 in the denominator of the differential operator $T(\cdot)$, and obtain the following boundary value problem:

$$\nabla \cdot \tilde{T}(u) = u \quad \text{in } \Omega, \tag{20}$$

$$\vec{\nu}_1 \cdot \tilde{T}(u) = \cos\gamma \quad \text{on } \partial\Omega_1, \tag{21}$$

$$\vec{\nu}_2 \cdot \tilde{T}(u) = \cos\gamma \quad \text{on } \partial\Omega_2, \tag{22}$$

where

$$\tilde{T}(u) = \frac{\nabla u}{|\nabla u|}.$$

This approximation is called the "steep slope approximation" [King et al. 1999] of the Laplace–Young boundary value problem, and unbounded closed-form solutions of this boundary value problem are known for two types of domains: the unbounded corner domain of Figure 2, left [King et al. 1999] and the circular cusp domains of Figure 4 [Aoki 2007]. Also, it has been shown that the exact solutions of this boundary value problem are good asymptotic approximations of the solutions of the original Laplace–Young equation on the same domains [Miersemann 1993; Aoki 2007]. We shall refer to this boundary value problem as the *asymptotic Laplace–Young boundary value problem*. Note that this boundary value problem is a rare case of a nonlinear PDE with nonlinear boundary conditions for which one can find closed-form solutions in some nontrivial domains.

**2.2.1.** *Closed-form solutions.* Corner domain (Figure 2, left, $\gamma + \alpha < \pi/2$): Let $u \in C^2(\Omega)$ be a solution of the boundary value problem (20)–(22) on the unbounded corner domain defined as in (3)–(5) with the boundary functions

$$f_1(x) = x\tan\alpha, \qquad f_2(x) = -x\tan\alpha.$$

If $\gamma + \alpha < \pi/2$, then $u$ is given as the following closed-form expression in terms of the polar coordinate variables $r$ and $\theta$:

$$u(r,\theta) = \frac{\cos\theta - \sqrt{k^2 - \sin^2\theta}}{kr}, \tag{23}$$

where $k = \sin\alpha/\cos\gamma$. This precise property of the asymptotic function in (8) was first observed in [King et al. 1999].

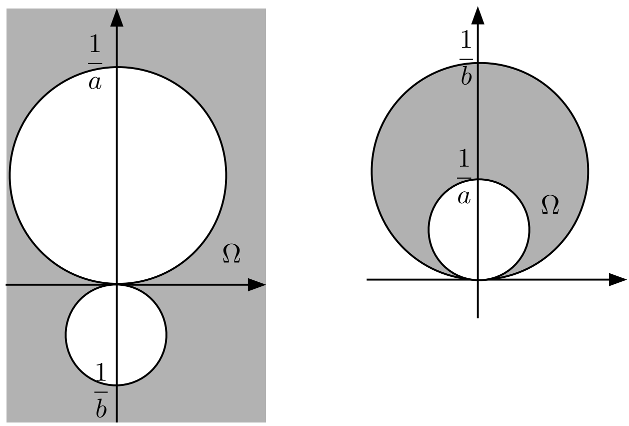



**Figure 4.** Circular cusp domains.

Circular cusp domain (Figure 4, $\gamma \neq \pi/2$): Let $u \in C^2(\Omega)$ be a solution of the boundary value problem (20)–(22) with $\gamma \neq \pi/2$ and with the domain defined as

$$\Omega := \begin{cases} \left\{(x, y) \in \mathbb{R}^2 \setminus \left(\overline{B}_{\frac{1}{2a}}\left(0, \frac{1}{2a}\right) \cup \overline{B}_{-\frac{1}{2b}}\left(0, \frac{1}{2b}\right)\right)\right\} & \text{for } b < 0, \\ \left\{(x, y) \in \left(B_{\frac{1}{2b}}\left(0, \frac{1}{2b}\right) \setminus \overline{B}_{\frac{1}{2a}}\left(0, \frac{1}{2a}\right)\right)\right\} & \text{for } b > 0, \end{cases}$$

where $B_r(x_o, y_o)$ is the open disc of radius $r$ centered at $(x_o, y_o)$, i.e.,

$$B_r(x_o, y_o) = \{(x, y) \in \mathbb{R}^2 : (x - x_o)^2 + (y - y_o)^2 < r^2\}.$$

A closed-form expression for $u$ is given by

$$u(p, q) = Ap^2 - 2\sqrt{1 - A^2(q - q_0)^2}\, p - A(q - q_0)^2 + Aq_0^2, \tag{24}$$

where

$$A = \frac{2\cos\gamma}{a - b}, \quad q_0 = \frac{a + b}{2},$$

and $p$ and $q$ are the coordinate variables of the tangent cylindrical coordinate system introduced in [Moon and Spencer 1961], depicted in Figure 5 and defined as

$$p = \frac{x}{x^2 + y^2}, \quad q = \frac{y}{x^2 + y^2}.$$

This closed-form solution of the asymptotic Laplace–Young equation first appears in [Aoki 2007]. Note that $\lim_{(x,y)\to(0,0)} p = \infty$ and the solution (24) behaves like $1/x^2$ as $x \to 0$, hence it exhibits a more severe singularity than the singularity of the asymptotic Laplace–Young PDE in a corner domain, which features a $1/r$ singularity.

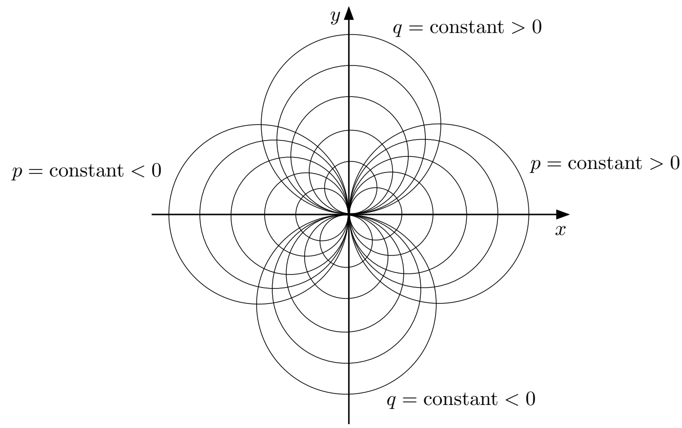



**Figure 5.** Tangent cylindrical coordinate system.

**2.2.2.** *Model Problems 3 and 4.* For the numerical experiments on the asymptotic Laplace–Young equation we consider the following model problems on the corner and cusp domains of Figure 3.
Let $u \in C^2(\Omega)$ be the function that satisfies the boundary value problem

$$\begin{aligned} \nabla \cdot \tilde{T}(u) &= u && \text{in } \Omega, \\ \vec{\nu}_1 \cdot \tilde{T}(u) &= \cos\gamma_1 && \text{on } \partial\Omega_1, \\ \vec{\nu}_2 \cdot \tilde{T}(u) &= \cos\gamma_2 && \text{on } \partial\Omega_2, \\ u &= u_{\text{exact}} && \text{on } \partial\Omega_3. \end{aligned}$$

Here with $\vec{\nu}_1, \vec{\nu}_2, \vec{\nu}_3$ are the exterior unit normal vectors on the boundaries $\partial\Omega_1$, $\partial\Omega_2$, $\partial\Omega_3$, while $\gamma_1$, $\gamma_2$ are the contact angles, and $u_{\text{exact}}$ is the closed-form solutions given in (23) or (24).

The bounded open domain $\Omega$ and boundaries $\partial\Omega_{1,2,3}$ are defined as in (16)–(17); see Figure 3. The boundary functions and the contact angles are chosen for each model problem as tabulated in Table 3, resulting in model problems with unbounded

| Name | $f_1(x)$ | $f_2(x)$ | $\gamma_1$ | $\gamma_2$ | | unbd? |
|---|---|---|---|---|---|---|
| MP 3 | $x\tan(\pi/7)$ | $-x\tan(\pi/7)$ | $\pi/6$ | $\pi/6$ | corner | yes |
| MP 4 | $-\sqrt{5^2-x^2}+5$ | $\sqrt{10^2-x^2}-10$ | $\pi/6$ | $\pi/6$ | circular cusp | yes |

**Table 3.** Model Problems 3 and 4: asymptotic Laplace–Young equation in domains with a corner and a cusp. MP 3 has $\alpha = \pi/7$ and $\gamma + \alpha < \pi/2$, resulting in a solution that is unbounded at (0,0). MP 4 has $\gamma \neq \pi/2$, and its solution is also unbounded at (0,0).

solutions. The closed-form solutions of these two model problems are given by (23) and (24), respectively.

## 3. Numerical method

In this section, we propose a numerical methodology to accurately find global numerical approximations of singular solutions of the Laplace–Young equation in domains with a corner or a cusp. The starting point of our approach is the finite volume element method (FVEM) [Bank and Rose 1987; Aoki and De Sterck 2011] or the Galerkin finite element method (FEM) [Strang and Fix 1973; Brenner and Scott 1994], and two simple but crucial additional steps are made to arrive at a method that can capture the singular behavior. The first step is to consider a change of variable, with the new solution variable being smoother than the capillary height variable $u$ and more amenable to accurate numerical approximation. The second step is to solve the PDE numerically in a new coordinate system, which allows us to accurately represent the discontinuous behavior of the new solution variable at the singular point. We describe these two crucial ingredients of our methodology along with the FEM and FVEM discretizations, and show in the numerical results of Section 4 that this approach leads to a global approximation method for singular solutions of the Laplace–Young equation that recovers the proper asymptotic behavior, and is more accurate and has better convergence properties than numerical methods that were considered previously.

**3.1.** ***Change of variable.*** From the asymptotic analysis results (10)–(15) we observe that the solutions we wish to approximate have the asymptotic behavior

$$\begin{aligned} u(x, y) &= O\left(\frac{1}{f_1(x) - f_2(x)}\right) && \text{as } x \to 0 \\ &= \frac{O(1)}{f_1(x) - f_2(x)} && \text{as } x \to 0. \end{aligned}$$

This implies that, if we transform the unknown function $u(x, y)$ as follows, the new unknown function $v(x, y)$ is a bounded function:

$$u(x, y) = \frac{v(x, y)}{f_1(x) - f_2(x)}.$$

We aim to approximate the solution of the boundary value problem, $u(x, y)$, by numerically approximating the new unknown function $v(x, y)$. Since $v(x, y)$ is bounded while $u(x, y)$ is unbounded, we expect a better quality of numerical approximation.

**3.2.** ***Change of coordinates.*** An appropriate choice of coordinate system is essential for the asymptotic analysis of unbounded solutions of the Laplace–Young equation, as shown in [Miersemann 1993; Scholz 2004; Aoki 2007; Aoki and Siegel 2012]. We have observed that an appropriate choice of coordinate system is also beneficial for the numerical approximation of unbounded solutions.

For MP 1, we can observe as follows that the new unknown function $v$ is discontinuous at the origin. From (9), we know that the solution $u$ of MP 1 behaves like $(\cos\theta - \sqrt{k^2 - \sin^2\theta})/kr$ near the origin $r = 0$. This gives that the new unknown function $v$ behaves like $(\cos\theta - \sqrt{k^2 - \sin^2\theta})/k$ near the origin. Hence, as $r \to 0$, $v$ approaches different values depending on the angle $\theta$, so the new unknown function $v$ has a jump discontinuity at the origin. Our idea is to expand the point of singularity on the boundary into a boundary line segment through a coordinate transformation in order to accurately approximate the discontinuous behavior of $v$.

For MP 2, since the boundaries for the cusp domain are curved boundaries, we would need special boundary elements (e.g., isoparametric elements) to accurately represent the cusp domain when approximating the unknown function through finite element approximation in the standard $(x, y)$ coordinate system. However, the change to $(s, t)$ coordinates introduced in [Aoki and Siegel 2012] and illustrated in Figure 6 transforms a cusp domain into a rectangular domain, and hence no special treatment is needed for curved boundaries.

We use this $(s, t)$ coordinate system for numerical simulation on domains with a corner or a cusp at $(0, 0)$. The $(s, t)$ coordinate transformation as depicted in Figure 6 is given by

$$t = \frac{2y - (f_1 + f_2)}{f_1 - f_2}, \quad s = x.$$

The Cartesian coordinates can be expressed using the above coordinate system as

$$x = s, \quad y = \frac{t(f_1(s) - f_2(s)) + (f_1(s) + f_2(s))}{2} = \frac{1+t}{2} f_1(s) + \frac{1-t}{2} f_2(s).$$

We have $y = f_1(x)$ when $t = 1$ and $y = f_2(x)$ when $t = -1$, so the domain of interest in the curvilinear $(s, t)$ coordinate system can be written as (see Figure 6)

$$\Omega = \{(s, t) \in \mathbb{R}^2 : 0 < s < 1,\ -1 < t < 1\}. \tag{25}$$

With some calculation, the left-hand side of the Laplace–Young PDE can be rewritten in the curvilinear coordinate system as

$$\begin{aligned}\nabla\cdot T(u) = {} & \frac{\partial}{\partial s}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}} + \frac{f_1'-f_2'}{f_1-f_2}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}} \\ & + \frac{\partial}{\partial t}\left(\frac{2}{f_1-f_2}\frac{u_y}{\sqrt{1+u_x^2+u_y^2}} + \left(-\frac{f_1'+f_2'}{f_1-f_2} - t\frac{f_1'-f_2'}{f_1-f_2}\right)\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}\right),\end{aligned}$$

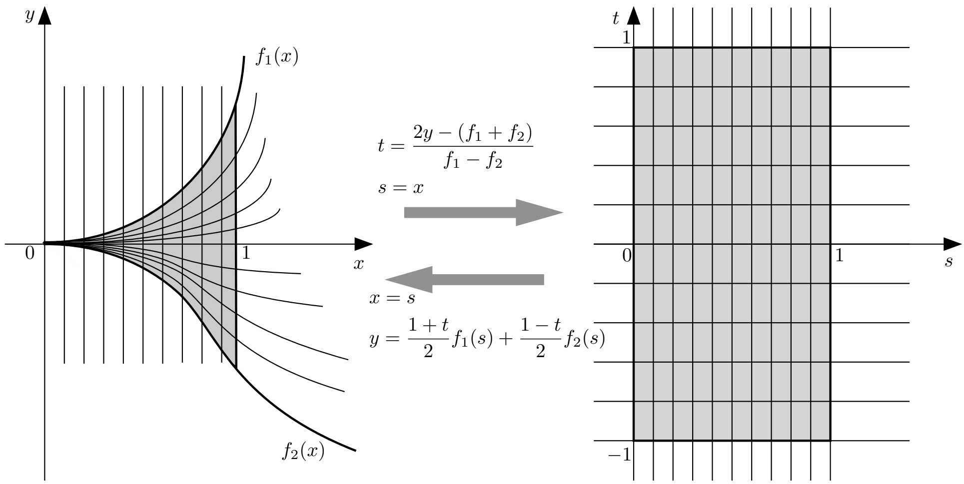



**Figure 6.** Coordinate transformation.

where

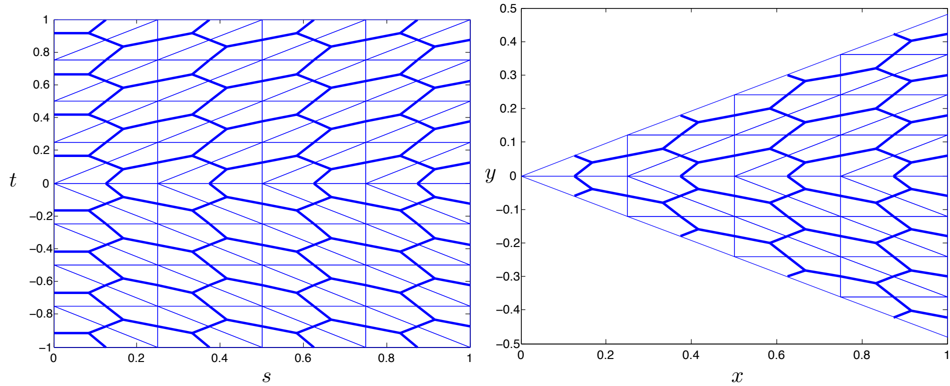
$$u_x = \frac{v_s}{f_1 - f_2} - \frac{v(f_1' - f_2')}{(f_1 - f_2)^2} - v_t \frac{(f_1' + f_2') + t(f_1' - f_2')}{(f_1 - f_2)^2}, \quad u_y = \frac{2v_t}{(f_1 - f_2)^2}.$$

The boundary conditions on $\partial\Omega_1$ and $\partial\Omega_2$ can be written as

$$\begin{aligned} \text{(26)} \quad \vec{\nu}_{1,2}\cdot T(u) &= \vec{\nu}_{1,2}\cdot\hat{s}\left(\frac{f_1 - f_2}{2}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}\right) \\ &\quad + \vec{\nu}_{1,2}\cdot\hat{t}\left(\frac{u_y}{\sqrt{1+u_x^2+u_y^2}} + \frac{-(f_1'+f_2') - t(f_1'-f_2')}{2}\frac{u_x}{\sqrt{1+u_x^2+u_y^2}}\right) \\ &= \sqrt{1+f_{1,2}'(s)^2}\cos\gamma_{1,2} \quad \text{on } \partial\Omega_{1,2}. \end{aligned}$$

The boundary condition for boundary $\partial\Omega_3$ of MPs 1 and 2 is as in (26) but with zero on the right-hand side. The left-hand side and the boundary conditions of the asymptotic Laplace–Young PDE in the $(s, t)$ coordinate system can be obtained by just neglecting the 1 in the denominator in the expressions above. Note that the point $(x, y) = (0, 0)$ corresponds to the line segment $(s = 0, t \in [-1, 1])$ in the $(s, t)$ coordinate system.

**3.3.** ***Discretized boundary value problem.*** In the numerical results of Section 4 we approximate the new unknown function $v(s, t)$ in the new coordinate variables $s$ and $t$ numerically on the Cartesian grid in $(s, t)$-space, as shown in Figure 7(a), and for comparison we also perform some calculations on the corner domain of Figure 7(b) without a change of coordinates. We now describe the Galerkin finite




(a) $(s, t)$ space

(b) a corner domain in $(x, y)$ space (without change of coordinates)

**Figure 7.** Finite elements and control volumes for the numerical methods. The thin lines give the finite element triangulation, which is used in both the FEM and the FVEM. The thick lines give the control volumes that are used in the FVEM. The grid in panel (a) can be used for corner domains or for cusp domains (depending on the boundary functions $f_1$ and $f_2$ that enter into the coordinate transformation formulas), and the grid in panel (b) is used for comparison simulations for corner domains (without coordinate transformation).

element method (FEM) and the finite volume element method (FVEM) discretizations.

**3.3.1.** *Galerkin finite element method discretization.* We follow the construction of the finite element space presented in Chapter 3 of Brenner and Scott [1994]. Let $N_{\text{node}}$ be the number of nodes created by finite element triangulation of the domain and $\mathcal{N}$ be the set of indices of the nodes, i.e., $\mathcal{N} = \{1, 2, \ldots, N_{\text{node}}\}$. The triangulation of the domain is as depicted in Figure 7(a) (or Figure 7(b) for the corner problem without a change of coordinates). Also, we let $\mathcal{N}_{\text{Dirichlet}}$ be the indices of the nodes on the boundary with Dirichlet boundary condition. That is to say, for MPms 3 and 4,

$$(s_i, t_i) \in \overline{\partial\Omega_3} \Rightarrow i \in \mathcal{N}_{\text{Dirichlet}},$$

where $(s_i, t_i)$ is the location of the $i$-th node, and for MPs 1 and 2 $\mathcal{N}_{\text{Dirichlet}} = \varnothing$ since there is no Dirichlet boundary. Let $\phi_i(s, t)$ be the standard continuous piecewise linear nodal basis function (tent function) that corresponds to node $i$ in the finite element triangulation on domain $\Omega$. We have

$$\phi_i(s_j, t_j) = \delta_{i,j},$$

where $\delta_{i,j}$ is the Kronecker delta function. We approximate the unknown function $v$ with a linear combination of these basis functions, i.e.,

$$v \approx v^h := \sum_{i=1}^{N_{\text{node}}} c_i \phi_i.$$

The $\{c_1, c_2, \ldots, c_{N_{\text{node}}}\}$ are the unknowns of the discretized boundary value problem. The Galerkin finite element discretization of MPs 1 and 2 can then be written as follows (the discretization of MPs 3 and 4 can be derived similarly):

$$\int_\Omega \left( \nabla \cdot T\left( \frac{\sum_{i=1}^{N_{\text{node}}} c_i \phi_i}{f_1(s) - f_2(s)} \right) \right) \phi_j \, \mathrm{d}A = \int_\Omega \frac{\sum_{i=1}^{N_{\text{node}}} c_i \phi_i}{f_1(s) - f_2(s)} \phi_j \, \mathrm{d}A \qquad \text{for } j \in \mathcal{N} \backslash \mathcal{N}_{\text{Dirichlet}}, \tag{27}$$

$$\frac{c_i}{f_1(s_i) - f_2(s_i)} = u_{\text{exact}}(s_i, t_i) \qquad \text{for } i \in \mathcal{N}_{\text{Dirichlet}}. \tag{28}$$

By the divergence theorem we can rewrite (27) as

$$\int_{\partial\Omega} \left( \nu \cdot T\left( \frac{\sum_{i=1}^{N_{\text{node}}} c_i \phi_i}{f_1(s) - f_2(s)} \right) \right) \phi_j \, \mathrm{d}l - \int_\Omega \left( T\left( \frac{\sum_{i=1}^{N_{\text{node}}} c_i \phi_i}{f_1(s) - f_2(s)} \right) \right) \cdot \nabla \phi_j \, \mathrm{d}A = \int_\Omega \sum_{i=1}^{N_{\text{node}}} c_i \frac{\phi_i \phi_j}{f_1(s) - f_2(s)} \, \mathrm{d}A \qquad \text{for } j \in \mathcal{N} \backslash \mathcal{N}_{\text{Dirichlet}}.$$

By imposing the boundary conditions (26), we obtain the following system of equations:

$$\int_\Omega \left( T\left( \frac{\sum_{i=1}^{N_{\text{node}}} c_i \phi_i}{f_1(s) - f_2(s)} \right) \right) \cdot \nabla \phi_j \, \mathrm{d}A - \sum_{i=1}^{N_{\text{node}}} c_i \int_\Omega \frac{\phi_i \phi_j}{f_1(s) - f_2(s)} \, \mathrm{d}A = \int_{\partial\Omega_1} \sqrt{1 + f_1'(s)^2} \cos\gamma_1 \phi_j \, \mathrm{d}l + \int_{\partial\Omega_2} \sqrt{1 + f_2'(s)^2} \cos\gamma_2 \phi_j \, \mathrm{d}l \qquad \text{for } j \in \mathcal{N} \backslash \mathcal{N}_{\text{Dirichlet}}. \tag{29}$$

After some calculation we can rewrite (29) together with (28) as the following system of nonlinear equations:

$$
\begin{aligned}
(30)\quad & \int_{t=-1}^{1}\int_{s=0}^{1}(\phi_j)_s\left(\frac{f_1-f_2}{2}\frac{u_x^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}\right) \\
& +(\phi_j)_t\left(\frac{u_y^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}+\frac{-(f_1'+f_2')-t(f_1'-f_2')}{2}\frac{u_x^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}\right)\mathrm{d}s\,\mathrm{d}t \\
& \qquad -\sum_{i=1}^{N_{\text{node}}} c_i\int_{t=-1}^{1}\int_{s=0}^{1}\phi_i\phi_j\,\mathrm{d}s\,\mathrm{d}t \\
& =\int_{\partial\Omega_1}\sqrt{1+f_1'(s)^2}\cos\gamma_1\phi_j\,\mathrm{d}l+\int_{\partial\Omega_2}\sqrt{1+f_2'(s)^2}\cos\gamma_2\phi_j\,\mathrm{d}l \\
& \qquad\qquad \text{for } j\in\mathcal{N}\backslash\mathcal{N}_{\text{Dirichlet}}, \\
& c_i=u_{\text{exact}}(x_i,y_i) \qquad \text{for } i\in\mathcal{N}_{\text{Dirichlet}},
\end{aligned}
$$

where

$$
(31)\quad u_x^h=\sum_{i=1}^{N_{\text{node}}} c_i\left(\frac{(\phi_i)_s}{f_1-f_2}-\frac{(\phi_i)(f_1'-f_2')}{(f_1-f_2)^2}-(\phi_i)_t\frac{(f_1'+f_2')+t(f_1'-f_2')}{(f_1-f_2)^2}\right),
$$

$$
(32)\quad u_y^h=\sum_{i=1}^{N_{\text{node}}} c_i\frac{2(\phi_i)_t}{(f_1-f_2)^2}
$$

and $(\phi_i)_s$ and $(\phi_i)_t$ are the partial derivatives of $\phi_i$ with respect to $s$ and $t$. We can construct a system of nonlinear equations by integrating each of the terms in (30) numerically. Note that although we are integrating the unbounded functions $v^h\phi_j/(f_1(s)-f_2(s))$, due to the change of coordinates the area element $\mathrm{d}A$ becomes $(f_1(s)-f_2(s))/2\,\mathrm{d}s\,\mathrm{d}t$, hence the integrand becomes $2v^h\phi_j$, a piecewise quadratic polynomial; hence we avoid singular integration. We solve this system of nonlinear equations with the Levenberg–Marquardt method to obtain the unknowns $\{c_1, c_2, \dots, c_{N_{\text{node}}}\}$. This gives a numerical approximation for $v$, and hence a numerical approximation of the solution of the boundary value problem $u$.

**3.3.2.** *Finite volume element method discretization.* The finite volume element method (FVEM) is a type of Petrov–Galerkin method that uses piecewise constant functions as test functions in the weak form, instead of using the finite element basis functions as in the Galerkin FEM. The test functions for the FVEM are chosen as follows:

$$
\psi_j(s,t)=\begin{cases}1 & \text{if } (s,t)\in\Omega_j,\\ 0 & \text{otherwise,}\end{cases}
$$

where the $\Omega_j$ are control volumes constructed as in [Bank and Rose 1987] (note that in [Bank and Rose 1987] the control volumes are called “boxes”). As depicted in Figure 7(a) (and Figure 7(b)), the control volumes $\{\Omega_j\}_{j=1}^{N_{\text{node}}}$ are constructed by first computing the centroids of the finite element triangles, and then connecting those element centroids with the midpoints of the finite element triangle edges. This construction divides each finite element triangle into three quadrilaterals. The control volume $\Omega_j$ for finite element node $j$ is then constructed as the union of the quadrilaterals adjacent to node $j$.

By substituting the test functions $\phi_j$ by $\psi_j$ in the Galerkin finite element discretization (27) and after some calculation, we obtain the following system of nonlinear equations for the FVEM, where $u_x$ and $u_y$ are defined as in (31) and (32):

$$\begin{aligned}(33)\quad &\int_{\partial\Omega_j} \vec{\nu}\cdot\hat{s}\left(\frac{f_1-f_2}{2}\frac{u_x^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}\right)\\ &+\vec{\nu}\cdot\hat{t}\left(\frac{u_y^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}+\frac{-(f_1'+f_2')-t(f_1'-f_2')}{2}\frac{u_x^h}{\sqrt{1+(u_x^h)^2+(u_y^h)^2}}\right)\mathrm{d}l\\ &\qquad -\sum_{i=1}^{N_{\text{node}}} c_i \iint_{\Omega_j}\phi_i\,\mathrm{d}s\,\mathrm{d}t\\ &=\int_{\partial\Omega_1\cap\partial\Omega_j}\sqrt{1+f_1'(s)^2}\cos\gamma_1\,\mathrm{d}l+\int_{\partial\Omega_2\cap\partial\Omega_j}\sqrt{1+f_2'(s)^2}\cos\gamma_2\,\mathrm{d}l\\ &\qquad \text{for } j\in\mathcal{N}\setminus\mathcal{N}_{\text{Dirichlet}}.\end{aligned}$$

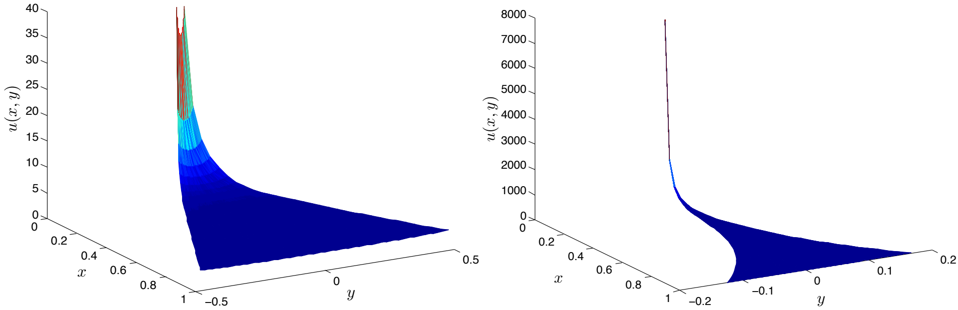
Again, we avoid singular integration by the change of coordinates, hence the integration can be done numerically without any special treatment for singular integration. We solve the resulting system of nonlinear equations using the Levenberg–Marquardt method.

Note that we choose the triangulations of Figures 7(a) and 7(b) symmetric with respect to the $t=0$ and $y=0$ axes, respectively. While this is not a requirement, we made this choice because some of our model problems are symmetric with respect to the $t=0$ and $y=0$ axes, and this choice of grid leads to numerical solutions that closely retain this symmetry.

The FEM is known to achieve optimality in the energy norm for linear elliptic PDEs, but it does not have a local conservation property. The FVEM has a local conservation property like the finite volume method; however, it does not necessarily produce an optimal approximation. We have conducted numerical experiments using both methods, and the results we obtained were very similar. For brevity, we mainly present the numerical experiment results obtained by the FVEM in this paper, except in a few places where we compare them with the Galerkin FEM.

## 4. Numerical results

We now show that the numerical approximations we obtain with the computational methodology proposed in Section 3 for singular solutions of the Laplace–Young




(a) MP 1-1: corner problem (b) MP 2a-1: cusp problem

**Figure 8.** MPs 1-1 and 2a-1. FVEM solution on the $(s, t)$-type grid of Figure 7(a) with $33 \times 65$ nodes. Surface plots of the unbounded capillary surfaces in the corner and cusp domains.

equation in domains with a corner or a cusp are accurate global approximations. As an initial illustration, surface plots for two numerical approximations of singular solutions of the Laplace–Young equation in domains with a corner and with a cusp are shown in Figure 8. In what follows, we first show how our numerical methods obtain accurate global solutions for unbounded solutions of the Laplace–Young equation in domains with a corner or a cusp, by comparing with known asymptotic expansions and formal asymptotic series. We then numerically investigate the convergence behavior of the methods we propose using known closed-form unbounded solutions for the asymptotic Laplace–Young equation. The numerical results confirm that the computational methods we propose are accurate and have good convergence properties, and that they can be used with confidence to numerically investigate open problems on asymptotic solutions of the Laplace–Young equation in Section 5.

**4.1.** ***Laplace–Young equation: asymptotic behavior.*** We now investigate how well our numerical solutions can approximate the singular behavior by comparing the numerical solutions to known asymptotic solutions for the Laplace–Young equation.

*MP 1: corner problem.* As given in (8), the leading-order term of the asymptotic series solution of the Laplace–Young equation at a sharp corner is known. In Figure 9, we plot a horizontal cross-section (a cross-section along the $x$-axis or $s$-axis; see Figure 6) of the numerical approximation and the asymptotic approximation in log-log scale. In Figure 10, we plot a vertical cross-section (a cross-section along the line $x = 1/2^5$ or $s = 1/2^5$; see Figure 6) of the numerical approximation and the asymptotic approximation.

In order to illustrate the crucial benefits of the change of variable and change of coordinates that are the essential building blocks of the numerical methodology we proposed in Section 3, we compare four different choices for obtaining the numerical approximation using the FVEM: with or without change of variable, and with or without change of coordinates. The only published work on numerical approximation of singular capillary surfaces [Scott et al. 2005] also uses the FVEM, but it does not use a change of variable nor a change of coordinates, and thus corresponds to Figures 9 and 10.

As can be seen in Figures 9 and 10, the change of variable and the change of coordinates proposed in Sections 3.1 and 3.2 are very beneficial for the accuracy of the numerical approximations on a domain with a sharp corner near the singularity. Note that we cannot conduct a numerical convergence study for these unbounded solutions of the Laplace–Young equation, as there is no known closed-form solution.

*MP 2: cusp problem.* We now consider the Laplace–Young equation in a domain with a cusp. Unbounded cusp solutions are known to have a more severe singularity than the sharp corner problem. The leading-order term of the asymptotic series solution is known; see (11). Also, as shown in Lemma 2.3 of [Aoki and Siegel 2012], the first two terms of the formal asymptotic series $\tilde{u}$ are known:

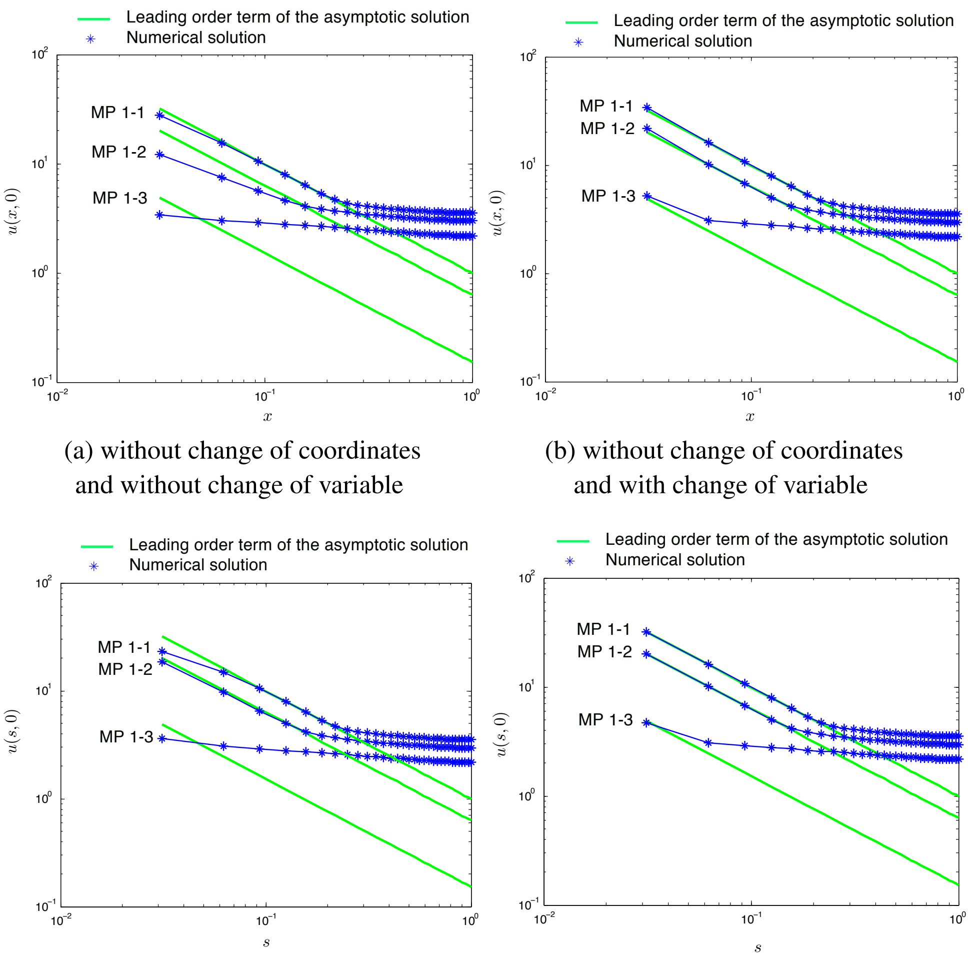
$$\tilde{u} = \frac{\cos\gamma_1 + \cos\gamma_2}{f_1(s) - f_2(s)} - \sqrt{1 - \left(\frac{\cos\gamma_1(t+1) + \cos\gamma_2(t-1)}{2}\right)^2}\,\frac{f_1'(s) - f_1'(s)}{f_1(s) - f_2(s)}. \tag{34}$$

The formal asymptotic series of a boundary value problem is a series that satisfies the PDE and the boundary condition asymptotically, but, as opposed to the case of an asymptotic expansion, a bound on the error has not been proven. (There is no $O(\cdot)$ term in (34), but there is one in the asymptotic expansion (11).)

As can be seen in Figure 11, the numerical solution we obtain for MP 2 with the change of variable and the change of coordinates proposed in Sections 3.1 and 3.2 accurately approximates the singular behavior.

Although it is not known if the second-order term of the formal asymptotic series of this problem is in fact the second-order term of the asymptotic series solution, it can be seen in Figure 11 that the numerical solution appears to match better with the second-order formal asymptotic series than with the first-order asymptotic solution. It is particularly interesting that the domain where the asymptotic approximation is a good approximation seems to expand by adding a second term to the asymptotic series.

**4.2.** ***Asymptotic Laplace–Young equation: convergence study.*** In the previous section, we have shown that the numerical approximations with a change of variable and a change of coordinates as proposed in Sections 3.1 and 3.2 exhibit the correct singular behavior for singular solutions of the Laplace–Young equation. Since our interest is to obtain global approximations which are accurate both at the singularity and away from the singularity, we now show that the numerical solution in fact



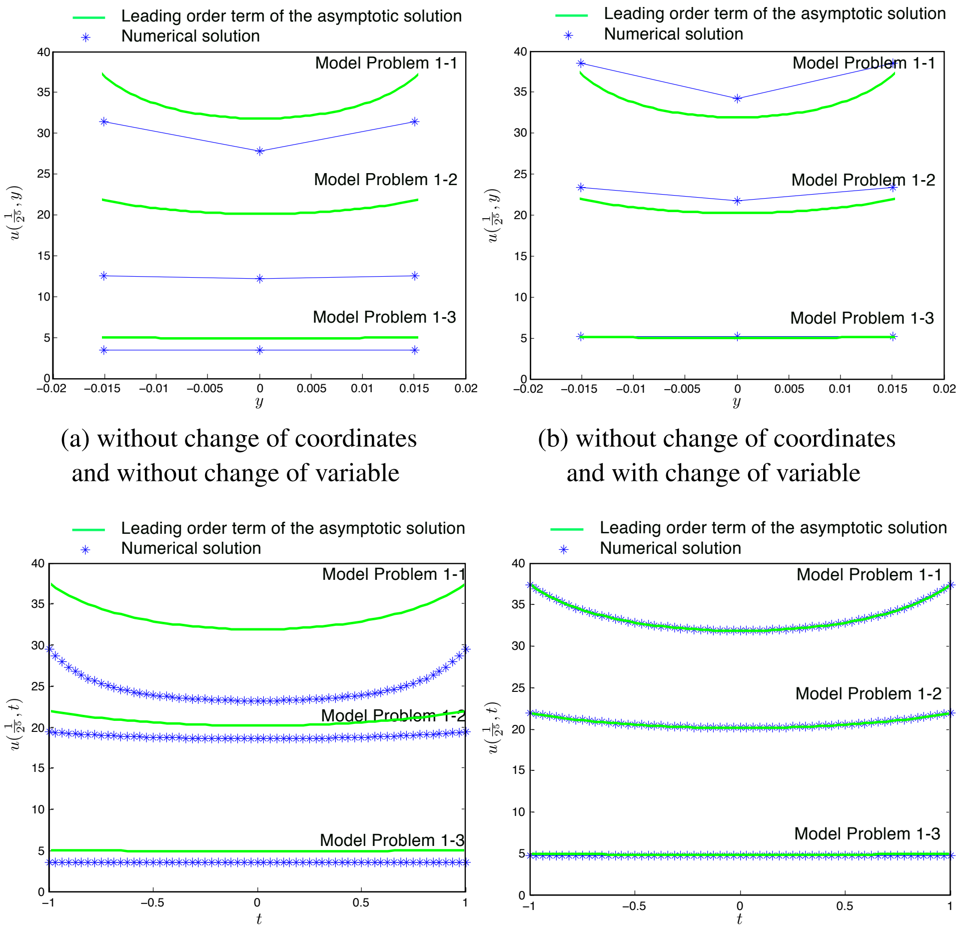

(a) without change of coordinates and without change of variable

(b) without change of coordinates and with change of variable

(c) with change of coordinates and without change of variable

(d) with change of coordinates and with change of variable

**Figure 9.** MPs 1-1, 1-2 and 1-3 (unbounded corner solutions). Panels (a) and (b) show FVEM solutions on the $(x, y)$-type grid of Figure 7(b) with 1089 nodes (no change of coordinates). Panels (c) and (d) show FVEM solutions on the $(s, t)$-type grid of Figure 7(a) with $33 \times 65$ nodes (with change of coordinates). Panels (a) and (c) are for computation of the original variable $u$, and panels (b) and (d) are for computation of the transformed variable $v$. The log-log plots show a comparison of the numerical solutions and the first-order asymptotic approximations in a horizontal cross section at $y = 0$ or $t = 0$. Panel (d) clearly gives the most accurate numerical solutions.




(a) without change of coordinates and without change of variable

(b) without change of coordinates and with change of variable

(c) with change of coordinates and without change of variable

(d) with change of coordinates and with change of variable

**Figure 10.** MPs 1-1, 1-2, and 1-3 (unbounded corner solutions). Panels (a) and (b) show FVEM solutions on the $(x, y)$-type grid of Figure 7(b) with 1089 nodes (no change of coordinates). Panels (c) and (d) show FVEM solutions on the $(s, t)$-type grid of Figure 7(a) with $33 \times 65$ nodes (with change of coordinates). Panels (a) and (c) are for computation of the original variable $u$, and panels (b) and (d) are for computation of the transformed variable $v$. The plots show a comparison of the numerical solutions and the first-order asymptotic approximations in a vertical cross section at $x = 1/2^5$ or $s = 1/2^5$ (the grid points closest to the singular point). Panel (d) clearly gives the most accurate numerical solutions.

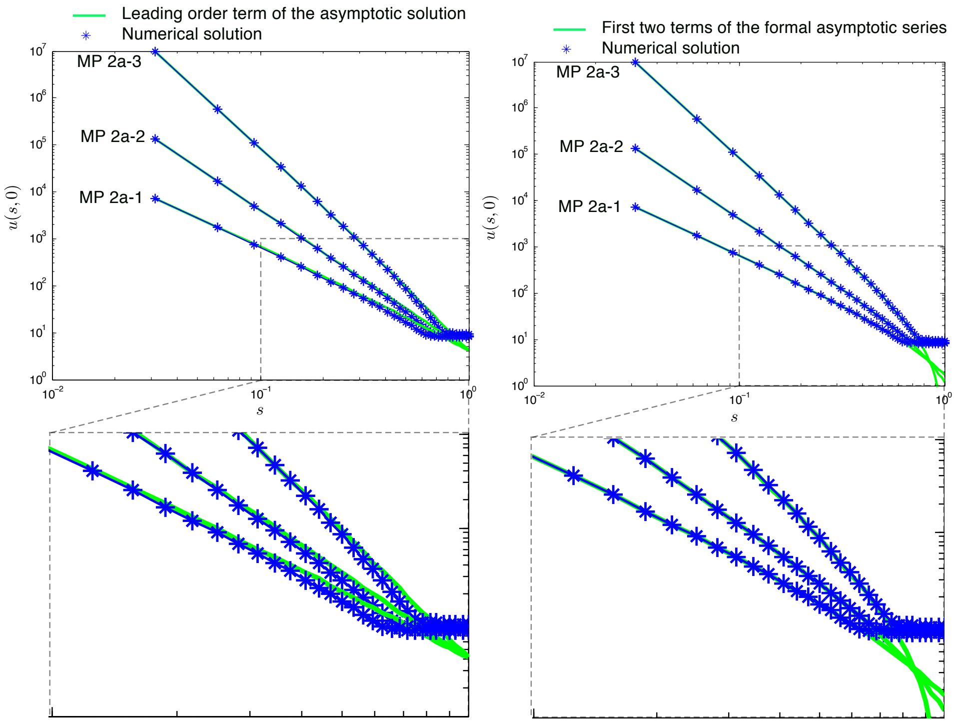



**Figure 11.** MPs 2a-1, 2a-2 and 2a-3 (unbounded cusp solutions). FVEM solutions on the $(s,t)$-type grid of Figure 7(a) with $33 \times 65$ nodes (with change of coordinates and with change of variable). The log-log plots in the left panels show a comparison of the numerical solutions with the first-order asymptotic solution in a horizontal cross section at $t = 0$. The log-log plots in the right panels show a comparison of the numerical solutions with the first two terms of the formal asymptotic series in a horizontal cross-section at $t = 0$. It is clear that accurate numerical solutions are obtained.

converges to the exact solution everywhere. It would be desirable to conduct a numerical convergence study for the Laplace–Young equation, but there is no known closed-form singular solution, and hence we cannot conduct a numerical convergence study. As we have discussed in Section 2.2, there are known exact solutions of the asymptotic Laplace–Young equation, and it is known that they have the same singular behavior as the corresponding solutions of the Laplace–Young equation. We therefore conduct a numerical convergence study for the asymptotic Laplace–Young equation in corner and cusp domains. Since the exact solution is in the $L_1$ function space but not in $L_2$, we conduct the convergence study in the $L_1$ norm.

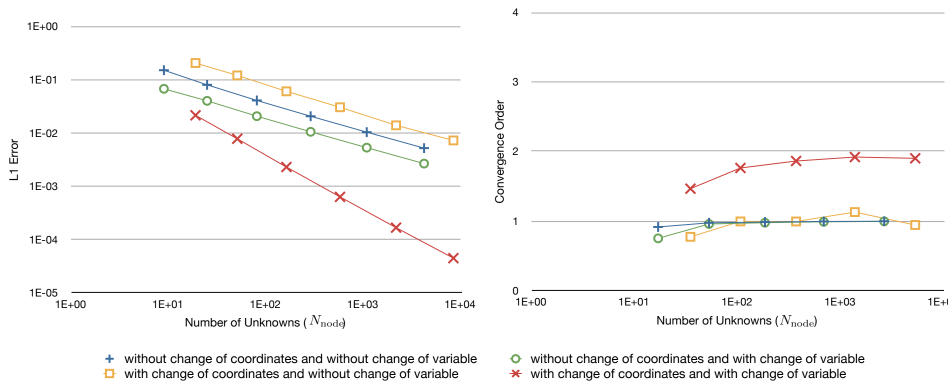



**Figure 12.** MP 3 (unbounded corner solution for asymptotic Laplace–Young). FVEM solutions on $(s, t)$-type grids (Figure 7(a)) and on $(x, y)$-type grids (Figure 7(b)), with and without change of variable. The plots show $L_1$ convergence of the numerical solutions obtained by the FVEM to the closed-form solution. The plots indicate that all four approaches converge, but it is clear that the method with change of variable and with change of coordinates converges significantly faster (with nearly second-order accuracy) than the other approaches.

*MP 3: corner problem.* As can be seen in Figure 12, the FVEM numerical approximation with change of variable and change of coordinates as proposed in Sections 3.1 and 3.2 converges to the closed-form solution nearly quadratically, whereas the other approaches (no change of variable or no change of coordinates) only converge linearly.

*MP 4: cusp problem.* We have also conducted a numerical convergence study for the circular cusp problem, where the solution has a more severe singularity than for the corner problem. For this problem, we have used both the Galerkin finite element method (FEM) and the finite volume element method (FVEM) to show that both numerical schemes work well with the change of variable and the change of coordinates proposed in Sections 3.1 and 3.2. As can be seen in Figure 13, both the FEM and the FVEM achieve near-quadratic convergence with the change of variable and change of coordinates, while only linear convergence can be achieved without change of variable.

## 5. Conjectures on open problems

As shown in the previous section, we can obtain a globally accurate approximation of unbounded solutions of the Laplace–Young equation using the numerical

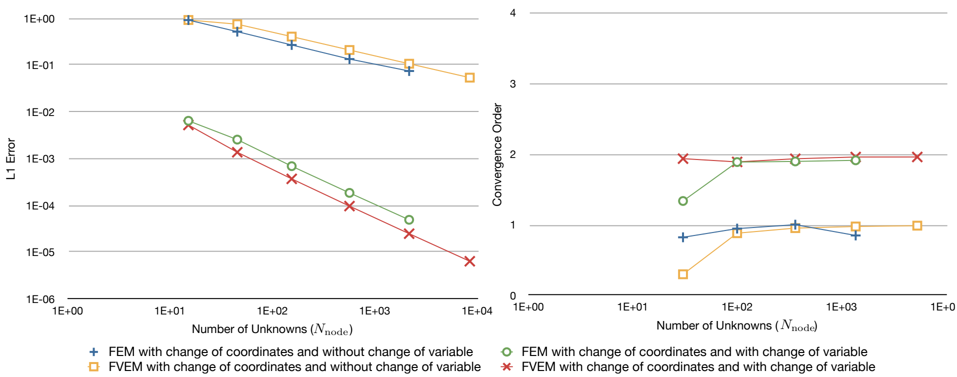



**Figure 13.** MP 4 (unbounded cusp solution for asymptotic Laplace–Young). FVEM and FEM solutions on $(s, t)$-type grids (Figure 7(a), with change of coordinates), with and without change of variable. The plots show $L_1$ convergence of the numerical solutions obtained by the FVEM and FEM to the closed-form solution. The plots indicate that all four approaches converge, but it is clear that the methods with change of variable converge significantly faster (with nearly second-order accuracy).

methodology proposed in Section 3. We now numerically approximate the solutions of two problems where the singular behavior is not known yet analytically. Our numerical results will allow us to formulate conjectures on asymptotic behavior for these open problems, which may guide further analytical study of these open problems.

**5.1.** ***Open problem 1: osculatory cusp with nonsupplementary contact angles.*** As stated in Section 2.1.2, the leading-order asymptotic behavior of the unbounded solution of the Laplace–Young equation at an osculatory cusp is not known: In summary, a proof for the leading-order asymptotic behavior could not be obtained in [Aoki and Siegel 2012] for the osculatory cusp because the authors were not able to determine the formal asymptotic series. As shown in Lemma 2.2 of that paper, the first two terms of the formal asymptotic series are known for the osculatory cusp case up to an additive constant in the coefficient of the second-order term, i.e.,

$$\tilde{u} = \frac{\cos\gamma_1 + \cos\gamma_2}{f_1(s) - f_2(s)} + \left(-\sqrt{1 - \left(\frac{\cos\gamma_1(t+1) + \cos\gamma_2(t-1)}{2}\right)^2} + C_1\right)\frac{f_1'(s) - f_1'(s)}{f_1(s) - f_2(s)}, \tag{35}$$

where $\tilde{u}$ asymptotically satisfies the boundary value problem, $C_1 = 0$ if the cusp is not an osculatory cusp, and $C_1$ is unknown if it is an osculatory cusp. One can see from the proofs in [Aoki and Siegel 2012] that the unknown additive constant $C_1$ is the elusive key to the proof of the leading-order behavior of the osculatory cusp problem. The coefficient $C_1$ is unknown and may depend on the specific functional form of the boundary functions $f_1(s)$ and $f_2(s)$.

Physical intuition suggests that the singular behavior of the unbounded capillary surface near a sharp corner or a cusp may be governed only by the distance between the two boundaries forming the sharp corner or cusp. In other words, one may think that the asymptotic behavior should only depend on $f_1(s)-f_2(s)$ and its derivatives and not on $f_1(s)$ and $f_2(s)$ separately. This would imply that the formal asymptotic series would be the same for the four MPs 2a-2 and 2b, since $f_1(s)-f_2(s)=7/24\,x^3$ for all these cases. If so, then $C_1=0$ is required also for the osculatory cusps of MPs 2b, since $C_1=0$ for the regular cusp of MP 2a-2. But it is also possible that $C_1$ depends on the precise functional form of $f_1(s)$ and $f_2(s)$.

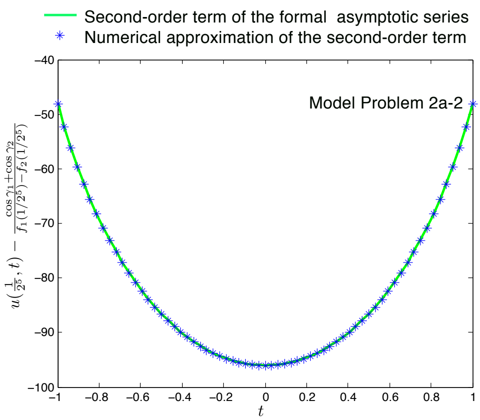
In order to investigate this, we now numerically approximate the second-order term of the formal asymptotic series by the following change of variable for the unknown function $u$:

$$u(s,t)=\frac{\cos\gamma_1+\cos\gamma_2}{f_1(s)-f_2(t)}+w(s,t)\frac{f_1'(s)-f_2'(s)}{f_1(s)-f_2(s)}. \tag{36}$$

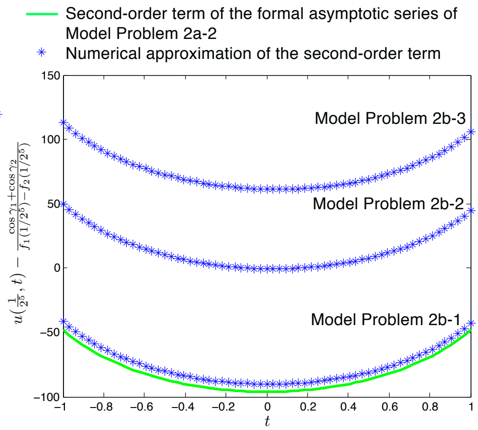
We numerically approximate the new unknown function $w(s,t)$ in $(s,t)$ coordinates, and we plot the second-order term $w(s,t)(f_1'(s)-f_2'(s))/(f_1(s)-f_2(s))$ (or equivalently, $u(s,t)-(\cos\gamma_1+\cos\gamma_2)/(f_1(s)-f_2(t))$) obtained from the numerical approximation in Figure 14.

As can be seen in Figure 14, the known second-order term of the formal asymptotic series for the regular cusp (MP 2a-2) is approximated correctly using the change of variable (36). Also, Figure 14 shows that the second-order term of the formal asymptotic series of the osculatory cusp case differs from the regular cusp case and is shifted up by constants, consistent with (35). These numerical results guide us in conjecturing that the additive constant $C_1$ of the coefficient of the second-order formal asymptotic series for the osculatory cusp changes depending on the leading-order term of the boundary functions $f_1(s)$ and $f_2(s)$, and is strictly greater than 0. The numerical evidence from Figure 14 indeed indicates that $C_1$ is not zero for osculatory cusps and that the asymptotic behavior depends on $f_1(s)$ and $f_2(s)$, and not just on the difference $f_1(s)-f_2(s)$. This conjecture on the unknown constant $C_1$ in (35), obtained from numerical investigation, can guide future analytical study of this case.

**5.2.** ***Open problem 2: infinite-curvature cusp with supplementary contact angles.*** Another open problem on the singular behavior of the Laplace–Young equation in a cusp domain is the infinite-curvature boundary cusp (i.e., $\lim_{x\to 0} f_1''(x)=\infty$ or $\lim_{x\to 0} f_2''(x)=\infty$) with supplementary contact angles (i.e., $\gamma_1+\gamma_2=\pi$). It was




(a) Regular Cusp: MP 2a-2



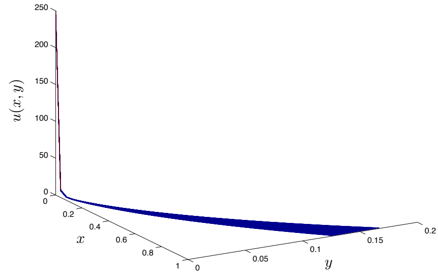

(b) Osculatory Cusp: MP 2b

**Figure 14.** MPs 2a-2 (unbounded cusp solution) and 2b (unbounded osculatory cusp solution, open problem). FVEM solutions on an $(s, t)$-type grid (Figure 7(a)) with $33 \times 65$ nodes, with change of variable. The plots show vertical cross sections at $s = 1/2^5$. The left panel shows how the numerical solution tracks the second-order term of the formal asymptotic series. The right panel supports the conjecture that $C_1 > 0$ in (35).

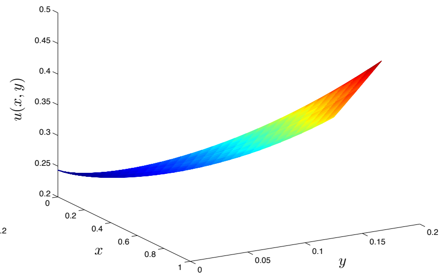
proven in [Aoki and Siegel 2012] that the cusp solution is bounded if the contact angles are supplementary angles and the boundaries forming the cusp have finite curvatures (but it is unbounded if the contact angles are not supplementary).

We conduct numerical experiments for MP 2c (infinite curvature cusp) without change of variable. Lemma 2.1 of [Aoki and Siegel 2012] gives that the solutions of MPs 2c-1 and 2c-3 are unbounded, and MP 2c-2 is the supplementary contact angle case with unknown behavior.

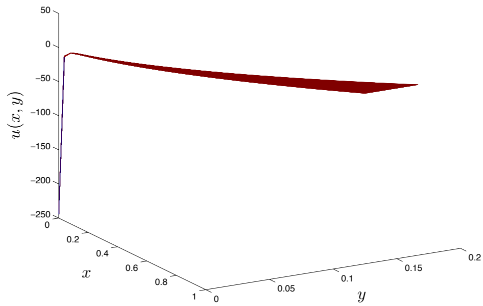
As can be seen in Figure 15, the numerical solution surface is bounded if the contact angles are supplementary for this case, where the boundaries forming a cusp have infinite curvature at the cusp. We have conducted various other numerical experiments; however, we were not able to find any evidence of unbounded solutions if the contact angles are supplementary angles. Guided by these numerical results we conjecture that the solution of the Laplace–Young equation in a domain with a cusp is always bounded if the contact angles of the boundaries forming the cusp are supplementary angles. We also note that, as an additional check on the validity of our numerical approach, we have conducted further numerical experiments with cusps with finite curvature boundaries and with the same contact angles as MPs 2c,




(a) MP 2c-1: $\gamma_1 + \gamma_2 = \pi - \frac{1}{180}\pi$




(b) MP 2c-2: $\gamma_1 + \gamma_2 = \pi$ (supplementary angles)




(c) MP 2c-3: $\gamma_1 + \gamma_2 = \pi + \frac{1}{180}\pi$

**Figure 15.** MPs 2c-1, 2c-2, and 2c-3 (infinite curvature cusp). It is known that the solutions for 2c-1 and 2c-3 are unbounded, but the behavior for 2c-2 is an open problem. FVEM solutions on an $(s, t)$-type grid (Figure 7(a)) with $33 \times 65$ nodes, with change of variable. Surface plots of the capillary surfaces are shown. The numerical result for MP 2c-2 supports the conjecture that the solution is bounded in this case.

and we have confirmed numerically the theoretical prediction that the solution is bounded for supplementary contact angles, and unbounded otherwise.

To conclude, we conjecture that the capillary surface in a cusp domain is bounded if the contact angles of the boundaries forming the cusp are supplementary angles, even if the curvatures of the boundaries are infinite. This conjecture on the open problem of the asymptotic behavior of capillary surfaces in domains with a cusp and supplementary contact angles, obtained from numerical investigation, can guide further analytical study of this case.

## 6. Conclusion

We have proposed a methodology for the numerical study of unbounded capillary surfaces in domains with a sharp corner or a cusp. The methodology was developed by incorporating knowledge obtained from asymptotic analysis into a finite element based approximation method. It contains two simple but important ingredients that are combined with the finite volume element method (FVEM) [Bank and Rose 1987; Aoki and De Sterck 2011] or the Galerkin finite element method (FEM) [Strang and Fix 1973; Brenner and Scott 1994]. The first ingredient is to consider a change of variable, with the new solution variable being smoother than the capillary height variable and more amenable to accurate numerical approximation. The second ingredient is to solve the PDE numerically in a new coordinate system that is inspired by asymptotic analysis work, which allows us to accurately represent the discontinuous behavior of the new solution variable at the singular point. We have shown in extensive numerical tests in domains with a sharp corner or a cusp that this approach leads to a global approximation method for singular solutions of the Laplace–Young equation that recovers the proper asymptotic behavior, and is more accurate and has better convergence properties than numerical methods that were considered for singular capillary surfaces before [Scott et al. 2005]. Although we have only considered the Laplace–Young equation and its steep slope approximation, it is likely that the methodology we have proposed can also be useful for other nonlinear elliptic PDEs with singularities. One important limitation of our approach is that in its present form it only works for problems with one singular point. Extension to problems with multiple singular points is a subject for further research.

The main mathematical contribution of this paper is that we were able to formulate conjectures for two open problems on the asymptotic behavior of capillary surfaces in domains with a cusp. These conjectures are derived from numerical investigation of these open problems using the numerical methodology we propose, and they may guide future analytical work on these open problems.

## Acknowledgement

The authors would like to thank Professor Chad Westphal for inspiring discussions, especially for his idea on log-log plots for the horizontal cross-sections of the solution surfaces (Figures 9 and 11).

## References

[Aoki 2007] Y. Aoki, "Analysis of asymptotic solutions for cusp problems in capillarity", Master's thesis, University of Waterloo, 2007, http://hdl.handle.net/10012/3352.

[Aoki and De Sterck 2011] Y. Aoki and H. De Sterck, "Augmented high order finite volume element method for elliptic PDEs in non-smooth domains: convergence study", *J. Comput. Appl. Math.* **235**:17 (2011), 5177–5187. MR 2012m:65377 Zbl 1221.65299

[Aoki and Siegel 2012] Y. Aoki and D. Siegel, "Bounded and unbounded capillary surfaces in a cusp domain", *Pacific J. Math.* **257**:1 (2012), 143–165. MR 2948462 Zbl 1248.35043

[Bank and Rose 1987] R. E. Bank and D. J. Rose, "Some error estimates for the box method", *SIAM J. Numer. Anal.* **24**:4 (1987), 777–787. MR 88j:65235 Zbl 0634.65105

[Brenner and Scott 1994] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Texts in Applied Mathematics **15**, Springer, New York, 1994. MR 95f:65001 Zbl 0804.65101

[Concus and Finn 1970] P. Concus and R. Finn, "On a class of capillary surfaces", *J. Analyse Math.* **23** (1970), 65–70. MR 42 #8765 Zbl 0257.76007

[Concus and Finn 1974] P. Concus and R. Finn, "On capillary free surfaces in a gravitational field", *Acta Math.* **132** (1974), 207–223. MR 58 #32327c Zbl 0382.76005

[Concus and Finn 1989] P. Concus and R. Finn, "Instability of certain capillary surfaces", *Manuscripta Math.* **63**:2 (1989), 209–213. MR 90d:53009 Zbl 0669.49022

[Finn 1986] R. Finn, *Equilibrium capillary surfaces*, Grundlehren der Mathematischen Wissenschaften **284**, Springer, New York, 1986. MR 88f:49001 Zbl 0583.35002

[Grisvard 1985] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics **24**, Pitman, Boston, 1985. MR 86m:35044 Zbl 0695.35060

[Hornung and Mittelmann 1990] U. Hornung and H. D. Mittelmann, "A finite element method for capillary surfaces with volume constraints", *J. Comput. Phys.* **87**:1 (1990), 126–136. MR 90m:76005 Zbl 0687.76015

[King et al. 1999] J. R. King, J. R. Ockendon, and H. Ockendon, "The Laplace–Young equation near a corner", *Quart. J. Mech. Appl. Math.* **52**:1 (1999), 73–97. MR 2000f:76017 Zbl 0932.76012

[Lancaster 2010] K. E. Lancaster, "A proof of the Concus–Finn conjecture", *Pacific J. Math.* **247**:1 (2010), 75–108. MR 2011m:53011 Zbl 05809032

[Miersemann 1993] E. Miersemann, "Asymptotic expansion at a corner for the capillary problem: the singular case", *Pacific J. Math.* **157**:1 (1993), 95–107. MR 93m:35039 Zbl 0796.76020

[Moon and Spencer 1961] P. Moon and D. E. Spencer, *Field theory handbook*, Springer, Berlin, 1961. 2nd ed. in 1988. MR 24 #B2131 Zbl 0097.39403

[Nigro et al. 2000] N. J. Nigro, B. P. Zellmer, D. Shangguan, and P. S. Lee, "A modified finite element method for determining equilibrium capillary surfaces of liquids with specified volumes", *International Journal for Numerical Methods in Fluids* **33**:6 (2000), 833–846. Zbl 0986.76044

[Norbury et al. 2005] J. Norbury, G. C. Sander, and C. F. Scott, "Corner solutions of the Laplace–Young equation", *Quart. J. Mech. Appl. Math.* **58**:1 (2005), 55–71. MR 2006b:76017 Zbl 1064.76020

[Polevikov 1999] V. K. Polevikov, "Methods for the numerical modeling of equilibrium capillary surfaces", *Differ. Uravn.* **35**:7 (1999), 975–981. MR 2000k:76099 Zbl 1064.76579

[Polevikov 2004] V. K. Polevikov, "Methods for numerical modeling of two-dimensional capillary surfaces", *Comput. Methods Appl. Math.* **4**:1 (2004), 66–93. MR 2005i:76045 Zbl 1221.76128

[Scholz 2001] M. Scholz, *Über das Verhalten von Kapillarflächen in Spitzen*, Ph.D. thesis, Universität Leipzig, 2001, http://www.people.imise.uni-leipzig.de/markus.scholz/pdf/p1.pdf.

[Scholz 2004] M. Scholz, "On the asymptotic behaviour of capillary surfaces in cusps", *Z. Angew. Math. Phys.* **55**:2 (2004), 216–234. MR 2005e:76023 Zbl 1058.35059

[Scott et al. 2005] C. F. Scott, G. C. Sander, and J. Norbury, "Computation of capillary surfaces for the Laplace–Young equation", *Quart. J. Mech. Appl. Math.* **58**:2 (2005), 201–212. MR 2006a:76019 Zbl 1072.76014

[Simon 1980] L. Simon, "Regularity of capillary surfaces over domains with corners", *Pacific J. Math.* **88**:2 (1980), 363–377. MR 82d:49043 Zbl 0467.35022

[Strang and Fix 1973] G. Strang and G. J. Fix, *An analysis of the finite element method*, Prentice-Hall, Englewood Cliffs, NJ, 1973. MR 56 #1747 Zbl 0356.65096

Received June 13, 2012. Revised September 18, 2012.

YASUNORI AOKI
DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF WATERLOO
200 UNIVERSITY AVE. WEST
WATERLOO N2L 3G1
CANADA
yaoki@uwaterloo.ca

HANS DE STERCK
DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF WATERLOO
200 UNIVERSITY AVE. WEST
WATERLOO N2L 3G1
CANADA
hdesterck@uwaterloo.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.35

# DUAL $R$-GROUPS OF THE INNER FORMS OF SL($N$)

KUOK FAI CHAO AND WEN-WEI LI

**We study the Knapp–Stein $R$-groups of the inner forms of SL($N$) over a nonarchimedean local field of characteristic zero, by using a restriction from the inner forms of GL($N$). As conjectured by Arthur, these $R$-groups are then shown to be naturally isomorphic to their dual avatars defined in terms of L-parameters. The 2-cocycles attached to $R$-groups can be described as well. The proofs are based on the results of K. Hiraga and H. Saito. We also construct examples to illustrate some new phenomena which do not occur in the case of SL($N$) or classical groups.**



## 1. Introduction

Let $G$ be a connected reductive group over a local field $F$ and $G(F)$ be the locally compact group of the $F$-points of $G$. The study of the tempered representations of $G(F)$ is a crucial ingredient of the monumental work of Harish-Chandra on his Plancherel formula. Denote by $\Pi_{\text{temp}}(G)$ the set of isomorphism classes of irreducible tempered representations, and by $\Pi_{2,\text{temp}}(G)$ its subset of representations which are square-integrable modulo the center. Roughly speaking, elements in $\Pi_{\text{temp}}(G)$ can be obtained as subrepresentations of $I_P^G(\sigma)$, where $P = MU$ is a parabolic subgroup, $\sigma \in \Pi_{2,\text{temp}}(M)$, and $I_P^G(\sigma)$ is the normalized parabolic induction. Assuming the knowledge of square-integrable representations, the study of $\Pi_{\text{temp}}(G)$ then boils down to that of the decomposition of $I_P^G(\sigma)$, for $P$ and $\sigma$ as above.

*MSC2010:* primary 22E50; secondary 11F70.
*Keywords:* R-group, local Langlands correspondence, intertwining operator.

Knapp, Stein, and Silberger (for the nonarchimedean case) described the decomposition of $I_P^G(\sigma)$ in terms of the *Knapp–Stein R-group* $R_\sigma$. More precisely, we have a central extension of groups

$$1 \to \mathbb{C}^\times \to \widetilde{R}_\sigma \to R_\sigma \to 1$$

defined using the normalized intertwining operators $R_P(w, \sigma)$. It is the set $\Pi_-(\widetilde{R}_\sigma)$ of the irreducible representations of $\widetilde{R}_\sigma$ by which $\mathbb{C}^\times$ acts by $z \mapsto z \cdot \mathrm{id}$ which governs the decomposition of $I_P^G(\sigma)$. Equivalently, we are given a cohomology class $\boldsymbol{c}_\sigma \in H^2(R_\sigma, \mathbb{C}^\times)$ attached to this central extension. The group $R_\sigma$ itself suffices to determine whether $I_P^G(\sigma)$ is reducible or not. To extract further information, such as the description of elliptic tempered representations, some knowledge about $\widetilde{R}_\sigma$ is also needed. We refer the reader to [Arthur 1993, Section 2] for details.

On the other hand, the tempered part of the local Langlands correspondence predicts a map $\phi \mapsto \Pi_\phi$ which assigns a finite subset $\Pi_\phi$ of $\Pi_{\mathrm{temp}}(G)$ to every bounded L-parameter $\phi \in \Phi_{\mathrm{bdd}}(G)$, taken up to equivalence, such that

$$\Pi_{\mathrm{temp}}(G) = \bigsqcup_{\phi \in \Phi_{\mathrm{bdd}}(G)} \Pi_\phi.$$

The internal structure of the tempered L-packets $\Pi_\phi$ is conjectured to be controlled by the $S$-group $S_\phi := Z_{\widehat{G}}(\mathrm{Im}(\phi))$. More precisely, following [Arthur 2006], one has to introduce a central extension

$$1 \to \widetilde{Z}_\phi \to \widetilde{\mathcal{S}}_\phi \to \mathcal{S}_\phi \to 1$$

of finite groups defined in terms of $S_\phi$. The L-packet $\Pi_\phi$ should be in bijection with a set $\Pi(\widetilde{\mathcal{S}}_\phi, \chi_G)$ of representations of $\widetilde{\mathcal{S}}_\phi$, where $\chi_G$ is a character of $\widetilde{Z}_\phi$ coming from Galois cohomology. The relevant definitions will be reviewed later in this article.

The tempered local Langlands correspondence is expected to behave well under normalized parabolic induction, namely, for $P = MU$ as above and $\phi_M \in \Phi_{\mathrm{bdd}}(M)$, we deduce $\phi \in \Phi_{\mathrm{bdd}}(G)$ by composing $\phi_M$ with the inclusion ${}^{\mathrm{L}}M \to {}^{\mathrm{L}}G$ of L-groups, which is well defined up to conjugacy. Then $\Pi_\phi$ should be the union of the irreducible constituents of $I_P^G(\sigma)$, where $\sigma$ ranges over the elements of $\Pi_{\phi_M}$. A natural question arises: is it possible to describe $R_\sigma$, or even $\widetilde{R}_\sigma$, in terms of the $S$-groups?

For archimedean $F$ this has been answered by Shelstad [1982]; in that case, the extension $\widetilde{R}_\sigma \to R_\sigma$ splits and $R_\sigma$ is abelian of exponent two. For general $F$ of characteristic zero, Arthur proposed a generalization [1989b, Section 7] as follows. For every $\phi \in \Phi_{\mathrm{bdd}}(G)$ coming from $\phi_M \in \Phi_{2,\mathrm{bdd}}(M)$ (that is, a parameter for $M$ which is square-integrable modulo the center), he introduced the *dual R-group* (also known as the *endoscopic R-group*) $R_\phi \simeq \mathcal{S}_\phi / \mathcal{S}_{\phi_M}$ and a subgroup $R_{\phi,\sigma} \subset R_\phi$ for every $\sigma \in \Pi_{\phi_M}$. Arthur conjectures a natural isomorphism

$$R_{\phi,\sigma} \simeq R_\sigma.$$

This has been verified for quasisplit classical groups and unitary groups by Arthur [2013] and Mok [2012], respectively. In their construction of L-packets, the dual $R$-groups play a pivotal role through the *local intertwining relations*; see [Arthur 2013, Chapter 2]. It turns out that in these cases, we have $R_\phi = R_{\phi,\sigma}$ and $\widetilde{R}_\sigma \to R_\sigma$ splits; see [Arthur 2013, Section 6.5]. Similar results were obtained independently in [Ban and Zhang 2005; Goldberg 2011; Ban and Goldberg 2012] for nonarchimedean $F$. For quaternionic unitary groups, see [Hanzer 2004].

We shall assume hereafter that $F$ is a nonarchimedean local field of characteristic zero.

Another good test ground for Arthur's conjectures is the group $\mathrm{SL}(N)$ and its inner forms. Indeed, the case $N = 2$ is the genesis of endoscopy [Labesse and Langlands 1979]; for general $N$, the local Langlands correspondence for the inner forms $G^\sharp$ of $\mathrm{SL}(N)$ is established in [Hiraga and Saito 2012], at least in the tempered case. This is based on the local Langlands correspondence for the inner forms $G$ of $\mathrm{GL}(N)$, which satisfies the following nice properties:

- The L-packets $\Pi_\phi$ for $G$ are all singletons.
- For any parabolic subgroup $P = MU$ and $\sigma \in \Pi_{\mathrm{temp}}(M)$, the induced representation $I_P^G(\sigma)$ is irreducible.

In fact, the latter property holds for all unitary $\sigma$, known as Tadić's property (U0) [Sécherre 2009].

The (tempered) local Langlands correspondence for $G^\sharp$ can be obtained by restriction from $G(F)$ to $G^\sharp(F)$; the procedure is somehow dual to the natural projection of L-groups

$$\mathbf{pr} : {}^{\mathrm{L}}G \twoheadrightarrow {}^{\mathrm{L}}G^\sharp.$$

The same recipe can be applied to any Levi subgroup $M$, with respect to $M^\sharp := M \cap G^\sharp$.

The method of restriction provides a convenient device, but we still have to study the internal structure of L-packets for $G^\sharp$ and their behavior under normalized parabolic induction. For the quasisplit case $G^\sharp = \mathrm{SL}(N)$, such issues can be addressed by the multiplicity-one property of Whittaker models. In that case, the Knapp–Stein $R$-groups are studied in depth in [Gelbart and Knapp 1981; 1982; Shahidi 1983; Tadić 1992; Goldberg 1994]. Roughly speaking, let $\sigma^\sharp \in \Pi_{2,\mathrm{temp}}(M^\sharp)$ which lies in the L-packet $\Pi_{\phi^\sharp}$. We may choose $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ so that $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$. Set $\pi := I_P^G(\sigma)$, which is irreducible. Then $R_{\sigma^\sharp}$ is described in terms of

$$X^G(\pi) = \{\eta \in (G(F)/G^\sharp(F))^D : \eta \otimes \pi \simeq \pi\}$$

and its analogue $X^M(\sigma)$ for the Levi subgroup $M$ with respect to $M^\sharp$, where $(G(F)/G^\sharp(F))^D$ means the group of continuous characters of $G(F)/G^\sharp(F)$. It is then easy to relate $R_{\sigma^\sharp}$ with $R_{\phi^\sharp}$, and we deduce a canonical isomorphism $R_{\phi^\sharp} = R_{\phi^\sharp,\sigma^\sharp} \simeq R_{\sigma^\sharp}$ as well as a splitting for $\widetilde{R}_{\sigma^\sharp} \twoheadrightarrow R_{\sigma^\sharp}$. Note that we used the notations $\phi^\sharp$, $\sigma^\sharp$, etc. to denote the objects attached to $G^\sharp$ and its Levi subgroups.

Whittaker models are no longer available for the nonquasisplit inner forms $G^\sharp$ of $\mathrm{SL}(N)$. What saves the day is that Hiraga and Saito [2012] defined a central extension

$$1 \to \mathbb{C}^\times \to S^G(\pi) \to X^G(\pi) \to 1$$

and related it to the central extension of $\mathcal{S}$-groups alluded to above. This allows us to study the internal structure of the L-packets obtained by restriction. In our main theorem, Theorem 6.2.4, we will prove, among other things, that there is

(i) a canonical isomorphism $R_{\phi^\sharp,\sigma^\sharp} \simeq R_{\sigma^\sharp}$ as conjectured by Arthur;

(ii) a "concrete" description of the dual $R$-groups for $G^\sharp$, namely,

$$R_{\phi^\sharp} \simeq X^G(\pi)/X^M(\sigma),$$
$$R_{\phi^\sharp,\sigma^\sharp} \simeq Z^M(\sigma)^\perp/X^M(\sigma);$$

(iii) a description of the class $\boldsymbol{c}_{\sigma^\sharp} \in H^2(R_{\sigma^\sharp}, \mathbb{C}^\times)$ attached to $\widetilde{R}_{\sigma^\sharp} \twoheadrightarrow R_{\sigma^\sharp}$, in terms of the obstruction for extending the representation $\rho$ of $\widetilde{\mathcal{S}}_{\phi^\sharp_M}$ to the preimage of $R_{\phi,\sigma^\sharp}$ in $\widetilde{\mathcal{S}}_{\phi^\sharp}$.

We refer the reader to Section 4 for unexplained notations. Note that the description of $\boldsymbol{c}_{\sigma^\sharp}$ is also conjectured by Arthur; see [Arthur 1996, Page 537; 2008, Section 3; 2013, Section 2.4] for further discussions.

Arthur's conjecture on $R$-groups for the inner forms of $\mathrm{SL}(N)$ is thus verified. The examples are probably more interesting, however. In Section 6.3 we will give conceptual constructions of $\phi^\sharp$ and $\sigma^\sharp$ as above such that

(i) $\widetilde{R}_{\sigma^\sharp} \twoheadrightarrow R_{\sigma^\sharp}$ is not split, or

(ii) $R_{\phi^\sharp,\sigma^\sharp} \subsetneq R_{\phi^\sharp}$.

Such phenomena do not occur to the quasisplit classical groups, the quaternionic unitary groups, or $\mathrm{SL}(N)$. The first example is perhaps more surprising, since $\widetilde{R}_{\sigma^\sharp} \twoheadrightarrow R_{\sigma^\sharp}$ always splits for generic inducing data. Keys [1987, Section 6] constructed a Knapp–Stein $R$-group with nonsplit cocycle in the nonconnected setting; our example seems to be the only known case for connected reductive groups. In both examples, the relation between $R_{\sigma^\sharp}$ and the $S$-groups is crucial.

In view of the possible applications to automorphic representations, one should also consider certain nontempered unitary representations, namely, those appearing in the *A-packets*; see Remark 6.2.6 for a short discussion.

Shortly after the release of the first version of our preprint, we were informed of an independent work of Choiy and Goldberg [2012] that treats the same problems except that of cocycles. Despite some overlap, their work has a completely different technical core, namely the transfer of Plancherel measures between inner forms, which should have wide-ranging applications.

*Organization of this article.* In Section 2, we recapitulate the formalism of normalized intertwining operators and Knapp–Stein $R$-groups. We follow the notations in [Arthur 1989a; 1993] closely. In particular, the $R$-group $R_\sigma$ is regarded as a quotient of the isotropy group $W_\sigma$, instead of a subgroup.

In Section 3, we set up a general formalism of restriction of representations. These results are scattered in [Shahidi 1983; Keys 1987; Tadić 1992; Hiraga and Saito 2012], just to mention a few. In view of the possible sequels of this work, the behavior under restriction of normalized intertwining operators is treated in generality.

A special assumption is made in Section 4 (Hypothesis 4.0.2), namely, that the parabolically induced representations in question should be irreducible. We are then able to deduce finer information on $R$-groups and their cocycles in this setting. The arguments are not too difficult, but require some careful manipulations.

In Section 5, we will specialize to the inner forms of SL($N$) and reformulate the results of Hiraga and Saito [2012] on the local Langlands correspondence and the $S$-groups. In order to study parabolic induction, we also have to generalize these results to the Levi subgroups.

In Section 6, we recapitulate Arthur's definition of dual $R$-groups via the omnipresent commutative diagram in Proposition 6.1.1. The results obtained earlier can then be easily assembled, and Arthur's conjecture on $R$-groups for the inner forms of SL($N$) (Theorem 6.2.4) follows.

## 2. Preliminaries

**2.1.** ***Conventions.*** *Local fields.* Throughout the article, $F$ always denotes a nonarchimedean local field of characteristic zero. We set.

- $\Gamma_F$: the absolute Galois group of $F$, defined with respect to a chosen algebraic closure $\bar{F}$;
- $W_F$: the Weil group of $F$;
- $\mathrm{WD}_F := W_F \times \mathrm{SU}(2)$: the Weil–Deligne group of $F$;
- $|\cdot| = |\cdot|_F$: the normalized absolute value of $F$;
- $q_F$: the cardinality of the residue field of $F$.

When discussing the canonical family of normalizing factors for GL($N$) and its inner forms, we will also fix a nontrivial additive character $\psi_F : F \to \mathbb{C}^\times$.

The usual Galois cohomology over $F$ is denoted by $H^\bullet(F, \cdot)$. The continuous cohomology of $W_F$ is denoted by $H^\bullet_{\mathrm{cont}}(W_F, \cdots)$; the groups of continuous cocycles are denoted by $Z^\bullet_{\mathrm{cont}}(W_F, \cdots)$.

*Groups and representations.* For an $F$-group scheme $G$, the group of its $F$-points is denoted by $G(F)$; subgroups of $G$ mean the closed subgroup schemes. The identity connected component of $G$ is denoted by $G^0$. The center of $G$ is denoted by $Z_G$. Centralizers (respectively normalizers) in $G$ are denoted by $Z_G(\cdot)$ (respectively $N_G(\cdot)$). The algebraic groups over $\mathbb{C}$ are identified with their $\mathbb{C}$-points.

The derived group of $G$ is denoted by $G_{\mathrm{der}}$. Now assume $G$ to be connected reductive. A simply connected cover of $G_{\mathrm{der}}$, which is unique up to isomorphism, is denoted by $G_{\mathrm{SC}} \to G_{\mathrm{der}}$. We denote the adjoint group of $G$ by $G_{\mathrm{AD}} := G/Z_G$. For every subgroup $H$ of $G$, we denote by $H_{\mathrm{sc}}$ (respectively $H_{\mathrm{ad}}$) the preimage of $H$ in $G_{\mathrm{SC}}$ (respectively image in $G_{\mathrm{AD}}$). The same formalism pertains to connected reductive $\mathbb{C}$-groups as well.

The definitions of the dual group ${}^{\mathrm{L}}G = \widehat{G} \rtimes W_F$ and the L-parameters will be reviewed in Section 3.5.

The symbol $\mathrm{Ad}(\cdots)$ denotes the adjoint action of an abstract group on itself, namely $\mathrm{Ad}(x) : g \mapsto gxg^{-1}$.

For any division algebra $D$ over $F$ and $n \in \mathbb{Z}_{\geq 1}$, we denote by $\mathrm{GL}_D(n)$ the group of invertible elements in $\mathrm{End}_D(D^n)$, where $D^n$ is viewed as a right $D$-module. It is also regarded as a connected reductive $F$-group.

The representations considered in this article are all over $\mathbb{C}$-vector spaces. For a connected reductive $F$-group $G$, we define the following.

- $\Pi(G)$: the set of equivalence classes of irreducible smooth representations of $G(F)$;
- $\Pi_{\mathrm{unit}}(G)$: the subset consisting of unitary (that is, unitarizable) representations;
- $\Pi_{\mathrm{temp}}(G)$: the subset consisting of tempered representations;
- $\Pi_{2,\mathrm{temp}}(G)$: the subset consisting of unitary representations which are square-integrable modulo the center.

For an abstract group $S$, we will also denote by $\Pi(S)$ the set of its irreducible representations up to equivalence.

The central character of $\pi \in \Pi(G)$ is denoted by $\omega_\pi$. The group of morphisms (respectively the set of isomorphisms) in the category of representations of $G(F)$ is denoted by $\mathrm{Hom}_G(\cdots)$ (respectively $\mathrm{Isom}_G(\cdots)$).

For any topological group $H$, we set

$$H^D := \{\chi : H \to \mathbb{C}^\times, \text{ continuous character}\}.$$

For any representation $\pi$ of $G(F)$ and any $\eta \in G(F)^D$, we write $\eta\pi := \eta \otimes \pi$ for abbreviation. Also note that $\pi$ and $\eta\pi$ have the same underlying $\mathbb{C}$-vector spaces. If $M$ is a subgroup of $G$ and $\pi$ is a smooth representation of $G(F)$, we shall denote the restriction of $\pi$ to $M(F)$ by $\pi|_M$.

*Combinatorics.* Let $G$ be a connected reductive $F$-group. We employ the following notations in this article. Let $M$ be a Levi subgroup.

- $\mathcal{P}(M)$: the set of parabolic subgroups of $G$ with Levi component $M$;
- $\mathcal{L}(M)$: the set of Levi subgroups of $G$ containing $M$;
- $\mathcal{F}(M)$: the set of parabolic subgroups of $G$ containing $M$;
- $W(M) := N_G(M)(F)/M(F)$: the Weyl group (in a generalized sense) relative to $M$;

The Levi decompositions are written as $P = MU$, where $U$ denotes the unipotent radical of $P$. For $M$ chosen, the opposite parabolic of $P = MU$ is denoted by $\bar{P} = M\bar{U}$. When we have to emphasize the role of $G$, the notations $\mathcal{P}^G(M)$, $\mathcal{L}^G(M)$, $\mathcal{F}^G(M)$, and $W^G(M)$ will be used.

Let $w \in W(M)$ with a representative $\tilde{w} \in G(F)$. For $\sigma \in \Pi(M)$, we define $\tilde{w}\sigma$ to be the representation on the same underlying vector space, with the new action

$$(\tilde{w}\sigma)(m) := \sigma(\tilde{w}^{-1} m \tilde{w}), \quad m \in M(F).$$

The equivalence class of $\tilde{w}\sigma$ depends only on $w \in W(M)$, and we will write $w\sigma$ instead, if there is no confusion.

Define $\mathcal{X}(G) := \mathrm{Hom}_{F-\mathrm{grp}}(G, \mathbb{G}_\mathrm{m})$ and $\mathfrak{a}_G := \mathrm{Hom}(\mathcal{X}(G), \mathbb{R})$. For every Levi subgroup $M$, there is a canonically split short exact sequence of finite-dimensional $\mathbb{R}$-vector spaces

$$0 \to \mathfrak{a}_G \to \mathfrak{a}_M \leftrightarrows \mathfrak{a}_M^G \to 0.$$

The linear duals of these spaces are denoted by $\mathfrak{a}_G^*$, etc. We also write $\mathfrak{a}_{G,\mathbb{C}} := \mathfrak{a}_G \otimes_\mathbb{R} \mathbb{C}$, etc. Sometimes we also write $\mathfrak{a}_P$ instead of $\mathfrak{a}_M$ if $P = MU$.

The Harish-Chandra map $H_G : G(F) \to \mathfrak{a}_G$ is the homomorphism characterized by

$$\langle \chi, H_G(x) \rangle = \log |\chi(x)|_F, \quad \chi \in \mathcal{X}(G).$$

For $\lambda \in \mathfrak{a}_{G,\mathbb{C}}^*$ and $\pi \in \Pi(G)$, we define $\pi_\lambda \in \Pi(G)$ by

$$\pi_\lambda := e^{\langle \lambda, H_G(\cdot) \rangle} \otimes \pi. \tag{1}$$

Fix a minimal parabolic subgroup $P_0 = M_0 U_0$ of $G$. We define $\Delta_0$, $\Delta_0^\vee$ to be the set of simple roots and coroots, which form bases of $(\mathfrak{a}_{M_0}^G)^*$ and $\mathfrak{a}_{M_0}^G$, respectively. The set of positive roots is denoted by $\Sigma_0$, and its subset of reduced roots by $\Sigma_0^{\mathrm{red}}$. They form a bona fide root system. For every $P = MU \supset P_0$, we define

$\Delta_P \subset \Sigma_P \subset \Sigma_P^{\mathrm{red}}$ by taking the set of nonzero restrictions to $(\mathfrak{a}_M^G)^*$ of elements in $\Delta_0 \subset \Sigma_0 \subset \Sigma_0^{\mathrm{red}}$. To each $\alpha \in \Sigma_P$ we may associate the coroot $\alpha^\vee \in \mathfrak{a}_M^G$: it is defined as the restriction of the coroot in $\Delta_0^\vee$. For a given $P$, the objects above are independent of the choice of $P_0$. We can emphasize the role of $G$ by using the notations $\Delta_P^G$, etc. whenever needed.

*Induction.* We always consider a parabolic subgroup $P = MU$ of $G$. The modulus character of $P(F)$ is denoted by $\delta_P$, that is,

$$(\text{left Haar measure}) = \delta_P \cdot (\text{right Haar measure}).$$

The usual smooth induction functor is denoted by $\mathrm{Ind}(\cdots)$. The normalized parabolic induction functor from $P$ to $G$ is denoted by $I_P^G(\,\cdot\,) := \mathrm{Ind}_P^G(\delta_P^{1/2} \otimes \cdot\,)$. Recall that for $\sigma \in \Pi(M)$ with underlying vector space $V_\sigma$, we use the usual model to realize $I_P^G(\sigma)$ as the space of functions $\varphi : G(F) \to V_\sigma$ such that $\varphi$ is invariant under right translation by an open compact subgroup of $G(F)$, and that $\varphi(umx) = \delta_P(m)^{1/2}\sigma(m)(\varphi(x))$ for all $m \in M(F)$, $u \in U(F)$. The group $G(F)$ acts on this function space by the right regular representation.

For $\sigma, \sigma' \in \Pi(M)$ and $f \in \mathrm{Hom}_M(\sigma, \sigma')$, the induced morphism is denoted by $I_P^G(f)$; it sends $\varphi$ to $f(\varphi)$.

**2.2.** ***Normalized intertwining operators.*** Our basic reference for normalized intertwining operators is [Arthur 1993]. Consider the following data.

- $G$: a connected reductive $F$-group.
- $M$: a Levi subgroup of $G$.
- $P, Q \in \mathcal{P}(M)$.
- $\sigma : M(F) \to \mathrm{Aut}_{\mathbb{C}}(V_\sigma)$: a smooth representation of $M(F)$ of finite length.
- $\lambda \in \mathfrak{a}_{M,\mathbb{C}}^*$.

For every $\alpha \in \Delta_P$, we denote by $r_\alpha$ the smallest positive rational number such that $r_\alpha \cdot \alpha^\vee$ lies in the lattice $H_M(M(F))$. We define

$$\check{\alpha} := r_\alpha \alpha^\vee. \tag{2}$$

By recalling (1), we form the normalized parabolic induction $I_P^G(\sigma_\lambda)$, $I_Q^G(\sigma_\lambda)$. Their underlying spaces are denoted by $I_P^G(V_{\sigma_\lambda})$, $I_Q^G(V_{\sigma_\lambda})$. The standard intertwining operator

$$J_{Q|P}(\sigma_\lambda) : I_P^G(\sigma_\lambda) \longrightarrow I_Q^G(\sigma_\lambda)$$

is defined by the absolutely convergent integral

$$(J_{Q|P}(\sigma_\lambda)\varphi)(x) = \int_{U_P(F)\cap U_Q(F)\backslash U_Q(F)} \varphi(ux)\,du, \quad x \in G(F), \tag{3}$$

when $\langle \mathrm{Re}(\lambda), \alpha^\vee \rangle \gg 0$ for all $\alpha \in \Sigma_P^{\mathrm{red}} \cap \Sigma_{\bar{Q}}^{\mathrm{red}}$; see [Waldspurger 2003, IV.1] for the precise meaning of absolute convergence. Recall that upon choosing a special maximal compact open subgroup $K \subset G(F)$ in good position relative to $M$, these induced representations can be realized on a vector space that is independent of $\lambda$. It is known that $J_{Q|P}(\sigma_\lambda)$ is a rational function in the variables

$$\{q_F^{-\langle \lambda, \check{\alpha} \rangle} : \alpha \in \Delta_P\}.$$

In particular, as a function in $\lambda$, $J_{Q|P}(\sigma_\lambda)$ admits a meromorphic continuation to $\mathfrak{a}^*_{M,\mathbb{C}}$. When $\sigma \in \Pi_{\mathrm{temp}}(M)$, it is known that (3) is absolutely convergent for $\langle \mathrm{Re}(\lambda), \alpha^\vee \rangle > 0$ for all $\alpha \in \Sigma_P^{\mathrm{red}} \cap \Sigma_{\bar{Q}}^{\mathrm{red}}$. Moreover, as a meromorphic family of operators, it satisfies $\mathrm{ord}_{\lambda=0}(J_{Q|P}(\sigma_\lambda)) \geq -1$.

Henceforth we assume $\sigma$ irreducible, that is, $\sigma \in \Pi(M)$. Take any $P \in \mathcal{P}(M)$. Define the $j$-functions as

$$j(\sigma_\lambda) := J_{P|\bar{P}}(\sigma_\lambda) J_{\bar{P}|P}(\sigma_\lambda). \tag{4}$$

It is known that $\lambda \mapsto j(\sigma_\lambda)$ a scalar-valued meromorphic function, which is not identically zero. Moreover, $j(\sigma_\lambda)$ is independent of $P$ and admits a product decomposition

$$j(\sigma_\lambda) = \prod_{\alpha \in \Sigma_P^{\mathrm{red}}} j_\alpha(\sigma_\lambda),$$

where $j_\alpha$ denotes the $j$-function defined relative to the Levi subgroup $M_\alpha \in \mathcal{L}(M)$ such that $\Sigma_M^{M_\alpha, \mathrm{red}} = \{\pm\alpha\}$.

Now assume $\sigma \in \Pi_{2,\mathrm{temp}}(M)$. In this paper, we define Harish-Chandra's $\mu$-function as the meromorphic function

$$\mu(\sigma_\lambda) := j(\sigma_\lambda)^{-1}.$$

Accordingly, $\mu$ also admits a product decomposition $\mu = \prod_\alpha \mu_\alpha$. It is analytic and nonnegative for $\lambda \in i\mathfrak{a}^*_M$. Note that our definitions of $j$-functions and $\mu$-functions depend on the choice of Haar measures on unipotent radicals. In particular, our $\mu$-function differs from that in [Waldspurger 2003, V.2] by some harmless constant.

**Definition 2.2.1** (cf. [Arthur 1989a, Section 2]). In this article, a family of normalizing factors is a family of meromorphic functions on the $\mathfrak{a}^*_{M,\mathbb{C}}$-orbits in $\Pi(M)$, for all Levi subgroup $M$ of $G$, written as

$$r_{Q|P}(\sigma_\lambda), \quad P, Q \in \mathcal{P}(M), \sigma_\lambda \in \mathfrak{a}^*_{M,\mathbb{C}}$$

satisfying the following conditions. First of all, we define the corresponding normalized intertwining operators as

$$R_{Q|P}(\sigma_\lambda) := r_{Q|P}(\sigma_\lambda)^{-1} J_{Q|P}(\sigma_\lambda),$$

which is a meromorphic family (in $\lambda$) of intertwining operators $I_P^G(\sigma_\lambda) \to I_Q^G(\sigma_\lambda)$.

We shall also assume that a family of normalizing factors is chosen for every proper Levi subgroup.

(R$_1$) For all $P$, $P'$, $P'' \in \mathscr{P}(M)$, we have $R_{P''|P}(\sigma_\lambda) = R_{P''|P'}(\sigma_\lambda) R_{P'|P}(\sigma_\lambda)$.

(R$_2$) If $\sigma \in \Pi_{\mathrm{unit}}(M)$, then

$$R_{Q|P}(\sigma_\lambda) = R_{P|Q}(\sigma_{-\bar{\lambda}})^*, \quad \lambda \in \mathfrak{a}^*_{M,\mathbb{C}}.$$

In particular, $R_{Q|P}(\sigma)$ is a well-defined unitary operator.

(R$_3$) This family is compatible with conjugacy, namely,

$$R_{gQg^{-1}|gPg^{-1}}(g\sigma_\lambda) = \ell(g) R_{Q|P}(\sigma_\lambda) \ell(g)^{-1}$$

for all $g \in G(F)$, where $\ell(g)$ is the map $\varphi(\,\cdot\,) \mapsto \varphi(g^{-1}\cdot\,)$.

(R$_4$) We have

$$r_{Q|P}(\sigma_\lambda) = \prod_{\alpha \in \Sigma_P^{\mathrm{red}} \cap \Sigma_{\bar{Q}}^{\mathrm{red}}} r^{\widetilde{M}_\alpha}_{\bar{P}_\alpha|P_\alpha}(\sigma_\lambda),$$

where $P_\alpha := P \cap M_\alpha$, and $r^{M_\alpha}_{\bar{P}_\alpha|P_\alpha}$ comes from the family of normalizing factors for $M_\alpha$.

(R$_5$) Let $S = LU \in \mathscr{F}(M)$ containing both $P$ and $Q$. Then $R_{Q|P}(\sigma_\lambda)$ is the operator deduced from $R^L_{P\cap L|Q\cap L}(\sigma_\lambda)$ by the functor $I_S^G(\,\cdot\,)$.

(R$_6$) The function $\lambda \mapsto r_{Q|P}(\sigma_\lambda)$ is rational in the variables $\{q_F^{-\langle\lambda,\check{\alpha}\rangle} : \alpha \in \Delta_P\}$.

(R$_7$) If $\sigma \in \Pi_{\mathrm{temp}}(M)$, the meromorphic function $\lambda \mapsto r_{Q|P}(\sigma_\lambda)$ is invertible whenever $\mathrm{Re}\langle\lambda, \alpha^\vee\rangle > 0$ for all $\alpha \in \Delta_P$.

Observe that (R$_2$) is equivalent to saying $r_{Q|P}(\sigma_\lambda) = \overline{r_{P|Q}(\sigma_{-\bar{\lambda}})}$ for $\sigma \in \Pi_{\mathrm{unit}}(M)$, as the unnormalized operators $J_{Q|P}(\sigma_\lambda)$ satisfy a similar condition. Similarly, (R$_3$) is equivalent to saying $r_{gQg^{-1}|gPg^{-1}}(g\sigma_\lambda) = r_{Q|P}(\sigma_\lambda)$.

The fundamental result about the normalizing factors is that they exist [Arthur 1989a, Theorem 2.1].

**Remark 2.2.2.** According to Langlands [1976, Appendix 2], there is a conjectural canonical family of normalizing factors $r_{Q|P}(\sigma_\lambda)$ in terms of local factors, namely,

$$r_{Q|P}(\sigma_\lambda) = \varepsilon(0, \rho^\vee_{Q|P} \circ \phi_\lambda, \psi_F)^{-1} L(0, \phi_\lambda, \rho^\vee_{Q|P}) L(1, \phi_\lambda, \rho^\vee_{Q|P})^{-1},$$

where

- $\phi_\lambda$ is the Langlands parameter for $\sigma_\lambda$;
- $\hat{\mathfrak{u}}_Q$ (resp. $\hat{\mathfrak{u}}_P$) denotes the Lie algebra of the unipotent radical of the dual parabolic subgroup $\hat{Q}$ (respectively $\hat{P}$) in $\widehat{G}$;

- $\rho_{Q|P}$ is the adjoint representation of ${}^{\mathrm{L}}M$ on $\hat{\mathfrak{u}}_Q/(\hat{\mathfrak{u}}_Q \cap \hat{\mathfrak{u}}_P)$, and $\rho^\vee_{Q|P}$ denotes its contragredient;
- $\psi_F : F \to \mathbb{C}^\times$ is a chosen nontrivial additive character.

We will invoke this description only in the case $G = \mathrm{GL}_F(n)$. In that case, the local factors in sight are essentially those associated with pairs $(\phi_1, \phi_2)$ where $\phi_1, \phi_2$ are among the L-parameters parametrizing the components of $\sigma$. Such Artin local factors are known to agree with their representation-theoretic avatars, say, those defined by Rankin–Selberg convolution or by the Langlands–Shahidi method.

**Remark 2.2.3.** The construction of normalizing factors can be reduced to the case where $M$ is a maximal proper Levi subgroup of $G$ and $\sigma \in \Pi_{2,\mathrm{temp}}(M)$, as illustrated in [Arthur 1989a]. Let us give a quick sketch of this reduction.

(i) In view of (R$_4$), we are led to the case $M$ maximal proper. Moreover, it suffices to verify (R$_3$) for the representatives in $G(F)$ of the elements in $W(M) := N_{G(F)}(M)/M(F)$, which has at most two elements.

(ii) Assume that $\sigma \in \Pi_{\mathrm{temp}}(M)$. By the classification of tempered representations, there exists a parabolic subgroup $R = M_R U_R$ of $M$ and $\tau \in \Pi_{\mathrm{temp}}(M_R)$ such that $\sigma \hookrightarrow I_R^M(\tau)$. The pair $(M, \tau)$ is unique up to conjugacy. There is a unique element $P(R)$ in $\mathcal{P}(M_R)$, characterized by the properties

- $P(R) \subset P$,
- $P(R) \cap M = R$.

Consequently, parabolic induction in stages gives $I_P^G I_R^M(\tau) = I_{P(R)}^G(\tau)$. The same construction works when $P$ is replaced by $Q$. Set

$$r_{Q|P}(\sigma_\lambda) := r_{Q(R)|P(R)}(\tau_\lambda).$$

In view of (R$_5$) together with parabolic induction in stages, we see that $R_{Q|P}(\sigma_\lambda)$ is the restriction of $R_{Q(R)|P(R)}(\sigma_\lambda)$ to $I_P^G(\sigma_\lambda)$. The required conditions can be readily verified.

(iii) For general $\sigma$, we may realize it as the Langlands quotient $I_R^M(\tau_\mu) \twoheadrightarrow \sigma$, where $R = M_R U_R$ is a parabolic subgroup of $M$, $\tau \in \Pi_{\mathrm{temp}}(M_R)$, and $\mu \in \mathfrak{a}^*_{M_R}$ satisfies $\mathrm{Re}\langle \mu, \beta^\vee \rangle > 0$ for all $\beta \in \Delta_R^M$. The triplet $(M, \tau, \mu)$ is again unique up to conjugacy. Let $P, Q \in \mathcal{P}(M)$. Define $P(R), Q(R) \in \mathcal{P}(M_R)$ as before and set

$$r_{Q|P}(\sigma_\lambda) := r_{Q(R)|P(R)}(\tau_{\lambda+\mu}).$$

Recall that

$$\mathrm{Ker}[I_R^M(\tau_\mu) \twoheadrightarrow \sigma] = \mathrm{Ker}(J^M_{\bar{R}|R}(\tau_\mu)).$$

The condition (R$_7$) applied to $\tau_\mu$ tells us that $\mathrm{Ker}(J^M_{\bar{R}|R}(\tau_\mu)) = \mathrm{Ker}(R^M_{\bar{R}|R}(\tau_\mu))$. The same is true for $\mu$ replaced by $\mu + \lambda$. Using (R$_5$), we see that $R_{Q(R)|P(R)}(\tau_{\lambda+\mu})$

factors into $R_{Q|P}(\sigma_\lambda)$ on $I_P^G(\sigma_\lambda)$. All conditions except (R$_2$) follow from this. The proof of (R$_2$) requires more effort to deal with the unitarizability of Langlands quotients; the reader can consult [Arthur 1989a, Page 30] for details.

(iv) Reverting to our original assumption that $M$ is maximal proper and

$$\sigma \in \Pi_{2,\mathrm{temp}}(M),$$

it clearly remains to verify conditions (R$_1$), (R$_2$), (R$_3$), (R$_6$), (R$_7$). Furthermore, one can reduce (R$_3$) to the assertion that $r_{w\bar{P}w^{-1}|wPw^{-1}}(w(\sigma_\lambda)) = r_{\bar{P}|P}(\sigma_\lambda)$, for $w \in W(G)$ being the nontrivial element in $W(M)$ if it exists.

**2.3.** ***Knapp–Stein R-groups.*** Fix a family of normalizing factors for $G$. Assume henceforth that $M$ is a Levi subgroup of $G$ and $\sigma \in \Pi_{2,\mathrm{temp}}(M)$. Define the isotropy group

$$W_\sigma := \{w \in W(M) : w\sigma \simeq \sigma\}.$$

Fix $P \in \mathcal{P}(M)$. For $w \in W(M)$ with a representative $\tilde{w} \in G(F)$, we define the operator $r_P(\tilde{w}, \sigma) \in \mathrm{Isom}_G(I_P^G(\sigma), I_P^G(\tilde{w}\sigma))$ by

$$r_P(\tilde{w}, \sigma) : I_P^G(\sigma) \xrightarrow{R_{w^{-1}Pw|P}(\sigma)} I_{w^{-1}Pw}^G(\sigma) \xrightarrow{[\ell(\tilde{w}):\phi \mapsto \phi(\tilde{w}^{-1}\cdot)]} I_P^G(\tilde{w}\sigma). \tag{5}$$

We notice the following property: for any $w, w' \in W(M)$ with representatives $\tilde{w}, \tilde{w}' \in G(F)$, we have

$$r_P(\tilde{w}\tilde{w}', \sigma) = r_P(\tilde{w}, \tilde{w}'\sigma) \circ r_P(\tilde{w}', \sigma). \tag{6}$$

Assume now $w \in W_\sigma$. Choose a representative $\tilde{w}$ of $w$ and $\sigma(\tilde{w}) \in \mathrm{Isom}_M(\tilde{w}\sigma, \sigma)$ to define the operator

$$R_P(\tilde{w}, \sigma) := I_P^G(\sigma(\tilde{w})) \circ r_P(\tilde{w}, \sigma).$$

The class $R_P(\tilde{w}, \sigma) \bmod \mathbb{C}^\times$ is independent of the choices of $\sigma(\tilde{w})$ and the representative $\tilde{w}$. We also have

$$\begin{gathered} R_P(\tilde{w}, \sigma) \in \mathrm{Aut}_G(I_P^G(\sigma)), \\ R_P(\tilde{w}, \sigma) R_P(\tilde{w}', \sigma) = R_P(\tilde{w}\tilde{w}', \sigma) \mod \mathbb{C}^\times, \quad w, w' \in W_\sigma. \end{gathered}$$

Now we can define the Knapp–Stein $R$-group as follows.

$$\begin{aligned} W_\sigma^0 &:= \{w \in W_\sigma : R_P(\tilde{w}, \sigma) \in \mathbb{C}^\times \mathrm{id}\}, \\ R_\sigma &:= W_\sigma / W_\sigma^0. \end{aligned}$$

We will also make use of the following alternative description of $R_\sigma$.

**Proposition 2.3.1.** *The subgroup $W_\sigma^0$ is the Weyl group of the root system on $\mathfrak{a}_M$ composed of the multiples of the roots in $\{\alpha \in \Sigma_P^{\mathrm{red}} : \mu_\alpha(\sigma) = 0\}$.*

*Given any Weyl chamber $\mathfrak{a}_\sigma^+ \subset \mathfrak{a}_M$ for the aforementioned root system, there is then a unique section $R_\sigma \hookrightarrow W_\sigma$ that sends $r \in R_\sigma$ to the representative $w \in W_\sigma$ such that $w\mathfrak{a}_\sigma^+ = \mathfrak{a}_\sigma^+$. Consequently, we can write $W_\sigma = W_\sigma^0 \rtimes R_\sigma$.*

In the literature, $R_\sigma$ is sometimes viewed as a subgroup of $W_\sigma$ in this manner; see, for example, [Goldberg 2006].

Write $V_\sigma$ (respectively $I_P^G(V_\sigma)$) for the underlying vector space of the representation $\sigma$ (respectively $I_P^G(\sigma)$). It follows that $w \mapsto R_P(\tilde{w}, \sigma)$ induces a projective representation of $R_\sigma$ on $I_P^G(V_\sigma)$, where $\tilde{w} \in G(F)$ is any representative of $w \in W_\sigma$. We denote this projective representation provisionally by $r \mapsto R_P(r, \sigma)$, for $r \in R_\sigma$.

There is a standard way to lift $R_P(\,\cdot\,, \sigma)$ to an authentic representation of some group $\widetilde{R}_\sigma$ which sits in a central extension

$$1 \to \mathbb{C}^\times \to \widetilde{R}_\sigma \to R_\sigma \to 1,$$

such that $\mathbb{C}^\times$ acts by $z \mapsto z \cdot \mathrm{id}$. Namely, we can set $\widetilde{R}_\sigma$ to be the group of elements $(r, M[r]) \in R_\sigma \times \mathrm{Aut}_{\mathbb{C}}(I_P^G(V_\sigma))$ such that $M[r] \bmod \mathbb{C}^\times$ gives $R_P(r, \sigma)$. The lifted representation, denoted by $\tilde{r} \mapsto R_P(\tilde{r}, \sigma)$, is then $\tilde{r} = (r, M[r]) \mapsto M[r]$. Such a central extension by $\mathbb{C}^\times$ that lifts $R_P(\,\cdot\,, \sigma)$ is unique up to isomorphism.

Note that the central extension above can also be described by the class

$$\boldsymbol{c}_\sigma \in H^2(R_\sigma, \mathbb{C}^\times)$$

coming from the $\mathbb{C}^\times$-valued 2-cycocle $c_\sigma$ defined by

$$R_P(\widetilde{r_1 r_2}, \sigma) = c_\sigma(r_1, r_2) R_P(\tilde{r}_1, \sigma) R_P(\tilde{r}_2, \sigma), \quad r_1, r_2 \in R_\sigma, \tag{7}$$

where we choose a preimage $\tilde{r} \in \widetilde{R}_\sigma$ for every $r \in R_\sigma$.

**Theorem 2.3.2** (Harish-Chandra; [Silberger 1978]). *Fix a preimage $\tilde{r} \in \widetilde{R}_\sigma$ for every $r \in R_\sigma$. Then the operators $\{R_P(\tilde{r}, \sigma) : r \in R_\sigma\}$ form a basis of* $\mathrm{End}_G(I_P^G(\sigma))$.

Following Arthur, we reformulate this fundamental result as follows. Let

$$\Pi_\sigma(G) := \{\text{irreducible constituents of } I_P^G(\sigma)\}/\simeq,$$
$$\Pi_-(\widetilde{R}_\sigma) := \{\rho \in \Pi(\widetilde{R}_\sigma) : \text{ for all } z \in \mathbb{C}^\times\ \rho(z) = z \cdot \mathrm{id}\}.$$

Note that $\Pi_\sigma(G)$, $\Pi_-(\widetilde{R}_\sigma)$ are both finite sets, and each $\rho \in \Pi_-(\widetilde{R}_\sigma)$ is finite-dimensional.

**Corollary 2.3.3.** *Let $\mathfrak{R}$ be the representation of $\widetilde{R}_\sigma \times G(F)$ on $I_P^G(V_\sigma)$ defined by*

$$\mathfrak{R}(\tilde{r}, x) = R_P(\tilde{r}, \sigma) I_P^G(\sigma, x), \quad \tilde{r} \in \widetilde{R}_\sigma, x \in G(F).$$

*Then there is a decomposition*

$$\mathcal{R} \simeq \bigoplus_{\rho \in \Pi_-(\widetilde{R}_\sigma)} \rho \boxtimes \pi_\rho, \tag{8}$$

*where $\rho \mapsto \pi_\rho$ is a bijection from $\Pi_-(\widetilde{R}_\sigma)$ to $\Pi_\sigma(G)$, characterized by (8).*

Consequently, $I_P^G(\sigma)$ is irreducible if and only if $R_\sigma = \{1\}$.

**Remark 2.3.4.** When $G$ is quasisplit and $\sigma$ is generic with respect to a given Whittaker datum for $M$, the work of Shahidi [1990] furnishes

(i) a canonical family of normalizing factors $r_{Q|P}(\sigma)$;
(ii) a canonically defined homomorphism $w \mapsto R_P(w, \sigma)$;
(iii) a canonical splitting of the central extension $1 \to \mathbb{C}^\times \to \widetilde{R}_\sigma \to R_\sigma \to 1$.

These properties are not expected in general. Indeed, we see in Example 6.3.3 that (iii) may fail.

**Remark 2.3.5.** The formalism above depends not only on $(M, \sigma)$, but also on the choice of $P \in \mathcal{P}(M)$. One can easily pass to another choice $Q \in \mathcal{P}(M)$ by transport of structure using $R_{Q|P}(\sigma)$. For example, one has

$$r_Q(\tilde{w}, \sigma) = R_{P|Q}(\tilde{w}\sigma)^{-1} r_P(\tilde{w}, \sigma) R_{P|Q}(\sigma),$$
$$R_Q(\tilde{w}, \sigma) = R_{P|Q}(\sigma)^{-1} R_P(\tilde{w}, \sigma) R_{P|Q}(\sigma)$$

for all $w \in W(M)$ with a representative $\tilde{w} \in G(F)$ and some chosen $\sigma(\tilde{w})$.

## 3. Restriction

Let $G$, $G^\sharp$ be connected reductive $F$-groups such that

$$G_{\mathrm{der}} \subset G^\sharp \subset G.$$

**3.1. *Restriction of representations.*** In this subsection, we will review the basic results in [Tadić 1992, Section 2; Hiraga and Saito 2012, Chapter 2] concerning the restriction of a smooth representation from $G(F)$ to $G^\sharp(F)$. The objects associated to $G^\sharp$ are endowed with the superscript $\sharp$, for example, $\pi^\sharp \in \Pi(G^\sharp)$.

**Proposition 3.1.1** [Silberger 1979; Tadić 1992, Lemma 2.1 and Proposition 2.2]. *Let $\pi \in \Pi(G)$. Then $\pi|_{G^\sharp}$ decomposes into a finite direct sum of smooth irreducible representations. Each irreducible constituent of $\pi|_{G^\sharp}$ has the same multiplicity.*

*Conversely, every $\pi^\sharp \in \Pi(G^\sharp)$ embeds into $\pi|_{G^\sharp}$ for some $\pi \in \Pi(G)$. If the central character $\omega_{\pi^\sharp}$ is unitary, one can choose $\pi$ so that $\omega_\pi$ is also unitary.*

**Proposition 3.1.2** [Tadić 1992, Corollary 2.5]. *Let $\pi_1, \pi_2 \in \Pi(G)$. The following are equivalent*:

(i) $\mathrm{Hom}_{G^\sharp}(\pi_1, \pi_2) \neq \{0\}$.

(ii) $\pi_1|_{G^\sharp} \simeq \pi_2|_{G^\sharp}$.

(iii) *There exists* $\eta \in (G(F)/G^\sharp(F))^D$ *such that* $\eta\pi_1 \simeq \pi_2$.

For $\pi \in \Pi(G)$, we define a finite "packet" of smooth irreducible representations of $G^\sharp(F)$ as

$$\Pi_\pi := \{\text{irreducible constituents of } \pi|_{G^\sharp}\}/\simeq .$$

Consequently, Proposition 3.1.2 implies that $\Pi(G^\sharp) = \bigsqcup_\pi \Pi_\pi$, when $\pi$ is taken over the $(G(F)/G^\sharp(F))^D$-orbits in $\Pi(G)$.

**Proposition 3.1.3** [Tadić 1992, Proposition 2.7]. *Let* $\pi \in \Pi(G)$ *and assume that* $\omega_\pi$ *is unitary. Let* $\boldsymbol{P}$ *be one of the following properties of smooth irreducible representations of* $G(F)$ *or* $G^\sharp(F)$:

(i) *unitary*,

(ii) *tempered*,

(iii) *square-integrable modulo the center*,

(iv) *cuspidal.*

*Then we have equivalences of the form*

$$\begin{aligned}[\pi \textit{ satisfies } \boldsymbol{P}] &\iff [\textit{for some } \pi^\sharp \in \Pi_\pi, \pi^\sharp \textit{ satisfies } \boldsymbol{P}] \\ &\iff [\textit{for all } \pi^\sharp \in \Pi_\pi, \pi^\sharp \textit{ satisfies } \boldsymbol{P}].\end{aligned}$$

Now comes the decomposition of $\pi|_{G^\sharp}$.

**Definition 3.1.4.** Let $\pi \in \Pi(G)$ with the underlying $\mathbb{C}$-vector space $V_\pi$. Note that $V_{\eta\pi} = V_\pi$ for all $\eta \in (G(F)/G^\sharp(F))^D$. We introduce the groups

$$\begin{aligned}X^G(\pi) &:= \{\eta \in (G(F)/G^\sharp(F))^D : \eta\pi \simeq \pi\},\\ S^G(\pi) &:= \langle I_\eta^G \in \mathrm{Isom}_G(\eta\pi, \pi) : \eta \in X^G(\pi)\rangle \subset \mathrm{Aut}_{G^\sharp}(\pi).\end{aligned}$$

Observe that $\mathrm{Isom}_G(\eta\pi, \pi)$ is a $\mathbb{C}^\times$-torsor by Schur's lemma, and an element $I_\eta^G \in S^G(\pi)$ uniquely determines $\eta$. The group law is given by composition in $\mathrm{Aut}_{\mathbb{C}}(V_\pi)$, namely, by

$$\begin{aligned}&\mathrm{Isom}_G(\eta\pi, \pi) \times \mathrm{Isom}_G(\eta'\pi, \pi)\\ &\qquad = \mathrm{Isom}_G(\eta'\eta\pi, \eta'\pi) \times \mathrm{Isom}_G(\eta'\pi, \pi) \to \mathrm{Isom}_G(\eta'\eta\pi, \pi)\end{aligned}$$

for all $\eta, \eta' \in X^G(\pi)$.

Thus we obtain a central extension of locally compact groups

$$1 \to \mathbb{C}^\times \to S^G(\pi) \to X^G(\pi) \to 1, \tag{9}$$

where the first arrow is $z \mapsto z \cdot \mathrm{id}$ and the second one is $I_\eta^G \mapsto \eta$.

Also note that $X^G(\pi) = X^G(\xi\pi)$, $S^G(\pi) = S^G(\xi\pi)$ for any $\xi \in (G(F)/G^\sharp(F))^D$.

Note that implicit in the notations above is the reference to $G^\sharp$, which is usually clear from the context. Indications to $G^\sharp$ will be given when necessary.

It is easy to see that $X^G(\pi)$ is finite abelian. As in the setting of $R$-groups, we define the finite set

$$\Pi_-(S^G(\pi)) := \{\rho \in \Pi(S^G(\pi)) : \text{ for all } z \in \mathbb{C}^\times, \rho(z) = z \cdot \mathrm{id}\}.$$

**Theorem 3.1.5** [Hiraga and Saito 2012, Lemma 2.5 and Corollary 2.7]. *Let $\mathfrak{S} = \mathfrak{S}(\pi)$ be the representation of $S^G(\pi) \times G^\sharp(F)$ on $V_\pi$ defined by*

$$\mathfrak{S}(I, x) = I \circ \pi(x), \quad I \in S^G(\pi), x \in G^\sharp(F).$$

*Then there is a decomposition*

$$\mathfrak{S} \simeq \bigoplus_{\rho \in \Pi_-(S^G(\pi))} \rho \boxtimes \pi^\sharp_\rho, \tag{10}$$

*where $\rho \mapsto \pi^\sharp_\rho$ is a bijection from $\Pi_-(S^G(\pi))$ to $\Pi_\pi$, characterized by (10).*

**3.2. *Relation to parabolic induction.*** Let $P$ be a parabolic subgroup of $G$ with a Levi decomposition $P = MU$. In this article, we denote systematically

$$P^\sharp := P \cap G^\sharp,$$
$$M^\sharp := M \cap M^\sharp.$$

Then $P^\sharp$ is a parabolic subgroup of $G^\sharp$ with Levi decomposition $P^\sharp = M^\sharp U$, since every unipotent subgroup of $G$ is contained in $G_{\mathrm{der}}$. The map $P \mapsto P^\sharp$ (respectively $M \mapsto M^\sharp$) induces a bijection between the parabolic subgroups (respectively Levi subgroups) of $G$ and $G^\sharp$, which leaves the unipotent radicals intact. We also have a canonical identification $W(M^\sharp) = W(M)$. In what follows, we will fix Haar measures on the unipotent radicals of parabolic subgroups of $G$ and $G^\sharp$, which are compatible with the identifications above.

Obviously, the modulus functions satisfy $\delta_P(m) = \delta_{P^\sharp}(m)$ for all $m \in M^\sharp(F)$.

**Lemma 3.2.1** [Tadić 1992, Lemma 1.1]. *Let $\sigma \in \Pi(M)$. Then we have the following isomorphism between smooth representations of $G^\sharp(F)$:*

$$I^G_P(\sigma)|_{G^\sharp} \longrightarrow I^{G^\sharp}_{P^\sharp}(\sigma|_{M^\sharp}),$$
$$\varphi \mapsto \varphi|_{G^\sharp(F)},$$

*which is functorial in $\sigma$.*

*Proof.* Upon recalling the definitions of $I^G_P(\sigma)$ and $I^{G^\sharp}_{P^\sharp}(\sigma|_{M^\sharp})$ as function spaces, the assertion follows from the canonical isomorphisms

$$P^\sharp(F)\backslash G^\sharp(F) = (P^\sharp\backslash G^\sharp)(F) \xrightarrow{\sim} (P\backslash G)(F) = P(F)\backslash G(F)$$

and the fact that $\delta_P|_{M^\sharp} = \delta_{P^\sharp}$. $\square$

The next result will not be used in this article; we include it only for the sake of completeness. Recall that the normalized Jacquet functor $r_P^G$ is the left adjoint of $I_P^G$. The same is true for $r_{P^\sharp}^{G^\sharp}$.

**Lemma 3.2.2.** *Let $\pi \in \Pi(G)$. The restriction of representations induces an isomorphism*

$$r_P^G(\pi)|_{M^\sharp} \xrightarrow{\sim} r_{P^\sharp}^{G^\sharp}(\pi|_{G^\sharp})$$

*between smooth representations of $M^\sharp(F)$, which is functorial in $\pi$.*

*Proof.* The claim is evident. $\square$

**Proposition 3.2.3.** *Let $M$ be a Levi subgroup of $G$. Then the inclusion map $M \hookrightarrow G$ induces an isomorphism between locally compact abelian groups*:

$$M(F)/M^\sharp(F) \xrightarrow{\sim} G(F)/G^\sharp(F).$$

*Proof.* The inclusion map induces an isomorphism $M/M^\sharp \hookrightarrow G/G^\sharp$ as $F$-tori. Hence the short exact sequence $1 \to M^\sharp \to M \to M/M^\sharp \to 1$ and its avatar for $G$ provide a commutative diagram of pointed sets with exact rows:

$$\begin{tikzcd}
1 \arrow[r] & G^\sharp(F) \arrow[r] & G(F) \arrow[r] & (G/G^\sharp)(F) \arrow[r] & H^1(F, G^\sharp) \\
1 \arrow[r] & M^\sharp(F) \arrow[r] \arrow[u] & M(F) \arrow[r] \arrow[u] & (M/M^\sharp)(F) \arrow[r] \arrow[u, equal] & H^1(F, M^\sharp). \arrow[u]
\end{tikzcd}$$

Fix a parabolic subgroup $P$ of $G$ with a Levi decomposition $P = MU$. The rightmost vertical arrow factorizes as

$$H^1(F, M^\sharp) \to H^1(F, P^\sharp) \to H^1(F, G^\sharp).$$

The first map is an isomorphism whose inverse is induced by $P^\sharp \twoheadrightarrow P^\sharp/U = M^\sharp$. It is well known that the second map is injective, hence so is the composition. A simple diagram chasing shows $M(F)/M^\sharp(F) \xrightarrow{\sim} G(F)/G^\sharp(F)$, as asserted. $\square$

**Corollary 3.2.4.** *The restriction map induces an isomorphism*

$$(G(F)/G^\sharp(F))^D \xrightarrow{\sim} (M(F)/M^\sharp(F))^D.$$

Here is a trivial but important consequence: any $\eta \in (M(F)/M^\sharp(F))^D$ is invariant under the $W(M)$-action.

**3.3. *Relation to intertwining operators.*** Let $M$ be a Levi subgroup of $G$. First of all, observe that there is a natural decomposition

$$\mathfrak{a}_M^* = \mathfrak{a}_{M^\sharp}^* \oplus \mathfrak{b}^*,$$

where

$$\mathfrak{b}^* := \mathscr{X}(M/M^\sharp) \otimes_{\mathbb{Z}} \mathbb{R} \hookrightarrow \mathfrak{a}_G^*.$$

Henceforth, we shall identify $\mathfrak{a}_{M^\sharp}^*$ as a vector subspace of $\mathfrak{a}_M^*$. We shall do the same for their complexifications. This is compatible with restrictions in the following sense:

$$(\sigma_\lambda)|_{M^\sharp} = (\sigma|_{M^\sharp})_\lambda, \quad \sigma \in \Pi(M),\ \lambda \in \mathfrak{a}_{M^\sharp,\mathbb{C}}^*.$$

**Lemma 3.3.1.** *Let $\sigma \in \Pi(M)$, $P, Q \in \mathscr{P}(M)$. For $\lambda \in \mathfrak{a}_{M^\sharp,\mathbb{C}}^*$ in general position, the following diagram is commutative*:

$$\begin{tikzcd} I_P^G(\sigma_\lambda)|_{G^\sharp} \arrow[r, "J_{Q|P}(\sigma_\lambda)"] \arrow[d, "\simeq"'] & I_Q^G(\sigma_\lambda)|_{G^\sharp} \arrow[d, "\simeq"] \\ I_{P^\sharp}^{G^\sharp}(\sigma_\lambda|_{M^\sharp}) \arrow[r, "J_{Q^\sharp|P^\sharp}(\sigma_\lambda|_{M^\sharp})"'] & I_{Q^\sharp}^{G^\sharp}(\sigma_\lambda|_{M^\sharp}), \end{tikzcd}$$

*where the vertical isomorphisms are those defined in Lemma 3.2.1.*

*Proof.* It suffices to check this for $\mathrm{Re}(\lambda)$ in the cone of absolute convergence of the integrals (3) defining $J_{Q|P}$ and $J_{Q^\sharp|P^\sharp}$. The commutativity then follows from (3) and the definition of the isomorphism in Lemma 3.2.1. □

**Proposition 3.3.2.** *Let $\sigma \in \Pi_{2,\mathrm{temp}}(M)$, $\sigma^\sharp \in \Pi_{2,\mathrm{temp}}(M^\sharp)$ such that $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$. Then, for all $\lambda \in \mathfrak{a}_{M^\sharp,\mathbb{C}}^*$, we have $\mu(\sigma_\lambda) = \mu(\sigma_\lambda^\sharp)$; more precisely,*

$$\mu_\alpha(\sigma_\lambda) = \mu_\alpha(\sigma_\lambda^\sharp) \quad \textit{for all } \alpha \in \Sigma_P^{\mathrm{red}} = \Sigma_{P^\sharp}^{\mathrm{red}},\ P \in \mathscr{P}(M).$$

*Proof.* In view of our choice of measures on unipotent radicals, the identities of $\mu$-functions follow from (4) and Lemma 3.3.1. □

**Proposition 3.3.3.** *Let $\sigma \in \Pi(M)$ and $\eta \in (G(F)/G^\sharp(F))^D$. Then we have*

$$j(\sigma) = j(\eta\sigma).$$

*In particular, for $\sigma \in \Pi_{2,temp}(M)$, we have $\mu(\sigma) = \mu(\eta\sigma)$.*

*Proof.* In view of the definition of $j$-function (4), it suffices to observe that for all $P$, $Q \in \mathcal{P}(M)$ and $\lambda \in \mathfrak{a}^*_{M,\mathbb{C}}$ in general position, the following diagram commutes:

$$\begin{tikzcd} \eta I_P^G(\sigma_\lambda) \arrow[r, "J_{Q|P}(\sigma_\lambda)"] \arrow[d, "\simeq"'] & \eta I_Q^G(\sigma_\lambda) \arrow[d, "\simeq"] \\ I_P^G(\eta\sigma_\lambda) \arrow[r, "J_{Q|P}(\eta\sigma_\lambda)"'] & I_Q^G(\eta\sigma_\lambda), \end{tikzcd}$$

where the vertical arrows are given by $\varphi(\cdot) \mapsto \eta(\cdot)\varphi(\cdot)$; note that we used the natural identification $\mathrm{Hom}_G(\pi_1, \pi_2) = \mathrm{Hom}_G(\eta\pi_1, \eta\pi_2)$ for all $\pi_1, \pi_2 \in \Pi(G)$. Indeed, the commutativity can be seen from the definition (3) of $J_{Q|P}(\cdot)$ when $\mathrm{Re}(\lambda)$ lies in the cone of absolute convergence. $\square$

**Theorem 3.3.4.** *One can choose a family of normalizing factors for $G$ such that*

$$r_{Q|P}(\sigma_\lambda) = r_{Q|P}(\eta\sigma_\lambda)$$

*for all $(M, \sigma)$, $P$, $Q \in \mathcal{P}(M)$ and $\eta \in (G(F)/G^\sharp(F))^D$, which is unitary. Given such a family of normalizing factors, one can define normalizing factors $r_{Q^\sharp|P^\sharp}(\sigma^\sharp_\lambda)$ for those $\sigma^\sharp$ such that $\omega_{\sigma^\sharp}$ is unitary by setting*

$$r_{Q^\sharp|P^\sharp}(\sigma^\sharp_\lambda) := r_{Q|P}(\sigma_\lambda), \quad \lambda \in \mathfrak{a}^*_{M^\sharp,\mathbb{C}}, \tag{11}$$

*where $\sigma \in \Pi(M)$ is as in Proposition 3.1.1 with $\omega_\sigma$ unitary.*

*Moreover, let $\sigma$, $\sigma^\sharp$ be as above and $\iota : \sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$ be an embedding. Let $P$, $Q \in \mathcal{P}(M)$, $w \in W(M)$ with a representative $\tilde{w} \in G^\sharp(F)$. The following diagrams of $G^\sharp(F)$-representations are commutative:*

$$\begin{tikzcd} I_P^G(\sigma_\lambda) \arrow[r, "R_{Q|P}(\sigma_\lambda)"] & I_Q^G(\sigma_\lambda) \\ I^{G^\sharp}_{P^\sharp}(\sigma^\sharp_\lambda) \arrow[u, hook] \arrow[r, "R_{Q^\sharp|P^\sharp}(\sigma^\sharp_\lambda)"'] & I^{G^\sharp}_{Q^\sharp}(\sigma^\sharp_\lambda), \arrow[u, hook] \end{tikzcd} \qquad \begin{tikzcd} I_P^G(\sigma_\lambda) \arrow[r, "r_P(\tilde{w},\sigma_\lambda)"] & I_P^G(\tilde{w}(\sigma_\lambda)) \\ I^{G^\sharp}_{P^\sharp}(\sigma^\sharp_\lambda) \arrow[u, hook] \arrow[r, "r_{P^\sharp}(\tilde{w},\sigma^\sharp_\lambda)"'] & I^{G^\sharp}_{P^\sharp}(\tilde{w}(\sigma^\sharp_\lambda)) \arrow[u, hook] \end{tikzcd}$$

*for $\lambda \in \mathfrak{a}^*_{M,\mathbb{C}}$ in general position, where the vertical arrows are given by*

$$I^{G^\sharp}_{P^\sharp}(\sigma^\sharp_\lambda) \xrightarrow{I^{G^\sharp}_{P^\sharp}(\iota)} I^{G^\sharp}_{P^\sharp}(\sigma_\lambda|_{M^\sharp}) \xrightarrow{\sim} I_P^G(\sigma_\lambda)|_{G^\sharp}.$$

Observe that the asserted invariance under $\eta$-twist is satisfied by Langlands' conjectural family of normalizing factors in Remark 2.2.2, since they are defined in terms of local factors through the adjoint representation of ${}^{\mathrm{L}}M$ on the Lie algebra of $\widehat{G}$.

*Proof.* We will show that $r_{Q|P}(\sigma_\lambda) = r_{Q|P}(\eta\sigma_\lambda)$ by reviewing the construction in Remark 2.2.3. More precisely, we will start from the square-integrable case and show that the equality is preserved throughout the inductive construction.

To begin with, suppose that $M$ is maximal proper and $\sigma \in \Pi_{2,\mathrm{temp}}(M)$. Suppose that $r_{Q|P}(\sigma_\lambda)$ is chosen so that (R$_1$), (R$_2$), (R$_3$), (R$_6$), (R$_7$) are satisfied. Put $r_{Q|P}(\eta\sigma_\lambda) := r_{Q|P}(\sigma_\lambda)$. Then all the conditions above except (R$_1$) are trivially satisfied for $\eta\sigma$. As for (R$_1$), all that we need to check is that when $Q = \bar{P}$,

$$r_{P|Q}(\sigma_\lambda) r_{Q|P}(\sigma_\lambda) = j(\eta\sigma_\lambda), \quad \lambda \in \mathfrak{a}^*_{M,\mathbb{C}}.$$

By Proposition 3.3.3, the right hand side is equal to $j(\sigma_\lambda)$, hence the equality holds. This completes the case of $\sigma \in \Pi_{2,\mathrm{temp}}(M)$.

Suppose now $\sigma \in \Pi_{\mathrm{temp}}(M)$. By the classification of tempered representations, we may write $\sigma \hookrightarrow I^M_R(\tau)$ for some parabolic subgroup $R = M_R U_R$ of $M$ and $\tau \in \Pi_{2,\mathrm{temp}}(M_R)$. Twisting everything by $\eta$, we obtain $\eta\sigma \hookrightarrow I^M_R(\eta\tau)$ in the classification of tempered representations. We have $r_{Q|P}(\sigma) = r_{Q(R)|P(R)}(\tau) = r_{Q(R)|P(R)}(\eta\tau)$ by the previous case; on the other hand, the inductive construction of normalizing factors says that $r_{Q|P}(\eta\sigma) = r_{Q(R)|P(R)}(\eta\tau)$, hence $r_{Q|P}(\eta\sigma) = r_{Q|P}(\sigma)$.

The case of general $\sigma$ is similar. We may write $\sigma$ as the Langlands quotient $I^M_R(\tau_\mu) \twoheadrightarrow \sigma$, where $R = M_R U_R$ is as before, $\tau \in \Pi_{\mathrm{temp}}(M_R)$, and $\mathrm{Re}\langle \mu, \alpha^\vee \rangle > 0$ for all $\alpha \in \Delta^M_R$. Twisting everything by $\eta$, we have $I^M_R(\eta\tau_\mu) \twoheadrightarrow \eta\sigma$, which is still a Langlands quotient. The inductive construction of normalizing factors says that $r_{Q|P}(\sigma) = r_{Q(R)|P(R)}(\tau_{\mu+\lambda})$. Repeating the arguments for the previous case, it follows that $r_{Q|P}(\sigma) = r_{Q|P}(\eta\sigma)$.

Now we can check that

$$r_{Q^\sharp|P^\sharp}(\sigma^\sharp_\lambda) := r_{Q|P}(\sigma_\lambda)$$

is well defined. Recall that $\omega_{\sigma^\sharp}$ and $\omega_\sigma$ are assumed to be unitary. If $\sigma'$ is another choice such that $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$ and $\omega_{\sigma'}$ is unitary, there exists $\eta$ such that $\sigma \simeq \eta\sigma'$. This would imply that $\eta|_{Z_G(F)}$ is unitary, hence so is $\eta$ itself. Therefore $r_{Q|P}(\eta\sigma_\lambda) = r_{Q|P}(\sigma_\lambda)$.

Finally, the commutativity of the diagram results from Lemma 3.3.1 and (11). $\square$

**3.4.** ***Relation to R-groups.*** In this subsection, we will fix

- a parabolic subgroup $P = MU$ of $G$;
- the corresponding parabolic subgroup $P^\sharp := P \cap G^\sharp = M^\sharp U$ of $G^\sharp$;
- $\sigma^\sharp \in \Pi_{2,\mathrm{temp}}(M^\sharp)$;
- $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ such that $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$.

Given $\sigma^\sharp$, the existence of such a $\sigma$ is guaranteed by Propositions 3.1.1 and 3.1.3.

**Lemma 3.4.1** (cf. [Goldberg 2006, Lemma 2.3]). *Under the identification $W(M) = W(M^\sharp)$, we have $W^0_\sigma = W^0_{\sigma^\sharp}$.*

*Proof.* Since both sides are generated by root reflections, it suffices to fix $\alpha \in \Sigma^{\mathrm{red}}_P = \Sigma^{\mathrm{red}}_{P^\sharp}$ and show that $s_\alpha \in W^0_\sigma$ if and only if $s_\alpha \in W^0_{\sigma^\sharp}$, where $s_\alpha$ denotes the root reflection with respect to $\alpha$.

By [Waldspurger 2003, Proposition IV.2.2], $\mu_\alpha(\sigma) = 0$ implies $s_\alpha\sigma \simeq \sigma$. The same is true for $\sigma^\sharp$ instead of $\sigma$. According to the description of $W^0_\sigma$ (respectively $W^0_{\sigma^\sharp}$) in terms of $\mu$-functions, we obtain

$$\mu_\alpha(\sigma) = 0 \Leftrightarrow s_\alpha \in W^0_\sigma \quad (\text{respectively } \mu_\alpha(\sigma^\sharp) = 0 \Leftrightarrow s_\alpha \in W^0_{\sigma^\sharp}).$$

On the other hand, Proposition 3.3.2 implies $\mu_\alpha(\sigma) = \mu_\alpha(\sigma^\sharp)$. The assertion follows immediately. □

**Definition 3.4.2.** Set

$$L(\sigma) := \{\eta \in (M(F)/M^\sharp(F))^D : \text{ there exists } w \in W(M),\, w\sigma \simeq \eta\sigma\},$$
$$L(\sigma^\sharp) := \{\eta \in (M(F)/M^\sharp(F))^D : \text{ there exists } w \in W(M),\, w\sigma \simeq \eta\sigma,\, w\sigma^\sharp \simeq \sigma^\sharp\}.$$

These are subgroups of $(M(F)/M^\sharp(F))^D$. Indeed, let

$$\eta, \eta' \in L(\sigma) \quad \text{and} \quad w, w' \in W(M)$$

such that $\eta\sigma \simeq w\sigma$, $\eta'\sigma \simeq w'\sigma$. Then one has

$$\eta'\eta\sigma \simeq \eta' w\sigma = w\eta'\sigma \simeq ww'\sigma. \tag{12}$$

Hence $\eta\eta' \in L(\sigma)$. The case of $L(\sigma^\sharp)$ is similar. Note that $X^M(\sigma) \subset L(\sigma^\sharp) \subset L(\sigma)$.

There is an obvious counterpart for the Weyl group, namely,

$$\overline{W}_\sigma := \{w \in W(M) : \text{ there exists } \eta \in (M(F)/M^\sharp(F))^D,\; w\sigma \simeq \eta\sigma\}.$$

It is clear that $\overline{W}_\sigma \supset W_\sigma$. On the other hand, Proposition 3.1.2 implies that $\overline{W}_\sigma \supset W_{\sigma^\sharp}$.

The following result is clear in view of the preceding definitions.

**Lemma 3.4.3.** *There is a homomorphism given by*

$$\overline{\Gamma} : \overline{W}_\sigma \longrightarrow L(\sigma)/X^M(\sigma) w \mapsto \textit{ the } [\eta \bmod X^M(\sigma)] \textit{ such that } w\sigma \simeq \eta\sigma,$$

*which satisfies the following.*

(i) $\overline{\Gamma}$ *is surjective.*

(ii) $\mathrm{Ker}(\overline{\Gamma}) = W_\sigma$.

(iii) *The preimage of $L(\sigma^\sharp)/X^M(\sigma)$ is $W_{\sigma^\sharp} W_\sigma$.*

**Definition 3.4.4.** Let

$$\Gamma : \overline{W}_\sigma / W_\sigma \xrightarrow{\sim} L(\sigma)/X^M(\sigma)$$

be the isomorphism obtained from $\overline{\Gamma}$ in the previous Lemma.

**Proposition 3.4.5** [Goldberg 2006, Proposition 3.2]. *Set*

$$R_\sigma[\sigma^\sharp] := (W_\sigma \cap W_{\sigma^\sharp})/W^0_{\sigma^\sharp},$$

*which is legitimate by Lemma 3.4.1. This is a subgroup of $R_{\sigma^\sharp}$.*

(i) *The homomorphism $\overline{\Gamma}$ induces an isomorphism*

$$\Gamma : R_{\sigma^\sharp}/R_\sigma[\sigma^\sharp] \xrightarrow{\sim} L(\sigma^\sharp)/X^M(\sigma).$$

(ii) *If $R_\sigma = \{1\}$, or equivalently, if $I_P^G(\sigma)$ is irreducible, then $\Gamma$ induces an isomorphism $R_{\sigma^\sharp} \xrightarrow{\sim} L(\sigma^\sharp)/X^M(\sigma)$. Consequently, $R_{\sigma^\sharp}$ is abelian in this case.*

*Proof.* Lemma 3.4.3 gives an isomorphism

$$W_{\sigma^\sharp}/(W_\sigma \cap W_{\sigma^\sharp}) \xrightarrow{\sim} L(\sigma^\sharp)/X^M(\sigma)$$

that can be viewed as a restriction of $\Gamma$. By Lemma 3.4.1, we can take the quotients by $W^0_\sigma = W^0_{\sigma^\sharp}$ on the left hand side. The first assertion follows immediately.

For the second assertion, it suffices to note that $R_\sigma[\sigma^\sharp]$ embeds into $R_\sigma$ as well, since $W^0_\sigma = W^0_{\sigma^\sharp}$. □

**3.5. *L-parameters.*** Let $G$ be a connected reductive $F$-group equipped with a quasisplit inner twist

$$\psi : G \times_F \overline{F} \to G^* \times_F \overline{F}.$$

We identify $\widehat{G}$ with $\widehat{G^*}$, thus ${}^{\mathrm{L}}G = {}^{\mathrm{L}}G^*$. The reader should recall that the definition of the complex reductive group $\widehat{G^*}$ and the $\Gamma_F$-action thereof depend on the choice of a $\Gamma_F$-stable splitting $(B^*, T^*, (E_\alpha)_{\alpha\in\Delta(B^*,T^*)})$ (also known as an $F$-splitting) of $G^*(\overline{F})$; see [Kottwitz 1984, Section 1]. These choices permit us to define a correspondence $M^* \leftrightarrow {}^{\mathrm{L}}M^*$ between the conjugacy classes of Levi subgroups of $G^*$ and their dual avatars inside ${}^{\mathrm{L}}G^*$. Using the inner twist $\psi$, it also makes sense to say if a Levi subgroup $M^*$ of $G^*$ comes from $G$; this notion only depends on the conjugacy classes of Levi subgroups.

For any Levi subgroup $M$ of $G$, there is a canonical bijection between $W^G(M)$ and $W^{\widehat{G}}(\widehat{M})$ coming from the bijection between roots and coroots.

An L-parameter for $G^*$ is a homomorphism

$$\phi : \mathrm{WD}_F \to {}^{\mathrm{L}}G^* = {}^{\mathrm{L}}G$$

such that

- $\phi$ is an L-homomorphism, that is, the composition of $\phi$ with the projection ${}^{\mathrm{L}}G \to W_F$ equals $\mathrm{WD}_F \to W_F$;
- $\phi$ is continuous;
- the projection of $\mathrm{Im}(\phi)$ to $\widehat{G}$ is formed of semisimple elements.

Two L-parameters $\phi_1$, $\phi_2$ are called equivalent, denoted by $\phi_1 \sim \phi_2$, if they are conjugate by $\widehat{G}$. We say that $\phi$ is bounded if the projection of $\mathrm{Im}(\phi)$ to $\widehat{G}$ is bounded (that is, relatively compact); this property depends only on the equivalence class of $\phi$.

Given an L-parameter $\phi$, we define

$$S_\phi := Z_{\widehat{G}}(\mathrm{Im}(\phi)).$$

The connected component $S_\phi^0$ is a connected reductive subgroup of $\widehat{G}$. We record the following basic properties.

(i) The Levi subgroups ${}^{\mathrm{L}}M^* \subset {}^{\mathrm{L}}M$ which contain $\mathrm{Im}(\phi)$ minimally are conjugate by $S_\phi^0$.

(ii) Letting ${}^{\mathrm{L}}M^*$ be a Levi subgroup containing $\mathrm{Im}(\phi)$ minimally, $Z_{\widehat{M^*}}^{\Gamma_F,0}$ is a maximal torus of $S_\phi^0$.

Indeed, these assertions follow from [Borel 1979, Proposition 3.6] and its proof applied to the subgroup $\mathrm{Im}(\phi)$ of ${}^{\mathrm{L}}G$.

So far, everything depends only on the quasisplit inner form $G^*$. We say that $\phi$ is $G$-relevant if $M_\phi^*$ corresponds to a Levi subgroup of $G$; in this case, we write $M_\phi^* = M_\phi$. Put

$$\begin{aligned}
\Phi(G) &:= \{\phi : \mathrm{WD}_F \to {}^{\mathrm{L}}G,\ \phi \text{ is a } G\text{-relevant L-parameter}\}/\sim,\\
\Phi_{\mathrm{bdd}}(G) &:= \{\phi \in \Phi(G) : \phi \text{ is bounded}\},\\
\Phi_{2,\mathrm{bdd}}(G) &:= \{\phi \in \Phi_{\mathrm{bdd}}(G) : M_\phi = G\}.
\end{aligned}$$

Since the relevance condition is vacuous if $G = G^*$, we have $\Phi(G) \subset \Phi(G^*)$, etc.

Now let $G^\sharp$ be a subgroup of $G$ such that $G_{\mathrm{der}} \subset G^\sharp \subset G$. We will study the lifting of L-parameters from $G^\sharp$ to $G$, which is in some sense dual to the restriction of representations. In what follows, the L-groups of $G$ and $G^\sharp$ will be defined using compatible choices of quasisplit inner twists and $F$-splittings.

There is a natural, $\Gamma_F$-equivariant central extension

$$1 \to \widehat{Z^\sharp} \to \widehat{G} \xrightarrow{\mathbf{pr}} \widehat{G^\sharp} \to 1$$

which is dual to $G^\sharp \to G$; here $\widehat{Z^\sharp}$ is the $\mathbb{C}$-torus dual to $G/G^\sharp$.

For $\phi \in \Phi(G)$, we shall set $\phi^\sharp := \mathbf{pr} \circ \phi \in \Phi(G^\sharp)$. When this equality holds, $\phi$ is called a lifting of $\phi^\sharp$.

**Theorem 3.5.1.** *For any $\phi^\sharp \in \Phi(G^\sharp)$, there exists a lifting $\phi \in \Phi(G)$ of $\phi^\sharp$ which is unique up to twists by $H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp)$. If $\phi^\sharp \in \Phi_{\mathrm{bdd}}(G^\sharp)$ (respectively $\phi^\sharp \in \Phi_{2,\mathrm{bdd}}(G^\sharp)$), then $\phi$ can be chosen so that $\phi \in \Phi_{\mathrm{bdd}}(G)$ (respectively $\phi \in \Phi_{2,\mathrm{bdd}}(G)$).*

Note that by local class field theory, $H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp)$ parametrizes the continuous characters of $(G(F)/G^\sharp(F))^D$.

*Proof.* The existential part is just [Labesse 1985, Théorème 8.1] and the uniqueness follows easily. Assume that $\phi^\sharp \in \Phi_{\mathrm{bdd}}(G^\sharp)$ (respectively $\phi^\sharp \in \Phi_{2,\mathrm{bdd}}(G^\sharp)$) and let $\phi$ be any lifting of $\phi^\sharp$; we have to show that there exists a continuous 1-cocycle $a : W_F \to \widehat{Z}^\sharp$ such that the twisted L-parameter $a\phi$ is bounded (respectively bounded and satisfying $M_{a\phi} = M_\phi = G$).

Note that there exists a central isogeny of connected reductive groups

$$G^\sharp \times C \to G$$

given by multiplication, where $C$ is some subtorus of $Z^0_G$. Hence $C \to G/G^\sharp$ is also an isogeny. By duality, we obtain a $\Gamma_F$-equivariant central isogeny of connected reductive complex groups

$$\widehat{G} \to \widehat{G^\sharp} \times \widehat{C}.$$

Let $\phi'$ be the composition of $\phi$ (projected to the $\widehat{G}$ component) with the aforementioned central isogeny. Let us show that $\mathrm{Im}(\phi')$ is bounded upon twisting $\phi$. The first component of $\phi'$ is automatically bounded since $\phi^\sharp$ is. On the other hand, $\widehat{C}$ is isogenous to $\widehat{Z}^\sharp$, therefore, upon replacing $\phi$ by $a\phi$ for some suitable 1-cocycle $a : W_F \to \widehat{Z}^\sharp$, the second component can be made bounded. Hence $a\phi$ is a bounded L-parameter.

To finish the proof, it remains to observe that, assuming $\phi^\sharp = \mathbf{pr} \circ \phi$, the preimage of $M^\sharp_{\phi^\sharp}$ in $\widehat{G}$ is equal to $M_\phi$. □

Here we record a construction related to inner forms which will be required later. Recall that the inner forms of $G^*$ are parametrized by $H^1(F, G^*_{\mathrm{AD}})$. Kottwitz [1984, Section 6] defined the "abelianization" map $\mathrm{ab}^1 : H^1(F, G^*_{\mathrm{AD}}) \to (Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}})^D$ between pointed sets. Hence we can associate to $G$ a character $\chi_G$ of $Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}}$, in the following way:

(13) $$\{\text{inner forms of } G^*\} = H^1(F, G^*_{\mathrm{AD}}) \xrightarrow{\mathrm{ab}^1} (Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}})^D, \quad G \mapsto [\chi_G : Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}} \to \mathbb{C}^\times].$$

## 4. Restriction, continued

This section is devoted to the study of restriction under parabolic induction. As before, we fix connected reductive $F$-groups $G$, $G^\sharp$ such that $G_{\mathrm{der}} \subset G^\sharp \subset G$. We also fix a Levi subgroup $M$ of $G$ and $P \in \mathcal{P}(M)$. The bijection between Levi subgroups (respectively parabolic subgroups) $M \mapsto M^\sharp$ (respectively $P \mapsto P^\sharp$) is defined in Section 3.2.

The normalizing factors for $G$, $G^\sharp$ are chosen as in Theorem 3.3.4 for the representations with unitary central character.

Let $\sigma \in \Pi(M)$. We shall make the following (rather restrictive) hypothesis on $\sigma$ throughout this section.

**Hypothesis 4.0.2.** We assume that $\pi := I_P^G(\sigma)$ is irreducible.

### 4.1. *Embedding of central extensions.*

**Proposition 4.1.1.** *Let $\sigma \in \Pi(M)$ and $\pi := I_P^G(\sigma) \in \Pi(G)$.*

(i) *Under the identification of Corollary 3.2.4, we have $X^M(\sigma) \hookrightarrow X^G(\pi)$.*

(ii) *Let $\omega \in X^M(\sigma)$, $I_\omega^M \in \mathrm{Isom}_M(\omega\sigma, \sigma)$. Define the operator $I_\omega^G$ as the composition of*

$$A_\omega : \omega I_P^G(\sigma) \to I_P^G(\omega\sigma), \quad \varphi \mapsto \omega(\cdot)\varphi(\cdot),$$

*with $I_P^G(I_\omega^M) : I_P^G(\omega\sigma) \xrightarrow{\sim} I_P^G(\sigma)$. Then $I_\omega^G \in \mathrm{Isom}_G(\omega\pi, \pi)$.*

(iii) *We have the following commutative diagram of groups with exact rows:*

$$\begin{tikzcd}
1 \arrow[r] & \mathbb{C}^\times \arrow[r] & S^G(\pi) \arrow[r] & X^G(\pi) \arrow[r] & 1 \\
1 \arrow[r] & \mathbb{C}^\times \arrow[r] \arrow[u, equal] & S^M(\sigma) \arrow[r] \arrow[u, hook] & X^M(\sigma) \arrow[r] \arrow[u, hook] & 1,
\end{tikzcd}$$

*where the arrow $S^M(\sigma) \to S^G(\pi)$ is the map $I_\omega^M \mapsto I_\omega^G$ defined above.*

*Proof.* It follows from the definition that $I_\omega^G \in S^G(\pi)$ for all $\omega \in X^M(\sigma)$, hence $X^M(\sigma) \subset X^G(\pi)$. On the other hand, $I_\omega^G$ is simply the map $\varphi \mapsto \omega(\cdot) I_\omega^M(\varphi(\cdot))$. It is clear that $I_\omega^M \mapsto I_\omega^G$ is a group homomorphism. The commutativity of the diagram is then clear. $\square$

We denote by $\boldsymbol{K}_0(\Pi_-(S^M(\sigma)))$ the space of virtual characters of $S^M(\sigma)$ generated by the elements of $\Pi_-(S^M(\sigma))$. Similarly, $\boldsymbol{K}_0(\Pi_\sigma)$ denotes the space of virtual characters of $M^\sharp(F)$ generated by the elements of $\Pi_\sigma$. The bijection $\rho \mapsto \pi_\rho^\sharp$ in Theorem 3.1.5 extends to an isomorphism $\boldsymbol{K}_0(\Pi_-(S^M(\sigma))) \xrightarrow{\sim} \boldsymbol{K}_0(\Pi_\sigma)$. Assuming $\pi = I_P^G(\sigma)$ irreducible, we have the analogous isomorphism

$$\boldsymbol{K}_0(\Pi_-(S^G(\pi))) \xrightarrow{\sim} \boldsymbol{K}_0(\Pi_\pi),$$

as well as the linear maps

$$\mathrm{Ind}_{S^M(\sigma)}^{S^G(\pi)} : \boldsymbol{K}_0(\Pi_-(S^M(\sigma))) \to \boldsymbol{K}_0(\Pi_-(S^G(\pi))),$$

$$I_{P^\sharp}^{G^\sharp} : \boldsymbol{K}_0(\Pi_\sigma) \to \boldsymbol{K}_0(\Pi_\pi),$$

given by the usual induction and normalized parabolic induction, respectively. Note that Lemma 3.2.1 is invoked here.

**Proposition 4.1.2.** *The following diagram commutes*:

$$\begin{tikzcd} \boldsymbol{K}_0(\Pi_-(S^G(\pi))) \arrow[r, "\simeq"] & \boldsymbol{K}_0(\Pi_\pi) \\ \boldsymbol{K}_0(\Pi_-(S^M(\sigma))) \arrow[r, "\simeq"'] \arrow[u, "\mathrm{Ind}_{S^M(\sigma)}^{S^G(\pi)}"] & \boldsymbol{K}_0(\Pi_\sigma). \arrow[u, "I_{P^\sharp}^{G^\sharp}"'] \end{tikzcd}$$

To prove this, some harmonic analysis on the groups $S^M(\sigma)$, $S^G(\pi)$ is needed. These groups are infinite; nonetheless, the usual theory carries over, as we are only concerned about the representations in $\Pi_-(S^G(\pi))$, $\Pi_-(S^M(\sigma))$ or their contragredients. The worried reader may reduce $S^G(\pi) \to X^G(\pi)$ (respectively $S^M(\sigma) \to X^M(\sigma)$) to a central extension by $\mu_m := \{z \in \mathbb{C}^\times : z^m = 1\}$ for some $m \in \mathbb{Z}$, which is always possible.

*Proof.* Let $\sigma^\sharp \in \Pi_\sigma$ and $\rho \in \Pi_-(S^M(\sigma))$ be the corresponding element. Define

$$\tau := \mathrm{Ind}_{S^M(\sigma)}^{S^G(\pi)}(\rho) \in \boldsymbol{K}_0(\Pi_-(S^G(\pi))),$$
$$\pi^\sharp := I_{P^\sharp}^{G^\sharp}(\sigma^\sharp) \in \boldsymbol{K}_0(\Pi_\pi).$$

We have to show that $\tau$ corresponds to $\pi^\sharp$. To begin with, set

$$\sigma[I_\omega^M] := \sigma(\,\cdot\,) \circ I_\omega^M : M(F) \to \mathrm{Aut}_{\mathbb{C}}(V_\sigma), \quad I_\omega^M \in S^M(\sigma),$$

where $V_\sigma$ is the underlying vector space of $\sigma$. Then $(\sigma[I_\omega^M], \sigma)$ is a smooth $\omega$-representation of $M(F)$, that is,

$$\sigma[I_\omega^M](xy) = \omega(y)\sigma[I_\omega^M](x)\sigma(y), \quad x, y \in M(F).$$

This notion appears in the study of automorphic induction, and more generally it fits into the formalism of twisted endoscopy; cf. [Lemaire 2010, Section 0.4]. It is easy to see that $\Theta_\sigma[I_\sigma^M] := \mathrm{Tr}\,\sigma[I_\sigma^M]$ is well defined as a distribution on $M(F)$. We may restrict $\sigma[I_\omega^M]$ to $M^\sharp(F)$; by abuse of notations, the corresponding distribution, which is also well defined by Proposition 3.1.1, is again denoted by $\Theta_\sigma[I_\sigma^M]$. The same definition applies to $\pi$.

Theorem 3.1.5 implies the following identity of distributions on $M^\sharp(F)$:

$$\Theta_{\sigma^\sharp} = \frac{1}{|X^M(\sigma)|} \sum_{\omega \in X^M(\sigma)} \mathrm{Tr}(\rho^\vee)(I_\omega^M) \cdot \Theta_\sigma[I_\omega^M], \tag{14}$$

where $\rho^\vee$ is the contragredient of $\rho$ and $I_\omega^M \in S^M(\sigma)$ is any preimage $\omega$; the summand does not depend on the choice of $I_\omega^M$.

Define $Z^M(\sigma)$ to be the subgroup of elements $\omega \in X^M(\sigma)$ such that every preimage of $\omega$ in $S^M(\sigma)$ is central. Define $Z^G(\pi)$ similarly. The sum in (14) can be taken over $Z^M(\sigma)$, since $\rho|_{\mathbb{C}^\times} = \mathrm{id}$ implies that $\mathrm{Tr}(\rho^\vee)$ is zero outside the center.

Let $\pi_1^\sharp \in \boldsymbol{K}_0(\Pi_\pi)$ be the character corresponding to $\tau$. By the same reasoning, there is an identity of distributions on $G^\sharp(F)$

$$\Theta_{\pi_1^\sharp} = \frac{1}{|X^G(\pi)|} \sum_{\eta \in Z^G(\pi)} \mathrm{Tr}(\tau^\vee)\left(I_\eta^G\right) \cdot \Theta_\pi[I_\eta^G], \tag{15}$$

where $I_\eta^G \in S^G(\pi)$ is any preimage of $\eta$, as before. It remains to show that $\Theta_{\pi_1^\sharp}(f^\sharp) = \Theta_{\pi^\sharp}(f^\sharp)$ for every $f^\sharp \in C_c^\infty(G^\sharp(F))$.

Choose a special maximal compact open subgroup $K \subset G(F)$ in good position relative to $M$, and set $K^\sharp := K \cap G^\sharp(F)$. Equip $K$ and $K^\sharp$ with appropriate Haar measures that are compatible with the Iwasawa decomposition; see [Waldspurger 2003, I.1]. The parabolic descent of characters implies

$$\Theta_{\pi^\sharp}(f^\sharp) = \Theta_{\sigma^\sharp}(f^\sharp_{P^\sharp}) = \frac{1}{|X^M(\sigma)|} \sum_{\omega \in Z^M(\sigma)} \mathrm{Tr}(\rho^\vee)\left(I_\omega^M\right) \cdot \Theta_\sigma[I_\sigma^M](f^\sharp_{P^\sharp}), \tag{16}$$

where

$$f^\sharp_{P^\sharp}(m) = \delta_P^{1/2}(m) \iint_{U(F)\times K^\sharp} f^\sharp(k^{-1}muk)\, du\, dk, \quad m \in M^\sharp(F).$$

Since $\tau = \mathrm{Ind}_{S^M(\sigma)}^{S^G(\pi)}(\rho)$, we have

$$\mathrm{Tr}(\tau^\vee)(I_\eta^G) = \begin{cases} \dfrac{1}{|X^M(\sigma)|} \displaystyle\sum_{\xi \in X^G(\pi)} \mathrm{Tr}(\rho^\vee)((I_\xi^G)^{-1} I_\eta^M I_\xi^G), & \text{if } \eta \in X^M(\sigma),\ I_\eta^M \mapsto I_\eta^G, \\ 0, & \text{otherwise,} \end{cases}$$

where $I_\xi^G \in S^G(\pi)$ is any preimage of $\xi$; cf. [Serre 1967, Proposition 20]. This may be rewritten as

$$\mathrm{Tr}(\tau^\vee)(I_\eta^G) = \begin{cases} \dfrac{|X^G(\pi)|}{|X^M(\sigma)|} \mathrm{Tr}(\rho^\vee)(I_\eta^M), & \text{if } \eta \in X^M(\sigma) \cap Z^G(\pi),\ I_\eta^M \mapsto I_\eta^G, \\ 0, & \text{otherwise.} \end{cases}$$

We claim that $\Theta_\pi[I_\omega^G](f^\sharp) = \Theta_\sigma[I_\omega^M](f^\sharp_{P^\sharp})$ if $I_\sigma^M \mapsto I_\sigma^G \in S^G(\pi)$. First of all, note that $\Theta_\pi[I_\omega^G]$ is the normalized parabolic induction of $\Theta_\sigma[I_\omega^M]$ in the setting of $\omega$-representations [Lemaire 2010, Sections 1.7 and 3.8]. Hence we have the parabolic descent of $\omega$-characters [Lemaire 2010, Théorème 3.8.2], namely,

$$\Theta_\pi[I_\omega^G](f) = \Theta_\sigma[I_\omega^M](f_{P,\omega}), \quad f \in C_c^\infty(G(F)),$$

where

$$f_{P,\omega}(m) = \delta_P^{1/2}(m) \iint_{U(F)\times K} \omega(k) f(k^{-1}muk)\, du\, dk, \quad m \in M(F).$$

To prove the claim, let us sketch how to "restrict" the $\omega$-character relation above to $G^\sharp(F)$. There exists a compact open subgroup $C \subset Z_G(F)$ such that

(i) $C \cap G^\sharp(F) = \{1\}$,

(ii) $C \subset K$,

(iii) $\omega$ and $\omega_\sigma$ are trivial on $C$, and

(iv) the multiplication maps $C \times G^\sharp(F) \hookrightarrow G(F)$ and $C \times M^\sharp(F) \hookrightarrow M(F)$ are submersive.

Define $\mathbb{1}_C$ to be the constant function 1 on $C$. Choose the unique Haar measure on $C$ such that the submersions above preserve measures locally. Given that $f^\sharp \in C_c^\infty(G^\sharp(F))$, we set $f = \mathrm{vol}(C)^{-1}\mathbb{1}_C \otimes f^\sharp$ on $C \times G^\sharp(F)$, and zero elsewhere. For such $f$, by inspecting the proof of [Lemaire 2010, Proposition 1.8.1], we may redefine $f_{P,\omega}$ by taking the double integral of $f(k^{-1}muk)$ over $U(F) \times K^\sharp$, so that $f_{P,\omega} = \mathrm{vol}(C)^{-1}\mathbb{1}_C \otimes f^\sharp_{P^\sharp}$ on $C \times M^\sharp(F)$ and zero elsewhere. Therefore

$$\Theta_\pi[I^G_\omega](f) = \Theta_\pi[I^G_\omega](f^\sharp),$$
$$\Theta_\sigma[I^M_\sigma](f_{P,\omega}) = \Theta_\sigma[I^M_\sigma](f^\sharp_{P^\sharp}).$$

Hence our claim follows.

All in all, (15) becomes

$$\Theta_{\pi_1^\sharp}(f^\sharp) = \frac{1}{|X^M(\sigma)|} \sum_{\omega \in X^M(\sigma)\cap Z^G(\pi)} \mathrm{Tr}(\rho^\vee)(I^M_\omega) \cdot \Theta_\sigma[I^M_\omega](f^\sharp_{P^\sharp}),$$

where $I^M_\sigma \in S^M(\sigma)$ is any preimage of $\omega$ and $I^M_\sigma \mapsto I^G_\sigma \in S^G(\pi)$. In comparison with (16), it suffices to show that $\Theta_\sigma[I^M_\omega](f^\sharp_{P^\sharp}) = 0$ if $\omega \in Z^M(\sigma)$ but $\omega \notin Z^G(\pi)$. Indeed, for $I \in S^G(\pi)$, we have

$$\Theta_\sigma[I^M_\omega](f^\sharp_{P^\sharp}) = \Theta_\pi[I^G_\omega](f^\sharp) = \mathrm{Tr}\,\mathfrak{S}(I^G_\omega, f^\sharp) = \mathrm{Tr}\,\mathfrak{S}(I^{-1}I^G_\omega I, f^\sharp),$$

since $\mathfrak{S}$ is a representation of $S^G(\pi) \times G^\sharp(F)$. Since $I$ is arbitrary and $\mathfrak{S}|_{\mathbb{C}^\times \times \{1\}} = \mathrm{id}$, we conclude that $\Theta_\sigma[I^M_\omega](f^\sharp_{P^\sharp}) \neq 0$ only if $\omega \in Z^G(\pi)$. $\square$

**4.2.** ***Description of R-groups.*** Let $w \in W(M)$ with a chosen representative

$$\tilde{w} \in G^\sharp(F).$$

Recall the operator $r_P(\tilde{w}, \sigma) : I_P^G(\sigma) \to I_P^G(\tilde{w}\sigma)$ defined in (5), which is the composition of $R_{w^{-1}Pw|P}(\sigma) : I_P^G(\sigma) \to I^G_{w^{-1}Pw|P}(\sigma)$ with the isomorphism

$$\ell(\tilde{w}) : I^G_{w^{-1}Pw}(\sigma) \to I_P^G(\tilde{w}\sigma), \quad \varphi(\,\cdot\,) \mapsto \varphi(\tilde{w}^{-1}\cdot).$$

For $\eta \in (G(F)/G^\sharp(F))^D$, recall the isomorphism $A_\eta$ defined as

$$\eta I_P^G(\sigma) \to I_P^G(\eta\sigma), \quad \varphi(\,\cdot\,) \mapsto \eta(\,\cdot\,)\varphi(\,\cdot\,).$$

Note that the representations $\sigma$, $\eta\sigma$, $\tilde{w}\sigma$, and $\eta\tilde{w}\sigma$ share the same underlying vector space $V_\sigma$. As usual, we will compose the operators above after appropriate twists by $\eta$ or $\tilde{w}$. For example, given $\eta, \eta'$, we may define $A_\eta A_{\eta'}$, which is equal to $A_{\eta\eta'} : \eta\eta' I_P^G(\sigma) \to I_P^G(\eta\eta'\sigma)$.

**Proposition 4.2.1.** *Let*

$$L(\sigma) \subset (M(F)/M^\sharp(F))^D$$

*be the subgroup defined in Definition 3.4.2. Upon identifying $M(F)/M^\sharp(F)$ and $G(F)/G^\sharp(F)$, we have*

$$L(\sigma) \subset X^G(\pi).$$

*If $\sigma \in \Pi_{2,\mathrm{temp}}(M)$, equality holds.*

*Proof.* Let $\eta \in L(\sigma)$. By definition, there exists $w \in W(M)$ with a representative $\tilde{w} \in G^\sharp(F)$ such that $\eta\sigma \simeq \tilde{w}\sigma$. Hence $\eta\pi \simeq I_P^G(\eta\sigma) \simeq I_P^G(\tilde{w}\sigma)$. There is also an isomorphism $r_P(\tilde{w}^{-1}, \tilde{w}\sigma) : I_P^G(\tilde{w}\sigma) \xrightarrow{\sim} I_P^G(\sigma)$. Hence $\eta\pi \simeq \pi$.

Assume $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ and let $\eta \in X^G(\pi)$. Then $I_P^G(\eta\sigma) \simeq I_P^G(\sigma)$. By [Waldspurger 2003, Proposition III.4.1], there exists $w \in W(M)$ with $w\sigma \simeq \eta\sigma$. □

Henceforth we take $w \in \overline{W}_\sigma$. Take any $\eta \in L(\sigma)$ whose class modulo $X^M(\sigma)$ equals $\overline{\Gamma}(w)$ (see Lemma 3.4.3). Then $\eta$ is of finite order by Proposition 4.2.1; in particular, $\eta$ is unitary. For any isomorphism

$$i : \eta\sigma \xrightarrow{\sim} \tilde{w}\sigma,$$

we deduce an isomorphism

$$\mathrm{Ad}(i) : S^M(\eta\sigma) \to S^M(\tilde{w}\sigma), \quad I \mapsto iIi^{-1} =: \mathrm{Ad}(i)I.$$

By the obvious identifications $S^M(\sigma) = S^M(\eta\sigma) = S^M(\tilde{w}\sigma)$ (without using $i$), one can view $\mathrm{Ad}(i)$ as an automorphism of $S^M(\sigma)$. It leaves $\mathbb{C}^\times$ intact and covers the identity map $X^M(\eta\sigma) \xrightarrow{\sim} X^M(\tilde{w}\sigma)$, and hence induces a bijection

$$\Pi_-(S^M(\sigma)) \to \Pi_-(S^M(\sigma)), \quad \rho \mapsto \rho \circ \mathrm{Ad}(i).$$

We shall write

$$\tilde{w}\rho := \rho \circ \mathrm{Ad}(i).$$

The puzzling notation will be justified by Lemma 4.2.3.

Henceforth, we adopt the following convention: for $I^G_\eta \in S^G(\pi)$, we regard $\mathrm{Ad}(I^G_\eta)$ as an automorphism of $S^M(\sigma)$ via the embedding $S^M(\sigma) \hookrightarrow S^G(\pi)$ provided by Proposition 4.1.1.

**Lemma 4.2.2.** *Let $\eta$, $\tilde{w}$, and $i : \eta\sigma \xrightarrow{\sim} \tilde{w}\sigma$ be as above. As automorphisms of $S^M(\sigma)$, we have*

$$\mathrm{Ad}(i) = \mathrm{Ad}(I^G_\eta)$$

*for every $I^G_\eta \in S^G(\pi)$ in the preimage of $\eta$.*

*Proof.* Given $\eta$, the assertion is independent of the choice of $I^G_\eta$. Let us consider the specific choice as follows:

$$I^G_\eta := r_P(\tilde{w}^{-1}, \tilde{w}\sigma) \circ I^G_P(i) \circ A_\eta : \eta I^G_P(\sigma) \to I^G_P(\sigma).$$

By the definition of the embedding $S^M(\sigma) \hookrightarrow S^G(\pi)$ and that of $A_\eta$, one sees that $\mathrm{Ad}(A_\eta)$ induces the identity map $S^M(\sigma) \xrightarrow{\sim} S^M(\eta\sigma)$. Similarly, the functorial properties of $r_P(\tilde{w}^{-1}, \cdot)$ imply that $\mathrm{Ad}(r_P(\tilde{w}^{-1}, \tilde{w}\sigma))$ induces the identity map $S^M(\tilde{w}\sigma) \xrightarrow{\sim} S^M(\sigma)$. One can readily check that $\mathrm{Ad}(I^G_P(i))$ induces $\mathrm{Ad}(i) : S^M(\eta\sigma) \xrightarrow{\sim} S^M(\tilde{w}\sigma)$. Hence the assertion follows. □

Before stating the next result, recall that $\sigma$, $\tilde{w}\sigma$, and $\eta\sigma$ share the same underlying space $V_\sigma$.

**Lemma 4.2.3.** *Let $\rho \in \Pi_-(S^M(\sigma))$ and $\sigma^\sharp \in \Pi_\sigma$. By identifying the groups $S^M(\sigma)$, $S^M(\tilde{w}\sigma)$, and $S^M(\eta\sigma)$, the following are equivalent*:

(i) $\rho \in \Pi_-(S^M(\sigma))$ *corresponds to* $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$.

(ii) $\rho \in \Pi_-(S^M(\tilde{w}\sigma))$ *corresponds to* $\tilde{w}\sigma^\sharp \hookrightarrow \tilde{w}\sigma|_{M^\sharp}$.

(iii) $\tilde{w}\rho \in \Pi_-(S^M(\eta\sigma))$ *corresponds to* $\tilde{w}\sigma^\sharp \hookrightarrow \eta\sigma|_{M^\sharp}$.

(iv) $\tilde{w}\rho \in \Pi_-(S^M(\sigma))$ *corresponds to* $\tilde{w}\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$.

*Proof.* Recall that $\rho$ corresponds to $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$ means that $\rho \boxtimes \sigma^\sharp \hookrightarrow \mathfrak{S}(\sigma)$, where $\mathfrak{S}(\sigma)$ is the $S^M(\sigma) \times M^\sharp(F)$-representation on $V_\sigma$ defined in Theorem 3.1.5.

The first two properties are equivalent by a transport of structure via $\mathrm{Ad}(\tilde{w}) : M^\sharp \to M^\sharp$. The last two properties are evidently equivalent. Finally, the equivalence between the second and the third properties follows by pulling $\rho$ back via $\mathrm{Ad}(i) : S^M(\eta\sigma) \xrightarrow{\sim} S^M(\tilde{w}\sigma)$. □

Now, recall that $Z^M(\sigma)$ is the projection to $X^M(\sigma)$ of the center of $S^M(\sigma)$. Temporarily fix a preimage $I^G_\eta$ for every $\eta \in X^G(\pi)$ and define

$$Z^M(\sigma)^\perp := \{\eta \in X^G(\pi) : \text{ for all } \omega \in Z^M(\sigma),\ I^G_\eta I^G_\omega = I^G_\omega I^G_\eta\} \supset X^M(\sigma). \tag{17}$$

**Proposition 4.2.4.** *Let $\sigma$, $w$, and $\eta$ be as before. For any $\sigma^\sharp \in \Pi_\sigma$ (respectively $\rho \in \Pi_-(S^M(\sigma))$), we have $\tilde{w}\sigma^\sharp \simeq \sigma^\sharp$ (respectively $\tilde{w}\rho \simeq \rho$) if and only if $\eta \in Z^M(\sigma)^\perp$.*

*Proof.* Let $\rho \in \Pi_-(S^M(\sigma))$ be the representation corresponding to $\sigma^\sharp \in \Pi_\sigma$ by Theorem 3.1.5. By Lemma 4.2.3, it suffices to show that $\tilde{w}\rho \simeq \rho$ if and only if $\eta \in Z_M(\sigma)^\perp$.

The elements in $\Pi_-(S^M(\sigma))$ are described by a variant of the Stone–von Neumann theorem for the central extension $1 \to \mathbb{C}^\times \to S^M(\sigma) \to X^M(\sigma) \to 1$. Namely, consider the data $(L, \rho_0)$ where

- $L$ is a maximal abelian subgroup of $S^M(\sigma)$;
- $\rho_0$ is an irreducible representation of $L$ such that $\rho_0(z) = z$ for all $z \in \mathbb{C}^\times \subset L$.

Then $\rho := \mathrm{Ind}_L^{S^M(\sigma)}(\rho_0)$ is an element of $\Pi_-(S^M(\sigma))$. Every $\rho \in \Pi_-(S^M(\sigma))$ arises in this way. Moreover, the isomorphism class of $\rho$ is determined by its central character. These facts are standard consequences of Mackey's theory. See [Kazhdan and Patterson 1984, 0.3] and the remark after Proposition 4.1.2.

We have $\tilde{w}\rho = \rho \circ \mathrm{Ad}(I_\eta^G)$ by Lemma 4.2.2. To conclude the proof, it suffices to show that $\mathrm{Ad}(I_\eta^G)$ fixes the central character of $\rho$ if and only if $\eta \in Z^M(\sigma)^\perp$. This is immediate. $\square$

**Corollary 4.2.5.** *Assume $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ and $\sigma^\sharp \in \Pi_\sigma$. Then we have $L(\sigma^\sharp) = Z^M(\sigma)^\perp$, and the map $\Gamma$ in Proposition 3.4.5 is an isomorphism*

$$\Gamma : R_{\sigma^\sharp} \to Z^M(\sigma)^\perp / X^M(\sigma), \quad wW^0_{\sigma^\sharp} \mapsto \eta X^M(\sigma)$$

*where $w \in W_{\sigma^\sharp}$ and $\eta \in Z^M(\sigma)^\perp$ satisfy the relation*

$$w\sigma \simeq \eta\sigma.$$

*Proof.* This results immediately from the definition of $L(\sigma^\sharp)$. $\square$

### 4.3. *Cocycles.*

**Definition 4.3.1.** Suppose for a moment that $H$ is a finite group and $N$ is a normal subgroup of $H$. Let $\rho$ be an irreducible representation of $N$ and assume that $h\rho := \rho \circ \mathrm{Ad}(h)^{-1} \simeq \rho$ for all $h \in H$. This is a necessary condition for extending $\rho$ to an irreducible representation of $H$, but not sufficient in general. Recall the following construction of an obstruction $\boldsymbol{c}_\rho \in H^2(H/N, \mathbb{C}^\times)$ for extending $\rho$, where $\mathbb{C}^\times$ is equipped with the trivial $H/N$-action. We can choose intertwining operators $\rho(h) \in \mathrm{Isom}_N(h\rho, \rho)$ for each $h \in H$, such that

$$\rho(nh) = \rho(n)\rho(h), \quad \rho(hn) = \rho(h)\rho(n)$$

for every $h \in H$, $n \in N$. Note that either of the equations above implies the other. There is a $\mathbb{C}^\times$-valued 2-cocycle $c_\rho$ characterized by

$$\rho(h_1 h_2) = c_\rho(h_1, h_2)\rho(h_1)\rho(h_2), \quad h_1, h_2 \in H. \tag{18}$$

One readily checks that $c_\rho$ factors through $H/N \times H/N$, and thus defines a class $\boldsymbol{c}_\rho \in H^2(H/N, \mathbb{C}^\times)$. This cohomology class only depends on $\rho$ itself.

The formalism can also be generalized to the case where $H, N$ are central extensions of finite groups by $\mathbb{C}^\times$, and $\rho(z) = z \cdot \mathrm{id}$ for all $z \in \mathbb{C}^\times$.

Let us return to the formalism of the previous subsection. In particular, we assume $P = MU \subset G$ and $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ with the underlying vector space $V_\sigma$. Set $\pi := I_P^G(\sigma)$ as usual. For every $\eta \in Z^M(\sigma)^\perp$, we fix $w \in W(M)$, a representative $\tilde{w} \in G^\sharp(F)$, and an isomorphism

$$i : \eta\sigma \xrightarrow{\sim} \tilde{w}\sigma.$$

Let $\rho \in \Pi_-(S^M(\sigma))$ be corresponding to $\sigma^\sharp \in \Pi_\sigma$. Proposition 4.2.4 implies that $\tilde{w}\rho \simeq \rho$ for every $\eta$ as above. Equivalently, $\rho \circ \mathrm{Ad}((I_\eta^G)^{-1}) \simeq \rho$ for every $I_\eta^G \in S^G(\pi)$ in the preimage of $\eta$ by Lemma 4.2.2. We will use the shorthand

$${}^\eta\rho := \rho \circ \mathrm{Ad}((I_\eta^G)^{-1}).$$

As in Definition 4.3.1, one considers the problem of extending $\rho$ to the preimage of $Z^M(\sigma)^\perp$ in $S^G(\pi)$. Recall that $Z^M(\sigma)^\perp / X^M(\sigma) = R_{\sigma^\sharp}$ by Corollary 4.2.5. The goal of this subsection is to describe the obstruction class $\boldsymbol{c}_\rho \in H^2(R_{\sigma^\sharp}, \mathbb{C}^\times)$ so obtained.

Recall that in Theorem 3.1.5, we have defined an $S^M(\sigma) \times M^\sharp(F)$-representation $\mathfrak{S} = \mathfrak{S}(\sigma)$ on $V_\sigma$. Analogously, we define $\mathfrak{S}(\eta\sigma)$ and $\mathfrak{S}(\tilde{w}\sigma)$; all of them are realized on $V_\sigma$. We fix an embedding $\iota : \rho \boxtimes \sigma^\sharp \hookrightarrow \mathfrak{S}(\sigma)$ of $S^M(\sigma) \times M^\sharp(F)$-representations. By Lemma 4.2.3, the same map gives $\iota : \rho \boxtimes \tilde{w}\sigma^\sharp \hookrightarrow \mathfrak{S}(\tilde{w}\sigma)$ and $\iota : \rho \boxtimes \sigma^\sharp \hookrightarrow \mathfrak{S}(\eta\sigma)$ with appropriate equivariances.

**Lemma 4.3.2.** *For $\eta$, $\tilde{w}$, and $\sigma^\sharp$ fixed as before, we define $\mathfrak{S}'(\eta\sigma)$ to be the $S^M(\eta\sigma) \times M^\sharp(F)$-representation on $V_\sigma$ defined by*

$$\mathfrak{S}'(\eta\sigma)(I, x) = \mathfrak{S}(\eta\sigma)(\mathrm{Ad}(I_\eta^G)^{-1} I, x), \quad I \in S^M(\eta\sigma), x \in M^\sharp(F).$$

*Then the map $\iota$ induces an embedding of $S^M(\eta\sigma) \times M^\sharp(F)$-representations*

$$\iota : {}^\eta\rho \boxtimes \sigma^\sharp \hookrightarrow \mathfrak{S}'(\eta\sigma),$$

*and there exists a unique equivariant isomorphism*

$$\alpha \boxtimes \sigma^\sharp(\tilde{w})^{-1} : {}^\eta\rho \boxtimes \sigma^\sharp \xrightarrow{\sim} \rho \boxtimes \tilde{w}\sigma^\sharp,$$

*for some* $\alpha \in \mathrm{Isom}_{S^M(\sigma)}({}^{\eta}\rho, \rho)$ *and* $\sigma^\sharp(\tilde{w}) \in \mathrm{Isom}_{M^\sharp}(\tilde{w}\sigma^\sharp, \sigma^\sharp)$, *which makes the following diagram commutative*:

$$\begin{tikzcd}
\mathfrak{S}'(\eta\sigma) \arrow[r, "i", "\simeq"'] & \mathfrak{S}(\tilde{w}\sigma) \\
{}^{\eta}\rho \boxtimes \sigma^\sharp \arrow[u, "\iota"] \arrow[r, "\alpha\boxtimes\sigma^\sharp(\tilde{w})^{-1}"'] & \rho \boxtimes \tilde{w}\sigma^\sharp. \arrow[u, "\iota"']
\end{tikzcd}$$

Observe that the pair $(\alpha, \sigma^\sharp(\tilde{w})^{-1})$ is unique up to $\{(z, z^{-1}) : z \in \mathbb{C}^\times\}$.

*Proof.* The $S^M(\eta\sigma)$-action on $\mathfrak{S}'(\eta\sigma)$ makes $i$ equivariant. The leftmost vertical arrow comes from the original embedding $\iota : \rho\boxtimes\sigma^\sharp \hookrightarrow \mathfrak{S}(\eta\sigma)$ by an $\mathrm{Ad}(I^G_\eta)^{-1}$-twist. The images of the vertical arrows are characterized as the $\sigma^\sharp$ (respectively $\tilde{w}\sigma^\sharp$)-isotypic parts under the $M^\sharp(F)$-action. Proposition 4.2.4 implies that $\sigma^\sharp \simeq \tilde{w}\sigma^\sharp$. Therefore there must exist an equivariant isomorphism ${}^{\eta}\rho \boxtimes \sigma^\sharp \xrightarrow{\sim} \rho \boxtimes \tilde{w}\sigma^\sharp$ that makes the diagram commute. Such an isomorphism must be of the form $\alpha \boxtimes \sigma^\sharp(\tilde{w})^{-1}$. □

**Lemma 4.3.3.** *Write* $r_P := r_P(\tilde{w}, \sigma)$ *and* $r_{P^\sharp} := r_{P^\sharp}(\tilde{w}, \sigma^\sharp)$. *There is a commutative diagram*

$$\begin{tikzcd}
I^G_P(\sigma) \arrow[r, "r_P"] & I^G_P(\tilde{w}\sigma) \\
\rho \boxtimes I^{G^\sharp}_{P^\sharp}(\sigma^\sharp) \arrow[u, "I^{G^\sharp}_{P^\sharp}(\iota)"] \arrow[r, "\mathrm{id}\boxtimes r_{P^\sharp}"'] & \rho \boxtimes I^{G^\sharp}_{P^\sharp}(\tilde{w}\sigma^\sharp) \arrow[u, "I^{G^\sharp}_{P^\sharp}(\iota)"']
\end{tikzcd}$$

*whose arrows are equivariant for the* $S^M(\tilde{w}\sigma) \times G^\sharp(F)$ *and* $S^M(\sigma) \times G^\sharp(F)$-*actions.*

*Proof.* Without loss of generality, we may assume $\rho = \mathrm{Hom}_{M^\sharp}(\sigma^\sharp, \sigma)$, that is, the multiplicity space. The embedding $\iota : \rho\boxtimes\sigma^\sharp \hookrightarrow \sigma$ can be taken to be $\epsilon \otimes v \mapsto \epsilon(v)$. Then the commutativity of the diagram follows by applying Theorem 3.3.4 to each $\epsilon \in \mathrm{Hom}_{M^\sharp}(\sigma^\sharp, \sigma)$. The equivariance of the horizontal arrows results from Theorem 3.3.4 and the functorial properties of $r_P(\tilde{w}, \cdot)$, $r_{P^\sharp}(\tilde{w}, \cdot)$. □

**Lemma 4.3.4.** *With the notations of Lemma 4.3.2, there is a commutative diagram*

$$\begin{tikzcd}
I^G_P(\eta\sigma) \arrow[r, "r_P(\tilde{w},\sigma)^{-1}\circ I^G_P(i)"] & I^G_P(\sigma) \\
{}^{\eta}\rho \boxtimes I^{G^\sharp}_{P^\sharp}(\sigma^\sharp) \arrow[u, "I^{G^\sharp}_{P^\sharp}(\iota)"] \arrow[r, "\alpha\boxtimes R_{P^\sharp}(\tilde{w},\sigma^\sharp)^{-1}"'] & \rho \boxtimes I^{G^\sharp}_{P^\sharp}(\sigma^\sharp), \arrow[u, "I^{G^\sharp}_{P^\sharp}(\iota)"']
\end{tikzcd}$$

*where we set* $R_{P^\sharp}(\tilde{w}, \sigma^\sharp) := \sigma^\sharp(\tilde{w}) \circ r_{P^\sharp}(\tilde{w}, \sigma^\sharp)$*, by using the pair* $(\alpha, \sigma^\sharp(\tilde{w})^{-1})$ *of isomorphisms in Lemma 4.3.2.*

*Proof.* This is the concatenation of the diagram in Lemma 4.3.3 and the one in Lemma 4.3.2, after applying $I_{P^\sharp}^{G^\sharp}(\cdot)$. ☐

**Proposition 4.3.5.** *Let* $\boldsymbol{c}_\rho$ *be the obstruction of extending* $\rho$ *to the preimage of* $Z^M(\sigma)^\perp$ *in* $S^G(\pi)$*, and* $\boldsymbol{c}_{\sigma^\sharp}$ *the class attached to* $R_{\sigma^\sharp}$ *in (7). Then we have*

$$\boldsymbol{c}_\rho = \boldsymbol{c}_{\sigma^\sharp}^{-1}$$

*in* $H^2(R_{\sigma^\sharp}, \mathbb{C}^\times)$*.*

*Proof.* Fix $\iota : \rho \boxtimes \sigma^\sharp \hookrightarrow \sigma$. Also fix a set of representatives $\tilde{w} \in G^\sharp(F)$ for each $w \in R_{\sigma^\sharp}$. For each $\eta$, together with the auxiliary choice $i : \eta\sigma \xrightarrow{\sim} \tilde{w}\sigma$, the top row of the diagram in Lemma 4.3.4 gives an operator $I_\eta^G \circ A_\eta^{-1} : I_P^G(\eta\sigma) \xrightarrow{\sim} I_P^G(\sigma)$ for some $I_\eta^G \in S^G(\pi)$. The isomorphism $A_\eta : \eta I_P^G(\sigma) \xrightarrow{\sim} I_P^G(\sigma)$ has no effect after restriction. Therefore Lemma 4.3.4 asserts that $I_\eta^G$ is pulled-back to $\alpha \boxtimes R_{P^\sharp}(\tilde{w}, \sigma^\sharp)^{-1}$ under $I_{P^\sharp}^{G^\sharp}(\iota)$.

Now we can forget $i$ and vary $I_\eta^G$ in the preimage of $\eta$ in $S^G(\pi)$, which is a $\mathbb{C}^\times$-torsor. Regard $\alpha = \alpha(I_\eta^G)$ as a function of $I_\eta^G$; it is well defined once we have pinned down the operator $\sigma^\sharp(\tilde{w})$ coupled with $\alpha$.

Suppose that $\eta$ is replaced by $\eta\omega$, where $\omega \in X^M(\sigma)$; accordingly, $I_\eta^G$ is replaced by $I_\eta^G I_\omega^G$, where $I_\omega^M \in S^M(\sigma)$ lies in the preimage of $\omega$ and $I_\omega^M \mapsto I_\omega^G$. This does not affect the chosen data $\tilde{w}$ and $\iota$. On the other hand, the diagram in Lemma 4.3.4 says that $\alpha \boxtimes R_{P^\sharp}(\tilde{w}, \sigma^\sharp)^{-1}$ is replaced by

$$\alpha \circ \rho(I_\omega^M) \boxtimes R_{P^\sharp}(\tilde{w}, \sigma^\sharp)^{-1}.$$

It follows that we can pin down the operators $\sigma^\sharp(\tilde{w})$, and introduce a well-defined function

$$I_\eta^G \mapsto \alpha(I_\eta^G) \in \mathrm{Isom}_{S^M(\sigma)}({}^\eta\rho, \rho),$$

for every $I_\eta^G$ in the preimage of $\eta \in Z^M(\sigma)^\perp$ in $S^G(\pi)$, such that

- $I_\eta^G$ is pulled-back to $\alpha(I_\eta^G) \boxtimes R_{P^\sharp}(\tilde{w}, \sigma^\sharp)^{-1}$ under $I_{P^\sharp}^{G^\sharp}(\iota)$;
- $\alpha(I_\eta^G I_\omega^G) = \alpha(I_\eta^G)\rho(I_\omega^M)$ for every $\omega \in X^M(\sigma)$ and $I_\omega^M$ in its preimage.

Such a family of intertwining operators meets the requirements of Definition 4.3.1. Thus the obstruction can be accounted by the $\mathbb{C}^\times$-valued 2-cocycle $c_\rho$ given by

$$\alpha(I_\xi^G I_\eta^G) = c_\rho(w_\xi, w_\eta)\alpha(I_\xi^G)\alpha(I_\eta^G), \quad \xi, \eta \in Z^M(\sigma)^\perp,$$

where $w_\eta \in R_{\sigma^\sharp}$ denotes the element determined by $\eta$ as in Corollary 4.2.5. The same is true for $w_\xi$.

On the other hand, write $w_\eta \mapsto \tilde{w}_\eta$ for the map that picks the chosen representative for $w_\eta \in R_{\sigma^\sharp}$. Equation (7) defines a 2-cocycle $c_{\sigma^\sharp}$. For every $\xi, \eta \in Z^M(\sigma)^\perp$, we obtain

$$R_{P^\sharp}(\tilde{w}_{\xi\eta}, \sigma^\sharp) = R_{P^\sharp}(\tilde{w}_{\eta\xi}, \sigma^\sharp) = c_{\sigma^\sharp}(w_\eta, w_\xi) R_{P^\sharp}(\tilde{w}_\eta, \sigma^\sharp) R_{P^\sharp}(\tilde{w}_\xi, \sigma^\sharp).$$

All in all, the pull-back of $I_\xi^G I_\eta^G$ by $I_{P^\sharp}^{G^\sharp}(\iota)$ equals

$$c_\rho(w_\xi, w_\eta) c_{\sigma^\sharp}(w_\eta, w_\xi)^{-1} \cdot (\text{the pull-back of } I_\xi^G) \circ (\text{the pull-back of } I_\eta^G).$$

Therefore $c_\rho(w_\xi, w_\eta) = c'_{\sigma^\sharp}(w_\xi, w_\eta) := c_{\sigma^\sharp}(w_\eta, w_\xi)$. It is routine to check that $c'_{\sigma^\sharp} : (R_{\sigma^\sharp})^2 \to \mathbb{C}^\times$ is also a 2-cocycle. Denote by $\boldsymbol{c}'_{\sigma^\sharp}$ the cohomology class of $c'_{\sigma^\sharp}$. It remains to show that $\boldsymbol{c}'_{\sigma^\sharp} = \boldsymbol{c}_{\sigma^\sharp}^{-1}$. We use the following observation: let $A$ be a finite abelian group acting trivially on $\mathbb{C}^\times$; we claim that there is an injective group homomorphism

$$\text{comm} : H^2(A, \mathbb{C}^\times) \to \text{Hom}\left(\bigwedge^2 A, \mathbb{C}^\times\right), \quad \boldsymbol{c} \mapsto [x \wedge y \mapsto c(y, x) c(x, y)^{-1}],$$

where $c$ is any 2-cocycle representing the class $\boldsymbol{c}$. Indeed, let

$$1 \to \mathbb{C}^\times \to \widetilde{A} \to A \to 1$$

be the central extension corresponding to $\boldsymbol{c}$. Then $(x, y) \mapsto c(y, x) c(x, y)^{-1}$ is just the commutator pairing of this central extension. The injectivity results from the elementary fact that such an extension splits if and only if $\widetilde{A}$ is commutative.

Apply this to $A = R_{\sigma^\sharp}$. Since $\text{comm}(\boldsymbol{c}'_{\sigma^\sharp}) = \text{comm}(\boldsymbol{c}_{\sigma^\sharp}^{-1})$, we deduce $\boldsymbol{c}'_{\sigma^\sharp} = \boldsymbol{c}_{\sigma^\sharp}^{-1}$, as asserted. □

## 5. The inner forms of SL($N$)

**5.1.** ***The groups.*** Fix $N \in \mathbb{Z}_{\geq 1}$ and let $G^* := \text{GL}_F(N)$. Let $A$ be a central simple algebra over $F$ of dimension $N^2$. There exist $n \in \mathbb{Z}_{\geq 1}$ and a central division algebra $D$ over $F$ satisfying

$$n^2 \cdot \dim_F D = N^2,$$

such that $A$ is isomorphic to $\text{End}_D(D^n)$. The division $F$-algebra $D$ is uniquely determined by $A$. We put

$$\begin{aligned} \text{Nrd} &:= \text{the reduced norm of } A, \\ \text{GL}_D(n) &:= A^\times, \\ \text{SL}_D(n) &:= \text{Ker}(\text{Nrd} : A^\times \to \mathbb{G}_{\text{m}}). \end{aligned}$$

We can regard $A^\times$ as a reductive $F$-group. It is well known that $A \mapsto A^\times$ induces a bijection between the central simple $F$-algebras of dimension $N^2$ and the inner forms of $G^*$. Given $A$, or equivalently, given $(n, D)$ as above, we shall always write

$$G := \mathrm{GL}_D(n).$$

Under an inner twist $\psi : G \times_F \overline{F} \xrightarrow{\sim} G^* \times_G \overline{F}$, the determinant map $\det : G^* \to \mathbb{G}_\mathrm{m}$ corresponds to $\mathrm{Nrd} : G \to \mathbb{G}_\mathrm{m}$. Since the parametrization of the inner forms of an $F$-group $G^*$ only depend on $G^*_{\mathrm{AD}}$, the map $A \mapsto \mathrm{SL}_D(n)$ establishes a bijection between the central simple $F$-algebras of dimension $N^2$ and the inner forms of $\mathrm{SL}_N(F)$. We write

$$G^\sharp := \mathrm{SL}_D(n) = G_{\mathrm{der}}.$$

Note that $G(F)/G^\sharp(F) = (G/G^\sharp)(F) = F^\times$, since $H^1(F, G^\sharp)$ is trivial by the Hasse principle.

As mentioned in Section 3.5, the inner twist $\psi$ gives a correspondence between Levi subgroups: the Levi subgroups of $G$ is of the form

$$M = \prod_{i=1}^{r} \mathrm{GL}_D(n_i), \quad n_1 + \cdots + n_r = n,$$

and the corresponding Levi subgroup of $G^*$, well defined up to conjugacy, is simply

$$M^* = \prod_{i=1}^{r} \mathrm{GL}_F(n_i \cdot \dim_F D).$$

The L-groups of $G$ and $G^\sharp$ are easily described. We have

$$\begin{aligned} \widehat{G} = \widehat{G^*} &= \mathrm{GL}(N, \mathbb{C}), \\ \widehat{G^\sharp} &= \mathrm{PGL}(N, \mathbb{C}), \\ \widehat{G}_{\mathrm{SC}} = (\widehat{G^\sharp})_{\mathrm{SC}} &= \mathrm{SL}(N, \mathbb{C}), \\ Z_{\widehat{G}_{\mathrm{SC}}} = Z_{(\widehat{G^\sharp})_{\mathrm{SC}}} &= \mu_N(\mathbb{C}) := \{z \in \mathbb{C}^\times : z^N = 1\}. \end{aligned}$$

These complex groups are endowed with the trivial Galois action, thus ${}^{\mathrm{L}}G = \widehat{G} \times W_F$ and ${}^{\mathrm{L}}G^\sharp = \widehat{G^\sharp} \times W_F$. The inclusion $G^\sharp \hookrightarrow G$ is dual to the quotient homomorphism $\mathrm{GL}(N, \mathbb{C}) \to \mathrm{PGL}(N, \mathbb{C})$.

It is also possible to describe the characters $\chi_G = \chi_{G^\sharp}$ in (13) explicitly. Observe that $\Gamma_F$ acts trivially on $Z_{\widehat{G}_{\mathrm{SC}}}$, and one can identify the Pontryagin dual of $Z_{\widehat{G}_{\mathrm{SC}}} = \mu_N(\mathbb{C})$, denoted by $Z^D_{\widehat{G}_{\mathrm{SC}}}$, with $\mathbb{Z}/N\mathbb{Z}$: a class $e \in \mathbb{Z}/N\mathbb{Z}$ corresponds to the character $z \mapsto z^e$. For the inner form $G = \mathrm{GL}_D(n)$ of $G^* = \mathrm{GL}_F(N)$, we have

$$\chi_G \in Z^D_{\widehat{G}_{\mathrm{SC}}} \text{ corresponds to } (n \bmod N) \in \mathbb{Z}/N\mathbb{Z}. \tag{19}$$

Later on, the results of Section 4 will be applied to the tempered representations of $G(F)$. This is justified by the following general result.

**Theorem 5.1.1** [Sécherre 2009]. *Let $P = MU$ be a parabolic subgroup of $G$ and $\sigma \in \Pi_{\mathrm{unit}}(M)$. Then $I_P^G(\sigma)$ is irreducible. In particular, Hypothesis 4.0.2 is satisfied by $\sigma$.*

Note that the tempered case is already established in [Deligne et al. 1984].

**5.2.** ***Local Langlands correspondences.*** This subsection is a summary of [Hiraga and Saito 2012, Chapter 11].

*Local Langlands correspondence for* $\mathrm{GL}_D(n)$. Using the local Langlands correspondence for $G^*$, we can define the notion of $G^*$-generic elements in $\Phi(G)$: a parameter $\phi \in \Phi(G) \subset \Phi(G^*)$ is called $G^*$-generic if it parametrizes a generic representation of $G^*(F)$. This defines a subset $\Phi_{G^*\text{-gen}}(G)$ of $\Phi(G)$.

**Theorem 5.2.1** [Hiraga and Saito 2012, Lemmas 11.1 and 11.2]. *Let $G = \mathrm{GL}_D(n)$ and $G^* = \mathrm{GL}_F(N)$ as in Section 5.1. There exists a subset $\Pi_{G^*\text{-gen}}(G)$ of $\Pi(G)$ satisfying*

- $\Pi_{G^*\text{-gen}}(G) \supset \Pi_{\mathrm{temp}}(G)$,
- $\Pi_{G^*\text{-gen}}(G)$ *is stable under twists by* $(G(F)/G^\sharp(F))^D$,

*and a canonically defined bijection between $\Pi_{G^*\text{-gen}}(G)$ and $\Phi_{G^*\text{-gen}}(G)$, denoted by $\pi \leftrightarrow \phi$, such that*

$$\begin{array}{ccc}
\Pi_{G^*\text{-gen}}(G) & \xleftrightarrow{\sim} & \Phi_{G^*\text{-gen}}(G) \\
\uparrow & & \uparrow \\
\Pi_{\mathrm{temp}}(G) & \xleftrightarrow{\sim} & \Phi_{\mathrm{bdd}}(G) \\
\uparrow & & \uparrow \\
\Pi_{2,\mathrm{temp}}(G) & \xleftrightarrow{\sim} & \Phi_{2,\mathrm{bdd}}(G).
\end{array}$$

*The correspondence satisfies the following compatibility properties.*

(i) *When $G = G^*$, the usual Langlands correspondence for $\mathrm{GL}_F(N)$ is recovered.*

(ii) *Given $\pi \leftrightarrow \phi$ and $\boldsymbol{a} \in H^1_{\mathrm{cont}}(W_F, Z_{\widehat{G}})$, let $\eta$ be the character of $G(F)$ deduced from $\boldsymbol{a}$ by local class field theory. Then we have $\omega\pi \leftrightarrow a\phi$.*

(iii) *Given a Levi subgroup $M = \prod_{i\in I} \mathrm{GL}_D(n_i)$ of $G$, let $\sigma := \boxtimes_{i\in I}\sigma_i \in \Pi_{\mathrm{temp}}(M)$. Let $\phi_M \in \Phi_{\mathrm{bdd}}(M)$ such that $\sigma \leftrightarrow \phi_M$ and let $\phi$ be the composition of $\phi_M$ with some L-embedding ${}^{\mathrm{L}}M \hookrightarrow {}^{\mathrm{L}}G$. Then, for any $P \in \mathcal{P}(M)$, we have*

$$I_P^G(\sigma) \leftrightarrow \phi.$$

Note that, in the last assertion, $I_P^G(\sigma)$ is irreducible according to Theorem 5.1.1.

The definitions of $\Pi_{G^*\text{-gen}}(G)$ and $\pi \leftrightarrow \phi$ are based upon the local Langlands correspondence for $G^*$ and the Jacquet–Langlands correspondence for essentially square-integrable representations. We refer the reader to [Hiraga and Saito 2012, Section 11] for details; the compatibility properties are also implicit. Only the tempered/bounded case of the theorem will be used in this article.

*Local Langlands correspondence for* $\mathrm{SL}_D(n)$. Let $G = \mathrm{GL}_D(n)$ and $G^\sharp = G_{\mathrm{der}} = \mathrm{SL}_D(n)$, so that the formalism in Section 3 is applicable. The idea is to define the packets $\Pi_{\phi^\sharp}$ via restriction, by combining the results in Sections 3.1 and 3.5. Let $\Phi_{G^*\text{-gen}}(G^\sharp)$ be the set of $\phi^\sharp \in \Phi(G^\sharp)$ such that $\phi \in \Phi_{G^*\text{-gen}}(G)$ for some lifting $\phi$ of $\phi^\sharp$ (hence for all liftings, since twisting by characters does not affect $G^*$-genericity). For any $\phi^\sharp \in \Phi_{G^*\text{-gen}}(G^\sharp)$, define the corresponding packet by

$$\Pi_{\phi^\sharp} := \Pi_\pi, \quad \pi \leftrightarrow \phi \text{ for some lifting } \phi \in \Phi_{G^*\text{-gen}}(G).$$

By Proposition 3.1.2 and Theorem 3.5.1, the definition of $\Pi_{\phi^\sharp}$ does not depend on the choice of lifting.

On the other hand, set

$$\Pi_{G^*\text{-gen}}(G^\sharp) = \bigsqcup_\pi \Pi_\pi,$$

where $\pi$ ranges over the $(G(F)/G^\sharp(F))^D$-orbits in $\Pi_{G^*\text{-gen}}(G)$. Our version of the local Langlands correspondence for $G^\sharp$ is stated as follows.

**Theorem 5.2.2** [Hiraga and Saito 2012, Chapter 12]. *We have*

$$\Pi_{\mathrm{temp}}(G^\sharp) \subset \Pi_{G^*\text{-gen}}(G^\sharp),$$

*and there is a decomposition*

$$\Pi_{G^*\text{-gen}}(G^\sharp) = \bigsqcup_{\phi^\sharp \in \Phi_{G^*\text{-gen}}(G^\sharp)} \Pi_{\phi^\sharp}, \tag{20}$$

*which restricts to*

$$\Pi_{\mathrm{temp}}(G^\sharp) = \bigsqcup_{\phi^\sharp \in \Phi_{\mathrm{temp}}(G^\sharp)} \Pi_{\phi^\sharp},$$

$$\Pi_{2,\mathrm{temp}}(G^\sharp) = \bigsqcup_{\phi^\sharp \in \Phi_{2,\mathrm{temp}}(G^\sharp)} \Pi_{\phi^\sharp}.$$

*Proof.* The assertion follows from Proposition 3.1.3, Theorems 5.2.1 and 3.5.1, and Proposition 3.1.2. □

Note that each packet $\Pi_{\phi^\sharp}$ is finite. From the endoscopic point of view, in order to justify the correspondence (20), one has to explicate

(i) the internal structure of the packets $\Pi_{\phi^\sharp}$,

(ii) their relation to $S$-groups,

(iii) the endoscopic character identities for $G^\sharp$.

We will recall the definition of $S$-groups (or, more precisely, their component groups, called $\mathcal{S}$-groups) in the next subsection, then summarize its relation to the internal structure of packets; this is one of the main results in [Hiraga and Saito 2012]. The character identities will not be used in this article; we refer the interested reader to [Hiraga and Saito 2012, Theorem 12.7].

*Normalizing factors.* Choose a nontrivial additive character $\psi_F : F \to \mathbb{C}^\times$. Now we can exhibit a canonical family of normalizing factors for $G$ and $G^\sharp$ with respect to $\psi_F$.

Let us begin with $G$. According to the construction in Remark 2.2.3, it suffices to consider the case of inducing representations $\sigma \in \Pi_{2,\mathrm{temp}}(M)$, where $M$ is a Levi subgroup of $G$. When $D = F$, or equivalently, $G = G^* = \mathrm{GL}_F(N)$, the formula in Remark 2.2.2 furnishes a family of normalizing factors in the tempered case, by the Langlands–Shahidi method. To pass to the nonquasisplit case, we use the preservation of $\mu$-functions by Jacquet–Langlands correspondence [Aubert and Plymen 2005, Theorem 7.2] (up to a harmless constant depending only on $D$ and $n$).

From Theorem 3.3.4, we deduce a canonical family of normalizing factors for $G^\sharp$, at least for the inducing representations $\sigma^\sharp$ whose central character is unitary. In what follows, the normalized intertwining operators for $G$ and $G^\sharp$ are assumed to be defined with respect to these factors.

**5.3.** ***Identification of S-groups.*** *Generalities.* To begin with, we summarize the definition of the $S$-groups in the nonquasisplit case by following [Arthur 2006].

**Definition 5.3.1.** Let $G$ be a connected reductive $F$-group. Choose a quasisplit inner twist $\psi : G \times_F \bar{F} \to G^* \times_F \bar{F}$ as well as an $F$-splitting for $G^*(\bar{F})$ to define the L-groups. Let $\phi \in \Phi(G^*)$. We set

$$\begin{aligned}
S_{\phi,\mathrm{ad}} &:= Z_{\widehat{G}}(\mathrm{Im}(\phi))/Z_{\widehat{G}}^{\Gamma_F} \xrightarrow{\sim} (Z_{\widehat{G}}(\mathrm{Im}(\phi))Z_{\widehat{G}})/Z_{\widehat{G}} \subset \widehat{G}_{\mathrm{AD}},\\
S_{\phi,\mathrm{sc}} &:= \text{ the preimage of } S_{\phi,\mathrm{ad}} \text{ in } \widehat{G}_{\mathrm{SC}},\\
\mathcal{S}_\phi &:= \pi_0(S_{\phi,\mathrm{ad}}, 1),\\
\widetilde{\mathcal{S}}_\phi &:= \pi_0(S_{\phi,\mathrm{sc}}, 1).
\end{aligned}$$

From the central extension $1 \to Z_{\widehat{G}_{\mathrm{SC}}} \to S_{\phi,\mathrm{sc}} \to S_{\phi,\mathrm{ad}} \to 1$, we obtain another central extension

$$1 \to \widetilde{Z}_\phi \to \widetilde{\mathcal{S}}_\phi \to \mathcal{S}_\phi \to 1,$$

where

$$\widetilde{Z}_\phi := Z_{\widehat{G}_{\mathrm{SC}}}/(Z_{\widehat{G}_{\mathrm{SC}}} \cap S^0_{\phi,\mathrm{sc}}) = \mathrm{Im}[Z_{\widehat{G}_{\mathrm{SC}}} \to \widetilde{\mathcal{S}}_\phi].$$

**Remark 5.3.2.** When $G$ is an inner form of SL($N$), we recover the definition of the modified $S$-groups in [Hiraga and Saito 2012].

The relevance condition of L-parameters intervenes in the following result. Recall that we have defined a character $\chi_G$ of $Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}}$ in (13).

**Lemma 5.3.3** [Hiraga and Saito 2012, Lemma 9.1]. *If $\phi \in \Phi(G)$, then*

$$\chi_G : Z^{\Gamma_F}_{\widehat{G}_{SC}} \to \mathbb{C}^\times$$

*is trivial on $Z_{\widehat{G}_{SC}} \cap S^0_{\phi,sc}$.*

By abuse of notations, the so-obtained character of $Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}}/(Z_{\widehat{G}_{\mathrm{SC}}} \cap S^0_{\phi,\mathrm{sc}}) \subset \widetilde{Z}_\phi$ is still denoted by $\chi_G$. Also note that $\chi_G$ depends only on $G_{\mathrm{AD}}$.

*Proof.* Let us reproduce the proof in [Hiraga and Saito 2012] here. Let $M = M_\phi$ be a minimal Levi subgroup of $G$ through which $\phi$ factorizes (we used the relevance condition here). By the recollections in Section 3.5, $Z^{\Gamma_F,0}_{\widehat{M}_{\mathrm{sc}}}$ is a maximal torus in $S^0_{\phi,\mathrm{sc}}$. Therefore

$$S^0_{\phi,\mathrm{sc}} \cap Z_{\widehat{G}_{\mathrm{SC}}} = Z^{\Gamma_F,0}_{\widehat{M}_{\mathrm{sc}}} \cap Z^{\Gamma_F}_{\widehat{G}_{\mathrm{SC}}},$$

and the last group is contained in $\mathrm{Ker}(\chi_G)$ by [Arthur 1999, Corollary 2.2]. □

Consider the familiar situation $G_{\mathrm{der}} \subset G^\sharp \subset G$ (cf. Section 3.5), so that we have the $\Gamma_F$-equivariant central extension

$$1 \to \widehat{Z}^\sharp \to \widehat{G} \xrightarrow{\mathbf{pr}} \hat{G}^\sharp \to 1.$$

Let $\phi \in \Phi(G)$ and $\phi^\sharp := \mathbf{pr} \circ \phi \in \Phi(G^\sharp)$. The definitions above pertain to $(G^\sharp, \phi^\sharp)$ as well. Set

$$X^G(\phi) := \{\boldsymbol{a} \in H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp) : \boldsymbol{a}\phi \sim \phi\}.$$

This is a finite abelian group (cf. the proof of Theorem 3.5.1).

**Lemma 5.3.4.** *Let $s \in S_{\phi^\sharp,\mathrm{ad}}$, regarded as an element of $\widehat{G^\sharp}/Z^{\Gamma_F}_{\widehat{G^\sharp}}$. Then $s$ determines a class $\boldsymbol{a} \in H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp)$ characterized by*

$$\tilde{s}\phi(w)\tilde{s}^{-1} = a(w)\phi(w), \quad w \in \mathrm{WD}_F, \tag{21}$$

*where*

- *$\tilde{s} \in \widehat{G}$ is a lifting of $s$,*
- *$a : W_F \to \widehat{Z}^\sharp$ is some 1-cocycle representing $\boldsymbol{a}$, inflated to $\mathrm{WD}_F$.*

*This induces an exact sequence*

$$\mathcal{S}_\phi \to \mathcal{S}_{\phi^\sharp} \to X^G(\phi) \to 1.$$

*Proof.* Choose a lifting $\tilde{s} \in \widehat{G}$. Since $s$ centralizes $\phi^\sharp$, there exists a continuous function $a : \mathrm{WD}_F \to \widehat{Z}^\sharp$ satisfying (21). It is straightforward to check that $a$ is inflated from a 1-cocycle $W_F \to \widehat{Z}^\sharp$. The 1-cocycle $a$ does depend on the choice of $\tilde{s}$, but its class $\boldsymbol{a} \in H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp)$ is uniquely determined by $s$; it is also obvious that $s \mapsto \boldsymbol{a}$ is a homomorphism. Conversely, every $s$ that satisfies (21) for some $\tilde{s}$, $a$ clearly belongs to $S_{\phi^\sharp,\mathrm{ad}}$. Hence the image of $s \mapsto \boldsymbol{a}$ equals $X^G(\phi)$, by the very definition of $X^G(\phi)$.

If $s$ is mapped to the trivial class in $X^G(\phi)$, we may choose $\tilde{s}$ so that $\tilde{s}\phi\tilde{s}^{-1} = \phi$; therefore $s$ comes from $S_{\phi,\mathrm{ad}}$, and vice versa. Hence we have an exact sequence of locally compact groups

$$S_{\phi,\mathrm{ad}} \to S_{\phi^\sharp,\mathrm{ad}} \to X^G(\phi) \to 1.$$

By a connectedness argument, we may pass from the $S$-groups to the $\mathscr{S}$-groups that give the asserted exact sequence. □

*The case of the inner forms of* SL$(N)$. Let us revert to the situation where

$$\begin{aligned} G &= \mathrm{GL}_D(n), \\ G^* &= \mathrm{GL}_F(N), \\ G^\sharp &= \mathrm{SL}_D(n), \\ \chi_G : Z_{\widehat{G}_{\mathrm{SC}}} &= \mu_N(\mathbb{C}) \to \mathbb{C}^\times. \end{aligned}$$

It is well known that $\mathscr{S}_\phi = \{1\}$ for every $\phi \in \Phi(G^*)$. Indeed, $Z_{\widehat{G}}(\mathrm{Im}(\phi))$ is a principal Zariski open subset in some linear subspace of $\mathrm{Mat}_{N\times N}(\mathbb{C})$, and thus is connected; so is its quotient by $Z_{\widehat{G}}^{\Gamma_F} = Z_{\widehat{G}} = \mathbb{C}^\times$.

Let $\phi^\sharp \in \Phi(G^\sharp)$ with a lifting $\phi \in \Phi(G)$. Hence Lemma 5.3.4 yields a canonical isomorphism

$$\mathscr{S}_{\phi^\sharp} \xrightarrow{\sim} X^G(\phi).$$

Assume henceforth that $\phi^\sharp \in \Phi_{G^*\text{-gen}}(G^\sharp)$, so $\phi \in \Phi_{G^*\text{-gen}}(G)$ as well. The local Langlands correspondence for $G$ (Theorem 5.2.1) is thus applicable. Since the local Langlands correspondence is compatible with twisting by characters, we have $X^G(\phi) = X^G(\pi)$, where $\pi \leftrightarrow \phi$. Therefore we deduce the natural isomorphism

$$\mathscr{S}_{\phi^\sharp} \xrightarrow{\sim} X^G(\pi), \quad \text{where } \pi \leftrightarrow \phi,\ \phi^\sharp = \mathbf{pr} \circ \phi. \tag{22}$$

Also observe that $\chi_G$ induces a character of $\widetilde{Z}_{\phi^\sharp}$ by Lemma 5.3.3, since $\Gamma_F$ acts trivially on $Z_{\widehat{G}_{\mathrm{SC}}}$.

**Theorem 5.3.5** [Hiraga and Saito 2012, Lemma 12.5]. *Let $\phi^\sharp \in \Phi_{G^*\text{-gen}}(G^\sharp)$ with a chosen lifting $\phi \in \Phi_{G^*\text{-gen}}(G)$. Let $\pi \in \Pi_{G^*\text{-gen}}(G)$ such that $\pi \leftrightarrow \phi$*

*by Theorem 5.2.1. Then there exists a homomorphism*

$$\Lambda : \widetilde{\mathcal{S}}_{\phi^\sharp} \to S^G(\pi)$$

*such that the following diagram is commutative with exact rows:*

$$\begin{tikzcd}
1 \arrow[r] & \widetilde{Z}_{\phi^\sharp} \arrow[r] \arrow[d, "\chi_G"'] & \widetilde{\mathcal{S}}_{\phi^\sharp} \arrow[r] \arrow[d, "\Lambda"] & \mathcal{S}_{\phi^\sharp} \arrow[r] \arrow[d, "\simeq"] & 1 \\
1 \arrow[r] & \mathbb{C}^\times \arrow[r] & S^G(\pi) \arrow[r] & X^G(\pi), \arrow[r] & 1
\end{tikzcd}$$

*where the rightmost vertical arrow is that of (22).*

*Moreover, $\Lambda$ is unique up to* $\mathrm{Hom}(X^G(\pi), \mathbb{C}^\times)$, *that is, up to the automorphisms of the lower central extension, and upon identifying $\mathcal{S}_{\phi^\sharp}$ and $X^G(\pi)$, this diagram is a push-forward of central extensions by $\chi_G$.*

Note that the assertions about uniqueness and the push-forward are evident; the upshot is the existence of $\Lambda$.

Let $\phi^\sharp$, $\phi$, $\pi$ be as above. Put

$$\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) := \{\rho \in \Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}) : \text{for all } z \in \widetilde{Z}_{\phi^\sharp},\ \rho(z) = \chi_G(z)\,\mathrm{id}\}.$$

The homomorphism $\Lambda$ in Theorem 5.3.5 induces a bijection

$$\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) \xrightarrow{\sim} \Pi_-(S^G(\pi)).$$

Recall that in the local Langlands correspondence for $G^\sharp$ (Theorem 5.2.2), the packet $\Pi_{\phi^\sharp}$ attached to $\phi^\sharp$ is defined as $\Pi_\pi$, the set of irreducible constituents of $\pi|_{G^\sharp}$. Combining Theorem 3.1.5 with Theorem 5.3.5, we arrive at the following description of the packet $\Pi_{\phi^\sharp}$.

**Corollary 5.3.6.** *Let $\phi^\sharp$, $\phi$, $\pi$ be as above. Let* $\mathrm{Hom}(X^G(\pi), \mathbb{C}^\times)$ *act on $\Pi_{\phi^\sharp}$ via the canonical isomorphisms $\Pi_{\phi^\sharp} = \Pi_\pi = \Pi_-(S^G(\pi))$. Then there is a bijection*

$$\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) \xrightarrow{\sim} \Pi_{\phi^\sharp},$$

*which is canonical up to the* $\mathrm{Hom}(X^G(\pi), \mathbb{C}^\times)$*-action on $\Pi_{\phi^\sharp}$.*

When $G$ is quasisplit, $\chi_G$ will be trivial and $\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) = \Pi(\mathcal{S}_{\phi^\sharp})$; the bijection in Corollary 5.3.6 can then be normalized by choosing a Whittaker datum for $G^\sharp$; cf. [Hiraga and Saito 2012, Chapter 3]. In general, however, there is no reason to expect a canonical choice of the bijection $\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) \xrightarrow{\sim} \Pi_{\phi^\sharp}$.

**5.4.** ***Generalization.*** Consider the following abstract setting.

- Let $M$, $M^\sharp$, $M_0^\sharp$ be connected reductive $F$-groups such that $M$ has a split inner form $M^*$, and
$$M_{\mathrm{der}} \subset M_0^\sharp \subset M^\sharp \subset M.$$

- For $\pi \in \Pi(M)$, let $S^M(\pi)$ and $X^M(\pi)$ be the groups defined in Section 3.1 relative to $M^\sharp$, and denote by $S_0^M(\pi)$, $X_0^M(\pi)$ the groups defined relative to $M_0^\sharp$.
- Assume that there are subsets $\Pi_{\mathrm{gen}}(M)$ and $\Phi_{\mathrm{gen}}(M)$ of $\Pi(M)$ and $\Phi(M)$, respectively, together with a "local Langlands correspondence" $\pi \leftrightarrow \phi$ between $\Pi_{\mathrm{gen}}(M)$ and $\Phi_{\mathrm{gen}}(M)$ that is compatible with twist by characters, as in Theorem 5.2.1. We define $\Phi_{\mathrm{gen}}(M^\sharp)$ (respectively $\Phi_{\mathrm{gen}}(M_0^\sharp)$) to be the set of L-parameters that lift to $\Pi_{\mathrm{gen}}(M^\sharp)$ (respectively $\Pi_{\mathrm{gen}}(M_0^\sharp)$) via Theorem 3.5.1.
- Assume that $\mathscr{S}_\phi = \{1\}$ for every $\phi \in \Phi_{\mathrm{gen}}(M)$.

Let $\phi \in \Phi_{\mathrm{gen}}(M)$ and $\pi \in \Pi_{\mathrm{gen}}(M)$ such that $\pi \leftrightarrow \phi$. As before, we deduce L-parameters $\phi^\sharp \in \Phi_{\mathrm{gen}}(M^\sharp)$ and $\phi_0^\sharp \in \Phi_{\mathrm{gen}}(M_0^\sharp)$. First of all, let $\bullet$ be one of the subscripts "ad" or "sc". We have

$$S_{\phi^\sharp,\bullet} \subset S_{\phi_0^\sharp,\bullet} \quad \text{and} \quad S^0_{\phi^\sharp,\bullet} \subset S^0_{\phi_0^\sharp,\bullet}.$$

In view of the definitions in Section 5.3, we deduce natural isomorphisms $\mu$, $\tilde{\mu}$ that fit into the following commutative diagram:

$$\begin{tikzcd}
\mathscr{S}_{\phi^\sharp} \arrow[rr, "\mu"] & & \mathscr{S}_{\phi_0^\sharp} \\
\widetilde{\mathscr{S}}_{\phi^\sharp} \arrow[u] \arrow[rr, "\tilde{\mu}"] & & \widetilde{\mathscr{S}}_{\phi_0^\sharp} \arrow[u] \\
\widetilde{Z}_{\phi^\sharp} \arrow[u] \arrow[rr, two heads, "\tilde{\mu}"] \arrow[dr, "\chi_M"'] & & \widetilde{Z}_{\phi_0^\sharp}. \arrow[u] \arrow[dl, "\chi_M"] \\
& \mathbb{C}^\times &
\end{tikzcd} \tag{23}$$

Secondly, we have $X^M(\pi) \subset X_0^M(\pi)$ as subgroups of $M(F)^D$. Consequently, $S^M(\pi) \subset S_0^M(\pi)$. Iterating the arguments for (22), we obtain the commutative diagram

$$\begin{tikzcd}
\mathscr{S}_{\phi^\sharp} \arrow[r, "\simeq"] \arrow[d, hook, "\mu"'] & X^M(\pi) \arrow[d, hook] \\
\mathscr{S}_{\phi_0^\sharp} \arrow[r, "\simeq"] & X_0^M(\pi).
\end{tikzcd} \tag{24}$$

**Theorem 5.4.1.** *Let $\phi, \phi^\sharp, \phi_0^\sharp$, and $\pi$ be as above. Assume that there exists a homomorphism $\Lambda_0 : \widetilde{\mathscr{S}}_{\phi_0^\sharp} \to S_0^M(\pi)$ such that the following diagram is commutative with exact rows*:

$$\begin{tikzcd} 1 \arrow[r] & \widetilde{Z}_{\phi_0^\sharp} \arrow[r] \arrow[d, "\chi_M"'] & \widetilde{\mathscr{S}}_{\phi_0^\sharp} \arrow[r] \arrow[d, "\Lambda_0"] & \mathscr{S}_{\phi_0^\sharp} \arrow[r] \arrow[d, "\simeq"] & 1 \\ 1 \arrow[r] & \mathbb{C}^\times \arrow[r] & S_0^M(\pi) \arrow[r] & X_0^M(\pi) \arrow[r] & 1. \end{tikzcd} \tag{25}$$

*Then, by setting $\Lambda := \Lambda_0 \circ \tilde{\mu} : \widetilde{\mathscr{S}}_{\phi^\sharp} \to S_0^M(\pi)$ (cf. (23)), the image of $\Lambda$ lies in $S^M(\pi)$ and the analogous diagram below is commutative*:

$$\begin{tikzcd} 1 \arrow[r] & \widetilde{Z}_{\phi^\sharp} \arrow[r] \arrow[d, "\chi_M"'] & \widetilde{\mathscr{S}}_{\phi^\sharp} \arrow[r] \arrow[d, "\Lambda"] & \mathscr{S}_{\phi^\sharp} \arrow[r] \arrow[d, "\simeq"] & 1 \\ 1 \arrow[r] & \mathbb{C}^\times \arrow[r] & S^M(\pi) \arrow[r] & X^M(\pi) \arrow[r] & 1. \end{tikzcd} \tag{26}$$

*Consequently, there is a bijection*

$$\Pi(\widetilde{\mathscr{S}}_{\phi^\sharp}, \chi_G) \xrightarrow{\sim} \Pi_{\phi^\sharp},$$

*which is canonical up to the* $\mathrm{Hom}(X^G(\pi), \mathbb{C}^\times)$*-action on* $\Pi_{\phi^\sharp}$.

*Proof.* Let $\tilde{s} \in \widetilde{\mathscr{S}}_{\phi^\sharp}$, denote by $s$ its image in $\mathscr{S}_{\phi^\sharp}$, and set $s_0 := \mu(s) \in \mathscr{S}_{\phi_0^\sharp}$. Let $\eta \in X_0^M(\pi)$ be the character coming from $\Lambda(\tilde{s}) := \Lambda_0(\tilde{\mu}(\tilde{s})) \in S_0^M(\pi)$. Then, by (25) and (23), $\eta$ is the image of $s_0$ under $\mathscr{S}_{\phi_0^\sharp} \xrightarrow{\sim} X_0^M(\pi)$; using (24), it is also the image of $s$ under $\mathscr{S}_{\phi^\sharp} \xrightarrow{\sim} X^M(\pi)$. If we can show $\Lambda(\tilde{s}) \in S^M(\pi)$ for all $\tilde{s}$, the rightmost square in (26) will commute. Since the square

$$\begin{tikzcd} S^M(\pi) \arrow[r, two heads] \arrow[d, hook] & X^M(\pi) \arrow[d, hook] \\ S_0^M(\pi) \arrow[r, two heads] & X_0^M(\pi) \end{tikzcd}$$

is commutative and cartesian for trivial reasons, it follows that $\Lambda(\tilde{s}) \in S^M(\pi)$. Hence the image of $\Lambda$ lies in $S^M(\pi)$. This also finishes the commutativity of the rightmost square in (26).

Consider the leftmost square in (26). It follows from (23) that, for all $z \in \widetilde{Z}_{\phi^\sharp}$, we have

$$\Lambda(z) = \Lambda_0(\tilde{\mu}(z)) = \chi_M(\tilde{\mu}(z)) = \chi_M(z).$$

Hence the leftmost square is commutative as well.

The bijection $\Pi(\widetilde{\mathcal{S}}_{\phi^\sharp}, \chi_G) \xrightarrow{\sim} \Pi_{\phi^\sharp}$ follows easily from the previous assertions, as in the proof of Corollary 5.3.6. □

**Remark 5.4.2.** The conditions of Theorem 5.4.1 are satisfied if $M$ is a Levi subgroup of $\mathrm{GL}_D(n)$, say of the form

$$M = \prod_{i \in I} \mathrm{GL}_D(n_i), \quad \sum_{i \in I} n_i = n$$

and

$$M_0^\sharp := M_{\mathrm{der}} = \prod_{i \in I} \mathrm{SL}_D(n_i).$$

We simply set $\Pi_{\mathrm{gen}}(M) := \Pi_{M^*\text{-gen}}(M)$ and $\Phi_{\mathrm{gen}}(M) := \Phi_{M^*\text{-gen}}(M)$ by a straightforward generalization of the definitions in Section 5.2. The correspondence $\pi \leftrightarrow \phi$ follows from that in Theorem 5.2.1, applied to each index $i \in I$. The group $M^\sharp$ can be any intermediate group between $M_0^\sharp$ and $M$, including the important case where

$$M^\sharp := M \cap \mathrm{SL}_D(n) = \left\{ (x_i)_{i \in I} \in M : \prod_{i \in I} \mathrm{Nrd}(x_i) = 1 \right\}.$$

Therefore, Theorem 5.3.5 and Corollary 5.3.6 can be generalized to the Levi subgroups of $\mathrm{SL}_D(n)$.

Indeed, it suffices to verify the commutativity of the diagram (25). Writing $\pi = \boxtimes_{i \in I} \pi_i$ and $\phi = (\phi_i)_{i \in I}$, the results in Section 5.3 applied to each $i \in I$ gives a commutative diagram similar to (25), except that its bottom row is the central extension

$$1 \to (\mathbb{C}^\times)^I \to \prod_{i \in I} S^{\mathrm{GL}_D(n_i)}(\pi_i) \to X_0^G(\pi) \to 1 \tag{27}$$

and $\chi_M$ is replaced by

$$\prod_{i \in I} \chi_{\mathrm{GL}_D(n_i)} : \widetilde{Z}_{\phi_0^\sharp} \to (\mathbb{C}^\times)^I.$$

To obtain the desired short exact sequence, it remains to take the push-forward of (27) by the multiplication map $(\mathbb{C}^\times)^I \to \mathbb{C}^\times$.

## 6. The dual $R$-groups

**6.1.** ***A commutative diagram.*** As in Section 5.1, we take

$$G = \mathrm{GL}_D(n), \quad G^* = \mathrm{GL}_F(N), \quad G^\sharp = \mathrm{SL}_D(n).$$

We also fix a Levi subgroup $M$ of $G$ and set $M^\sharp := M \cap G^\sharp$.

To define the dual groups ${}^{\mathrm{L}}G$, ${}^{\mathrm{L}}M$, etc., we fix a quasisplit inner twist $\psi : G \times_F \overline{F} \xrightarrow{\sim} G^* \times_F \overline{F}$ which restricts to a quasisplit inner twist $M \times_F \overline{F} \xrightarrow{\sim} M^* \times_F \overline{F}$, as well as an $F$-splitting for $G^*$ that is compatible with $M^*$. Therefore there is a canonical L-embedding ${}^{\mathrm{L}}M \hookrightarrow {}^{\mathrm{L}}G$. The same is true for ${}^{\mathrm{L}}G^\sharp$ and ${}^{\mathrm{L}}M^\sharp$.

As usual, the natural projections ${}^{\mathrm{L}}G \to {}^{\mathrm{L}}G^\sharp$ and ${}^{\mathrm{L}}M \to {}^{\mathrm{L}}M^\sharp$ are denoted by $\mathbf{pr}$. Put

$$A_{\widehat{M^\sharp}} := Z_{\widehat{M^\sharp}} = Z_{\widehat{M^\sharp}}^{\Gamma_F, 0} \hookrightarrow \widehat{G^\sharp}.$$

Consider $\phi_M \in \Phi_{2,\mathrm{bdd}}(M)$. Let $\phi$ be its composition with ${}^{\mathrm{L}}M \hookrightarrow {}^{\mathrm{L}}G$. Set

$$\phi_M^\sharp := \mathbf{pr} \circ \phi_M \in \Phi_{2,\mathrm{bdd}}(M^\sharp) \quad \text{and} \quad \phi^\sharp := \mathbf{pr} \circ \phi \in \Phi_{\mathrm{bdd}}(G^\sharp).$$

Every $\phi^\sharp \in \Phi_{\mathrm{bdd}}(G^\sharp)$ is obtained in this way (recall Theorem 3.5.1).

The construction of the dual $R$-group associated to $\phi^\sharp$, denoted by $R_{\phi^\sharp}$, is given as follows. Define

$$\begin{aligned}
N_{\phi^\sharp,\mathrm{ad}} &:= N_{S_{\phi^\sharp,\mathrm{ad}}}(A_{\widehat{M^\sharp}}),\\
\mathfrak{N}_{\phi^\sharp} &:= \pi_0(N_{\phi^\sharp,\mathrm{ad}}, 1),\\
W_{\phi^\sharp} &:= W(S_{\phi^\sharp,\mathrm{ad}}, A_{\widehat{M^\sharp}}) \hookrightarrow W^{\widehat{G}}(\widehat{M}),\\
W^0_{\phi^\sharp} &:= W(S^0_{\phi^\sharp,\mathrm{ad}}, A_{\widehat{M^\sharp}}) \lhd W_{\phi^\sharp},\\
R_{\phi^\sharp} &:= W_{\phi^\sharp} / W^0_{\phi^\sharp}.
\end{aligned}$$

The meaning of $W(\cdots, \cdots)$ is as follows: for any pair of complex groups $a \subset A$, the symbol $W(A, a)$ denotes the group $N_A(a)/Z_A(a)$. Note that $W^0_{\phi^\sharp}$ is the Weyl group associated to some root system, as $S^0_{\phi^\sharp,\mathrm{ad}}$ is connected and reductive.

Since the centralizer of $A_{\widehat{M^\sharp}}$ in the connected reductive group $S^0_{\phi^\sharp,\mathrm{ad}}$ is connected, there exists a canonical injection $W^0_{\phi^\sharp} \hookrightarrow \mathfrak{N}_{\phi^\sharp}$. From the results recalled in Section 3.5, the torus $A_{\widehat{M^\sharp}}$ is a maximal torus in $S^0_{\phi^\sharp,\mathrm{ad}}$. Using the conjugacy of maximal tori, one sees that the inclusion map $N_{\phi^\sharp,\mathrm{ad}} \hookrightarrow S_{\phi^\sharp,\mathrm{ad}}$ induces a canonical isomorphism $\mathfrak{N}_{\phi^\sharp} / W^0_{\phi^\sharp} \xrightarrow{\sim} \mathcal{S}_{\phi^\sharp}$.

On the other hand, we also have canonical injections

$$\begin{aligned}
\mathcal{S}_{\phi^\sharp_M} &\hookrightarrow \mathcal{S}_{\phi^\sharp},\\
\mathcal{S}_{\phi^\sharp_M} &\hookrightarrow \mathfrak{N}_{\phi^\sharp}.
\end{aligned}$$

The first one follows from the fact that

$$Z^{\Gamma_F}_{\widehat{M^\sharp}} = Z^{\Gamma_F}_{\widehat{G^\sharp}} Z^{\Gamma_F, 0}_{\widehat{M^\sharp}};$$

see [Arthur 1999, Lemma 1.1]. The injectivity of the second map follows; moreover, its image is characterized as the elements fixing $A_{\widehat{M^\sharp}}$ pointwise.

The relations among these groups are recapitulated in the following result.

**Proposition 6.1.1** [Arthur 1989b, Section 7]. *The groups above fit into a commutative diagram*

$$\begin{tikzcd}
 & & 1 \arrow[d] & 1 \arrow[d] & \\
 & & W^0_{\phi^\sharp} \arrow[r, equal] \arrow[d] & W^0_{\phi^\sharp} \arrow[d] & \\
1 \arrow[r] & \mathscr{S}_{\phi^\sharp_M} \arrow[r] \arrow[d, equal] & \mathfrak{N}_{\phi^\sharp} \arrow[r] \arrow[d] & W_{\phi^\sharp} \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & \mathscr{S}_{\phi^\sharp_M} \arrow[r] & \mathscr{S}_{\phi^\sharp} \arrow[r] \arrow[d] & R_{\phi^\sharp} \arrow[r] \arrow[d] & 1 \\
 & & 1 & 1 &
\end{tikzcd}$$

*whose rows and columns are exact.*

The arrow $\mathscr{S}_{\phi^\sharp} \to R_{\phi^\sharp}$ is uniquely determined by the other terms in this diagram; cf. the proof of Lemma 6.2.1 below.

The same constructions can be applied to $\phi$ and $\phi_M$. The corresponding objects are denoted by $W_\phi$, $W^0_\phi$, etc.

Upon identifying $W^{\widehat{G}}(\widehat{M})$ and $W^G(M)$, we can make $W_{\phi^\sharp}$ act on the tempered L-packet $\Pi_{\phi^\sharp}$. For any $\sigma^\sharp \in \Pi_{\phi^\sharp}$, define

$$\begin{aligned}
W_{\phi^\sharp,\sigma^\sharp} &:= \mathrm{Stab}_{W_{\phi^\sharp}}(\sigma^\sharp), \\
W^0_{\phi^\sharp,\sigma^\sharp} &:= \mathrm{Stab}_{W^0_{\phi^\sharp}}(\sigma^\sharp), \\
R_{\phi^\sharp,\sigma^\sharp} &:= W_{\phi^\sharp,\sigma^\sharp} / W^0_{\phi^\sharp,\sigma^\sharp}.
\end{aligned}$$

The last object $R_{\phi^\sharp,\sigma^\sharp}$, viewed as a subgroup of $R_{\phi^\sharp}$, is what we want to compare with the Knapp–Stein $R$-group $R_{\sigma^\sharp}$.

**6.2.** ***Identification of R-groups.*** Retain the notations of the previous subsection and fix a parabolic subgroup $P \in \mathscr{P}(M)$. We shall always make the identification $W^G(M) = W^{G^\sharp}(M^\sharp)$. The results in Section 4 are applicable to tempered representations of $M(F)$ by Theorem 5.1.1.

Henceforth, let $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ (respectively $\pi \in \Pi_{\mathrm{temp}}(G)$) be the representations corresponding to $\phi_M$ (respectively $\phi$) by Theorem 5.2.1. Then we have

$$\pi \simeq I^G_P(\sigma).$$

Recall that we have defined canonical isomorphisms

$$\mathcal{S}_{\phi^\sharp} \xrightarrow{\sim} X^G(\pi) \quad \text{and} \quad \mathcal{S}_{\phi^\sharp_M} \xrightarrow{\sim} X^M(\sigma)$$

in Section 5.3-5.4. By inspecting the construction in Lemma 5.3.4, we see that these two isomorphisms are compatible with the embeddings $\mathcal{S}_{\phi^\sharp_M} \hookrightarrow \mathcal{S}_\phi$ and $X^M(\sigma) \hookrightarrow X^G(\sigma)$. Therefore we obtain $\gamma : X^G(\pi)/X^M(\sigma) \xrightarrow{\sim} \mathcal{S}_{\phi^\sharp}/\mathcal{S}_{\phi^\sharp_M}$.

**Lemma 6.2.1.** *Define an isomorphism*

$$\widehat{\Gamma}^{-1} : X^G(\pi)/X^M(\sigma) \xrightarrow{\gamma} \mathcal{S}_{\phi^\sharp}/\mathcal{S}_{\phi^\sharp_M} \xrightarrow{\sim} W_{\phi^\sharp}/W^0_{\phi^\sharp} =: R_{\phi^\sharp},$$

*where the second arrow is given by Proposition 6.1.1. Then it is characterized by the equation*

$$\eta\sigma \simeq w\sigma$$

*whenever*

$$\widehat{\Gamma}^{-1}(\eta \bmod X^M(\sigma)) = w \bmod W^0_{\phi^\sharp},$$

*for all $\eta \in X^G(\pi)$ and $w \in W_{\phi^\sharp}$.*

As the notation suggests, we set $\widehat{\Gamma}$ to be the inverse of $\widehat{\Gamma}^{-1}$.

*Proof.* The equation $\eta\sigma \simeq w\sigma$ clearly characterizes $\widehat{\Gamma}^{-1}$. Let $\eta \in X^G(\pi)$ and $\boldsymbol{a} \in H^1_{\mathrm{cont}}(W_F, \widehat{Z}^\sharp)$ which corresponds to $\eta$, together with a chosen 1-cocycle $a$ in the cohomology class of $\boldsymbol{a}$. Since $L(\sigma) = X^G(\pi)$ by Proposition 4.2.1, there exists $w \in W^G(M)$ such that $\eta\sigma \simeq w\sigma$. On the dual side, it implies that there exists $t \in N_{\widehat{G}}(\widehat{M})$ representing $w$, such that

$$t\phi t^{-1} = a\phi.$$

This implies that $t \bmod Z_{\widehat{G}}$ belongs to $S_{\phi^\sharp,\mathrm{ad}}$, and its class $[t]$ in $\mathcal{S}_{\phi^\sharp}$ corresponds to $\eta$ (recall the construction in Lemma 5.3.4).

On the other hand, we have $t \in N_{\phi^\sharp,\mathrm{ad}}$ and its class [t] in $\mathfrak{N}_{\phi^\sharp}$ is mapped to $[t]$ under the arrow $\mathfrak{N}_{\phi^\sharp} \twoheadrightarrow \mathcal{S}_{\phi^\sharp}$ in Proposition 6.1.1. One can also apply the arrow $\mathfrak{N}_{\phi^\sharp} \twoheadrightarrow W_{\phi^\sharp}$ to [t]; since $W_{\phi^\sharp}$ is identified as a subgroup of $W^G(M)$, the image is simply $w$.

Upon some contemplation of the diagram in Proposition 6.1.1, one can see that the image of $[t]$ under $\mathcal{S}_{\phi^\sharp} \to R_{\phi^\sharp}$ is just $w$ modulo $W^0_{\phi^\sharp}$, completing the proof. □

Recall that we have defined the group $\overline{W}_\sigma \subset W^G(M)$. In view of Lemma 3.4.3 and Proposition 4.2.1, we have a canonical isomorphism

$$\Gamma : \overline{W}_\sigma/W_\sigma \xrightarrow{\sim} X^G(\pi)/X^M(\sigma).$$

This is to be compared with $\widehat{\Gamma} : W_{\phi^\sharp}/W^0_{\phi^\sharp} \xrightarrow{\sim} X^G(\pi)/X^M(\sigma)$.

**Proposition 6.2.2.** *We have*

(i) $\overline{W}_\sigma = W_{\phi^\sharp}$,

(ii) $W_\sigma = W^0_{\phi^\sharp}$,

(iii) $\Gamma = \widehat{\Gamma}$.

*In particular*, $R_{\phi^\sharp} \simeq X^G(\pi)/X^M(\sigma)$.

*Proof.* The first assertion follows from the definition of $\overline{W}_\sigma$ and Theorem 3.5.1. Hence $\Gamma$ and $\widehat{\Gamma}$ can be regarded as two surjective homomorphisms from $W_{\phi^\sharp}$ onto $X^G(\pi)/X^M(\sigma)$. However, they admit the same characterization (of the form $\eta\sigma \simeq w\sigma$) by Lemmas 6.2.1 and 3.4.3, and hence are equal. This proves the remaining two assertions. □

**Proposition 6.2.3.** *For all* $\sigma^\sharp \in \Pi_{\phi^\sharp}$, *we have that*

(i) $W_{\sigma^\sharp} = W_{\phi^\sharp,\sigma^\sharp}$,

(ii) $W^0_{\sigma^\sharp} = W^0_{\phi^\sharp,\sigma^\sharp}$,

(iii) *the restriction of* $\widehat{\Gamma}$ *to* $W_{\phi^\sharp,\sigma^\sharp}/W^0_{\phi^\sharp,\sigma^\sharp}$ *induces an isomorphism*

$$W_{\phi^\sharp,\sigma^\sharp}/W^0_{\phi^\sharp,\sigma^\sharp} \xrightarrow{\sim} Z^M(\sigma)^\perp/X^M(\sigma).$$

*In particular*, $R_{\phi^\sharp,\sigma^\sharp} = R_{\sigma^\sharp}$, *and the isomorphisms* $\Gamma$, $\widehat{\Gamma}$ *from these* $R$*-groups onto* $Z^M(\sigma)^\perp/X^M(\sigma)$ *coincide* (*recall Proposition 3.4.5 and Corollary 4.2.5*).

Remainder: the group $Z^M(\sigma)^\perp$ above is defined in (17).

*Proof.* Our proof is based on the previous result. The first assertion follows immediately from the disjointness of tempered $L$-packets. By Lemma 3.4.1 and the fact that $W_\sigma = W^0_\sigma$, we have

$$\begin{aligned} W^0_{\phi^\sharp,\sigma^\sharp} &= W^0_{\phi^\sharp} \cap W_{\sigma^\sharp} = W_\sigma \cap W_{\phi^\sharp} \\ &= W^0_{\sigma^\sharp} \cap W_{\sigma^\sharp} = W^0_{\sigma^\sharp}. \end{aligned}$$

The second assertion follows and the third assertion is then immediate from Proposition 4.2.4. □

Note that the proof for the isomorphism $\widehat{\Gamma} : R_{\phi^\sharp,\sigma^\sharp} \xrightarrow{\sim} Z^M(\sigma)^\perp/X^M(\sigma)$ is independent of the Knapp–Stein theory.

The behavior of the local Langlands correspondence (Theorem 5.2.2) for $G^\sharp$ and its Levi subgroups can now be summarized as follows.

**Theorem 6.2.4.** *Let* $G$, $G^\sharp$ *and* $P = MU$, $P^\sharp = M^\sharp U$ *be as before. For* $\phi^\sharp_M$ *in* $\Phi_{\mathrm{bdd}}(M^\sharp)$, *let* $\phi^\sharp \in \Phi_{\mathrm{bdd}}(G^\sharp)$ *be the composition of* $\phi^\sharp_M$ *with* ${}^{\mathrm{L}}M^\sharp \hookrightarrow {}^{\mathrm{L}}G^\sharp$.

(i) *For every* $\rho \in \Pi(\widetilde{\mathcal{S}}_{\phi_M^\sharp}, \chi_M)$, *parametrizing an irreducible representation* $\sigma^\sharp \in \Pi_{\phi_M^\sharp}$, *the map*

$$I_{P^\sharp}^{G^\sharp}(\sigma^\sharp),$$

*regarded as a virtual character of* $G^\sharp(F)$, *corresponds to that of* $\mathrm{Ind}_{\widetilde{\mathcal{S}}_{\phi_M^\sharp}}^{\widetilde{\mathcal{S}}_{\phi^\sharp}}(\rho)$.

(ii) *For any* $\sigma^\sharp$ *as above*, $I_{P^\sharp}^{G^\sharp}(\sigma^\sharp)$ *is irreducible if and only if* $Z^M(\sigma)^\perp = X^M(\sigma)$ *for some (equivalently, for any)* $\sigma \in \Pi_{\mathrm{temp}}(M)$ *such that* $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$.

(iii) *If* $\phi_M^\sharp \in \Phi_{2,\mathrm{bdd}}(M^\sharp)$, *we have natural isomorphisms*

$$R_{\phi^\sharp} \simeq X^G(\pi)/X^M(\sigma),$$
$$R_{\phi^\sharp,\sigma^\sharp} \simeq R_{\sigma^\sharp} \simeq Z^M(\sigma)^\perp/X^M(\sigma),$$

*where we set* $\pi := I_P^G(\sigma) \in \Pi_{\mathrm{temp}}(G)$, *for any choice of* $\sigma \in \Pi_{2,\mathrm{temp}}(M)$ *such that* $\sigma^\sharp \hookrightarrow \sigma|_{M^\sharp}$.

(iv) *For* $\phi_M^\sharp$, $\sigma^\sharp$, *and* $\rho$ *as above, the class* $\boldsymbol{c}_{\sigma^\sharp} \in H^2(R_{\sigma^\sharp}, \mathbb{C}^\times)$ *of* (7) *corresponds to* $\boldsymbol{c}_\rho^{-1}$, *where* $\boldsymbol{c}_\rho \in H^2(R_{\phi^\sharp,\sigma^\sharp}, \mathbb{C}^\times)$ *is the obstruction for extending* $\rho$ *to a representation of the preimage in* $\widetilde{\mathcal{S}}_{\phi^\sharp}$ *of* $R_{\phi^\sharp,\sigma^\sharp}$ (see Definition 4.3.1).

*If* $G^\sharp$ *is quasisplit*, $Z^M(\sigma)^\perp = X^G(\pi)$ *and* $\widetilde{R}_{\sigma^\sharp} \to R_{\sigma^\sharp}$ *splits.*

As mentioned in the Introduction, this settles Arthur's conjectures on $R$-groups for $G^\sharp$.

*Proof.* The first part is nothing but a special case of Proposition 4.1.2. The second part then results from the proof of Proposition 4.2.4; the independence of the choice of $\sigma$ is clear. The third part results from Propositions 6.2.2 and 6.2.3. The fourth part is the combination of Proposition 4.3.5 and Theorem 5.4.1.

Finally, $S^G(\pi)$ is commutative when $G^\sharp$ is quasisplit, as $\chi_G = 1$. Hence we have $Z^M(\sigma)^\perp = X^G(\pi)$ and $\rho$ can always be extended in that case. □

**Remark 6.2.5.** The decomposition of $I_{P^\sharp}^{G^\sharp}(\sigma^\sharp)$ depends on $\phi_M^\sharp$, but not on the element $\sigma^\sharp$. This is not expected to hold for other groups.

**Remark 6.2.6.** We have limited ourselves to the tempered representations. However, if the local Langlands correspondence (Theorem 5.2.2) and Theorem 5.3.5 can be extended to Arthur parameters $\psi^\sharp : \mathrm{WD}_F \times \mathrm{SU}(2) \to {}^{\mathrm{L}}G^\sharp$ (see [Arthur 1989b, Section 6]), our results should be applicable to Arthur packets $\Pi_{\psi^\sharp}$ as well, except the part concerning the Knapp–Stein $R$-groups $R_{\sigma^\sharp}$. Note that the crucial lifting Theorem 3.5.1 also holds for Arthur parameters; see [Labesse 1985, Remarque 8.2].

**6.3. *Examples.*** The next example on $R$-groups will be constructed using Steinberg representations, whose definitions are reviewed below.

**Definition 6.3.1.** For the moment, we assume $G$ to be any connected reductive $F$-group. Fix a minimal parabolic subgroup $P_0$ of $G$. The Steinberg representation $\mathrm{St}_G$ of $G$ is the virtual character of $G(F)$ given by

$$\mathrm{St}_G := \sum_{\substack{P \supset P_0 \\ P = MU}} (-1)^{\dim \mathfrak{a}_M^G} I_P^G(\delta_P^{-1/2} \mathbb{1}_M),$$

where the sum ranges over the parabolic subgroups $P$ containing $P_0$ and $\mathbb{1}_M$ denotes the trivial representation of $M(F)$.

The basic fact [Casselman 1973] is that $\mathrm{St}_G$ comes from a smooth irreducible representation in $\Pi_{2,\mathrm{temp}}(G)$, which we denote by the same symbol $\mathrm{St}_G$. It is clearly independent of the choice of $P_0$.

**Lemma 6.3.2.** *For $G$ as in Definition 6.3.1 and a subgroup $G^\sharp$ satisfying*

$$G_{\mathrm{der}} \subset G^\sharp \subset G,$$

*we have*

$$\mathrm{St}_G|_{G^\sharp} \simeq \mathrm{St}_{G^\sharp}.$$

*In particular, the group $X^G(\mathrm{St}_G)$ defined in Section 3.1 is trivial.*

*Proof.* Recall the bijection $P \mapsto P^\sharp := P \cap G^\sharp$ between the parabolic subgroups of $G$ and $G^\sharp$. Since $(\mathbb{1}_L)|_{L^\sharp} = \mathbb{1}_{L^\sharp}$ for any Levi subgroup $L$ of $G$, the first isomorphism follows by comparing the formulas defining $\mathrm{St}_G$ and $\mathrm{St}_{G^\sharp}$, together with Lemma 3.2.1. Hence the restriction of $\mathrm{St}_G$ to $G^\sharp$ is irreducible. It follows from Theorem 3.1.5 that $X^G(\mathrm{St}_G) = \{1\}$. □

Let us revert to the setting $G = \mathrm{GL}_D(n)$ and $G^\sharp = \mathrm{SL}_D(n)$.

**Example 6.3.3.** We now set out to construct an example in which $\widetilde{R}_{\sigma^\sharp} \to R_{\sigma^\sharp}$ does not split for every $\sigma^\sharp \hookrightarrow \sigma|_M$.

First of all, there exists $\mathrm{GL}_D(m)$, for some choice of $D$, $m$, and a representation $\tau \in \Pi_{2,\mathrm{temp}}(\mathrm{GL}_D(m))$ such that $S^{\mathrm{GL}_D(m)}(\sigma)$ is noncommutative. Indeed, for $m = 1$ and $D$ equal to the quaternion algebra over $F$, Arthur exhibits [2006, Page 215] an L-parameter $\phi_\tau \in \Phi_{2,\mathrm{temp}}(D^\times)$ such that

- $\widetilde{\mathcal{S}}_{\phi_\tau^\sharp}$ is isomorphic to the quaternion group of order 8;
- $\widetilde{Z}_{\phi_\tau}$ corresponds to $\{\pm 1\}$.

In fact, $\phi_\tau$ factors through a homomorphism $\mathrm{Gal}(K/F) \to \mathrm{PGL}(2, \mathbb{C})$, where $K$ is a biquadratic extension of $F$, whose image is generated by the elements $\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ in $\mathrm{PGL}(2, \mathbb{C})$.

Since $\chi_{D^\times}$ is injective on $\widetilde{Z}_{\phi_\tau}$ by (19), Theorem 5.3.5 entails that $S^{D^\times}(\tau)$ is noncommutative.

Take $\eta, \omega \in X^{\mathrm{GL}_D(m)}(\tau)$ so that their preimages in $S^{\mathrm{GL}_D(m)}(\tau)$ do not commute. Let $c$ (respectively $d$) be the order of $\eta$ (respectively $\omega$). Put $\mathrm{St} := \mathrm{St}_{\mathrm{GL}_D(m)}$ and take

$$\begin{aligned} M &:= \mathrm{GL}_D(m) \times \prod_{\substack{1 \leq i \leq c \\ 1 \leq j \leq d}} \mathrm{GL}_D(m), \\ G &:= \mathrm{GL}_D(m(cd+1)), \\ \sigma &:= \tau \boxtimes \mathop{\boxtimes}_{\substack{1 \leq i \leq c \\ 1 \leq j \leq d}} \eta^{i-1}\omega^{j-1}\mathrm{St}, \\ \pi &:= I_P^G(\sigma) \quad \text{for some } P \in \mathscr{P}(M). \end{aligned}$$

Here $X^M(\sigma)$ is defined relatively to $M^\sharp := G^\sharp \cap M$ where $G^\sharp = \mathrm{SL}_D(m(cd+1))$. The presence of St forces $X^M(\sigma)$ to be trivial, by Lemma 6.3.2. Hence $\sigma^\sharp := \sigma|_{M^\sharp}$ is irreducible, and it is parametrized by the 1-dimensional character $\chi_M : \widetilde{Z}_{\phi^\sharp_M} \to \mathbb{C}^\times$. According to Theorem 6.2.4, the central extension $\widetilde{R}_{\sigma^\sharp} \twoheadrightarrow R_{\sigma^\sharp}$ splits if and only if $\rho$ can be extended to $\widetilde{\mathscr{S}}_{\phi^\sharp}$. This is the case if and only if $S^G(\pi) \twoheadrightarrow X^G(\pi)$ splits, by Theorem 5.3.5. Hence it suffices to show the noncommutativity of $S^G(\pi)$.

Put $L := \mathrm{GL}_D(m) \times \mathrm{GL}_D(mcd) \in \mathscr{L}^G(M)$ and set $\nu := I^L_{P \cap L}(\sigma)$. We claim $\eta, \omega \in S^L(\nu)$. Indeed, the $\mathrm{GL}_D(m)$-component of $L$ does not cause any problem. As for the $\mathrm{GL}_D(mcd)$-component, take representatives $\tilde{w}_\eta$, $\tilde{w}_\eta$ in $\mathrm{SL}_D(mcd)$ of the cyclic permutations

$$\begin{aligned} w_\eta &: 1 \to \cdots \to c \to 1, \\ w_\omega &: 1 \to \cdots \to d \to 1 \end{aligned}$$

of the indexes $i$ and $j$, respectively. Then the intertwining operators $J_\eta$, $J_\omega$ are given by the operators in (5) using $\tilde{w}_\eta$, $\tilde{w}_\omega$. Furthermore, $J_\eta$ and $J_\omega$ commute with each other; this follows from (6) and the obvious fact that $\tilde{w}_\eta$ and $\tilde{w}_\omega$ can be chosen to commute.

From our choice of $\eta, \omega$, it follows that the preimages of $\eta, \omega$ in $S^L(\nu)$ do not commute. Since $S^L(\nu) \hookrightarrow S^G(\pi)$ by Proposition 4.1.1, $S^G(\pi)$ is noncommutative, as required.

**Example 6.3.4.** Now we set out to show that the inclusion $R_{\phi^\sharp, \sigma^\sharp} \subset R_{\phi^\sharp}$ is proper in general. By Theorem 6.2.4 and the notations therein, it amounts to showing that $Z^M(\sigma)^\perp \subsetneq X^G(\pi)$ in general.

As in the previous example, we take some $m \geq 1$, a central division $F$-algebra $D$, and $\tau \in \Pi_{2,\mathrm{temp}}(\mathrm{GL}_D(m))$ such that $X^{\mathrm{GL}_D(m)}(\tau)$ contains $\eta, \omega$ with noncommuting preimages in $S^{\mathrm{GL}_D(m)}(\sigma)$. Take another $\tau' \in \Pi_{2,\mathrm{temp}}(\mathrm{GL}_D(m))$ such that

$X^{\mathrm{GL}_D(m)}(\tau') = \langle \eta \rangle$. Denote by $d$ the order of $\omega$. We take

$$\begin{aligned} M &:= \mathrm{GL}_D(m) \times \prod_{1 \le j \le d} \mathrm{GL}_D(m), \\ G &:= \mathrm{GL}_D(m(d+1)), \\ \sigma &:= \tau \boxtimes \mathop{\boxtimes}_{1 \le j \le d} \omega^{j-1}\tau', \\ \pi &:= I_P^G(\sigma) \quad \text{for some } P \in \mathcal{P}(M). \end{aligned}$$

Therefore $X^M(\sigma) = \langle \eta \rangle$ (defined relative to $M^\sharp = M \cap G^\sharp$ with $G^\sharp := \mathrm{SL}_D(m(d+1))$ as before), and $S^M(\sigma)$ is commutative. In particular $Z^M(\sigma) = X^M(\sigma) = \langle \eta \rangle$.

On the other hand, a variant of the arguments in the previous example show that $\omega, \eta \in X^G(\pi)$ with noncommuting preimages in $S^G(\pi)$. Hence $\omega \in X^G(\pi)$ and $\omega \notin Z^M(\sigma)^\perp$, as required.

Note that such $\tau, \tau'$ do exist when $D$ is the quaternion algebra over $F$ and $m = 1$; in that case $\eta, \omega$ are identified with quadratic characters of $F^\times$. Indeed, a candidate of $\tau$ is given in the previous example. On the other hand, to construct $\tau'$ for a given $\eta$, we are reduced to constructing $\tau'' \in \Pi_{2,\mathrm{temp}}(\mathrm{GL}_F(2))$ with $X^{\mathrm{GL}_F(2)}(\tau'') = \{1, \eta\}$ and then taking $\tau$ to be the Jacquet–Langlands transfer of $\tau''$.

To finish the construction, let $E$ be the quadratic extension of $F$ determined by $\eta$ and let $\theta : E^\times \to \mathbb{C}^\times$ be a continuous character. Set $\tau'' := \mathrm{Ind}_{E/F}(\theta)$ (the local automorphic induction; cf. [Jacquet and Langlands 1970, Theorem 4.6]). Then $\eta\tau'' = \tau''$. From [Labesse and Langlands 1979, Pages 738–739], one sees that $\tau''$ is cuspidal and $|X^{\mathrm{GL}_F(2)}(\tau'')| = 2$ for general $\theta$, which suffices to conclude.

## Acknowledgements

We thank Kwangho Choiy and David Goldberg for communicating their results to us, as well as the referee for a meticulous reading. Li is grateful to Dipendra Prasad for his interest in this work and helpful comments.

## References

[Arthur 1989a] J. Arthur, "Intertwining operators and residues, I: Weighted characters", *J. Funct. Anal.* **84**:1 (1989), 19–84. MR 90j:22018 Zbl 0679.22011

[Arthur 1989b] J. Arthur, "Unipotent automorphic representations: conjectures", pp. 13–71 in *Orbites unipotentes et représentations, II: Groupes p-adiques et réels*, Astérisque **171–172**, Société Mathématique de France, Paris, 1989. MR 91f:22030 Zbl 0728.22014

[Arthur 1993] J. Arthur, "On elliptic tempered characters", *Acta Math.* **171**:1 (1993), 73–138. MR 94i:22038 Zbl 0822.22011

[Arthur 1996] J. Arthur, "On local character relations", *Selecta Math.* (*N.S.*) **2**:4 (1996), 501–579. MR 2000a:22017 Zbl 0923.11081

[Arthur 1999] J. Arthur, "On the transfer of distributions: weighted orbital integrals", *Duke Math. J.* **99**:2 (1999), 209–283. MR 2000i:22023 Zbl 0938.22019

[Arthur 2006] J. Arthur, "A note on $L$-packets", *Pure Appl. Math. Q.* **2**:1 (2006), 199–217. MR 2006k:22014 Zbl 1158.22017

[Arthur 2008] J. Arthur, "Problems for real groups", pp. 39–62 in *Representation theory of real reductive Lie groups* (Snowbird, UT, 2006), edited by J. Arthur et al., Contemp. Math. **472**, Amer. Math. Soc., Providence, RI, 2008. MR 2010c:22014 Zbl 1159.22003

[Arthur 2013] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, Colloquium Publications **61**, Amer. Math. Soc., Providence, RI, 2013.

[Aubert and Plymen 2005] A.-M. Aubert and R. Plymen, "Plancherel measure for $\mathrm{GL}(n, F)$ and $\mathrm{GL}(m, D)$: explicit formulas and Bernstein decomposition", *J. Number Theory* **112**:1 (2005), 26–66. MR 2006d:22018 Zbl 1063.22015

[Ban and Goldberg 2012] D. Ban and D. Goldberg, "$R$-groups and parameters", *Pacific J. Math.* **255**:2 (2012), 281–303. MR 2928553 Zbl 06049810

[Ban and Zhang 2005] D. Ban and Y. Zhang, "Arthur $R$-groups, classical $R$-groups, and Aubert involutions for $\mathrm{SO}(2n+1)$", *Compositio Math.* **141**:2 (2005), 323–343. MR 2006d:22019 Zbl 1131.11033

[Borel 1979] A. Borel, "Automorphic $L$-functions", pp. 27–61 in *Automorphic forms, representations and $L$-functions* (Corvallis, OR, 1977), vol. 2, edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR 81m:10056 Zbl 0412.10017

[Casselman 1973] W. Casselman, "The Steinberg character as a true character", pp. 413–417 in *Harmonic analysis on homogeneous spaces* (Williamstown, MA, 1972), edited by C. C. Moore, Proc. Sympos. Pure Math. **26**, Amer. Math. Soc., Providence, RI, 1973. MR 49 #3039 Zbl 0289.22017

[Deligne et al. 1984] P. Deligne, D. Kazhdan, and M.-F. Vignéras, "Représentations des algèbres centrales simples $p$-adiques", pp. 33–117 in *Représentations des groups réductifs sur un corps local*, edited by J.-N. Bernstein et al., Hermann, Paris, 1984. MR 86h:11044 Zbl 0583.22009

[Gelbart and Knapp 1981] S. S. Gelbart and A. W. Knapp, "Irreducible constituents of principal series of $\mathrm{SL}_n(k)$", *Duke Math. J.* **48**:2 (1981), 313–326. MR 82j:22018 Zbl 0473.22010

[Gelbart and Knapp 1982] S. S. Gelbart and A. W. Knapp, "$L$-indistinguishability and $R$ groups for the special linear group", *Adv. in Math.* **43**:2 (1982), 101–121. MR 83j:22009 Zbl 0493.22005

[Goldberg 1994] D. Goldberg, "$R$-groups and elliptic representations for $\mathrm{SL}_n$", *Pacific J. Math.* **165**:1 (1994), 77–92. MR 95h:22020 Zbl 0855.22016

[Goldberg 2006] D. Goldberg, "Reducibility for $\mathrm{SU}_n$ and generic elliptic representations", *Canad. J. Math.* **58**:2 (2006), 344–361. MR 2007d:22026 Zbl 1100.22011

[Goldberg 2011] D. Goldberg, "On dual $R$-groups for classical groups", pp. 159–185 in *On certain $L$-functions* (West Lafayette, IN, 2007), edited by J. Arthur et al., Clay Math. Proc. **13**, Amer. Math. Soc., Providence, RI, 2011. MR 2012e:22022 Zbl 05932913

[Goldberg and Choiy 2012] D. Goldberg and K. Choiy, "Transfer of $R$-groups between $p$-adic inner forms of $\mathrm{SL}_n$", preprint, 2012. arXiv 1211.5054

[Hanzer 2004] M. Hanzer, "$R$ groups for quaternionic Hermitian groups", *Glas. Mat. Ser. III* **39**:1 (2004), 31–48. MR 2005d:22008 Zbl 1056.22010

[Hiraga and Saito 2012] K. Hiraga and H. Saito, "On $L$-packets for inner forms of $SL_n$", *Mem. Amer. Math. Soc.* **215**:1013 (2012). MR 2918491 Zbl 1242.22023

[Jacquet and Langlands 1970] H. Jacquet and R. P. Langlands, *Automorphic forms on* GL(2), Lecture Notes in Mathematics **114**, Springer, Berlin, 1970. MR 53 #5481 Zbl 0236.12010

[Kazhdan and Patterson 1984] D. A. Kazhdan and S. J. Patterson, "Metaplectic forms", *Inst. Hautes Études Sci. Publ. Math.* **59** (1984), 35–142. MR 85g:22033 Zbl 0559.10026

[Keys 1987] C. D. Keys, "$L$-indistinguishability and $R$-groups for quasisplit groups: unitary groups in even dimension", *Ann. Sci. École Norm. Sup.* (4) **20** (1987), 31–64. MR 88m:22042 Zbl 0634.22014

[Kottwitz 1984] R. E. Kottwitz, "Stable trace formula: cuspidal tempered terms", *Duke Math. J.* **51**:3 (1984), 611–650. MR 85m:11080 Zbl 0576.22020

[Labesse 1985] J.-P. Labesse, "Cohomologie, $L$-groupes et fonctorialité", *Compositio Math.* **55**:2 (1985), 163–184. MR 86j:11117 Zbl 0595.12004

[Labesse and Langlands 1979] J.-P. Labesse and R. P. Langlands, "$L$-indistinguishability for SL(2)", *Canad. J. Math.* **31**:4 (1979), 726–785. MR 81b:22017 Zbl 0421.12014

[Langlands 1976] R. P. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics **544**, Springer, Berlin, 1976. MR 58 #28319 Zbl 0332.10018

[Lemaire 2010] B. Lemaire, "Caractères tordus des représentations admissibles", preprint, 2010. arXiv 1007.3576

[Mok 2012] C. P. Mok, "Endoscopic classification of representations of quasi-split unitary groups I", preprint, 2012. arXiv 1206.0882

[Sécherre 2009] V. Sécherre, "Proof of the Tadić conjecture (U0) on the unitary dual of $\mathrm{GL}_m(D)$", *J. Reine Angew. Math.* **626** (2009), 187–203. MR 2009k:22027 Zbl 1170.22009

[Serre 1967] J.-P. Serre, *Représentations linéaires des groupes finis*, Hermann, Paris, 1967. MR 38 #1190 Zbl 0189.02603

[Shahidi 1983] F. Shahidi, "Some results on $L$-indistinguishability for SL($r$)", *Canad. J. Math.* **35**:6 (1983), 1075–1109. MR 85i:22030 Zbl 0553.10024

[Shahidi 1990] F. Shahidi, "A proof of Langlands' conjecture on Plancherel measures; complementary series for $p$-adic groups", *Ann. of Math.* (2) **132**:2 (1990), 273–330. MR 91m:11095 Zbl 0780.22005

[Shelstad 1982] D. Shelstad, "$L$-indistinguishability for real groups", *Math. Ann.* **259**:3 (1982), 385–430. MR 84c:22017 Zbl 0506.22014

[Silberger 1978] A. J. Silberger, "The Knapp–Stein dimension theorem for $p$-adic groups", *Proc. Amer. Math. Soc.* **68**:2 (1978), 243–246. MR 58 #11245 Zbl 0348.22007

[Silberger 1979] A. J. Silberger, "Isogeny restrictions of irreducible admissible representations are finite direct sums of irreducible admissible representations", *Proc. Amer. Math. Soc.* **73**:2 (1979), 263–264. MR 80f:22017 Zbl 0368.22009

[Tadić 1992] M. Tadić, "Notes on representations of non-Archimedean SL($n$)", *Pacific J. Math.* **152**:2 (1992), 375–396. MR 92k:22029 Zbl 0724.22017

[Waldspurger 2003] J.-L. Waldspurger, "La formule de Plancherel pour les groupes $p$-adiques (d'après Harish-Chandra)", *J. Inst. Math. Jussieu* **2**:2 (2003), 235–333. MR 2004d:22009 Zbl 1029.22016

Received December 2, 2012.

KUOK FAI CHAO
SHANGHAI CENTER FOR MATHEMATICAL SCIENCES
FUDAN UNIVERSITY
NO. 220 HANDAN ROAD
SHANGHAI, 200433
CHINA
kchao@fudan.edu.cn

WEN-WEI LI
MORNINGSIDE CENTER OF MATHEMATICS
ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
55, ZHONGGUANCUN EAST ROAD
BEIJING, 100190
CHINA
wwli@math.ac.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.91

# AUTOMORPHISMS AND QUOTIENTS OF QUATERNIONIC FAKE QUADRICS

AMIR DŽAMBIĆ AND XAVIER ROULLEAU

**A $\mathbb{Q}$-homology quadric is a normal projective algebraic surface with the same Betti numbers as the smooth quadric in $\mathbb{P}^3$. A smooth $\mathbb{Q}$-homology quadric is either rational or of general type with vanishing geometric genus. Smooth minimal $\mathbb{Q}$-homology quadrics of general type are called fake quadrics. Here we study quaternionic fake quadrics, that is, fake quadrics whose fundamental group is an irreducible lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ derived from a division quaternion algebra over a real number field. We provide examples of quaternionic fake quadrics $X$ with a nontrivial automorphism group $G$ and compute the invariants of the quotient $X/G$ and of its minimal desingularization $Z$. In this way we provide examples of singular $\mathbb{Q}$-homology quadrics and minimal surfaces $Z$ of general type with $q = p_g = 0$ and $K^2 = 4$ or 2 which contain the maximal number of disjoint $(-2)$-curves. Conversely, we also show that if a smooth minimal surface of general type has the same invariant as $Z$ and same number of $(-2)$-curves, then we can construct geometrically a surface of general type with $c_1^2 = 8$, $c_2 = 4$.**

## 1. Introduction

In this paper we will be interested in *$\mathbb{Q}$-homology quadric surfaces*, which are normal projective algebraic surfaces with the same Betti numbers as the quadric surface in $\mathbb{P}^3$, that is, $b_1 = 0$ and $b_2 = 2$. A smooth $\mathbb{Q}$-homology quadric $S$ has the following numerical invariants: $p_g(S) = q(S) = 0$, $e(S) = c_2(S) = 4$, and $c_1^2(S) = 8$. By the classification theory of algebraic surfaces, such $S$ is either a Hirzebruch surface $\Sigma_n$ (with $\Sigma_0 = \mathbb{P}^1 \times \mathbb{P}^1$) or $S$ is of general type. The latter $S$ is either minimal or has at most one exceptional curve. Blowing down this $(-1)$-curve we obtain a *fake projective plane*, that is, a smooth minimal surface of general type with the same Betti numbers as the projective plane $\mathbb{P}^2$, being an example of a $\mathbb{Q}$-homology projective plane. By the analogy with fake projective planes, we define a *fake quadric* to be a minimal smooth $\mathbb{Q}$-homology quadric of general type (see [Barth et al. 2004, p. 231; Hirzebruch 1987, p. 780; Iskovskikh and Shafarevich 1989, p. 195]).

*MSC2010:* primary 14G35, 14J10, 14J29; secondary 14J50, 11F06, 11R52.

*Keywords:* $\mathbb{Q}$-homology quadrics, surfaces with $q = p_g = 0$, fake quadrics, surfaces of general type, automorphisms.

All known fake quadrics have $\mathbb{H} \times \mathbb{H}$, the product of two copies of the complex upper half plane, as the universal covering. Hence, such a fake quadric $X$ is of the form $X = \Gamma \backslash \mathbb{H} \times \mathbb{H}$, where $\Gamma$ is a torsion-free and cocompact discrete subgroup in $\mathrm{Aut}(\mathbb{H} \times \mathbb{H})$, the group of holomorphic automorphisms of $\mathbb{H} \times \mathbb{H}$. Essentially, we distinguish between two classes of such quotients according to the structure of $\Gamma$.

One class of fake quadrics consists of surfaces $\Gamma \backslash \mathbb{H} \times \mathbb{H}$ with the property that the group $\Gamma$ is reducible. By reducible we mean that there exists a subgroup of finite index $\Gamma' = \Gamma_1 \times \Gamma_2$ of $\Gamma$ such that the group $\Gamma_i$ acts on $\mathbb{H}$ and $C_i = \mathbb{H}/\Gamma_i$ is a smooth algebraic curve. This case is now well understood and the full classification of these fake quadrics, named also fake quadrics *isogenous to a higher product*, has been achieved by Bauer, Catanese and Grunewald in [Bauer et al. 2008]. In practice, this classification and construction is done geometrically by classifying triples $(C_1, C_2, G)$ of two smooth curves $C_i$ of general type and a group $G$, such that $G$ acts faithfully and freely on the surface $C_1 \times C_2$ and the quotient $(C_1 \times C_2)/G$ has the asked invariants.

In this paper we will focus on fake quadrics of the other class, which we call *quaternionic fake quadrics*. These fake quadrics are *Shimura surfaces*, that is, quotients of $\mathbb{H} \times \mathbb{H}$ by cocompact irreducible arithmetic lattices $\Gamma$ in $\mathrm{Aut}(\mathbb{H}) \times \mathrm{Aut}(\mathbb{H})$, defined by an indefinite quaternion algebra over a totally real number field. Within the general framework of Prasad and Yeung on fake compact symmetric Hermitian spaces the quaternionic fake quadrics belong to the class of so-called *arithmetic fake* $A_1$; see [Prasad and Yeung 2012].

Using the previous work of Kuga, the first quaternionic fake quadrics have been constructed in [Shavel 1978]. We know that these surfaces are rigid and thus that there are only a finite number of them, but at the moment we do not have a complete list of all these surfaces. We have a list of commensurability classes of fake quadrics defined by quaternion algebras over quadratic fields (see [Džambić 2013]).

The situation for quaternionic fake quadrics is very similar to the case of fake projective planes. By the theorem of Klingler (and also Yeung), all fake projective planes are quotients of the 2-dimensional complex unit ball $\mathbb{B}^2$ by cocompact arithmetic lattices $\Gamma \subset \mathrm{PU}(2, 1)$. This provides an arithmetic construction of these surfaces, but it is generally not easy to handle and construct these surfaces geometrically, for instance, as a quotient or ramified cover of some known surfaces.

In order to remedy this situation, Keum [2012; 2008; 2006] studied quotients of fake projective planes by groups of automorphisms. In this way, he obtained surfaces of general type with geometric genus $p_g = 0$ and was able to rebuild a fake projective plane by only knowing the properties of the quotient surface.

The aim of this paper is to study automorphisms of quaternionic fake quadrics and the quotients of these surfaces by groups of automorphisms. Let $X = \Gamma \backslash \mathbb{H} \times \mathbb{H}$

be a Shimura surface. We say that a curve $C \hookrightarrow X$ is a Shimura curve if it is a totally geodesic submanifold of $X$.

The first main result we obtain is the following:

**Theorem A.** *An automorphism of a smooth Shimura surface $X = \Gamma\backslash\mathbb{H}\times\mathbb{H}$ has only finitely many fixed points or it is an involution whose fixed point set is a disjoint union of smooth Shimura curves.*

*An automorphism of a quaternionic fake quadric $X$ has only finitely many fixed points. There exist quaternionic fake quadrics $X$ with automorphism group isomorphic to*

$$\mathbb{Z}/2\mathbb{Z},\ (\mathbb{Z}/2\mathbb{Z})^2,\ \mathbb{D}_4,\ \mathbb{D}_6,\ \mathbb{D}_8,\ \textit{or}\ \mathbb{D}_{10},$$

*where $\mathbb{D}_n$ is the dihedral group with order $2n$.*

Let us remark that the knowledge of surfaces of general type with $p_g = 0$ and a large automorphism group can be interesting to check whether the Bloch conjecture holds (see, for example, [Inose and Mizukami 1979]). The computations in [Džambić 2013] lead us to the conjecture that the order of the automorphism group of a quaternionic fake quadric is always less or equal 24 (see Section 4).

The second aim of this paper is to study the minimal desingularization of the quotient of a quaternionic fake quadric by a group of automorphisms, in order to obtain new surfaces with $p_g = 0$.

**Theorem B.** *Let $X$ be a quaternionic fake quadric and $G$ a finite group of automorphisms of $X$. The minimal desingularization $Z$ of the quotient $X/G$ has the following numerical invariants*:

| $G$ | $c_1^2(Z)$ | $c_2(Z)$ | Singularities on $X/G$ | Minimal | $\kappa(Z)$ |
|---|---|---|---|---|---|
| $\mathbb{Z}/2\mathbb{Z}$ | 4 | 8 | $4A_1$ | yes | 2 |
| $\mathbb{Z}/3\mathbb{Z}$ | 2 | 10 | $2A_{3,1}+2A_2$ | ? | 2 |
| $\mathbb{Z}/6\mathbb{Z}$ | $-4$ | 16 | $2A_{6,1}+2A_5$ | no | ? |
| $\mathbb{Z}/8\mathbb{Z}$ | $-2$ | 14 | $A_{8,3}+A_{8,5}$ | no | ? |
| $\mathbb{Z}/10\mathbb{Z}$ | $-12$ | 24 | $2A_{10,1}+2A_9$ | no | ? |
| $(\mathbb{Z}/2\mathbb{Z})^2$ | 2 | 10 | $6A_1$ | yes | 2 |
| $\mathbb{D}_4$ | 0 | 12 | $4A_1+A_{4,3}+A_{4,1}$ | no | $\geq 1$ |
| $\mathbb{D}_8$ | $-1$ | 13 | $4A_1+A_{8,3}+A_{8,5}$ | no | ? |

*Here*, *$\kappa$ indicates the Kodaira dimension of the surface $Z$.*

We obtain also results and restrictions for the groups $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/5\mathbb{Z}$ and $\mathbb{D}_3$. We note that the surfaces of general type we obtain have vanishing geometric genus and could be therefore interesting from the point of view of the classification of surfaces with $p_g = 0$. We intend to study these surfaces more closely, regarding, for instance, the fundamental groups in a future paper.

A curve $C$ on a surface is called nodal if $C \simeq \mathbb{P}^1$ and $C^2 = -2$. A nodal curve is the resolution of a nodal singularity. The surfaces $Z$ we obtain as the quotient of a fake quadric by an automorphism group $(\mathbb{Z}/2\mathbb{Z})^n$, $n \in \{1, 2\}$ have the maximum number of nodal curves (the so-called Miyaoka bound [1984]). If minimal, the surfaces obtained by taking a quotient by the groups $\mathbb{Z}/3\mathbb{Z}$ and $\mathbb{D}_3$ have also the maximum number of quotient singularities. Similarly to Keum's construction of fake projective planes, we can reverse the construction.

**Proposition C.** *Let $Z$ be a smooth minimal surface of general type with $q = p_g = 0$.*

(a) *Suppose that $c_1^2 = 4$, $2$, or $1$, $\mathrm{Pic}(Z)$ has no 2-torsion, and that there is a birational map $Z \to Y$ onto a surface containing $8 - c_1^2$ nodal singularities $A_1$. There exists a smooth minimal surface of general type $S$ with invariants $c_1^2 = 2c_2 = 8$ and a $(\mathbb{Z}/2\mathbb{Z})^m$-cover $S \to Y$ ramified over the nodes, with $m$ such that $2^m = 8/c_1^2$.*

(b) *Suppose that $c_1^2 = 2$, $\mathrm{Pic}(Z)$ has no 3-torsion, and that there is a birational map $Z \to Y$ onto a surface with $2A_{3,1} + 2A_2$ singularities. There exist a smooth surface $S$ with invariants $c_1^2 = 2c_2 = 8$ and a $(\mathbb{Z}/3\mathbb{Z})$-cover $Z \to Y$ ramified over the singularities of $Y$.*

The proof of part (a) of this proposition uses mainly the results of Dolgachev, Mendes Lopes, and Pardini [Dolgachev et al. 2002] and illustrates their theory. The proof of part (b) is more original because it mixes two types of singularities.

The paper is structured as follows: We begin recalling the known facts on quotients of surfaces (Section 2) and on quaternionic fake quadrics (Section 3). In Section 4, we provide examples of fake quadrics having a large group of automorphisms, we then compute the quotients surfaces (Section 5) and reverse the construction in the opposite direction: starting with a surface with the same invariants as the quotient, we construct a surface with $c_1^2 = 2c_2 = 8$ (Section 6).

## 2. Generalities on quotients of a surface

In this section we recall results from the theory of quotient surface singularities and their resolution. The main reference for these topics is [Barth et al. 2004]; see also [Roulleau 2012].

Let $S$ be a smooth algebraic surface and let $G$ be a group of automorphisms acting on $S$. We denote by $S/G$ the quotient surface and by $\pi : Z \to S/G$ the minimal desingularization map. If $G = \langle\sigma\rangle$ is cyclic, we will often write $S/\sigma$ to denote the quotient $S/\langle\sigma\rangle$.

**Proposition 2.4** (topological Lefschetz formula). *Let $\sigma$ be an automorphism acting on $S$ and $S^\sigma$ the fixed point set of $\sigma$. We have*

$$e(S^\sigma) = \sum_{j=0}^{j=4} (-1)^j \operatorname{Tr}(\sigma | H^i(S, \mathbb{Z})_{mt}),$$

*where $H^i(S, \mathbb{Z})_{mt}$ is the group $H^i(S, \mathbb{Z})$ modulo torsion.*

Note that for a fake quadric $X$ we have $q = p_g = 0$; thus

$$H^1(X, \mathbb{Z})_{mt} = \{0\}, \; H^2(X, \mathbb{Z}) \otimes \mathbb{C} = H^1(X, \Omega_X).$$

**Corollary 2.5.** *Let $X$ be a fake quadric and $\sigma$ an automorphism of order $n > 1$ acting on $G$. We have $e(X^\sigma) = 2$ or 4. If $\sigma = \tau^2$ for an automorphism $\tau$ (for example, if $n$ is prime to 2), we have $e(X^\sigma) = 4$.*

*Proof.* Since $X$ is a fake quadric, the space $H^1(X, \Omega_X)$ is 2-dimensional and is generated by the classes of 2-curves in the Néron–Severi group. As an automorphism preserves the canonical divisor, the invariant subspace of $H^1(X, \Omega_X)$ is at least one-dimensional. Therefore the trace of $\sigma$ on $H^1(X, \Omega_X)$ is 2 or 0. If we suppose that this action is not trivial, then 2 divides the order of $\sigma$, moreover we see that the action of $\sigma^2$ is always trivial. □

Let $\xi$ be a primitive $n$-th root of unity. Let us recall that for $1 \le q \le n-1$ coprime to $n$, the quotient of $\mathbb{C}^2$ by the action of

$$(x, y) \to (\xi x, \xi^q y)$$

has a unique singularity, called an $A_{n,q}$ singularity. For $n, m > 0$ two numbers, we write $[n, m]$ for $n - 1/m$. The $A_{n,q}$ singularity is resolved by a Hirzebruch–Jung string (see [Barth et al. 2004]), that is, a chain of smooth rational curves $C_1, \dots, C_k$ such that $C_i$ intersects $C_{i\pm1}$ transversally in one point for $2 \le i \le k-1$ and $C_i^2 = -n_i$ with integers $n_i \ge 2$ determined by the relation

$$\frac{n}{q} = [n_1, [n_2, \dots, [n_{k-1}, n_k] \dots ]].$$

As is conventional, we denote $A_{n,n-1}$ by $A_{n-1}$.

Let $S$ be a surface with $p_g = q = 0$ and let $\sigma$ be an automorphism of order $n \ge 2$ such that the fixed points of the $\sigma^k$, $k = 1, \dots, n-1$ are isolated.

**Proposition 2.6** (holomorphic Lefschetz fixed point formula [Atiyah and Singer 1968, p. 567]). *Let $S^\sigma$ be the fixed point set of $\sigma$. Then*

$$1 = \sum_{s \in S^\sigma} \frac{1}{\det(1 - d\sigma | T_{S,s})},$$

*where $d\sigma_s | T_{S,s}$ denotes the action of $\sigma$ on the tangent space $T_{S,s}$.*

Suppose moreover that the automorphism $\sigma$ has prime order $p$. Let $\xi$ be a primitive $p$-th root of unity. Let $r_i$ be the number of isolated fixed points of $\sigma$ whose image in $S/\sigma$ are $A_{p,i}$ singularities.

**Proposition 2.7** (Zhang's formula [2001, Lemma 1.6]). *We have*

$$\sum_{i=1}^{p-1} r_i a_i(p) = 1$$

*where*

$$a_i(p) = \frac{1}{p-1} \sum_{j=1}^{p-1} \frac{1}{(1-\xi^j)(1-\xi^{ij})}.$$

*In particular, we have*

$$a_1(p) = \tfrac{1}{12}(5-p), \quad a_2(p) = \tfrac{1}{24}(11-p), \quad a_3(5) = \tfrac{1}{4}, \quad a_4(5) = \tfrac{1}{2}.$$

Let $1 \le i < p$ and $1 \le k < p$ be such that $ik \equiv 1 \bmod p$. As $A_{p,i} = A_{p,k}$, the notations for $r_i$ and $r_k$ in Zhang's formula can be confusing. However, as $a_i(p) = a_k(p)$, there should be no trouble in taking the convention that $r_i + r_k$ is the total number of $A_{p,i} = A_{p,k}$ singularities, rather that choosing a representative $i$ or $k$ for every such pair $(i, k)$.

Let us recall that an automorphism of a vector space is called a reflection if all its eigenvalues but one are equal to 1. Let $S$ be a surface and $G$ an automorphism group acting on $S$. Suppose that for every automorphism of $G$ the fixed point set is finite. Let $s$ be a fixed point of $G$; recall (see [Barth et al. 2004]):

**Lemma 2.8.** *The action of the group $G$ on the tangent space $T_{S,s}$ is faithful and contains no reflections.*

In particular, if $G$ is cyclic of order $n$, the singularity type of the image of the fixed point $s$ in the quotient $S/G$ is always $A_{n,q}$ with $q$ prime to $n$.

**Lemma 2.9.** *The Euler number of $S/G$ is given by the formula*

$$e(S/G) = \frac{1}{|G|}(e(S) + \sum_{n \ge 2} (n-1)e(S_n)),$$

*where $S_n = \left\{ s \in S \;\middle|\; |\mathrm{Stab}(G, s)| = n \right\}$. The Euler number of the minimal resolution $Z$ is the sum of $e(S/G)$ and the number of irreducible components of the exceptional curves of the resolution $\pi : Z \to S/G$.*

Let $C_1, \ldots, C_k$ be the irreducible components of the one dimensional fibers of $\pi : Z \to S/G$. We have the relations $K_Z = \pi^* K_{S/G} - \sum_{i=1}^{i=k} a_i C_i$, for rational numbers $a_i$ such that $K_Z C_k = -2 - C_k^2$ and $C_k \pi^* K_{S/G} = 0$. Moreover, we have the equality $K_{S/G}^2 = K_S^2/|G|$, where $|G|$ is the order of $G$. As $K_S$ is ample, the canonical $\mathbb{Q}$-divisor $K_{S/G}$ is ample and $\pi^* K_{S/G}$ is nef. We remark also that $K_Z^2 \le K_{S/G}^2$.

Recall that the Kähler lemma implies that for a dominant rational map between varieties, the pull back map among the spaces of sections of sheaves of holomorphic forms are injective, therefore we obtain (see also [Roulleau 2012] for a proof in the case of surfaces):

**Lemma 2.10.** *Let $S$ be a surface with $q = p_g = 0$. The minimal resolution $Z$ of the quotient of $S$ by a group $G$ has also $q = p_g = 0$.*

Suppose that $S$ is moreover minimal of general type and the fixed points of automorphisms in $G$ are isolated.

**Lemma 2.11.** *If $K_Z^2 = 0$, the surface $Z$ has Kodaira dimension $\kappa \ge 1$. If $K_Z^2 > 0$, the surface $Z$ has Kodaira dimension $\kappa = 2$.*

*Proof.* (We follow the ideas from [Keum 2008].) The quotient surface has $q = p_g = 0$ and thus $\chi(\mathcal{O}_Z) = 1$. Let $m \ge 1$ be an integer. Then

$$-mK_Z\pi^* K_{S/G} = -mK_{S/G}^2 = -\frac{8}{|G|}m < 0;$$

therefore $H^0(Z, -mK_Z) = \{0\}$ for every $m \ge 1$. Let $m \ge 2$; then from the Serre duality and Riemann–Roch we obtain

$$H^0(Z, mK_Z) = \chi(\mathcal{O}_Z) + \frac{m(m-1)}{2} K_Z^2 + h^1(Z, mK_Z).$$

If $K_Z^2 > 0$, it immediately follows that $Z$ is of general type. If $K_Z^2 = 0$, the surface has $h^0(Z, 2K_Z) \ne 0$ and cannot be rational by the Castelnuovo criterion. Moreover, as $\chi = 1$ it cannot be a ruled surface. Suppose that $Z$ is an Enriques surface. As $K_Z^2 = 0$, it is a minimal surface, but this is impossible because $h^0(Z, 3K_Z) \ne 0$; therefore $\kappa > 0$. □

## 3. Automorphisms of smooth Shimura surfaces and generalities on quaternionic fake quadrics

Let us give a more detailed description of Shimura surfaces and quaternionic fake quadrics. First, recall that a lattice $\Gamma < \mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R}) \cong \mathrm{Aut}\,\mathbb{H} \times \mathrm{Aut}\,\mathbb{H}$ is irreducible if it is not commensurable with a product $\Gamma_1 \times \Gamma_2$ of two discrete subgroups $\Gamma_1, \Gamma_2 \subset \mathrm{PSL}_2(\mathbb{R})$. Equivalently, the image of $\Gamma$ under the projection onto one of the factors $\mathrm{PSL}_2(\mathbb{R})$ is a dense subgroup of $\mathrm{PSL}_2(\mathbb{R})$. Irreducible lattices in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ can be constructed arithmetically in the following way. Let $k$ be a totally real number field of degree $g = [k : \mathbb{Q}] \ge 2$, and let

$B = (\alpha, \beta)_k := \langle 1, i, j, ij \rangle_k$, with $i^2 = \alpha \in k$, $j^2 = \beta \in k$, $ij = -ji$, be a quaternion algebra over $k$ such that

$$B \otimes_{\mathbb{Q}} \mathbb{R} = \prod_{\rho \in \operatorname{Hom}(k, \mathbb{R})} B^\rho \cong M_2(\mathbb{R}) \times M_2(\mathbb{R}) \times \underbrace{H_{\mathbb{R}} \times \cdots \times H_{\mathbb{R}}}_{g-2}. \tag{3-1}$$

Here, $B^\rho = (\alpha^\rho, \beta^\rho)_{\mathbb{R}}$ and $H_{\mathbb{R}} = (-1, -1)_{\mathbb{R}}$ denotes the skew field of Hamiltonian quaternions. Let $\mathbb{O}_k$ be the ring of integers of $k$ and $\mathcal{O}$ a maximal order in $B$, that is, a maximal subring of $B$ which is a complete $\mathbb{O}_k$-lattice in $B$. Finally, let $\mathcal{O}^1$ be the subgroup of all elements in $\mathcal{O}$ of reduced norm one.

The isomorphism (3-1) induces an embedding of $\mathcal{O}^1$ into $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ by taking the element $\gamma \in \mathcal{O}^1$ to the pair $(\gamma^{\rho_1}, \gamma^{\rho_2}) \in \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, where $\gamma^{\rho_i}$ is the image of $\gamma$ in $B^{\rho_i}$. The group $\mathcal{O}^1$ then acts on $\mathbb{H} \times \mathbb{H}$ as a group of fractional linear transformations. Namely, if $(z, w) \in \mathbb{H} \times \mathbb{H}$ is a point and an element $\gamma \in \mathcal{O}^1$ is identified with two matrices $\gamma^{\rho_1}$ and $\gamma^{\rho_2} \in \mathrm{SL}_2(\mathbb{R})$, then

$$\gamma(z, w) = (\gamma^{\rho_1} z, \gamma^{\rho_2} w).$$

After dividing out by the ineffective kernel, one considers the group

$$\Gamma^1_{\mathcal{O}} = \mathcal{O}^1 / \{\pm 1\} \subset \mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$$

and it can be proven that $\Gamma^1_{\mathcal{O}}$ is an irreducible lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$. Moreover, this lattice is cocompact if and only if $B$ is a division quaternion algebra (see [Vignéras 1980, p. 104]). In general we say that a subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ is an *arithmetic lattice* if there exists $k, B, \rho_1, \rho_2, \mathcal{O}$ as above such that $\Gamma$ is commensurable with $\Gamma^1_{\mathcal{O}}$. Since $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ is a semisimple real Lie group of real rank 2, the famous arithmeticity theorem of Margulis [1991, Theorem (A), p. 298] (or see [Zimmer 1984, Theorem 6.1.2, p. 114]) states that any irreducible lattice $\Gamma$ in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ is an arithmetic lattice.

Let $\Gamma$ be irreducible and cocompact (arithmetic) lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$ and $X_\Gamma := \Gamma \backslash \mathbb{H} \times \mathbb{H}$ be the orbit space of the discontinuous action of $\Gamma$ on $\mathbb{H} \times \mathbb{H}$. Then, there is a natural structure of compact algebraic surface on $X_\Gamma$. Such a surface $X_\Gamma$ is called (*compact*) *Shimura surface* and can be seen as the compact analog of a Hilbert modular surface. We know that $X_\Gamma$ is smooth if $\Gamma$ is torsion free. The numerical invariants of a smooth $X_\Gamma$ are computed in [Matsushima and Shimura 1963]; see also [Shavel 1978]. It follows that a smooth $X_\Gamma$ is a fake quadric if and only if $c_2(X_\Gamma) = 4$ (see [Shavel 1978]).

Let us now study automorphisms of smooth Shimura surfaces $X_\Gamma = \Gamma \backslash \mathbb{H} \times \mathbb{H}$ where $\Gamma$ is a cocompact and irreducible torsion-free lattice in $\operatorname{Aut} \mathbb{H} \times \operatorname{Aut} \mathbb{H}$. Let $\mu \colon \mathbb{H} \times \mathbb{H} \to \mathbb{H} \times \mathbb{H}$ be the involution exchanging the two factors. The group $\operatorname{Aut}(\mathbb{H} \times \mathbb{H})$ is the semidirect product of $\operatorname{Aut} \mathbb{H} \times \operatorname{Aut} \mathbb{H}$ by the group generated by $\mu$. Let $N\Gamma$ (resp. $N_0\Gamma$) be the normalizer of $\Gamma$ in $\operatorname{Aut}(\mathbb{H} \times \mathbb{H})$ (resp. in $\operatorname{Aut} \mathbb{H} \times \operatorname{Aut} \mathbb{H}$);

$N_0\Gamma$ is a subgroup of $N\Gamma$ of index 1 or 2. The following result is crucial for our computations.

**Theorem 3.12.** *The automorphism group of the smooth Shimura surface $X_\Gamma$ is (isomorphic to) $N\Gamma/\Gamma$. An automorphism has only finitely many fixed points or it is an involution whose fixed point set is a union of smooth Shimura curves.*

*There exists an involution with a purely one-dimensional fixed point set if and only if $N\Gamma \neq N_0\Gamma$.*

*An automorphism of a fake quadric has only finitely many fixed points.*

*Proof.* Since $\mathbb{H}\times\mathbb{H}$ is the universal covering of $X_\Gamma$, every automorphism $\sigma$ of $X_\Gamma$ lifts to an automorphism $\gamma$ of $\mathbb{H}\times\mathbb{H}$; this $\gamma$ normalizes $\Gamma$ and two elements $\gamma,\ \gamma'$ both represent $\sigma$ if and only if $\gamma^{-1}\gamma'\in\Gamma$. Hence, $\sigma$ defines an element $\gamma\Gamma$ of the group $N\Gamma/\Gamma$. Conversely, for a class $\gamma\Gamma\in N\Gamma/\Gamma$, the map $\sigma: X\to X$ defined by $\sigma(\Gamma x)=\gamma\Gamma x=\gamma\Gamma\gamma^{-1}\gamma x=\Gamma\gamma x$ is an automorphism of $X$. We thus proved that $\mathrm{Aut}(X)=N\Gamma/\Gamma$.

We say that $\sigma\in\mathrm{Aut}(X)=N\Gamma/\Gamma$, with $\sigma=\gamma\Gamma\in N\Gamma/\Gamma$, is a factor preserving automorphism if $\gamma$ is in $N_0\Gamma$.

Let us denote by $\mathbb{F}_\Gamma$ a fundamental domain in $\mathbb{H}\times\mathbb{H}$ of $\Gamma$. Let $\sigma\in\mathrm{Aut}(X)$ be a nontrivial factor preserving automorphism and let $s$ be a fixed point, with representative $(z_1,z_2)$ in $\mathbb{F}_\Gamma$. Let $\gamma\in N_0\Gamma$ be a representative of $\sigma$ such that $\gamma(z_1,z_2)=(\gamma^{\rho_1}z_1,\gamma^{\rho_2}z_2)=(z_1,z_2)$. The point $s$ is an isolated fixed point of $\sigma$ if and only if $\gamma$ has finitely many fixed points in $\mathbb{F}_\Gamma$.

Since $(\gamma^{\rho_1}z_1,\gamma^{\rho_2}z_2)=(z_1,z_2)$, $z_1$ is a fixed point of $\gamma^{\rho_1}$ and $z_2$ is a fixed point of $\gamma^{\rho_2}$. The only automorphisms of $\mathbb{H}$ with fixed points in $\mathbb{H}$ are elliptic transformations or the identity. An elliptic transformation has a unique fixed point in $\mathbb{H}$.

By Shimizu's theorem [1963, Theorem 2], $\gamma^{\rho_1}$ is trivial if and only if $\gamma^{\rho_2}$ is trivial. Since we supposed that $\sigma$ is nontrivial, at least one — and thus both — of the $\gamma^{\rho_i}$ are elliptic elements of $\mathrm{PSL}_2(\mathbb{R})$. Thus the point $(z_1,z_2)$ is the unique fixed point of $\gamma$ in $\mathbb{H}\times\mathbb{H}$, therefore the point $s$ is an isolated fixed point of $\sigma$.

Suppose now that $\sigma\in\mathrm{Aut}(X)$ is not a factor preserving automorphism. Let $\gamma'\in N\Gamma$ a representative of $\sigma\in N\Gamma/\Gamma$. There exists $\gamma=(\gamma_1,\gamma_2)\in\mathrm{Aut}\,\mathbb{H}\times\mathrm{Aut}\,\mathbb{H}$ such that $\gamma'=\gamma\mu\in\mathrm{Aut}(\mathbb{H}\times\mathbb{H})$. Suppose that $\sigma$ has an infinite number of fixed points. Then by the above discussion, the factor preserving automorphism $\sigma^2$ (with representative $(\gamma\mu)^2=(\gamma_1\gamma_2,\gamma_2\gamma_1)$) must be the identity and $(\gamma_1\gamma_2,\gamma_2\gamma_1)$ must be in $\Gamma$. Let $s=\Gamma(z_1,z_2)$ be a fixed point of $\sigma$. There exists $\lambda\in\Gamma$ such that

$$(\gamma_1z_2,\gamma_2z_1)=\lambda(z_1,z_2).$$

After the change of the representative $\gamma'$ by $\lambda^{-1}\gamma'$, we can assume that $\lambda=1$, thus $z_2=\gamma_2z_1$, $\gamma_1\gamma_2z_1=z_1$ and $\gamma_1\gamma_2z_2=z_2$. Since $(\gamma_1\gamma_2,\gamma_2\gamma_1)$ is in the group $\Gamma$ in which a nontrivial element has no fixed points, we obtain that $\gamma_2\gamma_1 = \gamma_1\gamma_2 = 1$. Since $\gamma_1\gamma_2 = 1$, the point $(t, \gamma_2 t)$ (for $t \in \mathbb{H}$) satisfies

$$\gamma'(t, \gamma_2 t) = (t, \gamma_2 t),$$

therefore there are no isolated fixed points for $\sigma$ and its fixed point set is purely one-dimensional. The image of the disk $\Delta = \{(t, \gamma_2 t) \mid t \in \mathbb{H}\}$ in $X$ is a smooth Shimura curve (see, for instance, [Granath 2002, Chapter 7]) fixed by $\sigma$.

Assume now that $X_\Gamma$ is a quaternionic fake quadric and that the fixed locus $C$ of $\sigma$ is a smooth curve. The topological Lefschetz formula (see Corollary 2.5) implies that the genus of the irreducible components of $C$ is negative, thus the automorphism has only a finite number of fixed points. □

**Remark 3.13.** Note that according to Theorem 3.12 and the proof of Corollary 2.5, the quotient of a quaternionic fake quadric by a group $G$ is a $\mathbb{Q}$-homology quadric if and only if each automorphism $\sigma \in G$ has 4 fixed points, otherwise this quotient is a $\mathbb{Q}$-homology projective plane. All the cyclic groups $G$ which we will study in Section 5 give examples of $\mathbb{Q}$-homology quadrics, with the only possible exception for automorphisms of order 4.

## 4. Quaternionic fake quadrics with nontrivial automorphism groups

As already mentioned, a series of examples of quaternionic fake quadrics has been constructed in [Shavel 1978]. There, the author concentrates on arithmetic lattices $\Gamma \supseteq \Gamma^1_{\mathcal{O}}$ which are defined by quaternion algebras over real quadratic fields of class number one. More recently, in [Džambić 2013], more examples of quaternionic quadrics associated with quaternion algebras over quadratic fields have been found. In this section we will give examples of some known quaternionic fake quadrics together with their automorphism groups. We refer the reader to [Vignéras 1980] and [Deuring 1968] for generalities on arithmetic theory of quaternion algebras.

Let us first make a few general observations, before we discuss the examples in detail. For technical reasons it is more practical to consider the group $\mathrm{PGL}_2^+(\mathbb{R}) \times \mathrm{PGL}_2^+(\mathbb{R})$, where $\mathrm{PGL}_2^+(\mathbb{R}) = \mathrm{GL}_2^+(\mathbb{R})/\mathbb{R}^*$ and $\mathrm{GL}_2^+(\mathbb{R})$ is the group of all $2 \times 2$ matrices with positive determinant, instead of $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$. We identify $\mathrm{PGL}_2^+(\mathbb{R}) \times \mathrm{PGL}_2^+(\mathbb{R})$ with the group $\mathrm{Aut}\,\mathbb{H} \times \mathrm{Aut}\,\mathbb{H}$ of factor preserving holomorphic automorphisms.

From the point of view of Theorem 3.12 we will be interested only in automorphism subgroups $G \le N\Gamma/\Gamma = \mathrm{Aut}(X_\Gamma)$ of factor preserving automorphisms, that is, with $N\Gamma = N_0\Gamma < \mathrm{Aut}\,\mathbb{H} \times \mathrm{Aut}\,\mathbb{H}$, which we will do in the following. In all the considered examples the normalizers $N\Gamma$ will be normalizers of maximal orders and all such lattices can be described arithmetically as follows (see [Borel 1981]).

If $X_\Gamma$ is a quaternionic fake quadric, there is an associated tuple $(k, \rho_1, \rho_2, B, \mathcal{O})$ as described in Section 3. The quaternion algebra $B$ is for fixed $\rho_1, \rho_2$ uniquely determined (up to isomorphism) by the reduced discriminant $d_B = v_1 \cdots v_r$, the formal product over finite places $v_i$ of $k$ where $B$ is ramified, that is, $B \otimes_k k_{v_i} \ncong M_2(k_{v_i})$. This is a special case of H. Hasse's deep classification theorem for simple algebras over number fields (see [Deuring 1968, VII.5, Satz 9, p. 119] or [Vignéras 1980, Chapitre III, Théorème 3.1, p. 74] ). Hence, $(k, \rho_1, \rho_2, B, \mathcal{O}) = (k, \rho_1, \rho_2, d_B, \mathcal{O})$. In the following we will often abbreviate such a datum which determines the quaternion algebra $B$ with $B(k, d_B)$ or $B(k, v_1 \cdots v_r)$. Let us fix a datum $B(k, v_1 \cdots v_r)$ and let $B^+$ be the group of all $x \in B^*$ such that the reduced norm $\mathrm{Nrd}(x)$ is totally positive. It is known that

$$N\Gamma_{\mathcal{O}}^+ = \{x \in B^+ \mid x\mathcal{O}x^{-1} = \mathcal{O}\}/k^* \tag{4-1}$$

is a maximal lattice. $N\Gamma_{\mathcal{O}}^+$ contains $\Gamma_{\mathcal{O}}^1$ and $\Gamma_{\mathcal{O}}^1$ is normal in $N\Gamma_{\mathcal{O}}^+$ with $N\Gamma_{\mathcal{O}}^+/\Gamma_{\mathcal{O}}^1 \cong (\mathbb{Z}/2\mathbb{Z})^l$ an elementary abelian 2-group with $l \geq r$ and $r$ is the number of ramified places in $B$ (see [Shavel 1978], for instance). If the class number of $k$ is one (as will be the case in all the considered examples) there is an alternative description of $N\Gamma_{\mathcal{O}}^+$ as

$$\begin{aligned} N\Gamma_{\mathcal{O}}^+ = \Big\{\alpha = \varrho_1^{\epsilon_1} \cdots \varrho_r^{\epsilon_r} \lambda\tau \in B^* \;\Big|\; &\mathrm{Nrd}(\alpha) \text{ totally positive}, \\ &\tau \in k^*,\ \lambda \in \mathcal{O}^*,\ \epsilon_i \in \{0, 1\},\ \mathrm{Nrd}(\varrho_i) \text{ divides } d_B\Big\}/k^* \end{aligned} \tag{4-2}$$

(see [Shavel 1978, p. 223]). It follows that a quaternionic fake quadric $X_\Gamma$ with $\Gamma \supseteq \Gamma_{\mathcal{O}}^1$ will have an elementary abelian 2-group as the automorphism group $\mathrm{Aut}(X_\Gamma)$. All Shavel's examples will provide such automorphism groups.

***A fake quadric with automorphism group $\mathbb{Z}/2\mathbb{Z}$.*** There are examples of quaternionic fake quadrics $X_\Gamma$ whose automorphism group is $\mathbb{Z}/2\mathbb{Z}$ and, as mentioned above, they already appear in [Shavel 1978].

For example, let $k = \mathbb{Q}(\sqrt{2})$ and let $B = B(k, \mathfrak{p}_3\mathfrak{p}_7)$ be the (unique) quaternion algebra over $k$ which is ramified exactly at the two finite primes $\mathfrak{p}_3$ and $\mathfrak{p}_7$ of $k$ lying over the rational primes 3 and 7 respectively. Since $k$ has the class number one, there is the unique (up to conjugation) maximal order $\mathcal{O}$ in $B$. Consider the group $\Gamma_{\mathcal{O}}^1$. By [Shavel 1978, Proposition 4.7], $X_{\Gamma_{\mathcal{O}}^1}$ is smooth. By the already mentioned general result of Matsushima and Shimura [1963], $q(X_{\Gamma_{\mathcal{O}}^1}) = 0$. The Chern number $c_2(X_{\Gamma_{\mathcal{O}}^1})$ is computed via the volume formula of Shimizu (see [Shavel 1978, Theorem 3.1]). Since the prime 3 is inert and 7 is decomposed in $k$, this formula gives $c_2(X_{\Gamma_{\mathcal{O}}^1}) = 8$. The normalizer of $\Gamma_{\mathcal{O}}^1$ is $N\Gamma_{\mathcal{O}}^+$ and by [Shavel 1978, Proposition 1.3 and Proposition 1.4], we have

$$\mathrm{Aut}(X_{\Gamma_{\mathcal{O}}^1}) \cong L_1/L_2 = \langle[\mathfrak{p}_3], [\mathfrak{p}_7]\rangle \cong (\mathbb{Z}/2\mathbb{Z})^2,$$

where $L_1$ is the group of principal fractional ideals of type $(\mathfrak{p}_3)(\mathfrak{p}_7)I^2$ ($I$ a principal fractional ideal) for which one can find a totally positive generator and $L_2$ consists of all principal ideals of type $(a^2)$ with $a \in k$ (see also [Shimura 1967, Section 3.12]). Let $\Gamma_{\mathfrak{p}_3}$ be the kernel of the canonical homomorphism

$$N\Gamma_{\mathcal{O}}^{+} \longrightarrow L_1/L_2 \longrightarrow \langle [\mathfrak{p}_7] \rangle.$$

By Shavel's criterion [1978, Theorem 4.11], $\Gamma_{\mathfrak{p}_3}$ is torsion free and as $[\Gamma_{\mathfrak{p}_3} : \Gamma_{\mathcal{O}}^1] = 2$, $X_{\Gamma_{\mathfrak{p}_3}}$ is a fake quadric with $\mathrm{Aut}(X_{\Gamma_{\mathfrak{p}_3}}) \cong \mathbb{Z}/2\mathbb{Z}$.

***A fake quadric with automorphism group* $(\mathbb{Z}/2\mathbb{Z})^2$.** Consider again $k = \mathbb{Q}(\sqrt{2})$ and now the quaternion algebra $B = B(k, \mathfrak{p}_2\mathfrak{p}_5)$ over $k$ which is ramified exactly at the two finite places $\mathfrak{p}_2$ and $\mathfrak{p}_5$. Again there is the unique maximal order $\mathcal{O}$ in $B$ and as in the previous example, Shavel's results show that $X_{\Gamma_{\mathcal{O}}^1}$ is smooth. The prime 2 is ramified and 5 is inert in $k$ and therefore Shimizu's volume formula gives $c_2(X_{\Gamma_{\mathcal{O}}^1}) = 4$. Hence $X_{\Gamma_{\mathcal{O}}^1}$ is a fake quadric. With the same arguments as in the previous example $\mathrm{Aut}(X_{\Gamma_{\mathcal{O}}^1})$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$.

***A fake quadric with automorphism group of order* 20.** Consider $k = \mathbb{Q}(\sqrt{5})$ and the quaternion algebra $B = B(k, \mathfrak{p}_2\mathfrak{p}_5)$ over $k$ which is ramified exactly at the primes $\mathfrak{p}_2$ and $\mathfrak{p}_5$. In this case the group $\Gamma_{\mathcal{O}}^1$ (where $\mathcal{O}$ is again a maximal order in $B$), contains torsion elements of order 5 and no other torsions (see [Shavel 1978, Proposition 4.7 and Theorem 4.8]).[1] Volume formula of Shimizu gives in this case $c_2(X_{\Gamma_{\mathcal{O}}^1}) = 4/5$. Let us now give a torsion-free subgroup $\Gamma < \Gamma_{\mathcal{O}}^1$ of index 5. The corresponding surface $X_\Gamma$ will be a fake quadric. Since $\mathfrak{p}_2$ is ramified in $B$, there is a prime ideal $\mathfrak{P}_2$ in $\mathcal{O}$ lying over $\mathfrak{p}_2$ and satisfying $\mathfrak{P}_2^2 = \mathfrak{p}_2\mathcal{O}$. Let

$$\mathcal{O}^1(\mathfrak{P}_2) = \{x \in \mathcal{O}^1 \mid x \equiv 1 \bmod \mathfrak{P}_2\}$$

and $\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$ the image of $\mathcal{O}^1(\mathfrak{P}_2)$ in $\Gamma_{\mathcal{O}}^1$. The group $\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$ is a normal subgroup in $\Gamma_{\mathcal{O}}^1$ and the index can be computed via the localization of $B$ at $\mathfrak{p}_2$. Namely, observe first that $\Gamma_{\mathcal{O}}^1/\,\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$ is isomorphic to the factor group $\mathcal{O}^1/\mathcal{O}^1(\mathfrak{P}_2)$. This is because $-1$ is in $\mathcal{O}^1(\mathfrak{P}_2)$. Let $\mathcal{O}_{\mathfrak{p}_2}$ be the maximal order in $B_{\mathfrak{p}_2}$, that is, $\mathcal{O}_{\mathfrak{p}_2} = \mathcal{O} \otimes_{\mathbb{O}_k} \mathbb{O}_{k_{\mathfrak{p}_2}}$, where $\mathbb{O}_{k_{\mathfrak{p}_2}}$ is the ring of integers in $k_{\mathfrak{p}_2}$. Its maximal ideal $\widehat{\mathfrak{P}}_2$ is the topological closure of $\mathfrak{P}_2$. By the strong approximation property, $\mathcal{O}^1/\mathcal{O}^1(\mathfrak{P}_2) \cong \mathcal{O}_{\mathfrak{p}_2}^1/\mathcal{O}_{\mathfrak{p}_2}^1(\widehat{\mathfrak{P}}_2)$. Note that $B_{\mathfrak{p}_2}^1 = \mathcal{O}_{\mathfrak{p}_2}^1$, since $\mathcal{O}_{\mathfrak{p}_2}$ is the subring of $B_{\mathfrak{p}_2}$ consisting of elements whose reduced norm is less or equal 1. We use a theorem of C. Riehm [1970, Theorem 7] by which

$$\mathcal{O}_{\mathfrak{p}_2}^1/\mathcal{O}_{\mathfrak{p}_2}^1(\widehat{\mathfrak{P}}_2) \cong \ker\big((\mathcal{O}_{\mathfrak{p}_2}/\widehat{\mathfrak{P}}_2)^* \xrightarrow{\mathrm{Nr}} (\mathbb{O}_{k_{\mathfrak{p}_2}}/\mathfrak{p}_2)^*\big) \cong \ker\big(\mathbb{F}_{16}^* \xrightarrow{\mathrm{Nr}} \mathbb{F}_4^*\big) \cong \mathbb{Z}/5\mathbb{Z}$$

[1] In Theorem 4.8 of [Shavel 1978], the symbol $\left(\frac{\cdot}{p}\right)$ for $p = 2$ should be read as the Kronecker symbol; that is, $\left(\frac{d}{2}\right) = 1 \Leftrightarrow d \equiv \pm 1 \bmod 8$ and $= -1 \Leftrightarrow d \equiv \pm 3 \bmod 8$.

(Here Nr is the surjective homomorphism of multiplicative groups arising from the norm map for the field extension $\mathbb{F}_{16}/\mathbb{F}_4$.) Since $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$ is embedded in $\mathcal{O}^1_{\mathfrak{p}_2}(\widehat{\mathfrak{P}}_2)/\pm 1$ and the latter group is a pro-2-group (again by [Riehm 1970]) it cannot contain elements of order 5. Therefore, $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$ is a torsion-free group and $X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)}$ is a fake quadric. Since $\Gamma^1_{\mathcal{O}}$ contains a 5-torsion and $\Gamma^1_{\mathcal{O}}$ normalizes $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$, $X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)}$ contains an automorphism of order 5. In order to determine the full automorphism group $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)})$ we first need to find the normalizer of $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. By definition, elements of $N\Gamma^+_{\mathcal{O}}$ normalize $\mathcal{O}$, that is, $x\mathcal{O}x^{-1} = \mathcal{O}$. Let $\gamma \in \Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. Since the class number of $k$ is one, every two-sided $\mathcal{O}$-ideal is principal and we can choose $\Pi_2 \in \mathcal{O}$ such that $\Pi_2\mathcal{O} = \mathfrak{P}_2$. Moreover, as $\mathfrak{P}_2$ is uniquely determined by the property that the $\mathbb{O}_k$-ideal $\mathrm{Nrd}(\mathfrak{P}_2)$ is $\mathfrak{p}_2$, we can choose $\Pi_2$ such that $\mathrm{Nrd}(\Pi_2) = 2$. Then $\gamma = \pm(1 + m\Pi_2)$ with $m \in \mathcal{O}$. For $x \in N\Gamma^+_{\mathcal{O}}$ we have $x\gamma x^{-1} = 1 + xm\Pi_2 x^{-1} = 1 + m'x\Pi_2 x^{-1}$ with some $m' \in \mathcal{O}$. The element $x\Pi_2 x^{-1}$ lies in $\mathcal{O}$ and $\mathrm{Nrd}(x\Pi_2 x^{-1}) = \mathrm{Nrd}(\Pi_2) = 2$. Since $\mathfrak{P}_2 = \langle \Pi_2 \rangle$ is the unique prime ideal over 2, $x\Pi_2 x^{-1} \in \mathfrak{P}_2$ and $x\gamma x^{-1} \in \Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. It follows that the normalizer of $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$ is $N\Gamma^+_{\mathcal{O}}$. This leads to an exact sequence

$$1 \longrightarrow \Gamma^1_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2) \longrightarrow N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2) \longrightarrow N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}} \longrightarrow 1 \tag{4-3}$$

which we can write abstractly as

$$1 \longrightarrow \mathbb{Z}/5\mathbb{Z} \longrightarrow \mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\Pi_2)}) \longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \longrightarrow 1.$$

Let $\lambda \in \mathcal{O}^1$ satisfy $\lambda^5 = -1$, that is, $\lambda$ gives rise to a 5-torsion in $\Gamma^1_{\mathcal{O}}$. Then $\lambda$ satisfies the equation $\lambda^2 - \frac{1+\sqrt{5}}{2}\lambda + 1 = 0$ over $k$. We can assume that $\lambda$ generates $\Gamma^1_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. Let $g = \lambda + 1$. The reduced norm of $g$ is $\mathrm{Nrd}(g) = (\lambda+1)(\bar{\lambda}+1) = \mathrm{Nrd}(\lambda) + \mathrm{Trd}(\lambda) + 1 = 2 + \frac{1+\sqrt{5}}{2} = \frac{5+\sqrt{5}}{2}$, where Trd is the reduced trace and $x \mapsto \bar{x}$ is the standard involution of first kind on $B$. Since $\frac{5+\sqrt{5}}{2}$ is a totally positive generator of the prime ideal over 5, $g$ defines an element of $N\Gamma^+_{\mathcal{O}}$ (see (4-2)). On the other hand $g^2 = (\lambda+1)^2 = \lambda^2 + 2\lambda + 1 = \left(\frac{1+\sqrt{5}}{2}\lambda - 1\right) + 2\lambda + 1 = \left(\frac{5+\sqrt{5}}{2}\right)\lambda$. This shows that $g$ has order 10 in $N\Gamma^+_{\mathcal{O}}$ and hence gives an element of order 10 in $N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. Moreover the image of $g$ in $N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}$ is not trivial. Using the computer algebra system PARI, we can check that both ramified primes $\mathfrak{p}_2$ and $\mathfrak{p}_5$ are not split in $k(\sqrt{-2})$. This implies that $k(\sqrt{-2}) \subset B$ (see [Shavel 1978, Proposition 4.5]) and we can take $\sqrt{-2}$ as the generator $\Pi_2$ of $\mathfrak{P}_2$. Hence, $\Pi_2$, considered as an element of $N\Gamma^+_{\mathcal{O}}$, is of order 2 and the images of $g$ and $\Pi_2$ in $N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}$ generate this group.

**Lemma 4.14.** *Let $g$ and $\Pi_2$ be elements constructed above. Then in $N\Gamma^+_{\mathcal{O}}$ we have the relation $\Pi_2 g \Pi_2 = g^{-1}$ modulo $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$.*

*Proof.* The element $\Pi_2$ generates $\mathfrak{P}_2$. Consider $g$ and $\Pi_2$ as the elements of the localization $B_{\mathfrak{p}_2}$ of $B$ at $\mathfrak{p}_2$. This is a division quaternion algebra over $k_{\mathfrak{p}_2}$ and has a representation

$$B_{\mathfrak{p}_2} = L_{\mathfrak{p}_2} \oplus \Pi_2 L_{\mathfrak{p}_2},$$

where $L_{\mathfrak{p}_2}$ is the unique unramified quadratic extension of $k_{\mathfrak{p}_2}$ (see [Vignéras 1980, p. 34]). For every $t \in L_{\mathfrak{p}_2}$ we have $t\Pi_2 = \Pi_2\bar{t}$, where $\bar{t}$ is the Galois-conjugate of $t$ in $L_{\mathfrak{p}_2}$. The element $g$ lies in $k_{\mathfrak{p}_2}(\lambda) = k_{\mathfrak{p}_2}(\xi_5)$ which is an unramified quadratic extension of $k_{\mathfrak{p}_2}$, so $L_{\mathfrak{p}_2} = k_{\mathfrak{p}_2}(\lambda)$. Therefore $g \in L_{\mathfrak{p}_2}$ and $g\Pi_2 = \Pi_2\bar{g}$. Because $g\bar{g}$ is in $k^*$ we have that $\bar{g} = g^{-1}$ considered as an element of $N\Gamma^+_{\mathcal{O}} \subset B^*_{\mathfrak{p}_2}/k^*_{\mathfrak{p}_2}$. This gives a relation

$$\Pi_2 g \Pi_2 = g^{-1}$$

in $N\Gamma^+_{\mathcal{O}}$, since $\Pi_2^2 = 1$ in $N\Gamma^+_{\mathcal{O}}$. Also, $g$ and $\Pi_2 g \Pi_2 = g^{-1}$ are not equal modulo $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$ because this would imply that $g^2 \in \Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$. But as $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)$ is torsion-free and $g^2$ is of finite order, this is impossible. $\square$

**Proposition 4.15.** *With above notations we have*

$$\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)}) \cong \mathbb{D}_{10}.$$

*Proof.* By the above discussion, $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)})$ is of order 20 and is generated by elements $g$ of order 10 and $\Pi_2$ of order 2 satisfying $\Pi_2 g \Pi_2 = g^{-1}$. The only group of order 20 with these relations is $\mathbb{D}_{10}$. $\square$

***A fake quadric with automorphism group of order* 8.** We consider $k = \mathbb{Q}(\sqrt{5})$ and $B = B(k, \mathfrak{p}_2\mathfrak{p}_{11})$, the unique quaternion algebra ramified exactly at the primes $\mathfrak{p}_2$ and $\mathfrak{p}_{11}$. Since 2 is inert and 11 is decomposed in $k$, Shimizu's volume formula gives $c_2(X_{\Gamma^1_{\mathcal{O}}}) = \frac{4}{5\cdot 12}(4-1)(11-1) = 2$ as the value of the second Chern number of the quotient $X_{\Gamma^1_{\mathcal{O}}}$, where again $\Gamma^1_{\mathcal{O}}$ is the norm-1 group of a maximal order in $B$. As before, results of [Shavel 1978] show that $\Gamma^1_{\mathcal{O}}$ contains only torsion elements of order 2 and no other torsions (Here, observe that 2 is split in $\mathbb{Q}(\sqrt{-15})$, hence, by [Shavel 1978, Theorem 4.8] there are no elements of order 3 in $\Gamma^1_{\mathcal{O}}$, and note that there are no elements of order 5 because $11 \equiv 1 \bmod 5$ which implies that $\mathfrak{p}_{11}$ is split in $k(\xi_5)$ ). Since $\mathfrak{p}_{11}$ is ramified in $B$, there is the unique prime ideal $\mathfrak{P}_{11}$ in $\mathcal{O}$ such that $\mathfrak{P}_{11}^2 = \mathfrak{p}_{11}\mathcal{O}$. Consider the principal congruence subgroup

$$\mathcal{O}^1(\mathfrak{P}_{11}) = \{x \in \mathcal{O}^1 \mid x \equiv 1 \bmod \mathfrak{P}_{11}\}$$

and $\Gamma_{\mathcal{O}}(\mathfrak{P}_{11})$ its image in $\Gamma^1_{\mathcal{O}}$. It is a normal subgroup in $\Gamma^1_{\mathcal{O}}$. The quotient $\Gamma^1_{\mathcal{O}}/\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$ is isomorphic to $\mathcal{O}^1/\pm\mathcal{O}^1(\mathfrak{P}_{11})$ because $-1 \notin \mathcal{O}^1(\mathfrak{P}_{11})$. In order to compute the latter quotient we change over to the localization at the prime $\mathfrak{p}_{11}$. Let

$$B_{\mathfrak{p}_{11}} = B \otimes_k k_{\mathfrak{p}_{11}} = B \otimes_k \mathbb{Q}_{11}.$$

This is the unique division quaternion algebra over $\mathbb{Q}_{11}$. We write $\mathcal{O}_{\mathfrak{p}_{11}} = \mathcal{O} \otimes_{\mathbb{O}_k} \mathbb{Z}_{11}$

for its maximal order. As in the previous example let $\widehat{\mathfrak{P}}_{11}$ denote the prime ideal of $\mathcal{O}_{\mathfrak{p}_{11}}$. We have

$$\mathcal{O}^1/\mathcal{O}^1(\mathfrak{P}_{11}) \cong \mathcal{O}^1_{\mathfrak{p}_{11}}/\mathcal{O}^1_{\mathfrak{p}_{11}}(\widehat{\mathfrak{P}}_{11})$$

by the strong approximation theorem. By Riehm's result [1970, Theorem 7],

$$\mathcal{O}^1_{\mathfrak{p}_{11}}/\mathcal{O}^1_{\mathfrak{p}_{11}}(\widehat{\mathfrak{P}}_{11}) \cong \ker\big((\mathcal{O}_{\mathfrak{p}_{11}}/\widehat{\mathfrak{P}}_{11})^* \xrightarrow{\mathrm{Nr}} (\mathbb{O}_{k_{\mathfrak{p}_{11}}}/\mathfrak{p}_{11})^*\big) \cong \ker\big(\mathbb{F}^*_{121} \longrightarrow \mathbb{F}^*_{11}\big).$$

Since $\mathbb{F}_{121} = \mathbb{F}_{11}(\xi_{12})$, where $\xi_{12}$ denotes a primitive twelfth root of unity we conclude that $\mathcal{O}^1_{\mathfrak{p}_{11}}/\mathcal{O}^1_{\mathfrak{p}_{11}}(\widehat{\mathfrak{P}}_{11})$ is isomorphic to $\mu_{12} = \langle\xi_{12}\rangle$. Hence

$$\Gamma^1_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11}) \cong \mathcal{O}^1_{\mathfrak{p}_{11}}/\pm\mathcal{O}^1_{\mathfrak{p}_{11}}(\widehat{\mathfrak{P}}_{11}) \cong \mu_6 = \langle\xi_6\rangle.$$

Let us now define an intermediate group

$$\Gamma = \{x \in \Gamma^1_{\mathcal{O}} \mid x \mod \mathfrak{P}_{11} \in \langle\xi_6^2\rangle \subset \mu_6\}.$$

$\Gamma < \Gamma^1_{\mathcal{O}}$ is a subgroup of index 2, hence $c_2(X_\Gamma) = 4$. Moreover, $\Gamma$ is torsion-free since it cannot contain elements of order 2. For if an order-two element $x$ is in $\Gamma$, then its image $x \mod \mathfrak{P}_{11}$ in $\Gamma^1_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$ lies in a cyclic group $\langle\xi_6^2\rangle$ of order three, hence it must be the identity. But this means that $x$ is in $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$. On the other hand $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$ is torsion-free because it embeds in a pro-11 group $\mathcal{O}^1_{\mathfrak{p}_{11}}(\widehat{\mathfrak{P}}_{11})/\pm 1$. This contradicts the assumption on $x$. All this shows that $X_\Gamma$ is a fake quadric.

**Proposition 4.16.** *Let $N\Gamma^+_{\mathcal{O}}$ be defined as in (4-1). Then $N\Gamma^+_{\mathcal{O}}$ is the normalizer of $\Gamma$ and $N\Gamma^+_{\mathcal{O}}/\,\Gamma$ is isomorphic to $\mathbb{D}_4$.*

*Proof.* As a subgroup of index 2 in $\Gamma^1_{\mathcal{O}}$ the group $\Gamma$ is normal in $\Gamma^1_{\mathcal{O}}$. On the other hand, for the same reason as in the previous example, $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$ as well as $\Gamma^1_{\mathcal{O}}$ is normal subgroup in $N\Gamma^+_{\mathcal{O}}$. This already implies that $\Gamma$ is normal in $N\Gamma^+_{\mathcal{O}}$ because any conjugate of $\Gamma$ will be a subgroup between $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11})$ and $\Gamma^1_{\mathcal{O}}$ of index 2 in $\Gamma^1_{\mathcal{O}}$. There is only one such group, namely $\Gamma$, since $\Gamma^1_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}(\mathfrak{P}_{11}) \cong \mathbb{Z}/6\mathbb{Z}$. Similar exact sequence as (4-3) now shows that $\mathrm{Aut}(X_\Gamma)$ is an extension of $\mathbb{Z}/2\mathbb{Z}$ by the Klein's four group. Since the 2-torsions in $\Gamma^1_{\mathcal{O}}$ come from embeddings of fourth roots of unity into $\mathcal{O}$ there is $\lambda \in \mathcal{O}^1$ such that $\lambda^2 = -1$. Let $g = \lambda + 1$. Then, as $\mathrm{Trd}(\lambda) = 0$, we have $\mathrm{Nrd}(g) = (\lambda + 1)(\bar{\lambda} + 1) = 2$ and also $g^2 = (\lambda + 1)^2 = 2\lambda$ which implies that $g$ defines an element of order 4 in $N\Gamma^+_{\mathcal{O}}$ and hence an element of order 4 in $N\Gamma^+_{\mathcal{O}}/\,\Gamma$. Moreover, the image of $g$ in $N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}$ is not trivial. Since both prime divisors 2 and $\pi_{11}$ of the reduced discriminant do not split in $k(\sqrt{-\pi_{11}})$ (as can be checked using PARI, for instance), the element $\Pi_{11} = \sqrt{-\pi_{11}}$ is in $B$ and moreover $\Pi_{11}$ defines an element of $N\Gamma^+_{\mathcal{O}}$ of order 2 such that the images of $\Pi_{11}$ and $g$ in $N\Gamma^+_{\mathcal{O}}/\,\Gamma^1_{\mathcal{O}}$ generate this group. Same argument as in Lemma 4.14 gives a relation between $\Pi_{11}$ and $g$: consider $\Pi_{11}$ as the generator of the prime ideal $\mathfrak{P}_{11}$. Locally, $B_{\mathfrak{p}_{11}}$ can be written as $B_{\mathfrak{p}_{11}} = L_{\mathfrak{p}_{11}} \oplus \Pi_{11}L_{\mathfrak{p}_{11}}$, where $L_{\mathfrak{p}_{11}} = k_{\mathfrak{p}_{11}}(\xi_{12})$ is the unique unramified quadratic extension of $k_{\mathfrak{p}_{11}} \cong \mathbb{Q}_{11}$ with the multiplication rule $t\Pi_{11} = \Pi_{11}\bar{t}$ for all $t \in L_{\mathfrak{p}_{11}}$. The element $g$ is in $L_{\mathfrak{p}_{11}}$, namely $g = 1 + \xi_{12}^3$. Then $g\Pi_{11} = \Pi_{11}\bar{g} = \Pi_{11}(1+\overline{\xi_{12}}^3) = \Pi_{11}(1+\xi_{12}^9)$. In $N\Gamma_{\mathcal{O}}^+$ the relations $\bar{g} = g^{-1}$ and $\Pi_{11}^2 = 1$ hold, hence $\Pi_{11}g\Pi_{11} = g^{-1}$ in $N\Gamma_{\mathcal{O}}^+$. Also $g \neq g^{-1}$ modulo $\Gamma$, since otherwise $g^2$ would be in $\Gamma$ which is not possible because $g^2$ is torsion and $\Gamma$ torsion-free. $N\Gamma_{\mathcal{O}}^+/\,\Gamma$ is isomorphic to $\mathbb{D}_4$ which is the only group of order 8 generated by two elements $\Pi_{11}$ of order 2 and $g$ of order 4 with $h \neq g^2$ and $\Pi_{11}g\Pi_{11} = g^{-1}$. □

**Remark 4.17.** Considering $k = \mathbb{Q}(\sqrt{13})$, the quaternion algebra $B = B(k, \mathfrak{p}_2\mathfrak{p}_3)$, and $\Gamma = \Gamma_{\mathcal{O}}^1(\mathfrak{P}_3)$, the arguments as in the examples before will show that $X_{\Gamma_{\mathcal{O}}^1(\mathfrak{P}_3)}$ is a fake quadric whose automorphism group is isomorphic to $\mathbb{D}_4$.

***A fake quadric with automorphism group*** $\mathbb{D}_6$**.** This time we consider the quadratic field $k = \mathbb{Q}(\sqrt{2})$ and the quaternion algebra $B = B(k, \mathfrak{p}_2\mathfrak{p}_3)$. The norm-1 group $\Gamma_{\mathcal{O}}^1$ of a maximal order in $B$ contains torsion elements of order 3, but no elements of order 2, because $\mathfrak{p}_3$ is decomposed in $k(\sqrt{-1})$. The second Chern number of the quotient $X_{\Gamma_{\mathcal{O}}^1}$ is $c_2(X_{\Gamma_{\mathcal{O}}^1}) = (9-1)/6 = 4/3$. Let $\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$ be the principal congruence subgroup corresponding to the prime ideal $\mathfrak{P}_2 \subset \mathcal{O}$, defined by the relation $\mathfrak{P}_2^2 = \mathfrak{p}_2\mathcal{O}$. Again by Riehm's theorem and with arguments as in Section 4, $\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$ is torsion-free normal subgroup in $\Gamma_{\mathcal{O}}^1$ of index 3, hence $X_{\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)}$ is a fake quadric. The automorphism group $\mathrm{Aut}(X_{\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)})$ is isomorphic to the factor group

$$N\Gamma_{\mathcal{O}}^+/\,\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2).$$

which is an extension of $\Gamma_{\mathcal{O}}^1/\,\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2) \cong \mathbb{Z}/3\mathbb{Z}$ by $N\Gamma_{\mathcal{O}}^+/\,\Gamma_{\mathcal{O}}^1 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

**Proposition 4.18.** *We have* $\mathrm{Aut}(X_{\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)}) \cong \mathbb{D}_6$.

*Proof.* Let $\lambda \in \mathcal{O}^1$ be an element with $\lambda^3 = -1$ and $g = \lambda + 1$. Such $\lambda$ exists since $\Gamma_{\mathcal{O}}^1$ contains 3-torsions. We can take $\pm\lambda$ to be the generator of $\Gamma_{\mathcal{O}}^1/\,\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$. Since $\mathrm{Trd}(\lambda) = 1$, we have $\mathrm{Nrd}(g) = 3$ which implies that $g$ defines an element in $N\Gamma_{\mathcal{O}}^+$. Additionally $g^2 = \lambda^2 + 2\lambda + 1 = 3\lambda$ which means that $g$ has order 6 considered as an element of $N\Gamma_{\mathcal{O}}^+$. The totally positive element $\pi_2 = 2 + \sqrt{2} \in k$ generates $\mathfrak{p}_2$ and since neither $\pi_3 = 3$ nor $\pi_2$ are split in $k(\sqrt{-\pi_2})$, $\Pi_2 = \sqrt{-\pi_2}$ lies in $B$ and defines an element in $N\Gamma_{\mathcal{O}}^+$ of order 2 such that the classes of $g$ and $\Pi_2$ in $N\Gamma_{\mathcal{O}}^+/\,\Gamma_{\mathcal{O}}^1$ generate this group. In particular, $\Pi_2$ is a generator of $\mathfrak{P}_2$. Locally $B_{\mathfrak{p}_2} = L_{\mathfrak{p}_2} \oplus \Pi_2 L_{\mathfrak{p}_2}$, where $L_{\mathfrak{p}_2} = \mathbb{Q}_2(\xi_6)$ is the unramified quadratic extension of $k_{\mathfrak{p}_2} \cong \mathbb{Q}_2$. As in previous examples, $g$ lies in $L_{\mathfrak{p}_2}$ and $\Pi_2 g\Pi_2 = \bar{g} = g^{-1}$ in $N\Gamma_{\mathcal{O}}^+$. This gives a relation $\Pi_2 g\Pi_2 = g^{-1}$ in $N\Gamma_{\mathcal{O}}^+/\,\Gamma_{\mathcal{O}}^1(\mathfrak{P}_2)$. As $\Pi_2$ is not a power of $g$, the finite group generated by $g$ and $\Pi_2$ is isomorphic to $\mathbb{D}_6$. □

***Automorphism groups of order* 16 *and* 24.** There are more examples of quaternionic fake quadrics with a nontrivial automorphism group. For instance, all examples in Shavel's paper have $\mathbb{Z}/2\mathbb{Z}$ or $(\mathbb{Z}/2\mathbb{Z})^2$ as the full group of automorphisms. As in previous examples we show

**Proposition 4.19.** *Let* $B(\mathbb{Q}(\sqrt{2}), \mathfrak{p}_2, \mathfrak{p}_7)$ *be the indefinite quaternion algebra over* $k = \mathbb{Q}(\sqrt{2})$ *with reduced discriminant* $d_B = \mathfrak{p}_2\mathfrak{p}_7$ *and* $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)$ *the congruence subgroup in* $\Gamma^1_{\mathcal{O}}$ *corresponding to a maximal order* $\mathcal{O}$ *in* $B$ *with respect to the prime ideal* $\mathfrak{P}_7$ *of* $\mathcal{O}$ *lying over the ramified prime* $\mathfrak{p}_7$. *Then* $X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)}$ *is a fake quadric with the automorphism group* $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)}) \cong \mathbb{D}_8$.

*Proof.* The proof goes along the same lines as in the examples before. By Riehm's Theorem, $\Gamma^1_{\mathcal{O}}/\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7) \cong \mathbb{Z}/4\mathbb{Z}$ and we obtain $c_2(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)}) = 4$ by Shimizu's formula. By Shavel's criterion for the existence of torsions, we find that the maximal order $\mathcal{O}$ contains a primitive eighth root of unity $\lambda$ which leads to an element of order 4 in $\Gamma^1_{\mathcal{O}}$. We can take $\lambda$ as a generator of this quotient. As in the examples before take $g = 1 + \lambda$. Then, as $\lambda$ satisfies $\lambda^2 - \sqrt{2}\lambda + 1 = 0$ over $k$, $\mathrm{Nrd}(g) = \mathrm{Nrd}(\lambda + 1) = 2 + \sqrt{2}$, hence $g$ defines an element in $N\Gamma^+_{\mathcal{O}}$. We have $g^2 = \lambda^2 + 2\lambda + 1 = \sqrt{2}\lambda + 2\lambda = (2 + \sqrt{2})\lambda$. Hence, $g$ is an element of order 8 in $N\Gamma^+_{\mathcal{O}}$ and its image in $N\Gamma^+_{\mathcal{O}}/\Gamma^1_{\mathcal{O}}$ is not trivial. The rational prime 7 is split in $k$, so there are two possible choices of $\mathfrak{p}_7$. Fix a prime $\mathfrak{p}_7 = \langle\pi_7\rangle$ ($\pi_7 = 3 + \sqrt{2}$ say). Both $\pi_7$ as well as $\pi_2$ are ramified in $k(\sqrt{-\pi_7})$, hence $\sqrt{-\pi_7} \in B$ defines an element $\Pi_7 \in B$ which defines an order-2 element in $N\Gamma^+_{\mathcal{O}}$. As in the previous examples we have $\Pi_7 g \Pi_7 = \bar{g}$ because locally in $B_{\mathfrak{p}_7}$, $\Pi_7 = \sqrt{-\pi_7}$ generates the unique prime ideal of the maximal order $\mathcal{O}_{\mathfrak{p}_7}$ and $g$ lies in the unramified quadratic extension $L_{\mathfrak{p}_7} = \mathbb{Q}_7(\xi_8)$. This gives a relation $\Pi_7 g \Pi_7 = g^{-1}$ in $N\Gamma^+_{\mathcal{O}}/\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)$. Also $\Pi_7$ is not a power of $g$ modulo $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_7)$ since the reduced norms of $\Pi_7$ and $g$ are different primes. The only group of order 16 with these relations is $\mathbb{D}_8$. □

Let us finally sketch the construction of a fake quadric with an automorphism group of order 24.

**Proposition 4.20.** *Let* $B(\mathbb{Q}(\sqrt{3}), \mathfrak{p}_2, \mathfrak{p}_3)$ *be the indefinite quaternion algebra over* $k = \mathbb{Q}(\sqrt{3})$ *ramified over the prime ideals* $\mathfrak{p}_2$ *and* $\mathfrak{p}_3$ *and let* $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3) \triangleleft \Gamma^1_{\mathcal{O}}$ *be the principal congruence subgroup with respect to the principal ideal* $\mathfrak{P}_2\mathfrak{P}_3$ *of a maximal order* $\mathcal{O} \subset B$ *lying over* $\mathfrak{p}_2\mathfrak{p}_3$. *Then* $X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)}$ *is a fake quadric with* $|\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)})| = 24$. *The automorphism group* $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)})$ *contains a cyclic subgroup of order* 12.

**Remark 4.21.** The full automorphism group in this case has order 24. To our knowledge, this is the largest known automorphism group of a fake quadric. The precise abstract group structure of $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)})$ is not known to us, since the local method, used in previous examples does not apply directly in this case.

*Proof.* That $X_{\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)}$ has the correct numerical invariants follows again from Riehm's Theorem, Shimizu's formula and the observation that for the index we have $[\Gamma^1_{\mathcal{O}} : \Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)] = [\Gamma^1_{\mathcal{O}} : \Gamma^1_{\mathcal{O}}(\mathfrak{P}_2)][\Gamma^1_{\mathcal{O}} : \Gamma^1_{\mathcal{O}}(\mathfrak{P}_3)]$. By Shavel's criterion, $B$ contains $k(\xi_{12})$ where $\xi_{12}$ is a primitive twelfth root of unity, hence there is an element $\lambda \in \mathcal{O}$ with $\lambda^6 = -1$. To show that $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)$ is torsion-free we have to exclude the existence of 6-torsions in $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)$. But since the reduced trace of $\lambda$ is $\pm\sqrt{3}$ which is not congruent 2 modulo $\mathfrak{p}_2\mathfrak{p}_3$, $\lambda$ is not contained in $\Gamma^1_{\mathcal{O}}(\mathfrak{P}_2\mathfrak{P}_3)$. The element $g = \lambda + 1$ has norm $\mathrm{Nrd}(g) = 2 + \sqrt{3}$ which is a totally positive unit of $\mathbb{O}_k$ unit, hence $g$ lies in $\Gamma^+_{\mathcal{O}} = \mathcal{O}^+/\mathbb{O}^*_k$, where $\mathcal{O}^+$ denotes the group of all units whose reduced norm is totally positive. The group $\Gamma^+_{\mathcal{O}}$ is an index-2-extension of $\Gamma^1_{\mathcal{O}}$ since the fundamental unit $2 + \sqrt{3}$ is totally positive. Also $g^2 = (2 + \sqrt{3})\lambda$ which shows that $g$ has order 12 in $\Gamma^+_{\mathcal{O}} \lhd N\Gamma^+_{\mathcal{O}}$. The image of $g$ in $N\Gamma^+_{\mathcal{O}}/\Gamma^1_{\mathcal{O}}$ is not trivial and the discussion in [Shavel 1978, pp. 223–224] shows that $N\Gamma^+_{\mathcal{O}}/\Gamma^*_{\mathcal{O}}$ is generated by the class of an element $\Pi \in N\Gamma^+_{\mathcal{O}}$ with $\mathrm{Nrd}(\Pi) = \pi_2\pi_3$ where $\mathfrak{p}_2 = \langle\pi_2\rangle$, $\mathfrak{p}_3 = \langle\pi_3\rangle$ (note that the generators $\pi_2$ and $\pi_3$ cannot be chosen totally positive). Therefore, $\mathrm{Aut}(X_{\Gamma^1_{\mathcal{O}}((\mathfrak{P}_2\mathfrak{P}_3))})$ is of order 24 and is an extension of $\mathbb{Z}/6\mathbb{Z}$ by the Klein's four group. □

## 5. Computations of the quotient surfaces

Let $S$ be a quaternionic fake quadric, $G$ a group of automorphisms of $S$, $S/G$ the quotient surface and let $\pi : Z \to S/G$ be the minimal desingularization map.

Let us first study the case where $G$ is generated by an involution $\sigma$.

**Proposition 5.22.** *An involution $\sigma$ has 4 fixed points. The invariants of $Z$ are*

$$K_Z^2 = 4, \quad c_2 = 8, \quad q = p_g = 0, \quad h^{1,1} = 5.$$

*The surface $Z$ is minimal of general type.*

*Proof.* By Lefschetz's formula (Proposition 2.6), $1 = \sum_{s=\sigma(s)} \frac{1}{4}$, therefore $\sigma$ has 4 fixed points. Their images in $S/\sigma$ are 4 $A_1$ singularities, resolved by 4 $(-2)$-curves on $Z$. The invariants of $Z$ are easy to compute.

The surface $Z$ is of general type and is minimal because $K_Z$ is the pullback of the nef divisor $K_{S/G}$. □

**Proposition 5.23.** *Let $G = \langle\sigma\rangle \cong \mathbb{Z}/3\mathbb{Z}$. The singularities of the quotient surface $S/G$ are $2A_{3,1} + 2A_{3,2}$. The resolution $Z$ has general type, and*

$$K_Z^2 = 2, \quad c_2 = 10, \quad q = p_g = 0.$$

*Proof.* We use the notations of Zhang's formula (Proposition 2.7). In this case this formula gives $r_1 + r_2 = 4$. An $A_{3,1}$ singularity is resolved by a $(-3)$-curve, and we have

$$K_Z^2 = \frac{8}{3} - \frac{r_1}{3}.$$

Therefore $r_1 = 2$ and $r_2 = 2$. The singularities of $S/G$ are $2A_{3,1} + 2A_{3,2}$. Moreover, as $q = p_g = 0$, we have $c_2 = 10$. $Z$ is of general type by Lemma 2.11. □

**Proposition 5.24.** *There is no quaternionic fake quadric $S$ with $G = (\mathbb{Z}/3\mathbb{Z})^2 \subset \operatorname{Aut} S$.*

*Proof.* Let $\sigma_1, \sigma_2$ be the two generating elements of $G$. Let $p$ be one of the 4 fixed points of $\sigma_1$ (see Proposition 5.23). Since $\sigma_1$ and $\sigma_2$ commute, the set of fixed points of $\sigma_1$ is sent to itself by $\sigma_2$, indeed there are two orbits of 2 elements because of the different local type of the action of $\sigma_1$. Now $\sigma_2$ has order three, hence it acts trivially on these 2 orbits and the conclusion is that there are 4 fixed points for the action of the whole group $G$. The faithful action of $G$ on the tangent space of $p$ can be diagonalized, hence there are elements with one eigenvalue equal to 1, contradicting Lemma 2.8 and Theorem 3.12. □

**Proposition 5.25.** *Let $G = \mathbb{Z}/4\mathbb{Z}$. The singularities of the quotient $S/G$ are $2A_{4,1} + 2A_{4,3}$ or $A_1 + 2A_{4,3}$. The invariants of the resolution $Z$ are*

$$K_Z^2 = 0, \quad c_2 = 12, \quad q = p_g = 0$$

*in the first case, and in the second case $Z$ is minimal and satisfies*

$$K_Z^2 = 2, \quad c_2 = 10, \quad q = p_g = 0.$$

**Remark 5.26.** Proposition 5.34 gives an example of the first case.

*Proof.* Let $s$ be a fixed point of an order 4 automorphism $\sigma$ acting on $S$. As the involution $\sigma^2$ has only isolated fixed points, the eigenvalues of $\sigma$ acting on $T_{S,s}$ cannot be $(i, -1)$ or $(-i, -1)$. Let $a, b, c$ be the number of fixed points such that the eigenvalues of $\sigma$ are $(i, i)$, $(-i, -i)$ and $(i, -i)$ respectively. The Lefschetz holomorphic fixed point formula implies

$$-\frac{a}{2i} + \frac{b}{2i} + \frac{c}{2} = 1 \quad \text{and} \quad a + b + c = 4 \text{ or } 2,$$

thus there are two cases:

(1) $a = b = 1$ and $c = 2$. The singularities of $S/G$ are $2A_{4,1} + 2A_{4,3}$.

(2) $a = b = 0$ and $c = 2$. In this case, the singularities of $S/G$ are $A_1 + 2A_{4,3}$ because $\sigma^2$ has 4 fixed points.

An $A_{4,1}$ singularity is resolved by a $(-4)$-curve $C_k$ and an $A_{4,3}$ singularity is resolved by a chain of three $(-2)$-curves and we have

$$K_Z = \pi^* K_{S/\sigma} - \sum_{k=1}^{k=2} \tfrac{1}{2} C_k,$$

thus $K_Z^2 = \frac{8}{4} - 2 = 0$ in the first case. Additionally,

$$e(S/\sigma) = \tfrac{1}{4}(4 + (4-1)4) = 4,$$

thus $c_2(Z) = 4 + 8 = 12$. The invariants in the second case are computed in a similar way. □

**Proposition 5.27.** *Let $G = \mathbb{Z}/5\mathbb{Z}$. The singularities of $S/G$ are $4A_{5,2}$ or $A_{5,1} + 2A_{5,2} + A_{5,4}$ or $2A_{5,1} + 2A_{5,4}$. The invariants of the surface $Z$ are, respectively,*

$$\begin{matrix} K_Z^2 = 0 \\ c_2 = 12 \end{matrix} \quad \textit{or} \quad \begin{matrix} K_Z^2 = -1 \\ c_2 = 13 \end{matrix} \quad \textit{or} \quad \begin{matrix} K_Z^2 = -2 \\ c_2 = 14, \end{matrix}$$

*and in any case $q = p_g = 0$.*

**Remark 5.28.** (1) In Proposition 5.31 below, we give an example of a surface such that the quotient by an order 5 automorphism has $2A_{5,1} + 2A_{5,4}$ singularities.

(2) By the same kind of arguments as for $(\mathbb{Z}/3\mathbb{Z})^2$ (see Proposition 5.24), there is no fake quadric $S$ with $(\mathbb{Z}/5\mathbb{Z})^2 \subset \operatorname{Aut} S$.

*Proof.* Using the notations of Proposition 2.7, the number of fixed points $r_1 + r_2 + r_3 + r_4$ equals 4. As $e(S/\sigma) = \frac{1}{5}(4 + (5-1)4) = 4$, Zhang's formula yields

$$(a_1, \dots, a_4) = \left(0, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{2}\right),$$

with

$$\sum 4a_i r_i = r_2 + r_3 + 2r_4 = 4.$$

Thus $r_1 = r_4$. Therefore the possibilities for $(r_1, r_2, r_3, r_4)$ are $(0, i, j, 0)$ with $i + j = 4$, or $(1, i, j, 1)$ with $i + j = 2$, or $(2, 0, 0, 2)$. The singularities on the quotient are, respectively,

$$4A_{5,2} \quad \text{or} \quad A_{5,1} + 2A_{5,2} + A_{5,4} \quad \text{or} \quad 2A_{5,1} + 2A_4.$$

A singularity $A_{5,i}$ $(i = 1, \dots, 4)$ contributes (respectively)

$$-\tfrac{9}{5}, -\tfrac{2}{5}, -\tfrac{2}{5}, 0$$

to $K_Z^2$. Thus the self-intersection number is

$$K_Z^2 = \tfrac{1}{5}(8 - 9r_1 - 2(r_2 + r_3)),$$

and according to the possible tuples $(r_1, \ldots, r_4)$ as above: $K_Z^2 = 0$, or $K_Z^2 = -1$, or $K_Z^2 = -2$. As $e(S/G) = 4$, we get $c_2 = 12$, 13, or 14 according to the three possible singular loci.

Let us justify our computation of $K_Z^2$. An $A_{5,1}$-singularity is resolved by a $(-5)$-curve $C_5$, thus we have to add $-\frac{3}{5}C_5$ to the canonical divisor. This contributes $\left(-\frac{3}{5}C_5\right)^2 = -\frac{9}{5}$ to $K_Z^2$. On the other hand, an $A_{5,2}$-singularity is resolved by a chain of two curves $C_2, C_3$ with $C_k^2 = -k$. We have to add $-\frac{2}{5}C_3 - \frac{1}{5}C_2$ to $\pi^* K_{S/G}$, and the contribution to $K_Z^2$ is

$$\left(\tfrac{2}{5}C_3 + \tfrac{1}{5}C_2\right)^2 = -\tfrac{2}{5}.$$

Finally, note that $A_{5,3} = A_{5,2}$, and the $A_{5,4}$-singularity does not contribute to $K_Z^2$. □

**Proposition 5.29.** *If $G = \mathbb{Z}/6\mathbb{Z}$, then $S/G$ has singularities $2A_{6,1} + 2A_{6,5}$. The minimal resolution $Z$ has invariants*

$$K_Z^2 = -4, \quad c_2 = 16, \quad q = p_g = 0.$$

*Proof.* Let $s$ be a fixed point of an order 6 automorphism $\sigma$. Let $\alpha$ be a primitive third root of unity. By Lemma 2.8, the action of $\sigma$ on $T_{S,s}$ has eigenvalues $(-\alpha, (-\alpha)^a)$ or $(-\alpha^2, (-\alpha^2)^a)$ with $a = 1$ or 5. Let $r_1$, $r_2$ and $r_3$ be respectively the number of fixed points of $\sigma$ with eigenvalues $(-\alpha, -\alpha)$, $(-\alpha^2, -\alpha^2)$ and $(-\alpha, -\alpha^5)$. Lefschetz fixed point formula (Proposition 2.6) implies the relation

$$\frac{r_1}{(1+\alpha)^2} + \frac{r_2}{(1+\alpha^2)^2} + r_3 = 1,$$

therefore $r_1 = r_2$ and $-r_1 + r_3 = 1$. By Corollary 2.5, $\sigma$ has 2 or 4 fixed points. The only possibility for $(r_1, r_3)$ is therefore $(1, 2)$. The singularities are $2A_{6,1} + 2A_{6,5}$ and the minimal resolution $Z$ of $S/\sigma$ has $K_Z^2 = \frac{8}{6} - 2 \cdot \frac{8}{3} = -4$. Moreover $e(Z) = \frac{1}{6}(4 + 5 \cdot 4) + 2 + 2 \cdot 5 = 16$. □

Let us study the case $G = \mathbb{Z}/8\mathbb{Z}$.

**Proposition 5.30.** *Let $\sigma$ be an order 8 element acting on $S$. The singularities of $S/\sigma$ are $2A_{8,3} + 2A_{8,5}$. The resolution $Z$ of the quotient surface is a surface with*

$$K_Z^2 = -2, \quad c_2(Z) = 14, \quad q = p_g = 0.$$

*Proof.* Let $p$ be a fixed point of $\sigma$ and let $\xi_{(p)}$ be a primitive eighth root of unity such that $\sigma$ acts on $T_{S,p}$ with eigenvalues $\xi_{(p)}$ and $\xi_{(p)}^{q_p}$ for $q_p \in \{0, \ldots, 7\}$. There are no reflections, so we have $\xi_{(p)}^j \neq 1$ and $\xi_{(p)}^{jq_p} \neq 1$ for $j = 1, \ldots, 7$, thus $q_p$ is prime to 2: $q_p \in \{1, 3, 5, 7\}$. Let $a_1, a_3, a_5$ and $a_7$ be the number of fixed points $p$ with $q_p = 1, 3, 5$ or 7 respectively. We have $\sum a_i = 2$ or 4. By summing over the powers $\sigma^k$ for $k = 1, \ldots, 7$ in the formula of the holomorphic Lefschetz theorem, we get

$$7=\sum_{p\in S^{\sigma}}\sum_{k=1}^{k=7}\frac{1}{\det(1-d\sigma^k|T_{S,p})},$$

and thus

$$7=\sum_{u=0}^{u=3}\sum_{k=1}^{k=7}\frac{a_{2u+1}}{(1-\xi^k)(1-\xi^{k(2u+1)})}=\tfrac{7}{4}a_1+\tfrac{5}{4}a_3+\tfrac{9}{4}a_5+\tfrac{21}{4}a_7.$$

The possibilities for $(a_1,\dots,a_4)$ are $(4,0,0,0)$, $(2,1,1,0)$, $(1,0,0,1)$ and $(0,2,2,0)$.

For $t^2$ of order 4, we have seen that the singularities of $S/\sigma^2$ are $2A_{4,1}+2A_{4,3}$ or $A_1+2A_{4,3}$. Thus the only possibility for $(a_1,\dots,a_4)$ is $(0,2,2,0)$, and the singularities of $S/\sigma$ are $2A_{8,3}+2A_{8,5}$. The Euler number of $S/\sigma$ is

$$e(S/\sigma)=\tfrac{1}{8}(4+7\cdot 4)=4.$$

Since $\frac{8}{3}=3-\frac{1}{3}$ and $\frac{8}{5}=2-\frac{1}{3-\frac{1}{2}}$ we get

$$e(Z)=4+2\cdot 2+2\cdot 3=14.$$

It is easy to check that a singularity $A_{8,3}$ decreases $K_Z^2$ by 1 and a singularity $A_{8,5}$ decreases $K_Z^2$ by $\frac{1}{2}$, thus we obtain $K_Z^2=\frac{8}{8}-2\cdot 1-2\cdot\frac{1}{2}=-2$. □

**Proposition 5.31.** *Let $S$ be a fake quadric with $G=\mathbb{Z}/10\mathbb{Z}\subset\mathrm{Aut}(S)$. The singularities of the quotient surface $S/G$ are $2A_{10,1}+2A_{10,9}$. The resolution $Z$ has the invariants*

$$K^2=-12,\quad c_2=24,\quad q=p_g=0.$$

*Proof.* Let $\sigma$ be an automorphism of order 10 acting on $S$. It has 2 or 4 fixed points. As the involution $\sigma^5$ has 4 fixed points, $\sigma$ cannot have 2 fixed points. Therefore

$$e(S/G)=\frac{1}{10}(4+(10-1)4)=4.$$

Let $\xi$ be a primitive fifth root of unity and $p$ a fixed point. There exist $a=a(p)$ and $b=b(p)$ integers invertible mod 5 such that the action of $\sigma$ on $T_{S,p}$ has eigenvalues $(-\xi^a,-\xi^{ba})$. The Lefschetz holomorphic fixed point formula yields

$$1=\sum_{p\in S^{\sigma}}\frac{1}{(1+\xi^a)(1+\xi^{ab})}.$$

For $b=1,2,3,4$, the sum

$$c(b)=\sum_{a=1}^{a=4}\frac{1}{(1+\xi^a)(1+\xi^{ab})}$$

is equal to $-4$, 1, 1, 6, respectively. Recall again that $A_{10,3} = A_{10,7}$. For $k \in \{1, 3, 9\}$, let $r_k$ be the number of points in $S^\sigma$ giving an $A_{10,k}$ singularity. The Lefschetz fixed point formula gives

$$4 = -4r_1 + r_3 + 6r_9.$$

Taking care of the relation $r_1 + r_3 + r_9 = 4$, we have the following possibilities for $(r_1, r_3, r_9)$: (0, 4, 0), (1, 2, 1) and (2, 0, 2).

The resolution of an $A_{10,3}$-singularity is a chain of 3 curves $C_2, C_2', C_4$ with intersection numbers $(-2)-(-2)-(-4)$. We have to add $-\frac{1}{5}(C_2 + C_2' + C_4)$ to $\pi^* K_{S/G}$. Each singularity contributes $(-\frac{1}{5}(C_2 + C_2' + C_4))^2 = -\frac{6}{5}$ to $K_Z^2$.

Similarly, the resolution of an $A_{10,1}$-singularity is a $(-10)$-curve $C_{10}$. An $A_{10,1}$-singularity decreases $K_{S/G}^2$ by $(-\frac{8}{10}C_{10})^2 = -\frac{32}{5}$.

When the singularities of $S/G$ are respectively $4A_{10,3}$, $A_{10,1} + 2A_{10,3} + A_{10,9}$ and $2A_{10,1} + 2A_{10,9}$, we have: $K_Z^2 = \frac{8}{10} - 4 \cdot \frac{6}{5} = -4$, $K_Z^2 = \frac{8}{10} - \frac{32}{5} - 2 \cdot \frac{6}{5} - 0 = -8$ and $K_Z^2 = \frac{8}{10} - 2 \cdot \frac{32}{5} = -12$, respectively. The Euler number of $Z$ is respectively $4 + 4 \cdot 2 = 12$, $4 + 1 + 2 \cdot 2 + 9 = 18$ and $4 + 2 + 2 \cdot 9 = 24$. Only the last case is possible because 12 has to divide $K_Z^2 + e(Z)$. □

**Proposition 5.32.** *Let $G = (\mathbb{Z}/2\mathbb{Z})^2$. The quotient surface $S/G$ contains 6 $A_1$ singularities. The surface $Z$ is minimal of general type and has the invariants*

$$K_Z^2 = 2, \quad c_2 = 10, \quad q = p_g = 0.$$

*Proof.* A faithful representation of $G$ on a 2-dimensional space contains reflections, therefore by Lemma 2.8, there are no points fixed by the whole $G$. The group $G$ contains 3 involutions. Each of these involutions has 4 isolated fixed points whose image in $X$ are $2A_1$ singularities. Thus there are $6A_1$ singularities on $S/G$ and we have

$$e(Z) = e(S/G) + 6 = \tfrac{1}{4}(4 + 12) + 6 = 10.$$

Moreover, $K_Z = \pi^* K_{S/G}$ is nef and $K_{S/G}^2 = K_S^2/4 = 2$. By Lemma 2.10, we have $q = p_g = 0$. □

**Remark 5.33.** (a) Fabrizio Catanese and Miles Reid pointed out to us that a minimal surface of general type with $c_1^2 = 2c_2 = 8$, $p_g = 0$ and automorphism group containing $G = (\mathbb{Z}/2\mathbb{Z})^3$ such that each involution has only isolated points must deform, therefore $(\mathbb{Z}/2\mathbb{Z})^3$ cannot be a subgroup of the automorphism group of a quaternionic fake quadric, which is a rigid surface. The complete argument is as follows: The minimal resolution $Z$ of the quotient $Y = X/G$ of a fake quadric $X$ by $G$ would be a numerical Godeaux surface, that is, a surface with $c_1^2 = 1$ and with the maximal number of nodal curves, being equal to 7. We do not know whether such a surface exists, but from coding theory one would have a covering $S \to Y$ with group $G$ ramified only over the nodes.

The covering surface $S$ would have $c_1^2=8$, $p_g=0$. However, the Kuranishi family of deformations for the surface $Z$ has dimension greater or equal to the expected dimension, which is $8=10\chi-2c_1^2$, and the 7 nodes impose at most 1 condition each, because of the morphism of the global deformation space to the local deformation space of the singularities. Therefore this family of 7-nodal numerical Godeaux surfaces would have 1 modulus, and therefore also the above surfaces $S$ would vary in moduli. However, quotients of the bidisk by an irreducible subgroup are rigid, for instance, by a theorem of Jost and Yau.

(b) For $G=\mathbb{Z}/4\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, the quotient surface $S/G$ has singularities $2A_1+2A_3$ and the desingularization $Z$ has invariants $K_Z^2=1$, $c_2=11$, $q=p_g=0$. We do not know if a fake quadric $S$ with such automorphism subgroup exists.

**Proposition 5.34.** *Let $G=\mathbb{D}_4$ acting on the fake quadric $S$. The singularities of $S/G$ are $4A_1+A_{4,3}+A_{4,1}$. The resolution $Z$ of the quotient surface has invariants*

$$K_Z^2=0,\quad c_2(Z)=12,\quad q=p_g=0.$$

*The elements of order 4 in $\mathbb{D}_4$ have 4 fixed points.*

*Proof.* Let $t$ and $a$ be the generators of $\mathbb{D}_4$ such that $t^4=1$, $a^2=1$ and $at=t^3a$. The elements of order 4 are $t$ and $t^3$. The elements of order 2 are $a$, $ta$, $t^2a$, $t^3a$ and $t^2$.

There cannot be a point of $S$ that is fixed by the whole group $G$ because any faithful 2-dimensional representation of $G$ contains a reflection $(x,y)\to(x,-y)$ and thus such a point would lie on a curve fixed by an involution, but an automorphism of $S$ has only isolated fixed points.

First case: Suppose that $t$ has 4 fixed points, $\mathrm{Fix}(t)=\{p_1,ap_1,p_2,ap_2\}$. The Euler number of $S/G$ is

$$e(S/G)=\tfrac{1}{8}(4+(2-1)(4\cdot 4)+(4-1)4)=4.$$

The singularities on $S/G$ are $4A_1+A_{4,3}+A_{4,1}$ and therefore

$$e(Z)=4+4+3+1=12.$$

Moreover $K_Z^2=\frac{8}{8}+(-\frac{1}{2})^2(-4)=0$.

Second case: Suppose that $t$ has 2 fixed points, $\mathrm{Fix}(t)=\{p_1,ap_1\}$. The Euler number of $e(S/G)$ would be

$$\tfrac{1}{8}(4+(2-1)18+(4-1)2)=\tfrac{7}{2},$$

but this is not an integer. $\square$

**Proposition 5.35.** *Suppose that the dihedral group* $\mathbb{D}_8$ *of order* 16 *acts on fake quadric* $S$. *The singularities of* $S/\mathbb{D}_8$ *are* $4A_1 + A_{8,3} + A_{8,5}$. *The resolution* $Z$ *of the quotient surface has invariants*

$$K_Z^2 = -1, \quad c_2(Z) = 13, \quad q = p_g = 0.$$

*Proof.* Let $t$ and $a$ be generators of $\mathbb{D}_8$ such that $t^8 = a^2 = 1$ and $at = t^7a$. Order 8 elements in $G$ are $t, t^3, t^5, t^7$; order 4 elements are $t^2$, $t^6$; order 2 elements are $a$, $ta$, $t^2a$, $t^3a$, $t^4a$, $t^5a$, $t^6a$, $t^7a$ and $t^4$.

By the discussion on order 8 elements, $t$ has 4 fixed points, say $p_1$, $ap_1$, $p_2$, $ap_2$. Let $p$ be a fixed point of an involution $\sigma \neq t^4$. The orbit of $p$ under $G$ consists of 8 elements, each is a fixed point of an involution $\neq t^4$. The quotient surface has $\frac{1}{8} \cdot 8 \cdot 4A_1 + A_{8,3} + A_{8,5}$ singularities. We have

$$e(S/G) = \tfrac{1}{16}(4 + 1 \cdot (8 \cdot 4) + 7 \cdot 4) = 4,$$

and $e(Z) = 4+4+2+3 = 13$. Moreover $K_Z^2 = \frac{8}{16} - 1 - \frac{1}{2} = -1$. □

## 6. Reconstruction of a surface knowing its quotient

Miyaoka [1984] gives a bound on the number of disjoint $(-2)$-curves on a minimal smooth surface $Y$. This implies in particular that if $c_1^2 = 4,\ 2$ or 1 and $\chi(\mathcal{O}_Y) = 1$, there are at most 4, 6 and 7 such curves respectively. The surfaces with $c_1^2 = 4,\ 2$ we obtained as quotient of quaternionic fake quadrics reach that bound. For the cases $c_1^2 = 2$ these surfaces seem to be the first known ones with that property.

Dolgachev, Mendes Lopes, and Pardini [2002] study rational surfaces with the maximal number of $(-2)$-curves. For that aim they use and develop the theory of $(\mathbb{Z}/2\mathbb{Z})^n$-covers ramified over $A_1$ singularities. Using their results, we obtain:

**Proposition 6.36.** *Let* $Y$ *be a smooth minimal surface of general type with* $q = p_g = 0$ *and* ${}_2\mathrm{Pic}(Y) = 0$.

(a) *If* $c_1(Y)^2 = 4$, $c_2(Y) = 8$ *and* $Y$ *contains* 4 *disjoint* $(-2)$*-curves* $C_1, \dots, C_4$, *then there exists a double cover of* $Y$ *ramified over the curves* $C_i$. *The minimal model of this covering has invariants* $c_1^2 = 2c_2 = 8$ *and* $q \leq 1$.

(b) *If* $c_1(Y)^2 = 2$, $c_2(Y) = 10$ *and* $Y$ *contains* 6 *disjoint* $(-2)$*-curves* $C_1, \dots, C_6$, *then there exists a bidouble cover of* $Y$ *ramified over the curves* $C_i$. *The minimal model of this covering has invariants* $c_1^2 = 2c_2 = 8$ *and* $q \leq 2$.

(c) *If* $c_1(Y)^2 = 1$, $c_2(Y) = 11$ *and* $Y$ *contains* 7 *disjoint* $(-2)$*-curves* $C_1, \dots, C_7$, *then there exists a* $(\mathbb{Z}/2\mathbb{Z})^3$*-cover of* $Y$ *ramified over the curves* $C_i$. *The minimal model of this covering has invariants* $c_1^2 = 2c_2 = 8$ *and* $q \leq 2$.

Let $\mathbb{F}_2$ be the field with 2 elements. Let $C_1, \dots, C_k$ be $k$ $(-2)$-curves on a smooth surface $Y$. Let

$$\psi : {\mathbb{F}_2}^k \to \mathrm{Pic}(Y) \otimes \mathbb{F}_2$$

be the homomorphism sending $v = (v_1, \dots, v_k)$ to $\sum v_i C_i$. We say that the curve $C_j$ appears in the kernel $\ker\psi$ if there is a vector $v = (v_1, \dots, v_k)$ in $\ker\psi$ such that $v_j = 1$. For $v$ in $\ker\psi$, we denote by $L_v$ an element of $\mathrm{Pic}(Y)$ such that $2L_v = \sum v_i C_i$ (we sometimes identify elements of $\mathbb{F}_2$ with 0, 1 in $\mathbb{Z}$). We have:

**Proposition 6.37** [Dolgachev et al. 2002, Proposition 2.3]. *Suppose that* ${}_2\mathrm{Pic}(Y)$ *is zero. There exists a unique smooth connected Galois cover* $\pi : Z \to Y$ *such that the Galois group of* $\pi$ *is* $G = \mathrm{Hom}(\ker\psi, \mathbb{G}_m)$, *the branch locus of* $\pi$ *is the union of the* $C_i$ *appearing in* $\ker\psi$ *and the surface* $\bar{Z}$ *obtained by contracting the* $(-1)$*-curves over the* $(-2)$*-curves in* $Y$ *has invariants*

$$K_{\bar{Z}}^2 = 2^r K_Y^2 \, c_2(\bar{Z}) = \chi(\mathcal{O}_{\bar{Z}}) = 2^r \chi(\mathcal{O}_Y) - k2^{r-3} \quad \textit{and} \quad \kappa(\bar{Z}) = \kappa(Y),$$

*where* $r = \dim V$.

*Proof of Proposition 6.36.* We have to prove that for our surface $Y$, $\ker\psi$ has the required dimension and that all the curves appear in $\ker\psi$. For $c_1^2(Y) = 4$ and 2, we have $b_2(Y) = h^{1,1}(Y) = 6$, 8 and 9 respectively. As we supposed that ${}_2\mathrm{Pic}(Y) = 0$, the space $\mathrm{Pic}(Y) \otimes \mathbb{F}_2$ is $h^{1,1}$-dimensional. As $p_g = 0$, it has moreover a nondegenerate intersection pairing and therefore the dimension of a totally isotropic space in $\mathrm{Pic}(Y) \otimes \mathbb{F}_2$ is at most $\left[h^{1,1}/2\right] = 3$, 4, and 4 dimensional respectively. The image of $\psi$ is the totally isotropic space generated by the curves $C_i$, therefore the dimension $r$ of $\ker\psi$ is at least 1, 2 and 3 respectively.

A smooth double cover of a surface with $n$ nodes can exist only if $n$ is divisible by 4 (see [Dolgachev et al. 2002]). Therefore the vectors $v = (v_1, \dots, v_k)$ in $\ker\psi$ (of dimension $\leq 7$) have weight 4, that is, the number of indices $j$ such that $v_j = 1$ is 4.

In case (a), $\ker\psi$ is one-dimensional, generated by $w_1 = (1, 1, 1, 1)$. For (b), as every vector in $\ker\psi$ has weight 4, by [Beauville 1980, Lemme 1], we have $k \geq 2^r - 1$ and thus $r \leq 2$ and $r \leq 3$ respectively. Moreover, it is easy to check that in the case (b), the space $\ker\psi$ is (up to permutation of the basis vectors) generated by $w_1 = (1, 1, 1, 1, 0, 0)$ and $w_2 = (1, 1, 0, 0, 1, 1)$.

In case (c) [Beauville 1980, Lemme 1] implies that $\ker\psi$ is (up to permutation) generated by $w_1 = (1, 0, 0, 1, 1, 0, 1)$, $w_2 = (0, 1, 0, 1, 0, 1, 1)$ and $w_3 = (0, 0, 1, 0, 1, 1, 1)$.

The surface $\bar{Z}$ obtained by contracting the $(-1)$-curves over the $(-2)$-curves $C_i$ is minimal because no surface with $c_1^2 = 3c_2 = 9$ has an order 2 automorphism. □

Let us give a bound on the irregularity.

**Lemma 6.38.** *Let $Y$ be a surface of general type with $\chi = 1$ and $q = 0$ containing a 2-divisible set of 4 $(-2)$-curves. Let $Y' \to Y$ be the double cover. Then $q(Y') \leq 1$.*

*Proof.* As $q(Y) = 0$, the involution $\sigma$ on $Y'$ given by the cover $Y' \to Y$ acts as multiplication by $-1$ on $H^0(Y', \Omega_{Y'})$. Therefore, $\sigma$ acts trivially on $\bigwedge^2 H^0(Y', \Omega_{Y'})$. As $p_g(Y) = 0$, the map $\bigwedge^2 H^0(Y', \Omega_{Y'}) \to H^0(Y', \bigwedge^2 \Omega_{Y'})$ must be 0. Let $Y' \to Y''$ be the blow-down map of the 4 $(-1)$-curves over the 4 nodal curves of $Y$. If $q(Y'') \geq 1$, Castelnuovo–De Franchis Theorem implies that there is a fibration onto a curve $B$ of genus $q(Y'')$. By [Zucconi 2003], we get that $q(Y'') \leq 2$ and if $q(Y'') = 2$, then $Y''$ is an étale bundle of genus 2 fibers onto a genus 2 curve $B$ and $K^2_{Y''} = 8$. In that case, there is a commutative diagram

$$\begin{array}{ccc} Y'' & \to & X \\ \downarrow & & \downarrow \\ B & \to & \mathbb{P}^1 \end{array}$$

where the vertical maps are genus 2 fibrations and $X$ is the surface obtained by contracting the 4 $(-2)$-curves on $Y$. This diagram is obtained from $B \to \mathbb{P}^1$ by taking base change and normalizing. Since $Y'' \to X$ is unramified in codimension 1, the 6 fibers of $X \to \mathbb{P}^1$ occurring at the 6 branch points of $B \to \mathbb{P}^1$ are double. Since $X$ has only 4 singular points, $X \to \mathbb{P}^1$ has at least two double fibers contained in the smooth locus of X, but a multiple fiber on a genus 2 fibration cannot exist (because of the adjunction formula). Thus $q \leq 1$. □

Let us now consider a smooth minimal surface of general type $Z$ with $K^2 = 2$, $c_2 = 10$, $q = p_g = 0$ such that there is a birational map onto a surface $Y$ with singularities $2A_{3,1} + 2A_{3,2}$.

**Proposition 6.39.** *Suppose that ${}_3\mathrm{Pic}(Z) = 0$. There exists a smooth triple cover $X$ of $Y$ ramified precisely over the singularities of $Y$. The surface $X$ is of general type and has invariants $c_1^2 = 2c_2 = 8$.*

*Proof.* Let $D_1, D_2$ be the $(-3)$-curves over the singularities $A_{3,1}$ and let $D_3, \ldots, D_6$ be the $(-2)$-curves over the singularities $A_{3,2}$, with indices satisfying $D_3 D_4 = D_5 D_6 = 1$. Let $W \to Y$ be the blow-up at the intersection points of $D_3$, $D_4$ and of $D_5$, $D_6$. Let $C_1, \ldots, C_6$ be the strict transforms of the $D_i$ in $W$. Let

$$\psi : {\mathbb{F}_3}^6 \to \mathrm{Pic}(W) \otimes \mathbb{F}_3 = H^2(W, \mathbb{F}_3)$$

be the homomorphism sending $v = (v_1, \ldots, v_k)$ to $\sum v_i C_i$. The image of $\psi$ is a totally isotropic subspace in $H^2(W, \mathbb{F}_3)$. As $b_2(W) = 10$, this image is at most 5-dimensional and therefore $\dim \ker \psi \geq 1$. Let $v = (v_1, \ldots, v_6) \in \ker \psi$, $v \neq 0$. We choose the representatives of $\mathbb{F}_3$ in $\{0, 1, 2\}$. There exists a unique invertible sheaf $L$ such that

$$3L = \sum v_i C_i.$$

Let $T$ be the triple cover of $W$ ramified over the $r$ curves $C_i$ such that $v_i \neq 0$. The surface $T$ is smooth outside the curves $C_i$ with $v_i = 2$. Let $R$ be the minimal resolution of $T$ and let $f : R \to W$ be the composite map. By [Urzúa 2010, Propositions 2.2, 4.1 and 4.3], the invariants of $R$ are

$$K_R =_{\mathrm{num}} f^*\big(K_W + \tfrac{2}{3}\Sigma\big), \quad c_2(R) = 3c_2(W) - 4r, \quad \chi(\mathcal{O}_R) = 3\chi(\mathcal{O}_W) - \tfrac{1}{3}r,$$

where $\Sigma$ is the sum of the $r$ curves $C_i$ such that $v_i \neq 0$. Therefore $r = 3$ or 6 and

$$K_R^2 = 0, \quad c_2(R) = 36 - 4r, \quad \chi(\mathcal{O}_W) = 3 - \tfrac{1}{3}r.$$

As there are at least 3 curves $C_i$ in the branch locus, one of the curves $C_3, \dots, C_6$ is in that branch locus. Say it is $C_3$. Let $E$ be the exceptional curve going through $C_3$. As $C_3E = C_4E = 1$ and $E \sum v_i C_i$ is divisible by 3, it forces $C_4$ to be also in the branch locus and thus $r = 6$ (and $\dim \ker \psi = 1$). The inverse image of the 6 $(-3)$-curves are $(-1)$-curves. By the formula giving $K_R$, the inverse image of the two exceptional curves are $(-3)$-curves meeting two $(-1)$-curves. We can therefore effectuate 8 blow-downs and we obtain a fake quadric. It has general type because $Y$ has general type, it is minimal because the quotient of a fake plane by an order 3 automorphism with 4 isolated fixed points has $4A_2$ singularities. □

## Acknowledgements

The authors warmly thank Fabrizio Catanese and Miles Reid for pointing out an error in a previous version. We are indebted to Fabrizio Catanese for his explanations on that point (see Remark 5.33 (a)) and to Håkan Granath for his help to correct that error. We thank also Ingrid Bauer, Margarida Mendes Lopes, and Rita Pardini for many useful discussions, and last but not least, we thank the referee for many useful comments and suggestions, in particular for the insistence to develop the $\mathbb{Q}$-homology quadrics point of view (see Remark 3.13). Part of this research was done during Roulleau's stay in Strasbourg University, and in the Instituto Superior Technico under grant FCT SFRH/BPD/72719/2010 and project Geometria Algebrica PTDC/MAT/099275/2008.

## References

[Atiyah and Singer 1968] M. F. Atiyah and I. M. Singer, "The index of elliptic operators, III", *Ann. of Math.* (2) **87** (1968), 546–604. MR 38 #5245 Zbl 0164.24301

[Barth et al. 2004] W. P. Barth, K. Hulek, C. A. M. Peters, and A. Van de Ven, *Compact complex surfaces*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **4**, Springer, Berlin, 2004. MR 2004m:14070 Zbl 1036.14016

[Bauer et al. 2008] I. C. Bauer, F. Catanese, and F. Grunewald, "The classification of surfaces with $p_g = q = 0$ isogenous to a product of curves", *Pure Appl. Math. Q.* **4**:2, Part 1 (2008), 547–586. MR 2009a:14046 Zbl 1151.14027

[Beauville 1980] A. Beauville, "Sur le nombre maximum de points doubles d'une surface dans $\mathbf{P}^3$ $(\mu(5) = 31)$", pp. 207–215 in *Journées de Géometrie Algébrique* (Angers, 1979), edited by A. Beauville, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. MR 82k:14037 Zbl 0445.14016

[Borel 1981] A. Borel, "Commensurability classes and volumes of hyperbolic 3-manifolds", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **8**:1 (1981), 1–33. MR 82j:22008 Zbl 0473.57003

[Deuring 1968] M. Deuring, *Algebren*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete **41**, Springer, Berlin, 1968. MR 37 #4106 Zbl 0159.04201

[Dolgachev et al. 2002] I. Dolgachev, M. Mendes Lopes, and R. Pardini, "Rational surfaces with many nodes", *Compositio Math.* **132**:3 (2002), 349–363. MR 2003g:14049 Zbl 1059.14050

[Džambić 2013] A. Džambić, "Fake quadrics from irreducible lattices acting on the product of upper half planes", preprint, 2013. arXiv 1305.5174

[Granath 2002] H. Granath, *On quaternionic Shimura surfaces*, thesis, Chalmers Tekniska Hogskola, Göteborg, 2002. MR 2715512

[Hirzebruch 1987] F. Hirzebruch, "Kommentare", pp. 755–804 in *Gesammelte Abhandlungen*, vol. I, Springer, Berlin, 1987. MR 89h:01106 Zbl 0627.01044

[Inose and Mizukami 1979] H. Inose and M. Mizukami, "Rational equivalence of 0-cycles on some surfaces of general type with $p_g = 0$", *Math. Ann.* **244**:3 (1979), 205–217. MR 82c:14009 Zbl 0444.14006

[Iskovskikh and Shafarevich 1989] V. A. Iskovskikh and I. R. Shafarevich, "Алгебраические поверхности", pp. 131–263 in *Алгебраическая геометрия, II*, edited by I. R. Shafarevich, Itogi Nauki Tekh., Ser. Sovrem. Probl. Mat., Fundam. Napravleniya **35**, Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1989. Translated as "Algebraic surfaces" in *Algebraic geometry, II*, Encyclopaedia of Math. Sci. **35**, 127–262, Springer, Berlin, 1996. MR 91f:14029 Zbl 0733.14015

[Keum 2006] J. Keum, "A fake projective plane with an order 7 automorphism", *Topology* **45**:5 (2006), 919–927. MR 2008b:14065 Zbl 1099.14031

[Keum 2008] J. Keum, "Quotients of fake projective planes", *Geom. Topol.* **12**:4 (2008), 2497–2515. MR 2009g:14042 Zbl 1222.14088

[Keum 2012] J. Keum, "Toward a geometric construction of fake projective planes", *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **23**:2 (2012), 137–155. MR 2924897 Zbl 1248.14044

[Margulis 1991] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **17**, Springer, Berlin, 1991. MR 92h:22021 Zbl 0732.22008

[Matsushima and Shimura 1963] Y. Matsushima and G. Shimura, "On the cohomology groups attached to certain vector valued differential forms on the product of the upper half planes", *Ann. of Math.* (2) **78** (1963), 417–449. MR 27 #5274 Zbl 0141.38704

[Miyaoka 1984] Y. Miyaoka, "The maximal number of quotient singularities on surfaces with given numerical invariants", *Math. Ann.* **268**:2 (1984), 159–171. MR 85j:14060 Zbl 0521.14013

[Prasad and Yeung 2012] G. Prasad and S.-K. Yeung, "Nonexistence of arithmetic fake compact Hermitian symmetric spaces of type other than $A_n$ $(n \leq 4)$", *J. Math. Soc. Japan* **64**:3 (2012), 683–731. MR 2965425 Zbl 1266.22015

[Riehm 1970] C. Riehm, "The norm 1 group of a $\mathfrak{p}$-adic division algebra", *Amer. J. Math.* **92** (1970), 499–523. MR 41 #6860 Zbl 0199.37601

[Roulleau 2012] X. Roulleau, "Quotients of Fano surfaces", *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **23**:3 (2012), 325–349. MR 2960841 Zbl 1252.14023

[Shavel 1978] I. H. Shavel, "A class of algebraic surfaces of general type constructed from quaternion algebras", *Pacific J. Math.* **76**:1 (1978), 221–245. MR 58 #28002 Zbl 0422.14022

[Shimizu 1963] H. Shimizu, "On discontinuous groups operating on the product of the upper half planes", *Ann. of Math.* (2) **77** (1963), 33–71. MR 26 #2641 Zbl 0218.10045

[Shimura 1967] G. Shimura, "Construction of class fields and zeta functions of algebraic curves", *Ann. of Math.* (2) **85** (1967), 58–159. MR 34 #4268 Zbl 0204.07201

[Urzúa 2010] G. Urzúa, "Arrangements of curves and algebraic surfaces", *J. Algebraic Geom.* **19**:2 (2010), 335–365. MR 2011e:14072 Zbl 1192.14033

[Vignéras 1980] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Mathematics **800**, Springer, Berlin, 1980. MR 82i:12016 Zbl 0422.12008

[Zhang 2001] D.-Q. Zhang, "Automorphisms of finite order on rational surfaces", *J. Algebra* **238**:2 (2001), 560–589. MR 2002c:14064 Zbl 1057.14053

[Zimmer 1984] R. J. Zimmer, *Ergodic theory and semisimple groups*, Monographs in Mathematics **81**, Birkhäuser, Basel, 1984. MR 86j:22014 Zbl 0571.58015

[Zucconi 2003] F. Zucconi, "Surfaces with $p_g = q = 2$ and an irrational pencil", *Canad. J. Math.* **55**:3 (2003), 649–672. MR 2004d:14055 Zbl 1053.14042

Received January 18, 2012. Revised May 24, 2013.

AMIR DŽAMBIĆ
INSTITUT FÜR MATHEMATIK
JOHANN WOLFGANG GOETHE UNIVERSITÄT
ROBERT-MAYER-STR. 6–8
D-60325 FRANKFURT AM MAIN
GERMANY

dzambic@math.uni-frankfurt.de

XAVIER ROULLEAU
LABORATOIRE DE MATHÉMATIQUES ET APPLICATIONS
UNIVERSITÉ DE POITIERS
TÉLÉPORT 2 - BP 30179
86962 FUTUROSCOPE CHASSENEUIL
FRANCE

roulleau@math.univ-poitiers.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.121

# DISTANCE OF BRIDGE SURFACES FOR LINKS WITH ESSENTIAL MERIDIONAL SPHERES

YEONHEE JANG

**Bachman and Schleimer gave an upper bound for the distance of a bridge surface of a knot in a 3-manifold which admits an essential surface in the exterior. Here we give a sharper upper bound for the distance of a bridge surface of a link when the manifold admits an essential meridional sphere in the exterior.**

## 1. Introduction

Let $L$ be a link in a closed orientable 3-manifold $M$. A closed orientable surface $F$ embedded in $M$ is called a *Heegaard surface* of $M$ if it cuts $M$ into two handlebodies. We call this decomposition a *Heegaard splitting* of $M$. We say that $L$ is in *bridge position* with respect to a Heegaard surface $F$ if the intersection of $L$ and each handlebody is *trivial*, namely, the intersection together with some arcs on $F$ bounds mutually disjoint disks. We call $F$ a $(g, n)$*-bridge surface* (or a *bridge surface* in brief) of $L$, where $g$ is the genus of $F$ and $n$ is the half of the number $|L \cap F|$ of the components of $L \cap F$. In particular, we call a $(0, n)$-bridge surface an $n$*-bridge sphere* of $L$. Throughout this paper, we assume $n \geq 3$ for all $n$-bridge spheres.

Since the *distance* of a Heegaard splitting was introduced in [Hempel 2001] as a measure of complexity, it has been studied by various authors; see, for example, [Evans 2006; Hartshorn 2002; Kobayashi and Rieck 2009; Scharlemann and Tomova 2006]. This concept can be generalized to the distance of bridge surfaces of links in closed orientable 3-manifolds (see Section 2 for details). As generalizations of results from [Hartshorn 2002; Scharlemann and Tomova 2006], Bachman and Schleimer [2005] and Tomova [2007] gave upper bounds for the distance of a bridge surface of a knot in a 3-manifold when there exist essential surfaces in the knot exterior and alternate bridge surfaces, respectively, in terms of their Euler characteristics. Ido [2013] gave a refinement of the upper bound of [Tomova 2007] in the case where the genus of the bridge surface is 0.

This research is partially supported by a Grant-in-Aid for JSPS Research Fellowships for Young Scientists.
*MSC2010:* 57M25.
*Keywords:* bridge surfaces for links, distance of bridge surfaces.

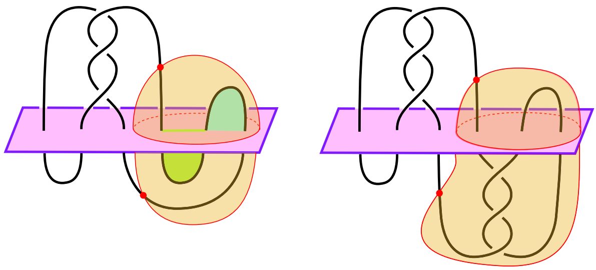


**Figure 1.** $d_{BS}(L, F) = 0$ and $d_T(L, F) = 1$.

In Theorem 1.1 and Corollary 1.3 below we give a refinement of Bachman and Schleimer's upper bound for the distance of bridge surfaces under some extra assumptions. (For detailed definitions, see Section 2.) For a surface $S$ in $M$, we denote by $S_L$ the surface $\mathrm{Cl}(S \setminus N(L))$, where $N(L)$ is a regular neighborhood of $L$ in $M$.

**Theorem 1.1.** *Let $L$ be a link in a closed orientable irreducible 3-manifold $M$ which is in bridge position with respect to a Heegaard surface $F$. Suppose that there exists a c-essential sphere $S$ in $M$ intersecting $L$ transversely in at least 4 points. Then $d_{BS}(L, F) \leq -\chi(S_L) = |\partial S_L| - 2$.*

Bachman and Schleimer's upper bound in this setting is $-\chi(S_L) + 2$, which equals $|\partial S_L|$.

We will denote by $d_{BS}(L, F)$ and $d_T(L, F)$ the definitions of distance given in [Bachman and Schleimer 2005] and [Tomova 2007], which disagree slightly. In general, it is easy to see that $d_{BS}(L, F) \leq d_T(L, F) \leq d_{BS}(L, F) + 2$. If we focus on bridge spheres for links in the 3-sphere $S^3$, we have:

**Proposition 1.2.** *For an $n$-bridge sphere $F$ of a link $L$ in $S^3$,*

- *if $d_{BS}(L, F) \geq 1$, then $d_T(L, F) = d_{BS}(L, F)$, and*
- *if $d_{BS}(L, F) = 0$, then $d_T(L, F) = 0$ or 1.*

The links and the 3-bridge spheres in Figure 1 give examples for which the two distances do not coincide, since $d_{BS}(L, F) = 0$ and $d_T(L, F) = 1$. In fact, this always holds when $L$ is nonsplit and either $L$ is composite or $F$ is perturbed.

The following is a direct consequence of Theorem 1.1 and Proposition 1.2.

**Corollary 1.3.** *Let $L$ be a link in the 3-sphere $S^3$ and $F$ an $n$-bridge sphere of $L$. Suppose that there exists a c-essential sphere $S$ in $M$ intersecting $L$ transversely in at least 4 points. Then $d_T(L, F) \leq -\chi(S_L) = |\partial S_L| - 2$.*

As a consequence of Theorem 1.1 and Corollary 1.3, we obtain:

**Corollary 1.4.** *Let $L$ be an arborescent link in the 3-sphere $S^3$. Then $d_{BS}(L, F) \leq 2$ and $d_T(L, F) \leq 2$ for any minimal bridge sphere $F$ of $L$.*

**Corollary 1.5.** *Let $L$ be a link in the 3-sphere $S^3$ and $F$ a minimal bridge sphere such that $d_{BS}(L, F) > 2$ or $d_T(L, F) > 2$. Then $L$ is a hyperbolic link and the double branched covering $M_2(L)$ of $S^3$ branched along $L$ is a hyperbolic manifold.*

Corollary 1.5 implies Corollary 6.2 of [Bachman and Schleimer 2005], which asserts the hyperbolicity of links admitting bridge surfaces with distance greater than 2. In fact, arborescent links are known to be hyperbolic except for some special cases (see [Bonahon and Siebenmann 2010; Futer and Guéritaud 2009; Jang 2011, Proposition 3]). On the other hand, the double branched covering $M_2(L)$ of $S^3$ branched along an arborescent link $L$ is a graph manifold, and hence not hyperbolic. Thus, the latter assertion in Corollary 1.5 is meaningful. We remark that, in fact, the hyperbolicity of $M_2(L)$ implies the hyperbolicity of the link $L$ (see [Kojima 1996; 1998]). Also, we conjecture that the assumptions on the minimality of the bridge spheres in Corollaries 1.4 and 1.5 are unnecessary. Specifically, we make the following conjectures:

(1) $d_{BS}(L, F) \leq 2$ and $d_T(L, F) \leq 2$ for any bridge sphere $F$ of an arborescent link $L$ in the 3-sphere $S^3$.

(2) For a link $L$ in $S^3$ which admits a bridge sphere $F$ such that $d_{BS}(L, F) > 2$ or $d_T(L, F) > 2$, the link $L$ is a hyperbolic link and the double branched covering $M_2(L)$ of $S^3$ branched along $L$ is a hyperbolic manifold.

Statements (1) and (2) are known to be true except for 3-bridge Montesinos links (see the proof of Corollaries 1.4 and 1.5). In fact, they are true if any nonminimal bridge sphere of a 3-bridge Montesinos link has distance at most 2 (or if any nonminimal bridge sphere of a 3-bridge Montesinos link is perturbed, which implies that the distance is at most 1).

## 2. Definitions and notation

Our conventions mostly follow [Bachman and Schleimer 2005], though we modify some of the definitions since we treat only meridional spheres in this paper, while Bachman and Schleimer treated more general surfaces.

Throughout this paper, $M$ is a closed orientable 3-manifold and $L$ is a link in $M$. We denote the manifold $\mathrm{Cl}(M \setminus N(L))$ by $M_L$. For a surface $F$ embedded in $M$ that intersects $L$ transversely, we denote the surface $F \cap M_L$ by $F_L$ and call it a *meridional surface* (with respect to $L$). A simple closed curve on $F_L$ is *inessential* on $F_L$ if it bounds a disk on $F_L$ or it bounds an annulus on $F_L$ together with a boundary component of $F_L$. We say that the curve is *essential* on $F_L$ if it is not inessential on $F_L$. A *compressing disk* for $F_L$ is a disk $D$ embedded in $M_L$ so that $F \cap D = \partial D$ is an essential simple closed curve on $F_L$. A *cut-disk* for $F_L$ is the intersection $D^c = D \cap M_L$, where $D(\subset M)$ is a disk such that $D \cap F = \partial D$ is an

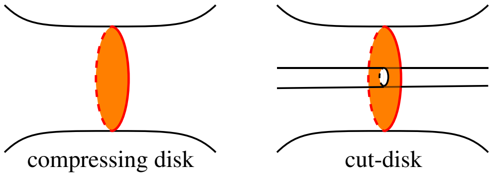



**Figure 2.** c-disks.

essential simple closed curve on $F_L$ and $|D \cap N(L)|$ is a meridian disk (i.e., $D$ intersects $L$ transversely in one point). A *c-disk* for $F_L$ is either a compressing disk or a cut-disk for $F_L$ (see Figure 2). We say that a surface $F \subset M$ is *c-essential* if there are no c-disks for $F_L$, $F_L$ is not boundary parallel in $M_L$ and $F$ is not a 2-sphere that bounds a 3-ball in $M_L$.

Let $L$ be a link in $M$ which is in a bridge position with respect to a Heegaard surface $F$ of $M$. We denote by $\mathscr{C}(F_L)$ the *curve complex* of $F_L$, that is, each vertex of $\mathscr{C}(F_L)$ corresponds to the isotopy class of an essential simple loop in $F_L$ and $k+1$ distinct vertices form a $k$-simplex if and only if there are mutually disjoint representatives of the corresponding isotopy classes. For two vertices $v$ and $v'$ of $\mathscr{C}(F_L)$, we denote by $d(v, v')$ the number of 1-simplexes in the shortest path (of 1-simplexes) connecting $v$ and $v'$. For two sets $A$ and $B$ of vertices of $\mathscr{C}(F_L)$, we define $d(A, B)$ by the minimum of $\{d(v, v') \mid v \in A, v' \in B\}$. Let $V_0$ and $V_1$ be the closures of the two components of $M \setminus F$, and let $H_i = V_i \cap M_L$ $(i = 0, 1)$. For each $i = 0, 1$, we denote by $\mathscr{D}_{BS}(H_i)$ (resp. $\mathscr{D}_T(H_i)$) the set of the vertices of $\mathscr{C}(F_L)$ with representatives bounding c-disks (resp. compressing disks) in $H_i$. We define the distances $d_{BS}(L, F)$ and $d_T(L, F)$ of $L$ with respect to $F$ as $d(\mathscr{D}_{BS}(H_0), \mathscr{D}_{BS}(H_1))$ and $d(\mathscr{D}_T(H_0), \mathscr{D}_T(H_1))$, respectively.

Let $V$ be a handlebody and $T$ the union of trivial arcs properly embedded in $V$. We say that a finite graph $\Sigma$ properly embedded in $V$ is a *spine* of $(V, T)$ if $V \setminus \Sigma$ is homeomorphic to $\partial V \times [0, 1)$ and the projection $V \setminus \Sigma \cong \partial V \times [0, 1) \to [0, 1)$ has no maxima on $T$. Let $L$ be a link in $M$ which is in a bridge position with respect to a Heegaard surface $F$ of $M$, and let $V_0$ and $V_1$ be the closures of the two components of $M \setminus F$. For each $i = 0, 1$, let $\Sigma_i$ be the spine of $(V_i, L \cap V_i)$ and let $p_i : V_i \setminus \Sigma_i (\cong \partial V_i \times [0, 1)) \to [0, 1)$ be the projection as above. Define maps $\varphi_0 : [0, 1) \to \left(0, \frac{1}{2}\right]$ and $\varphi_1 : [0, 1) \to \left[\frac{1}{2}, 1\right)$ by $\varphi_0(t) = \frac{1}{2}(1-t)$ and $\varphi_1(t) = \frac{1}{2}(1+t)$. A *sweep-out* of $F$ with respect to $L$ is a map $h : M \to [0, 1]$ defined by $h(\Sigma_i) = i$ and $h|_{V_i \setminus \Sigma_i} = \varphi_i \circ p_i$ $(i = 0, 1)$.

## 3. Proof of the main theorem

In this section, we prove Theorem 1.1, and also Proposition 1.2 and Corollary 1.3.

*Proof of Theorem 1.1.* If $d_{BS}(L, F) \leq 1$, then $d_{BS}(L, F) < -\chi(S_L) = |\partial S_L| - 2$ always holds since $|\partial S_L| \geq 4$ by the hypothesis. Hence, we may assume that $d_{BS}(L, F) \geq 2$. Let $H_0$, $H_1$, $\Sigma_0$ and $\Sigma_1$ be as in the previous section, and let $h : M \to [0, 1]$ be a sweep-out of $F$ with respect to $L$. Set $F_L(t) = h^{-1}(t) \cap M_L$. Let $H_0(t)$ be the closure of the component of $M_L \setminus F_L(t)$ that contains $\Sigma_0$, and $H_1(t)$ the closure of $M_L \setminus H_0(t)$. Let $\epsilon_0$ be chosen just larger than the radius of $N(L)$ but small enough so that $S$ meets $H_0(\epsilon_0)$ and $H_1(1-\epsilon_0)$ in c-disks for $F_L(\epsilon_0)$ and $F_L(1-\epsilon_0)$. Then the surface $F_L(t)$ is homeomorphic to $F_L$ for every value $t \in [\epsilon_0, 1-\epsilon_0]$, and we can take a homeomorphism

$$\Phi : \bigcup_{t=\epsilon_0}^{1-\epsilon_0} F_L(t) \to F_L \times [\epsilon_0, 1-\epsilon_0]$$

such that $\Phi(F_L(t)) = F_L \times \{t\}$. Let $\pi = \mathrm{pr}_1 \circ \Phi$, where $\mathrm{pr}_1 : F_L \times [\epsilon_0, 1-\epsilon_0] \to F_L$ is the projection onto the first factor. Hence, for a loop $\gamma$ on $F_L(t)$, the image $\pi(\gamma)$ is a loop on $F_L$.

*Note:* The results referred to throughout this proof are from [Bachman and Schleimer 2005].

We assume that the essential meridional sphere $S$ is in *standard position* as in the proof of the main theorem of that reference. Namely,

- Each boundary component of $S_L$ lies on $\partial F_L(t)$ for some $t \in (\epsilon_0, 1-\epsilon_0)$. If some boundary component of $S$ is contained in $\partial F_L(t)$, we consider $t$ a critical value for $S$.
- All critical points of $h|_{S_L}$ are nondegenerate (i.e., maxima, minima, or saddles). We will refer to any such critical point whose height is between $\epsilon_0$ and $1-\epsilon_0$ and to any meridional boundary component as a *critical submanifold* (of $S$).
- The heights of any two critical submanifolds of $S$ are distinct.

Let $t_0$ be the supremum of $t \in [\epsilon_0, 1-\epsilon_0]$ such that there is a loop in $S \cap F_L(t)$ which bounds a c-disk for $F_L(t)$ in $H_0(t)$. Likewise, let $t_1$ be the infimum of $t \in [\epsilon_0, 1-\epsilon_0]$ such that some loop in $S \cap F_L(t)$ bounds a c-disk for $F_L(t)$ in $H_1(t)$. Since $d_{BS}(L, F) \geq 2$, we may assume that $\epsilon_0 < t_0 < t_1 < 1-\epsilon_0$ by Claims 5.4–5.6.

Choose $\epsilon > 0$ sufficiently small so that there is no critical values in $[t_0-\epsilon, t_0+\epsilon]$ and in $[t_1-\epsilon, t_1+\epsilon]$ other than $t_0$ and $t_1$. By the definition of $t_0$, there is a loop $\gamma_0 \subset S \cap F_L(t_0-\epsilon)$ which bounds a c-disk for $F_L(t_0-\epsilon)$ in $H_0(t_0-\epsilon)$. Similarly, there is a loop $\gamma_1 \subset S \cap F_L(t_1+\epsilon)$ which bounds a c-disk for $F_L(t_1+\epsilon)$ in $H_0(t_1+\epsilon)$.

We see that $S \cap F_L(t_0+\epsilon)$ contains a loop essential on $S_L$. To this end, assume on the contrary that every component of $S \cap F_L(t_0+\epsilon)$ is inessential on $S_L$. By the definition of $t_0$, a component of $S \cap F_L(t_0+\epsilon)$ is inessential also on $F_L(t_0+\epsilon)$ since, otherwise, $S \cap H_0(t_0+\epsilon)$ is a c-disk. Note that there is no essential spheres or decomposing spheres for $L$ by the assumption that $d_{BS}(L, F) \geq 2$ together with Theorem 1. Hence, we can isotope $S$ so that $S_L \subset H_1(t_0+\epsilon)$, which is impossible by Claim 5.2. Similarly, we can see that $S \cap F_L(t_1-\epsilon)$ contains a loop essential on $S_L$. Cut $S_L$ along loops on $S \cap F_L(t_0+\epsilon)$ and $S \cap F_L(t_1-\epsilon)$ which are essential on $S_L$. Let $S'$ be the closure of one of the components which meets both $F_L(t_0+\epsilon)$ and $F_L(t_1-\epsilon)$. Note that every loop on $S_L$ is separating since $S$ is a sphere, and that every component of $S_L \setminus S'$ contains at least two boundary components of $S_L$. Thus, the Euler characteristic $\chi(S')$ is bigger than or equal to $\chi(S_L)+2$.

Let $\alpha_0$ (resp. $\alpha_1$) be a component of $\partial S' \cap F_L(t_0+\epsilon)$ (resp. $\partial S' \cap F_L(t_1-\epsilon)$). By Claim 5.9, every loop of $S \cap F_L(t)$ for every regular value $t \in [t_0, t_1]$ of $h|_S$ is either essential on both $F_L(t)$ and $S_L$ or inessential on both $F_L(t)$ and $S_L$. In particular, the loops $\alpha_0$ and $\alpha_1$ are essential also on $F_L(t_0+\epsilon)$ and $F_L(t_1-\epsilon)$, respectively. Since we chose a sufficiently small $\epsilon$, we may assume that the images $\pi(\gamma_0)$ and $\pi(\alpha_0)$ on $F_L$ are disjoint. Similarly, we assume that $\pi(\gamma_1)$ and $\pi(\alpha_1)$ on $F_L$ are disjoint. By Claim 5.7 and Lemma 5.12, we see that the distance $d_{BS}(\pi(\alpha_0), \pi(\alpha_1))$ is bounded above by the number of *essential* critical submanifolds on $S'$. (Here, an essential critical submanifold is a critical submanifold $P$ of $S'$ such that neither of the boundary components of a small horizontal neighborhood of $P$ in $S'$ does not bound a disk on $S'$. See [Bachman and Schleimer 2005] for detail.) Note that the number of essential critical submanifolds on $S'$ equals $-\chi(S')$.

Hence, we have

$$\begin{aligned} d_{BS}(\pi(\gamma_0), \pi(\gamma_1)) &\leq d_{BS}(\pi(\gamma_0), \pi(\alpha_0)) + d_{BS}(\pi(\alpha_0), \pi(\alpha_1)) + d_{BS}(\pi(\alpha_1), \pi(\gamma_1)) \\ &\leq d_{BS}(\pi(\alpha_0), \pi(\alpha_1)) + 2 \\ &\leq -\chi(S') + 2 \\ &\leq -\chi(S_L). \end{aligned}$$

This completes the proof of Theorem 1.1. □

*Proof of Proposition 1.2.* Let $V_0$ and $V_1$ be the closures of the two components of $S^3 \setminus F$, and let $H_i = \mathrm{Cl}(V_i \setminus N(L))$ $(i = 0, 1)$.

We first assume that $d_{BS}(L, F) = n \geq 1$, and let $c_0, \ldots, c_n$ are essential loops on $F_L$ realizing the distance $d_{BS}(L, F)$. Namely, $c_0$ and $c_n$ bounds c-disk in $H_0$ and $H_1$, respectively, and $c_{i-1} \cap c_i = \varnothing$ for $i = 1, \ldots, n$. Assume that $c_0$ bounds a cut-disk $D^c$ in $H_0$. Since $V_0$ is a 3-ball by the hypothesis and $c_0$ is essential in $F_L$, $H_0 \setminus D^c$ has two components $H_0^1$ and $H_0^2$ neither of which is homeomorphic to a solid torus, and $c_1$ lies on $\partial H_0^1$ and $\partial H_0^2$, say $\partial H_0^1$. Then, we can find a compressing disk $D$ for $F_L$ in $\partial H_0^2$, disjoint from $D^c \cup c_1$, and we replace $c_0$ with $\partial D$. Similarly, in the case where $c_n$ bounds a cut-disk in $H_1$, we can replace $c_n$ with a loop $c_n'$ which bounds a compressing disk in $H_1$ and is disjoint from $c_{n-1}$. Hence, we have $d_T(L, F) = n = d_{BS}(L, F)$.

Assume that $d_{BS}(L, F) = 0$. Then there is a loop $c$ which bounds c-disks in both $H_0$ and $H_1$. By using an argument similar to that for the previous case, we can find loops $c'$ and $c''$ that bound compressing disks in $H_0$ and $H_1$, respectively, and are mutually disjoint. Hence, we have $d_T(L, F) \leq 1$. □

*Proof of Corollary 1.3.* By Proposition 1.2, we have $d_T(L, F) = \max\{1, d_{BS}(L, F)\}$. Since $d_{BS}(L, F) \leq -\chi(S_L)$ by Theorem 1.1 and $-\chi(S_L) \geq 2$ by the hypothesis, we have $d_T(L, F) \leq -\chi(S_L)$. □

## 4. Applications

In this section, we prove Corollaries 1.4 and 1.5.

A (2-string) *trivial tangle* is a pair of a 3-ball and the union of two arcs *trivially embedded* in the 3-ball, that is, the arcs together with some arcs on the boundary of the 3-ball bound disjoint disks. A *rational tangle* is an ambient isotopy class of a trivial tangle with its boundary fixed. It is well known that rational tangles can be parametrized by rational numbers, called the *slopes* of rational tangles. A *Montesinos pair* is a pair of a 3-manifold and a 1-submanifold which is built from the pair, called a *hollow Montesinos pair*, (illustrated in either half of Figure 3) by plugging some of the holes with rational tangles of finite slopes.

An *arborescent link* is a link in the 3-sphere $S^3$ obtained by gluing some Montesinos pairs in their boundaries. In particular, we call a link obtained from a hollow Montesinos pair of the form shown on the left in Figure 3 by plugging the holes with rational tangles of finite slopes $r_1, r_2, \ldots, r_m$ a *Montesinos link*, and denote it by $M_1(r_1, r_2, \ldots, r_m)$. We call $m$ the *length* of the Montesinos link $M_1(r_1, r_2, \ldots, r_m)$ when neither of $r_1, r_2, \ldots, r_m$ is an integer. Similarly, we denote by $M_2(r_1, r_2, \ldots, r_m)$ the arborescent link obtained from a hollow Montesinos pair of the form shown on the right in Figure 3 by plugging the holes with rational tangles of finite slopes $r_1, r_2, \ldots, r_m$.

**Lemma 4.1.** *Let $L$ be an arborescent link in $S^3$ which has bridge index at least* 3, *and suppose that $L$ does not admit an essential Conway sphere* (*i.e.*, *a c-essential sphere in $S^3$ intersecting $L$ transversely in exactly* 4 *points*). *Then $L$ is equivalent to a Montesinos link of length* 3 *as illustrated in Figure 4. In that figure*, *each circle with a rational number $r_i$* ($i = 1, 2, 3$) *inside represents a rational tangle of slope $r_i$.*

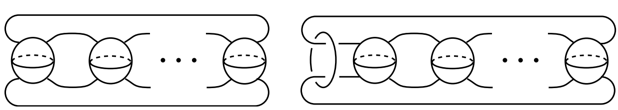


**Figure 3.** Hollow Montesinos pairs.

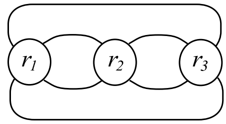



**Figure 4.** A Montesinos link $M_1(r_1, r_2, r_3)$.

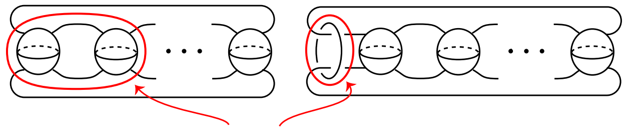



**Figure 5.** Essential Conway spheres in Montesinos pairs.

*Proof.* Let $L$ be an arborescent link in $S^3$ and suppose that $L$ does not admit an essential Conway sphere. Then $L$ is obtained from a Montesinos pair of one of the forms shown in Figure 3 by plugging the holes with rational tangles of finite slopes (see [Bonahon and Siebenmann 2010, Theorem 3.4] or [Jang 2011, Theorem 4]). That is, $L$ is equivalent to a Montesinos link $M_1(r_1, r_2, \ldots, r_{m_1})$ or an arborescent link $M_2(r_1, r_2, \ldots, r_{m_2})$ for some rational numbers $r_i$'s. Moreover, the $m_1$ and $m_2$ cannot be bigger than 3 and 1, respectively, since otherwise $L$ admits an essential Conway sphere as illustrated in Figure 5, which contradicts the hypothesis.

We note that $M_2(r_1)$ is equivalent to the Montesinos link $M_1(-1/2, 1/2, -1/r_1)$. Moreover, we can easily see that $M_1(r_1, r_2, \ldots, r_{m_1})$ admits a 2-bridge presentation if the length of $M_1(r_1, r_2, \ldots, r_{m_1})$ is 1 or 2, which contradicts the assumption that the bridge index of $L$ is at least 3. Thus, $L$ is equivalent to a Montesinos link of length 3, which is the desired result. □

*Proof of Corollary 1.4.* Let $L$ be an arborescent link in $S^3$ and $F$ a bridge sphere of $L$. If there is an essential tori or an essential Conway sphere in the complement of $L$, then the distances $d_{BS}(L, F)$ and $d_T(L, F)$ are at most 2 by [Bachman and Schleimer 2005, Theorem 5.1] together with Theorem 1.1 and Corollary 1.3. Thus, in the rest of the proof, we assume that there is no essential tori or essential Conway spheres. By Lemma 4.1, the link $L$ is equivalent to a Montesinos link of length 3 (see Figure 4).

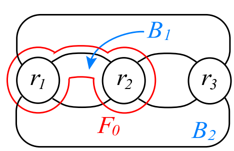



**Figure 6.** A 3-bridge sphere for a Montesinos link.

Assume that $F$ is a minimal bridge sphere (that is, a 3-bridge sphere) of $L$. By [Jang 2013], we may assume that $F$ is (equivalent to) the 3-bridge sphere $F_0$ in Figure 6 without loss of generality. Let $B_1$ be the 3-ball bounded by $F$ containing two of the three rational tangles and $B_2$ the other 3-ball bounded by $F$ (see Figure 6), and let $H_i$ be the closure of $B_i \setminus N(L)$ $(i = 1, 2)$. Let $c_0$, $c_1$ and $c_2$ be the loops on $F_L$ as illustrated in Figure 7. Then $c_0$ bounds a cut-disk in $H_1$, $c_2$ bounds a compressing disk in $H_2$, and $c_1$ is disjoint from $c_0 \cup c_2$. These imply $d_{BS}(L, F) \leq 2$. Moreover, by Proposition 1.2, we have $d_T(L, F) \leq 2$. □

*Proof of Corollary 1.5.* If the distances are greater than two, then, by [Bachman and Schleimer 2005, Theorem 5.1] together with Theorem 1.1 and Corollary 1.3, there is no essential tori in the exterior of $L$, no essential Conway spheres for $L$, no essential spheres nor essential annuli. By [Bachman and Schleimer 2005, Corollary 6.2], $L$ is a hyperbolic link. Moreover, the double branched cover $M_2(L)$ of $S^3$ branched along $L$ has a trivial JSJ decomposition. Thus, $M_2(L)$ is either a Seifert fibered space or a hyperbolic manifold. In the former case, we obtain that either $L$ is a Montesinos link or the complement of $L$ admits a Seifert fibration, which contradicts Corollary 1.4 or the fact that $L$ is hyperbolic, accordingly. Hence, $M_2(L)$ must be hyperbolic. □

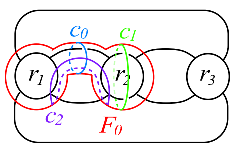



**Figure 7.** Curves realizing distance 2.

## Acknowledgements

The author would like to thank Professor Tsuyoshi Kobayashi and Professor Makoto Sakuma for their helpful comments. She also appreciates the referee for careful reading and valuable suggestions.

## References

[Bachman and Schleimer 2005] D. Bachman and S. Schleimer, "Distance and bridge position", *Pacific J. Math.* **219**:2 (2005), 221–235. MR 2007a:57028 Zbl 1086.57011

[Bonahon and Siebenmann 2010] F. Bonahon and L. C. Siebenmann, "New geometric splittings of classical knots, and the classification and symmetries of arborescent knots", preprint, 2010, http://www-bcf.usc.edu/~fbonahon/Research/Preprints/BonSieb.pdf.

[Evans 2006] T. Evans, "High distance Heegaard splittings of 3-manifolds", *Topology Appl.* **153**:14 (2006), 2631–2647. MR 2007j:57020 Zbl 1107.57011

[Futer and Guéritaud 2009] D. Futer and F. Guéritaud, "Angled decompositions of arborescent link complements", *Proc. Lond. Math. Soc.* (3) **98**:2 (2009), 325–364. MR 2009m:57008 Zbl 1163.57003

[Hartshorn 2002] K. Hartshorn, "Heegaard splittings of Haken manifolds have bounded distance", *Pacific J. Math.* **204**:1 (2002), 61–75. MR 2003a:57037 Zbl 1065.57021

[Hempel 2001] J. Hempel, "3-manifolds as viewed from the curve complex", *Topology* **40**:3 (2001), 631–657. MR 2002f:57044 Zbl 0985.57014

[Ido 2013] A. Ido, "An estimation of Hempl distance for bridge spheres", preprint, 2013.

[Jang 2011] Y. Jang, "Classification of 3-bridge arborescent links", *Hiroshima Math. J.* **41**:1 (2011), 89–136. MR 2012i:57011 Zbl 1231.57006

[Jang 2013] Y. Jang, "Classification of 3-bridge spheres of 3-bridge arborescent links", *J. Math. Soc. Japan* **65**:1 (2013), 97–136. MR 3034400 Zbl 06152429

[Kobayashi and Rieck 2009] T. Kobayashi and Y. Rieck, "Manifolds admitting both strongly irreducible and weakly reducible minimal genus Heegaard splittings", *Comm. Anal. Geom.* **17**:4 (2009), 637–649. MR 2011d:57054 Zbl 1235.57014

[Kojima 1996] S. Kojima, "Nonsingular parts of hyperbolic 3-cone-manifolds", pp. 115–122 in *Topology and Teichmüller spaces* (Katinkulta, 1995), edited by S. Kojima et al., World Scientific, Singapore, 1996. MR 99i:57028 Zbl 0928.57010

[Kojima 1998] S. Kojima, "Deformations of hyperbolic 3-cone-manifolds", *J. Differential Geom.* **49**:3 (1998), 469–516. MR 2000d:57023 Zbl 0990.57004

[Scharlemann and Tomova 2006] M. Scharlemann and M. Tomova, "Alternate Heegaard genus bounds distance", *Geom. Topol.* **10** (2006), 593–617. MR 2007b:57040 Zbl 1128.57022

[Tomova 2007] M. Tomova, "Multiple bridge surfaces restrict knot distance", *Algebr. Geom. Topol.* **7** (2007), 957–1006. MR 2008h:57014 Zbl 1142.57005

Received July 4, 2012. Revised May 30, 2013.

YEONHEE JANG
DEPARTMENT OF MATHEMATICS
NARA WOMEN'S UNIVERSITY
KITAUOYA NISHIMACHI
NARA 630-8526
JAPAN

yeonheejang@cc.nara-wu.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.131

# NORMAL STATES OF TYPE III FACTORS

YASUYUKI KAWAHIGASHI, YOSHIKO OGATA AND ERLING STØRMER

*Dedicated to Masamichi Takesaki on the occasion of his eightieth birthday.*

**Let $M$ be a factor of type III with separable predual and with normal states $\varphi_1, \dots, \varphi_k, \omega$ with $\omega$ faithful. Let $A$ be a finite-dimensional $C^*$-subalgebra of $M$. Then it is shown that there is a unitary operator $u \in M$ such that $\varphi_i \circ \operatorname{Ad} u = \omega$ on $A$ for $i = 1, \dots, k$. This follows from an embedding result of a finite-dimensional $C^*$-algebra with a faithful state into $M$ with finitely many given states. We also give similar embedding results of $C^*$-algebras and von Neumann algebras with faithful states into $M$. Another similar result for a factor of type $\mathrm{II}_1$ instead of type III holds.**

## 1. Introduction

Let $M$ be a factor of type III with separable predual. Then two nonzero projections $e$ and $f$ in $M$ are equivalent, that is, there exists a partial isometry $v \in M$ such that $v^*v = e$, $vv^* = f$. If, furthermore, $e$ and $f$ are different from the identity operator 1, then there is a unitary operator $u \in M$ such that $u^*eu = f$. This shows that there is an abundance of unitaries in $M$, so one might expect stronger results arising from these unitaries. That is what is done in the present paper. We show that if $\varphi$ and $\omega$ are faithful normal states in $M$ and $A \subset M$ is a finite-dimensional $C^*$-algebra, then there exists a unitary operator $u \in M$ such that the restrictions $\varphi \circ \operatorname{Ad} u|_A$ and $\omega|_A$ are equal, where $\operatorname{Ad} u$ is the inner automorphism $x \mapsto u^*xu$ of $M$. (See Corollary 2.2 for a more precise and general statement.)

This actually follows from an embedding result of a finite-dimensional $C^*$-algebra $A$ with a faithful state into $M$ with finitely given normal states. This result is then applied to obtain a similar result for the $C^*$-algebra of the compact operators on a separable Hilbert space. Furthermore, we have more general embedding results in Section 3 for $C^*$-algebras and von Neumann algebras with faithful states into a

---

Kawahigashi is supported in part by the Global COE Program "The research and training center for new development in mathematics", the Mitsubishi Foundation Research Grants and the Grants-in-Aid for Scientific Research, JSPS. Ogata is supported in part by the Inoue Foundation and the Grants-in-Aid for Scientific Research, JSPS.
*MSC2010:* 46L30.
*Keywords:* von Neumann algebra, type III factor, normal state.

type III factor $M$ such that a finite number of normal states on $M$ coincide after the embedding.

If $M$ is not of type III, the corresponding result is false in general, but if $M$ is a factor of $\mathrm{II}_1$, $\omega = \tau$ is the trace and $A \cong M_n(\mathbb{C})$, the matrix algebra of complex $n \times n$-matrices, then the corresponding result to the unitary equivalence on $A$ holds for $\omega = \tau$ and any $\varphi$. This will be shown in Section 4.

There exist results of a similar nature to the ones above in the literature. In [Connes and Størmer 1978], it has been shown that if $M$ is of type $\mathrm{III}_1$ and $\varepsilon > 0$ then there is a unitary operator $u \in M$ with

$$\|\varphi \circ \operatorname{Ad} u - \omega\| < \varepsilon.$$

If one takes a pointwise weak limit point of the automorphisms of the form $\operatorname{Ad} u$ in the above, then one finds a completely positive unital map $\pi : M \to M$ with $\varphi \circ \pi = \omega$.

In the nonseparable case, it has recently been shown by Ando and Haagerup [2013] that for some factors of type $\mathrm{III}_1$ constructed as ultraproducts, all faithful normal states are unitarily equivalent.

In the $C^*$-algebra case it has been shown in [Kishimoto et al. 2003] that if $\varphi$ and $\omega$ are pure states of a separable $C^*$-algebra $A$ with the same kernel for their GNS-representations, then there is an asymptotically inner automorphism $\alpha$ of $A$ such that $\varphi \circ \alpha = \omega$.

Our result gives an exact equality for two states, not an approximate one, but only on a finite-dimensional $C^*$-subalgebra $A$.

## 2. Factors of type III

In this section we state and prove our main result.

**Theorem 2.1.** *Let $M$ be a type III factor with separable predual and $\varphi_1, \dots, \varphi_k$ normal states on $M$. Let $A$ be a finite-dimensional $C^*$-algebra and $\rho$ a faithful state on $A$. Then there exists a unital injective homomorphism $\pi : A \to M$ with*

$$\varphi_i \circ \pi = \rho, \quad i = 1, \dots, k.$$

After proving this theorem, we will prove that it implies the following corollary.

**Corollary 2.2.** *Let $M$ be a factor of type III with separable predual. Let $A$ be a finite-dimensional $C^*$-subalgebra of $M$. Let $\varphi_1, \dots, \varphi_k$ and $\omega$ be normal states on $M$ and assume that $\omega$ is faithful. Then there exists a unitary operator $u \in M$ such that*

$$\varphi_i \circ \operatorname{Ad} u|_A = \omega|_A, \quad i = 1, \dots, k.$$

Before starting preliminaries of our proof of Theorem 2.1, we give an outline of our method for the case $A \cong M_d(\mathbb{C})$.

After diagonalizing the density matrix of $\rho$, what we have to find is a system of matrix units $\{e_{ij}\}$ in $M$ for which we have $\varphi_n(e_{ij}) = \delta_{ij}\lambda_i$ for all $n = 1, \dots, k$ and $i, j = 1, \dots, d$, where the $\lambda_i$ are eigenvalues of the density matrix of $\rho$. We first choose $e_{ii}$ satisfying this condition. Then we choose $e_{12}, e_{13}, \dots, e_{1d}$ inductively so that we have various identities saying that the values of certain linear functionals applied to a certain partial isometry are all zero at each induction step. This is done by a version of a noncommutative Lyapunov theorem, and what we need is a special case of [Akemann and Anderson 1991, Theorem 2.5(1)]. Since the statement and its proof are short, we include them here in the form we need, for the sake of convenience of the reader.

**Lemma 2.3.** *Let $M$ be a nonatomic von Neumann algebra and $\Phi : M \to \mathbb{C}^n$ a $\sigma$-weakly continuous linear map. Then for any $a \in M_{+,1}$, there exists a projection $p \in M$ such that $\Phi(p) = \Phi(a)$.*

*Proof.* Let

$$D := \{x \in M_{+,1} \mid \Phi(x) = \Phi(a)\},$$

where $M_{+,1}$ denotes the positive operators in the unit ball of $M$. Then $D$ is a nonempty $\sigma$-weakly compact convex set. Therefore, by the Krein–Milman theorem, there exists an extremal point $b$ of $D$. We show $b$ is a projection. If $b$ were not a projection, then there exists $\delta \in \left(0, \frac{1}{2}\right)$ such that the spectral projection $p$ of $b$ corresponding to $(\delta, 1-\delta)$ is nonzero. By the assumption on $M$, $pM_{\mathrm{sa}}p$ is an infinite-dimensional real linear space while its range with respect to $\Phi$ is finite-dimensional. This implies the existence of a nonzero $y \in pM_{\mathrm{sa}}p$ such that $\Phi(y) = 0$. Setting $t := \delta/\|y\|$, we have $b \pm ty \in D$. As we have $b = (b+ty)/2 + (b-ty)/2$, this contradicts the fact that $b$ is extremal in $D$. □

We now construct appropriate matrix units by induction on the size of matrix units.

**Lemma 2.4.** *Let $M$ be a type III factor with separable predual and $\varphi_1, \dots, \varphi_n$ normal states on $M$. Let $\lambda_i > 0$, $i = 1, \dots, m$ with $\sum_i \lambda_i = 1$. Then there exists a system of matrix units $\{e_{ij}\}_{i,j=1,\dots,m}$ such that*

$$\varphi_l(e_{ij}) = \delta_{ij}\lambda_i \quad \textit{for all } l = 1, \dots, m.$$

*Proof.* For a projection $p \in M$ satisfying $0 \le \varphi_l(p) = \lambda < 1$ for $l = 1, \dots, n$ and $0 \le t \le 1-\lambda$, there exists a projection $q$ orthogonal to $p$ such that $\varphi_l(q) = t$. To see this, we consider a $\sigma$-weakly continuous liner map $\Phi : M_{\bar{p}} \to \mathbb{C}^n$, where we write $\bar{p} = 1 - p$, given by $\Phi(x) = (\varphi_l(x))_{l=1}^n$, and apply Lemma 2.3 for $a = t\bar{p}/(1-\lambda)$.

Using this fact inductively, we have $\{e_{ii}\}$.

We next define partial isometries $u_{i1}$, $i = 1, \ldots, m$, inductively such that $e_{ij} = u_{i1}u_{j1}^*$ satisfy the conditions of the lemma. Let $u_{11} = e_{11}$ and assume that we have found $u_{i1}$, $i = 1, \ldots, k$ with $k < m$. Let $v$ be a partial isometry in $M$ with $v^*v = e_{11}$, $vv^* = e_{k+1,k+1}$. Then define a map

$$\Phi : e_{11}Me_{11} \to \mathbb{C}^{nk}$$
$$\Phi(x) := (\varphi_l(vxu_{j1}^*))_{l=1,\ldots,n,\, j=1,\ldots,k}.$$

This map $\Phi$ is $\sigma$-weakly continuous and linear, so by using Lemma 2.3 with $a = e_{11}/2$, we obtain a projection $p \in e_{11}Me_{11}$ such that $\Phi(p) = \Phi(e_{11})/2$. Define

$$u_{k+1,1} := vp - v(1-p).$$

Since $p \leq e_{11}$, an easy computation shows that $u_{k+1,1}^* u_{k+1,1} = e_{11}$, $u_{k+1,1}u_{k+1,1}^* = e_{k+1,k+1}$. Let $e_{k+1,j} = u_{k+1,1}u_{j1}^*$ and $e_{j,k+1} = u_{j1}u_{k+1,1}^*$. Then the $e_{ij}$, $i, j \leq k+1$, form a set of matrix units, and using the definition of $\Phi$ and that $\Phi(p) = \Phi(e_{11})/2$, we get for all $l$

$$\begin{aligned} \varphi_l(u_{k+1,1}u_{j1}^*) &= \varphi_l\big((2vp - v)u_{j1}^*\big) \\ &= 2\varphi_l(vpu_{j1}^*) - \varphi_l(vu_{j1}^*) \\ &= 0. \end{aligned}$$

Thus

$$\varphi_l(e_{j,k+1}) = \varphi_l(u_{j1}u_{k+1,1}^*) = \overline{\varphi_l(u_{k+1,1}u_{j1}^*)} = 0,$$

completing the proof of the lemma. □

*Proof of Theorem 2.1.* First we consider the case $A = M_m(\mathbb{C})$. We choose a system of matrix units $\{v_{ij}\}_{i,j=1,\ldots,m}$ of $A = M_m(\mathbb{C})$ which diagonalizes the density matrix $D_\rho$ of $\rho$, that is, $D_\rho = \sum_{i=1}^m \lambda_i v_{ii}$. As $\rho$ is faithful, we have $\lambda_i > 0$ for all $i$. By Lemma 2.4, we obtain a system of matrix units $\{e_{ij}\}_{i,j=1,\ldots,m}$ in $M$ satisfying

$$\varphi_n(e_{ij}) = \delta_{ij}\lambda_i, \quad n = 1, \ldots, k, \quad i, j = 1, \ldots, m. \tag{1}$$

Define

$$\pi : M_m(\mathbb{C}) \to M, \quad \pi(v_{ij}) = e_{ij}.$$

Then $\pi$ gives a unital homomorphism satisfying the desired condition.

For the general case $A \simeq \bigoplus_{k=1}^b M_{n_k}(\mathbb{C})$, let $m = \sum_{k=1}^b n_k$. Let $\hat{\rho}$ be a faithful extension of $\rho$ to $M_m(\mathbb{C})$. Applying the above result to $M_m(\mathbb{C})$ and $\hat{\rho}$, there exists a unital homomorphism $\hat{\pi} : M_m(\mathbb{C}) \to M$ such that

$$\varphi_n \circ \hat{\pi} = \hat{\rho}, \quad n = 1, \ldots, k.$$

The restriction $\pi := \hat{\pi}|_A$ gives a unital homomorphism from $A$ to $M$ satisfying $\varphi_n \circ \pi = \rho$, for $n = 1, \ldots, k$. □

*Proof of Corollary 2.2.* Let $p$ be the unit of $A$. Considering $A \oplus \mathbb{C}(1-p)$ instead of $A$, we may assume that $A$ contains the unit of $M$ from the beginning.

First we consider the case $A \simeq M_m(\mathbb{C})$, $m \in \mathbb{N}$. Let $\{f_{ij}\}_{i,j=1,\dots,m}$, $\{v_{ij}\}_{i,j=1,\dots,m}$ be systems of matrix units of $A$ and $M_m(\mathbb{C})$, respectively. Let $\gamma : M_m(\mathbb{C}) \to A$ be an isomorphism given by $\gamma(v_{ij}) = f_{ij}$.

Then $\rho := \omega \circ \gamma$ is a faithful state on $M_m(\mathbb{C})$. From Theorem 2.1, there exists a unital homomorphism $\pi : M_m(\mathbb{C}) \to M$ such that $\varphi_n \circ \pi = \rho$, $n = 1, \dots, k$. The algebras $A$ and $\pi(M_m(\mathbb{C}))$ are subalgebras of $M$ isomorphic to $M_m(\mathbb{C})$ with complete sets of matrix units $\{f_{ij}\}$ and $\{\pi(v_{ij})\}$. As in [Haagerup and Musat 2011, Lemma 2.1], if $v \in M$ is a partial isometry with $v^*v = \pi(v_{11})$ and $vv^* = f_{11}$, then $u := \sum_{i=1}^m \pi(v_{i1})v^* f_{1i}$ is a unitary in $M$ satisfying $uf_{ij}u^* = \pi(v_{ij})$. Hence we have

$$\varphi_n \circ \operatorname{Ad} u(f_{ij}) = \varphi_n(\pi(v_{ij})) = \rho(v_{ij}) = \omega \circ \gamma(v_{ij}) = \omega(f_{ij}),$$

that is, $\varphi_n \circ \operatorname{Ad} u|_A = \omega|_A$ for $n = 1, \dots, k$.

For the general case $A \simeq \bigoplus_{l=1}^b M_{n_l}(\mathbb{C})$, let $\{f_{ij}^{(l)}\}_{ij=1,\dots,n_l}$ be a system of matrix units of $M_{n_l}(\mathbb{C})$ for each $l = 1, \dots, b$. As $M$ is of type III, for all $l = 1, \dots, b$, the nonzero projections $f_{11}^{(1)}$ and $f_{11}^{(l)}$ are mutually equivalent. Hence, there exist partial isometries $v^{(l)} \in M$ such that $v^{(l)*}v^{(l)} = f_{11}^{(l)}$ and $v^{(l)}v^{(l)*} = f_{11}^{(1)}$. Set $w_{(k,i)(l,j)} := f_{i1}^{(k)} v^{(k)*} v^{(l)} f_{1j}^{(l)}$, for $k, l = 1, \dots, b$, $i = 1, \dots, n_k$, and $j = 1, \dots, n_l$. Then we have

$$\begin{aligned}
w_{(k,i)(l,j)}^* &= f_{j1}^{(l)} v^{(l)*} v^{(k)} f_{1i}^{(k)} = w_{(l,j)(k,i)}, \\
w_{(k,i)(l,j)} w_{(l',j')(k',i')} &= f_{i1}^{(k)} v^{(k)*} v^{(l)} f_{1j}^{(l)} f_{j'1}^{(l')} v^{(l')*} v^{(k')} f_{1i'}^{(k')} \\
&= \delta_{ll'} \delta_{jj'} f_{i1}^{(k)} v^{(k)*} v^{(l)} f_{11}^{(l)} v^{(l)*} v^{(k')} f_{1i'}^{(k')} \\
&= \delta_{ll'} \delta_{jj'} f_{i1}^{(k)} v^{(k)*} v^{(l)} v^{(l)*} v^{(k')} f_{1i'}^{(k')} \\
&= \delta_{ll'} \delta_{jj'} w_{(ki),(k'i')}, \\
\sum_{(k,i)} w_{(k,i)(k,i)} &= \sum_{i,k} f_{i1}^{(k)} v^{(k)*} v^{(k)} f_{1i}^{(k)} = \sum_{(k,i)} f_{ii}^{(k)} = 1.
\end{aligned}$$

Hence $\{w_{(k,i)(l,j)}\}_{(k,i),(l,j)}$ give a system of matrix units of a $C^*$-subalgebra $B$ of $M$ isomorphic to $M_m$, for $m := \sum_{k=1}^b n_k$. As $w_{(ki)(kj)} = f_{i1}^{(k)} f_{1j}^{(k)} = f_{ij}^{(k)}$, $\{w_{(k,i)(l,j)}\}$ is an extension of $\{f_{ij}^{(k)}\}$ and $A$ is a subalgebra of $B$. We apply the above argument to $B \simeq M_m(\mathbb{C})$ and obtain a unitary $u$ in $M$ such that $\varphi_i \circ \operatorname{Ad} u|_B = \omega|_B$. In particular, we obtain $\varphi_i \circ \operatorname{Ad} u|_A = \omega|_A$ for $i = 1, \dots, k$. □

## 3. Embedding of operator algebras with faithful states

The above theorem can be extended to the algebra of the compact operators as follows.

**Theorem 3.1.** *Let $K(\mathcal{H})$ denote the set of all the compact operators on a separable Hilbert space $\mathcal{H}$. Let $\rho$ be a faithful state on $K(\mathcal{H})$. Let $M$ be a factor of type III with separable predual, $\varphi_1, \varphi_2, \dots, \varphi_k$ normal states on $M$. Then there exists a homomorphism $\pi$ of $K(\mathcal{H})$ into $M$ such that*

$$\varphi_n \circ \pi = \rho, \quad n = 1, \dots, k.$$

*Proof.* We may assume that $\mathcal{H}$ is infinite-dimensional and $\varphi_1$ is faithful — for example, by adding a faithful state to the set of all the $\varphi_i$.

Let $\{v_{ij}\}$ be a system of matrix units of $K(\mathcal{H})$ diagonalizing the density matrix $D_\rho$ of $\rho$, that is, $D_\rho = \sum_{i=1}^{\infty} \lambda_i v_{ii}$. As $\rho$ is faithful, we have $\lambda_i > 0$ for all $i$.

We claim that there exists a system of matrix units $\{e_{ij}\}_{i,j\in\mathbb{N}}$ in $M$ satisfying

$$\varphi_n(e_{ij}) = \delta_{ij}\lambda_i, \quad n = 1, \dots, k, \quad i, j = 1, 2, \dots. \tag{2}$$

This is proved in the same way as in the proof of Theorem 2.1. □

A slight rewriting of the above theorem gives the following:

**Corollary 3.2.** *Let $B(\mathcal{H})$ be the set of all the bounded operators on a separable Hilbert space $\mathcal{H}$ and $\rho$ a faithful normal state on $B(\mathcal{H})$. Let $M$ be a factor of type III with separable predual and $\varphi_1, \varphi_2, \dots, \varphi_k$ normal states on $M$. Then there exists a homomorphism $\pi$ of $B(\mathcal{H})$ into $M$ such that*

$$\varphi_n \circ \pi = \rho, \quad n = 1, \dots, k.$$

We now consider an embedding of a $C^*$-algebra with a faithful state into a type III factor with finitely many normal states.

**Theorem 3.3.** *For a $C^*$-algebra $A$ and a faithful state $\omega$ on $A$, the following conditions are equivalent*:

(i) *The Hilbert space $\mathcal{H}_\omega$ in the GNS triple $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ of $\omega$ is separable and $\Omega_\omega$ is separating for $\pi_\omega(A)''$.*

(ii) *There exists a representation $(\mathcal{H}, \rho)$ of $A$ on a separable Hilbert space $\mathcal{H}$ and a faithful normal state $\sigma$ on $B(\mathcal{H})$ with $\omega = \sigma \circ \rho$.*

(iii) *For any factor $M$ of type III with separable predual and its normal states $\varphi_1, \dots, \varphi_n$, there exists an injective homomorphism $\gamma : A \to M$ with $\varphi_j \circ \gamma = \omega$ for all $j = 1, \dots, n$.*

*Proof.* Suppose condition (i) holds. Then $\Omega_\omega$ is cyclic for $\pi_\omega(A)'$. Therefore, using the separability of $\mathcal{H}_\omega$, we have a sequence $\{x_n\}_{n=1}^{\infty} \subset (\pi_\omega(A)')_1$ such that $\{x_n\Omega_\omega : n \in \mathbb{N}\}$ spans $\mathcal{H}_\omega$. Let $x_0 := \sqrt{1 - \sum_n x_n^* x_n / 2^n}$, and define a state $\sigma$ on $B(\mathcal{H}_\omega)$ given by the density matrix $\sum_{n=0}^{\infty} |x_n\Omega_\omega\rangle\langle x_n\Omega_\omega| / 2^n$. This $\sigma$ is faithful and normal. Let $\rho = \pi_\omega$. We can check $\sigma \circ \rho = \omega$. Hence (ii) holds.

Now suppose condition (ii) holds. We show (iii). By Theorem 3.1, we have an injective homomorphism $\pi : K(\mathcal{H}) \to M$ such that $\sigma|_{K(\mathcal{H})} = \varphi \circ \pi$. We denote the extension of $\pi$ to $B(\mathcal{H})$ by $\hat{\pi}$. Then from the way we have constructed $\pi$, we obtain $\sigma = \varphi \circ \hat{\pi}$. Define $\gamma := \hat{\pi} \circ \rho : A \to M$. Then we obtain $\varphi \circ \gamma = \varphi \circ \hat{\pi} \circ \rho = \sigma \circ \rho = \omega$.

Finally suppose condition (iii) holds, and we show this implies (i). To see this, fix a factor $M$ of type III with a faithful normal state $\varphi$, and let $(\mathcal{H}_\varphi, \pi_\varphi, \Omega_\varphi)$ be its GNS triple. We obtain $\gamma$ as in (iii). Let $K := \overline{\pi_\varphi \circ \gamma(A)\Omega_\varphi}$ and let $\beta$ be the restriction of $\pi_\varphi \circ \gamma$ to $K$. Then $(K, \beta, \Omega_\varphi)$ is the GNS triple of $\omega$. As $\Omega_\varphi$ is separating for $\pi_\varphi(M)$, it is separating for $\beta(A)''$, and (i) holds. □

As an immediate corollary, we obtain the following:

**Corollary 3.4.** *Let $N$ be a von Neumann algebra with separable predual and $\psi$ a faithful normal state on $N$. Then for any factor $M$ of type III with separable predual and a normal state $\varphi$ on $M$, there exists an injective homomorphism $\pi : N \to M$ with $\varphi \circ \pi = \psi$.*

Another easy corollary is as follows, by a well-known result on the KMS condition [Bratteli and Robinson 1997, Corollary 5.3.9].

**Corollary 3.5.** *Suppose that we have a $C^*$-algebra $A$, a state $\varphi$ on $A$, and a one-parameter automorphism group $\{\alpha_t\}_{t\in\mathbb{R}}$ such that these satisfy the KMS condition. Then the pair $(A, \varphi)$ satisfies the (equivalent) conditions in Theorem 3.3.*

**Remark 3.6.** Note that a general faithful state on a $C^*$-algebra $A$ does not satisfy condition (i) of Theorem 3.3 at all, as shown in [Takesaki 1974] by an example due to Pedersen. The $C^*$-algebra used by Takesaki is a very basic one, $C([0, 1]) \otimes M_2(\mathbb{C})$. A slight modification of the argument there also works for a simple $C^*$-algebra $A_\theta \otimes M_2(\mathbb{C})$, where $A_\theta$ is the irrational rotation $C^*$-algebra.

In Theorem 3 of the same paper, Takesaki gives a sufficient condition for our condition (i) in Theorem 3.3 and calls it the quasi-KMS condition, but it seems difficult to check this condition for a given example.

**Remark 3.7.** In all the above cases, we considered embeddings into a type III factor, but actually any properly infinite von Neumann algebra with separable predual works. This is because if we have a properly infinite von Neumann algebra and normal states on it, we simply restrict the states on a type III factor which is found as a subalgebra of the original von Neumann algebra. It is easy to see that if a von Neumann algebra with separable predual has a finite direct summand, this type of embedding is impossible, so actually this embeddability characterize proper infiniteness of a von Neumann algebra with separable predual.

## 4. Factors of type $\mathrm{II}_1$

The direct analogue of Theorem 2.1 for finite factors is trivially false. For example, if $M$ is of type $\mathrm{II}_1$ with trace $\tau$ and $\rho$ is not a trace on $A$, then the conclusion of Theorem 2.1 for $\varphi_1 = \tau$ is clearly false. However, if we restrict the choice of $\omega$ in Corollary 2.2, we obtain a positive result.

**Theorem 4.1.** *Let $\varphi_1, \dots, \varphi_k$ be normal states on a factor $M$ of type $II_1$ with the unique trace $\tau$. Let $A$ be a $C^*$-subalgebra of $M$ isomorphic to $M_m(\mathbb{C})$ with $1 \in A$. Then there exists a unitary operator $u \in M$ satisfying $\varphi_i \circ \mathrm{Ad}\, u|_A = \tau|_A$ for $i = 1, \dots, k$.*

*Proof.* We may assume that $\varphi_1 = \tau$ is the unique trace on $M$. We proceed as in the proof of Theorem 2.1. The only difference is that we take $\tau(e_{ii}) = 1/m$ instead of the proof of Lemma 2.4. □

## Acknowledgements

Y.K. thanks the CMTP in Rome and Université Paris VII, and Y.O. thanks University of Oregon and University of Oslo for hospitality during their stays when parts of this work were done. Y.K. is grateful to R. Longo and G. Skandalis for useful remarks, Y.O. to N. Christopher Phillips for helpful discussions and E.S. to S. Neshveyev.

## References

[Akemann and Anderson 1991] C. A. Akemann and J. Anderson, *Lyapunov theorems for operator algebras*, Mem. Amer. Math. Soc. **458**, Amer. Math. Soc., Providence, RI, 1991. MR 92e:46113 Zbl 0769.46036

[Ando and Haagerup 2013] H. Ando and U. Haagerup, "Ultraproducts of von Neumann algebras", preprint, 2013. arXiv 1212.5457

[Bratteli and Robinson 1997] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics, II: Equilibrium states, models in quantum statistical mechanics*, 2nd ed., Springer, Berlin, 1997. MR 98e:82004 Zbl 0903.46066

[Connes and Størmer 1978] A. Connes and E. Størmer, "Homogeneity of the state space of factors of type $\mathrm{III}_1$", *J. Functional Analysis* **28**:2 (1978), 187–196. MR 57 #10435 Zbl 0408.46048

[Haagerup and Musat 2011] U. Haagerup and M. Musat, "Factorization and dilation problems for completely positive maps on von Neumann algebras", *Comm. Math. Phys.* **303**:2 (2011), 555–594. MR 2012e:46125 Zbl 1220.46044

[Kishimoto et al. 2003] A. Kishimoto, N. Ozawa, and S. Sakai, "Homogeneity of the pure state space of a separable $C^*$–algebra", *Canad. Math. Bull.* **46**:3 (2003), 365–372. MR 2004f:46072 Zbl 1066.46052

[Takesaki 1974] M. Takesaki, "Faithful states on a $C^*$–algebra", *Pacific J. Math.* **52** (1974), 605–610. MR 50 #8096 Zbl 0295.46092

Received February 27, 2013. Revised May 30, 2013.

YASUYUKI KAWAHIGASHI
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
THE UNIVERSITY OF TOKYO
KOMABA, TOKYO, 153-8914
JAPAN

and

KAVLI IPMU (WPI)
THE UNIVERSITY OF TOKYO
5-1-5 KASHIWANOHA, KASHIWA, 277-8583
JAPAN

yasuyuki@ms.u-tokyo.ac.jp

YOSHIKO OGATA
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
THE UNIVERSITY OF TOKYO
KOMABA, TOKYO, 153-8914
JAPAN

yoshiko@ms.u-tokyo.ac.jp

ERLING STØRMER
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OSLO
P.O. BOX 1053
BLINDERN, NO-0316 OSLO
NORWAY

erlings@math.uio.no

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.141

# EIGENVALUES AND ENTROPIES UNDER THE HARMONIC-RICCI FLOW

YI LI

**In this paper, the author discusses the eigenvalues and entropies under the harmonic-Ricci flow, which is the Ricci flow coupled with the harmonic map flow. We give an alternative proof of results for compact steady and expanding harmonic-Ricci breathers. In the second part, we derive some monotonicity formulas for eigenvalues of the Laplacian under the harmonic-Ricci flow. Finally, we obtain the first variation of the shrinker and expanding entropies of the harmonic-Ricci flow.**



## 1. Introduction

Since the successful application of the Ricci flow to topological and geometric problems, several analogous flows have been studied, including the harmonic-Ricci flow [List 2006; Müller 2012], connection Ricci flow [Streets 2008], Ricci–Yang–Mills flow [Streets 2007; 2010; Young 2008], and renormalization group flows [He et al. 2008; Li 2012; Oliynyk et al. 2006; Streets 2009]. In this article, we study the

*MSC2010:* 53C44, 35K55.
*Keywords:* Eigenvalue, entropies, harmonic-Ricci flow, harmonic-Ricci breathers.

eigenvalue problems of the harmonic-Ricci flow, which is the following coupled system:

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + 4\,du(t)\otimes du(t), \tag{1-1}$$

$$\frac{\partial}{\partial t} u(t) = \Delta_{g(t)} u(t). \tag{1-2}$$

For convenience, we introduce a new symmetric 2-tensor $\mathcal{S}_{g(t),u(t)}$ whose components $S_{ij}$ are defined by

$$S_{ij} := R_{ij} - 2\partial_i u \partial_j u.$$

Its trace is $S_{g(t),u(t)} := g^{ij} S_{ij} = R_{g(t)} - 2|\nabla_{g(t)} u(t)|^2_{g(t)}$.

Suppose that $M$ is a compact Riemannian manifold. For any Riemannian metric $g$ and any smooth functions $u$, $f$, we have a number of functionals:

$$\begin{aligned}
\mathcal{F}(g,u,f) &= \int_M (R_g + |\nabla_g f|^2_g - 2|\nabla_g u|^2_g) e^{-f}\, dV_g,\\
\mathcal{E}(g,u,f) &= \int_M (R_g - 2|\nabla_g u|^2_g) e^{-f}\, dV_g,\\
\mathcal{F}_k(g,u,f) &= \int_M (kR_g + |\nabla_g f|^2_g - 2k|\nabla_g u|^2_g) e^{-f}\, dV_g.
\end{aligned}$$

List [2006] and Müller [2012] showed that, as in the case of Perelman's $\mathcal{F}$-functional, under the evolution equation

$$\begin{aligned}
&\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + 4\,du(t)\otimes du(t),\\
&\frac{\partial}{\partial t} u(t) = \Delta_{g(t)} u(t),\\
&\frac{\partial}{\partial t} f(t) = -\Delta_{g(t)} f(t) - R_{g(t)} + |\nabla_{g(t)} f(t)|^2_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)},
\end{aligned} \tag{1-3}$$

the evolution equation for the $\mathcal{F}$-functional is

$$\begin{aligned}
\frac{d}{dt}\mathcal{F}(g(t),u(t),f(t)) &= 2\int_M |\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)}\\
&\quad + 4\int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)},
\end{aligned} \tag{1-4}$$

which is nonnegative. Based on (1-4), we derive the following.

**Theorem 1.1.** *Under the evolution equation* (1-3), *one has*

$$\begin{aligned}
&\frac{d}{dt}\mathcal{E}(g(t),u(t),f(t))\\
&\qquad = 2\int_M |\mathcal{S}_{g(t),u(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} + 4\int_M |\Delta_{g(t)} u(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)},
\end{aligned} \tag{1-5}$$

*and*

$$\begin{aligned}
(1\text{-}6)\quad & \frac{d}{dt}\mathcal{F}_k(g(t), u(t), f(t)) \\
&= 2(k-1)\int_M |\mathcal{S}_{g(t),u(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} + 2\int_M |\mathcal{S}_{g(t),u(t)} \\
&\quad + \nabla^2_{g(t)} f(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} + 4(k-1)\int_M |\Delta_{g(t)} u(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 4\int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)}.
\end{aligned}$$

As a corollary we give a new proof of the following result.

**Corollary 1.2.** *There is no compact steady harmonic-Ricci breather unless the manifold $(M, g(t))$ is Ricci-flat and $u(t)$ is a constant.*

To deal with the expanding harmonic-Ricci breather, we need the functionals

$$\mathcal{L}_+(g, u, \tau, f) = \tau^2 \int_M \Big(R_g + \frac{n}{2\tau} + \Delta_g f - 2|\nabla_g u|^2_g\Big) e^{-f}\, dV_g,$$

$$\mathcal{L}_{+,k}(g, u, \tau, f) = \tau^2 \int_M \Big(k\Big(R_g + \frac{n}{2\tau}\Big) + \Delta_g f - 2k|\nabla_g u|^2_g\Big) e^{-f}\, dV_g.$$

Under the evolution equation

$$\begin{aligned}
&\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + 4\, du(t)\otimes du(t), \\
&\frac{\partial}{\partial t} u(t) = \Delta_{g(t)} u(t), \\
&\frac{\partial}{\partial t} f(t) = -\Delta_{g(t)} f(t) + |\nabla_{g(t)} f(t)|^2_{g(t)} - R_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)}, \\
&\frac{d}{dt}\tau(t) = 1,
\end{aligned}$$

we have:

**Theorem 1.3.** *Under the evolution equation, one has*

$$\begin{aligned}
(1\text{-}7)\quad & \frac{d}{dt}\mathcal{L}_+(g(t), u(t), \tau(t), f(t)) \\
&= 2\tau(t)^2 \int_M \Big|\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) + \frac{1}{2\tau(t)} g(t)\Big|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 4\tau(t)^2 \int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)}
\end{aligned}$$

*and*

$$\begin{aligned}(1\text{-}8)\quad &\frac{d}{dt}\mathscr{L}_{+,k}(g(t),u(t),\tau(t),f(t))\\ &= 2\tau(t)^2\int_M |\mathcal{S}_{g(t),u(t)}+\nabla^2_{g(t)}f(t)+\frac{1}{2\tau(t)}g(t)|^2_{g(t)}e^{-f(t)}dV_{g(t)}\\ &\quad+2(k-1)\tau(t)^2\int_M |\mathcal{S}_{g(t),u(t)}+\frac{1}{2\tau(t)}g(t)|^2_{g(t)}e^{-f(t)}dV_{g(t)}\\ &\quad+4\tau(t)^2\int_M |\Delta_{g(t)}u(t)-\langle du(t),df(t)\rangle_{g(t)}|^2_{g(t)}e^{-f(t)}dV_{g(t)}\\ &\quad+4(k-1)\tau(t)^2\int_M |\Delta_{g(t)}u(t)|^2_{g(t)}e^{-f(t)}dV_{g(t)}.\end{aligned}$$

As a corollary, we obtain a new proof of the following.

**Corollary 1.4.** *There is no expanding harmonic-Ricci breather on compact Riemannian manifolds unless the manifold $M$ is an Einstein manifold and $u(t)$ a constant.*

The second part of this paper focuses on the eigenvalue of the Laplacian operator under the harmonic-Ricci flow.

**Theorem 1.5.** *If $(g(t),u(t))$ is a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and $\lambda(t)$ denotes the eigenvalue of the Laplacian $\Delta_{g(t)}$ with eigenfunction $f(t)$,*

$$\begin{aligned}(1\text{-}9)\quad &\frac{d}{dt}\lambda(t)\cdot\int_M f(t)^2\,dV_{g(t)}\\ &=\lambda(t)\int_M S_{g(t),u(t)}f(t)^2\,dV_{g(t)}-\int_M S_{g(t),u(t)}|\nabla_{g(t)}f|^2_{g(t)}\,dV_{g(t)}\\ &\quad+2\int_M\langle\mathcal{S}_{g(t),u(t)},df(t)\otimes df(t)\rangle_{g(t)}\,dV_{g(t)}.\end{aligned}$$

Equation (1-9) is a general formula to describe the evolution of $\lambda(t)$ under the harmonic-Ricci flow. Under a curvature assumption, we can derive some monotonicity formulas for the eigenvalue $\lambda(t)$. Set

$$S_{\min}(0) := \min_{x\in M} S_{g(0),u(0)}(x), \qquad (1\text{-}10)$$

the minimum of $S_{g(t),u(t)}$ over $M$ at the time 0.

**Theorem 1.6.** *Let $(g(t),u(t))_{t\in[0,T]}$ be a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and let $\lambda(t)$ denote the eigenvalue of the Laplacian $\Delta_{g(t)}$. Suppose that $\mathcal{S}_{g(t),u(t)}-\alpha S_{g(t),u(t)}g(t)\geq 0$ along the harmonic-Ricci flow for some $\alpha\geq\frac{1}{2}$.*

(1) *If* $S_{\min}(0) \geq 0$, $\lambda(t)$ *is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

(2) *If* $S_{\min}(0) > 0$, *the quantity*

$$\left(1 - \frac{2}{n} S_{\min}(0)\, t\right)^{n\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for* $T \leq n/(2S_{\min}(0))$.

(3) *If* $S_{\min}(0) < 0$, *the quantity*

$$\left(1 - \frac{2}{n} S_{\min}(0)\, t\right)^{n\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

**Corollary 1.7.** *Let* $(g(t), u(t))_{t\in[0,T]}$ *be a solution of the harmonic-Ricci flow on a compact Riemannian surface* $\Sigma$ *and let* $\lambda(t)$ *denote the eigenvalue of the Laplacian* $\Delta_{g(t)}$.

(1) *Suppose that* $\mathrm{Ric}_{g(t)} \leq \epsilon\, du(t) \otimes du(t)$ *where*

$$\epsilon \leq 4\frac{1-\alpha}{1-2\alpha}, \qquad \alpha > \tfrac{1}{2}.$$

  (i) *If* $S_{\min}(0) \geq 0$, $\lambda(t)$ *is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

  (ii) *If* $S_{\min}(0) > 0$, *the quantity*

$$(1 - S_{\min}(0)\, t)^{2\alpha} \lambda(t)$$

  *is nondecreasing along the harmonic-Ricci flow for* $T \leq 1/S_{\min}(0)$.

  (iii) *If* $S_{\min}(0) < 0$, *the quantity*

$$(1 - S_{\min}(0)\, t)^{2\alpha} \lambda(t)$$

  *is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

(2) *Suppose that*

$$|\nabla_{g(t)} u(t)|^2_{g(t)}\, g(t) \geq 2\, du(t) \otimes du(t).$$

  (i) *If* $S_{\min}(0) \geq 0$, $\lambda(t)$ *is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

  (ii) *If* $S_{\min}(0) > 0$, *the quantity*

$$(1 - S_{\min}(0)\, t)\lambda(t)$$

  *is nondecreasing along the harmonic-Ricci flow for* $T \leq 1/S_{\min}(0)$.

(iii) *If* $S_{\min}(0) < 0$, *the quantity*

$$(1 - S_{\min}(0)t)\lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

When we restrict to the Ricci flow, we obtain:

**Corollary 1.8.** *Let* $(g(t))_{t\in[0,T]}$ *be a solution of the Ricci flow on a compact Riemannian surface* $\Sigma$ *and let* $\lambda(t)$ *denote the eigenvalue of the Laplacian* $\Delta_{g(t)}$.

(1) *If* $R_{\min}(0) \geq 0$, $\lambda(t)$ *is nondecreasing along the Ricci flow for any* $t \in [0, T]$.

(2) *If* $R_{\min}(0) > 0$, *the quantity* $(1 - R_{\min}(0)t)\lambda(t)$ *is nondecreasing along the Ricci flow for* $T \leq 1/R_{\min}(0)$.

(3) *If* $R_{\min}(0) < 0$, *the quantity* $(1 - R_{\min}(0)t)\lambda(t)$ *is nondecreasing along the Ricci flow for any* $t \in [0, T]$.

**Remark 1.9.** Let $(g(t))_{t\in[0,T]}$ be a solution of the Ricci flow on a compact Riemannian surface $\Sigma$ with nonnegative scalar curvature and let $\lambda(t)$ denote the eigenvalue of the Laplacian $\Delta_{g(t)}$. Then $\lambda(t)$ is nondecreasing along the Ricci flow for any $t \in [0, T]$.

Since

$$\mu(g, u) := \inf\left\{\mathcal{F}(g, u, f) \,\middle|\, f \in C^\infty(M),\ \int_M e^{-f}\, dV_g = 1\right\} \tag{1-11}$$

is the smallest eigenvalue of the operator $\Delta_{g,u} := -4\Delta_g + R_g - 2|\nabla_g u|_g^2$, we can consider the evolution equation for this eigenvalue under the harmonic-Ricci flow. To the operator $\Delta_{g,u}$ we associate a functional

$$\lambda_{g,u}(f) := \int_M f\Delta_{g,u} f\, dV_g. \tag{1-12}$$

When $f$ is an eigenfunction of the operator $\Delta_{g,u}$ with the eigenvalue $\lambda$ and normalized by $\int_X f^2\, dV_g = 1$, we obtain $\lambda_{g,u}(f) = \lambda$. Hence it suffices to study the evolution equation for $(d/dy)\lambda_{g,u}(f)$ under the harmonic-Ricci flow.

**Theorem 1.10.** *Suppose that* $(g(t), u(t))$ *is a solution of the harmonic-Ricci flow on a compact Riemannian manifold* $M$ *and* $f(t)$ *is an eigenfunction of* $\Delta_{g(t),u(t)}$, *that is,* $\Delta_{g(t),u(t)} f(t) = \lambda(t) f(t)$ (*where* $\lambda(t)$ *is only a function of time* $t$), *with the normalized condition* $\int_M f(t)^2\, dV_{g(t)} = 1$. *Then we have*

$$\begin{aligned}\frac{d}{dt}\lambda(t) = \frac{d}{dt}\lambda_{g,u}(f(t)) = \int_M & 2\langle \mathcal{S}_{g(t),u(t)}, df(t)\otimes df(t)\rangle_{g(t)}\, dV_{g(t)} \\ &+ \int_M f(t)^2\big(|\mathcal{S}_{g(t),u(t)}|_{g(t)}^2 + 2|\Delta_{g(t)}u(t)|_{g(t)}^2\big)\, dV_{g(t)}.\end{aligned} \tag{1-13}$$

List [2006] proved the nonnegativity of the operator $\mathcal{S}_{g(t),u(t)}$ is preserved by the harmonic-Ricci flow. Hence we get the following.

**Corollary 1.11.** *If* $\mathrm{Ric}_{g(0)} - 2du(0) \otimes du(0) \geq 0$, *the eigenvalues of the operator* $\Delta_{g(t),u(t)}$ *are nondecreasing under the harmonic-Ricci flow.*

**Remark 1.12.** If we choose $u(t) \equiv 0$, we obtain X. Cao's result [2007].

There is another expression for $d\lambda(t)/dt$.

**Theorem 1.13.** *Suppose that* $(g(t), u(t))$ *is a solution of the harmonic-Ricci flow on a compact Riemannian manifold* $M$ *and* $f(t)$ *is an eigenfunction of* $\Delta_{g(t),u(t)}$, *that is,* $\Delta_{g(t),u(t)} f(t) = \lambda(t) f(t)$ (*where* $\lambda(t)$ *is only a function of time* $t$), *with the normalized condition* $\int_M f(t)^2 \, dV_{g(t)} = 1$. *Then we have*

$$
\begin{aligned}
\frac{d}{dt}\lambda(t) = \frac{d}{dt}\lambda_{g,u}(f(t)) &= \frac{1}{2}\int_M |\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)}\varphi(t)|^2_{g(t)} e^{-\varphi(t)} \, dV_{g(t)} \\
&+ \frac{1}{4}\int_M |\mathcal{S}_{g(t),u(t)}|^2_{g(t)} e^{-\varphi(t)} \, dV_{g(t)} + \int_M |\langle du(t), d\varphi(t)\rangle_{g(t)}|^2 e^{-\varphi(t)} \, dV_{g(t)} \\
&+ 2\int_M |\nabla^2_{g(t)} u(t)|^2_{g(t)} e^{-\varphi(t)} \, dV_{g(t)} - \int_M \Delta_{g(t)}(|\nabla_{g(t)} u(t)|^2_{g(t)}) e^{-\varphi(t)} \, dV_{g(t)} \\
&+ \frac{1}{4}\int_M |\mathcal{S}_{g(t),u(t)} + 4du(t) \otimes du(t)|^2_{g(t)} e^{-\varphi(t)} \, dV_{g(t)},
\end{aligned} \tag{1-14}
$$

*where* $f(t)^2 = e^{-\varphi(t)}$.

**Remark 1.14.** When $u \equiv 0$, (1-14) reduces to J. Li's formula [2007].

Suppose that $M$ is a compact manifold of dimension $n$. For any Riemannian metric $g$, any smooth functions $u$, $f$, and any positive number $\tau$, we define

$$
\mathcal{W}_{\pm}(g, u, f, \tau) := \int_M [\tau(S_g + |\nabla_g f|^2_g) \mp f \pm n] \frac{e^{-f}}{(4\pi\tau)^{n/2}} \, dV_g. \tag{1-15}
$$

Set

$$
\mu_{\pm}(g, u, \tau) := \inf\left\{ \mathcal{W}_{\pm}(g, u, f, \tau) \,\middle|\, f \in C^\infty(M), \int_M \frac{e^{-f}}{(4\pi\tau)^{n/2}} \, dV_g = 1 \right\},
$$

$$
\nu_{-}(g, u) := \inf\{\mu_{-}(g, u, \tau) \mid \tau > 0\}, \quad \nu_{+}(g, u) := \sup\{\mu_{+}(g, u, \tau) \mid \tau > 0\}.
$$

The first variation of $\nu_{\pm}(g(s), u(s))$ is the following.

**Theorem 1.15.** *Suppose that* $(M, g)$ *is a compact Riemannian manifold and* $u$ *a smooth function on* $M$. *Let* $h$ *be any symmetric covariant* 2*-tensor on* $M$ *and set* $g(s) := g + sh$. *Let* $v$ *be any smooth function on* $M$ *and* $u(s) := u + sv$. *If*

*$\nu_\pm(g(s), u(s)) = \mathscr{W}_\pm(g(s), u(s), f_\pm(s), \tau_\pm(s))$ for some smooth functions $f_\pm(s)$ with $\int_M e^{-f_\pm(s)}\, dV_g/(4\pi\tau_\pm(s))^{n/2} = 1$ and constants $\tau_\pm(s) > 0$,*

$$
\begin{aligned}
\text{(1-16)}\quad \frac{d}{ds}\Big|_{s=0}\nu_\pm(g(s), u(s)) &= 4\tau_\pm \int_M v(\Delta_g u - \langle du, df_\pm\rangle_g)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g \\
&\quad - \tau_\pm \int_M \left(\langle h, \mathcal{S}_{g,u}\rangle_g + \langle h, \nabla_g^2 f\rangle_g \pm \frac{\operatorname{tr}_g h}{2\tau_\pm}\right)\frac{e^{-f_\pm}\, dV_g}{(4\pi\tau_\pm)^{n/2}},
\end{aligned}
$$

*where $f_\pm := f_\pm(0)$ and $\tau_\pm := \tau_\pm(0)$. In particular, the critical points of $\nu_\pm(\,\cdot\,,\cdot\,)$ satisfy*

$$
\mathcal{S}_{g,u} + \nabla_g^2 f \pm \frac{1}{2\tau_\pm} g = 0, \qquad \Delta_g u = \langle du, df_\pm\rangle_g.
$$

*Consequently, if $\mathscr{W}_\pm(g, u, f, \tau)$ and $\nu_\pm(g, u)$ achieve their extremum, $(M, g)$ is a gradient expanding and shrinker harmonic-Ricci soliton according to the sign.*

**Corollary 1.16.** *Suppose that $(M, g)$ is a compact Riemannian manifold and $u$ a smooth function on $M$. Let $h$ be any symmetric covariant 2-tensor on $M$ and set $g(s) := g + sh$. Let $v$ be any smooth function on $M$ and $u(s) := u + sv$. If $\nu_\pm(g(s), u(s)) = \mathscr{W}_\pm(g(s), u(s), f_\pm(s), \tau_\pm(s))$ for some smooth function $f_\pm(s)$ with $\int_M e^{-f_\pm(s)}\, dV/(4\pi\tau_\pm(s))^{n/2} = 1$ and a constant $\tau_\pm(s) > 0$, and $(g, u)$ is a critical point of $\nu_\pm(\,\cdot\,,\cdot\,)$, then*

$$
\operatorname{Ric}_g = \mp\frac{1}{2\tau_\pm} g, \qquad f_\pm \equiv \text{constant}, \qquad u \equiv \text{constant}.
$$

*Thus, if $\mathscr{W}_\pm(g, u, \cdot\,, \cdot)$ achieve their minimum and $(g, u)$ is a critical point of $\nu_\pm(\,\cdot\,,\cdot\,)$, $(M, g)$ is an Einstein manifold and $u$ is a constant function.*

**Remark 1.17.** In the situation of Corollary 1.16, by normalization, we my choose $f_\pm = n/2$ and $u = 0$.

## 2. Notation and commuting identities

Let $M$ be a compact Riemannian manifold of dimension $n$. For any vector bundle $E$ over $M$, we denote by $\Gamma(M, E)$ the space of smooth sections of $E$. Set

$$
\begin{aligned}
\odot^2(M) &:= \{v = (v_{ij}) \in \Gamma(M, T^*M \otimes T^*M) \mid v_{ij} = v_{ji}\}, \\
\odot^2_+(M) &:= \{g = (g_{ij}) \in \odot^2(M) \mid g_{ij} > 0\}.
\end{aligned}
$$

Thus $\odot^2(M)$ is the space of all symmetric covariant 2-tensors on $M$, while $\odot^2_+(M)$ is the space of all Riemannian metrics on $M$. The space of all smooth functions on $M$ is denoted by $C^\infty(M)$.

For a given Riemannian metric $g \in \bigodot^2_+(M)$, the corresponding Levi-Civita connection $\Gamma_g = (\Gamma^k_{ij})$ is given by

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{k\ell}(\partial_i g_{j\ell} + \partial_j g_{il} - \partial_\ell g_{ij}) \tag{2-1}$$

where $\partial_i := \partial/\partial x^i$ for a local coordinate system $\{x^1, \ldots, x^n\}$. The Riemann tensor $\mathrm{Rm}_g = (R^k_{ijl})$ is determined by

$$R^k_{ij\ell} = \partial_i \Gamma^k_{j\ell} - \partial_j \Gamma^k_{i\ell} + \Gamma^k_{ip}\Gamma^p_{j\ell} - \Gamma^k_{jp}\Gamma^p_{i\ell}. \tag{2-2}$$

The Ricci curvature $\mathrm{Ric}_g = (R_{ij})$ is

$$R_{ij} = g^{k\ell} R^\ell_{kij}. \tag{2-3}$$

The scalar curvature $R_g$ of the metric $g$ now is given by

$$R_g = g^{ij} R_{ij}. \tag{2-4}$$

For any tensor $A = (A^{k_1\cdots k_q}_{j_1\cdots j_p})$ the covariant derivative of $A$ is

$$\nabla_i A^{k_1\cdots k_q}_{j_1\cdots j_p} = \partial_i A^{k_1\cdots k_q}_{j_1\cdots j_p} - \sum_{r=1}^{p} \Gamma^m_{ij_r} A^{k_1\cdots k_q}_{j_1\cdots m\cdots j_p} + \sum_{s=1}^{q} \Gamma^{k_s}_{im} A^{k_1\cdots m\cdots k_q}_{j_1\cdots j_p}.$$

Next we recall the Ricci identity:

$$\nabla_i \nabla_j A^{\ell_1\cdots\ell_q}_{k_1\cdots k_p} - \nabla_j \nabla_i A^{\ell_1\cdots\ell_q}_{k_1\cdots k_p} = \sum_{r=1}^{q} R^{l_r}_{ijm} A^{\ell_1\cdots m\cdots\ell_q}_{k_1\cdots k_p} - \sum_{s=1}^{p} R^m_{ijk_s} A^{\ell_1\cdots\ell_q}_{k_1\cdots m\cdots k_p}.$$

In particular, for any smooth function $f \in C^\infty(M)$, we have

$$\nabla_i \nabla_j f = \nabla_j \nabla_i f.$$

The Bianchi identities are

$$0 = R_{ijk\ell} + R_{iklj} + R_{i\ell jk}, \tag{2-5}$$

$$0 = \nabla_q R_{ijk\ell} + \nabla_i R_{jqk\ell} + \nabla_j R_{qik\ell}, \tag{2-6}$$

and the contracted Bianchi identities are

$$0 = 2\nabla^j R_{ij} - \nabla_i R_g, \tag{2-7}$$

$$0 = \nabla_i R_{jk} - \nabla_j R_{ik} + \nabla^\ell R_{\ell kij}. \tag{2-8}$$

## 3. Harmonic-Ricci flow and the evolution equations

Motivated by the static Einstein vacuum equation, List [2006] introduced the harmonic-Ricci flow (originally called the Ricci flow coupled with the harmonic map flow). This flow is similar to the Ricci flow and is given by the coupled system

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}_{g(t)} + 4\,du(t) \otimes du(t), \tag{3-1}$$

$$\frac{\partial}{\partial t} u(t) = \Delta_{g(t)} u(t) \tag{3-2}$$

for a family of Riemannian metrics $g(t)$ and a family of smooth functions $u(t)$. Locally, we have

$$\frac{\partial}{\partial t} g_{ij} = -2R_{ij} + 4\partial_i u \cdot \partial_j u, \qquad \frac{\partial}{\partial t} u = \Delta_{g(t)} u(t). \tag{3-3}$$

Introduce a new symmetric tensor field $\mathscr{S}_{g(t),u(t)} = (S_{ij}) \in \bigodot^2(M)$,

$$S_{ij} := R_{ij} - 2\partial_i u \cdot \partial_j u. \tag{3-4}$$

Then its trace $S_{g(t),u(t)}$ is equal to

$$S_{g(t),u(t)} = g^{ij} S_{ij} = R_{g(t)} - 2|\nabla_{g(t)} u(t)|^2_{g(t)}. \tag{3-5}$$

The evolution equation for $R_{g(t)}$ is

$$\begin{aligned} \frac{\partial}{\partial t} R_{g(t)} = \Delta_{g(t)} R_{g(t)} + 2|\,\mathrm{Ric}_{g(t)}\,|^2_{g(t)} + 4|\Delta_{g(t)} u(t)|^2_{g(t)} \\ - 4|\nabla^2_{g(t)} u(t)|^2_{g(t)} - 8\langle \mathrm{Ric}_{g(t)}, du(t) \otimes du(t)\rangle_{g(t)}. \end{aligned} \tag{3-6}$$

Also, we have the evolution equation for $|\nabla_{g(t)} u|^2_{g(t)}$,

$$\frac{\partial}{\partial t} |\nabla_{g(t)} u(t)|^2_{g(t)} = \Delta_{g(t)} |\nabla_{g(t)} u(t)|^2_{g(t)} - 2|\nabla^2_{g(t)} u(t)|^2_{g(t)} - 4|\nabla_{g(t)} u(t)|^4_{g(t)}, \tag{3-7}$$

and the evolution equation for $S_{g(t),u(t)}$,

$$\frac{\partial}{\partial t} S_{g(t),u(t)} = \Delta_{g(t)} S_{g(t),u(t)} + 2|\mathscr{S}_{g(t),u(t)}|^2_{g(t)} + 4|\Delta_{g(t)} u(t)|^2_{g(t)}. \tag{3-8}$$

## 4. Entropies for harmonic-Ricci flow

Motivated by Perelman's entropy, List [2006] introduced a similar functional for the harmonic-Ricci flow:

$$\bigodot^2_+(M) \times C^\infty(M) \times C^\infty(M) \to \mathbb{R}, \quad (g, u, f) \mapsto \mathscr{F}(g, u, f)$$

where

$$\mathcal{F}(g,u,f) := \int_M (R_g + |\nabla_g f|_g^2 - 2|\nabla_g u|_g^2) e^{-f}\, dV_g. \tag{4-1}$$

He also showed that if $(g(t), u(t), f(t))$ satisfies the system

$$\begin{aligned} \frac{\partial}{\partial t} g(t) &= -2\,\mathrm{Ric}_{g(t)} + 4\, du(t) \otimes du(t) - 2\nabla^2_{g(t)} f(t), \\ \frac{\partial}{\partial t} u(t) &= \Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}, \\ \frac{\partial}{\partial t} f(t) &= -\Delta_{g(t)} f(t) - R_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)}, \end{aligned} \tag{4-2}$$

the evolution of the entropy is given by

$$\begin{aligned} &\frac{d}{dt}\mathcal{F}(g(t),u(t),f(t)) \\ &\quad = 2\int_M \Big( |\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t)|^2_{g(t)} \\ &\qquad\qquad + 2|\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} \Big) e^{-f(t)}\, dV_{g(t)} \\ &\quad \geq 0. \end{aligned} \tag{4-3}$$

**Remark 4.1.** The system (4-2) is equivalent to

$$\begin{aligned} \frac{\partial}{\partial t} g(t) &= -2\,\mathrm{Ric}_{g(t)} + 4\, du(t) \otimes du(t), \\ \frac{\partial}{\partial t} u(t) &= \Delta_{g(t)} u(t), \\ \frac{\partial}{\partial t} f(t) &= -\Delta_{g(t)} f(t) - R_{g(t)} + |\nabla_{g(t)} f(t)|^2_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)}. \end{aligned} \tag{4-4}$$

The same evolution of the entropy holds for system (4-4).

In particular, the entropy is nondecreasing and the equality holds if and only if $(g(t), u(t), f(t))$ satisfies

$$\begin{aligned} \mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) &= 0, \\ \Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)} &= 0. \end{aligned} \tag{4-5}$$

**Definition 4.2.** The $\mathcal{E}$-functional is defined as

$$\bigodot^2_+(M) \times C^\infty(M) \times C^\infty(M) \to \mathbb{R}, \quad (g,u,f) \mapsto \mathcal{E}(g,u,f),$$

where

$$\mathcal{E}(g,u,f) := \int_M (R_g - 2|\nabla_g u|_g^2) e^{-f}\, dV_g. \tag{4-6}$$

**Proposition 4.3.** *Under the evolution equation (4-4), one has*

$$\frac{d}{dt}\mathscr{E}(g(t),u(t),f(t)) = 2\int_M |\mathscr{S}_{g(t),u(t)}|^2_{g(t)} e^{-f(t)}\,dV_{g(t)} + 4\int_M |\Delta_{g(t)}u(t)|^2_{g(t)} e^{-f(t)}\,dV_{g(t)}. \tag{4-7}$$

*Proof.* Since $S_{g(t),u(t)} = R_{g(t)} - 2|\nabla_{g(t)}u(t)|^2_{g(t)}$ and

$$\begin{aligned}\frac{\partial}{\partial t}S_{g(t),u(t)} &= \Delta_{g(t)}S_{g(t),u(t)} + 2|\mathscr{S}_{g(t),u(t)}|^2_{g(t)} + 4|\Delta_{g(t)}u(t)|^2_{g(t)},\\ \frac{\partial}{\partial t}dV_{g(t)} &= -S_{g(t),u(t)}\,dV_{g(t)},\end{aligned}$$

we have

$$\begin{aligned}&\frac{d}{dt}\mathscr{E}(g(t),u(t),f(t))\\ &= \int_M \Big(\frac{\partial}{\partial t}S_{g(t),u(t)}\Big)e^{-f(t)}\,dV_{g(t)} + \int_M S_{g(t),u(t)}\frac{\partial}{\partial t}\big(e^{-f(t)}\,dV_{g(t)}\big)\\ &= \int_M \big(\Delta_{g(t)}S_{g(t),u(t)} + 2|\mathscr{S}_{g(t),u(t)}|^2_{g(t)} + 4|\Delta_{g(t)}u(t)|^2_{g(t)}\big)e^{-f(t)}\,dV_{g(t)}\\ &\qquad + \int_M S_{g(t),u(t)}\Big(-\frac{\partial}{\partial t}f(t) - S_{g(t),u(t)}\Big)e^{-f(t)}\,dV_{g(t)}\\ &= 2\int_M |\mathscr{S}_{g(t),u(t)}|^2_{g(t)}e^{-f(t)}\,dV_{g(t)} + 4\int_M |\Delta_{g(t)}u(t)|^2_{g(t)}e^{-f(t)}\,dV_{g(t)}\\ &\quad - \int_M S_{g(t),u(t)}\Big(\Delta_{g(t)}f(t) - |\nabla_{g(t)}f(t)|^2_{g(t)} + \frac{\partial}{\partial t}f(t) + S_{g(t),u(t)}\Big)e^{-f(t)}\,dV_{g(t)},\end{aligned}$$

which implies (4-7). □

**Definition 4.4.** For any $k \geq 1$ we define

$$\mathscr{F}_k(g,u,f) := \int_M \big(kR_g + |\nabla_g f|^2_g - 2k|\nabla_g u|^2_g\big)e^{-f}\,dV_g. \tag{4-8}$$

Using the definition, it is easy to show that

$$\mathscr{F}_k(g,u,f) = (k-1)\mathscr{E}(g,u,f) + \mathscr{F}(g,u,f). \tag{4-9}$$

When $k = 1$, this is the $\mathscr{F}$-functional.

**Theorem 4.5.** *Under the evolution equation (4-4), one has*

$$
\begin{aligned}
\text{(4-10)}\quad & \frac{d}{dt}\mathcal{F}_k(g(t), u(t), f(t)) \\
&= 2(k-1)\int_M |\mathcal{S}_{g(t),u(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} + 2\int_M |\mathcal{S}_{g(t),u(t)} \\
&\qquad + \nabla^2_{g(t)} f(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} + 4(k-1)\int_M |\Delta_{g(t)} u(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\qquad + 4\int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)}.
\end{aligned}
$$

*Furthermore, the monotonicity is strict unless $g(t)$ is Ricci-flat, $u(t)$ is constant, and $f(t)$ is constant.*

*Proof.* It immediately follows from (4-3) and (4-7). □

Set

$$
\text{(4-11)}\qquad \mu_k(g, u) := \inf\left\{\mathcal{F}_k(g, u, f)\,\middle|\, f \in C^\infty(M),\ \int_M e^{-f}\, dV_g = 1\right\}.
$$

Then $\mu_k(g, u)$ is the lowest eigenvalue of $-4\Delta_g + k(R_g - 2|\nabla_g u|^2_g)$.

## 5. Compact steady harmonic-Ricci breathers

In this section we give an alternative proof on some results on compact steady harmonic-Ricci breathers that were proved in [List 2006; Müller 2012].

**Definition 5.1.** A solution $(g(t), u(t))$ of the harmonic-Ricci flow (1-1)–(1-2) is called a *harmonic-Ricci breather* if there exist $t_1 < t_2$, a diffeomorphism $\psi : M \to M$, and a constant $\alpha > 0$ such that

$$
g(t_2) = \alpha\psi^* g(t_1), \qquad u(t_2) = \psi^* u(t_1).
$$

The cases $\alpha < 1$, $\alpha = 1$, and $\alpha > 1$, correspond to *shrinking*, *steady*, and *expanding harmonic-Ricci breathers*.

**Theorem 5.2.** *If $(g(t), u(t))$ is a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$, the lowest eigenvalue $\mu_k(g(t), u(t))$ of the operator $-4\Delta_{g(t)} + k(R_{g(t)} - 2|\nabla_{g(t)} u(t)|^2_{g(t)})$ is nondecreasing under the harmonic-Ricci flow. The monotonicity is strict unless $g(t)$ is Ricci-flat and $u(t)$ is constant.*

*Proof.* The proof is similar to that given in [Li 2007]. For any $t_1 < t_2$, suppose that

$$
\mu_k(g(t_2), u(t_2)) = \mathcal{F}_k(g(t_2), u(t_2), f_k(t_2))
$$

for some smooth function $f_k(x)$. Being an initial value, $f_k(x) = f_k(x, t_2)$ for some smooth function $f_k(x, t)$ satisfying the evolution equation (4-4). The monotonicity formula (4-10) implies $\mu_k(g(t_2), u(t_2)) \geq \mathcal{F}_k(g(t_1), u(t_1), f_k(t_1)) \geq \mu_k(g(t_1), u(t_1))$. This completes the proof. □

**Corollary 5.3.** *On a compact Riemannian manifold, the lowest eigenvalues of* $-\Delta_{g(t)} + (1/2)(R_{g(t)} - 2|\nabla_{g(t)} u(t)|^2_{g(t)})$ *are nondecreasing under the harmonic-Ricci flow.*

*Proof.* Since $\mu_2(g(t), u(t))/4$ is the lowest eigenvalue of this operator, the result immediately follows from Theorem 5.2. □

**Corollary 5.4.** *There is no compact steady harmonic-Ricci breather unless the manifold* $(M, g(t))$ *is Ricci-flat and* $u$ *is a constant.*

*Proof.* If $(g(t), u(t))$ is a steady harmonic-Ricci breather, then, for $t_1 < t_2$ given in the definition, we have

$$\mu_k(g(t_1), u(t_1)) = \mu_k(g(t_2), u(t_2)).$$

Hence, using Theorem 5.2, for any $t \in [t_1, t_2]$, we must have

$$\frac{d}{dt}\mu_k(g(t), u(t)) \equiv 0.$$

Thus $(M, g(t))$ is Ricci-flat and $u(t)$ is constant. □

## 6. Compact expanding harmonic-Ricci breathers

Inspired by [Li 2007], we define a new functional

$$\bigodot{}^2_+(M) \times C^\infty(M) \times C^\infty(\mathbb{R}) \times C^\infty(M) \to \mathbb{R}, \quad (g, u, \tau, f) \mapsto \mathcal{W}_+(g, u, \tau, f),$$

where $(\tau = \tau(t), t \in \mathbb{R})$.

$$\mathcal{W}_+(g, u, \tau, f) := \tau^2 \int_M \Big(R_g + \frac{n}{2\tau} + \Delta_g f - 2|\nabla_g u|^2_g\Big) e^{-f}\, dV_g. \tag{6-1}$$

Similarly, we define a family of functionals

$$\mathcal{W}_{+,k}(g, u, \tau, f) := \tau^2 \int_M \Big(k\Big(R_g + \frac{n}{2\tau}\Big) + \Delta_g f - 2k|\nabla_g u|^2_g\Big) e^{-f}\, dV_g. \tag{6-2}$$

It's clear that $\mathcal{W}_{+,1}(g, u, \tau, f) = \mathcal{W}_+(g, u, \tau, f)$.

**Lemma 6.1.** *One has*

$$\mathcal{W}_+(g, u, \tau, f) = \tau^2 \mathcal{F}(g, u, f) + \frac{n}{2}\tau \int_M e^{-f}\, dV_g,$$

$$\mathcal{W}_{+,k}(g, u, \tau, f) = \tau^2 \mathcal{F}_k(g, u, f) + \frac{kn}{2}\tau \int_M e^{-f}\, dV_g,$$

$$\mathcal{W}_{+,k}(g, u, \tau, f) = \mathcal{W}_+(g, u, \tau, f) + (k-1)\Big(\tau^2 \mathcal{E}(g, u, f) + \frac{n}{2}\tau \int_M e^{-f}\, dV_g\Big).$$

*Proof.* Since $\Delta(e^{-f}) = (-\Delta f + |\nabla f|^2)e^{-f}$, it follows that

$$\begin{aligned}
\mathcal{W}_+(g, u, \tau, f) - \tau^2\mathcal{F}(g, u, f) \\
&= \frac{n}{2}\tau \int_M e^{-f}\, dV_g + \tau^2 \int_M (\Delta_g f - |\nabla_g f|_g^2)e^{-f}\, dV_g \\
&= \frac{n}{2}\tau \int_M e^{-f}\, dV_g + \tau^2 \int_M \Delta_g(e^{-f})\, dV_g = \frac{n}{2}\tau \int_M e^{-f}\, dV_g.
\end{aligned}$$

We can similarly prove the remaining two relations. □

**Theorem 6.2.** *Under the coupled system*

$$\begin{aligned}
\frac{\partial}{\partial t} g(t) &= -2\operatorname{Ric}_{g(t)} + 4\, du(t) \otimes du(t) - 2\nabla^2_{g(t)} f(t), \\
\frac{\partial}{\partial t} u(t) &= \Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}, \\
\frac{\partial}{\partial t} f(t) &= -\Delta_{g(t)} f(t) - R_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)}, \\
\frac{d}{dt}\tau(t) &= 1,
\end{aligned}$$

*the first variation formula for* $\mathcal{W}_+(g(t), u(t), \tau(t), f(t))$ *is*

$$\begin{aligned}
\text{(6-3)}\quad & \frac{d}{dt}\mathcal{W}_+(g(t), u(t), \tau(t), f(t)) \\
&= 2\tau(t)^2 \int_M \Big|\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) + \frac{1}{2\tau(t)} g(t)\Big|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 4\tau(t)^2 \int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)},
\end{aligned}$$

*and the first variation formula for* $\mathcal{W}_{+,k}(g(t), u(t), \tau(t), f(t))$ *is*

$$\begin{aligned}
\text{(6-4)}\quad & \frac{d}{dt}\mathcal{W}_{+,k}(g(t), u(t), \tau(t), f(t)) \\
&= 2\tau(t)^2 \int_M \Big|\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) + \frac{1}{2\tau(t)} g(t)\Big|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 2(k-1)\tau(t)^2 \int_M \Big|\mathcal{S}_{g(t),u(t)} + \frac{1}{2\tau(t)} g(t)\Big|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 4\tau(t)^2 \int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\
&\quad + 4(k-1)\tau(t)^2 \int_M |\Delta_{g(t)} u(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)}.
\end{aligned}$$

*Proof.* Under this coupled system, we first observe that

$$\frac{d}{dt}\left(\int_M e^{-f(t)}\, dV_{g(t)}\right) = 0. \tag{6-5}$$

In fact, from $\frac{\partial}{\partial t}\, dV_{g(t)} = -S_{g(t),u(t)} - \Delta_{g(t)} f(t)\, dV_{g(t)}$ we obtain

$$\begin{aligned}\frac{d}{dt}\left(\int_M e^{-f(t)} dV_{g(t)}\right) &= \int_M \left(-\frac{\partial}{\partial t} f(t)\cdot dV_{g(t)} + \frac{\partial}{\partial t} dV_{g(t)}\right) e^{-f(t)} \\ &= \int_M [\Delta_{g(t)} f(t) + S_{g(t),u(t)} - S_{g(t),u(t)} - \Delta_{g(t)} f(t)] e^{-f(t)} dV_{g(t)} \\ &= 0.\end{aligned}$$

Lemma 6.1 and the identity (6-5) imply

$$\begin{aligned}&\frac{d}{dt}\mathcal{W}_+(g(t), u(t), \tau(t), f(t)) \\ &\quad = \tau(t)^2 \frac{d}{dt}\mathcal{F}(g(t), u(t), f(t)) + 2\tau(t)\mathcal{F}(g(t), u(t), f(t)) + \frac{n}{2}\int_M e^{-f(t)}\, dV_{g(t)} \\ &\quad = 2\tau(t)^2 \int_M |\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t)|^2_{g(t)} e^{-f(t)}\, dV_{g(t)} \\ &\qquad\qquad + 4\tau(t)^2 \int_M |\Delta_{g(t)} u(t) - \langle du(t), df(t)\rangle_{g(t)}|^2 e^{-f(t)}\, dV_{g(t)} \\ &\qquad + 2\tau(t)\int_M (S_{g(t),u(t)} + |\nabla_{g(t)} f(t)|^2_{g(t)}) e^{-f(t)}\, dV_{g(t)} + \frac{n}{2}\int_M e^{-f(t)}\, dV_{g(t)},\end{aligned}$$

which is (6-3). Using Lemma 6.1 and the same method, we can prove (6-4). □

**Remark 6.3.** Under the coupled system

$$\begin{aligned}\frac{\partial}{\partial t} g(t) &= -2\,\mathrm{Ric}_{g(t)} + 4\, du(t)\otimes du(t), \\ \frac{\partial}{\partial t} u(t) &= \Delta_{g(t)} u(t), \\ \frac{\partial}{\partial t} f(t) &= -\Delta_{g(t)} f(t) + |\nabla_{g(t)} f(t)|^2_{g(t)} - R_{g(t)} + 2|\nabla_{g(t)} u(t)|^2_{g(t)}, \\ \frac{d}{dt}\tau(t) &= 1,\end{aligned}$$

the same formulas (6-3) and (6-4) hold for $\mathcal{W}_+$ and $\mathcal{W}_{+,k}$.

Define

(6-6) $$\mu_+(g, u, \tau) := \inf\left\{\mathcal{W}_+(g, u, \tau, f) \,\middle|\, f \in C^\infty(M),\ \int_M e^{-f}\, dV_g = 1\right\}.$$

**Lemma 6.4.** *For any* $\alpha > 0$, *one has*

(6-7) $$\mu_+(\alpha g, u, \alpha\tau) = \alpha\mu_+(g, u, \tau).$$

*Proof.* Set $\bar{g} := \alpha g$; then $R_{\bar{g}} = \alpha^{-1} R_g$, $\Delta_{\bar{g}} f = \alpha^{-1}\Delta_g f$, and $|\nabla_{\bar{g}} u|^2_{\bar{g}} = \alpha^{-1}|\nabla_{g(t)} u|^2_g$. Hence

$$\begin{aligned}\mathcal{W}_+(\bar g, u, \alpha\tau, f) &= \alpha^2\tau^2 \int_M \Big(R_{\bar g} + \frac{n}{2\alpha\tau} + \Delta_{\bar g} f - 2|\nabla_{\bar g} u|^2_{\bar g}\Big) e^{-f}\, dV_{\bar g} \\ &= \alpha\tau^2 \int_M \Big(R_g + \frac{n}{2\tau} + \Delta_g f - 2|\nabla_{g(t)} u|^2_g\Big) \alpha^{n/2} e^{-f}\, dV_g.\end{aligned}$$

Since $f \mapsto f - (n/2)\ln\alpha$ is one-to-one and onto, by taking the infimum, we derive $\mu_+(\alpha g, u, \alpha\tau) = \alpha\mu_+(g, u, \tau)$. □

**Definition 6.5.** A solution $(g(t), u(t))$ of the harmonic-Ricci flow is called a *harmonic-Ricci soliton* if there exists a one-parameter family of diffeomorphisms $\psi_t : M \to M$, satisfying $\psi_0 = \mathrm{id}_M$, and a positive scaling function $\alpha(t)$ such that

$$g(t) = \alpha(t)\psi_t^* g(0), \quad u(t) = \psi_t^* u(0).$$

The cases $(\partial/\partial t)\alpha(t) = \dot\alpha < 0$, $\dot\alpha = 0$, and $\dot\alpha > 0$ correspond to *shrinking*, *steady*, and *expanding harmonic-Ricci solitons*, respectively. If the diffeomorphisms $\psi_t$ are generated by a (possibly time-dependent) vector field $X(t)$ that is the gradient of some function $f(t)$ on $M$, the soliton is called a *gradient harmonic-Ricci soliton* and $f$ is called the *potential of the harmonic-Ricci soliton*.

Müller [2012] showed that if $(g(t), u(t))$ is a gradient harmonic-Ricci soliton with potential $f$,

$$\begin{aligned}0 &= \mathrm{Ric}_{g(t)} - 2du(t) \otimes du(t) + \nabla^2_{g(t)} f(t) + cg(t), \\ 0 &= \Delta_{g(t)} u(t) - \langle \nabla_{g(t)} u(t), \nabla_{g(t)} f(t)\rangle_{g(t)}\end{aligned}$$

for some constant $c$.

**Corollary 6.6.** *There is no expanding breather on compact Riemannian manifolds other than expanding gradient harmonic-Ricci solitons.*

*Proof.* The proof is similar to that given in [Li 2007]. Suppose there is an expanding breather on a compact Riemannian manifold $M$. Then, by definition, we have

$$g(t_2) = \alpha\Phi^* g(t_1), \quad u(t_2) = \Phi^* u(t_1)$$

for some $t_1 < t_2$, where $\Phi$ be a diffeomorphism and the constant $\alpha > 1$. Let $f_+(x)$ be a smooth function where $\mathcal{W}_+(g(t_2), u(t_2), \tau(t_2), f(t_2))$ attains its minimum. Then there exists a smooth function $f_+(x, t) : M \times [t_1, t_2] \to \mathbb{R}$ with initial value $f_+(x, t_2) = f_+(x)$ that satisfies the coupled system in Remark 6.3. Define a linear function

$$\tau : [t_1, t_2] \to (0, +\infty), \quad \tau(t_2) = T + t_2$$

where $T$ is a constant. By the monotonicity formula, we have

$$\begin{aligned}\mu_+(g(t_2), u(t_2), \tau(t_2)) &= \mathcal{W}_+(g(t_2), u(t_2), \tau(t_2), f_+(t_2))\\ &\geq \mathcal{W}_+(g(t_1), u(t_1), \tau(t_1), f_+(t_1))\\ &\geq \mu_+(g(t_1), u(t_1), \tau(t_1)).\end{aligned}$$

Lemma 6.4 and the diffeomorphic invariant property of the functionals shows

$$\mu_+(g(t_1), u(t_1), \tau(t_1)) \leq \alpha\mu_+(g(t_1), u(t_1), \tau(t_1)),$$

which yields

$$\mu_+(g(t_1), u(t_1), \tau(t_1)) \geq 0,$$

since $\alpha > 1$.

If we impose an additional condition $\tau(t_2) = \alpha\tau(t_1)$ and $\tau(t_1) = T + t_1$, we have

$$\tau(t) = \frac{\alpha(t-t_1)-(t-t_2)}{\alpha-1}, \quad T = \frac{t_2-\alpha t_1}{\alpha-1}.$$

Then

$$\frac{\tau(t_2)^{n/2}}{V_{g(t_2)}} = \frac{[\alpha(t_2-t_1)/(\alpha-1)]^{n/2}}{\alpha^{n/2}V_{g(t_1)}} = \frac{\tau(t_1)^{n/2}}{V_{g(t_1)}}.$$

The mean value theorem tells us that there exists a time $\bar{t} \in [t_1, t_2]$ with

$$\begin{aligned}0 &= \frac{d}{dt}\Big|_{t=\bar{t}} \log \frac{\tau(t)^{n/2}}{V_{g(t)}}\\ &= \frac{V_{g(\bar{t})}}{\tau(\bar{t})^{n/2}} \cdot \frac{(n/2)\tau(\bar{t})^{n/2-1}V_{g(\bar{t})} - \tau(\bar{t})^{n/2}(d/dt)|_{t=\bar{t}}V_{g(t)}}{V^2_{g(\bar{t})}}\\ &= \frac{n}{2\tau(\bar{t})} - \frac{1}{V_{g(\bar{t})}}\frac{\partial}{\partial t}\Big|_{t=\bar{t}}V_{g(\bar{t})}.\end{aligned}$$

From the evolution equation for the volume element $dV_{g(t)}$, we have

$$\frac{d}{dt}V_{g(t)} = \int_M \frac{\partial}{\partial t}\,dV_{g(t)} = \int_M (-S_{g(t),u(t)} - \Delta_{g(t)}f(t))\,dV_{g(t)} = -\int_M S_{g(t),u(t)}\,dV_{g(t)}.$$

Putting these together yields

$$0 = \frac{n}{2\tau(\bar{t})} + \frac{1}{V_{g(\bar{t})}}\int_M S_{g(\bar{t}),u(\bar{t})}\,dV_{g(\bar{t})} = \frac{1}{V_{g(\bar{t})}}\int_M \Big(S_{g(\bar{t}),u(\bar{t})} + \frac{n}{2\tau(\bar{t})}\Big)\,dV_{g(\bar{t})}.$$

If we set $\bar{f} = \log V_{g(\bar{t})}$,

$$0 = \mathcal{W}_+(g(\bar{t}), u(\bar{t}), \tau(\bar{t}), \bar{f}) \geq \mu_+(g(\bar{t}), u(\bar{t}), \tau(\bar{t})).$$

By the monotonicity of $\mu_+$ we obtain

$$0 \le \mu_+(g(t_1), u(t_1), \tau(t_1)) \le \mu_+(g(\bar{t}), u(\bar{t}), \tau(\bar{t})) \le 0$$

Hence $\mu_+(g(t_1), u(t_1), \tau(t_1)) = \mu_+(g(t_2), u(t_2), \tau(t_2)) = 0$ and $\mathcal{W}_+ = 0$ on the interval $[t_1, t_2]$. This indicates that the first variation of $\mathcal{W}_+$ must vanish. So the expanding breather is a gradient soliton, that is,

$$\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) + \frac{1}{2\tau(t)} g(t) = 0.$$

Moreover, in this case $\Delta_{g(t)} u(t) = \langle du(t), df(t)\rangle_{g(t)}$. □

Because of (6-7), we define

$$\mu_{+,k}(g, u, \tau) := \inf\left\{ \mathcal{W}_{+,k}(g, u, \tau, f) \,\middle|\, f \in C^{+\infty}(M),\ \int_M e^{-f} dV_g = 1 \right\}. \tag{6-8}$$

Due to Lemma 6.4, we still have

$$\mu_{+,k}(\alpha g, u, \alpha\tau) = \alpha \mu_{+,k}(g, u, \tau). \tag{6-9}$$

**Corollary 6.7.** *If $(g(t), u(t))$ is an expanding harmonic-Ricci breather on compact Riemannian manifolds, $M$ is an Einstein manifold and $u(t)$ is constant.*

*Proof.* Using the same method as in Corollary 6.6 and $\mu_{+,k}$, we can show that the first variation of $\mathcal{W}_{+,k}$ must vanish. Hence, from (6-4), one has

$$\begin{aligned}
\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)} f(t) + \frac{1}{2\tau(t)} g(t) &= 0,\\
\mathcal{S}_{g(t),u(t)} + \frac{1}{2\tau(t)} g(t) &= 0,\\
\Delta_{g(t)} u(t) &= \langle du(t), df(t)\rangle_{g(t)},\\
\Delta_{g(t)} u(t) &= 0.
\end{aligned}$$

The above four equations can be reduced to the coupled equation

$$\mathcal{S}_{g(t),u(t)} + \frac{1}{2\tau(t)} g(t) = 0 = \Delta_{g(t)} u(t),$$

which indicates that $u(t)$ is a constant and $\mathrm{Ric}_{g(t)} = -(1/(2\tau(t)))g(t)$. □

## 7. Eigenvalues of the Laplacian under the harmonic-Ricci flow

In this section we consider the eigenvalues of the Laplacian $\Delta_{g(t)}$ under the harmonic-Ricci flow

$$\frac{\partial}{\partial t}g(t) = -2\,\mathrm{Ric}_{g(t)} + 4\,du(t)\otimes du(t), \tag{7-1}$$

$$\frac{\partial}{\partial t}u(t) = \Delta_{g(t)}u(t). \tag{7-2}$$

Suppose that $\lambda(t)$, which is a function of time $t$ only, is an eigenvalue of the Laplacian $\Delta_{g(t)}$ with an eigenfunction $f(t) = f(x,t)$, that is,

$$-\Delta_{g(t)}f(t) = \lambda(t)f(t). \tag{7-3}$$

Taking the derivative with respect to $t$, we get

$$-\Big(\frac{\partial}{\partial t}\Delta_{g(t)}\Big)f(t) - \Delta_{g(t)}\Big(\frac{\partial}{\partial t}f(t)\Big) = \Big(\frac{d}{dt}\lambda(t)\Big)f(t) + \lambda(t)\frac{\partial}{\partial t}f(t).$$

Integrating the above equation with $f$ yields

$$\begin{aligned}-\int_M f(t)\Big(\frac{\partial}{\partial t}\Delta_{g(t)}\Big)f(t)\,dV_{g(t)} - \int_M f(t)\Delta_{g(t)}\Big(\frac{\partial}{\partial t}f(t)\Big)\,dV_{g(t)}\\ = \frac{d}{dt}\lambda(t)\cdot\int_M f(t)^2\,dV_{g(t)} + \lambda(t)\int_M f(t)\frac{\partial}{\partial t}f(t)\,dV_{g(t)}.\end{aligned}$$

Since

$$\begin{aligned}-\int_M f(t)\Delta\Big(\frac{\partial}{\partial t}f(t)\Big)\,dV_{g(t)} &= -\int_M \Delta_{g(t)}f(t)\cdot\frac{\partial}{\partial t}f(t)\,dV_{g(t)}\\ &= \lambda(t)\int_M f(t)\frac{\partial}{\partial t}f(t)\,dV_{g(t)},\end{aligned}$$

it follows that

$$\frac{d}{dt}\lambda(t)\cdot\int_M f(t)^2\,dV_{g(t)} = -\int_M f(t)\Big(\frac{\partial}{\partial t}\Delta_{g(t)}\Big)f(t)\,dV_{g(t)}. \tag{7-4}$$

If we set $v_{ij} = -2R_{ij} + 4\partial_i u\partial_j u$,

$$\frac{\partial}{\partial t}\Gamma_{ij}^k = \tfrac{1}{2}g^{k\ell}(\partial_i v_{\ell j} + \partial_j v_{il} - \partial_\ell v_{ij}).$$

We temporarily omit all subscripts $t$. Multiplying with $g^{ij}$ on both sides, we obtain

$$\begin{aligned}g^{ij}\frac{\partial}{\partial t}\Gamma_{ij}^k &= \tfrac{1}{2}g^{kl}(2\nabla^i v_{li} - \nabla_l(g^{ij}v_{ij})) = g^{kl}\nabla^i v_{il} + \nabla^k S\\ &= g^{kl}\nabla^i(-2R_{il} + 4\nabla_i u\nabla_l u) + \nabla^k(R - 2|\nabla u|^2)\end{aligned}$$

$$
\begin{aligned}
&= -\nabla^k R + 4\Delta u \cdot \nabla^k u + 4\nabla_i u \cdot \nabla^i \nabla^k u + \nabla^k R - 4\nabla^k \nabla^i u \cdot \nabla_i u \\
&= 4\Delta u \cdot \nabla^k u.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\partial}{\partial t}(\Delta f) &= \frac{\partial}{\partial t}(g^{ij}\nabla_i \nabla_j f) \\
&= \Big(\frac{\partial}{\partial t} g^{ij}\Big)\nabla_i \nabla_j f + g^{ij}\Big[\partial_i \partial_j \frac{\partial f}{\partial t} - \Big(\frac{\partial}{\partial t}\Gamma^k_{ij}\Big)\partial_k f - \Gamma^k_{ij}\partial_k \frac{\partial f}{\partial t}\Big] \\
&= \Big(\frac{\partial}{\partial t} g^{ij}\Big)\nabla_i \nabla_j f + \Delta_{g(t)}\Big(\frac{\partial}{\partial t} f\Big) - g^{ij}\Big(\frac{\partial}{\partial t}\Gamma^k_{ij}\Big)\nabla_k f \\
&= (2R_{ij} - 4\nabla_i u \nabla_j u)\nabla^i \nabla^j f - 4\Delta u \cdot \nabla^k u \nabla_k f + \Delta_{g(t)}\Big(\frac{\partial}{\partial t} f\Big).
\end{aligned}
$$

Plugging this into (7-4), we derive

$$
\begin{aligned}
&\frac{d}{dt}\lambda(t) \cdot \int_M f(t)^2\, dV_{g(t)} \\
&\quad = -2\int_M R_{ij}\nabla^i \nabla^j f\, dV + 4\int_M f\nabla^i u \nabla^j u \nabla_i \nabla_j f\, dV + 4\int_M f\Delta u \cdot \nabla^k u \nabla_k f\, dV.
\end{aligned}
$$

The first term can be rewritten as

$$
\begin{aligned}
-2\int_M f R_{ij}\nabla^i \nabla^j f\, dV &= \int_M \nabla^i (2 f R_{ij})\nabla^j f\, dV \\
&= 2\int_M (\nabla^i f \cdot R_{ij} + f \cdot \nabla^i R_{ij})\nabla^j f\, dV \\
&= 2\int_M R_{ij}\nabla^i f \nabla^j f\, dV + \int_M f \nabla_j R \nabla^j f\, dV \\
&= 2\int_M R_{ij}\nabla^i f \nabla^j f\, dV - \int_M R\nabla_j (f\nabla^j f)\, dV \\
&= \lambda \int_R f^2\, dV - \int_M R|\nabla f|^2\, dV + 2\int_M R_{ij}\nabla^i f \nabla^j f\, dV.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\Big(\frac{d}{dt}\lambda(t)\Big)\int_M f(t)^2\, dV_{g(t)} \\
&= \lambda(t)\int_M R_{g(t)} f(t)^2\, dV_{g(t)} + 2\int_M R_{ij}\nabla^i f \nabla^j f\, dV - \int_M R_{g(t)}|\nabla_{g(t)} f(t)|^2_{g(t)}\, dV_{g(t)} \\
&\qquad + 4\int_M f(\nabla^i u \nabla^j u \nabla_i \nabla_j f + \Delta u \nabla^k u \nabla_k f)\, dV.
\end{aligned}
$$

On the other hand,

$$\begin{aligned}
&\int_M f\nabla^i u\nabla^j u\nabla_i\nabla_j f\,dV\\
&\quad=-\int_M \nabla_i(f\nabla^i u\nabla^j u)\nabla_j f\,dV\\
&\quad=-\int_M (\nabla_i f\nabla^i u\nabla^j u+f\Delta u\nabla^j u+f\nabla^i u\nabla_i\nabla^j u)\nabla_j f\,dV\\
&\quad=-\int_M f\Delta u\langle\nabla u,\nabla f\rangle\,dV-\int_M \nabla^i u\nabla^j u\nabla_i f\nabla_j f\,dV-\int_M f\nabla^i u\nabla^j f\nabla_i\nabla_j u\,dV
\end{aligned}$$

and therefore

$$\begin{aligned}
\frac{d}{dt}\lambda(t)\int_M f(t)^2 dV_{g(t)}=\lambda(t)\int_M R_{g(t)}f(t)^2 dV_{g(t)}-4\int_M f\nabla^i u\nabla^j f\nabla_i\nabla_j u\,dV\\
+2\int_M S_{ij}\nabla^i f\nabla_j f\,dV-\int_M R_{g(t)}|\nabla_{g(t)}f(t)|^2_{g(t)}\,dV_{g(t)}.
\end{aligned}$$

The last term here can be simplified as follows:

$$\begin{aligned}
-\int_M & f\nabla^i u\nabla^j f\nabla_i\nabla_j u\,dV\\
&=\int_M \nabla^j(f\nabla_i u\nabla_j f)\nabla^i u\,dV\\
&=\int_M (\nabla^j f\nabla_i u\nabla_j f+f\nabla^j\nabla_i u\nabla_j f+f\nabla_i u\Delta f)\nabla^i u\,dV\\
&=\int_M |\nabla u|^2|\nabla f|^2\,dV+\int_M f\Delta f|\nabla u|^2\,dV+\int_M f\nabla^i u\nabla^j f\nabla_i\nabla_j u\,dV.
\end{aligned}$$

Consequently,

$$-2\int_M f\nabla^i u\nabla^j f\nabla_i\nabla_j u\,dV=\int_M |\nabla u|^2|\nabla f|^2\,dV-\lambda\int_M f^2|\nabla u|^2\,dV.$$

Therefore we derive the following.

**Theorem 7.1.** *If $(g(t),u(t))$ is a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and $\lambda(t)$ denotes the eigenvalue of the Laplacian $\Delta_{g(t)}$, then*

$$\begin{aligned}
\text{(7-5)}\quad \frac{d}{dt}\lambda(t)\cdot\int_M f(t)^2\,dV_{g(t)}&\\
=\lambda(t)\int_M S_{g(t),u(t)}f(t)^2\,dV_{g(t)}&-\int_M S_{g(t),u(t)}|\nabla_{g(t)}f(t)|^2_{g(t)}\,dV_{g(t)}\\
&+2\int_M \langle\mathcal{S}_{g(t),u(t)},df(t)\otimes df(t)\rangle\,dV_{g(t)}.
\end{aligned}$$

We set

$$S_{\min}(0) := \min_{x\in M} S(x,0). \tag{7-6}$$

**Theorem 7.2.** *Let $(g(t), u(t))_{t\in[0,T]}$ be a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and let $\lambda(t)$ denote the eigenvalue of the Laplacian $\Delta_{g(t)}$. Suppose that $\mathcal{S}_{g(t),u(t)} - \alpha S_{g(t),u(t)} g(t) \geq 0$ along the harmonic-Ricci flow for some $\alpha \geq \frac{1}{2}$.*

(1) *If $S_{\min}(0) \geq 0$, $\lambda(t)$ is nondecreasing along the harmonic-Ricci flow for any $t \in [0, T]$.*

(2) *If $S_{\min}(0) > 0$, the quantity*

$$\left(1 - \frac{2}{n} S_{\min}(0)\, t\right)^{n\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for $T \leq n/(2S_{\min}(0))$.*

(3) *If $S_{\min}(0) < 0$, the quantity*

$$\left(1 - \frac{2}{n} S_{\min}(0)\, t\right)^{n\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for any $t \in [0, T]$.*

*Proof.* By Theorem 7.1, we have

$$\frac{d}{dt}\lambda(t) \geq \frac{\int_M S_{g(t),u(t)} f(t)^2\, dV_{g(t)}}{\int_M f(t)^2\, dV_{g(t)}} \lambda(t) + (2\alpha - 1) \frac{\int_M S_{g(t),u(t)} |\nabla_{g(t)} f(t)|^2_{g(t)}}{\int_M f(t)^2\, dV_{g(t)}}.$$

By definition we have $-f(t)\Delta_{g(t)} = \lambda(t) f(t)$. Integrating both sides yields that $\lambda(t) \geq 0$. Since

$$\frac{\partial}{\partial t} S_{g(t),u(t)} = \Delta_{g(t)} S_{g(t),u(t)} + 2|\mathcal{S}_{g(t),u(t)}|^2_{g(t)} + 4|\Delta_{g(t)} u(t)|^2_{g(t)}$$

and $|\mathcal{S}_{g(t),u(t)}|^2 \geq (1/n) S^2_{g(t),u(t)}$, it follows that

$$\frac{\partial}{\partial t} S_{g(t),u(t)} \geq \Delta_{g(t)} S_{g(t),u(t)} + \frac{2}{n} S^2_{g(t),u(t)}.$$

The corresponding ODE

$$\frac{d}{dt} a(t) = \frac{2}{n} a(t)^2, \quad a(t) = S_{\min}(0)$$

has the solution

$$a(t) = \frac{S_{\min}(0)}{1 - (2/n) S_{\min}(0)\, t}.$$

Then the maximum principle implies $S_{g(t),u(t)} \geq a(t)$ and hence, using the assumption that $2\alpha - 1 \geq 0$,

$$\frac{d}{dt}\lambda(t) \geq a(t)\lambda(t) + (2\alpha - 1)a(t)\frac{\int_M |\nabla_{g(t)} f(t)|^2_{g(t)}\, dV_{g(t)}}{\int_M f(t)^2\, dV_{g(t)}}.$$

By integration by parts, we note that

$$\int_M |\nabla f|^2\, dV = -\int_M f \cdot \Delta f\, dV = \lambda \int_M f^2\, dV,$$

which shows that

$$\frac{d}{dt}\lambda(t) \geq a(t)\lambda(t) + (2\alpha - 1)a(t)\lambda = 2\alpha a(t)\lambda(t)$$

and

$$\frac{d}{dt}\left(\lambda(t) \cdot \exp\left(-2\alpha \int_0^t a(\tau)\, d\tau\right)\right) \geq 0.$$

This inequality clearly implies the desired result. If $S_{\min}(0) \geq 0$, by the nonnegativity of $\mathcal{S}_{g(t)}$ preserved along the harmonic-Ricci flow, we conclude that $d\lambda(t)/dt \geq 0$. □

**Corollary 7.3.** *Let $(g(t), u(t))_{t\in[0,T]}$ be a solution of the harmonic-Ricci flow on a compact Riemannian surface $\Sigma$ and let $\lambda(t)$ denote the eigenvalue of the Laplacian $\Delta_{g(t)}$.*

(1) *Suppose that* $\mathrm{Ric}_{g(t)} \leq \epsilon du(t) \otimes du(t)$ *where*

(7-7) $$\epsilon \leq 4\frac{1-\alpha}{1-2\alpha}, \quad \alpha > \tfrac{1}{2}.$$

- (i) *If $S_{\min}(0) \geq 0$, $\lambda(t)$ is nondecreasing along the harmonic-Ricci flow for any $t \in [0, T]$.*
- (ii) *If $S_{\min}(0) > 0$, the quantity*

$$(1 - S_{\min}(0)\, t)^{2\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for $T \leq 1/S_{\min}(0)$.*

- (iii) *If $S_{\min}(0) < 0$, the quantity*

$$(1 - S_{\min}(0)\, t)^{2\alpha} \lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for any $t \in [0, T]$.*

(2) *Suppose that*

(7-8) $$|\nabla_{g(t)} u(t)|^2_{g(t)}\, g(t) \geq 2du(t) \otimes du(t).$$

- (i) *If $S_{\min}(0) \geq 0$, $\lambda(t)$ is nondecreasing along the harmonic-Ricci flow for any $t \in [0, T]$.*

(ii) *If* $S_{\min}(0) > 0$, *the quantity*

$$(1 - S_{\min}(0)t)\lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for* $T \leq 1/S_{\min}(0)$.

(iii) *If* $S_{\min}(0) < 0$, *the quantity*

$$(1 - S_{\min}(0)t)\lambda(t)$$

*is nondecreasing along the harmonic-Ricci flow for any* $t \in [0, T]$.

*Proof.* As above, we always omit subscripts $t$. In the surface case, we have $R_{ij} = \frac{1}{2}Rg_{ij}$. Then

$$\begin{aligned} T_{ij} := S_{ij} - \alpha S g_{ij} &= \frac{R}{2}g_{ij} - 2\nabla_i u \nabla_j u - \alpha(R - 2|\nabla u|^2)g_{ij} \\ &= \left(\tfrac{1}{2} - \alpha\right)Rg_{ij} - 2\nabla_i u \nabla_j u + 2\alpha|\nabla u|^2 g_{ij}. \end{aligned}$$

For any vector $V = (V^i)$, we calculate

$$\begin{aligned} T_{ij}V^iV^j &= \left(\tfrac{1}{2} - \alpha\right)R|V|^2 - 2(\nabla_i u V^i)^2 + 2\alpha|\nabla u|^2|V|^2 \\ &\geq \left(\tfrac{1}{2} - \alpha\right)R|V|^2 - 2|\nabla u|^2|V|^2 + 2\alpha|\nabla u|^2|V|^2. \end{aligned}$$

If $R_{ij} \leq \epsilon \nabla_i u \nabla_j u$, then $T_{ij}V^iV^j = [(\frac{1}{2} - \alpha)\epsilon - 2 + 2\alpha]\,|\nabla u|^2|V|^2 \geq 0$.

For the second case, we note that

$$\begin{aligned} T_{ij}V^iV^j &= R_{ij}V^iV^j - 2\nabla_i u V^i \nabla_j u V^j - \frac{R}{2}|V|^2 + |\nabla u|^2|V|^2 \\ &\geq R_{ij}V^iV^j - |\nabla u|^2|V|^2 - \frac{R}{2}|V|^2 + |\nabla u|^2|V|^2 = 0. \end{aligned}$$

Hence the corresponding results follow by Theorem 7.2. □

When we consider the Ricci flow, we have the following two results derived from Corollary 7.3.

**Corollary 7.4.** *Let* $(g(t))_{t\in[0,T]}$ *be a solution of the Ricci flow on a compact Riemannian surface* $\Sigma$ *and let* $\lambda(t)$ *denote the eigenvalue of the Laplacian* $\Delta_{g(t)}$.

(1) *If* $R_{\min}(0) \geq 0$, $\lambda(t)$ *is nondecreasing along the Ricci flow for any* $t \in [0, T]$.

(2) *If* $R_{\min}(0) > 0$, *the quantity* $(1 - R_{\min}(0)t)\lambda(t)$ *is nondecreasing along the Ricci flow for* $T \leq 1/R_{\min}(0)$.

(3) *If* $R_{\min}(0) < 0$, *the quantity* $(1 - R_{\min}(0)t)\lambda(t)$ *is nondecreasing along the Ricci flow for any* $t \in [0, T]$.

**Remark 7.5.** Let $(g(t))_{t\in[0,T]}$ be a solution of the Ricci flow on a compact Riemannian surface $\Sigma$ with nonnegative scalar curvature and let $\lambda(t)$ denote the eigenvalue of the Laplacian $\Delta_{g(t)}$. Then $\lambda(t)$ is nondecreasing along the Ricci flow for $t \in [0, T]$.

## 8. Eigenvalues of the Laplacian-type under the harmonic-Ricci flow

Recall that

$$\mu(g, u) = \mu_1(g, u) = \inf\left\{\mathscr{F}(g, u, f) \;\middle|\; \int_M e^{-f}\, dV_g = 1\right\}. \tag{8-1}$$

We showed that $\mu(g, u)$ is the smallest eigenvalue of the operator

$$-4\Delta_g + R_g - 2|\nabla_g u|_g^2.$$

Inspired by [Cao 2007; 2008], we define a Laplacian-type operators associated with quantities $g, u, c$:

$$\Delta_{g,u,c} := -\Delta_g + c(R_g - 2|\nabla_g u|_g^2), \tag{8-2}$$

$$\Delta_{g,u} := \Delta_{g,u,\frac{1}{2}} = -\Delta_g + \tfrac{1}{2}(R_g - 2|\nabla_g u|_g^2). \tag{8-3}$$

Then $\mu(g, u)$ is the smallest eigenvalue of the operator $4\Delta_{g,u,1/4}$.

To the operator $\Delta_{g,u}$ we associate the functional

$$C^\infty(M) \to \mathbb{R}, \quad f \mapsto \lambda_{g,u}(f) := \int_M f\,\Delta_{g,u} f\, dV_g. \tag{8-4}$$

When $f$ is an eigenfunction of the operator $\Delta_{g,u}$ with the eigenvalue $\lambda$, that is, $\Delta_{g,u} f = \lambda f$ and is normalized by $\int_M f^2\, dV_g = 1$, we obtain $\lambda_{g,u}(f) = \lambda$. The next lemma will deal with the evolution equation for $\lambda(f(t))$, where $f(t)$ is an eigenfunction of $\Delta_{g(t),u(t)}$ and the couple $(g(t), u(t))$ satisfies the harmonic-Ricci flow. Set

$$v_{ij} := -2S_{ij} = -2R_{ij} + 4\partial_i u \cdot \partial_j u, \quad v := g^{ij} v_{ij}. \tag{8-5}$$

The symmetric tensor field thus obtained is denoted by $\mathscr{V}_{g(t),u(t)} = (v_{ij})$.

**Lemma 8.1.** *Suppose that $(g(t), u(t))$ is a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and $f(t)$ is an eigenfunction of $\Delta_{g(t),u(t)}$, that is, $\Delta_{g(t),u(t)} f(t) = \lambda(t) f(t)$ (where $\lambda(t)$ is only a function of time $t$ only), with the normalized condition*

$$\int_M f(t)^2\, dV_{g(t)} = 1.$$

*Then we have*

$$
\begin{aligned}
(8\text{-}6)\quad & \frac{d}{dt}\lambda_{g(t),u(t)}(f(t) \\
&= \int_M f(t)\big(\nabla^i v_{ik} - \tfrac{1}{2}\nabla_k v\big)\nabla^k f(t)\,dV_{g(t)} - \int_M f^2(t)\frac{\partial}{\partial t}|\nabla_{g(t)}u(t)|^2_{g(t)}\,dV_{g(t)} \\
&\qquad + \int_M\bigg(\langle \mathcal{V}_{g(t),u(t)}, \nabla^2_{g(t)} f(t)\rangle_{g(t)} + \frac{1}{2}\Big(\frac{\partial}{\partial t}R_{g(t)}\Big)f(t)\bigg)f(t)\,dV_{g(t)}.
\end{aligned}
$$

Before proving the lemma, we recall a formula that is an immediate consequence of the evolution equation:

$$
\begin{aligned}
(8\text{-}7)\quad & \frac{\partial}{\partial t}(\Delta_{g(t)} f) \\
&= -g^{ip}g^{jq}v_{pq}\nabla_i\nabla_j f - g^{ij}g^{k\ell}\nabla_i v_{j\ell}\nabla_k f + \tfrac{1}{2}\langle \nabla_{g(t)}v_{g(t)}, \nabla_{g(t)} f(t)\rangle_{g(t)}
\end{aligned}
$$

where the metric $g(t)$ evolves by $\partial g_{ij}/\partial t = v_{ij}$.

*Proof.* Using (8-7) and integration by parts, we get

$$
\begin{aligned}
&\frac{d}{dt}\lambda_{g(t),u(t)}(f(t)) \\
&= \frac{\partial}{\partial t}\int_M\bigg(-\Delta_{g(t)} f(t) + \bigg(\frac{R_{g(t)}}{2} - |\nabla_{g(t)}u(t)|^2_{g(t)}\bigg)f(t)\bigg)f(t)\,dV_{g(t)} \\
&= \int_M\big(g^{ip}g^{jq}v_{pq}\nabla_i\nabla_j f + g^{ij}g^{kl}\nabla_i v_{jl}\nabla_k f - \tfrac{1}{2}\langle\nabla_{g(t)}v_{g(t)},\nabla_{g(t)} f(t)\rangle_{g(t)}\big)f(t)\,dV_{g(t)} \\
&\quad + \int_M\bigg(-\Delta_{g(t)}\big(\frac{\partial}{\partial t}f(t)\big) + \bigg(\frac{R_{g(t)}}{2} - |\nabla_{g(t)}u(t)|^2_{g(t)}\bigg)\frac{\partial}{\partial t}f(t) \\
&\quad + \bigg(\frac{\partial}{\partial t}(\tfrac{1}{2}R_{g(t)}) - \frac{\partial}{\partial t}(|\nabla_{g(t)}u(t)|^2_{g(t)})\bigg)f(t)\bigg)f(t)\,dV_{g(t)} \\
&\qquad + \int_M\bigg(-\Delta_{g(t)} f(t) + \bigg(\frac{R_{g(t)}}{2} - |\nabla_{g(t)}u|^2_{g(t)}\bigg)f(t)\bigg)\frac{\partial}{\partial t}(f(t)\,dV_{g(t)}) \\
&= \int_M\bigg(g^{ip}g^{jq}v_{pq}\nabla_i\nabla_j f + \frac{1}{2}\Big(\frac{\partial}{\partial t}R_{g(t)}\Big)f(t)\bigg)f(t)\,dV_{g(t)} \\
&\quad + \int_M(g^{ij}g^{kl}\nabla_i v_{jl}\nabla_k f - \tfrac{1}{2}g^{kl}\nabla_l v\nabla_k f)f(t)\,dV_{g(t)} \\
&\quad + \int_M \Delta_{g(t),u(t)} f(t)\bigg(\frac{\partial}{\partial t}f(t)\,dV_{g(t)} + \frac{\partial}{\partial t}(f(t)\,dV_{g(t)})\bigg) \\
&\qquad\qquad - \int_M\frac{\partial}{\partial t}(|\nabla_{g(t)}u(t)|^2_{g(t)})f(t)^2\,dV_{g(t)}.
\end{aligned}
$$

Since $f(t)$ is an eigenfunction of $\Delta_{g(t),u(t)}$, it follows that

$$
\int_M \Delta_{g(t),u(t)} f(t)\bigg(\frac{\partial}{\partial t}f(t)\,dV_{g(t)} + \frac{\partial}{\partial t}(f(t)\,dV_{g(t)})\bigg) = \lambda(t)\frac{\partial}{\partial t}\int_M f(t)^2\,dV_{g(t)} = 0
$$

by the normalization condition. This completes the proof. □

Using (3-6), we find that the first term in the right side of (8-6) can be written as

$$
\begin{aligned}
&\int_M \Big(v_{ij}\nabla^i\nabla^j f+\frac{1}{2}\Big(\frac{\partial}{\partial t}R_{g(t)}\Big)f(t)\Big)f(t)\,dV_{g(t)}\\
&=\int_M \big(-2f(t)\langle \mathrm{Ric}_{g(t)},\nabla^2_{g(t)}f(t)\rangle_{g(t)}+4f(t)\langle du(t)\otimes du(t),\nabla^2_{g(t)}f(t)\rangle_{g(t)}\big)\,dV_{g(t)}\\
&\qquad+\int_M \big((\tfrac{1}{2}\Delta_{g(t)}R_{g(t)}+|\mathrm{Ric}_{g(t)}|^2_{g(t)})f(t)^2+2f(t)^2|\Delta_{g(t)}u(t)|^2_{g(t)}\\
&\qquad\qquad-2f(t)^2|\nabla^2_{g(t)}u(t)|^2_{g(t)}-4f(t)^2\langle \mathrm{Ric}_{g(t)},du(t)\otimes du(t)\rangle_{g(t)}\big)\,dV_{g(t)}\\
&=\int_M \big(-2f(t)\langle \mathrm{Ric}_{g(t)},\nabla^2_{g(t)}f(t)\rangle_{g(t)}+(\tfrac{1}{2}\Delta_{g(t)}R_{g(t)}+|\mathrm{Ric}_{g(t)}|^2_{g(t)})f(t)^2\big)\,dV_{g(t)}\\
&\qquad+\int_M \big(4f(t)\langle du\otimes du,\nabla^2_{g(t)}f(t)\rangle_{g(t)}-4f^2\langle du(t)\otimes du(t),\mathrm{Ric}_{g(t)}\rangle_{g(t)}\\
&\qquad\qquad+2f(t)^2|\Delta_{g(t)}u(t)|^2_{g(t)}-2f(t)^2|\nabla^2_{g(t)}u(t)|^2_{g(t)}\big)\,dV_{g(t)}
\end{aligned}
$$

For the second term in (8-6), using the contracted Bianchi identities, one has

$$
\begin{aligned}
&\int_M (g^{ij}\nabla_i v_{jk}-\tfrac{1}{2}\nabla_k v)\nabla^k f\cdot f(t)\,dV_{g(t)}\\
&\qquad=\int_M \big(g^{ij}\nabla_i(-2R_{jk}+4\partial_j u\partial_k u)\\
&\qquad\qquad-\tfrac{1}{2}\nabla_k(-2R_{g(t)}+4|\nabla_{g(t)}u(t)|^2_{g(t)})\big)\nabla^k f\cdot f(t)\,dV_{g(t)}\\
&\qquad=\int_M 4f(t)\Delta_{g(t)}u(t)\langle\nabla_{g(t)}u(t),\,\nabla_{g(t)}f(t)\rangle_{g(t)}\,dV_{g(t)}\\
&\qquad\qquad+\int_M (4g^{ij}\nabla_j u\cdot\nabla_i\nabla_k u-2\nabla_k|\nabla_{g(t)}u(t)|^2_{g(t)})\nabla^k f\cdot f(t)\,dV_{g(t)}\\
&\qquad=\int_M 4f(t)\Delta_{g(t)}u(t)\langle\nabla_{g(t)}u(t),\,\nabla_{g(t)}f(t)\rangle_{g(t)}\,dV_{g(t)}
\end{aligned}
$$

where in the last step we use the identity $\nabla_k|\nabla u|^2=2g^{pq}\nabla_k\nabla_p u\cdot\nabla_q u$. Therefore

$$
\begin{aligned}
&\frac{d}{dt}\lambda_{g(t),u(t)}(f(t)) \qquad (8\text{-}8)\\
&=\int_M \big(-2f(t)\langle \mathrm{Ric}_{g(t)},\nabla^2_{g(t)}f(t)\rangle_{g(t)}+(\tfrac{1}{2}\Delta_{g(t)}R_{g(t)}+|\mathrm{Ric}_{g(t)}|^2_{g(t)})f(t)^2\big)\,dV_{g(t)}\\
&\qquad+4f(t)\int_M \big(\langle du(t)\otimes du(t),\,\nabla^2_{g(t)}f(t)\rangle_{g(t)}-f(t)\langle du(t)\otimes du(t),\mathrm{Ric}_{g(t)}\rangle_{g(t)}\\
&\qquad\qquad+2f(t)^2|\Delta_{g(u)}u(t)|^2_{g(t)}-2f(t)^2|\nabla^2_{g(t)}u(t)|^2_{g(t)}\\
&\qquad\qquad+4f(t)\Delta_{g(t)}u(t)\langle\nabla_{g(t)}u(t),\,\nabla_{g(t)}f(t)\rangle_{g(t)}\big)\,dV_{g(t)}\\
&\qquad-\int_M (\Delta_{g(t)}|\nabla_{g(t)}u(t)|^2_{g(t)}-2|\nabla^2_{g(t)}u(t)|^2_{g(t)}-4|\nabla_{g(t)}u(t)|^4_{g(t)})f(t)^2\,dV_{g(t)}.
\end{aligned}
$$

The above evolution equation can be simplified as follows.

**Theorem 8.2.** *Suppose* $(g(t), u(t))$ *is a solution of the harmonic-Ricci flow on a compact Riemannian manifold* $M$ *and* $f(t)$ *is an eigenfunction of* $\Delta_{g(t),u(t)}$, *that is*, $\Delta_{g(t),u(t)} f(t) = \lambda(t) f(t)$ (*where* $\lambda(t)$ *is only a function of time* $t$ *only*), *with the normalized condition* $\int_M f(t)^2 dV_{g(t)} = 1$. *Then we have*

$$\frac{d}{dt}\lambda_{g(t),u(t)}(f(t)) = \int_M 2\langle \mathcal{S}_{g(t)}, df(t) \otimes df(t)\rangle_{g(t)}\, dV_{g(t)} + \int_M f(t)^2\big(|\mathcal{S}_{g(t)}|^2_{g(t)} + 2|\Delta_{g(t)} u(t)|^2_{g(t)}\big)\, dV_{g(t)}. \tag{8-9}$$

*Proof.* Calculate

$$\begin{aligned}
&\int_M 4 f(t)\Delta_{g(t)} u(t)\langle \nabla_{g(t)} u(t), \nabla_{g(t)} f(t)\rangle_{g(t)}\, dV_{g(t)} \\
&\qquad = -4\int_M \nabla_i u[\nabla^i f \cdot \langle \nabla u, \nabla f\rangle + f(\nabla^i\langle \nabla u, \nabla f\rangle)]\, dV \\
&\qquad = -4\int_M |\langle \nabla u, \nabla f\rangle|^2\, dV_g - 4\int_M f\nabla_i u(\langle \nabla^i\nabla u, \nabla f\rangle + \langle \nabla u, \nabla^i\nabla f)\, dV.
\end{aligned}$$

By the same method, we have

$$\begin{aligned}
&\int_M -\Delta_{g(t)}|\nabla_{g(t)} u(t)|^2_{g(t)} f(t)^2\, dV_{g(t)} \\
&\qquad = -\int_M |\nabla u|^2(2f\Delta f + 2|\nabla f|^2)\, dV \\
&\qquad = -2\int_M |\nabla f|^2|\nabla u|^2\, dV - 2\int_M f\Delta f|\nabla u|^2\, dV.
\end{aligned}$$

However,

$$\begin{aligned}
\int_M f\Delta f|\nabla u|^2\, dV &= \int_M -\nabla_i f\cdot\nabla^i(f|\nabla u|^2)\, dV \\
&= -\int_M \nabla_i f(\nabla^i f|\nabla u|^2 + f\nabla^i|\nabla u|^2)\, dV \\
&= -\int_M |\nabla u|^2|\nabla f|^2\, dV - \int_M f\nabla_i f\cdot\nabla^i|\nabla u|^2\, dV.
\end{aligned}$$

Therefore we arrive at

$$\begin{aligned}
&\int_M -\Delta_{g(t)}|\nabla_{g(t)} u(t)|^2_{g(t)} f(t)^2\, dV_{g(t)} \\
&\qquad = 2\int_M f\nabla_i f\cdot\nabla^i|\nabla u|^2\, dV \\
&\qquad = 4\int_M f(t)\langle du(t)\otimes df(t), \nabla^2_{g(t)} u(t)\rangle_{g(t)}\, dV_{g(t)}.
\end{aligned}$$

Using the contracted Bianchi identities, we may simplify the term $\int_M \frac{1}{2} f^2 \Delta R\, dV$ as follows:

$$\begin{aligned}
&\int_N \frac{f(t)^2}{2} \Delta_{g(t)} R_{g(t)}\, dV_{g(t)} \\
&= -\frac{1}{2} \int_M \nabla_i R \cdot \nabla^i (f^2)\, dV \\
&= -\int_M \nabla_i R \cdot f \nabla^i f\, dV = -2 \int_M \nabla^k R_{ki} \cdot f \nabla^i f\, dV \\
&= 2 \int_M R_{ki} \nabla^k (f \nabla^j f)\, dV = 2 \int_M R_{ki} (\nabla^k f \cdot \nabla^j f + f \nabla^k \nabla^j f)\, dV \\
&= 2 \int_M \langle \mathrm{Ric}_{g(t)}, df(t) \otimes df(t) \rangle_{g(t)}\, dV_{g(t)} + 2 \int_M f(t) \langle \mathrm{Ric}_{g(t)}, \nabla^2_{g(t)} f(t) \rangle_{g(t)}\, dV_{g(t)}.
\end{aligned}$$

Hence (8-8) becomes

$$\begin{aligned}
&\frac{d}{dt} \lambda_{g(t),u(t)}(f(t)) \\
&\quad = \int_M \big( 2 \langle \mathrm{Ric}_{g(t)}, df(t) \otimes df(t) \rangle_{g(t)} + |\mathrm{Ric}_{g(t)}|^2_{g(t)} f(t)^2 \big)\, dV_{g(t)} \\
&\qquad + \int_M \big( 2 |\Delta_{g(t)} u(t)|^2_{g(t)} + 4 |\nabla_{g(t)} u(t)|^4_{g(t)} \big) f(t)^2\, dV_{g(t)} \\
&\qquad - \int_M 4 f(t)^2 \langle du(t) \otimes du(t), \mathrm{Ric}_{g(t)} \rangle_{g(t)}\, dV_{g(t)} \\
&\qquad - \int_M 4 |\langle \nabla_{g(t)} u(t), \nabla_{g(t)} f(t) \rangle_{g(t)}|^2\, dV_{g(t)} \\
&\quad = \int_M 2 \langle \mathcal{S}_{g(t)}, df(t) \otimes df(t) \rangle_{g(t)}\, dV_{g(t)} \\
&\qquad\quad + \int_M f(t)^2 \big( |\mathrm{Ric}_{g(t)} - 2 du(t) \otimes du(t)|^2_{g(t)} + 2 |\Delta_{g(t)} u(t)|^2_{g(t)} \big)\, dV_{g(t)}
\end{aligned}$$

where, by definition, $S_{ij} = R_{ij} - 2 \partial_i u \partial_j u$. □

List [2006] proved that the nonnegativity of the operator $\mathcal{S}_{g(t)}$ is preserved by the harmonic-Ricci flow. Hence we get the following.

**Corollary 8.3.** *If* $\mathrm{Ric}_{g(0)} - 2du(0) \otimes du(0) \geq 0$, *the eigenvalues of the operator* $\Delta_{g(t),u(t)}$ *are nondecreasing under the harmonic-Ricci flow.*

**Remark 8.4.** If we choose $u(t) \equiv 0$, we obtain Cao's result [2007].

## 9. Another formula for $\frac{d}{dt}\lambda(f(t))$

In this section we give another formula for $\frac{d}{dt}\lambda(f(t))$ using a method similar to that in [Li 2007]. Recall the formula

$$\frac{d}{dt}\lambda_{g(t),u(t)}(f(t)) = \int_M 2\langle \mathcal{S}_{g(t),u(t)}, df(t)\otimes df(t)\rangle_{g(t)}\, dV_{g(t)} + \int_M f(t)^2\big(|\mathcal{S}_{g(t),u(t)}|^2_{g(t)} + 2|\Delta_{g(t)}u(t)|^2_{g(t)}\big)\, dV_{g(t)}.$$

Consider the function $\varphi$ determined by $f^2(t) = e^{-\varphi(t)}$. Then we have

$$df = \frac{-e^{\varphi}d\varphi}{2f}, \qquad \frac{\nabla f}{f} = -\frac{\nabla\varphi}{2}, \qquad \frac{\Delta f}{f} = -\tfrac{1}{2}\Delta\varphi + \tfrac{1}{4}|\nabla\varphi|^2.$$

Hence

$$2\frac{d}{dt}\lambda_{g(t),u(t)}(f(t)) = \int_M \langle \mathcal{S}_{g(t),u(t)}, d\varphi(t)\otimes d\varphi(t)\rangle_{g(t)} e^{-\varphi(t)}\, dV_{g(t)} + 2\int_M \big(|\mathcal{S}_{g(t),u(t)}|^2_{g(t)} + 2|\Delta_{g(t)}u(t)|^2_{g(t)}\big)e^{-\varphi}\, dV_{g(t)}.$$

Using integration by parts and contracted Bianchi identities yields

$$\begin{aligned}
&\int_M \langle \mathcal{S}_{g(t),u(t)}, d\varphi(t)\otimes d\varphi(t)\rangle_{g(t)} e^{-\varphi(t)}\, dV_{g(t)}\\
&\quad = \int_M S_{ij}\nabla^i\varphi\nabla^j\varphi e^{-\varphi}\, dV = -\int_M S_{ij}\nabla^j\varphi\nabla^i(e^{-\varphi})\, dV\\
&\quad = \int_M e^{-\varphi}\nabla^i(S_{ij}\nabla^j\varphi)\, dV\\
&\quad = \int_M \nabla^i S_{ij}\cdot\nabla^j\varphi\cdot e^{-\varphi}\, dV + \int_M S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\, dV\\
&\quad = \int_M \nabla^i R_{ij}\cdot\nabla^j\varphi\cdot e^{-\varphi}\, dV_g + \int_M S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\, dV\\
&\qquad\qquad + \int_M \nabla^i(-2\nabla_i u\nabla_j u)\nabla^j\varphi\cdot e^{-\varphi}\, dV_g\\
&\quad = \frac{1}{2}\int_M R\Delta(e^{-\varphi})\, dV + \int_M S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\, dV - 2\int_M (\nabla^i u\nabla_j u)\nabla^i\nabla^j(e^{-\varphi})\, dV.
\end{aligned}$$

Thus

$$\begin{aligned}
&\int_M S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\, dV\\
&\qquad = \int_M S_{ij}\nabla^i\varphi\nabla^j\varphi e^{-\varphi}\, dV - \frac{1}{2}\int_M R\Delta(e^{-\varphi})\, dV + 2\int_M (\nabla^i u\nabla_j u)\nabla^i\nabla^j(e^{-\varphi}).
\end{aligned}$$

On the other hand, one gets

$$\begin{aligned}\int_M |\nabla^2_{g(t)}\varphi(t)|^2_{g(t)} e^{-\varphi(t)}\, dV_{g(t)} &= \int_M \nabla_i\nabla_j\varphi\nabla^i\nabla_j\varphi\cdot e^{-\varphi}\, dV\\ &= -\int_M \nabla_j\varphi\cdot\nabla_i\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\, dV - \int_M \nabla_j\varphi\cdot\nabla^i\nabla^j\varphi\cdot\nabla_i(e^{-\varphi})\, dV\\ &= -\int_M \nabla_j\varphi\cdot\nabla_i\nabla^j\nabla^i\varphi\cdot e^{-\varphi}\, dV - \int_M \nabla_j\varphi\cdot\nabla^i\nabla^j\varphi\cdot\nabla_i(e^{-\varphi})\, dV.\end{aligned}$$

Since

$$\begin{aligned}\int_M \nabla_j\varphi\cdot\nabla^i\nabla^j\varphi\cdot\nabla_i(e^{-\varphi})\, dV &= -\int_M \nabla^i(\nabla_j\varphi\cdot\nabla_i(e^{-\varphi}))\nabla^j\varphi\, dV\\ &= -\int_M \nabla^j\varphi\cdot\nabla^i\nabla_j\varphi\cdot\nabla_i(e^{-\varphi})\, dV - \int_M |\nabla\varphi|^2\Delta(e^{-\varphi})\, dV,\end{aligned}$$

which implies

$$\int_M \nabla_j\varphi\cdot\nabla^i\nabla^j\varphi\cdot\nabla_i(e^{-\varphi})dV = -\frac{1}{2}\int_M |\nabla\varphi|^2\Delta(e^{-\varphi})\, dV,$$

it follows that

$$\int_M |\nabla^2\varphi|^2 e^{-\varphi}\, dV = -\int_M \nabla_j\varphi\cdot\nabla_i\nabla^j\nabla^i\varphi\cdot e^{-\varphi}\, dV + \frac{1}{2}\int_M |\nabla\varphi|^2\Delta(e^{-\varphi})\, dV.$$

By the Ricci identity the term $\nabla^i\nabla^j\nabla^i\varphi$ equals

$$\begin{aligned}\nabla_i\nabla^j\nabla^i\varphi &= g^{jk}g^{il}\nabla_i\nabla_k\nabla_l\varphi = g^{jk}g^{il}(\nabla_k\nabla_i\nabla_l\varphi - R^p_{ikl}\nabla_p\varphi)\\ &= \nabla^j\nabla_i\nabla^i\varphi - g^{jk}g^{il}R_{iklp}\nabla^p\varphi\\ &= \nabla^j\Delta\varphi + g^{jk}g^{il}R_{ikpl}\nabla^p\varphi = \nabla^j\Delta\varphi + g^{jk}R_{kp}\nabla^p\varphi.\end{aligned}$$

Hence

$$\begin{aligned}&-\int_M \nabla_j\varphi\cdot\nabla_i\nabla^j\nabla^i\varphi\cdot e^{-\varphi}\, dV\\ &\quad= -\int_M \nabla_i\varphi\cdot\nabla^j\Delta\varphi\cdot e^{-\varphi}\, dV - \int_M R_{kp}\nabla^k\varphi\cdot\nabla^p\varphi e^{-\varphi}\, dV\\ &\quad= \int_M \nabla^j\Delta\varphi\cdot\nabla_j(e^{-\varphi}) + \int_M R_{kp}\nabla^k\varphi\cdot\nabla^p(e^{-\varphi})\, dV\\ &\quad= -\int_M \Delta\varphi\cdot\Delta(e^{-\varphi}) - \int_M e^{-\varphi}(\nabla^p R_{kp}\cdot\nabla^k\varphi + R_{kp}\nabla^p\nabla^k\varphi)\\ &\quad= -\int_M \Delta(e^{-\varphi})\cdot\Delta\varphi\, dV + \frac{1}{2}\int_M \nabla_k R\cdot\nabla^k(e^{-\varphi})\, dV - \int_M e^{-\varphi}R_{kp}\nabla^k\nabla^p\varphi\, dV\\ &\quad= -\int_M \Delta(e^{-\varphi})(\Delta\varphi + \tfrac{1}{2}R) - \int_M R_{kp}\nabla^k\nabla^p\varphi\cdot e^{-\varphi}\, dV.\end{aligned}$$

Putting those formulas together, we obtain

$$\begin{aligned}
&\int_M 2S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\,dV+\int_M|\nabla^2\varphi|^2e^{-\varphi}\,dV\\
&\qquad=\int_M S_{ij}\nabla^i\nabla_j\varphi\cdot e^{-\varphi}\,dV+\int_M(-2\nabla_iu\nabla_ju)\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\,dV\\
&\qquad\qquad-\int_M\Delta(e^{-\varphi})\Big(\Delta\varphi+\frac{R}{2}-\tfrac{1}{2}|\nabla\varphi|^2\Big)\,dV\\
&\qquad=\int_M S_{ij}\nabla^i\varphi\nabla^j\varphi\cdot e^{-\varphi}\,dV-\int_M\Delta(e^{-\varphi})(\Delta\varphi+R-\tfrac{1}{2}|\nabla\varphi|^2)\,dV\\
&\qquad\qquad+2\int_M(\nabla_iu\nabla_ju\cdot\nabla^i\nabla^j(e^{-\varphi})-\nabla_iu\nabla_ju\cdot\nabla^i\nabla^j\varphi\cdot e^{-\varphi})\,dV.
\end{aligned}$$

Since $f$ is an eigenfunction of $\lambda$, it induces

$$\lambda=-\frac{\Delta f}{f}+\frac{R}{2}-|\nabla u|^2=\tfrac{1}{2}\Delta\varphi-\tfrac{1}{4}|\nabla\varphi|^2+\frac{R}{2}-|\nabla u|^2,$$

and therefore

$$\begin{aligned}
&\int_M 2S_{ij}\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\,dV+\int_M|\nabla^2\varphi|^2e^{-\varphi}\,dV\\
&\qquad=\int_M S_{ij}\nabla^i\varphi\nabla^j\varphi\cdot e^{-\varphi}\,dV-2\int_M\Delta(|\nabla u|^2)\cdot e^{-\varphi}\,dV\\
&\qquad\qquad+2\int_M\nabla_iu\nabla_j(\nabla^i\nabla^j(e^{-\varphi})-\nabla^i\nabla^j\varphi\cdot e^{-\varphi})\,dV.
\end{aligned}$$

Plugging this into the expression of $\dfrac{d}{dt}\lambda(f(t))$ yields

$$\begin{aligned}
&2\frac{d}{dt}\lambda_{g(t),u(t)}(f(t))\\
&\quad=\int_M S_{ij}\nabla^i\varphi\nabla^j\varphi\cdot e^{-\varphi}\,dV+\int_M|\mathcal{S}|^2e^{-\varphi}\,dV+\int_M|\mathcal{S}|^2e^{-\varphi}\,dV+4\int_M|\Delta u|^2e^{-\varphi}\,dV\\
&\quad=\int_M|\mathcal{S}_{g(t),u(t)}+\nabla^2_{g(t)}\varphi(t)|^2_{g(t)}e^{-\varphi(t)}\,dV_{g(t)}+\int_M|\mathcal{S}_{g(t),u(t)}|^2_{g(t)}e^{-\varphi(t)}\,dV_{g(t)}\\
&\qquad+4\int_M|\Delta_{g(t)}u(t)|^2_{g(t)}e^{-\varphi(t)}\,dV_{g(t)}+2\int_M\Delta_{g(t)}|\nabla_{g(t)}u(t)|^2_{g(t)}e^{-\varphi(t)}\,dV_{g(t)}\\
&\qquad+2\int_M\nabla_iu\nabla_ju\big(-\nabla^i\nabla^j(e^{-\varphi})+\nabla^i\nabla^j\varphi\cdot e^{-\varphi}\big)\,dV
\end{aligned}$$

Now define

$$\begin{aligned}
I&:=\int_M(\nabla_iu\nabla_ju\cdot\nabla^i\nabla^j\varphi)e^{-\varphi}dV=-\int_M\nabla^i(\nabla_iu\nabla_ju\cdot e^{-\varphi})\nabla^j\varphi dV\\
&=-\int_M\nabla^j\varphi(\Delta u\cdot\nabla_ju\cdot e^{-\varphi}+\nabla_iu\nabla^i\nabla_ju\cdot e^{-\varphi}-\nabla_iu\nabla_ju\nabla^i\varphi\cdot e^{-\varphi})dV\\
&=-\int_M\nabla_ju\nabla^j\varphi\Delta u\cdot e^{-\varphi}dV-\int_M\nabla_iu\nabla^j\varphi\nabla^i\nabla_ju\cdot e^{-\varphi}dV+\int_M|\langle du,d\varphi\rangle|^2e^{-\varphi}dV,
\end{aligned}$$

$$\begin{aligned}
II :&= \int_M \nabla_i u \nabla_j u \nabla^i \nabla^j (e^{-\varphi})\, dV = \int_M \nabla^i \nabla^j (\nabla_i u \nabla_j u) e^{-\varphi}\, dV \\
&= \int_M \nabla^i (\nabla^j \nabla_i u \cdot \nabla_j u + \nabla_i u \Delta u) e^{-\varphi}\, dV \\
&= \int_M (\Delta \nabla^i u \cdot \nabla_i u + \nabla^i \Delta u \cdot \nabla_i u + |\nabla^2 u|^2 + |\Delta u|^2) e^{-\varphi}\, dV,
\end{aligned}$$

$$\begin{aligned}
III :&= \int_M \Delta(|\nabla u|^2) e^{-\varphi} dV = 2 \int_M \nabla^i (\nabla_i \nabla_j u \cdot \nabla^j u) e^{-\varphi} dV \\
&= 2 \int_M (\Delta \nabla_j u \cdot \nabla^j u + |\nabla^2 u|^2) e^{-\varphi} dV.
\end{aligned}$$

If we set

$$B := 2(III + I - II),$$

then

$$\begin{aligned}
\frac{B}{2} &= \int_M \big(\Delta \nabla_i u \cdot \nabla^i u - \nabla_i \Delta u \cdot \nabla^i u + |\nabla^2 u|^2 - |\Delta u|^2 + |\langle du, d\varphi \rangle|^2 \\
&\qquad\qquad - \nabla_i u \cdot \nabla^i \varphi \cdot \Delta u - \nabla_i u \cdot \nabla^j \varphi \cdot \nabla^i \nabla_j u \big) e^{-\varphi}\, dV \\
&= \int_M (R_{ij} \nabla^i u \nabla^j u + |\nabla^2 u|^2 - |\Delta u|^2 + |\langle du, d\varphi \rangle|^2 \\
&\qquad\qquad - \nabla_i u \cdot \nabla^i \varphi \cdot \Delta u - \nabla_i u \cdot \nabla^j \varphi \cdot \nabla^i \nabla_j u) e^{-\varphi}\, dV.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
-\int_M \nabla_i u \cdot \nabla^i \varphi \cdot \Delta u \cdot e^{-\varphi}\, dV &= \int_M (\nabla_i u \cdot \Delta u) \nabla^i (e^{-\varphi})\, dV \\
&= -\int_M \nabla^i (\nabla_i u \cdot \Delta u) e^{-\varphi}\, dV \\
&= \int_M (-|\Delta u|^2 - \nabla_i u \cdot \nabla^i \Delta u) e^{-\varphi}\, dV
\end{aligned}$$

and

$$\begin{aligned}
-\int_M \nabla_i u \nabla^j \varphi \nabla^i \nabla_j u \cdot e^{-\varphi}\, dV &= \int_M \nabla_i u \nabla^i \nabla_j u \nabla^j (e^{-\varphi})\, dV \\
&= -\int_M \nabla^j (\nabla_i u \nabla^i \nabla_j u) e^{-\varphi}\, dV \\
&= \int_M (-|\nabla^2 u|^2 - \nabla_i u \Delta \nabla^i u) e^{-\varphi}\, dV.
\end{aligned}$$

Therefore

$$\frac{B}{2} = \int_M \big(-2|\Delta u|^2 + |\langle du, d\varphi \rangle|^2 - 2\langle \nabla u, \nabla \Delta u \rangle \big) e^{-\varphi}\, dV. \tag{9-1}$$

By definition,

$$\Delta(|\nabla u|^2) = \Delta(\nabla^i u \cdot \nabla_i u) = 2\nabla^i u \cdot \Delta\nabla_i u + 2|\nabla^2 u|^2.$$

So

$$\begin{aligned}\Delta|\nabla u|^2 &= 2|\nabla^2 u|^2 + 2(\nabla_i \Delta u + R_{ij}\nabla^j u)\nabla^i u\\ &= 2|\nabla^2 u|^2 + 2R_{ij}\nabla^i u \cdot \nabla^j u + 2\langle \nabla u, \nabla\Delta u\rangle.\end{aligned}$$

Plugging this into (9-1) yields

$$\frac{B}{2} = \int_M \bigl(-2|\Delta u|^2 + |\langle du, d\varphi\rangle|^2 + 2|\nabla^2 u|^2 - \Delta|\nabla u|^2 + 2R_{ij}\nabla^i u\nabla^j u\bigr)e^{-\varphi}\, dV.$$

Since

$$\begin{aligned}2R_{ij}\nabla^i u\nabla^j u &= 2(S_{ij} + 2\nabla_i u\nabla_j u)\nabla^i u\nabla^j u\\ &= 2S_{ij}\nabla^i u\nabla^j u + 4|\nabla u|^4\\ &= \tfrac{1}{4}|\mathcal{S} + 4\, du\otimes du|^2 - \tfrac{1}{4}|\mathcal{S}|^2,\end{aligned}$$

it follows that

$$\begin{aligned}\frac{B}{2} &= III + I - II\\ &= \int_M \bigl(|\langle du, d\varphi\rangle|^2 - 2|\Delta u|^2 - \tfrac{1}{4}|\mathcal{S}|^2 + 2|\nabla^2 u|^2 + \tfrac{1}{4}|\mathcal{S} + 4\, du\otimes du|^2\bigr)e^{-\varphi}\, dV - III.\end{aligned}$$

Hence

$$B = \int_M \bigl(-4|\Delta u|^2 + 2|\langle du, d\varphi\rangle|^2 - \tfrac{1}{2}|\mathcal{S}|^2 + 4|\nabla^2 u|^2 + \tfrac{1}{2}|\mathcal{S} + 4\, du\otimes du|^2\bigr)e^{-\varphi}\, dV - 2III.$$

**Theorem 9.1.** *Suppose that $(g(t), u(t))$ is a solution of the harmonic-Ricci flow on a compact Riemannian manifold $M$ and $f(t)$ is an eigenfunction of $\Delta_{g(t),u(t)}$, that is, $\Delta_{g(t),u(t)} f(t) = \lambda(t) f(t)$ (where $\lambda(t)$ is only a function of time $t$), with the normalized condition $\int_M f(t)^2\, dV_{g(t)} = 1$. Then we have*

$$\begin{aligned}\frac{d}{dt}\lambda(t) &= \frac{d}{dt}\lambda_{g(t),u(t)}(f(t))\\ &= \tfrac{1}{2}\int_M |\mathcal{S}_{g(t),u(t)} + \nabla^2_{g(t)}\varphi(t)|^2_{g(t)} e^{-\varphi(t)} dV_{g(t)} + \frac{1}{4}\int_M |\mathcal{S}_{g(t),u(t)}|^2_{g(t)} e^{-\varphi(t)} dV_{g(t)}\\ &\qquad + \int_M |\langle du(t), d\varphi(t)\rangle_{g(t)}|^2 e^{-\varphi(t)}\, dV_{g(t)} + 2\int_M |\nabla^2_{g(t)} u(t)|^2_{g(t)} e^{-\varphi(t)}\, dV_{g(t)}\\ &\qquad + \frac{1}{4}\int_M |\mathcal{S}_{g(t),u(t)} + 4\, du(t)\otimes du(t)|^2_{g(t)} e^{-\varphi(t)}\, dV_{g(t)}\\ &\qquad - \int_M \Delta_{g(t)}(|\nabla_{g(t)} u(t)|^2_{g(t)}) e^{-\varphi(t)}\, dV_{g(t)}.\end{aligned}$$

**Remark 9.2.** When $u \equiv 0$, this equation reduces to Li's formula [2007].

## 10. The first variation of expander and shrinker entropies

Suppose that $M$ is a closed manifold of dimension $n$. We define

$$\mathcal{W}_\pm : \bigodot^2_+(M) \times C^\infty(M) \times C^\infty(M) \times \mathbb{R}^+ \to \mathbb{R}, \quad (g, u, f, \tau) \mapsto \mathcal{W}_\pm(g, u, f, \tau)$$

where

$$\mathcal{W}_\pm(g, u, f, \tau) := \int_M \big(\tau(S_{g,u} + |\nabla_g f|^2_g) \mp f \pm n\big) \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g. \tag{10-1}$$

Set

$$\mu_\pm(g, u, \tau) := \inf\left\{\mathcal{W}_\pm(g, u, f, \tau) \,\middle|\, f \in C^\infty(M),\ \int_M \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g = 1\right\},$$
$$\nu_\pm(g, u) := \sup\{\mu_\pm(g, u, \tau) \mid \tau > 0\}.$$

**Lemma 10.1.** *Suppose $\nu_\pm(g, u) = \mathcal{W}_\pm(g, u, f_\pm, \tau_\pm)$ for some functions $f_\pm$ and constants $\tau_\pm$ satisfying*

$$\int_M \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g = 1, \quad \tau_\pm > 0.$$

*Then we must have*

$$\tau_\pm(-2\Delta_g f_\pm + |\nabla_g f_\pm|^2_g - S_{g,u}) \pm f_\pm \mp n + \nu_\pm(g, u) = 0,$$
$$\int_M \frac{f_\pm e^{-f_\pm}}{(4\pi\tau)^{n/2}} dV_g = \frac{n}{2} \mp \nu_\pm(g, u).$$

*Proof.* Since $g$ and $u$ are fixed, we consider the corresponding Lagrangian multiplier function

$$\mathfrak{L}_\pm(f, \tau; \lambda) := \mathcal{W}_\pm(g, u, f, \tau) - \lambda\left(\int_M \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g - 1\right).$$

Then the variation of $\mathfrak{L}_\pm$ in $f$ direction is

$$\delta_f \mathfrak{L}_\pm(f, \tau; \lambda) = \int_M \big(2\tau\nabla^i f \nabla_i(\delta f) \mp \delta f + \lambda\delta f\big) \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g$$
$$- \int_M \big(\tau(S_{g,u} + |\nabla_g f|^2_g) \mp f \pm n\big)\delta f \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g.$$

By the divergence theorem, we calculate

$$\int_M \nabla^i f \cdot \nabla_i(\delta f) \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g = -\int_M \nabla_i(\nabla^i f \frac{e^{-f}}{(4\pi\tau)^{n/2}})\delta f\, dV_g$$
$$= -\int_M (\Delta_g f - |\nabla_g f|^2_g)\delta f \frac{e^{-f}}{(4\pi\tau)^{n/2}}\, dV_g.$$

Hence

$$\delta_f \mathfrak{L}_\pm(f,\tau;\lambda)=\int_M \big(\tau(-2\Delta_g f+|\nabla_g f|_g^2-S_{g,u})\pm f\mp n\mp 1+\lambda\big)\delta f\,\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV.$$

This implies that

$$\tau_\pm(-2\Delta_g f_\pm+|\nabla_g f_\pm|_g^2-S_{g,u})\pm f_\pm\mp n\mp 1+\lambda_\pm=0.$$

Since $f_\pm$ satisfies the normalized condition, it follows that

$$0=\lambda_\pm\mp 1+\int_M \big(\tau_\pm(-2\Delta_g f_\pm+|\nabla_g f_\pm|_g^2-S_{g,u})\pm f_\pm\mp n\big)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g.$$

From the identity

$$\int_M \Delta_g f\,\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV_g=\int_M |\nabla_g f|_g^2\,\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV_g$$

and the definition (10-1), we obtain

$$\nu_\pm(g,u)=\mathscr{W}_\pm(g,u,f_\pm,\tau_\pm)=\lambda_\pm\mp 1,$$

and, consequently,

$$\tau_\pm(-2\Delta_g f_\pm+|\nabla_g f_\pm|_g^2-S_{g,u})\pm f_\pm\mp n+\nu_\pm(g,u)=0.$$

The variation of $\mathfrak{L}_\pm$ with respect to $\tau$ indicates that

$$\begin{aligned}\delta_\tau \mathfrak{L}_\pm(f,\tau;\lambda)&=\int_M \delta\tau(S_{g,u}+|\nabla_g f|_g^2)\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV_g-\lambda\int_M\Big(-\frac{n}{2}\frac{\delta\tau}{\tau}\Big)\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV_g\\&\quad+\int_M\Big(-\frac{n}{2}\frac{\delta\tau}{\tau}\Big)\big(\tau(S_{g,u}+|\nabla_g f|_g^2)\mp f\pm n\big)\frac{e^{-f}}{(4\pi\tau)^{n/2}}\,dV_g\\&=\int_M \delta\tau\bigg(\Big(1-\frac{n}{2}\Big)(S_{g,u}+|\nabla_g f|_g^2)+\frac{n}{2\tau}(\lambda\pm f\mp n)\bigg)\frac{e^{-f}\,dV_g}{(4\pi\tau)^{n/2}}.\end{aligned}$$

Using the first proved equation, we have

$$\begin{aligned}0&=\int_M\Big((\nu_\pm(g,u)\pm f_\pm\mp n)\Big(1-\frac{n}{2}\Big)+\frac{n}{2}(\nu_\pm(g,u)\pm f_\pm\mp n\pm 1)\Big)\frac{e^{-f_\pm}\,dV_g}{(4\pi\tau_\pm)^{n/2}}\\&=\int_M\Big(\nu_\pm\pm f_\pm\mp\frac{n}{2}\Big)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g\end{aligned}$$

and therefore we obtain the second one. □

For a symmetric 2-tensor $h=(h_{ij})\in\bigodot^2(M)$, we set

$$g(s):=g+sh$$

Then the variation of $g(s)$ is

$$\frac{\partial}{\partial s}\Big|_{s=0} R_{g(s)} = -h^{ij}R_{ij} + \nabla^i\nabla^j h_{ij} - \Delta_g(\mathrm{tr}_g\, h). \tag{10-2}$$

**Theorem 10.2.** *Suppose that $(M, g)$ is a compact Riemannian manifold and $u$ a smooth function on $M$. Let $h$ be any symmetric covariant $2$-tensor on $M$ and set $g(s) := g + sh$. Let $v$ be any smooth function on $M$ and $u(s) := u + sv$. If*

$$\nu_\pm(g(s), u(s)) = \mathcal{W}_\pm(g(s), u(s),\, f_\pm(s),\, \tau_\pm(s))$$

*for some smooth functions $f_\pm(s)$ with*

$$\int_M e^{-f_\pm(s)}\, dV/(4\pi\tau_\pm(s))^{n/2} = 1$$

*and constants $\tau_\pm(s) > 0$,*

$$\frac{d}{ds}\Big|_{s=0}\nu_\pm(g(s), u(s)) = -\tau_\pm\int_M\left(\langle h, \mathcal{S}_{g,u}\rangle_g + \langle h, \nabla_g^2 f\rangle_g \pm \frac{1}{2\tau_\pm}\mathrm{tr}_g\, h\right)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g$$
$$+4\tau_\pm\int_M v(\Delta_g u - \langle du, df_\pm\rangle_g)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g,$$

*where $f_\pm := f_\pm(0)$ and $\tau_\pm := \tau_\pm(0)$. In particular, the critical points of $\nu_\pm(\,\cdot\,,\cdot\,)$ satisfy*

$$\mathcal{S}_{g,u} + \nabla_g^2 f \pm \frac{1}{2\tau_\pm} g = 0, \qquad \Delta_g u = \langle du, df_\pm\rangle_g.$$

*Consequently, if $\mathcal{W}_\pm(g, u, f, \tau)$ and $\nu_\pm(g, u)$ achieve their minimums, $(M, g)$ is a gradient expanding and shrinker harmonic-Ricci soliton according to the sign.*

*Proof.* By definition, one has

$$\begin{aligned}
&\frac{d}{ds}\nu_\pm(g(s), u(s)) \\
&\quad= \frac{d}{ds}\mathcal{W}_\pm(g(s), u(s),\, f_\pm(s),\, \tau_\pm(s)) \\
&\quad= \int_M\left(\frac{\partial}{\partial s}\tau_\pm(s)(S_{g(s),u(s)} + |\nabla_{g(s)} f_\pm(s)|^2_{g(s)})\right)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\, dV_{g(s)} \\
&\qquad+ \int_M\left(\tau_\pm(s)\frac{\partial}{\partial s}(S_{g(s),u(s)} + |\nabla_{g(s)} f_\pm(s)|^2_{g(s)}) \mp \frac{\partial}{\partial s} f_\pm(s)\right)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\, dV_{g(s)} \\
&\qquad\qquad+ \int_M\big(\tau_\pm(s)(S_{g(s),u(s)} + |\nabla_{g(s)} f_\pm(s)|^2_{g(s)}) \mp f_\pm(s) \pm n\big) \\
&\qquad\qquad\qquad\cdot\frac{\partial}{\partial s}\left(\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\, dV_{g(s)}\right).
\end{aligned}$$

Since

$$\begin{aligned}\frac{\partial}{\partial s}S_{g(s),u(s)} &= \frac{\partial}{\partial s}R_{g(s)} - 2\frac{\partial}{\partial s}|\nabla_{g(s)}u(s)|^2_{g(s)} \\ &= \frac{\partial}{\partial s}R_{g(s)} - 2\left(\frac{\partial}{\partial s}g^{ij}\right)\nabla_i u\nabla_j u - 4g^{ij}\frac{\partial}{\partial s}\nabla_i u\cdot\nabla_j u \\ &= \frac{\partial}{\partial s}R_{g(s)} - 2(-g^{ip}g^{jq}h_{pq})\nabla_i u\nabla_j u - 4g^{ij}\nabla_i\left(\frac{\partial}{\partial s}u\right)\nabla_j u \\ &= \frac{\partial}{\partial s}R_{g(s)} + 2h_{pq}\nabla^p u\nabla^q u - 4\nabla_i\left(\frac{\partial}{\partial t}u\right)\nabla^i u\end{aligned}$$

and

$$\begin{aligned}&\frac{\partial}{\partial s}\left(\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)}\right) \\ &= \left(-\frac{\partial}{\partial s}f_\pm(s) - \frac{n}{2\tau_\pm(s)}\frac{\partial}{\partial s}\tau_\pm(s)\right)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)} + \frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\frac{\partial}{\partial s}\,dV_{g(s)} \\ &= \left(-\frac{\partial}{\partial s}f_\pm(s) - \frac{n}{2\tau_\pm(s)}\frac{\partial}{\partial s}\tau_\pm(s) + \tfrac{1}{2}\operatorname{tr}_g h\right)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)},\end{aligned}$$

it follows that

$$\begin{aligned}&\frac{d}{ds}\nu_\pm(g(s),u(s)) \\ &\quad= \int_M \frac{\partial}{\partial s}\tau_\pm(s)(S_{g(s),u(s)} + |\nabla_{g(s)}f_\pm(s)|^2_{g(s)})\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)} \\ &\qquad+ \int_M\left(\tau_\pm(s)\left(\frac{\partial}{\partial s}R_{g(s)} + 2h_{pq}\nabla^p u\nabla^q u - 4\nabla_i\left(\frac{\partial}{\partial s}u\right)\nabla^i u\right.\right. \\ &\qquad\left.\left.- h_{pq}\nabla^p f\nabla^q f + 2\nabla_i\left(\frac{\partial}{\partial s}f\right)\nabla^i f\right) \mp \frac{\partial}{\partial s}f_\pm(s)\right)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)} \\ &\qquad+ \int_M\left(-\frac{\partial}{\partial s}f_\pm(s) - \frac{n}{2\tau_\pm(s)}\frac{\partial}{\partial s}\tau_\pm(s) + \tfrac{1}{2}\operatorname{tr}_g h\right) \\ &\qquad\qquad\cdot\big(\tau_\pm(s)(S_{g(s),u(s)} + |\nabla_{g(s)}f_\pm(s)|^2_{g(s)}) \mp f_\pm(s) \pm n\big)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_{g(s)}.\end{aligned}$$

From the equalities

$$\begin{aligned}\int_M \Delta_g\operatorname{tr}_g h\cdot e^{-f}\,dV_g &= \int_M \operatorname{tr}_g h\cdot\Delta_g(e^{-f})\,dV_g \\ &= \int_M \operatorname{tr}_g h(-\Delta_g f + |\nabla_g, f|^2_g)e^{-f}\,dV_g,\end{aligned}$$

$$\int_M \nabla^i\nabla^j h_{ij}\cdot e^{-f}\,dV_g = \int_M h_{ij}\nabla^i\nabla^j(e^{-f})\,dV$$
$$= \int_M h_{ij}(-\nabla^i\nabla^j f + \nabla^i f\nabla^j f)e^{-f}\,dV_g,$$
$$\int_M \nabla_i\left(\frac{\partial}{\partial s}f\right)\nabla^i f\, e^{-f}\,dV_g = \int_M -\frac{\partial}{\partial s}f(\Delta_g f - |\nabla_g f|_g^2)e^{-f}\,dV_g,$$
$$\int_M \Delta_g(e^{-f})\,dV_g = \int_M(-\Delta_g f + |\nabla_g f|_g^2)e^{-f}\,dV_g,$$

and Lemma 10.1, we obtain

$$\begin{aligned}
&\frac{d}{ds}\bigg|_{s=0}\nu_\pm(g(s),u(s))\\
&\quad= \int_M \frac{\partial}{\partial s}\bigg|_{s=0}\tau_\pm(s)(S_{g,u}+|\nabla_g f|_g^2)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g\\
&\qquad+\int_M\bigg(\tau_\pm\bigg(-h^{ij}R_{ij}+\nabla^i\nabla_j h_{ij}-\Delta_g(\operatorname{tr}_g h)+2h_{pq}\nabla^p u\nabla^q u\\
&\qquad\qquad-4\nabla_i v\nabla^i u - h_{pq}\nabla^p f\nabla^q f + 2\nabla_i\bigg(\frac{\partial}{\partial s}\bigg|_{s=0}f(s)\bigg)\nabla^i f\bigg)\mp\frac{\partial}{\partial s}\bigg|_{s=0}f(s)\bigg)\\
&\qquad\cdot\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g+\int_M\bigg(-\frac{\partial}{\partial s}\bigg|_{s=0}f_\pm(s)-\frac{n}{2\tau_\pm}(s)\frac{\partial}{\partial s}\bigg|_{s=0}\tau_\pm(s)+\tfrac{1}{2}\operatorname{tr}_g h\bigg)\\
&\qquad\qquad\cdot\big(\tau_\pm(S_{g,u}+|\nabla_g f_\pm|_g^2)\mp f_\pm\pm n\big)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g.
\end{aligned}$$

If we denote by $B$ the last term and by $A$ the remaining terms,

$$\begin{aligned}
A = \int_M\bigg(&\frac{\partial}{\partial s}\bigg|_{s=0}\tau_\pm(s)(|\nabla_g f_\pm|_g^2+S_{g,u})\\
&-\tau_\pm(h^{ij}\nabla_i\nabla_j f_\pm + h^{ij}S_{ij}+4\nabla_i v\cdot\nabla^i u)\mp\frac{\partial}{\partial s}f_\pm\bigg)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g\\
&+\int_M \tau_\pm(\Delta_g f_\pm - |\nabla_g f_\pm|_g^2)\bigg(\operatorname{tr}_g h - 2\frac{\partial}{\partial s}\bigg|_{s=0}f(s)\bigg)\frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\,dV_g.
\end{aligned}$$

The normalized condition

$$1 = \int_M \frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_g$$

implies

$$0 = \int_M\bigg(-\frac{\partial}{\partial s}\bigg|_{s=0}f_\pm(s)-\frac{n}{2\tau_\pm}\frac{\partial}{\partial s}\bigg|_{s=0}\tau_\pm(s)+\tfrac{1}{2}\operatorname{tr}_g h\bigg)\frac{e^{-f_\pm(s)}}{(4\pi\tau_\pm(s))^{n/2}}\,dV_g.$$

From Lemma 10.1, we conclude that

$$\tau_\pm S_{g,u} - \tau_\pm(|\nabla_g f_\pm|_g^2 - 2\Delta_g f_\pm) = \pm f_\pm \mp n + \nu_\pm(g,u).$$

Therefore

$$\tau_\pm(S_{g,u} + |\nabla_g f_\pm|_g^2) \mp f_\pm \pm n = 2\tau_\pm(|\nabla_g f_\pm|_g^2 - \Delta_g f_\pm) + \nu_\pm(g,u).$$

Plugging this into the definition of $B$ yields

$$\begin{aligned} B &= \int_M \left( -\frac{\partial}{\partial s}\bigg|_{s=0} f_\pm(s) - \frac{n}{2\tau_\pm}\frac{\partial}{\partial s}\bigg|_{s=0} \tau_\pm(s) + \tfrac{1}{2}\operatorname{tr}_g h \right) \\ &\qquad\qquad \cdot \big(2\tau_\pm(|\nabla_g f_\pm|_g^2 - \Delta_g f_\pm) + \nu_\pm(g,u)\big) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g \\ &= \int_M \left( -\frac{\partial}{\partial s}\bigg|_{s=0} f_\pm(s) - \frac{n}{2\tau_\pm}\frac{\partial}{\partial s}\bigg|_{s=0} \tau_\pm(s) + \tfrac{1}{2}\operatorname{tr}_g h \right) \\ &\qquad\qquad \cdot \big(2\tau_\pm(|\nabla_g f_\pm|_g^2 - \Delta_g f_\pm)\big) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g \\ &= \int_M \left( -\frac{\partial}{\partial s}\bigg|_{s=0} f_\pm(s) + \tfrac{1}{2}\operatorname{tr}_g h \right) 2\tau_\pm(|\nabla_g f_\pm|_g^2 - \Delta_g f_\pm) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g, \end{aligned}$$

where we use the fact that

$$\int_M \Delta_g(e^{-f})dV_g = 0.$$

Hence $B$ cancels with the last term in $A$. Therefore the above variation equals

$$\begin{aligned} &\frac{d}{ds}\bigg|_{s=0} \nu_\pm(g(s),u(s)) \\ &\quad = \int_M \left( \frac{\partial}{\partial s}\bigg|_{s=0} \tau_\pm(s) \left( |\nabla_g f_\pm|_g^2 + S_{g,u} \pm \frac{n}{2\tau_\pm} \right) - \tau_\pm \left( h^{ij}\nabla_i\nabla_j f + h^{ij} S_{ij} \right.\right. \\ &\qquad\qquad \left.\left. \pm \frac{1}{2\tau_\pm}\operatorname{tr}_g h + 4v(\langle du, df\rangle - \Delta_g u) \right)\right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g. \end{aligned}$$

To prove the theorem, it is sufficient to show that

$$\int_M \left( |\nabla_g f_\pm|_g^2 + S_{g,u} \pm \frac{n}{2\tau_\pm} \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV = 0.$$

Since $M$ is compact, we have

$$0 = \int_M \Delta_g(e^{-f_\pm}) = \int_M (-\Delta_g f_\pm + |\nabla_g f_\pm|_g^2) e^{-f_\pm}\, dV.$$

Hence

$$\int_M \left( |\nabla_g f_\pm|^2 + S_{g,u} \pm \frac{n}{2\tau_\pm} \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV$$
$$= \int_M \left( 2\Delta_g f_\pm - |\nabla_g f|_g^2 + S_{g,u} \pm \frac{n}{2\sigma_\pm} \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV.$$

Lemma 10.1 now indicates

$$\begin{aligned}\int_M \left( |\nabla_g f_\pm|^2 + S_{g,u} \pm \frac{n}{2\tau_\pm} \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV \\ &= \int_M \left( \frac{\pm f_\pm \mp n + \nu_\pm(g,u)}{\tau_\pm} \pm \frac{n}{2} \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV \\ &= \int_M \frac{1}{\tau_\pm} \left( \pm f_\pm \mp \frac{n}{2} + \nu_\pm(g,u) \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV \\ &= \frac{1}{\tau_\pm} \left( \pm\frac{n}{2} - \nu_\pm(g,u) \mp \frac{n}{2} + \nu_\pm(g,u) \right) = 0.\end{aligned}$$

The sign $+$ corresponds to the gradient expanding soliton and the sign $-$ to the gradient shrinker soliton. $\square$

**Corollary 10.3.** *Suppose that $(M, g)$ is a compact Riemannian manifold and $u$ is a smooth function on $M$. Let $h$ be any symmetric covariant 2-tensor on $M$ and set $g(s) := g + sh$. Let $v$ be any smooth function on $M$ and $u(s) := u + sv$. If $\nu_\pm(g(s), u(s)) = \mathcal{W}_\pm(g(s), u(s), f_\pm(s), \tau_\pm(s))$ for some smooth function $f_\pm(s)$ with $\int_M e^{-f_\pm(s)}\, dV/(4\pi\tau_\pm(s))^{n/2} = 1$ and a constant $\tau_\pm(s) > 0$, and $(g, u)$ is a critical point of $\nu_\pm(\cdot, \cdot)$, then*

$$\mathcal{S}_{g,u} = \mp\frac{1}{2\tau_\pm} g, \qquad f_\pm \equiv \text{constant}.$$

*Thus, if $\mathcal{W}_\pm(g, u, \cdot, \cdot)$ achieve their minimum and $(g, u)$ is a critical point of $\nu_\pm(\cdot, \cdot)$, $(M, g, u)$ satisfies the static Einstein vacuum equation.*

*Proof.* According to Lemma 10.1 and Theorem 10.2, we have

$$\tau_\pm(-2\Delta_g f_\pm + |\nabla_g f_\pm|_g^2 - S_{g,u}) \pm f_\pm \mp n$$
$$= -\nu_\pm = -\int_M \left( \tau_\pm(S_{g,u} + |\nabla_g f|_g^2) \mp f_\pm \pm n \right) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g,$$

and hence

$$\begin{aligned}2\Delta_g f_\pm - |\nabla_g f_\pm|_g^2 + S_{g,u} &= \int_M (S_{g,u} + |\nabla_g f_\pm|_g^2) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g \\ &= \int_M (S_{g,u} + \Delta_g f_\pm) \frac{e^{-f_\pm}}{(4\pi\tau_\pm)^{n/2}}\, dV_g \\ &= \mp \frac{n}{2\tau_\pm} = S_{g,u} + \Delta_g f_\pm.\end{aligned}$$

From this we get $\Delta_g f_\pm = |\nabla_g f_\pm|_g^2$. After integrating on both sides, the functions $f_\pm$ must be constant, which implies $\mathcal{S}_g \pm (1/(2\tau_\pm))g = 0$. □

**Remark 10.4.** In the situation of Corollary 10.3, by normalization, we my choose $f_\pm = n/2$.

## Acknowledgements

The author thanks Professor Kefeng Liu, who teaches the author mathematics, Professor Hongwei Xu, and other staff members at the Center of Mathematical Science, where part of this work was done.

## References

[Cao 2007] X. Cao, "Eigenvalues of $(-\Delta + R/2)$ on manifolds with nonnegative curvature operator", *Math. Ann.* **337**:2 (2007), 435–441. MR 2007g:53071 Zbl 1105.53051

[Cao 2008] X. Cao, "First eigenvalues of geometric operators under the Ricci flow", *Proc. Amer. Math. Soc.* **136**:11 (2008), 4075–4078. MR 2009f:53098 Zbl 1166.58007

[He et al. 2008] C.-L. He, S. Hu, D.-X. Kong, and K. Liu, "Generalized Ricci flow. I. Local existence and uniqueness", pp. 151–171 in *Topology and physics*, Nankai Tracts Math. **12**, World Sci. Publ., Hackensack, NJ, 2008. MR 2010k:53098

[Li 2007] J.-F. Li, "Eigenvalues and energy functionals with monotonicity formulae under Ricci flow", *Math. Ann.* **338**:4 (2007), 927–946. MR 2008c:53068 Zbl 1127.53059

[Li 2012] Y. Li, "Generatlized Ricci flow I: higher dericative estimates for compact manifolds", *Analysis & PDE* **5**:4 (2012), 747–775.

[List 2006] B. List, *Evolution of an extended Ricci flow system*, Ph.D. thesis, Fachbereich Mathematik und Informatik der Freie Universität Berlin, 2006, Available at http://www.diss.fu-berlin.de/2006/180/index.html.

[Müller 2012] R. Müller, "Ricci flow coupled with harmonic map flow", *Ann. Sci. Éc. Norm. Supér.* (4) **45**:1 (2012), 101–142. MR 2961788

[Oliynyk et al. 2006] T. Oliynyk, V. Suneeta, and E. Woolgar, "A gradient flow for worldsheet nonlinear sigma models", *Nuclear Phys. B* **739**:3 (2006), 441–458. MR 2006m:81185 Zbl 1109.81058

[Streets 2007] J. D. Streets, *Ricci Yang–Mills flow*, Ph.D. thesis, Duke University, 2007, Available at http://www.math.uci.edu/~jstreets/papers/StreetsThesis.pdf.

[Streets 2008] J. Streets, "Regularity and expanding entropy for connection Ricci flow", *J. Geom. Phys.* **58**:7 (2008), 900–912. MR 2009f:53105 Zbl 1144.53326

[Streets 2009] J. Streets, "Singularities of renormalization group flows", *J. Geom. Phys.* **59**:1 (2009), 8–16. MR 2010a:53143 Zbl 1153.53329

[Streets 2010] J. Streets, "Ricci Yang–Mills flow on surfaces", *Adv. Math.* **223**:2 (2010), 454–475. MR 2011c:53164 Zbl 1190.53069

[Young 2008] A. N. Young, *Modified Ricci flow on a principal bundle*, Ph.D. thesis, The University of Texas at Austin, 2008, Available at http://search.proquest.com/docview/193674070.

Received August 27, 2012. Revised January 6, 2013.

YI LI
DEPARTMENT OF MATHEMATICS
JOHNS HOPKINS UNIVERSITY
3400 N. CHARLES STREET
BALTIMORE, MD 21218
UNITED STATES

yli@math.jhu.edu
*Current address*:
DEPARTMENT OF MATHEMATICS
SHANGHAI JIAO TONG UNIVERSITY
800 DONG CHUAN ROAD, MIN HANG DISTRICT
SHANGHAI, 200240
CHINA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.185

# QUANTUM EXTREMAL LOOP WEIGHT MODULES AND MONOMIAL CRYSTALS

MATHIEU MANSUY

**In this paper we construct a new family of representations for the quantum toroidal algebra $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$, which are $\ell$-extremal in the sense of Hernandez. We construct extremal loop weight modules associated to level 0 fundamental weights $\varpi_\ell$ when $n = 2r+1$ is odd and $\ell = 1$, $r+1$ or $n$. To do this, we relate monomial realizations of level 0 extremal fundamental weight crystals to integrable representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$, and we introduce promotion operators for the level 0 extremal fundamental weight crystals. By specializing the quantum parameter, we get finite-dimensional modules of quantum toroidal algebras at roots of unity. In general, we give a conjectural process to construct extremal loop weight modules from monomial realizations of crystals.**



## 1. Introduction

Let us consider a finite-dimensional simple Lie algebra $\mathfrak{g}$ and its associated quantum affine algebra $\mathcal{U}_q(\hat{\mathfrak{g}})$. Beck [1994] and Drinfeld [1987] proved that $\mathcal{U}_q(\hat{\mathfrak{g}})$ has two realizations: first as the quantized enveloping algebra of the affine Lie algebra $\hat{\mathfrak{g}}$ and second as the Drinfeld quantum affinization of the quantum group $\mathcal{U}_q(\mathfrak{g})$.

*MSC2010:* 20G42, 81R10.
*Keywords:* quantum toroidal algebras, extremal loop weight modules, monomial crystals, promotion operators.

The representation theory of the quantum affine algebras has been intensively studied (see, among others, [Akasaka and Kashiwara 1997; Beck and Nakajima 2004; Chari and Pressley 1991; 1995; Frenkel and Mukhin 2001; Frenkel and Reshetikhin 1999; Lusztig 1993; Nakajima 2001]). Kashiwara [1994] has defined a class of integrable representations $V(\lambda)$ of these algebras, called extremal weight modules, parametrized by an integrable weight $\lambda$ and with crystal basis $\mathcal{B}(\lambda)$. When $\lambda$ is dominant, $V(\lambda)$ is the simple integrable module of highest weight $\lambda$. But in general $V(\lambda)$ is not simple and it is neither of highest weight nor of lowest weight. These representations were the subject of numerous papers (see [Beck 2002; Beck and Nakajima 2004; Hernandez and Nakajima 2006; Kashiwara 1994; 2002b; Naito and Sagaki 2003; 2006; Nakajima 2004]) and are particularly important because they have finite-dimensional quotients for some special weight $\lambda$. Kashiwara has proved in this way the existence of crystal bases for the finite-dimensional fundamental representations of $\mathcal{U}_q(\hat{\mathfrak{g}})$ (for a special choice of the spectral parameter).

The quantum affine algebra $\mathcal{U}_q(\hat{\mathfrak{g}})$ is also a quantum Kac–Moody algebra and thus can be affinized again by the Drinfeld quantum affinization process. One gets a toroidal (or double affine) quantum algebra $\mathcal{U}_q(\mathfrak{g}^{\mathrm{tor}})$ which is not a quantum Kac–Moody algebra anymore and can not be affinized again by this process (it can be viewed as "the terminal object" in this construction). These algebras were first introduced by Ginzburg, Kapranov and Vasserot [Ginzburg et al. 1995] in type A and then in the general context [Jing 1998; Nakajima 2001]. In type A, they are in Schur–Weyl duality with elliptic Cherednik algebras [Varagnolo and Vasserot 1996].

The representation theory of these algebras has been intensively studied (see for example [Feigin et al. 2011a; 2011b; 2012; 2013; Hernandez 2005; 2009; 2011; Miki 2000; Varagnolo and Vasserot 1998] and references therein). In the spirit of works of Kashiwara, Hernandez [2009] proposed the definition of extremal loop weight modules for $\mathcal{U}_q(\mathfrak{g}^{\mathrm{tor}})$. The main motivation is to construct finite-dimensional representations of the quantum toroidal algebra at roots of unity. He constructs the first example of such a module for $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ which is neither of $\ell$-highest weight nor of $\ell$-lowest weight. This module is generated by an $\ell$-weight vector of $\ell$-weight an analogue of the level 0 fundamental weight $\varpi_1 = \Lambda_1 - \Lambda_0$. By specializing the quantum parameter $q$ at roots of unity, he obtains finite-dimensional representations of the quantum toroidal algebra at roots of unity.

In the present paper, we construct a new family of extremal loop weight modules for the quantum toroidal algebra $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$[1]: we define extremal loop weight

[1] After this paper appeared on the arXiv, the constructions in [Feigin et al. 2013] were brought to our attention by H. Nakajima. Some of the representations constructed in this paper (the $V(Y_{1,0}Y_{0,1}^{-1})$) are also defined in [Feigin et al. 2013] from another point of view and are called *vector representations* there.

modules associated to the level 0 fundamental weight $\varpi_\ell = \Lambda_\ell - \Lambda_0$ when $n = 2r+1$ is odd and $\ell = 1, r+1$ or $n$ (Theorem 4.1). We call them the extremal fundamental loop weight modules. This construction is based on the monomial realizations of level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$. We relate these monomial crystals with integrable representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ by studying their combinatorics: we introduce promotion operators for $\mathcal{B}(\varpi_\ell)$ $(1 \le \ell \le n)$. We describe them in terms of monomials. These operators play an important role in our work: on the one hand, at the level of crystals, they are used to check that these monomial crystals are closed when $\ell = 1, r+1$ or $n$ (see Definition 3.6 for this notion, related to the theory of $q$-characters). On the other hand, at the level of representations, they enable us to define the action of the quantum toroidal algebra. We show that the representations we constructed are irreducible and, as modules over the horizontal quantum affine subalgebra, they are isomorphic to the fundamental extremal weight modules $V(\varpi_\ell)$. We give explicit formulas for the action from the associated monomial crystal. By specializing the quantum parameter $q$ at roots of unity, we get new irreducible finite-dimensional representations of the quantum toroidal algebra at roots of unity. When $\ell$ is not equal to $1, r+1$ or $n$, the corresponding monomial crystals are not closed and it is not possible to make the same construction. We give a conjectural process to define other extremal loop weight modules in this situation: as an example, we construct an extremal loop weight module of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ associated to the weight $2\varpi_1$.

Let us describe the methods used in this paper in more detail. Kashiwara [2003] and Nakajima [2003] have defined a crystal $\mathcal{M}$, called the monomial crystal, whose vertices are Laurent monomials. They determined monomial realizations of crystals of finite type. These results have been extended in [Hernandez and Nakajima 2006] to the level 0 extremal weight $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystals $\mathcal{B}(\varpi_\ell)$ $(1 \le \ell \le n)$: if $n = 2r+1$ is odd, it is isomorphic to a sub-$\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystal $\mathcal{M}_\ell$ of $\mathcal{M}$.

The monomials occurring in these realizations of crystals can be interpreted at the level of representation theory. In fact Frenkel and Reshetikhin [1999] defined a correspondence between $\ell$-weights (eigenvalues of the Cartan subalgebra for the Drinfeld realization) and these monomials. Motivated by these facts, Hernandez [2009] used the monomial $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-crystal $\mathcal{M}_1$ to construct an integrable representation of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ whose $\ell$-weights are the monomials occurring in this crystal. He defined in this way the first example of extremal loop weight modules for $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$. We use the same technical feature in this paper. We propose to relate the monomial $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystals $\mathcal{M}_\ell$ (where $n = 2r+1$ is supposed to be odd) with integrable representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$.

Let us outline the main steps of the construction of extremal fundamental loop weight modules associated to $\mathcal{M}_\ell$. It is based on the combinatorial study of these crystals. The cyclic symmetry of the Dynkin diagram of type $A_n^{(1)}$ has a counterpart at the level of crystals. Actually, these symmetry properties are already known for the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystals of finite type, and translated into the existence of promotion operators (see [Bandlow et al. 2010; Fourier et al. 2009; Okado and Schilling 2008; Schilling 2008; Shimozono 2002] and references therein). Here we introduce promotion operators for the level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$ $(1 \leq \ell \leq n)$. We improve these operators in the monomial realizations $\mathcal{M}_\ell$ of [Hernandez and Nakajima 2006]. In particular, we get a new description of these monomial crystals.

A monomial set is not in general the set of $\ell$-weights of an integrable representation. In fact, it must satisfy combinatorial properties related to the theory of $q$-characters (see [Frenkel and Mukhin 2001; Frenkel and Reshetikhin 1999]). This leads us to introduce the notion of closed monomial set (Definition 3.6). It gives a necessary condition for a set to be the set of $\ell$-weights of an integrable representation. Finally, we determine when the monomial crystal $\mathcal{M}_\ell$ is closed, using promotion operators: this is the case if and only if $\ell = 1, r+1$ or $n$ (Theorem 3.22).

When $\mathcal{M}_\ell$ is closed, we construct an associated integrable $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module whose set of $\ell$-weights consists of monomials occurring in $\mathcal{M}_\ell$. For that, we paste together some finite-dimensional representations of the various vertical quantum affine subalgebras of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$. The existence of promotion operators for $\mathcal{M}_\ell$ involves that it defines a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module structure. Furthermore we check that the representations obtained in this way do satisfy the definition of extremal loop weight modules. They are irreducible, isomorphic to the level 0 fundamental extremal representations $V(\varpi_\ell)$ as modules over the horizontal quantum affine subalgebra. Moreover the action of the quantum toroidal algebra on them is explicitly known, given from the associated crystal. Finally by specializing the quantum parameter $q$ at roots of unity, we get finite-dimensional representations of the quantum toroidal algebra at roots of unity.

When the monomial crystal $\mathcal{M}_\ell$ is not closed, there is no integrable representation of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ whose set of $\ell$-weights consists of monomials occurring in it. The idea is to consider instead of $\mathcal{M}_\ell$ a closed crystal containing it and to apply the preceding methods to this crystal. We treat an example of such a construction: we define a representation of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ which satisfies the definition of extremal loop weight modules.

Let us now describe briefly the organization of this paper.

In Section 2 we recall the definitions of quantum affine algebras $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ and quantum toroidal algebras $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ and we briefly review their representation theory. In particular one defines the extremal weight modules and the extremal loop weight modules. Section 3 is devoted to the study of monomial crystals. We recall its definition and we introduce the notion of closed monomial set (Definition 3.6). We introduce promotion operators for the level 0 fundamental extremal weight crystals. As a consequence, we determine when $\mathcal{M}_\ell$ is closed (Theorem 3.22). In Section 4 we construct a new family of representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ (the extremal fundamental loop weight modules) when $n$ is odd and $\mathcal{M}_\ell$ is closed (Theorem 4.1). We check that these representations satisfy the definition of extremal loop weight modules (Theorem 4.7) and we give formulas for the action (Theorem 4.12). We get finite-dimensional representations of the quantum toroidal algebra at roots of unity by specializing the quantum parameter $q$ at roots of unity (Theorem 4.18). In Section 5 we treat an example where the considered monomial crystal is not closed. We construct a representation of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ associated to the level 0 weight $2\varpi_1$. In Section 6 other possible developments and applications of these results are discussed.

## 2. Background

We recall the main definitions and general properties about the representation theory of quantum affine algebras and quantum toroidal algebras of type $A$.

**2A.** ***Cartan matrix.*** Let $C = (C_{i,j})_{0\le i,j\le n}$ be a Cartan matrix of type $A_n^{(1)}$ ($n \ge 2$),

$$C = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2 & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 2 & -1 \\ -1 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

**Remark 2.1.** The case $n = 1$ is not studied in the article and is particular. In this case, the Cartan matrix is

$$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

and involves $-2$. Furthermore, the quantum toroidal algebra $\mathcal{U}_q(\mathrm{sl}_2^{\mathrm{tor}})$ requires a special definition with different possible choices of the quantized Cartan matrix (see [Hernandez 2011, Remark 4.1]).

Set $I = \{0, \dots, n\}$ and $I_0 = \{1, \dots, n\}$. In particular, $(C_{i,j})_{i,j\in I_0}$ is the Cartan matrix of finite type $A_n$. In the following, $I$ will be often identified with the set $\mathbb{Z}/(n+1)\mathbb{Z}$. Consider the $(n+2)$-dimensional vector space

$$\mathfrak{h} = \mathbb{Q}h_0 \oplus \mathbb{Q}h_1 \oplus \cdots \oplus \mathbb{Q}h_n \oplus \mathbb{Q}d$$

and the linear functions $\alpha_i$ (the simple roots) and $\Lambda_i$ (the fundamental weights) on

$\mathfrak{h}$ given by $(i, j \in I)$,

$$\alpha_i(h_j) = C_{j,i}, \quad \alpha_i(d) = \delta_{0,i},$$
$$\Lambda_i(h_j) = \delta_{i,j}, \quad \Lambda_i(d) = 0.$$

Denote by $\Pi = \{\alpha_0, \dots, \alpha_n\} \subset \mathfrak{h}^*$ the set of simple roots and by $\Pi^\vee = \{h_0, \dots, h_n\} \subset \mathfrak{h}$ the set of simple coroots. Let $P = \{\lambda \in \mathfrak{h}^* \mid \lambda(h_i) \in \mathbb{Z} \text{ for any } i \in I\}$ be the weight lattice and $P^+ = \{\lambda \in P \mid \lambda(h_i) \geq 0 \text{ for any } i \in I\}$, the semigroup of dominant weights. Let $Q = \bigoplus_{i \in I} \mathbb{Z}\alpha_i \subset P$ (the root lattice) and $Q^+ = \sum_{i \in I} \mathbb{N}\alpha_i \subset Q$. For $\lambda, \mu \in \mathfrak{h}^*$, write $\lambda \geq \mu$ if $\lambda - \mu \in Q^+$.

Set $\mathfrak{h}_0 = \mathbb{Q}h_1 \oplus \cdots \oplus \mathbb{Q}h_n$ and $\Pi_0 = \{\alpha_1, \dots, \alpha_n\}$, $\Pi_0^\vee = \{h_1, \dots, h_n\}$. Then $(\mathfrak{h}_0, \Pi_0, \Pi_0^\vee)$ is a realization of $(C_{i,j})_{i,j \in I_0}$ (see [Kac 1990]). We define as above the associated weight lattice $P_0$, its subset $P_0^+$ of dominant weights, and the root lattice $Q_0$.

Denote by $W$ the affine Weyl group: it is the subgroup of $\mathrm{GL}(\mathfrak{h}^*)$ generated by the simple reflections $s_i \in \mathrm{GL}(\mathfrak{h}^*)$ defined by $s_i(\lambda) = \lambda - \lambda(h_i)\alpha_i$ $(i \in I)$. The Weyl group of finite type $W_0$ is the subgroup of $W$ generated by the $s_i$ with $i \in I_0$.

Let $c = h_0 + \cdots + h_n$ and $\delta = \alpha_0 + \cdots + \alpha_n$. We have

$$\{\omega \in P \mid \omega(h_i) = 0 \text{ for all } i \in I\} = \mathbb{Q}\delta.$$

Put $P_{\mathrm{cl}} = P/\mathbb{Q}\delta$ and denote by $\mathrm{cl} : P \to P_{\mathrm{cl}}$ the canonical projection. Denote by $P^0 = \{\lambda \in P \mid \lambda(c) = 0\}$ the set of level 0 weights.

**2B. *Quantum affine algebra $\mathcal{U}_q(\hat{\mathfrak{sl}}_{n+1})$.*** In this article $q = e^t \in \mathbb{C}^*$ $(t \in \mathbb{C})$ is not a root of unity and is fixed. For $l \in \mathbb{Z}, r \geq 0, m \geq m' \geq 0$ we set

$$[l]_q = \frac{q^l - q^{-l}}{q - q^{-1}} \in \mathbb{Z}[q^{\pm 1}],$$
$$[r]_q! = [r]_q[r-1]_q \dots [1]_q,$$
$$\begin{bmatrix} m \\ m' \end{bmatrix}_q = \frac{[m]_q!}{[m-m']_q!\,[m']_q!}.$$

**Definition 2.2.** The quantum affine algebra $\mathcal{U}_q(\hat{\mathfrak{sl}}_{n+1})$ is the $\mathbb{C}$-algebra with generators $k_h$ $(h \in \mathfrak{h})$, $x_i^{\pm}$ $(i \in I)$ and relations

$$k_h k_{h'} = k_{h+h'}, \, k_0 = 1,$$
$$k_h x_j^{\pm} k_{-h} = q^{\pm\alpha_j(h)} x_j^{\pm},$$
$$[x_i^+, x_j^-] = \delta_{i,j} \frac{k_i - k_i^{-1}}{q - q^{-1}},$$
$$(x_i^{\pm})^{(2)} x_{i+1}^{\pm} - x_i^{\pm} x_{i+1}^{\pm} x_i^{\pm} + x_{i+1}^{\pm} (x_i^{\pm})^{(2)} = 0.$$

Here we use the notation $k_i^{\pm 1} = k_{\pm h_i}$ and for all $r \geq 0$ we set $(x_i^{\pm})^{(r)} = (x_i^{\pm})^r/[r]_q!$. One defines a coproduct on $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ by setting

$$\Delta(k_h) = k_h \otimes k_h,$$
$$\Delta(x_i^+) = x_i^+ \otimes 1 + k_i^+ \otimes x_i^+, \quad \Delta(x_i^-) = x_i^- \otimes k_i^- + 1 \otimes x_i^-.$$

Let $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ be the subalgebra of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ generated by the $x_i^{\pm}$ and $k_h$, for $h \in \sum \mathbb{Q} h_i$. This has $P_{\mathrm{cl}}$ as a weight lattice.

For $J \subset I$ denote by $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})_J$ the subalgebra of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ generated by the $x_i^{\pm}, k_{ph_i}$ for $i \in J$, $p \in \mathbb{Q}$. If $J = I_0$, $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})_{I_0}$ is the quantum group of finite type associated to the data $(\mathfrak{h}_0, \Pi_0, \Pi_0^\vee)$, also denoted by $\mathcal{U}_q(\mathrm{sl}_{n+1})$. In particular, a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module has a structure of $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module. If $J = \{i\}$ with $i \in I$, $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})_J$ is isomorphic to $\mathcal{U}_q(\mathrm{sl}_2)$ and denoted by $\mathcal{U}_i$. So a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module has also a structure of $\mathcal{U}_q(\mathrm{sl}_2)$-module.

Let $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})^+$ (resp. $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})^-$, $\mathcal{U}_q(\mathfrak{h})$) be the subalgebra of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ generated by the $x_i^+$ (resp. the $x_i^-$, the $k_h$). We have a triangular decomposition of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ (see [Lusztig 1993]):

**Theorem 2.3.** *We have an isomorphism of vector spaces*

$$\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1}) \simeq \mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})^- \otimes \mathcal{U}_q(\mathfrak{h}) \otimes \mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})^+.$$

**2C.** ***Representations of*** $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$**.** For $V$ a representation of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ and $\nu \in P$, the weight space $V_\nu$ of $V$ is

$$V_\nu = \{v \in V \mid k_h \cdot v = q^{\nu(h)} v, \forall h \in \mathfrak{h}\}.$$

Set $\mathrm{wt}(V) = \{\nu \in P \mid V_\nu \neq \{0\}\}$.

For $\lambda \in P$, a representation $V$ is said to be of highest weight $\lambda$ if there is $v \in V_\lambda$ such that for all $i \in I$, $x_i^+ \cdot v = 0$ and $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1}) \cdot v = V$. Furthermore there is a unique simple highest weight module of highest weight $\lambda$.

**Definition 2.4.** A representation $V$ is said to be integrable if

(i) it admits a weight space decomposition $V = \bigoplus_{\nu \in P} V_\nu$,

(ii) all the $x_i^{\pm}$ $(i \in I)$ are locally nilpotent.

**Remark 2.5.** This definition differs from the one given in [Hernandez 2009]. In fact we require that the following additional conditions be satisfied:

(iii) $V_\nu$ is finite-dimensional for any $\nu \in P$.

(iv) $V_{\nu \pm N\alpha_i} = \{0\}$ for all $\nu \in P$, $N \gg 0$, $i \in I$.

These conditions are implied by the previous ones for the highest weight modules.

**Theorem 2.6** [Lusztig 1993]. *The simple highest weight module of highest weight $\lambda$ is integrable if and only if $\lambda$ is dominant. We denote it $V(\lambda)$ $(\lambda \in P^+)$.*

For an integrable representation $V$ of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ with finite-dimensional weight spaces, one defines the usual character

$$\chi(V) = \sum_{\nu \in P} \dim(V_\nu) e(\nu) \in \prod_{\nu \in P} \mathbb{Z} e(\nu).$$

Similar definitions hold for the quantum group $\mathcal{U}_q(\mathrm{sl}_{n+1})$. In this case, the integrable simple highest weight modules are parametrized by $P_0^+$ and denoted by $V_0(\lambda)$ ($\lambda \in P_0^+$). Further they are finite-dimensional (see [Lusztig 1993; Rosso 1991]). Let $\mathscr{C}$ be the category of integrable finite-dimensional representations of $\mathcal{U}_q(\mathrm{sl}_{n+1})$ and $\mathscr{R}$ its Grothendieck ring.

**Theorem 2.7** [Lusztig 1993; Rosso 1991]. *The category $\mathscr{C}$ is a semisimple tensor category and the simple objects of $\mathscr{C}$ are the $(V_0(\lambda))_{\lambda \in P_0^+}$. Furthermore $\chi$ induces a ring morphism*

$$\chi : \mathscr{R} \to \bigoplus_{\nu \in P_0} \mathbb{Z} e(\nu),$$

*where the product on the right is defined by $e(\mu)e(\nu) = e(\mu + \nu)$.*

We do not recall here the theory of crystal bases of quantum groups, we just refer to [Kashiwara 1994; 2002a; 2002b]. Let us remind only that for $\lambda \in P^+$, the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V(\lambda)$ has a crystal basis $\mathscr{B}(\lambda)$. In the same way we denote by $\mathscr{B}_0(\lambda)$ the crystal basis of the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module $V_0(\lambda)$ ($\lambda \in P_0^+$). When we want to distinguish crystals of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$, $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})_J$ with $J \subset I$ and $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$, we call it respectively a $P$-crystal or an $I$-crystal, a $J$-crystal and a $P_{\mathrm{cl}}$-crystal.

**2D.** ***Extremal weight modules.*** In this section we recall the definition and some properties of extremal weight modules for the quantum affine algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ given by Kashiwara [1994; 2002b]. All of these hold for general quantum Kac–Moody algebras and in particular for $\mathcal{U}_q(\mathrm{sl}_{n+1})$.

**Definition 2.8.** For an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V$ and $\lambda \in P$, a vector $v \in V_\lambda$ is called extremal of weight $\lambda$ if there are vectors $\{v_w\}_{w \in W}$ such that $v_{\mathrm{Id}} = v$ and

$$x_i^{\pm} \cdot v_w = 0 \quad \text{and} \quad (x_i^{\mp})^{(\pm w(\lambda)(h_i))} \cdot v_w = v_{s_i(w)} \quad \text{if } \pm w(\lambda)(h_i) \geq 0.$$

Note that if the vector $v$ is extremal of weight $\lambda$, then for $w \in W$, $v_w$ is extremal of weight $w(\lambda)$.

**Remark 2.9.** The definition of extremal vector can be rewritten as follows (see [Kashiwara 1994]): for an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V$, a weight vector $v$ of weight $\lambda$ is called $i$-extremal if $x_i^+ \cdot v = 0$ or $x_i^- \cdot v = 0$. In this case we set $S_i(v) = (x_i^-)^{(\lambda(h_i))} \cdot v$ or $S_i(v) = (x_i^+)^{(-\lambda(h_i))} \cdot v$ respectively. Then a weight vector $v$ is extremal if, for any $l \geq 0$, $S_{i_1} \circ \cdots \circ S_{i_l}(v)$ is $i$-extremal for any $i, i_1, \dots, i_l \in I$. We set in that case

$$W \cdot v = \{S_{i_1} \circ \cdots \circ S_{i_l}(v)v \mid l \in \mathbb{N}, i_1, \dots, i_l \in I\}.$$

The notion of extremal elements in a crystal $\mathcal{B}$ can be defined in the same way.

**Definition 2.10.** For $\lambda \in P$, the extremal weight module $V(\lambda)$ of extremal weight $\lambda$ is the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module generated by a vector $v_\lambda$ with the defining relations that $v_\lambda$ is extremal of weight $\lambda$.

**Example 2.11.** If $\lambda$ is dominant, $V(\lambda)$ is the simple highest weight module of highest weight $\lambda$.

**Theorem 2.12** [Kashiwara 1994]. *For $\lambda \in P$, the module $V(\lambda)$ is integrable and has a crystal basis $\mathcal{B}(\lambda)$.*

Set $\lambda = \varpi_\ell$, where $1 \leq \ell \leq n$ and $\varpi_\ell$ is the level 0 fundamental weight $\varpi_\ell = \Lambda_\ell - \Lambda_0$.

**Theorem 2.13** [Kashiwara 2002b]. *Let $1 \leq \ell \leq n$.*

(i) *$V(\varpi_\ell)$ is an irreducible $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module.*

(ii) *Any nonzero integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module generated by an extremal weight vector of weight $\varpi_\ell$ is isomorphic to $V(\varpi_\ell)$.*

Let $w$ be an element of $W$ such that $w(\varpi_\ell) = \varpi_\ell + \delta$. Such an element exists and is not unique (see [Kashiwara 2002b]). It defines a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-automorphism (also called $P_{\mathrm{cl}}$-automorphism in the following) of the restricted $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module $V(\varpi_\ell)$, which sends $v$ to $v_w$. It is of weight $\delta$, and denoted by $z_\ell$. Let us define the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module

$$W(\varpi_\ell) = V(\varpi_\ell)/(z_\ell - 1)V(\varpi_\ell).$$

**Theorem 2.14** [Kashiwara 2002b]. *Let $1 \leq \ell \leq n$.*

(i) *$W(\varpi_\ell)$ is a finite-dimensional irreducible $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module.*

(ii) *For any $\mu \in \mathrm{wt}(V(\varpi_\ell))$,*

$$W(\varpi_\ell)_{\mathrm{cl}(\mu)} \simeq V(\varpi_\ell)_\mu.$$

(iii) *$V(\varpi_\ell)$ is isomorphic to $W(\varpi_\ell)_{\mathit{aff}}$ as a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module.*

Here $M_{\mathrm{aff}}$ is the affinization of an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module $M$: this is the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module with a weight space decomposition $M_{\mathrm{aff}} = \bigoplus_{\nu \in P} (M_{\mathrm{aff}})_\nu$ defined by

$$(M_{\mathrm{aff}})_\nu = M_{\mathrm{cl}(\nu)}$$

and with the obvious action of $x_i^{\pm}$. Note also that we have an isomorphism of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-modules

$$M_{\mathrm{aff}} \simeq \mathbb{C}[z, z^{-1}] \otimes M,$$

where $x_i^{\pm}$ act on the right side by $z^{\pm\delta_{i,0}} x_i^{\pm}$. In the same way one defines the affinization $\mathcal{B}_{\mathrm{aff}}$ of a $P_{\mathrm{cl}}$-crystal $\mathcal{B}$. For an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module $M$ with associated $P_{\mathrm{cl}}$-crystal $\mathcal{B}$, the affinization $M_{\mathrm{aff}}$ has a $P$-crystal $\mathcal{B}_{\mathrm{aff}}$.

**2E. *Quantum toroidal algebra* $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$.** In this section, we recall the definition and the main properties of the quantum toroidal algebra $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ (without central charge).

**Definition 2.15** [Ginzburg et al. 1995]. The quantum toroidal algebra $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ is the $\mathbb{C}$-algebra with generators $x_{i,r}^{\pm}$ $(i \in I, r \in \mathbb{Z})$, $k_h$ $(h \in \mathfrak{h})$, $h_{i,m}$ $(i \in I, m \in \mathbb{Z}-\{0\})$ and the following relations $(i, j \in I, r, r', r_1, r_2 \in \mathbb{Z}, m \in \mathbb{Z}-\{0\})$:

$$k_h k_{h'} = k_{h+h'}, \quad k_0 = 1, \quad [k_h, h_{j,m}] = 0, \quad [h_{i,m}, h_{j,m'}] = 0,$$

$$k_h x_{j,r}^{\pm} k_{-h} = q^{\pm\alpha_j(h)} x_{j,r}^{\pm},$$

$$[h_{i,m}, x_{j,r}^{\pm}] = \pm\frac{1}{m}[mC_{i,j}]_q x_{j,m+r}^{\pm},$$

$$[x_{i,r}^{+}, x_{j,r'}^{-}] = \delta_{ij} \frac{\phi_{i,r+r'}^{+} - \phi_{i,r+r'}^{-}}{q - q^{-1}}, \tag{1}$$

$$x_{i,r+1}^{\pm} x_{j,r'}^{\pm} - q^{\pm C_{ij}} x_{j,r'}^{\pm} x_{i,r+1}^{\pm} = q^{\pm C_{ij}} x_{i,r}^{\pm} x_{j,r'+1}^{\pm} - x_{j,r'+1}^{\pm} x_{i,r}^{\pm},$$

$$\begin{aligned} x_{i,r_1}^{\pm} x_{i,r_2}^{\pm} x_{i\pm1,r'}^{\pm} - (q+q^{-1}) x_{i,r_1}^{\pm} x_{i\pm1,r'}^{\pm} x_{i,r_2}^{\pm} + x_{i\pm1,r'}^{\pm} x_{i,r_1}^{\pm} x_{i,r_2}^{\pm} \\ = -x_{i,r_2}^{\pm} x_{i,r_1}^{\pm} x_{i\pm1,r'}^{\pm} + (q+q^{-1}) x_{i,r_2}^{\pm} x_{i\pm1,r'}^{\pm} x_{i,r_1}^{\pm} - x_{i\pm1,r'}^{\pm} x_{i,r_2}^{\pm} x_{i,r_1}^{\pm}, \end{aligned}$$

and $[x_{i,r_1}^{\pm}, x_{j,r_2}^{\pm}] = 0$ if $i \neq j, j \pm 1$. Here for all $i \in I$ and $m \in \mathbb{Z}$, $\phi_{i,m}^{\pm} \in \mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ is determined by the formal power series in $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})[\![z^{\pm1}]\!]$

$$\phi_i^{\pm}(z) = \sum_{m\geq0} \phi_{i,\pm m}^{\pm} z^{\pm m} = k_i^{\pm1} \exp\biggl(\pm(q-q^{-1}) \sum_{m'\geq1} h_{i,\pm m'} z^{\pm m'}\biggr),$$

and $\phi_{i,m}^{+} = 0$ for $m < 0$, $\phi_{i,m}^{-} = 0$ for $m > 0$.

There is an algebra morphism $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1}) \to \mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ defined by $k_h \mapsto k_h$, $x_i^{\pm} \mapsto x_{i,0}^{\pm}$ $(h \in \mathfrak{h}, i \in I)$. Its image is called the horizontal quantum affine subalgebra of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ and is denoted by $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$. In particular, a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V$ has also a structure of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module. We denote by $\mathrm{Res}(V)$ the restricted $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module obtained from $V$.

As said above, the quantum affine algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ has another realization in terms of Drinfeld generators [Beck 1994; Drinfeld 1987]: this is the $\mathbb{C}$-algebra with generators $x^{\pm}_{i,r}$ ($i \in I_0, r \in \mathbb{Z}$), $k_h$ ($h \in \mathfrak{h}_0$), $h_{i,m}$ ($i \in I_0, m \in \mathbb{Z}-\{0\}$) and the same relations as in Definition 2.15. It is isomorphic to the subalgebra $\mathcal{U}^v_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ generated by the $x^{\pm}_{i,r}$, $k_h$, $h_{i,m}$ ($i \in I_0, r \in \mathbb{Z}, h \in \mathfrak{h}_0, m \in \mathbb{Z}-\{0\}$). $\mathcal{U}^v_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ is called the vertical quantum affine subalgebra of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$.

For all $j \in I$, set $I_j = I-\{j\}$ and define the subalgebra $\mathcal{U}^{v,j}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ generated by the $x^{\pm}_{i,r}$, $k_h$, $h_{i,m}$ ($i \in I_j, r \in \mathbb{Z}, h \in \bigoplus_{i\in I_j} \mathbb{Q}h_i, m \in \mathbb{Z}-\{0\}$). In particular $\mathcal{U}^{v,0}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ is the vertical quantum affine subalgebra $\mathcal{U}^v_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$. All the $\mathcal{U}^{v,j}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ for various $j \in I$ are isomorphic: in fact let $\theta$ be the automorphism of the Dynkin diagram of type $A^{(1)}_n$ corresponding to the rotation such that $\theta(k) = k+1$, where $I$ is identified to the set $\mathbb{Z}/(n+1)\mathbb{Z}$. It defines an automorphism $\theta_{\mathfrak{h}}$ of $\mathfrak{h}$ by sending $h_i, d$ to $h_{\theta(i)}, d$ ($i \in I$). For all $j \in J$, let $\theta^{(j)}$ be the automorphism of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ which sends $x^{\pm}_{i,r}$, $k_h$, $h_{i,m}$ to $x^{\pm}_{\theta^j(i),r}$, $k_{\theta^j_{\mathfrak{h}}(h)}$, $h_{\theta^j(i),m}$ respectively (where $i \in I$, $h \in \mathfrak{h}$, $r \in \mathbb{Z}$, $m \in \mathbb{Z}-\{0\}$). It gives by restriction an isomorphism of algebras between $\mathcal{U}^v_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ and $\mathcal{U}^{v,j}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$, still denoted by $\theta^{(j)}$ in the following. If $V$ is a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module, we denote by $V^{(j)}$ the induced $\mathcal{U}^{v,j}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$-module.

For $i \in I$, the subalgebra $\hat{\mathcal{U}}_i$ generated by the $x^{\pm}_{i,r}, h_{i,m}, k_{ph_i}$ ($r \in \mathbb{Z}, m \in \mathbb{Z}-\{0\}$, $p \in \mathbb{Q}$) is isomorphic to $\mathcal{U}_q(\hat{\mathrm{sl}}_2)'$.

We have a triangular decomposition of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$.

**Theorem 2.16** [Miki 2000; Nakajima 2001]. *We have an isomorphism of vector spaces*

$$\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1}) \simeq \mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})^- \otimes \mathcal{U}_q(\hat{\mathfrak{h}}) \otimes \mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})^+,$$

*where $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})^{\pm}$ (resp. $\mathcal{U}_q(\hat{\mathfrak{h}})$) is generated by the $x^{\pm}_{i,r}$ (resp. the $k_h$, the $h_{i,m}$).*

## 2F. *Representations of* $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$.

**Definition 2.17.** A representation $V$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ is said to be integrable if $\mathrm{Res}(V)$ is integrable as a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module.

**Definition 2.18.** A representation $V$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ is said to be of $\ell$-highest weight if there is $v \in V$ such that

(i) $V = \mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})^- \cdot v$,

(ii) $\mathcal{U}_q(\hat{\mathfrak{h}}) \cdot v = \mathbb{C}v$,

(iii) for any $i \in I, r \in \mathbb{Z}$, $x^+_{i,r} \cdot v = 0$.

For $\gamma \in \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})$ an algebra morphism, by Theorem 2.16 we have a corresponding Verma module $M(\gamma)$ and a simple representation $V(\gamma)$ which are $\ell$-highest weight. Then we have:

**Theorem 2.19** [Miki 2000; Nakajima 2001]. *The simple representations $V(\gamma)$ of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ are integrable if there is $(\lambda, (P_i)_{i\in I}) \in P^+ \times (1+u\mathbb{C}[u])^I$ satisfying $\gamma(k_h) = q^{\lambda(h)}$ and the following relation in $\mathbb{C}[\![z^{\pm 1}]\!]$, for each $i \in I$:*

$$\gamma(\phi_i^{\pm}(z)) = q^{\deg(P_i)} \frac{P_i(zq^{-1})}{P_i(zq)}.$$

The polynomials $P_i$ are called Drinfeld polynomials and the representation $V(\gamma)$ is then denoted by $V(\lambda, (P_i)_{i\in I})$. Such a representation is also integrable in the sense of [Hernandez 2009], that is, $V(\lambda, (P_i)_{i\in I})$ satisfies conditions (iii) and (iv) of Remark 2.5.

The Kirillov–Reshetikhin module associated to $k \geq 0$, $a \in \mathbb{C}^*$ and $0 \leq \ell \leq n$ is the simple integrable representation of weight $k\Lambda_\ell$ with the $n$-tuple

$$P_i(u) = \begin{cases} (1-ua)(1-uaq^2)\dots(1-uaq^{2(k-1)}) & \text{for } i = \ell, \\ 1 & \text{for } i \neq \ell. \end{cases}$$

If $k = 1$, it is also called the fundamental module.

Consider an integrable representation $V$ of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$. As the subalgebra $\mathcal{U}_q(\hat{\mathfrak{h}})$ is commutative, we have a decomposition of the weight spaces $V_\nu$ in simultaneous generalized eigenspaces

$$V_\nu = \bigoplus_{\substack{\nu \in P \\ \gamma \in \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})}} V_{(\nu,\gamma)},$$

where $V_{(\nu,\gamma)} = \{x \in V : \exists p \in \mathbb{N}, \forall i \in I, \forall m \geq 0, (\phi_{i,\pm m}^{\pm} - \gamma(\phi_{i,\pm m}^{\pm}))^p \cdot x = 0\}$. If $V_{(\nu,\gamma)} \neq \{0\}$, then $(\nu, \gamma)$ is called an $\ell$-weight of $V$.

**Definition 2.20.** A $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V$ is weighted if the Cartan subalgebra $\mathcal{U}_q(\hat{\mathfrak{h}})$ acts on $V$ by diagonalizable operators. The module $V$ is thin if it is weighted and the joint spectrum is simple.

The terminology is different in [Feigin et al. 2011a; 2011b; 2012; 2013]: a thin module is called tame.

**Definition 2.21** [Frenkel and Reshetikhin 1999; Hernandez 2005; Nakajima 2001]. The $q$-character of an integrable representation $V$ of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ with finite-dimensional $\ell$-weight spaces is defined by the formal sum

$$\chi_q(V) = \sum_{\substack{\nu \in P \\ \gamma \in \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})}} \dim(V_{(\nu,\gamma)}) e(\nu, \gamma).$$

Furthermore if the weight spaces of $V$ are finite-dimensional we have

$$\chi(\mathrm{Res}(V)) = \beta(\chi_q(V)),$$

where $\mathrm{Res}(V)$ still denotes the restricted $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module obtained from $V$, and

$$\beta : \prod_{\substack{\nu \in P \\ \gamma \in \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})}} \mathbb{Z}e(\nu, \gamma) \to \prod_{\nu \in P} \mathbb{Z}e(\nu)$$

is $\mathbb{Z}$-linear such that $\beta(e(\nu, \gamma)) = e(\nu)$ for all $(\nu, \gamma) \in P \times \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})$.

**Proposition 2.22** [Frenkel and Reshetikhin 1999; Hernandez 2005; Nakajima 2001]. *Let $V$ be an integrable representation of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ and consider an $\ell$-weight $(\nu, \gamma) \in P \times \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}), \mathbb{C})$ of $V$. Then there exist polynomials $Q_i(z), R_i(z) \in \mathbb{C}[z]$ $(i \in I)$ of constant term $1$ such that*

$$\sum_{m \geq 0} \gamma(\phi^{\pm}_{i,\pm m}) z^{\pm m} = q^{\deg(Q_i) - \deg(R_i)} \frac{Q_i(zq^{-1}) R_i(zq)}{Q_i(zq) R_i(zq^{-1})} \tag{2}$$

*in $\mathbb{C}[\![z^{\pm 1}]\!]$. Furthermore, if $V$ has a finite composition series*

$$L_0 = \{0\} \subset L_1 \subset L_2 \subset \cdots \subset L_k = V$$

*such that $L_{j+1}/L_j \simeq V(\lambda_j, (P_i^j)_{i \in I})$, where the roots of $P_i^j$ are in $q^{\mathbb{Z}}$ for all $i \in I$, $0 \leq j \leq k-1$, then*

(ii) *there exist $\omega \in P^+, \alpha \in Q^+$ satisfying $\nu = \omega - \alpha$,*

(iii) *the zeros of the polynomials $Q_i(z), R_i(z)$ are in $q^{\mathbb{Z}}$.*

If $V$ is a Kirillov–Reshetikhin module, one reduces to the case where the defining parameter $a$ is in $q^{\mathbb{Z}}$ by twisting the action by the automorphisms $t_b$ of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_{n+1})$ given by $(b \in \mathbb{C}^*)$

$$t_b(x^{\pm}_{i,r}) = b^r x^{\pm}_{i,r}, \quad t_b(h^{\pm}_{i,m}) = b^m h^{\pm}_{i,m}, \quad t_b(k_h) = k_h.$$

Consider formal variables $Y^{\pm 1}_{i,l}$, $e^{\nu}$ $(i \in I,\ l \in \mathbb{Z},\ \nu \in P)$ and let $A$ be the group of monomials of the form $m = e^{\omega(m)} \prod_{i \in I, l \in \mathbb{Z}} Y_{i,l}^{u_{i,l}(m)}$, where $u_{i,l}(m) \in \mathbb{Z}$, $\omega(m) \in P$ are such that

$$\sum_{l \in \mathbb{Z}} u_{i,l}(m) = \omega(m)(h_i).$$

For example, $e^{\pm \Lambda_i} Y^{\pm 1}_{i,l} \in A$ and $A_{i,l} = e^{\alpha_i} Y_{i,l-1} Y_{i,l+1} Y^{-1}_{i-1,l} Y^{-1}_{i+1,l} \in A$. A monomial $m$ is said to be $J$-dominant $(J \subset I)$ if for all $j \in J$ and $l \in \mathbb{Z}$ we have $u_{j,l}(m) \geq 0$. An $I$-dominant monomial is said to be dominant.

**Remark 2.23.** Let us fix a monomial $m \in A$ and consider monomials $m'$ which are products of $m$ with various $A^{\pm 1}_{i,l}$ $(i \in I, l \in \mathbb{Z})$. By [Hernandez and Nakajima 2006, Remark 2.1], $\omega(m')$ is uniquely determined by $\omega(m)$ and $u_{i,l}(m')$. So in the following when we are in this situation, the term $e^{\omega(m')}$ will be safely omitted.

Let $V$ be an integrable $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module such that for all $\ell$-weight $(\nu,\gamma)$ of $V$, the roots of the associated polynomials $Q_i(z)$ and $R_i(z)$ are in $q^{\mathbb{Z}}$ for all $i \in I$. For $(\nu,\gamma) \in P \times \mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}),\mathbb{C})$ an $\ell$-weight of $V$, one defines the monomial $m_{(\nu,\gamma)} = e^{\nu} \prod_{i\in I, l\in\mathbb{Z}} Y_{i,l}^{u_{i,l}-v_{i,l}}$, where

$$Q_i(z) = \prod_{l\in\mathbb{Z}} (1-zq^l)^{u_{i,l}} \quad \text{and} \quad R_i(z) = \prod_{l\in\mathbb{Z}} (1-zq^l)^{v_{i,l}}.$$

We denote $V_{(\nu,\gamma)} = V_{m_{(\nu,\gamma)}}$. We rewrite the $q$-character of an integrable representation $V$ with finite-dimensional $\ell$-weight spaces by the formal sum

$$\chi_q(V) = \sum_m \dim(V_m) m \in \mathbb{Z}[\![e^{\nu}, Y_{i,l}^{\pm 1}]\!]_{\nu\in P, i\in I, l\in\mathbb{Z}}.$$

Let us denote by $\mathcal{M}(V)$ the set of monomials occurring in $\chi_q(V)$.

By this correspondence between $\ell$-weights and monomials due to Frenkel and Reshetikhin [1999], the $I$-tuple of Drinfeld polynomials with zeros in $q^{\mathbb{Z}}$ are identified with the dominant monomials. In particular for a dominant monomial $m$, one denotes by $V(m)$ the simple module of $\ell$-highest weight $m$. For example $V(e^{k\Lambda_\ell} Y_{\ell,s} Y_{\ell,s+2} \dots Y_{\ell,s+2(k-1)})$ is the Kirillov–Reshetikhin module associated to $k \geq 0$, $a = q^s \in \mathbb{C}^*$ ($s \in \mathbb{Z}$) and $\ell \in I$, and $V(e^{\Lambda_\ell} Y_{\ell,s})$ is the fundamental module associated to $a = q^s \in \mathbb{C}^*$ ($s \in \mathbb{Z}$) and $\ell \in I$.

Similar results hold for the quantum affine algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ due to Chari and Pressley [1994]. In this case, the simple integrable representations are finite-dimensional and denoted $V_0((P_i)_{i\in I_0})$ in the following. Note that the weights $\lambda \in P_0$ can be omitted here because they are completely determined by the Drinfeld polynomials $(P_i)_{i\in I_0}$,

$$\lambda = \deg(P_1)\Lambda_1 + \dots + \deg(P_n)\Lambda_n.$$

In the same way if $V$ is a Kirillov–Reshetikhin module of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$, its $\ell$-weights can be only considered as elements of $\mathrm{Hom}(\mathcal{U}_q(\hat{\mathfrak{h}}_0),\mathbb{C})$ (where $\mathcal{U}_q(\hat{\mathfrak{h}}_0)$ is the subalgebra of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ generated by $k_h$ ($h \in \mathfrak{h}_0$) and $h_{i,m}$ ($i \in I_0, m \in \mathbb{Z} - \{0\}$)). They still satisfy the relations in (2). By twisting the action on $V$ by an automorphism $t_b$ of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ for some $b \in \mathbb{C}^*$, it can be parametrized as above by Laurent monomials in $\mathbb{Z}[Y_{i,l}^{\pm 1}]_{i\in I_0, l\in\mathbb{Z}}$. The weight $\omega(m) \in P_0$ of a monomial $m \in \mathbb{Z}[Y_{i,l}^{\pm 1}]_{i\in I_0, l\in\mathbb{Z}}$ can be omitted here because it is completely determined by the $u_{i,l}(m)$ ($i \in I_0, l \in \mathbb{Z}$).

Recall that the Kirillov–Reshetikhin modules $V_0(Y_{\ell,s} Y_{\ell,s+2} \dots Y_{\ell,s+2(k-1)})$ ($k \geq 0$, $s \in \mathbb{Z}$, $\ell \in I_0$) over $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ can be obtained from the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-modules $V_0(k\Lambda_\ell)$ as follows: there exist evaluation morphisms $\mathrm{ev}_a : \mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})' \to \mathcal{U}_q(\mathrm{sl}_{n+1})$ ($a \in \mathbb{C}^*$) which send $x_{i,0}, k_h$ on $x_i, k_h$ respectively ($i \in I_0, h \in \mathfrak{h}_0$). So the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module $V_0(Y_{\ell,s} Y_{\ell,s+2} \dots Y_{\ell,s+2(k-1)})$ is obtained by pulling back the action of $\mathcal{U}_q(\mathrm{sl}_{n+1})$ on $V_0(k\Lambda_\ell)$ by $\mathrm{ev}_a$ for some $a \in \mathbb{C}^*$. In particular,

$V_0(Y_{\ell,s}Y_{\ell,s+2}\dots Y_{\ell,s+2(k-1)})$ is irreducible as a $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module, isomorphic to $V_0(k\Lambda_\ell)$.

We have defined irreducible finite-dimensional $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-modules $W(\varpi_\ell)$ ($\ell \in I_0$) in Section 2D. One can determine them in terms of the Drinfeld realization. For that we need the following additional result.

**Lemma 2.24** [Nakajima 2004]. *Let $v$ be a vector of an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module of weight $\lambda \in \mathrm{cl}(P^0)$ such that for all $i \in I_0$, $\lambda(h_i) \geq 0$. Then the following conditions are equivalent*:

(i) *$v$ is an extremal vector.*

(ii) *$x^+_{i,r} \cdot v = 0$ for all $i \in I_0, r \in \mathbb{Z}$.*

As a direct consequence of these results, $W(\varpi_\ell)$ is isomorphic to a fundamental representation $V_0((1-\delta_{\ell,i}au)_{i\in I_0})$ for a special choice of the spectral parameter $a \in \mathbb{C}^*$ (see [Nakajima 2004, Remark 3.3] for an expression of it). In particular for this spectral parameter, one deduces that $V_0((1-\delta_{\ell,i}au)_{i\in I_0})$ has a crystal basis.

Let $\mathscr{C}_l$ be the category of finite-dimensional $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-modules (of type 1) and $\mathscr{R}_l$ its Grothendieck ring. Recall that $\mathscr{C}_l$ is an abelian monoidal category, not semisimple, with as simple objects the $V_0((P_i)_{i\in I_0})$ and $\mathscr{R}_l$ is the polynomial ring over $\mathbb{Z}$ in the classes $[V_0((1-\delta_{\ell,i}au)_{i\in I_0})]$ ($\ell \in I_0, a \in \mathbb{C}^*$) (see [Chari and Pressley 1994; Frenkel and Reshetikhin 1999]). As in [Hernandez and Leclerc 2010], we consider $\mathscr{C}_{l,\mathbb{Z}}$ the full subcategory of $\mathscr{C}_l$ whose objects $V$ satisfy:

> For every composition factor $S$ of $V$, the roots of the Drinfeld polynomials $(P_i(u))_{i\in I_0}$ belong to $q^{\mathbb{Z}}$.

This is also an abelian monoidal category, not semisimple and the Grothendieck ring $\mathscr{R}_{l,\mathbb{Z}}$ of $\mathscr{C}_{l,\mathbb{Z}}$ is the subring of $\mathscr{R}_l$ generated by the classes $[V_0(Y_{\ell,s})]$ with $\ell \in I_0, s \in \mathbb{Z}$ (see [Frenkel and Mukhin 2001]).

**Theorem 2.25** [Frenkel and Reshetikhin 1999]. *$\chi_q$ induces a ring morphism $\chi_q : \mathscr{R}_{l,\mathbb{Z}} \to \mathbb{Z}[Y^{\pm 1}_{i,l}]_{i\in I_0, l\in\mathbb{Z}}$, called the morphism of $q$-characters. Furthermore we have the commutative diagram*

$$\begin{array}{ccc}
\mathscr{R}_{l,\mathbb{Z}} & \xrightarrow{\ \chi_q\ } & \mathbb{Z}[Y^{\pm 1}_{i,l}]_{i\in I_0,l\in\mathbb{Z}} \\
\Big\downarrow{\scriptstyle \mathrm{Res}} & & \Big\downarrow{\scriptstyle \beta} \\
\mathscr{R} & \xrightarrow{\ \chi\ } & \bigoplus_{\nu\in P_0} \mathbb{Z}e(\nu)
\end{array}$$

*where the ring morphism* $\mathrm{Res} : \mathscr{R}_{l,\mathbb{Z}} \to \mathscr{R}$ *is the restriction and $\beta$ is defined by $\beta(m) = e(\omega(m))$.*

One does not have an expression of $q$-character of a representation in general. But explicit formulas exist for the fundamental modules and the Kirillov–Reshetikhin modules over $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ and $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$, given in terms of tableaux [Hernandez 2011; Nakajima 2003].

**2G. *Extremal loop weight modules.*** We recall the notion of extremal loop weight modules for $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$. The main motivation for this is the construction of finite-dimensional representations of the quantum toroidal algebra as in the theory of Kashiwara, but at roots of unity in this case.

**Definition 2.26** [Hernandez 2009]. An extremal loop weight module of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ is an integrable representation $V$ such that there is an $\ell$-weight vector $v \in V$ satisfying:

(i) $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \cdot v = V$.

(ii) $v$ is extremal for $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$.

(iii) $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \cdot w$ is finite-dimensional for all $w \in V$ and $j \in I$.

**Example 2.27.** If $m$ is dominant, the simple $\ell$-highest weight module $V(m)$ of $\ell$-highest weight $m$ is an extremal loop weight module.

An example of such a representation which is neither of $\ell$-highest weight nor of $\ell$-lowest weight is given in [Hernandez 2009]. The goal of this article is to construct a new family of extremal loop weight modules, called *extremal fundamental loop weight modules*.

## 3. Study of the monomial crystals $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$

We will relate in our paper the monomial realizations $\mathcal{M}_\ell$ of level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$ ($1 \leq \ell \leq n$) of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ with integrable representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$. In this section, we study the combinatorics of these monomial realizations, the main point being the use of promotion operators for level 0 extremal fundamental weight crystals introduced below. This is the first step of the construction of integrable modules associated to $\mathcal{M}_\ell$.

In Section 3A, one gives the definition of the monomial $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystal $\mathcal{M}$ [Kashiwara 2003; Nakajima 2003]. This definition holds when the considered Cartan matrix has no odd cycle. So it does not work for $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ when $n$ is even, and we have to assume that $n = 2r + 1$ ($r \geq 1$) is odd until the end of the article. Following [Hernandez and Nakajima 2006], we recall the monomial realization of $\mathcal{B}(\varpi_\ell)$ ($1 \leq \ell \leq n$): it is isomorphic to the sub-$\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystal

$$\mathcal{M}_\ell = \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$$

of $\mathcal{M}$ generated by the monomial $e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}$ (with $d_\ell$ equal to $\min(\ell, n+1-\ell)$). Furthermore we define the notions of a $q$-closed monomial set and of a monomial set closed by the Kashiwara operators, respectively related to the theory of $q$-characters and to the combinatorics of crystals.

The monomial crystals $\mathcal{M}_\ell$ have already been studied in [Hernandez and Nakajima 2006]: the monomials occurring in these crystals are explicitly given for $1 \le \ell \le n$ and their automorphisms $z_\ell$ are described in terms of monomials. We recall these results in Section 3B.

In Section 3C, we introduce promotion operators for level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$ $(1 \le \ell \le n)$. The promotion operators were introduced in [Shimozono 2002] for the Young tableaux realization of the finite $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystals $\mathcal{B}_0(k\Lambda_\ell)$ $(k \in \mathbb{N}^*,\ 1 \le \ell \le n)$ and studied in numerous papers (see [Bandlow et al. 2010; Fourier et al. 2009; Okado and Schilling 2008; Schilling 2008; Shimozono 2002] and references therein). It is the counterpart at the level of crystals of the cyclic symmetry of the Dynkin diagram of type $A_n^{(1)}$. After recalling these definitions, we extend the promotion operators for the level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$. Finally we specify the promotion operator of $\mathcal{B}(\varpi_\ell)$ in its monomial realization $\mathcal{M}_\ell$.

In Section 3D, we use promotion operators to obtain a new description of $\mathcal{M}_\ell$. In particular, we improve results given in [Hernandez and Nakajima 2006] for these crystals. Furthermore we determine the $\ell \in I_0$ for which the monomial crystals $\mathcal{M}_\ell$ are closed (Theorem 3.22): this is the case if and only if $\ell = 1, r+1$ or $n$.

**3A.** ***Monomial crystals.*** In this section we define the monomial crystal $\mathcal{M}$ of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ when $n = 2r+1$ is supposed to be odd, following [Kashiwara 2003; Nakajima 2003]. Monomial realizations of the crystals $\mathcal{B}(\lambda)$ with $\lambda \in P$, in particular of $\mathcal{B}(\varpi_\ell)$ $(1 \le \ell \le n)$, are studied in [Hernandez and Nakajima 2006; Kashiwara 2003; Nakajima 2003]. They are obtained as subcrystals of $\mathcal{M}$ generated by a monomial. We recall these results here. Finally we introduce new notions of $q$-closed monomial set and of monomial set closed by the Kashiwara operators.

As we have said above, the definition of the monomial crystal $\mathcal{M}$ requires that the considered Cartan matrix $C$ is without odd cycle. So we assume that $n = 2r+1$ $(r \ge 1)$ is odd until the end of the article. In particular there is a function $s : I \to \{0,1\}, i \mapsto s_i$ such that $C_{i,j} = -1$ implies $s_i + s_j = 1$.

Consider the subgroup $\mathcal{M} \subset A$ defined by

$$\mathcal{M} = \{m \in A \mid u_{i,l}(m) = 0 \text{ if } l \equiv s_i + 1 \bmod 2\}.$$

Following [Kashiwara 2003; Nakajima 2003], let us define $\mathrm{wt} : \mathcal{M} \to P$ by

$$\mathrm{wt}(m) = \omega(m),$$

and $\varepsilon_i, \varphi_i, p_i, q_i : \mathcal{M} \to \mathbb{Z} \cup \{\infty\} \cup \{-\infty\}$ for $i \in I$ by ($m \in \mathcal{M}$)

$$\varphi_{i,L}(m) = \sum_{l \le L} u_{i,l}(m), \quad \varphi_i(m) = \max\{\varphi_{i,L}(m) \mid L \in \mathbb{Z}\} \ge 0,$$
$$\varepsilon_{i,L}(m) = -\sum_{l \ge L} u_{i,l}(m), \quad \varepsilon_i(m) = \max\{\varepsilon_{i,L}(m) \mid L \in \mathbb{Z}\} \ge 0,$$
$$\begin{aligned} p_i(m) &= \max\{L \in \mathbb{Z} \mid \varepsilon_{i,L}(m) = \varepsilon_i(m)\} \\ &= \max\Big\{L \in \mathbb{Z} \;\Big|\; \sum_{l<L} u_{i,l}(m) = \varphi_i(m)\Big\}, \\ q_i(m) &= \min\{L \in \mathbb{Z} \mid \varphi_{i,L}(m) = \varphi_i(m)\} \\ &= \min\Big\{L \in \mathbb{Z} \;\Big|\; -\sum_{l>L} u_{i,l}(m) = \varepsilon_i(m)\Big\}. \end{aligned}$$

Then we define $\tilde{e}_i, \tilde{f}_i : \mathcal{M} \to \mathcal{M} \cup \{0\}$ for $i \in I$ by

$$\tilde{e}_i \cdot m = \begin{cases} 0 & \text{if } \varepsilon_i(m) = 0, \\ mA_{i,p_i(m)-1} & \text{if } \varepsilon_i(m) > 0, \end{cases}$$
$$\tilde{f}_i \cdot m = \begin{cases} 0 & \text{if } \varphi_i(m) = 0, \\ mA^{-1}_{i,q_i(m)+1} & \text{if } \varphi_i(m) > 0. \end{cases}$$

**Theorem 3.1** [Kashiwara 2003; Nakajima 2003]. *The crystal* $(\mathcal{M}, \mathrm{wt}, \varepsilon_i, \varphi_i, \tilde{e}_i, \tilde{f}_i)$ *is a* $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$*-crystal, called the monomial crystal.*

**Remark 3.2.** When $n$ is even, the Dynkin diagram of type $A_n^{(1)}$ is not bipartite. In this case, $(\mathcal{M}, \mathrm{wt}, \varepsilon_i, \varphi_i, \tilde{e}_i, \tilde{f}_i)$ does not satisfy the axioms of crystal (see [Kashiwara 2003]). Other crystal structures are defined on (a subset of) $A$ in [Kashiwara 2003]. But the monomials used are different with those occurring in the theory of $q$-characters of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-modules and it is not useful for what we will do in the next sections.

For $m \in \mathcal{M}$ denote by $\mathcal{M}(m)$ the connected subcrystal of $\mathcal{M}$ generated by $m$. As it is explained above, the weight of a monomial $m' \in \mathcal{M}(m)$ is determined by $\omega(m)$ and $u_{i,l}(m')$ (Remark 2.23). So we will omit the term $e^{\omega(m')}$ and we just specify the weight of the monomial $m$. For $J \subset I$ and $m \in \mathcal{M}$, denote by $\mathcal{M}_J(m)$ the set of monomials obtained from $m$ by applying the Kashiwara operators $\tilde{e}_i, \tilde{f}_i$ for $i \in J$. It is a connected sub-$J$-crystal of $\mathcal{M}(m)$ generated by $m$.

For $p \in \mathbb{Z}$ and $\alpha \in \mathbb{Z}\delta$, let $\tau_{2p,\alpha}$ be the map $\tau_{2p,\alpha} : \mathcal{M} \to \mathcal{M}$ defined by

$$\tau_{2p,\alpha}\Big(e^{\lambda} \prod Y_{i,n}^{u_{i,n}}\Big) = e^{\lambda+\alpha} \prod Y_{i,n+2p}^{u_{i,n}}.$$

This is a $P_{\mathrm{cl}}$-crystal automorphism of the crystal $\mathcal{M}$, also called shift automorphism in the following.

The following result was proved in [Kashiwara 2003; Nakajima 2003] when $m$ is a dominant monomial and is generalized in [Hernandez and Nakajima 2006] for all $m \in \mathcal{M}$.

**Theorem 3.3.** *For $m \in \mathcal{M}$, the crystal $\mathcal{M}(m)$ is isomorphic to a connected component of the crystal $\mathcal{B}(\lambda)$ of an extremal weight module for some $\lambda \in P$.*

It was shown in [Kashiwara 2002b] that the fundamental extremal crystals $\mathcal{B}(\varpi_\ell)$ are connected for all $\ell \in I_0$. Let

$$d_\ell = \min(\ell, n+1-\ell)$$

be the distance between the nodes 0 and $\ell$ in the Dynkin diagram of type $A_n^{(1)}$. We have the following monomial realization of $\mathcal{B}(\varpi_\ell)$.

**Theorem 3.4** [Hernandez and Nakajima 2006]. *Set $M = e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}$ for $\ell \in I_0$. Then $M$ is extremal in $\mathcal{M}$ and $\mathcal{M}(M) \simeq \mathcal{B}(\varpi_\ell)$ as $P$-crystals.*

One can define in the same way the monomial crystal $\mathcal{M}_0$ associated to $\mathcal{U}_q(\mathrm{sl}_{n+1})$. It can be done for all $n \geq 2$, the Cartan matrix of type $A_n$ being without cycle. As it is said above, the weights of monomials are completely determined by the powers of variables $Y_{i,l}^{\pm 1}$ in this case. So they can be safely omitted. For $m \in \mathcal{M}_0$, we denote by $\mathcal{M}_0(m)$ the subcrystal of $\mathcal{M}_0$ generated by $m$. We have:

**Proposition 3.5** [Kashiwara 2003; Nakajima 2003]. *The $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystals $\mathcal{M}_0(Y_{i,k})$ and $\mathcal{B}_0(\Lambda_i)$ are isomorphic for all $i \in I_0$ and $k \in \mathbb{Z}$.*

For $i \in I$, set $\Xi_i : \mathcal{M} \to \mathcal{M}$ the map sending the variables $Y_{j,*}^{\pm 1}, e^\nu$ to 1 for all $j \neq i$ and $\nu \in P$, and $Y_{i,*}^{\pm 1}$ to themselves. Another map will be used below: $\Xi^i : \mathcal{M} \to \mathcal{M}$, which sends the variables $Y_{j,*}^{\pm 1}$ to themselves if $j \neq i$ and $Y_{i,*}^{\pm 1}, e^\nu$ to 1 for all $\nu \in P$. These two maps are also defined in [Frenkel and Mukhin 2001] and denoted by $\beta_{\{i\}}$ and $\beta_{I_i}$ respectively.

**Definition 3.6.** (i) A set of monomials $\mathcal{S} \subset \mathcal{M}$ is said to be $q$-closed in the direction $i$ ($i \in I$) if for all $m \in \mathcal{S}$ there exists a finite subset

$$\mathcal{S}_m \subset \mathcal{S} \cap \left( m \cdot \prod_{l \in \mathbb{Z}} A_{i,l}^{\mathbb{Z}} \right),$$

which contains $m$, and a sequence $(n_s)_{s \in \mathcal{S}_m}$ of positive integers such that $\Xi_i\left(\sum_{s \in \mathcal{S}_m} n_s \cdot s\right)$ is the $q$-character of a representation of $\hat{\mathcal{U}}_i$.

(ii) A set of monomials $\mathcal{S}$ is said to be $J$-$q$-closed ($J \subset I$), or simply $q$-closed if $J = I$, if $\mathcal{S}$ is $q$-closed in the direction $i$ for all $i \in J$.

(iii) A set of monomials $\mathcal{S} \subset \mathcal{M}$ is said to be $J$-closed by the Kashiwara operators ($J \subset I$), or simply closed by the Kashiwara operators if $J = I$, if the operators $\tilde{e}_i$, $\tilde{f}_i$ preserve $\mathcal{S}$ for all $i \in J$.

(iv) A set of monomials $\mathcal{S} \subset \mathcal{M}$ which is $J$-$q$-closed and $J$-closed by the Kashiwara operators ($J \subset I$), is called a $J$-closed set. If $J = I$, it is simply called a closed monomial set.

**Remark 3.7.** (i) The definition of a $q$-closed set is inspired by the theory of $q$-characters and the algorithm of [Frenkel and Mukhin 2001]. In particular, it involves $q$-characters of $\mathcal{U}_q(\hat{\mathrm{sl}}_2)'$-modules. Let us recall that in this case, the image of $\chi_q : \mathcal{R}_{l,\mathbb{Z}} \to \mathbb{Z}[Y_{1,l}^{\pm 1}]_{l \in \mathbb{Z}}$ is known (see [Frenkel and Reshetikhin 1999]): it is equal to the subring $\mathbb{Z}[(Y_{1,l} + Y_{1,l+2}^{-1})]_{l \in \mathbb{Z}}$ of $\mathbb{Z}[Y_{1,l}^{\pm 1}]_{l \in \mathbb{Z}}$ generated by the $Y_{1,l} + Y_{1,l+2}^{-1}$ ($l \in \mathbb{Z}$).

(ii) The notion of a $q$-closed set holds also for the monomial $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal $\mathcal{M}_0$. Further it extends naturally when $q$ is specialized at roots of unity, by using the theory of $q$-characters at roots of unity [Frenkel and Mukhin 2002].

Let $V$ be an integrable $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module such that for all $\ell$-weight $(\nu, \gamma)$ of $V$, $V_{(\nu,\gamma)}$ is finite-dimensional and the roots of the associated polynomials $Q_i(z)$ and $R_i(z)$ are in $q^{\mathbb{Z}}$ for all $i \in I$. Then the monomial set $\mathcal{M}(V)$ is $q$-closed. Note that the Frenkel–Mukhin algorithm need not necessarily hold for $V$: for example, it does not work for the simple finite-dimensional $\mathcal{U}_q(\hat{\mathrm{sl}}_3)'$-module $V_0(Y_{1,0}^2 Y_{2,3}) \simeq V_0(Y_{1,0}Y_{2,3}) \otimes V_0(Y_{1,0})$ considered in [Hernandez and Leclerc 2010], but $\mathcal{M}(V_0(Y_{1,0}^2 Y_{2,3}))$ is $q$-closed.

In general, $\mathcal{M}(V)$ is not closed by the Kashiwara operators: for example, the $q$-character of the $\mathcal{U}_q(\hat{\mathrm{sl}}_2)'$-module $V_0(Y_{1,2}Y_{1,0}^2)$ contains the monomial $Y_{1,0}$ but does not contain $Y_{1,2}^{-1}$. However, it holds for the fundamental $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-modules. In fact by using the tableaux sum expressions of their $q$-characters given in [Nakajima 2003], we have:

**Proposition 3.8** [Nakajima 2003]. *Let $V_0(Y_{i,k})$ be a fundamental representation of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ ($i \in I_0$, $k \in \mathbb{Z}$). Then the monomial sets $\mathcal{M}_0(Y_{i,k})$ and $\mathcal{M}(V_0(Y_{i,k}))$ are equal.*

In particular by Proposition 3.5, $\mathcal{M}(V_0(Y_{i,k}))$ has a $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal structure isomorphic to $\mathcal{B}_0(\Lambda_i)$. As a consequence:

**Corollary 3.9.** *For all $1 \le i \le n$ and $k \in \mathbb{Z}$, the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal $\mathcal{M}_0(Y_{i,k})$ is closed.*

Finally, let us give an example of a monomial crystal which is not $q$-closed. Consider the $\mathcal{U}_q(\mathrm{sl}_2)$-crystal $\mathcal{M}_0(Y_{1,4}Y_{1,0})$:

$$Y_{1,4}Y_{1,0} \to Y_{1,6}^{-1}Y_{1,0} \to Y_{1,6}^{-1}Y_{1,2}^{-1}.$$

If $\mathcal{M}_0(Y_{1,4}Y_{1,0})$ is $q$-closed, it should contain $\mathcal{M}(V_0(Y_{1,4}Y_{1,0}))$. This is not the case, the $q$-character of $V_0(Y_{1,4}Y_{1,0})$ being

$$\chi_q(V_0(Y_{1,4}Y_{1,0})) = Y_{1,4}Y_{1,0} + Y_{1,6}^{-1}Y_{1,0} + Y_{1,4}Y_{1,2}^{-1} + Y_{1,6}^{-1}Y_{1,2}^{-1}.$$

**3B.** ***Description of the monomial crystal*** $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$**.** Assume that $n = 2r+1$ is odd with $r \geq 1$. The monomial crystals $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ are studied in [Hernandez and Nakajima 2006, Section 4]: the monomials occurring in these crystals are explicitly described and the automorphisms $z_\ell$ are given in terms of monomials. We recall these results here.

To describe the monomial crystals $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$, we assume in this section that $\ell \leq r+1$ (as in [Hernandez and Nakajima 2006]). Let us begin by explaining why we can do that. We need the notion of twisted isomorphism of crystals (this definition appears in [Bandlow et al. 2010]).

**Definition 3.10.** Let $\mathcal{B}$ and $\mathcal{B}'$ be crystals over two isomorphic Dynkin diagrams $D$ and $D'$ with vertices respectively indexed by $I$ and $I'$ and let $\theta : I \to I'$ be an isomorphism from $D$ to $D'$. Then $\phi$ is a $\theta$-twisted isomorphism if $\phi : \mathcal{B} \to \mathcal{B}'$ is a bijection map and for all $b \in \mathcal{B}$ and $i \in I$,

$$\tilde{f}_{\theta(i)} \cdot \phi(b) = \phi(\tilde{f}_i \cdot b) \quad \text{and} \quad \tilde{e}_{\theta(i)} \cdot \phi(b) = \phi(\tilde{e}_i \cdot b).$$

Let $\iota$ be the automorphism of the Dynkin diagram of type $A_n^{(1)}$ such that $\iota(k) = -k$ ($k \in I$), where $I$ is identified to the set $\mathbb{Z}/(n+1)\mathbb{Z}$. It defines an automorphism $\iota_{\mathfrak{h}}$ of $\mathfrak{h}$ by sending $h_i, d$ to $h_{\iota(i)}, d$ for all $i \in I$. Let $\psi : \mathcal{M} \to \mathcal{M}$ be the map defined by ($r \in \mathbb{Q}$)

$$\psi\Big(e^{r\delta}\prod(e^{\Lambda_i}Y_{i,n})^{u_{i,n}}\Big) = e^{r\delta}\prod(e^{\Lambda_{-i}}Y_{-i,n})^{u_{i,n}}.$$

Then we show easily that $\psi$ is an $\iota$-twisted automorphism of the $P$-crystal $\mathcal{M}$. Furthermore it induces an $\iota$-twisted isomorphism

$$\psi : \mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1}) \longrightarrow \mathcal{M}(e^{\varpi_{n+1-\ell}}Y_{n+1-\ell,0}Y_{0,\ell}^{-1})$$

between the monomial crystals $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ and $\mathcal{M}(e^{\varpi_{n+1-\ell}}Y_{n+1-\ell,0}Y_{0,\ell}^{-1})$ for all $1 \leq \ell \leq r+1$.

So one can assume that $1 \leq \ell \leq r+1$. In this case, $d_\ell = \ell$ and we study the crystal $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ (see [Hernandez and Nakajima 2006]). One defines the monomials

$$\boxed{k}_p = Y_{k-1,p+k}^{-1}Y_{k,p+k-1} \quad \text{for } 1 \leq k \leq n+1,\ p \in \mathbb{Z},$$

with $Y_{n+1,p} = Y_{0,p}$ by convention. By Remark 2.23, the terms $e^{\omega(m')}$ can be safely omitted for all $m' \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$. Set $M_0 = e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}$ and

$$\begin{aligned} M_j &= Y_{\ell,2j} Y_{0,n-\ell+1+2j}^{-1} Y_{j,\ell+j}^{-1} Y_{j,n-\ell+1+j} \\ &= \big(\boxed{1}_{\,n-\ell+2j} \boxed{2}_{\,n-\ell+2j-2} \cdots \boxed{j}_{\,n-\ell+2}\big) \times \big(\boxed{j+1}_{\,\ell-1} \boxed{j+2}_{\,\ell-3} \cdots \boxed{\ell}_{\,1-\ell+2j}\big) \\ &= \prod_{p=1}^{j} \boxed{p}_{\,n-\ell-2p+2j+2} \times \prod_{p=j+1}^{\ell} \boxed{p}_{\,\ell+1-2p+2j} \end{aligned}$$

with $0 \le j \le \ell$. In particular, $M_\ell = Y_{\ell,n+1} Y_{0,n+1+\ell}^{-1} = \tau_{n+1,-\ell\delta}(M_0)$ and $M_1 = \tau_{2,-\delta}(M_0)$ for $\ell = r+1$. One defines other monomials as follows: for $j \in \mathbb{Z}$ and a Young tableau of shape $(\ell)$ $T = (1 \le i_1 < i_2 < \cdots < i_\ell \le n+1)$ we set

$$m_{T;j} = \prod_{p=1}^{j} \boxed{i_p}_{\,n-\ell-2p+2j+2} \times \prod_{p=j+1}^{\ell} \boxed{i_p}_{\,\ell+1-2p+2j} \quad \text{for } 0 \le j \le \ell-1, \tag{3}$$

and $m_{T;j+\ell} = \tau_{n+1,-\ell\delta}(m_{T;j})$. Note that $M_j = m_{T;j}$ with $T = (1, 2, \ldots, \ell)$.

By Theorem 3.4, $\mathcal{M}(M_0)$ and $\mathcal{B}(\varpi_\ell)$ are isomorphic as $P$-crystals. Furthermore:

**Proposition 3.11** [Hernandez and Nakajima 2006]. (i) *$\mathcal{M}_{I_0}(M_j)$ consists of $m_{T;j}$ for various sequences $T$.*

(ii) *We have the equality of $I_0$-crystals*

$$\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}) = \bigsqcup_{k \in \mathbb{Z}} (\tau_{n+1,-\ell\delta})^k \Big( \bigsqcup_{j=0}^{\ell-1} \mathcal{M}_{I_0}(M_j) \Big). \tag{4}$$

(iii) *The map*

$$\sigma : \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}) \to \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$$

*defined by $\sigma(m_{T;j}) = m_{T;j+1}$ is a $P_{\mathrm{cl}}$-crystal automorphism and equals $z_\ell^{-1}$.*

(iv) *The Kashiwara operators $\tilde{e}_i$, $\tilde{f}_i$ are described in terms of tableaux: for $i \neq 0$ we have $\tilde{e}_i \cdot m_{T;j} = m_{T';j}$ or $0$. Here $T'$ is obtained from $T$ by replacing $i+1$ by $i$. If it is not possible (that is, when we have both $i+1$ and $i$ in $T$ or when $i+1$ does not occur in $T$), then it is zero. Similarly $\tilde{f}_i \cdot m_{T;j} = m_{T'';j}$ or $0$, where $T''$ is given by replacing $i$ by $i+1$. For the action of $\tilde{e}_0$, $\tilde{f}_0$, we have*

$$\tilde{e}_0 \cdot m_{T;j} = \begin{cases} 0 & \text{if } i_1 \neq 1 \text{ or } i_\ell = n+1, \\ m_{(i_2,\ldots,i_\ell,n+1);j-1} & \text{if } i_1 = 1 \text{ and } i_\ell \neq n+1, \end{cases}$$

$$\tilde{f}_0 \cdot m_{T;j} = \begin{cases} 0 & \text{if } i_1 = 1 \text{ or } i_\ell \neq n+1, \\ m_{(1,i_1,\ldots,i_{\ell-1});j+1} & \text{if } i_1 \neq 1 \text{ and } i_\ell = n+1. \end{cases}$$

**Proposition 3.12.** *There is a bijection given by $\Xi^0$ between $\mathcal{M}_{I_0}(M_0)$ and $\mathcal{M}(V)$, where $V = V_0(\Xi^0(M_0))$ is the fundamental representation of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ associated to $Y_{\ell,0}$. In particular the monomial crystal $\mathcal{M}_{I_0}(M_0)$ is $I_0$-closed.*

*Proof.* By the previous description, $\mathcal{M}_{I_0}(M_0)$ consists of the monomials $m_{T;0}$ for various sequences $T$. By applying the map $\Xi^0$, they are sent to the monomials

$$m_T = \prod_{p=1}^{\ell} \boxed{i_p}_{\,\ell+1-2p}$$

with $T = (1 \leq i_1 < \cdots < i_\ell \leq n+1)$ and where we set

$$\boxed{1}_{\,p} = Y_{1,p} \quad \text{and} \quad \boxed{n+1}_{\,p} = Y^{-1}_{n,p+n+1}$$

for all $p \in \mathbb{Z}$. They are exactly the monomials occurring in the tableaux sum expressions of $q$-characters of fundamental modules of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ (see [Nakajima 2003]). So the image of $\mathcal{M}_{I0}(M_0)$ by $\Xi^0$ is equal to $\mathcal{M}(V_0(Y_{\ell,0}))$. Further this set is $I_0$-$q$-closed by definition, and $\mathcal{M}_{I_0}(M_0)$ is also $I_0$-closed ☐

Now let us consider the monomial crystal $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y^{-1}_{0,d_\ell})$ with $1 \leq \ell \leq n$. We determine in the next proposition when $z_\ell$ has the particular form of a shift.

**Proposition 3.13.** *The automorphism $z_\ell$ of $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y^{-1}_{0,d_\ell})$ has the special form of a shift $\tau_{p,\alpha}$ ($p \in \mathbb{Z}, \alpha \in \mathbb{Z}\delta$) if and only if $\ell = 1, n$ or $\ell = r+1$. Moreover, we have $z_1 = z_n = \tau_{-n-1,\delta}$ and $z_{r+1} = \tau_{-2,\delta}$.*

*Proof.* Assume that $\ell \leq r+1$. We have seen that $z_\ell = \sigma^{-1}$. So it suffices to determine when $\sigma$ is a shift. We have the equality $\sigma^\ell = \tau_{n+1,-\ell\delta}$. Hence if $\ell = 1$, $\sigma = \tau_{n+1,-\delta}$ is a shift. Assume that $\ell = r+1$. In this case, $M_1 = \tau_{2,-\delta}(M_0) = \sigma(M_0)$. As the crystal $\mathcal{M}(M_0)$ is connected and $\sigma$ and $\tau_{2,-\delta}$ are automorphisms of crystals, we have $\sigma = \tau_{2,-\delta}$. For the other cases, $\sigma$ is explicitly known and is not a shift. As $\psi$ and shift automorphisms commute, the result follows for $\ell > r+1$. ☐

**3C.** ***Affinized promotion operators and monomial crystals $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y^{-1}_{0,d_\ell})$.*** In this section, we introduce promotion operators for the level 0 extremal fundamental weight crystals. We describe them in the monomial realizations of $\mathcal{B}(\varpi_\ell)$ $(1 \leq \ell \leq n)$.

Let us begin by some definitions and properties about the promotion operators (see [Bandlow et al. 2010; Fourier et al. 2009; Schilling 2008; Shimozono 2002] and references therein for more details). In type $A_n$, the highest weight crystal $\mathcal{B}_0(\lambda)$ of highest weight $\lambda \in P_0^+$ can be realized by the semistandard Young tableaux of shape $(\lambda)$. The weight function wt is defined by the content of tableaux, that is, $\mathrm{wt}(T) := (w_1(T), \ldots, w_{n+1}(T))$, where $w_i(T)$ is the number of letters $i$ occurring in the tableau $T$. It can be viewed as an element of $P_0$ in the following way: set $\epsilon_i = \Lambda_i - \Lambda_{i-1}$ for $2 \leq i \leq n$, $\epsilon_1 = \Lambda_1$ and $\epsilon_{n+1} = -\epsilon_1 - \cdots - \epsilon_n$. In particular,

$\alpha_i = \epsilon_i - \epsilon_{i+1}$, $\Lambda_i = \epsilon_1 + \cdots + \epsilon_i$ $(1 \leq i \leq n)$, and we have $P = \mathbb{Z}\epsilon_1 + \cdots + \mathbb{Z}\epsilon_{n+1}$. Then $\mathrm{wt}(T)$ corresponds to the element

$$w_1(T)\epsilon_1 + \cdots + w_{n+1}(T)\epsilon_{n+1} \in P_0$$

for all Young tableau $T$.

**Definition 3.14.** Let $\mathcal{B}_0 = \mathcal{B}_0(\lambda)$ be a highest weight $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal of highest weight $\lambda \in P_0^+$. A promotion operator pr on $\mathcal{B}_0$ is an operator $\mathrm{pr} : \mathcal{B}_0 \to \mathcal{B}_0$ such that

(i) pr shifts the content: if $\mathrm{wt}(T) = (w_1, \ldots, w_{n+1})$ is the content of $T \in \mathcal{B}_0$, then $\mathrm{wt}(\mathrm{pr}(T)) = (w_{n+1}, w_1, \ldots, w_n)$;

(ii) the promotion operator has order $n+1$ : $\mathrm{pr}^{n+1} = \mathrm{id}$;

(iii) $\mathrm{pr} \circ \tilde{e}_i = \tilde{e}_{i+1} \circ \mathrm{pr}$ and $\mathrm{pr} \circ \tilde{f}_i = \tilde{f}_{i+1} \circ \mathrm{pr}$ for $i \in \{1, 2, \ldots, n-1\}$.

Given a promotion operator pr on a highest weight $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal $\mathcal{B}_0(\lambda)$ $(\lambda \in P_0^+)$, one defines an associated affine $P_{\mathrm{cl}}$-crystal by setting

$$\tilde{e}_0 := \mathrm{pr}^{-1} \circ \tilde{e}_1 \circ \mathrm{pr} \quad \text{and} \quad \tilde{f}_0 := \mathrm{pr}^{-1} \circ \tilde{f}_1 \circ \mathrm{pr}.$$

We denote the $P_{\mathrm{cl}}$-crystal hence obtained by $\mathcal{B}_0(\lambda)'_{\mathrm{aff}}$.

It was shown in [Shimozono 2002] that the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-crystal $\mathcal{B}_0(\lambda)$ $(\lambda \in P_0)$ has a unique promotion operator pr when $\lambda$ is rectangular (that is, of the form $k\Lambda_\ell$ with $\ell \in I_0$ and $k \in \mathbb{N}^*$), given by the Schützenberger jeu-de-taquin process. Furthermore the affine $P_{\mathrm{cl}}$-crystal $\mathcal{B}_0(k\Lambda_\ell)'_{\mathrm{aff}}$ obtained by using the promotion operator pr is isomorphic to the crystal basis of a Kirillov–Reshetikhin module associated to $\ell \in I_0, k \in \mathbb{N}^*$ (for a special choice of the spectral parameter $a \in \mathbb{C}^*$ see [Kang et al. 1992]).

From the affine $P_{\mathrm{cl}}$-crystal $\mathcal{B}_0(k\Lambda_\ell)'_{\mathrm{aff}}$, let us consider its affinization $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ (see also [Kashiwara 2002b]): this is the $P$-crystal with vertices in $\{z^s T \mid s \in \mathbb{Z}, T \in \mathcal{B}_0(k\Lambda_\ell)'_{\mathrm{aff}}\}$ such that for all $s \in \mathbb{Z}$ and $T \in \mathcal{B}_0(k\Lambda_\ell)'_{\mathrm{aff}}$,

$$\mathrm{wt}(z^s T) = \mathrm{wt}(T) + s\delta, \quad \tilde{e}_i \cdot z^s T = z^{s+\delta_{i,0}}(\tilde{e}_i \cdot T), \quad \tilde{f}_i \cdot z^s T = z^{s-\delta_{i,0}}(\tilde{f}_i \cdot T).$$

Assume in the following that $\ell \leq r+1$ (the case $\ell > r+1$ is studied at the end of this section). We introduce the affinized promotion operator on $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$.

**Definition 3.15.** Let us consider the crystal of finite type $\mathcal{B}_0(k\Lambda_\ell)$ (with $k \in \mathbb{N}$ and $\ell \leq r+1$), pr its associated promotion operator and $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ its affinization. The affinized promotion operator on $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ is the operator $\mathrm{pr}_{\mathrm{aff}} : \mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}} \to \mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ such that for all $T \in \mathcal{B}_0(k\Lambda_\ell)'_{\mathrm{aff}}$ and $s \in \mathbb{Z}$,

$$\mathrm{pr}_{\mathrm{aff}}(z^s T) = z^{s - w_{n+1}(T)}\mathrm{pr}(T).$$

One checks easily the following statements.

**Lemma 3.16.** *The affinized promotion operator* $\mathrm{pr}_{\mathrm{aff}}$ *of* $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ *shifts the content. It satisfies*

$$\mathrm{pr}_{\mathrm{aff}} \circ \tilde{e}_i = \tilde{e}_{i+1} \circ \mathrm{pr}_{\mathrm{aff}} \quad \textit{and} \quad \mathrm{pr}_{\mathrm{aff}} \circ \tilde{f}_i = \tilde{f}_{i+1} \circ \mathrm{pr}_{\mathrm{aff}}$$

*for* $i \in \{0, 1, \ldots, n\}$ (*where* $\tilde{e}_{n+1}, \tilde{f}_{n+1}$ *are understood to be* $\tilde{e}_0, \tilde{f}_0$ *respectively*). *It has infinite order*, *the weight of* $\mathrm{pr}_{\mathrm{aff}}^{n+1}$ *being* $-k\ell\delta$.

Recall that one has defined an automorphism $\theta$ of the Dynkin diagram of type $A_n^{(1)}$ corresponding to a rotation such that $\theta(i) = i+1$ ($i \in I$). Then by the above Lemma, $\mathrm{pr}_{\mathrm{aff}}$ is a $\theta$-twisted automorphism of $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$. Furthermore as the $P$-crystals $\mathcal{B}(\varpi_\ell)$ and $\mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}}$ are isomorphic (see [Kashiwara 2002b]), the affinized promotion operator $\mathrm{pr}_{\mathrm{aff}} : \mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}} \to \mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}}$ induces a $\theta$-twisted automorphism of the level 0 fundamental extremal weight crystal $\mathcal{B}(\varpi_\ell)$ ($\ell \le r+1$). We call it the promotion operator of $\mathcal{B}(\varpi_\ell)$, also denoted by $\mathrm{pr}_{\mathrm{aff}}$.

We want to describe the promotion operators of the crystals $\mathcal{B}(\varpi_\ell)$ in the monomial realizations when $\ell \le r+1$. To do that, let $\phi : \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) \to \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ be the map such that

$$\phi\Big(\prod Y_{i,l}^{u_{i,l}}\Big) = \prod Y_{i+1,l+1}^{u_{i,l}},$$

the terms $e^\nu$ being safely omitted in the definition by Remark 2.23. Denote by

$$\varphi : \mathcal{B}(\varpi_\ell) \simeq \mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}} \to \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$$

the isomorphism of $P$-crystals between $\mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}}$ and $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$. It is explicitly given by

$$\varphi : z^s T \in \mathcal{B}_0(\Lambda_\ell)_{\mathrm{aff}} \mapsto m_{T;-s} \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) \quad (s \in \mathbb{Z}, T \in \mathcal{B}_0(\Lambda_\ell)).$$

The following result relates the map $\phi : \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) \to \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ to the promotion operator $\mathrm{pr}_{\mathrm{aff}}$ of $\mathcal{B}(\varpi_\ell)$ introduced above.

**Proposition 3.17.** *Assume that* $\ell \le r+1$. *The following diagram commutes*:

$$\begin{tikzcd} \mathcal{B}(\varpi_\ell) \arrow[r, "\varphi"] \arrow[d, "\mathrm{pr}_{\mathrm{aff}}"'] & \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) \arrow[d, "\phi"] \\ \mathcal{B}(\varpi_\ell) \arrow[r, "\varphi"] & \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}). \end{tikzcd}$$

*Proof.* For $1 \le k \le n+1$ and $p \in \mathbb{Z}$, we have

$$\begin{aligned} \phi\Big(\boxed{k}_{\,p}\Big) &= \phi(Y_{k-1,p+k}^{-1} Y_{k,p+k-1}) = Y_{k,p+k+1}^{-1} Y_{k+1,p+k} \\ &= \begin{cases} \boxed{k+1}_{\,p} & \text{if } k \le n, \\ \boxed{1}_{\,p+n+1} & \text{if } k = n+1. \end{cases} \end{aligned}$$

Fix $j \in \mathbb{Z}$ and a Young tableau

$$T = (1 \leq i_1 < i_2 < \cdots < i_\ell \leq n+1)$$

of shape $(\Lambda_\ell)$. If $i_\ell \neq n+1$, we have

$$\begin{aligned}\phi(m_{T;j}) &= \phi\Bigg(\prod_{p=1}^{j} \boxed{i_p}_{\,n-\ell-2p+2j+2} \times \prod_{p=j+1}^{\ell} \boxed{i_p}_{\,\ell+1-2p+2j}\Bigg)\\ &= \prod_{p=1}^{j} \boxed{i_p+1}_{\,n-\ell-2p+2j+2} \times \prod_{p=j+1}^{\ell} \boxed{i_p+1}_{\,\ell+1-2p+2j}\\ &= m_{\mathrm{pr}(T);j} = \varphi(\mathrm{pr}_{\mathrm{aff}}(z^{-j}T)).\end{aligned}$$

Assume that $i_\ell = n+1$. Then

$$\begin{aligned}&\phi(m_{T;j})\\ &= \prod_{p=1}^{j} \boxed{i_p+1}_{\,n-\ell-2p+2j+2} \times \prod_{p=j+1}^{\ell-1} \boxed{i_p+1}_{\,\ell+1-2p+2j} \times \boxed{1}_{\,n-\ell+2j+2}\\ &= \prod_{p=2}^{j+1} \boxed{i_{p-1}+1}_{\,n-\ell-2p+2(j+1)+2} \times \prod_{p=j+2}^{\ell} \boxed{i_{p-1}+1}_{\,\ell+1-2p+2(j+1)} \times \boxed{1}_{\,n-\ell+2(j+1)}\\ &= m_{\mathrm{pr}(T);j+1} = \phi(z^{-j-1}\mathrm{pr}(T)) = \varphi(\mathrm{pr}_{\mathrm{aff}}(z^{-j}T)).\end{aligned}$$

□

**Remark 3.18.** It follows in particular that $\phi$ is a $\theta$-twisted automorphism of $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$, since $\mathcal{B}(\varpi_\ell)$ and $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ are connected and $\mathrm{pr}_{\mathrm{aff}}$ is a $\theta$-twisted automorphism.

The case $\ell \geq r+1$ is similar to the previous one. The affinized promotion operator of $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ ($k \in \mathbb{N}^*$) is the operator

$$\tilde{\mathrm{pr}}_{\mathrm{aff}} : \mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}} \longrightarrow \mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$$

such that for all $T \in \mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$ and $s \in \mathbb{Z}$,

$$\tilde{\mathrm{pr}}_{\mathrm{aff}}(z^s T) = z^{s+k-w_{n+1}(T)}\mathrm{pr}(T).$$

Note that the definition of the affinized promotion operator is different to the one when $\ell \leq r+1$. This provides to the automorphism $\iota_{\mathfrak{h}}$ of $\mathfrak{h}$.

Let us consider the map

$$\psi \circ \phi \circ \psi^{-1} : \mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,n+1-\ell}^{-1}) \to \mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,n+1-\ell}^{-1}).$$

It is a $\theta^{-1}$-twisted automorphism of $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,n+1-\ell}^{-1})$ such that

$$\psi \circ \phi \circ \psi^{-1}\Big(\prod Y_{i,l}^{u_{i,l}}\Big) = \prod Y_{i-1,l+1}^{u_{i,l}}.$$

It can be related to the promotion operator $\tilde{\mathrm{pr}}_{\mathrm{aff}}$ of $\mathcal{B}_0(k\Lambda_\ell)_{\mathrm{aff}}$: one can check that $\psi \circ \phi \circ \psi^{-1} = \tilde{\mathrm{pr}}_{\mathrm{aff}}^{-1}$.

**3D.** ***Application of promotion operators to the study of*** $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$**.** In this section, we use promotion operators to obtain a new description of the monomial crystal $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$, improving results given in [Hernandez and Nakajima 2006]. Moreover, we determine the $\ell$ for which the crystals $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ are closed.

Assume first that $\ell \le r+1$ (where $n = 2r+1$ is still supposed to be odd). Let us begin with the following remarks. The monomials $\phi^j(Y_{\ell,0} Y_{0,\ell}^{-1}) = Y_{\ell+j,j} Y_{j,\ell+j}^{-1}$ will have a particular importance in the construction of extremal fundamental loop weight modules. One can give them in terms of Young tableaux, thanks to the $\theta$-twisted automorphism $\phi$ of $\mathcal{M}$:

- If $j$ is such that $\ell + j \le n+1$, $Y_{\ell+j,j} Y_{j,\ell+j}^{-1} \in \mathcal{M}_{I_0}(M_0)$ and is equal to $m_{T;0}$ with $T = (j+1, j+2, \ldots, j+\ell)$.
- If $1 \le j \le \ell - 1$, then $Y_{j,n-\ell+j+1} Y_{n-\ell+j+1,n+j+1}^{-1} \in \mathcal{M}_{I_0}(M_j)$ and is equal to $m_{T;j}$ with $T = (1, 2, \ldots, j, n-\ell+j+2, \ldots, n+1)$.

We will have to consider the finite sub-$I_j$-crystals of $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$,

$$\mathcal{M}_{I_j}(Y_{\ell+j,j+k(n+1)} Y_{j,\ell+j+k(n+1)}^{-1})$$

for $j \in I$ and $k \in \mathbb{Z}$: this is the sub-$I_j$-crystal of $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ generated by the monomial $Y_{\ell+j,j+k(n+1)} Y_{j,\ell+j+k(n+1)}^{-1}$. Note that one of these crystals can be obtained from another one by application of powers of $\phi$.

**Proposition 3.19.** *Let $\ell \le r+1$. We have the equality of sets*

$$\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \bigcup_{k \in \mathbb{Z}} (\tau_{n+1,-\ell\delta})^k \left( \bigcup_{j=0}^{n} \mathcal{M}_{I_j}(Y_{\ell+j,j} Y_{j,\ell+j}^{-1}) \right).$$

*Proof.* As $Y_{\ell+j,j} Y_{j,\ell+j}^{-1} \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ for all $0 \le j \le n$ and $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is connected,

$$\bigcup_{j=0}^{n} \mathcal{M}_{I_j}(Y_{\ell+j,j} Y_{j,\ell+j}^{-1}) \subset \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$$

as sets.

Let us fix $m \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$. The monomial $m$ is of the form $m_{T;j}$ with $T = (1 \le i_1 < i_2 < \cdots < i_\ell \le n+1)$ and $j \in \mathbb{Z}$. By application of the shift automorphism, one can assume that $0 \leq j \leq \ell-1$. So we have to show that $m_{T;j} \in \bigcup_{j=0}^{n} \mathcal{M}_{I_j}(Y_{\ell+j,j}Y_{j,\ell+j}^{-1})$. If $j=0$, we have $m_{T;0} \in \mathcal{M}_{I_0}(Y_{\ell,0}Y_{0,\ell}^{-1})$. Assume that $1 \leq j \leq \ell-1$ and set $s = i_{j+1}-1$. Then $T = (i_1 < \cdots < i_j < s+1 < i_{j+2} < \cdots < i_\ell)$ and by application of the Kashiwara operators $\tilde{e}_1, \ldots, \tilde{e}_{s-1}, \tilde{e}_{s+2}, \ldots, \tilde{e}_n$ on $m_{T;j}$, we show that

$$m_{T;j} \in \mathcal{M}_{I_s}(m_{T';j}) \quad \text{with } T' = (1 < \cdots < j < s+1 < \cdots < s+\ell-j).$$

By applying $\tilde{e}_1, \ldots, \tilde{e}_{j-1}, \tilde{e}_{s+\ell-j+1}, \ldots, \tilde{e}_n$ and $\tilde{e}_0$ on $m_{T';j}$, it is sent on

$$m_{T'';0} \quad \text{with } T'' = (s+1 < \cdots < s+\ell) \quad \text{if } s+\ell \leq n+1,$$

and on

$$m_{T'';u} \text{ with } u = s+\ell-n-1, T'' = (1 < \cdots < u < s+1 < \cdots < n+1) \text{ otherwise.}$$

Furthermore $m_{T'';u} = \phi^s(Y_{\ell,0}Y_{0,\ell}^{-1}) = Y_{\ell+s,s}Y_{s,\ell+s}^{-1}$ by the above remark and $m_{T;j}$ is also contained in $\mathcal{M}_{I_s}(Y_{\ell+s,s}Y_{s,\ell+s}^{-1})$. □

**Remark 3.20.** One of the questions treated in [Hernandez and Nakajima 2006, Section 4] is to give an explicit description of monomials occurring in $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ ($\ell \leq r+1$). Actually by the shift automorphism and the description given in that reference, all the monomials in $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ can be obtained from the monomials occurring in $\bigsqcup_{j=0}^{\ell-1}\mathcal{M}_{I_0}(M_j)$ (see (4)). So this description requires knowing

$$\left[\ell\binom{n+1}{\ell}\right]$$

monomials to obtain all the other ones. The preceding proposition improves this result. In fact to determine all the vertices of $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$, it suffices to know the monomials occurring in the $I_{\{0,1\}}$-crystal $\mathcal{M}_{I_{\{0,1\}}}(Y_{\ell,0}Y_{0,\ell}^{-1})$ and to apply $\phi$. Further a monomial $m_{T';0} \in \mathcal{M}_{I_{\{0,1\}}}(Y_{\ell,0}Y_{0,\ell}^{-1})$ is such that $T'$ has the form $T' = (1 < i_2 < \cdots < i_\ell)$. So by Proposition 3.19, only $\binom{n}{\ell-1}$ monomials are sufficient to determine all the vertices of $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$.

The following lemma will be useful.

**Lemma 3.21.** *Assume that $\ell = 1$ or $\ell = r+1$ and set $p = n+1$ or $p = 2$ respectively. We have the equality of $I_j$-crystals* ($0 \leq j \leq n$)

$$\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1}) = \bigsqcup_{k\in\mathbb{Z}}(\tau_{p,-\delta})^k\big(\mathcal{M}_{I_j}(Y_{\ell+j,j}Y_{j,\ell+j}^{-1})\big). \tag{5}$$

*Proof.* By Proposition 3.13, the automorphism $z_\ell$ has the special form of a shift in the considered cases. Further we know that $(\tau_{-p,\delta})^\ell = \tau_{-n-1,\ell\delta}$. Using (4), we obtain

$$\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})=\bigsqcup_{k\in\mathbb{Z}}(\tau_{p,-\delta})^{k\ell}\left(\bigsqcup_{j=0}^{\ell-1}(\tau_{p,-\delta})^{j}(\mathcal{M}_{I_0}(M_0))\right).$$

As $\tau_{p,-\delta}$ and $\phi$ commute, (5) follows. $\square$

Similar equalities of crystals can be given for $\ell\geq r+1$, by using the automorphism $\psi$. Now we are able to determine the $\ell\in I_0$ for which the monomial crystal $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ is closed.

**Theorem 3.22.** *The monomial crystal* $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ *is closed if and only if* $\ell=1,r+1$ *or* $n$.

*Proof.* Let us begin by the case $\ell\leq r+1$. Assume that the crystal $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ is $q$-closed for $2\leq\ell\leq r$. Consider the monomial

$$M_j=Y_{\ell,2j}Y_{0,n-\ell+1+2j}^{-1}Y_{j,\ell+j}^{-1}Y_{j,n-\ell+1+j}\in\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$$

with $j\neq 0$. We have $\Xi_j(M_j)=Y_{j,\ell+j}^{-1}Y_{j,n-\ell+1+j}$. By Definition 3.6, there exists a subset

$$\mathcal{S}_{M_j}\subset\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})\cap\left(M_j\cdot\prod_{l\in\mathbb{Z}}A_{j,l}^{\mathbb{Z}}\right)$$

containing $M_j$ such that its image $\Xi_j(\mathcal{S}_{M_j})$ by $\Xi_j$ is the set of $\ell$-weights of a representation of $\hat{\mathcal{U}}_j$. By the theory of $q$-characters of $\hat{\mathcal{U}}_j$-modules, we should have

$$Y_{j,\ell+j}^{-1}Y_{j,n-\ell+1+j}\,\Xi_j(A_{j,\ell+j-1})\in\Xi_j(\mathcal{S}_{M_j}).$$

Furthermore, the map $\Xi_j$ is injective when it is restricted on the set of monomials $M_j\cdot\big(\prod_{l\in\mathbb{Z}}A_{j,l}^{\mathbb{Z}}\big)$: indeed, we have for all monomial $M_j\cdot\big(\prod_{l\in\mathbb{Z}}A_{j,l}^{u_{j,l}}\big)$,

$$\Xi_j\left(M_j\cdot\prod_{l\in\mathbb{Z}}A_{j,l}^{u_{j,l}}\right)=\Xi_j(M_j)\cdot\prod_{l\in\mathbb{Z}}\Xi_j(A_{j,l}^{u_{j,l}})=\Xi_j(M_j)\cdot\prod_{l\in\mathbb{Z}}\tau_{\{j\}}(A_{j,l}^{u_{j,l}}),$$

where $\tau_{\{j\}}$ is the map defined in Definition 3.2 of [Frenkel and Mukhin 2001] (the second equality is a consequence of Lemma 3.5 of the same reference). The injectivity is a consequence of the injectivity of $\tau_{\{j\}}$ (see [Frenkel and Mukhin 2001, Lemma 3.3]).

So the monomial

$$\begin{aligned}m&=M_j\cdot A_{j,\ell+j-1}\\&=Y_{\ell,2j}Y_{0,n-\ell+1+2j}^{-1}Y_{j,\ell+j-2}Y_{j-1,\ell+j-1}^{-1}Y_{j+1,\ell+j-1}^{-1}Y_{j,n-\ell+1+j}\end{aligned}$$

should occur in $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$. But this is not the case, $m$ being not of the form (3). Hence $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ is not $q$-closed when $2\leq\ell\leq r$.

Now assume that $\ell = 1$ or $\ell = r+1$. In this case, $z_\ell = \tau_{-p,\delta}$ with $p = n+1$ or $p = 2$ respectively and by the above lemma

$$\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \bigsqcup_{k\in\mathbb{Z}} (\tau_{p,-\delta})^k \big(\mathcal{M}_{I_j}(Y_{\ell+j,j} Y_{j,\ell+j}^{-1})\big)$$

as $I_j$-crystals ($0 \le j \le n$). By Proposition 3.12, the finite crystal $\mathcal{M}_{I_0}(M_0)$ is $I_0$-closed. As the $I_j$-crystals $\mathcal{M}_{I_j}(Y_{\ell+j,j} Y_{j,\ell+j}^{-1})$ can be obtained from $\mathcal{M}_{I_0}(M_0)$ by application of powers of $\phi$, they are also $I_j$-closed for all $0 \le j \le n$. Then by the above equalities $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is closed.

Finally, the result follows for all the $\ell \in I_0$ by using the $\iota$-twisted automorphism $\psi$ (which preserves the notion of $q$-closed monomial set). $\square$

## 4. Extremal fundamental loop weight modules for $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ when $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is closed

Assume that $n = 2r+1$ ($r \ge 1$) is odd and $\mathcal{M}_\ell = \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is closed (it holds if and only if $\ell = 1, r+1$ or $n$). In this section, we relate the monomial $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystals $\mathcal{M}_\ell$ with integrable representations of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$.

In Section 4A, we construct a new infinite family of representations $V_\ell$ of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ (Theorem 4.1). We call these representations the extremal fundamental loop weight modules. Let us give the outline of this construction: consider the vector space $V_\ell$ freely generated by the monomials occurring in $\mathcal{M}_\ell$. For all $0 \le j \le n$, we define an action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on it, denoted by $V_\ell^{(j)}$, such that

$$V_\ell^{(j)} = \bigoplus_{k\in\mathbb{Z}} V_k^{(j)},$$

where $V_k^{(j)}$ is a subvector space endowed with a structure of a simple $\ell$-highest weight $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module. We show that it defines a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module structure in this way on $V_\ell$, the compatibility between the action of various vertical subalgebras being a consequence of the existence of promotion operators on $\mathcal{M}_\ell$. Furthermore the $q$-character of $V_\ell$ is the sum of monomials occurring in $\mathcal{M}_\ell$ with multiplicity one.

In Section 4B, we study these representations: we show that $V_\ell$ is irreducible and it is an extremal loop weight module, generated by an extremal vector of $\ell$-weight $e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}$. Furthermore explicit formulas are given for the action of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on $V_\ell$. It is remarkable that these formulas are expressed only from the associated monomial crystal and are "universal" in the following sense: the action on all the extremal fundamental loop weight modules $V_\ell$ is completely determined by these formulas and by the data of the corresponding monomial crystals $\mathcal{M}_\ell$. This sheds new light on the link between monomial crystals and the theory of $q$-characters already expected in [Hernandez and Nakajima 2006]. All these sentences hold for the fundamental $\ell$-highest weight modules $V_0(Y_{\ell,0})$ of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ with the corresponding monomial crystals $\mathcal{M}_0(Y_{\ell,0})$.

In Section 4C, we specialize $q$ at a root of unity $\epsilon$. We obtain new irreducible finite-dimensional representations of the quantum toroidal algebra $\mathcal{U}_\epsilon(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$.

**4A.** ***Construction of the extremal fundamental loop weight modules.*** Let us begin with the main result of this section.

**Theorem 4.1.** *Assume that $n = 2r + 1$ is odd and $\ell = 1, r + 1$ or $n$. There exists a thin representation of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ whose $q$-character is the sum of all monomials occurring in $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ with multiplicity one. It is denoted by $V_\ell = V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ and called the extremal fundamental loop weight module of $\ell$-weight $e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}$.*

To construct these representations, let us start with results about the fundamental modules $V_0(Y_{\ell,k})$ of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ ($n \in \mathbb{N}^*$, $1 \le \ell \le n$, $k \in \mathbb{Z}$). As it is said above, it is isomorphic to the fundamental highest weight $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module $V_0(\Lambda_\ell)$. So we begin by recalling some well-known facts about $V_0(\Lambda_\ell)$, which will be useful.

**Lemma 4.2.** *All the weight spaces of the fundamental highest weight $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module $V_0(\Lambda_\ell)$ ($1 \le \ell \le n$) are of dimension one. Furthermore the Weyl group of finite type $W_0$ acts transitively on* $\mathrm{wt}(V_0(\Lambda_\ell))$.

**Proposition 4.3.** *Let $V_0(Y_{\ell,k})$ be a fundamental module of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ ($\ell \in I_0, k \in \mathbb{Z}$). Then $V_0(Y_{\ell,k})$ is a thin $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module which admits a basis $(v_m)$ indexed by the vertices of the monomial crystal $\mathcal{M}_0(Y_{\ell,k})$, such that for all $m \in \mathcal{M}_0(Y_{\ell,k})$ and $i \in I_0$, $v_m$ is of $\ell$-weight $m$ and*

$$x_{i,0}^+ \cdot v_m = v_{\tilde{e}_i \cdot m}, \quad x_{i,0}^- \cdot v_m = v_{\tilde{f}_i \cdot m},$$

*where $v_0 = 0$ by convention.*

*Proof.* It is known that $\mathrm{Res}(V_0(Y_{\ell,k}))$ is the fundamental highest weight $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module $V_0(\Lambda_\ell)$. By the preceding Lemma, its weight spaces are all of dimension one. In particular its $\ell$-weight spaces are also of dimension one and $V_0(Y_{\ell,k})$ is a thin $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module.

Furthermore $\mathrm{Res}(V_0(Y_{\ell,k}))$ is the extremal weight module of extremal weight $\Lambda_\ell$, generated by an extremal vector $v$ of weight $\Lambda_\ell$. Hence, there exists $\{v_w\}_{w \in W_0}$ such that $v_{\mathrm{Id}} = v$ and

$$x_{i,0}^{\pm} \cdot v_w = 0 \quad \text{and} \quad (x_{i,0}^{\mp})^{(\pm w(\Lambda_\ell)(h_i))} \cdot v_w = v_{s_i(w)} \quad \text{if } \pm w(\Lambda_\ell)(h_i) \ge 0.$$

By the above lemma for all $\nu \in \mathrm{wt}(\mathrm{Res}(V_0(Y_{\ell,k})))$, there exists $w$ such that $\nu = w(\Lambda_\ell)$. Then the corresponding vector $v_w$ is nonzero of weight $\nu$. As all the weight spaces of $V_0(Y_{\ell,k})$ are of dimension one, $\{v_w\}_{w\in W_0}$ generates $V_0(Y_{\ell,k})$ as a vector space. Furthermore for all $w, w' \in W_0$,

$$w(\Lambda_\ell) = w'(\Lambda_\ell) \iff v_w = v_{w'}.$$

In fact, we have (see [Bourbaki 1968, Chapter V.3.3, Proposition 2])

$$\begin{aligned} w(\Lambda_\ell) = w'(\Lambda_\ell) \iff w^{-1}w'(\Lambda_\ell) = \Lambda_\ell \iff w^{-1}w' \in \langle s_i, i \in I_0 - \{\ell\}\rangle \\ \iff w' \in w \cdot \langle s_i, i \in I_0 - \{\ell\}\rangle. \end{aligned}$$

Fix an $\ell$-weight $m \in \mathcal{M}(V_0(Y_{\ell,k})) = \mathcal{M}_0(Y_{\ell,k})$. By what we have said above, one can define $v_m$ as the unique vector $v_w$ ($w \in W_0$) such that $w(\Lambda_\ell) = \mathrm{wt}(m)$. Then $\{v_m \mid m \in \mathcal{M}_0(Y_{\ell,k})\}$ is a basis of $V_0(Y_{\ell,k})$. Furthermore as the weight subspaces and the $\ell$-weight subspaces of $V_0(Y_{\ell,k})$ coincide and are of dimension one, $v_m$ is also an $\ell$-weight vector of $\ell$-weight $m$ for all $m \in \mathcal{M}_0(Y_{\ell,k})$.

We determine the action of $\mathcal{U}_q^h(\hat{\mathrm{sl}}_{n+1})$ on this basis. Fix $m \in \mathcal{M}_{I_0}(Y_{\ell,k})$. For all $i \in I_0$, we have $\mathrm{wt}(m)(h_i) = 0, \pm 1$. Assume that $\mathrm{wt}(m)(h_i) = 0$. Then on the one hand $x_{i,0}^{\pm} \cdot v_m = 0$ by definition of the family $\{v_w\}_{w\in W_0}$. And on the other hand $\tilde{e}_i \cdot m = 0$ and $\tilde{f}_i \cdot m = 0$ by the description of the crystal $\mathcal{M}_0(Y_{\ell,k})$ recalled above. Now assume that $\mathrm{wt}(m)(h_i) = \pm 1$. The vector $S_i(v_m) = x_{i,0}^{\mp} \cdot v_m$ is of the form $v_{m'}$ with $m' \in \mathcal{M}_0(Y_{\ell,k})$ such that

$$\mathrm{wt}(m') = s_i(\mathrm{wt}(m)) = \mathrm{wt}(m) \mp \alpha_i.$$

But the description of $\mathcal{M}_0(Y_{\ell,k})$ shows that the unique monomial of weight $\mathrm{wt}(m) \mp \alpha_i$ is $\tilde{f}_i \cdot m$ (resp. $\tilde{e}_i \cdot m$). Hence $m'$ is equal to $\tilde{f}_i \cdot m$ (resp. $\tilde{e}_i \cdot m$). Finally we have shown that for all $i \in I_0$ and $m \in \mathcal{M}_0(Y_{\ell,k})$,

$$x_{i,0}^{+} \cdot v_m = v_{\tilde{e}_i \cdot m} \quad \text{and} \quad x_{i,0}^{-} \cdot v_m = v_{\tilde{f}_i \cdot m}. \qquad \square$$

In particular, the action of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ on the fundamental modules $V_0(Y_{\ell,k})$ is determined by the combinatorics of monomial crystals $\mathcal{M}_0(Y_{\ell,k})$; in fact, the action of operators $x_{i,r}^{\pm}$ ($1 \le i \le n$, $r \in \mathbb{Z}$) deduces from the action of the $x_{i,0}^{\pm}$ (given by $\mathcal{M}_0(Y_{\ell,k})$) and the action of $h_{i,r}$ (given by the $\ell$-weights $m \in \mathcal{M}_0(Y_{\ell,k})$) from (1).

Let us begin the construction of extremal fundamental loop weight modules. Assume that $n = 2r+1$ is odd and $\ell \le r+1$ (the case $\ell > r+1$ is discussed below at Remark 4.5). Consider the monomial $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystal $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$, supposed to be closed. This is the case if and only if $\ell = 1$ or $\ell = r+1$ (Theorem 3.22). Set $p = n+1$ or $p = 2$ respectively.

Denote by $\mathscr{E}$ (resp. $\mathscr{E}_{j,k}$ for $0 \le j \le n$ and $k \in \mathbb{Z}$) the set of monomials occurring in the crystal $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ (resp. in $(\tau_{p,-\delta})^k \big(\mathcal{M}_{I_j}(Y_{\ell+j,j} Y_{j,\ell+j}^{-1})\big) = \mathcal{M}_{I_j}(Y_{\ell+j,j+kp} Y_{j,\ell+j+kp}^{-1})$). By (5), one has $\mathscr{E} = \bigsqcup_{k\in\mathbb{Z}} \mathscr{E}_{j,k}$ for all $0 \le j \le n$.

Let

$$V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \bigoplus_{m \in \mathcal{E}} \mathbb{C} v_m \tag{6}$$

be the vector space freely generated by elements of $\mathcal{E}$. For all $0 \leq j \leq n$ and $k \in \mathbb{Z}$ set $V_k^{(j)} = \bigoplus_{m \in \mathcal{E}_{j,k}} \mathbb{C} v_m$ the subspace of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ of dimension $\dim(V_0(\Lambda_\ell))$. In particular, we have

$$V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \bigoplus_{k \in \mathbb{Z}} V_k^{(j)}.$$

This decomposition can be compared to the equalities of crystals (5).

We endow the vector space $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ with a structure of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module as follows ($0 \leq j \leq n$): for all $k \in \mathbb{Z}$, let $(v_m)$ be the basis of the $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V_0(Y_{\ell,j+kp})^{(j)}$ defined in Proposition 4.3, indexed by the set of monomials

$$\Xi^j\big(\mathcal{M}_{I_j}(Y_{\ell+j,j+kp} Y_{j,\ell+j+kp}^{-1})\big) = \{\Xi^j(m) \mid m \in \mathcal{M}_{I_j}(Y_{\ell+j,j+kp} Y_{j,\ell+j+kp}^{-1})\}.$$

Let us define an isomorphism of vector spaces between $V_k^{(j)}$ and $V_0(Y_{\ell,j+kp})^{(j)}$ by

$$\begin{aligned} V_k^{(j)} &\longrightarrow V_0(Y_{\ell,j+kp})^{(j)} \\ v_m &\mapsto v_{\Xi^j(m)}. \end{aligned}$$

We endow the vector space $V_k^{(j)}$ with a structure of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module by pulling back the action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on $V_0(Y_{\ell,j+kp})^{(j)}$. By direct sum, $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module, denoted by $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)}$.

**Proposition 4.4.** *There exists a structure of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module on $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ such that the induced $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module is isomorphic to $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)}$ for all $j \in I$. Furthermore the $q$-character of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is*

$$\chi_q\big(V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})\big) = \sum_{m \in \mathcal{E}} m,$$

*where $\mathcal{E}$ is the set of the monomials occurring in $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$.*

*Proof.* To define an action of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$, we determine the action of the subalgebras $\hat{\mathcal{U}}_i$ for all $i \in I$. For that, let $j \in I$ be such that $j \neq i$. The action of $\hat{\mathcal{U}}_i$ on $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is the restriction of the action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)}$. Furthermore we set for all $h \in \mathfrak{h}$ and $m \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$,

$$k_h \cdot v_m = q^{\mathrm{wt}(m)(h)} v_m.$$

The definition of the action of $\hat{\mathcal{U}}_i$ ($i \in I$) is independent of the choice of $j \in I$, $j \neq i$: for $m \in \mathcal{E}$, the action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on the vector $v_m$ is determined by the sub-$I_j$-crystal $\mathcal{M}_{I_j}(m)$ of $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ and by the $\ell$-weight $\Xi^j(m)$. So the action of $\hat{\mathcal{U}}_i$ on $v_m$ is determined by the action of $\tilde{e}_i$ and $\tilde{f}_i$ on $m$ and by the $\ell$-weight $\Xi_i(m)$, which are independent of the choice of $j$.

We show that this action endows $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ with a structure of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module. We fix two indices $i_1, i_2 \in I$ and we check the relations satisfied by $\hat{\mathcal{U}}_{i_1}$ and $\hat{\mathcal{U}}_{i_2}$. The indices $i_1$ and $i_2$ are in the same connected subset $I_j$ of the set of vertices of the Dynkin diagram ($j \in I$). By construction, the action of $\hat{\mathcal{U}}_{i_1}$ and $\hat{\mathcal{U}}_{i_2}$ are restrictions of the action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ on $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$. As $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)}$ is a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module, the relations between $\hat{\mathcal{U}}_{i_1}$ and $\hat{\mathcal{U}}_{i_2}$ are satisfied and $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module.

By construction the induced $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module obtained from $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ by restriction is isomorphic to $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)}$ for all $j \in I$. Furthermore the $\ell$-weight of $v_m$ is $\Xi_i(m)$ for the action of $\hat{\mathcal{U}}_i$ ($i \in I$). So $m$ is the $\ell$-weight of $v_m$ and the $q$-character of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is the sum of monomials occurring in $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ with multiplicity one. ☐

**Remark 4.5.** Let us consider the case $\ell > r+1$. $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,n+1-\ell}^{-1})$ is a monomial crystal closed for $\ell = n$. We show in the same way as above that there exists also a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_n} Y_{n,0} Y_{0,1}^{-1})$ whose $q$-character is the sum of monomials occurring in $\mathcal{M}(e^{\varpi_n} Y_{n,0} Y_{0,1}^{-1})$ with multiplicity one.

Actually this $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module is related to the previous one for $\ell = 1$ as follows. We have defined an automorphism $\iota$ of the Dynkin diagram of type $A_n^{(1)}$. It induces an algebra automorphism of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ we still denote $\iota$, which sends $x_{i,r}^{\pm}, h_{i,m}, k_h$ to $x_{\iota(i),r}^{\pm}, h_{\iota(i),m}, k_{\iota_{\mathfrak{h}}(h)}$ ($i \in I, r \in \mathbb{Z}, m \in \mathbb{Z} - \{0\}, h \in \mathfrak{h}$). Let us denote by $V(e^{\varpi_n} Y_{n,0} Y_{0,1}^{-1})^{\iota}$ the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module obtained from $V(e^{\varpi_n} Y_{n,0} Y_{0,1}^{-1})$ by twisting the action by $\iota$. Then we show easily that $V(e^{\varpi_n} Y_{n,0} Y_{0,1}^{-1})^{\iota}$ and $V(e^{\varpi_1} Y_{1,0} Y_{0,1}^{-1})$ are isomorphic.

**4B.** ***Study of the extremal fundamental loop weight modules.*** In this section, we study the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-modules $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$, where $n = 2r+1$ is supposed to be odd and $\ell = 1, n$ or $\ell = r+1$. We set $p = n+1$ or $p = 2$ respectively.

**Proposition 4.6.** *The $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is integrable. Moreover, it satisfies properties* (iii) *and* (iv) *of Remark 2.5 with weight subspaces of dimension one.*

*Proof.* Assume first that $\ell = 1, r+1$. The $q$-character of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is known: this is the sum of monomials occurring in $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ with multiplicity one. Furthermore one has the equality of $I_0$-crystals

$$\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \bigsqcup_{k \in \mathbb{Z}} (\tau_{p,-\delta})^k \big(\mathcal{M}_{I_0}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})\big).$$

For all $m \in \mathcal{M}_{I_0}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ and $k \in \mathbb{Z}$, $\mathrm{wt}((\tau_{p,-\delta})^k(m)) = \mathrm{wt}(m) - k\delta$. So to prove that the weight spaces of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ are of dimension one, we have to show that the weights of monomials occurring in $\mathcal{M}_{I_0}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ are different to each other. More precisely, it is sufficient to show that the sum

$$\sum_{m \in \mathcal{M}_{I_0}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})} e\big(\mathrm{wt}(\Xi^0(m))\big) \in \bigoplus_{\nu \in P_0} \mathbb{Z} e(\nu)$$

is without multiplicity. This follows from the above results: it is the character of the $\mathcal{U}_q(\mathrm{sl}_{n+1})$-module $V_0(\Lambda_\ell)$.

For all $j \in I$, the representation $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is completely reducible as a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module and we have

$$V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})^{(j)} = \bigoplus_{p \in \mathbb{Z}} V_0(Y_{\ell, j+kp})^{(j)}. \tag{7}$$

As the representations $V_0(Y_{\ell,j+kp})$ are all integrable, it holds for $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$. Furthermore $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ satisfies the stronger property (iv) of Remark 2.5: in fact the representations $V_0(Y_{\ell,j+kp})$ are all isomorphic as $\mathcal{U}_q(\mathrm{sl}_{n+1})$-modules and satisfy property (iv). Hence we have $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})_{\nu+N\alpha_i} = \{0\}$ for all $\nu \in P$, $i \in I$, $N \gg 0$.

Finally, the case $\ell = n$ is deduced from the case $\ell = 1$ by the $\iota$-twisted automorphism $\psi$. □

**Theorem 4.7.** *The $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is an extremal loop weight module generated by the vector $v_{e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}}$ of $\ell$-weight $e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}$.*

*Proof.* We treat the case $\ell = 1, r+1$ (the case $\ell = n$ can be deduced from $\ell = 1$ by using $\psi$). The formulas in (7) imply immediately the third point of Definition 2.26. The first two points are consequences of the following lemmas. □

**Lemma 4.8.** *Let $\mathcal{M}'$ be a sub-$\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-crystal of $\mathcal{M}$. Assume $V$ is a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module with basis $(v_m)_{m \in \mathcal{M}'}$ satisfying*

$$\mathrm{wt}(v_m) = \mathrm{wt}(m), \quad (x_i^+)^{(k)} \cdot v_m = v_{\tilde{e}_i^k \cdot m} \quad \text{and} \quad (x_i^-)^{(k)} \cdot v_m = v_{\tilde{f}_i^k \cdot m} \tag{8}$$

*for all $m \in \mathcal{M}', i \in I$ and $k \in \mathbb{N}$, where $v_0 = 0$ by convention. If the monomial $m$ is extremal of weight $\lambda$, then the vector $v_m$ is an extremal vector of weight $\lambda$. Furthermore if the crystal $\mathcal{M}'$ is connected, then the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V$ is cyclic generated by any $v_m$ with $m \in \mathcal{M}'$.*

*Proof.* Assume that $m$ is extremal of weight $\lambda$: there exists $\{m_w\}_{w \in W}$ such that $m_{\mathrm{Id}} = m$ and

$$\begin{aligned} &\tilde{e}_i \cdot m_w = 0 \quad \text{and} \quad (\tilde{f}_i)^{w(\lambda)(h_i)} \cdot m_w = m_{s_i(w)} \quad \text{if } w(\lambda)(h_i) \geq 0, \\ &\tilde{f}_i \cdot m_w = 0 \quad \text{and} \quad (\tilde{e}_i)^{-w(\lambda)(h_i)} \cdot m_w = m_{s_i(w)} \quad \text{if } w(\lambda)(h_i) \leq 0. \end{aligned} \tag{9}$$

For all $w \in W$, set $v_w = v_{m_w}$. By (8) and (9), $\{v_w\}_{w\in W}$ satisfies $v_{\mathrm{Id}} = v_m$ and

$$x_i^{\pm} \cdot v_w = 0 \quad \text{if } \pm w(\lambda)(h_i) \geq 0 \quad \text{and} \quad (x_i^{\mp})^{(\pm w(\lambda)(h_i))} \cdot v_w = v_{s_i(w)}.$$

Hence the vector $v_m$ is extremal of weight $\lambda$.

Assume that the crystal $\mathcal{M}'$ is connected and fix $m \in \mathcal{M}'$. For $m' \in \mathcal{M}'$, there exists a product $s$ of Kashiwara operators such that $s(m) = m'$. Consider the corresponding operator $S \in \mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$ at the level of $V$, that is, $S$ has the same expression as $s$, where the operators $\tilde{e}_i^k$ (resp. $\tilde{f}_i^k$) are replaced by $(x_i^+)^{(k)}$ (resp. $(x_i^-)^{(k)}$) in the product ($k \in \mathbb{N}, i \in I$). By (8), $S(v_m) = v_{s(m)} = v_{m'}$ and the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V$ is cyclic generated by $v_m$. $\square$

**Lemma 4.9.** *Assume that $\ell = 1, r+1$. For the action of $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$, the $\ell$-weight vector $v_{e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}} \in V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is an extremal vector of weight $\varpi_\ell$. Furthermore*

$$V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}) = \mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \cdot v_{e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}}.$$

*Proof.* Let us begin to show that the basis $(v_m)$ of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ introduced in (6) satisfies the properties in (8). For all $m \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$, $v_m$ is an $\ell$-weight vector of $\ell$-weight $m$ and $\mathrm{wt}(v_m) = \mathrm{wt}(m)$. Fix $i \in I$ and let $j \in I$ be such that $j \neq i$. As a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module, $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ is completely reducible (see (7)) and there exists $k \in \mathbb{Z}$ such that $v_m \in V_0(Y_{\ell,j+kp})^{(j)}$. As the properties in (8) are satisfied in $V_0(Y_{\ell,j+kp})^{(j)}$ (Proposition 4.3), it holds on $v_m$ for $i \in I$.

From there the result is a direct consequence of the above lemma and the fact that $e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}$ is extremal in the connected crystal $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1})$ (Theorem 3.4). $\square$

**Proposition 4.10.** *The $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is irreducible and is a simple $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module. Also,* $\mathrm{Res}(V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}))$ *is isomorphic to $V(\varpi_\ell)$.*

*Proof.* Let $V$ be a nontrivial sub-$\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. As the weight spaces of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ are all of dimension one, there exists $m \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ such that $v_m \in V$. By Lemma 4.8, $v_m$ generates $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ and $V = V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. Hence $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is simple as a $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module and as a $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module. Furthermore, $\mathrm{Res}(V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}))$ is an integrable $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module generated by the extremal vector

$$v_{e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}}$$

of weight $\varpi_\ell$. Then by Theorem 2.13, $\mathrm{Res}(V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}))$ is isomorphic to $V(\varpi_\ell)$. $\square$

Some readers may expect that the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ can be obtained from the extremal weight module $V(\varpi_\ell)$ by an evaluation morphism, but this is not the case for the following reasons (which generalize arguments given in [Hernandez 2009]): $\mathrm{Res}(V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1}))$ is isomorphic to the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module $V(\varpi_\ell)$. In particular,

$$\tau_{p,-\delta} : V(\varpi_\ell) \to V(\varpi_\ell), v_m \mapsto v_{\tau_{p,-\delta}(m)} \quad \text{for all } m \in \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$$

is a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-automorphism of $V(\varpi_\ell)$ (with $p = n+1$ or $p = 2$ if $\ell = 1, n$ or $\ell = r+1$ respectively). If $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is obtained from an evaluation morphism $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \to \mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})$, $\tau_{p,-\delta}$ should induce an automorphism of $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. But it does not commute with the action of the $x_{i,r}^{\pm}, h_{i,r}$ for $i \in I$ and $r \in \mathbb{Z} - \{0\}$. In the same way, $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ can not be obtained from an evaluation morphism $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \to \mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ $(j \in I)$. In fact, $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ is completely reducible as a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module and is a direct sum of fundamental modules (see (7)). But it is a simple $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module.

**Remark 4.11.** Let us denote $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$ the quantum toroidal algebra without derivation element, that is, this is the subalgebra of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ generated by $x_{i,r}^{\pm}$ $(i \in I, r \in \mathbb{Z})$, $h_{i,m}$ $(i \in I, m \in \mathbb{Z} - \{0\})$ and $k_h$ $(h \in \sum \mathbb{Q} h_i)$. An automorphism $\Psi$ of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$ which exchanges vertical and horizontal quantum affine subalgebras is defined in [Miki 1999]. Denote by $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})^{\Psi}$ the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$-module obtained from $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ by twisting the action by $\Psi$. It would be interesting to determine if $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})^{\Psi}$ is already known, for example if it is of $\ell$-highest weight. Actually this is not the case: for the vertical quantum affine subalgebra $\mathcal{U}_q^v(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$, it is an integrable and cyclic module which is reducible. Further as a $\mathcal{U}_q^h(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$-module, it is a completely reducible, direct sum of irreducible finite-dimensional representations. So $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})^{\Psi}$ cannot be an $\ell$-highest weight module or an $\ell$-lowest weight module.

From now on, let $\mathcal{M}_0'$ be a subcrystal of $\mathcal{M}_0$ over $\mathcal{U}_q(\mathrm{sl}_{n+1})$ (resp. $\mathcal{M}'$ subcrystal of $\mathcal{M}$ over $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$). Let us consider the vector space $V$ with basis $(v_m)$ indexed by the vertices of $\mathcal{M}_0'$ (resp. $\mathcal{M}'$). We define an action of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ (resp. $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$) on $V$ by the formulas

$$\begin{aligned} x_{i,r}^{+} \cdot v_m &= q^{r(p_i(m)-1)} v_{\tilde{e}_i \cdot m}, \\ x_{i,r}^{-} \cdot v_m &= q^{r(q_i(m)+1)} v_{\tilde{f}_i \cdot m}, \\ \phi_{i,\pm s}^{\pm} \cdot v_m &= \pm(q-q^{-1})\big(\varphi_i(m) q^{\pm s(q_i(m)+1)} - \varepsilon_i(m) q^{\pm s(p_i(m)-1)}\big) v_m, \\ k_h \cdot v_m &= q^{\mathrm{wt}(m)(h)} v_m, \end{aligned} \tag{10}$$

with $r \in \mathbb{Z}$, $s > 0$, $i \in I_0$ (resp. $i \in I$) and $h \in \mathfrak{h}_0$ (resp. $h \in \mathfrak{h}$), and where $v_0 = 0$ by convention. Note that $p_i(m)$ is well defined only if $\varepsilon_i(m) > 0$ or equivalently if $\tilde{e}_i \cdot m \neq 0$ and $q_i(m)$ is well defined only if $\varphi_i(m) > 0$ or equivalently if $\tilde{f}_i \cdot m \neq 0$. Then, these expressions make sense.

**Theorem 4.12.** (i) *Set $n \in \mathbb{N}^*$, $1 \le \ell \le n$. Assume that $\mathcal{M}'_0 = \mathcal{M}_0(Y_{\ell,k})$. Then the formulas in (10) endow $V$ with a structure of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module isomorphic to the fundamental module $V_0(Y_{\ell,k})$.*

(ii) *Assume that $n = 2r+1$ is odd and $\ell = 1, r+1, n$. Set $\mathcal{M}' = \mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. Then the formulas in (10) endow $V$ with a structure of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module isomorphic to the extremal fundamental loop weight module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$.*

*Proof.* The action of the horizontal quantum affine subalgebra and the action of the Cartan subalgebra are known on the basis $(v_m)_{m \in \mathcal{M}'}$ for the $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-module $V_0(Y_{\ell,k})$ and for the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. From (1) it is straightforward to deduce the action of the $x_{i,r}^{\pm}$ on these modules ($r \in \mathbb{Z}$). We obtain the formulas in (10) given only from the corresponding monomial crystal. □

**Remark 4.13.** In [Hernandez 2011], the algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_\infty)$ is introduced as the quantum affinization of $\mathcal{U}_q(\mathrm{sl}_\infty)$. It is defined by the same generators and relations as in Definition 2.15 with the infinite Cartan matrix $C = (C_{i,j})_{i,j \in \mathbb{Z}}$ such that

$$C_{i,i} = 2, \quad C_{i,i+1} = -1, \quad C_{i+1,i} = -1, \quad C_{i,j} = 0$$

if $i - j \notin \{-1, 0, 1\}$. The representation theory of $\mathcal{U}_q(\hat{\mathrm{sl}}_\infty)$ is similar to the one of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$: the simple $\ell$-highest weight modules are parametrized by Drinfeld polynomials. In particular, the fundamental modules can be defined and they are the inductive limit of the fundamental modules for the quantum affine algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ when $n \to \infty$ (see Theorem 3.8 and Proposition 3.11 in [Hernandez 2011]). So, the previous results about the fundamental modules of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$ extend directly to the case of the fundamental modules of $\mathcal{U}_q(\hat{\mathrm{sl}}_\infty)$.

**Remark 4.14.** As we have said, relations between monomial crystals and the set of monomials occurring in the $q$-character of representations are known and have combinatorial origin (see [Hernandez 2011; Hernandez and Nakajima 2006; Nakajima 2003]). The above results, in particular Theorem 4.12, give one way to better understand the representation theoretical meaning of this narrow link expected in [Hernandez and Nakajima 2006]. In fact, the formulas in (10) hold for all the fundamental $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})'$-modules $V_0(Y_{\ell,k})$ and for all the extremal fundamental loop weight $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-modules $V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$. Hence the knowledge of these representations is reduced to the one of the corresponding crystals $\mathcal{M}_0(Y_{\ell,k})$ and $\mathcal{M}(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$ respectively, which is totally combinatorial.

**Example 4.15.** Assume that $n = 3$ and $\ell = 1$. We study the extremal fundamental loop weight module $V(e^{\varpi_1} Y_{1,0} Y_{0,1}^{-1})$ for $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$. Let us consider the monomial crystal $\mathcal{M}(e^{\varpi_1} Y_{1,0} Y_{0,1}^{-1})$. It is closed and $p = 4$ in this case. Using the notation introduced above, $\mathcal{E} = \bigsqcup_{k \in \mathbb{Z}} \mathcal{E}_{0,k}$ and we have

$$\mathcal{E}_{0,0} = \{e^{\varpi_1} Y_{1,0} Y_{0,1}^{-1}, Y_{2,1} Y_{1,2}^{-1}, Y_{3,2} Y_{2,3}^{-1}, Y_{0,3} Y_{3,4}^{-1}\}.$$

$\mathscr{E}_{0,k}$ can be obtained from $\mathscr{E}_{0,0}$ by applying $\tau_{4k,-\delta}$. In the same way, we obtain $\mathscr{E}_{j,k}$ by applying $\phi^{j+4k}$ to $\mathscr{E}_{0,0}$. Then the $q$-character of the extremal fundamental loop weight module $V(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})$ is

$$\chi_q(V(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})) = \sum_{k\in\mathbb{Z}}\big(e^{\varpi_1-k\delta}Y_{1,4k}Y_{0,1+4k}^{-1} + Y_{2,1+4k}Y_{1,2+4k}^{-1} + Y_{3,2+4k}Y_{2,3+4k}^{-1} + Y_{0,3+4k}Y_{3,4+4k}^{-1}\big).$$

Furthermore the action is explicitly given by the crystal $\mathcal{M}(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})$ and by the formulas in (10). This module was already constructed in [Hernandez 2009].

**Remark 4.16.** After this paper appeared on the arXiv, the constructions in [Feigin et al. 2013] were brought to our attention by H. Nakajima: some representations over the $d$-deformation $\mathcal{U}_{q,d}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ of the quantum toroidal algebra are obtained as the quantum version of a module over a Lie algebra of difference operators. They are called vector representations in [Feigin et al. 2013]. Our works give another way to define these representations. Actually, let

$$\omega : \mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}}) \to \mathcal{U}_{q^{-1}}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$$

be the map sending $x_{i,r}^{\pm}, h_{i,m}, k_h$ to $x_{i,r}^{\mp}, h_{i,m}, k_h$ $(i\in I, r\in\mathbb{Z}, m\in\mathbb{Z}-\{0\}, h\in\mathfrak{h})$. This $\omega$ extends to an isomorphism of algebras. For $u\in\mathbb{C}^*$, let

$$[V(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})]_u$$

be the $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module obtained from $V(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})$ by twisting the action by $t_{uq^{-1}}\circ\omega$. Then $[V(e^{\varpi_1}Y_{1,0}Y_{0,1}^{-1})]_u$ is isomorphic to a vector representation where we specialized the parameter $d$ at 1 (this representation is denoted by $V^{(2)}(u)$ in [Feigin et al. 2013]).

**Example 4.17.** Assume that $n=3$ and $\ell=2$. Let us study the extremal fundamental loop weight module $V(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$ of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$. Consider the closed monomial crystal $\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$. In this case, $p=2$ and we have

$$\mathscr{E}_{0,0} = \left\{\begin{array}{c} e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1},\, Y_{1,1}Y_{2,2}^{-1}Y_{3,1}Y_{0,2}^{-1},\, Y_{1,1}Y_{3,3}^{-1}, \\ Y_{3,1}Y_{1,3}^{-1},\, Y_{1,3}^{-1}Y_{2,2}Y_{3,3}^{-1}Y_{0,2},\, Y_{2,4}^{-1}Y_{0,2} \end{array}\right\}.$$

To describe all the monomials occurring in $\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$, it is sufficient to consider only the sub-$I_{\{0,1\}}$-crystal

$$Y_{2,0}Y_{0,2}^{-1} \xrightarrow{2} Y_{1,1}Y_{2,2}^{-1}Y_{3,1}Y_{0,2}^{-1} \xrightarrow{3} Y_{1,1}Y_{3,3}^{-1}$$

and to apply the $\theta$-twisted automorphism $\phi$ (Remark 3.20). The $q$-character of $V(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$ is

$$\begin{aligned}\chi_q(V(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1}))\\ =\sum_{k\in\mathbb{Z}}e^{\varpi_2-k\delta}\big(&Y_{2,2k}Y_{0,2+2k}^{-1}+Y_{1,1+2k}Y_{2,2+2k}^{-1}\\ &+Y_{3,1+2k}Y_{0,2+2k}^{-1}+Y_{1,3+2k}^{-1}Y_{3,1+2k}+Y_{3,1+2k}Y_{1,3+2k}^{-1}\\ &+Y_{1,3+2k}^{-1}Y_{2,2+2k}+Y_{3,3+2k}^{-1}Y_{0,2+2k}+Y_{2,4+2k}^{-1}Y_{0,2+2k}\big),\end{aligned}$$

and the action of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ on $V(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$ is explicitly given by the crystal $\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$ and the formulas in (10).

**4C.** ***Finite-dimensional representations at roots of unity.*** The existence of shift automorphisms for $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ is related to finite-dimensional representations of quantum toroidal algebras at roots of unity. We explain that in this section.

So assume that $n=2r+1$ is odd ($r\geq 1$) and $\ell=1,n$ or $\ell=r+1$. In this case $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ is closed and its automorphism $z_\ell$ has the special form of a shift $\tau_{-p,\delta}$ with $p=n+1$ or $p=2$ respectively.

Set $L\geq 1$ and $\epsilon$ a primitive $(pL)$-root of unity (we assume also that $p\neq 2$ or $L>1$ in the following). Let $\mathcal{U}_\epsilon(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$ be the algebra defined as $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ with $\epsilon$ instead of $q$ (without divided powers and derivation element).

For $N\in\mathbb{N}^*$, let

$$\Gamma_N:\mathbb{Z}[Y_{i,l}^{\pm 1}]_{i\in I,l\in\mathbb{Z}}\to\mathbb{Z}[Y_{i,\bar{l}}^{\pm 1}]_{i\in I,\bar{l}\in\mathbb{Z}/N\mathbb{Z}}$$

be the map defined by sending the variables $Y_{i,l}^{\pm 1}$ to $Y_{i,\bar{l}}^{\pm 1}$ ($i\in I,l\in\mathbb{Z}$). Set $\mathcal{S}_\epsilon$ the image of a monomial set $\mathcal{S}$ by $\Gamma_{(pL)}$.

Consider the monomial set $\mathcal{E}_\epsilon$. By the existence of the shift automorphism $\tau_{-p,\delta}$, we have

$$\mathcal{E}_\epsilon=\bigsqcup_{0\leq k\leq L-1}(\tau_{p,-\delta})^k((\mathcal{E}_{j,k})_\epsilon)$$

with $j\in I$. One checks easily that $\mathcal{E}_\epsilon$ is closed.

Specializing the representations $V(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})$ at a root of unity $\epsilon$, we obtain:

**Theorem 4.18.** *Assume that $\epsilon$ is a primitive $(pL)$-root of unity. There is an irreducible $\mathcal{U}_\epsilon(\mathrm{sl}_{n+1}^{\mathrm{tor}})'$-module $V(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})_\epsilon$ of dimension $L\binom{n+1}{\ell}$ such that*

$$\chi_\epsilon(V(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})_\epsilon)=\sum_{m\in\mathcal{E}_\epsilon}m.$$

*Furthermore there exists a basis $(v_m)$ of $V(e^{\varpi_\ell}Y_{\ell,0}Y_{0,d_\ell}^{-1})_\epsilon$ indexed by $\mathcal{E}_\epsilon$ such that* *the action on it is given by*

$$\begin{aligned}
x_{i,r}^{+} \cdot v_m &= \epsilon^{r(p_i(m)-1)} v_{\tilde{e}_i \cdot m},\\
x_{i,r}^{-} \cdot v_m &= \epsilon^{r(q_i(m)+1)} v_{\tilde{f}_i \cdot m},\\
\phi_{i,\pm s}^{\pm} \cdot v_m &= \pm(\epsilon-\epsilon^{-1})\big(\varphi_i(m)\epsilon^{\pm s(q_i(m)+1)} - \varepsilon_i(m)\epsilon^{\pm s(p_i(m)-1)}\big) v_m,\\
k_i^{\pm} \cdot v_m &= \epsilon^{\pm(\varphi_i(m)-\varepsilon_i(m))} v_m.
\end{aligned}$$

## 5. Extremal loop weight modules for $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ when the considered monomial crystal is not closed

In this section, we still assume that $n = 2r+1$ is odd and we discuss the case where the considered monomial crystal $\mathcal{M}'$ is not closed. It is not possible here to construct an integrable module whose $q$-character is a sum of monomials occurring in $\mathcal{M}'$. In fact some monomials miss and we have to consider a larger closed monomial crystal $\bar{\mathcal{M}}'$ containing it. It is obtained from $\mathcal{M}'$ by adding other monomial crystals. But its structure is more complicated than $\mathcal{M}'$ and it is difficult for us to construct systematically a possible representation of $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ associated to $\bar{\mathcal{M}}'$.

So we propose to treat an example of such a construction. Assume $n = 3$ and consider the crystal $\mathcal{M}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$, which is not closed. We determine a closed monomial crystal $\bar{\mathcal{M}}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ containing it and we construct a representation $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ such that its $q$-character is the sum of monomials occurring with multiplicity one in $\bar{\mathcal{M}}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ (Theorem 5.6). We will see that the definition of extremal loop weight module is satisfied by $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$.

Section 5A, we study the crystal $\mathcal{M}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ and determine a closed monomial crystal $\bar{\mathcal{M}}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$, containing it.

The construction of the $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ is done in Section 5B. The process is the same as in the preceding section: we consider the vector space freely generated by the vertices $m$ of $\bar{\mathcal{M}}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ and we define an action of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ by pasting together some finite-dimensional representations of the vertical quantum affine subalgebras $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$ $(j \in I)$.

In Section 5C, we study the representation $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$: it is an integrable representation of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ which is thin and irreducible. Furthermore $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ is an extremal loop weight module of $\ell$-weight $e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1}$.

In Section 5D, we specialize $q$ at roots of unity $\epsilon$. We get finite-dimensional representations of the quantum toroidal algebra $\mathcal{U}_\epsilon(\mathrm{sl}_4^{\mathrm{tor}})'$.

**Remark 5.1.** It could be interesting to construct other extremal fundamental loop weight modules of $\ell$-weight $e^{\varpi_\ell} Y_{\ell,0} Y_{0,\ell}^{-1}$ with $2 \leq \ell \leq r$ in the same way. The first crystal $\mathcal{M}(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ which is not closed is obtained for $n=5$ and $\ell=2$. We are led to consider the closed crystal

$$\bar{\mathcal{M}}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})=\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})\oplus\bigoplus_{s\in\mathbb{N}^*}\mathcal{M}(e^{2\Lambda_1-s\delta}Y_{1,1}Y_{1,-1+6s}Y_{0,2}^{-1}Y_{0,6s}^{-1}),$$

which contains $\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})$. The maps $\phi$ and $\tau_{6,-2\delta}$ are automorphisms of it and the $P_{\mathrm{cl}}$-crystals

$$\mathcal{M}(e^{\varpi_2}Y_{2,0}Y_{0,2}^{-1})/(\tau_{6,-2\delta})\quad\text{and}\quad\mathcal{M}(e^{2\Lambda_1-s\delta}Y_{1,1}Y_{1,-1+6s}Y_{0,2}^{-1}Y_{0,6s}^{-1})/(\tau_{6,-2\delta})$$

have 30 vertices and 36 vertices respectively.

The example we propose to treat in this section is simpler than the case of the extremal fundamental loop weight modules and we focus only on this situation for the sake of clarity and simplicity.

**5A.** ***Study of the monomial*** $\mathcal{U}_q(\hat{\mathfrak{sl}}_4)$***-crystal*** $\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$**.** We refer to the Appendix for explicit descriptions of all the crystals considered in this section. Let us study the monomial crystal $\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$: the maps $\phi$ and $\tau_{4,-2\delta}$ are automorphisms of $\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. Further straightforward computations lead to the following result.

**Proposition 5.2.** (i) *We have the equality of sets*

$$\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})=\bigcup_{k\in\mathbb{Z}}(\tau_{4,-2\delta})^k\left(\bigcup_{j=0}^{3}\mathcal{M}_{I_j}(Y_{1+j,1+j}Y_{1+j,-1+j}Y_{j,2+j}^{-1}Y_{j,j}^{-1})\right).$$

(ii) *For all $j\in I$, the monomial crystal $\mathcal{M}_{I_j}(Y_{1+j,1+j}Y_{1+j,-1+j}Y_{j,2+j}^{-1}Y_{j,j}^{-1})$ is $I_j$-$q$-closed. More precisely, we have the bijection of monomial sets*

$$\Xi^j:\mathcal{M}_{I_j}(Y_{1+j,1+j}Y_{1+j,-1+j}Y_{j,2+j}^{-1}Y_{j,j}^{-1})\longrightarrow\mathcal{M}(V_0(Y_{1,1+j}Y_{1,-1+j})^{(j)}),$$

*where $V_0(Y_{1,1+j}Y_{1,-1+j})$ is the simple $\ell$-highest weight representation of $\mathcal{U}_q(\hat{\mathfrak{sl}}_4)'$ of $\ell$-highest weight $Y_{1,1+j}Y_{1,-1+j}$.*

(iii) *For all $j\in I$, the $I_j$-crystal $\mathcal{M}_{I_j}(Y_{1+j,1+j}Y_{1+j,-1+j}Y_{j,2+j}^{-1}Y_{j,j}^{-1})$ is not $q$-closed: the monomial $\phi^j(Y_{1,-1}Y_{3,5}^{-1}Y_{0,0}^{-1}Y_{0,4})$ occurs in this crystal, but it is not the case of $\phi^j(Y_{1,-1}Y_{3,5}^{-1}Y_{0,0}^{-1}Y_{0,4}\cdot A_{0,-1})=\phi^j(Y_{1,5}Y_{1,-1}Y_{0,6}^{-1}Y_{0,0}^{-1})$.*

Hence, we are led to consider the crystal $\mathcal{M}(e^{2\varpi_1+\delta}Y_{1,1}Y_{1,-5}Y_{0,2}^{-1}Y_{0,-4}^{-1})$, which is also not closed. More generally we have to deal with all the monomial crystals $\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1})$ with $s\in\mathbb{N}$. We set

$$\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})=\bigoplus_{s\in\mathbb{N}}\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1}).$$

For all $(k,s) \in \mathbb{Z}\times\mathbb{N}$ and $j \in I$, denote by

- $\mathcal{M}^1_{j,k,s}$ the sub-$I_j$-crystal of $\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s})$ generated by the monomial $\phi^{j+4k}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s})$,
- $\mathcal{M}^2_{j,k,s}$ the sub-$I_j$-crystal of $\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s})$ generated by the monomial $\phi^{j+4k}(Y_{1,1}Y^{-1}_{1,-3-4s}Y_{2,-4-4s}Y^{-1}_{0,2})$.

**Proposition 5.3.** (i) *For all $s \in \mathbb{N}$ and $j \in I$, one has the equality of $I_j$-crystals*

$$\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s}) = \bigoplus_{k\in\mathbb{Z}}(\mathcal{M}^1_{j,k,s}\oplus\mathcal{M}^2_{j,k,s}).$$

(ii) *For all $j \in I$, $k \in \mathbb{Z}$ and $s \geq 1$, the monomial crystal $\mathcal{M}_{j,k,s} = \mathcal{M}^1_{j,k,s}\oplus\mathcal{M}^2_{j,k,s-1}$ is $I_j$-$q$-closed. More precisely, we have the bijection of monomial sets*

$$\Xi^j : \mathcal{M}_{j,k,s} \longrightarrow \mathcal{M}\big(V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k-4s})^{(j)}\big),$$

*where $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k-4s})$ is the $\ell$-highest weight representation of $\mathcal{U}_q(\hat{\mathrm{sl}}_4)'$ of $\ell$-highest weight $Y_{1,1+j+4k}Y_{1,-1+j+4k-4s}$.*

The proof of these statements is straightforward. As a consequence of these results, we have:

**Corollary 5.4.** *The monomial crystal $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y^{-1}_{0,2}Y^{-1}_{0,0})$ is closed.*

**Proposition 5.5.** *$\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y^{-1}_{0,2}Y^{-1}_{0,0})$ is a monomial realization of the $P$-crystal $\mathcal{B}(2\varpi_1)$. Further, the monomials $M_s = e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s}$ are extremal of weight $2\varpi_1+s\delta$ ($s \in \mathbb{N}$).*

*Proof.* The monomial crystal $\mathcal{M}(e^{2\varpi_1}Y^2_{1,1}Y^{-2}_{0,2})$ is isomorphic to the connected component of $\mathcal{B}(2\varpi_1)$ generated by $v_{2\varpi_1}$ [Hernandez and Nakajima 2006, Proposition 3.1]. One checks that the map

$$\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y^{-1}_{0,2}Y^{-1}_{0,-4s}) \longrightarrow \mathcal{M}(e^{2\varpi_1}Y^2_{1,1}Y^{-2}_{0,2})$$

which sends the monomial $M_s$ to the extremal element $e^{2\varpi_1}Y^2_{1,1}Y^{-2}_{0,2}$ is an isomorphism of $P_{\mathrm{cl}}$-crystals for all $s \in \mathbb{N}$. Then the result is a direct consequence of the description of the crystal $\mathcal{B}(2\varpi_1)$ given in [Beck and Nakajima 2004]: all the connected components of $\mathcal{B}(2\varpi_1)$ are isomorphic to each other modulo shift of weight by $\delta$. ☐

**5B.** ***Construction of the $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_4)$-module $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y^{-1}_{0,2}Y^{-1}_{0,0})$.*** Let us give the main result of this section.

**Theorem 5.6.** *There exists a thin representation of $\mathcal{U}_q(\mathrm{sl}^{\mathrm{tor}}_4)$ whose $q$-character is the sum of monomials occurring in $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y^{-1}_{0,2}Y^{-1}_{0,0})$ with multiplicity one. It is denoted by $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y^{-1}_{0,2}Y^{-1}_{0,0})$.*

The construction of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is analogous to the one of $V(e^{\varpi_\ell}Y_{\ell,0}Y_{0,\ell}^{-1})$ in Theorem 4.1: we paste together the finite-dimensional representations $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k})^{(j)}$ and $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k-4s})^{(j)}$ of $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$ with $j \in I$, $k \in \mathbb{Z}$ and $s \in \mathbb{N}^*$.

Let us begin by recalling some well-known facts about the Kirillov–Reshetikhin module $V_0(\Xi^0(M))$ over $\mathcal{U}_q(\hat{\mathrm{sl}}_4)'$ with $M = e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}$. It is irreducible as a $\mathcal{U}_q(\mathrm{sl}_4)$-module, isomorphic to $V_0(2\Lambda_1)$. In particular, $V_0(\Xi^0(M))$ is an extremal weight module of extremal weight $2\Lambda_1$ and there exist vectors $v_{\phi^j(M)}$ ($j = 0, \dots, 3$) such that $v_M$ is an $\ell$-highest weight vector of $V_0(\Xi^0(M))$ and

$$\begin{aligned}(x_{i,0}^-)^{(2)} \cdot v_{\phi^{i-1}(M)} &= v_{\phi^i(M)} && \text{for } i = 1, \dots, 3,\\ (x_{i,0}^+)^{(2)} \cdot v_{\phi^i(M)} &= v_{\phi^{i-1}(M)} && \text{for } i = 1, \dots, 3,\\ x_{i,0}^\pm \cdot v_{\phi^j(M)} &= 0 && \text{in the other cases.}\end{aligned}$$

Set

$$v_{\tilde{f}_1\cdot M} := x_{1,0}^- \cdot v_M, \quad v_{\tilde{f}_2\tilde{f}_1\cdot M} := x_{2,0}^- x_{1,0}^- \cdot v_M, \quad v_{\tilde{f}_3\tilde{f}_2\tilde{f}_1\cdot M} := x_{3,0}^- x_{2,0}^- x_{1,0}^- \cdot v_M,$$
$$v_{\tilde{f}_2\cdot\phi(M)} := x_{2,0}^- \cdot v_{\phi(M)}, \quad v_{\tilde{f}_3\tilde{f}_2\cdot\phi(M)} := x_{3,0}^- x_{2,0}^- \cdot v_{\phi(M)},$$
$$v_{\tilde{f}_3\cdot\phi^2(M)} := x_{3,0}^- \cdot v_{\phi^2(M)}.$$

These vectors form a basis $(v_m)$ of $V_0(\Xi^0(M))$, indexed by the monomials occurring in $\mathcal{M}_{I_0}(M)$. Furthermore for all $m \in \mathcal{M}_{I_0}(M)$, $v_m$ is an $\ell$-weight vector of $\ell$-weight $\Xi^0(m)$.

The other finite-dimensional representations of $\mathcal{U}_q(\hat{\mathrm{sl}}_4)'$ we have to consider are $V_0(\Xi^0(M_s))$ with $M_s = e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1}$ and $s \in \mathbb{N}^*$. The following two points are well known:

(i) $V_0(\Xi^0(M_s))$ is an irreducible $\mathcal{U}_q(\hat{\mathrm{sl}}_4)'$-module isomorphic to

$$V_0(Y_{1,1}) \otimes V_0(Y_{1,-1-4s}).$$

(ii) $\mathrm{Res}(V_0(\Xi^0(M_s)))$ is a completely reducible $\mathcal{U}_q(\mathrm{sl}_4)$-module isomorphic to $V_0(2\Lambda_1) \oplus V_0(\Lambda_2)$.

Furthermore there exist vectors $v_{\phi^j(M_s)}$ ($j = 0, \dots, 3$) such that $v_{M_s}$ is an $\ell$-highest weight vector of $V_0(\Xi^0(M_s))$ and

$$\begin{aligned}(x_{i,0}^-)^{(2)} \cdot v_{\phi^{i-1}(M_s)} &= v_{\phi^i(M_s)} && \text{for } i = 1, \dots, 3,\\ (x_{i,0}^+)^{(2)} \cdot v_{\phi^i(M_s)} &= v_{\phi^{i-1}(M_s)} && \text{for } i = 1, \dots, 3,\\ x_{i,0}^\pm \cdot v_{\phi^j(M_s)} &= 0 && \text{in the other cases.}\end{aligned}$$

To complete this family of vectors to a basis of $V_0(\Xi^0(M_s))$, the following example is used.

**Example 5.7.** Let $a, b \in \mathbb{Z}$ be such that $a \neq b$ and $a \neq b \pm 2$. Consider the $\mathcal{U}_q(\hat{\mathrm{sl}}_2)'$-module $V_0(Y_{1,a}Y_{1,b})$. This module was already studied in [Hernandez 2010]. We have

$$\chi_q(V_0(Y_{1,a}Y_{1,b})) = Y_{1,a}Y_{1,b} + Y_{1,a}Y_{1,b+2}^{-1} + Y_{1,a+2}^{-1}Y_{1,b} + Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}.$$

In particular, it was shown that there exists a basis

$$\{v_{Y_{1,a}Y_{1,b}}, v_{Y_{1,a+2}^{-1}Y_{1,b}}, v_{Y_{1,a}Y_{1,b+2}^{-1}}, v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}}\},$$

where the action of the Drinfeld generators on it is given by

$$\begin{aligned}
x_r^+ \cdot v_{Y_{1,a}Y_{1,b}} &= 0,\\
x_r^- \cdot v_{Y_{1,a}Y_{1,b}} &= \\
&\quad \frac{q^{b-1}-q^{a+1}}{q^b-q^a} q^{r(a+1)} v_{Y_{1,a+2}^{-1}Y_{1,b}} + \frac{q^{b+1}-q^{a-1}}{q^b-q^a} q^{r(b+1)} v_{Y_{1,a}Y_{1,b+2}^{-1}},\\
x_r^+ \cdot v_{Y_{1,a+2}^{-1}Y_{1,b}} &= q^{r(a+1)} v_{Y_{1,a}Y_{1,b}},\\
x_r^- \cdot v_{Y_{1,a+2}^{-1}Y_{1,b}} &= q^{r(b+1)} v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}},\\
x_r^+ \cdot v_{Y_{1,a}Y_{1,b+2}^{-1}} &= q^{r(b+1)} v_{Y_{1,a}Y_{1,b}},\\
x_r^- \cdot v_{Y_{1,a}Y_{1,b+2}^{-1}} &= q^{r(a+1)} v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}},\\
x_r^+ \cdot v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}} &= \\
&\quad \frac{q^{b-1}-q^{a+1}}{q^b-q^a} q^{r(b+1)} v_{Y_{1,a+2}^{-1}Y_{1,b}} + \frac{q^{b+1}-q^{a-1}}{q^b-q^a} q^{r(a+1)} v_{Y_{1,a}Y_{1,b+2}^{-1}},\\
x_r^- \cdot v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}} &= 0,
\end{aligned}$$

and with $v_m$ of $\ell$-weight $m$ for $m = Y_{1,a}Y_{1,b}, \dots, Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}$. Note that the basis used in [Hernandez 2010] is renormalized here and we have

$$(x_0^-)^{(2)} \cdot v_{Y_{1,a}Y_{1,b}} = v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}}, \quad (x_0^+)^{(2)} \cdot v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}} = v_{Y_{1,a}Y_{1,b}}.$$

As $a \neq b \pm 2$, it is well known that the $\mathcal{U}_q(\hat{\mathrm{sl}}_2)'$-module $V_0(Y_{1,a}Y_{1,b})$ is isomorphic to $V_0(Y_{1,a}) \otimes V_0(Y_{1,b})$. Furthermore the $\mathcal{U}_q(\mathrm{sl}_2)$-module $\mathrm{Res}(V_0(Y_{1,a}Y_{1,b}))$ is not irreducible, but it is cyclic generated by one of the vectors $v_{Y_{1,a}Y_{1,b+2}^{-1}}$ or $v_{Y_{1,a+2}^{-1}Y_{1,b}}$.

Set $M_s^1 = Y_{1,3}^{-1}Y_{1,-1-4s}Y_{2,2}Y_{0,-4s}^{-1}$ and $M_s^2 = Y_{1,1}Y_{1,1-4s}^{-1}Y_{2,-4s}Y_{0,2}^{-1}$. Let $v_{M_s^1}$ and $v_{M_s^2} \in V_0(\Xi^0(M))$ of $\ell$-weight $M_s^1$ and $M_s^2$, respectively, be such that

$$x_{1,0}^- \cdot v_{M_s} = \frac{q^{-2-4s}-q^4}{q^{-1-4s}-q^3} v_{M_s^1} + \frac{q^{-4s}-q^2}{q^{-1-4s}-q^3} v_{M_s^2}.$$

Set

$$v_{\tilde{f}_2\cdot M_s^u} := x_{2,0}^- \cdot v_{M_s^u}, \quad v_{\tilde{f}_3\tilde{f}_2\cdot M_s^u} := x_{3,0}^- x_{2,0}^- \cdot v_{M_s^u}$$

with $u = 1, 2$. In the same way, one can define

$$v_{\phi(M_s^u)}, \quad v_{\tilde{f}_3\cdot\phi(M_s^u)} \quad \text{and} \quad v_{\phi^2(M_s^u)}$$

for $u = 1, 2$. We check that these vectors form a basis $(v_m)$ of $V_0(\Xi^0(M_s))$, indexed by the monomials occurring in $\mathcal{M}_{0,0,s}$. Moreover $v_m$ is an $\ell$-weight vector of $\ell$-weight $\Xi^0(m)$ for all $m$.

By twisting the action of $\mathcal{U}_q(\hat{\mathrm{sl}}_4)'$ on $V_0(Y_{1,1}Y_{1,-1})$ and $V_0(Y_{1,1}Y_{1,-1-4s})$ by $\theta^{(j)}$ and $t_b$ for some $b \in \mathbb{C}^*$, we obtain for all $j \in I, k \in \mathbb{Z}$ and $s \in \mathbb{N}^*$

- the $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$-modules $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k})^{(j)}$, called modules of type KR below;
- the $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$-modules $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k-4s})^{(j)}$, called modules of type $s$-TP below. The modules of type $s$-TP for various $s \in \mathbb{N}^*$ are called modules of type TP.

From the construction done above, we get bases $(v_m)$ of these modules indexed by the monomial crystals $\mathcal{M}_{I_j}(\phi^{j+4k}(Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}))$ (resp. $\mathcal{M}_{j,k,s}$) with analogous properties as the previous ones. In particular, the action on a vector $v_m$ is completely determined by the action of the horizontal quantum affine subalgebra on it and by its $\ell$-weight $m$.

We begin the construction of the $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. Denote by $\mathcal{E}$ the set of monomials occurring in $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ and for all $j \in I$, $k \in \mathbb{Z}$ and $s \in \mathbb{N}^*$, $\mathcal{E}_{j,k,0}$ (resp. $\mathcal{E}_{j,k,s}$) the set of monomials corresponding to $\mathcal{M}_{j,k,0}^1$ (resp. $\mathcal{M}_{j,k,s}$). We have for all $0 \le j \le 3$,

$$\mathcal{E} = \bigsqcup_{k\in\mathbb{Z}} \mathcal{E}_{j,k,0} \sqcup \bigsqcup_{k\in\mathbb{Z},s\in\mathbb{N}^*} \mathcal{E}_{j,k,s}.$$

Let

$$V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}) = \bigoplus_{m\in\mathcal{E}} \mathbb{C}v_m$$

be the vector space freely generated by $\mathcal{E}$. For $0 \le j \le 3$, $k \in \mathbb{Z}$ and $s \in \mathbb{N}^*$, set $V_k^{(j)} = \bigoplus_{m\in\mathcal{E}_{j,k,0}} \mathbb{C}v_m$ (resp. $V_{k,s}^{(j)} = \bigoplus_{m\in\mathcal{E}_{j,k,s}} \mathbb{C}v_m$). Then for all $0 \le j \le 3$,

$$V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}) = \bigoplus_{k\in\mathbb{Z}} V_k^{(j)} \oplus \bigoplus_{k\in\mathbb{Z},s\in\mathbb{N}^*} V_{k,s}^{(j)}.$$

We endow $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ with a structure of $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$-module for all $j \in I$ as follows: for $k \in \mathbb{Z}$ and $s \in \mathbb{N}^*$, $V_k^{(j)}$ (resp. $V_{k,s}^{(j)}$) is isomorphic to $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k})^{(j)}$ (resp. $V_0(Y_{1,1+j+4k}Y_{1,-1+j+4k-4s})^{(j)}$) by identifying the corresponding bases. So $V_k^{(j)}$ (resp. $V_{k,s}^{(j)}$) is endowed with a structure of $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$-module, and $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ also by direct sum. We denote it by $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})^{(j)}$.

**Proposition 5.8.** *There exists a $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module structure on the vector space $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ such that for all $j \in I$ the induced $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module is isomorphic to $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})^{(j)}$. Furthermore the $q$-character of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is*

$$\chi_q\big(V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})\big) = \sum_{m\in\mathscr{E}} m,$$

*where $\mathscr{E}$ is the set of monomials occurring in $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$.*

*Proof.* The process is the same as in Theorem 4.1: to define an action of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$, we determine the action of the subalgebras $\hat{\mathcal{U}}_i$ for all $i \in I$. For that, let $j \in I$ be such that $j \neq i$. Then the action of $\hat{\mathcal{U}}_i$ on $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is the restriction of the action of $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$ on $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})^{(j)}$. We check that this is independent of the choice of $j \neq i$.

Let us show that this action endows $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ with a structure of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module. For that, we have to distinguish two types of monomials:

- the $m$ such that there is no $s, s' \in \mathbb{N}$ with $s \neq s'$ and $m \in \mathscr{E}_{j,k,s} \cap \mathscr{E}_{j',k',s'}$ for some $0 \leq j, j' \leq 3$ and $k, k' \in \mathbb{Z}$. For such a monomial, the defined action on $v_m$ comes from the same type of modules, that is, only of modules of type KR or only on modules of type $s$-TP for one $s \in \mathbb{N}^*$;
- the $m$ such that there is $s, s' \in \mathbb{N}$ with $s \neq s'$ and $m \in \mathscr{E}_{j,k,s} \cap \mathscr{E}_{j',k',s'}$ for some $0 \leq j, j' \leq 3$ and $k, k' \in \mathbb{Z}$. For such a monomial, the defined action on $v_m$ comes from two different types of modules, that is, of modules of type KR and of type TP or of modules of type $s$-TP and of type $s'$-TP with $s \neq s'$.

For the first ones, the same process as in Theorem 4.1 (using promotion operator) implies that the defining relations of $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ hold on it. For the other ones, this is more complicated. Such a monomial is of the form $m = \phi^{j+4k}(Y_{1,-1-4s}Y_{3,5}^{-1}Y_{0,4}Y_{0,-4s}^{-1})$ with $0 \leq j \leq 3$, $k \in \mathbb{Z}$, $s \in \mathbb{N}^*$. The promotion operator implies some relations on $v_m$ but not all and we check directly that they are satisfied. We do not detail the calculations here. $\square$

### 5C. *Study of the $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$.*

**Proposition 5.9.** *The $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is integrable. Moreover, it satisfies property* (iv) *of Remark 2.5.*

*Proof.* For all $j \in I$, $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ is completely reducible as a $\mathcal{U}_q^{v,j}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module and we have

$$V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})^{(j)} = \bigoplus_{\substack{s\in\mathbb{N}\\ k\in\mathbb{Z}}} V_0(Y_{1,1+j+4k} Y_{1,-1+j+4k-4s})^{(j)}. \tag{11}$$

The representations occurring in the direct sum on the right-hand side are integrable. Hence $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ is an integrable $\mathcal{U}_q(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-module. Furthermore the modules of type KR are all isomorphic as $\mathcal{U}_q(\mathrm{sl}_4)$-modules and satisfy property (iv) of Remark 2.5; the same is true of the modlules of type TP. Therefore, we have

$$V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})_{\nu+N\alpha_i} = \{0\} \quad \text{for all } \nu \in P, i \in I, N \gg 0. \qquad \square$$

**Remark 5.10.** The weight spaces of $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ have infinite dimension and property (iii) of Remark 2.5 does not hold. However its $\ell$-weight spaces are all of dimension one.

The main result of this section is the following:

**Theorem 5.11.** *Set $M = e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1}$. The representation $V(M)$ is an extremal loop weight module generated by the vector $v_M$ of $\ell$-weight $M$.*

*Proof.* The third point of Definition 2.26 is a consequence of (11). For the first two points, we use the following results. $\square$

**Lemma 5.12.** *Let $V$ be a $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$-module with basis $(v_m)_{m\in\mathcal{M}'}$ indexed by a subcrystal $\mathcal{M}'$ of $\mathcal{M}$. Assume that $M \in \mathcal{M}'$ is extremal of weight* $\mathrm{wt}(M)$ *and for all $i \in I$ and $m \in W \cdot M$,*

$$\mathrm{wt}(v_m) = \mathrm{wt}(m), \quad x_i^{\pm} \cdot v_m = 0$$

*and*

$$(x_i^{\mp})^{(\pm\mathrm{wt}(m)(h_i))} \cdot v_m = v_{S_i(m)} \quad \textit{if } \pm\mathrm{wt}(m)(h_i) \geq 0.$$

*Then $v_M$ is an extremal vector of weight* $\mathrm{wt}(M)$.

*Proof.* The proof is analogous to the one of Lemma 4.8. $\square$

**Corollary 5.13.** *Set*

$$M_s = e^{2\varpi_1+s\delta} Y_{1,1} Y_{1,-1-4s} Y_{0,2}^{-1} Y_{0,-4s}^{-1} \quad (s \in \mathbb{N}).$$

*Then $v_{M_s}$ is an extremal vector of $V(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1})$ of weight $2\varpi_1 + s\delta$ for the horizontal quantum affine subalgebra $\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$.*

*Proof.* By construction of the basis $(v_m)$ of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$, we have $\mathrm{wt}(v_m)=\mathrm{wt}(m)$ for all $m$. Furthermore a monomial in $W\cdot M_s$ is of the form $\phi^{j+4k}(M_s)$ with $j\in I_0$ and $k\in\mathbb{Z}$ and we have $\mathrm{wt}(\phi^{j+4k}(M_s))=2\Lambda_{j+1}-2\Lambda_j-2k\delta$,

$$
\begin{aligned}
(x_{j,0}^{+})^{(2)}\cdot v_{\phi^{j+4k}(M_s)} &= v_{\phi^{j-1+4k}(M_s)} = v_{S_j(\phi^{j+4k}(M_s))},\\
(x_{j+1,0}^{-})^{(2)}\cdot v_{\phi^{j+4k}(M_s)} &= v_{\phi^{j+1+4k}(M_s)} = v_{S_{j+1}(\phi^{j+4k}(M_s))},\\
x_i^{\pm}\cdot v_{\phi^{j+4k}(M_s)} &= 0 \quad \text{in the other cases.}
\end{aligned}
$$

Hence the hypotheses of the above lemma are satisfied and $v_{M_s}$ is extremal of weight $2\varpi_1+s\delta$ for the horizontal quantum affine subalgebra $\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$. $\square$

**Proposition 5.14.** *The representation $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is cyclic as a $\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$-module generated by the vector $v=v_{e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}}$.*

*Proof.* Consider the sub-$\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$-module $V$ generated by $v$. By construction of the basis $(v_m)$ of

$$V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}),$$

$v_m\in V$ for all $m\in\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. By a recursive argument, assume for one $s\in\mathbb{N}$ we have $v_m\in V$ for all $m\in\mathcal{M}(e^{2\varpi_1-t\delta}Y_{1,1}Y_{1,-1-4t}Y_{0,2}^{-1}Y_{0,-4t}^{-1})$ with $0\le t\le s$. In particular $v_{Y_{1,-1-4s}Y_{3,5}^{-1}Y_{0,4}Y_{0,-4s}^{-1}}$ is in $V$ and by Example 5.7,

$$x_{0,0}^{-}\cdot v_{Y_{1,-1-4s}Y_{3,5}^{-1}Y_{0,4}Y_{0,-4s}^{-1}} = v_{e^{2\varpi_1-(s+1)\delta}Y_{1,5}Y_{1,-1-4s}Y_{0,6}^{-1}Y_{0,-4s}^{-1}}\in V.$$

In the same way

$$v_{\phi^k(Y_{1,-1-4s}Y_{3,5}^{-1}Y_{0,4}Y_{0,-4s}^{-1})} \quad \text{and} \quad v_{\phi^k(e^{2\varpi_1-(s+1)\delta}Y_{1,5}Y_{1,-1-4s}Y_{0,6}^{-1}Y_{0,-4s}^{-1})}$$

are in $V$ for any $k\in\mathbb{Z}$. All $v_m$ with $m\in\mathcal{M}(e^{2\varpi_1-(s+1)\delta}Y_{1,5}Y_{1,-1-4s}Y_{0,6}^{-1}Y_{0,-4s}^{-1})$ can be obtained from these vectors by action of $\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$: this is straightforward from Example 5.7 and the construction of the basis $(v_m)$. $\square$

**Proposition 5.15.** *The $\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is irreducible.*

*Proof.* Let $V$ be a nontrivial sub-$\mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$-module of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. As the $\ell$-weight spaces are of dimension one, there exists $s\in\mathbb{N}$ and a monomial $m\in\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1})$ such that $v_m\in V$.

If $s=0$, we have already shown that $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ is cyclic generated by $v_m$ and $V=V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. Assume that $s\in\mathbb{N}^*$. By Example 5.7 and the construction of $(v_m)$, there exists $x\in\mathcal{U}_q^h(\mathrm{sl}_4^{\mathrm{tor}})$ such that

$$x\cdot v_m = v_{e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1}}.$$

Furthermore

$$\hat{\mathcal{U}}_1\cdot v_{e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1}}$$

is the simple $\ell$-highest weight $\hat{\mathcal{U}}_1$-module of $\ell$-highest weight $Y_{1,1}Y_{1,-1-4s}$ and there exists $y \in \mathcal{U}_q(\mathrm{sl}_4^{\mathrm{tor}})$ such that

$$y \cdot v_m = v_{Y_{1,1}Y_{1,1-4s}^{-1}Y_{2,-4s}Y_{0,2}^{-1}}$$

with $Y_{1,1}Y_{1,1-4s}^{-1}Y_{2,-4s}Y_{0,2}^{-1} \in \mathcal{M}(e^{2\varpi_1+(s-1)\delta}Y_{1,1}Y_{1,-1-4(s-1)}Y_{0,2}^{-1}Y_{0,-4(s-1)}^{-1})$. Repeating this argument, one shows that the vector $v_{e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}}$ is in $V$. By the above proposition we get

$$V = V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}). \qquad \square$$

**Proposition 5.16.** *The $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-module* $\mathrm{Res}(V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}))$ *has a crystal basis isomorphic to $\mathcal{B}(2\varpi_1)$.*

*Proof.* Set $K = \mathbb{C}(q)$ with $q$ an indeterminate and let $A$ be the subring of $K$ consisting of rational functions in $K$ without pole at $q = 0$. We normalize the basis $(v_m)$ of the $\mathbb{C}(q)$-vector space $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ as follows. For all $m \in \bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$, let $w_m$ be the vector defined by $w_m = (1/q)v_m$ if there exists $k \in \mathbb{Z}$, $s \in \mathbb{N}^*$ such that $m = \phi^k(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1})$ and $w_m = v_m$ otherwise. Set $\mathcal{B} = (w_m)_m$ and $\mathcal{L} = \bigoplus_m Aw_m$. We check directly that $(\mathcal{L}, \mathcal{B})$ is a crystal basis of the $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-module $\mathrm{Res}(V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}))$ isomorphic to $\mathcal{B}(2\varpi_1)$. We do not detail the calculations. $\square$

**Remark 5.17.** All these results suggest that $\mathrm{Res}(V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}))$ is isomorphic to the extremal weight $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-module $V(2\varpi_1)$. One expects to prove such a result for all the extremal loop weight modules constructed by the conjectural process given above.

**5D.** ***Finite-dimensional representations at roots of unity.*** Set $L \geq 1$ and let $\epsilon$ be a primitive $(4L)$-root of unity.

Denote by $\mathcal{E}_s$ the set of monomials occurring in

$$\mathcal{M}(e^{2\varpi_1+s\delta}Y_{1,1}Y_{1,-1-4s}Y_{0,2}^{-1}Y_{0,-4s}^{-1})$$

for all $s \in \mathbb{N}$. Consider $\mathcal{E}'$ the subset of $\mathcal{E}$ defined by

$$\mathcal{E}' = \bigsqcup_{0 \leq s \leq L-1} \mathcal{E}_s.$$

Let $\mathcal{E}_\epsilon$ and $\mathcal{E}'_\epsilon$ be the images of the sets $\mathcal{E}$ and $\mathcal{E}'$ respectively by the map $\Gamma_{(4L)}$: $\mathcal{E}'_\epsilon$ is a finite monomial set of cardinality $16L^2$.

**Theorem 5.18.** *Assume that $\epsilon$ is a primitive $4L$-root of unity. There exists an irreducible $\mathcal{U}_\epsilon(\mathrm{sl}_4^{\mathrm{tor}})'$-module $V_\epsilon$ of dimension $16L^2$ such that*

$$\chi_\epsilon(V_\epsilon) = \sum_{m \in \mathcal{E}'_\epsilon} m.$$

*Proof.* The main difficulty is to specialize $q$ at $\epsilon$ in the $\mathcal{U}_q^{v,j}(\mathrm{sl}_4^{\mathrm{tor}})$-modules of type TP. In fact, these modules can be undefined or reducible after specialization. To better understand these phenomena, let us study the specialized $\mathcal{U}_\epsilon(\hat{\mathrm{sl}}_2)'$-module $V_0(Y_{1,a}Y_{1,b})_\epsilon$ with $a, b \in \mathbb{Z}$. This representation is well defined if $a \notin b + 4L\mathbb{Z}$. Assume that in the following and study $V_0(Y_{1,a}Y_{1,b})_\epsilon$. If $a \notin b \pm 2 + 4L\mathbb{Z}$, this representation is irreducible. If $a \in b + 2 + 4L\mathbb{Z}$, it is not irreducible: in fact

$$\mathcal{U}_\epsilon(\hat{\mathrm{sl}}_2)' \cdot v_{Y_{1,a}Y_{1,b}} = \mathbb{C}v_{Y_{1,a}Y_{1,b}} \oplus \mathbb{C}v_{Y_{1,a+2}^{-1}Y_{1,b}} \oplus \mathbb{C}v_{Y_{1,a+2}^{-1}Y_{1,b+2}^{-1}}$$

is an irreducible submodule of $V_0(Y_{1,a}Y_{1,b})_\epsilon$.

By our study of the $\mathcal{U}_\epsilon(\hat{\mathrm{sl}}_2)'$-module $V_0(Y_{1,a}Y_{1,b})_\epsilon$, one can specialize $q$ at $\epsilon$ in the defining relations of the action on the basis $(v_m)$ of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. Moreover one checks that

$$\mathcal{U}_\epsilon(\mathrm{sl}_4^{\mathrm{tor}})' \cdot v_{Y_{1,1}Y_{1,-1-4L}Y_{0,2}^{-1}Y_{0,-4L}^{-1}} = \bigoplus_{m \in \mathscr{E}-\mathscr{E}'} \mathbb{C}v_m$$

is a sub-$\mathcal{U}_\epsilon(\mathrm{sl}_4^{\mathrm{tor}})'$-module of $V(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})_\epsilon$. By taking the quotient, we obtain a $\mathcal{U}_\epsilon(\mathrm{sl}_4^{\mathrm{tor}})'$-module

$$V_\epsilon = \bigoplus_{m \in \mathscr{E}'} \mathbb{C}v_m$$

which is irreducible: this is straightforward with the formulas of the action. □

## 6. Further possible developments and applications

In this last section, we give other promising directions to study the extremal loop weight modules for quantum toroidal algebras of general types. Moreover we give some possible applications of the results obtained in this article. This will be done in further papers.

In our construction of level 0 extremal loop weight modules in type $A$, monomial realizations of crystals and promotion operators on the finite crystals have a crucial role. Let us give some results which suggest that a similar construction is possible in other types. In [Hernandez and Nakajima 2006], an explicit description of monomial realizations of level 0 extremal fundamental weight crystals of quantum affine algebras is given for all the nonexceptional types. The automorphisms $z_\ell$ are determined in these cases. Furthermore in other types, there exists also symmetry properties for crystals arising from automorphisms of the associated Dynkin diagram (analogue of promotion operators in type $A$). Using that, a combinatorial process allows to obtain Kirillov–Reshetikhin crystals from crystals of finite type (see [Fourier et al. 2009; Kang et al. 1992; Okado and Schilling 2008]). These symmetry properties will be useful for a similar construction of extremal loop weight modules in other types.

As we have seen, the extremal fundamental loop weight modules

$$V(e^{\varpi_\ell} Y_{\ell,0} Y_{0,d_\ell}^{-1})$$

($n = 2r+1$ and $\ell = 1, r+1$ or $n$) are completely reducible as $\mathcal{U}_q^{v,0}(\mathrm{sl}_{n+1}^{\mathrm{tor}})$-modules: they are direct sums of fundamental modules of $\mathcal{U}_q(\hat{\mathrm{sl}}_{n+1})$. Similar vector spaces are considered in [Chari and Greenstein 2003] for the quantum affine algebra $\mathcal{U}_q(\hat{\mathfrak{g}})$ associated to a simple Lie algebra $\mathfrak{g}$ over $\mathbb{C}$. In fact for a finite-dimensional representation $V$ of $\mathcal{U}_q(\hat{\mathfrak{g}})'$, the vector space $V \otimes_{\mathbb{C}} \mathbb{C}[z, z^{-1}]$ is endowed with a structure of $\mathcal{U}_q(\hat{\mathfrak{g}})$-module by using the grading of this algebra. So the action is very different to the one defined in this article and we do not have a way to extend this action for the quantum toroidal algebra $\mathcal{U}_q(\mathfrak{g}^{\mathrm{tor}})$. But it would be interesting to study an analogous construction for the quantum toroidal algebra $\mathcal{U}_q(\mathfrak{g}^{\mathrm{tor}})$. We can expect to construct other examples of extremal loop weight modules by this process.

Let us explain another approach to construct extremal loop weight modules which could be fruitful. Let $\mathfrak{g}$ be a Kac–Moody algebra. For an integral weight $\lambda$, one defines

$$\lambda_+ = \sum_{\lambda(h_i) \geq 0} \lambda(h_i)\Lambda_i$$

and $\lambda_- = \lambda_+ - \lambda$. To study the extremal weight module $V(\lambda)$, Kashiwara [1994] considers the tensor product $V'(\lambda) = V(\lambda_+) \otimes V(\lambda_-)$ of the simple highest weight module $V(\lambda_+)$ and the simple lowest weight module $V(\lambda_-)$. By analogy, it would be interesting to define an action of the quantum affinization $\mathcal{U}_q(\hat{\mathfrak{g}})$ on the tensor product of simple $\ell$-highest weight modules and simple $\ell$-lowest weight modules, in the spirit of [Hernandez 2005; 2007; Feigin et al. 2011a; 2011b; 2012; 2013]. This will be studied in a further paper.

Another possible direction is to study the finite-dimensional representations of double affine Hecke algebras (or Cherednik algebras) at roots of unity obtained from the new finite-dimensional representations of $\mathcal{U}_\epsilon(\mathrm{sl}_{n+1}^{\mathrm{tor}})$ defined above, via Schur–Weyl duality [Varagnolo and Vasserot 1996].

In this article, we have defined promotion operators for the level 0 extremal fundamental weight crystals $\mathcal{B}(\varpi_\ell)$ in type $A_n$ ($n \geq 2$ odd, $1 \leq \ell \leq n$). It will be interesting to discuss the existence of promotion operators for other level 0 extremal weight crystals and the uniqueness of them in the spirit of [Shimozono 2002].

## Appendix

We describe here the monomial crystal

$$\bar{\mathcal{M}}(e^{2\varpi_1} Y_{1,1} Y_{1,-1} Y_{0,2}^{-1} Y_{0,0}^{-1}) = \bigoplus_{s \in \mathbb{N}} \mathcal{M}(e^{2\varpi_1 + s\delta} Y_{1,1} Y_{1,-1-4s} Y_{0,2}^{-1} Y_{0,-4s}^{-1}).$$

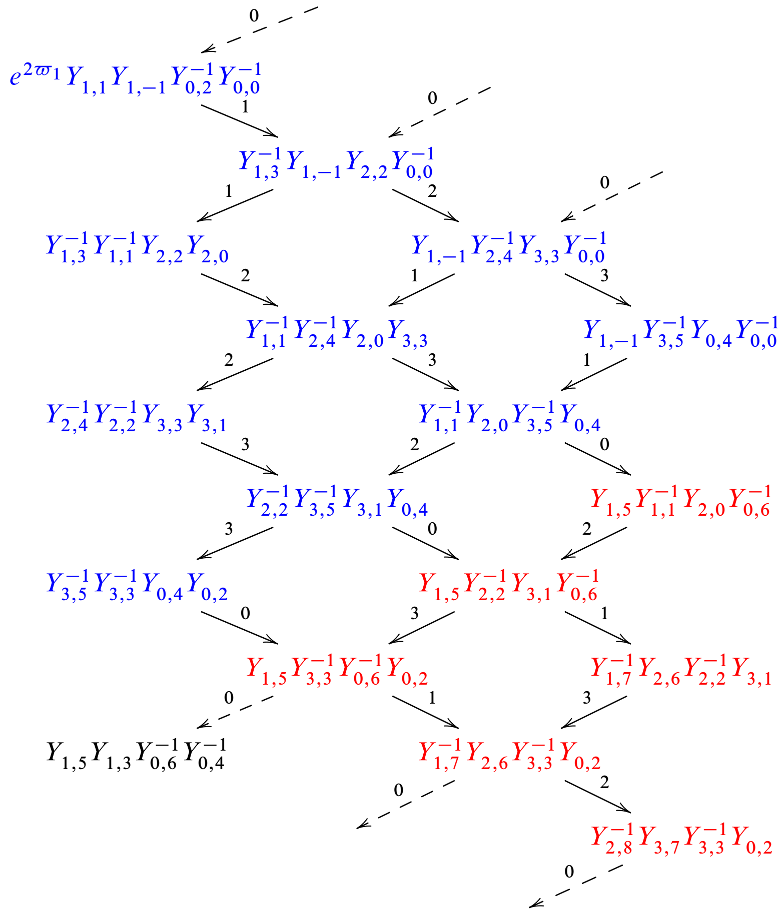



**Figure 1.** The $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-crystal $\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$.

More precisely, we show in Figures 1 and 2 the two connected components

$$\mathcal{M}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1}) \quad \text{and} \quad \mathcal{M}(e^{2\varpi_1+\delta}Y_{1,1}Y_{1,-5}Y_{0,2}^{-1}Y_{0,-4}^{-1})$$

of $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$. Recall that all the connected components of the latter are isomorphic modulo shift of weight by $\delta$. Furthermore the map $\tau_{4,-2\delta}$ is an automorphism of these crystals and we only give a part of them. The full crystals are obtained by applying the automorphism $\tau_{4,-2\delta}$. The sub-$I_0$-crystals

$$\mathcal{M}_{0,0,0}^{1} = \mathcal{M}_{I_0}(Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$$

and

$$\mathcal{M}_{0,0,1} = \mathcal{M}_{I_0}(Y_{1,1}Y_{1,-5}Y_{0,2}^{-1}Y_{0,-4}^{-1}) \oplus \mathcal{M}_{I_0}(Y_{1,1}^{-1}Y_{1,5}Y_{2,0}Y_{0,6}^{-1})$$

are explicitly given.

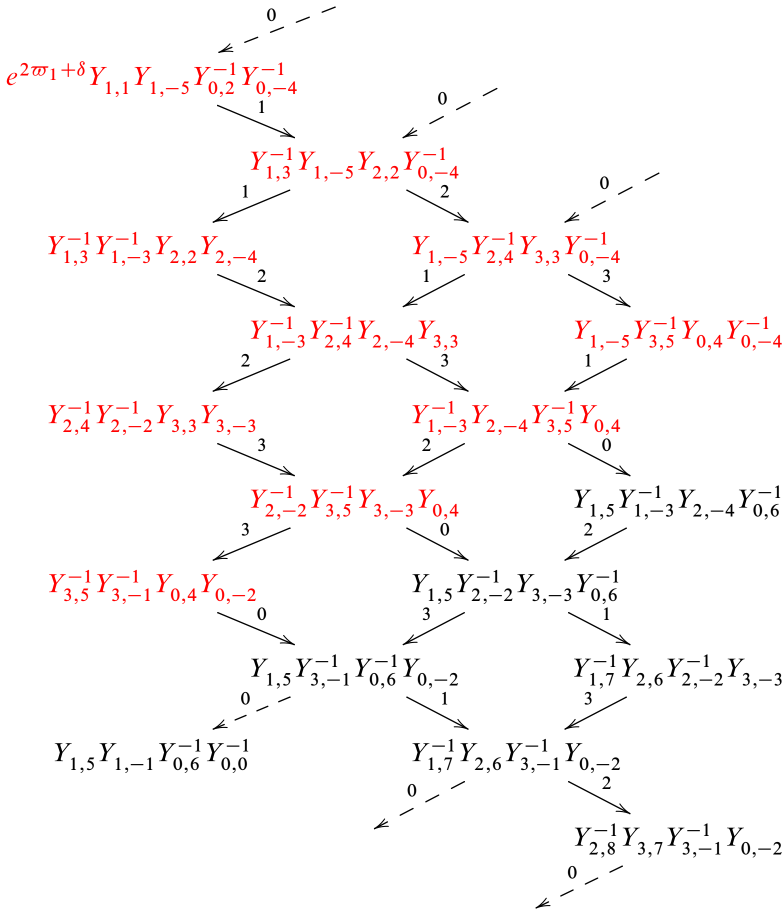



**Figure 2.** The $\mathcal{U}_q(\hat{\mathrm{sl}}_4)$-crystal $\mathcal{M}(e^{2\varpi_1+\delta}Y_{1,1}Y_{1,-5}Y_{0,2}^{-1}Y_{0,-4}^{-1})$.

Note that the $\theta$-twisted automorphism $\phi$ of $\bar{\mathcal{M}}(e^{2\varpi_1}Y_{1,1}Y_{1,-1}Y_{0,2}^{-1}Y_{0,0}^{-1})$ can be viewed as a descent of one diagonal in these crystals.

## Acknowledgements

I am grateful to my advisor, David Hernandez, for suggesting this problem to me, for his encouragement and precious comments. I would like to thank Anne Schilling and Cedric Lecouvey for their comments on the promotion operators and for pointing out some references to me. I want to thank Hiraku Nakajima for interesting questions about a first version of this work. I thank Alexandre Bouayad for accurate discussions and numerous observations, Xin Fang and Dragos Fratila for useful discussions. Finally the referee deserves thanks for careful reading and many precious comments.

## References

[Akasaka and Kashiwara 1997] T. Akasaka and M. Kashiwara, "Finite-dimensional representations of quantum affine algebras", *Publ. Res. Inst. Math. Sci.* **33**:5 (1997), 839–867. MR 99d:17017 Zbl 0915.17011

[Bandlow et al. 2010] J. Bandlow, A. Schilling, and N. M. Thiéry, "On the uniqueness of promotion operators on tensor products of type $A$ crystals", *J. Algebraic Combin.* **31**:2 (2010), 217–251. MR 2011i:17021 Zbl 1228.05286

[Beck 1994] J. Beck, "Braid group action and quantum affine algebras", *Comm. Math. Phys.* **165**:3 (1994), 555–568. MR 95i:17011 Zbl 0807.17013

[Beck 2002] J. Beck, "Crystal structure of level zero extremal weight modules", *Lett. Math. Phys.* **61**:3 (2002), 221–229. MR 2003k:17016 Zbl 1023.17005

[Beck and Nakajima 2004] J. Beck and H. Nakajima, "Crystal bases and two-sided cells of quantum affine algebras", *Duke Math. J.* **123**:2 (2004), 335–402. MR 2005e:17020 Zbl 1062.17006

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie, chapitres IV–VI*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. Translated as *Lie groups and Lie algebras*, Chapters 4–6, Springer, Berlin, 2002. MR 39 #1590 Zbl 0186.33001

[Chari and Greenstein 2003] V. Chari and J. Greenstein, "Quantum loop modules", *Represent. Theory* **7** (2003), 56–80. MR 2004b:17044 Zbl 1030.17010

[Chari and Pressley 1991] V. Chari and A. Pressley, "Quantum affine algebras", *Comm. Math. Phys.* **142**:2 (1991), 261–283. MR 93d:17017 Zbl 0739.17004

[Chari and Pressley 1994] V. Chari and A. Pressley, *A guide to quantum groups*, Cambridge University Press, Cambridge, 1994. MR 95j:17010 Zbl 0839.17009

[Chari and Pressley 1995] V. Chari and A. Pressley, "Quantum affine algebras and their representations", pp. 59–78 in *Representations of groups* (Banff, AB, 1994), edited by B. N. Allison and G. H. Cliff, CMS Conf. Proc. **16**, Amer. Math. Soc., Providence, RI, 1995. MR 96j:17009 Zbl 0855.17009

[Drinfeld 1987] V. G. Drinfeld, "Новая реализация янгианов и квантовых аффинных алгебр", *Dokl. Akad. Nauk SSSR* **269** (1987), 13–17. Translated as "A new realization of Yangians and quantized affine algebras" in *Sov. Math. Dokl.* **36**:2 (1988), 212–216. MR 88j:17020 Zbl 0667.16004

[Feigin et al. 2011a] B. Feigin, E. Feigin, M. Jimbo, T. Miwa, and E. Mukhin, "Quantum continuous $\mathfrak{gl}_\infty$: semiinfinite construction of representations", *Kyoto J. Math.* **51**:2 (2011), 337–364. MR 2012k:17024 Zbl 05907269

[Feigin et al. 2011b] B. Feigin, E. Feigin, M. Jimbo, T. Miwa, and E. Mukhin, "Quantum continuous $\mathfrak{gl}_\infty$: tensor products of Fock modules and $\mathcal{W}_n$-characters", *Kyoto J. Math.* **51**:2 (2011), 365–392. MR 2012c:17028 Zbl 05907270

[Feigin et al. 2012] B. Feigin, M. Jimbo, T. Miwa, and E. Mukhin, "Quantum toroidal $\mathfrak{gl}_1$-algebra: plane partitions", *Kyoto J. Math.* **52**:3 (2012), 621–659. MR 2959950 Zbl 06081386

[Feigin et al. 2013] B. Feigin, M. Jimbo, T. Miwa, and E. Mukhin, "Representations of quantum toroidal $\mathfrak{gl}_n$", *J. Algebra* **380** (2013), 78–108. MR 3023228

[Fourier et al. 2009] G. Fourier, M. Okado, and A. Schilling, "Kirillov–Reshetikhin crystals for nonexceptional types", *Adv. Math.* **222**:3 (2009), 1080–1116. MR 2010j:17028 Zbl 1239.17007

[Frenkel and Mukhin 2001] E. Frenkel and E. Mukhin, "Combinatorics of $q$-characters of finite-dimensional representations of quantum affine algebras", *Comm. Math. Phys.* **216**:1 (2001), 23–57. MR 2002c:17023 Zbl 1051.17013

[Frenkel and Mukhin 2002] E. Frenkel and E. Mukhin, "The $q$-characters at roots of unity", *Adv. Math.* **171**:1 (2002), 139–167. MR 2004a:17018 Zbl 1020.17009

[Frenkel and Reshetikhin 1999] E. Frenkel and N. Reshetikhin, "The $q$-characters of representations of quantum affine algebras and deformations of $\mathcal{W}$-algebras", pp. 163–205 in *Recent developments in quantum affine algebras and related topics* (Raleigh, NC, 1998), edited by N. Jing and K. C. Misra, Contemp. Math. **248**, Amer. Math. Soc., Providence, RI, 1999. MR 2002f:17022 Zbl 0973.17015

[Ginzburg et al. 1995] V. Ginzburg, M. Kapranov, and É. Vasserot, "Langlands reciprocity for algebraic surfaces", *Math. Res. Lett.* **2**:2 (1995), 147–160. MR 96f:11086 Zbl 0914.11040

[Hernandez 2005] D. Hernandez, "Representations of quantum affinizations and fusion product", *Transform. Groups* **10**:2 (2005), 163–200. MR 2006k:17025 Zbl 1102.17009

[Hernandez 2007] D. Hernandez, "Drinfeld coproduct, quantum fusion tensor category and applications", *Proc. Lond. Math. Soc.* (3) **95**:3 (2007), 567–608. MR 2008k:17017 Zbl 1133.17010

[Hernandez 2009] D. Hernandez, "Quantum toroidal algebras and their representations", *Selecta Math.* (*N.S.*) **14**:3-4 (2009), 701–725. MR 2010a:17027 Zbl 1222.17017

[Hernandez 2010] D. Hernandez, "Simple tensor products", *Invent. Math.* **181**:3 (2010), 649–675. MR 2011k:17027 Zbl 1221.17015

[Hernandez 2011] D. Hernandez, "The algebra $\mathcal{U}_q(\hat{\mathrm{sl}}_\infty)$ and applications", *J. Algebra* **329** (2011), 147–162. MR 2012h:17019 Zbl 05925840

[Hernandez and Leclerc 2010] D. Hernandez and B. Leclerc, "Cluster algebras and quantum affine algebras", *Duke Math. J.* **154**:2 (2010), 265–341. MR 2011g:17027 Zbl 05788166

[Hernandez and Nakajima 2006] D. Hernandez and H. Nakajima, "Level 0 monomial crystals", *Nagoya Math. J.* **184** (2006), 85–153. MR 2008e:17012 Zbl 1201.17008

[Jing 1998] N. Jing, "Quantum Kac–Moody algebras and vertex representations", *Lett. Math. Phys.* **44**:4 (1998), 261–271. MR 99j:17043 Zbl 0911.17006

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990. MR 92k:17038 Zbl 0716.17022

[Kang et al. 1992] S.-J. Kang, M. Kashiwara, K. C. Misra, T. Miwa, T. Nakashima, and A. Nakayashiki, "Perfect crystals of quantum affine Lie algebras", *Duke Math. J.* **68**:3 (1992), 499–607. MR 94j:17013 Zbl 0774.17017

[Kashiwara 1994] M. Kashiwara, "Crystal bases of modified quantized enveloping algebra", *Duke Math. J.* **73**:2 (1994), 383–413. MR 95c:17024 Zbl 0794.17009

[Kashiwara 2002a] M. Kashiwara, *Bases cristallines des groupes quantiques*, edited by C. Cochet, Cours Spécialisés **9**, Société Mathématique de France, Paris, 2002. MR 2004g:17013 Zbl 1066.17007

[Kashiwara 2002b] M. Kashiwara, "On level-zero representations of quantized affine algebras", *Duke Math. J.* **112**:1 (2002), 117–175. MR 2002m:17013 Zbl 1033.17017

[Kashiwara 2003] M. Kashiwara, "Realizations of crystals", pp. 133–139 in *Combinatorial and geometric representation theory* (Seoul, 2001), edited by S.-J. Kang and K.-H. Lee, Contemp. Math. **325**, Amer. Math. Soc., Providence, RI, 2003. MR 2004g:17014 Zbl 1066.17006

[Lusztig 1993] G. Lusztig, *Introduction to quantum groups*, Progress in Mathematics **110**, Birkhäuser, Boston, 1993. MR 94m:17016 Zbl 0788.17010

[Miki 1999] K. Miki, "Toroidal braid group action and an automorphism of toroidal algebra $U_q(\mathrm{sl}_{n+1,\mathrm{tor}})$ $(n \geq 2)$", *Lett. Math. Phys.* **47**:4 (1999), 365–378. MR 2000c:17027 Zbl 1022.17009

[Miki 2000] K. Miki, "Representations of quantum toroidal algebra $U_q(\mathrm{sl}_{n+1,\mathrm{tor}})$ $(n \geq 2)$", *J. Math. Phys.* **41**:10 (2000), 7079–7098. MR 2001j:17033 Zbl 1028.17011

[Naito and Sagaki 2003] S. Naito and D. Sagaki, "Path model for a level-zero extremal weight module over a quantum affine algebra", *Int. Math. Res. Not.* **2003**:32 (2003), 1731–1754. MR 2004c:17033 Zbl 1080.17008

[Naito and Sagaki 2006] S. Naito and D. Sagaki, "Path model for a level-zero extremal weight module over a quantum affine algebra, II", *Adv. Math.* **200**:1 (2006), 102–124. MR 2007a:17022 Zbl 1091.17009

[Nakajima 2001] H. Nakajima, "Quiver varieties and finite-dimensional representations of quantum affine algebras", *J. Amer. Math. Soc.* **14**:1 (2001), 145–238. MR 2002i:17023 Zbl 0981.17016

[Nakajima 2003] H. Nakajima, "$t$-analogs of $q$-characters of quantum affine algebras of type $A_n$, $D_n$", pp. 141–160 in *Combinatorial and geometric representation theory* (Seoul, 2001), edited by S.-J. Kang and K.-H. Lee, Contemp. Math. **325**, Amer. Math. Soc., Providence, RI, 2003. MR 2004h:17018 Zbl 1098.17013

[Nakajima 2004] H. Nakajima, "Extremal weight modules of quantum affine algebras", pp. 343–369 in *Representation theory of algebraic groups and quantum groups* (Tokyo, 2001), edited by T. Shoji et al., Adv. Stud. Pure Math. **40**, Math. Soc. Japan, Tokyo, 2004. MR 2005g:17036 Zbl 1088.17008 arXiv math/0204183

[Okado and Schilling 2008] M. Okado and A. Schilling, "Existence of Kirillov–Reshetikhin crystals for nonexceptional types", *Represent. Theory* **12** (2008), 186–207. MR 2009c:17022 Zbl 1243.17009

[Rosso 1991] M. Rosso, *Représentations des groupes quantiques*, pp. 443–483, Exp. 744, Astérisque **201-203**, Société Mathématique de France, Paris, 1991, Available at http://eudml.org/doc/110147. MR 93b:17050 Zbl 0752.17016

[Schilling 2008] A. Schilling, "Combinatorial structure of Kirillov–Reshetikhin crystals of type $D_n^{(1)}$, $B_n^{(1)}$, $A_{2n-1}^{(2)}$", *J. Algebra* **319**:7 (2008), 2938–2962. MR 2009e:17026 Zbl 1201.17010

[Shimozono 2002] M. Shimozono, "Affine type A crystal structure on tensor products of rectangles, Demazure characters, and nilpotent varieties", *J. Algebraic Combin.* **15**:2 (2002), 151–187. MR 2002m:17005 Zbl 1106.17305

[Varagnolo and Vasserot 1996] M. Varagnolo and E. Vasserot, "Schur duality in the toroidal setting", *Comm. Math. Phys.* **182**:2 (1996), 469–483. MR 98a:17024 Zbl 0879.17007

[Varagnolo and Vasserot 1998] M. Varagnolo and E. Vasserot, "Double-loop algebras and the Fock space", *Invent. Math.* **133**:1 (1998), 133–159. MR 99g:17035 Zbl 0904.17014

Received September 26, 2012. Revised April 26, 2013.

MATHIEU MANSUY
UNIVERSITÉ PARIS-DIDEROT PARIS 7
IMJ-PRG CNRS UMR 7586
BÂTIMENT SOPHIE GERMAIN
CASE 7012
75205 PARIS CEDEX 13
FRANCE

mansuy@math.jussieu.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 267, No. 1, 2014
dx.doi.org/10.2140/pjm.2014.267.243

# LEFSCHETZ FIBRATIONS WITH SMALL SLOPE

NAOYUKI MONDEN

**We construct Lefschetz fibrations over $S^2$ which do not satisfy the slope inequality. This disproves a conjecture of Hain.**

## 1. Introduction

Lefschetz fibrations have been an active area of research ever since the remarkable work in [Donaldson 1999] and [Gompf and Stipsicz 1999] revealed a close connection between them and symplectic 4-manifolds. In this paper, we consider the geography problem of Lefschetz fibrations over $S^2$, which derives from that of complex surfaces fibred over curves.

We are interested in two kinds of geography problems. Let $\sigma$ and $e$ be the signature and the Euler characteristic of a closed oriented smooth 4-manifold $X$, respectively. For an almost complex closed 4-manifold $X$, we set $K^2 := 3\sigma + 2e$ and $\chi_h := (\sigma + e)/4$ (the *holomorphic Euler characteristic*).

One is the geography problem for complex surfaces: the characterization of pairs $(K^2, \chi_h)$ corresponding to minimal complex surfaces. It is well known that any minimal complex surface of general type satisfies $K^2 > 0$, $\chi_h > 0$, the *Noether inequality* $2\chi_h - 6 \leq K^2$ and the *Bogomolov–Miyaoka–Yau inequality* $K^2 \leq 9\chi_h$ (see [Barth et al. 1984], for example). The above geography problem can be extended to the symplectic 4-manifolds. However, Fintushel and Stern [1998] constructed Lefschetz fibration which does not satisfy the Noether inequality. In particular, for most pairs $(p, q)$ satisfying $p < 2q - 6$, there exists a minimal symplectic 4-manifold with $p = K^2$ and $q = \chi_h$ (see [Gompf and Stipsicz 1999]). On the other hand, no examples of a minimal symplectic 4-manifold with $K^2 > 9\chi_h$ have been found yet.

The other is the geography problem for complex surfaces fibred over curves. Hereafter, we assume $g \geq 2$. Let $f : S \to C$ be a relatively minimal holomorphic genus-$g$ fibration, where $S$ is a complex surface and $C$ is a complex curve of genus $k$. We define relative numerical invariants $\chi_f := \chi_h - (g-1)(k-1)$ and $K_f^2 := K^2 - 8(g-1)(k-1)$ for $f : S \to C$. Then, we have two inequalities $\chi_f \geq 0$

*MSC2010:* primary 57R20; secondary 57N13, 14D06, 20F38.
*Keywords:* Lefschetz fibration, mapping class group, slope.

and $K_f^2 \geq 0$, known as *Beauville's inequality* (see [Beauville 1979]) and *Arakelov's inequality* (see [Arakelov 1971]), respectively. For $\chi_f \neq 0$, which is equivalent to the fact that $f$ is not a holomorphic bundle, we define $\lambda_f$ to be the quotient $K_f^2/\chi_f$, called the slope of $f$. Xiao [1987] proved that $4-4/g \leq \lambda_f \leq 12$ (that is, $(4-4/g)\chi_f \leq K_f^2 \leq 12\chi_f$). The former inequality is called the *slope inequality*. For a relatively minimal genus-$g$ Lefschetz fibration, $\chi_f$, $K_f^2$ and the slope $\lambda_f$ are defined in the same way as for complex surfaces fibred over curves. To the author's knowledge, the slope of all known Lefschetz fibrations over $S^2$ is greater than or equal to $4-4/g$.

**Conjecture 1.1** (Hain; see [Amorós et al. 2000, Question 5.10; Endo and Nagami 2005, Conjecture 4.12]). *For every relatively minimal genus-g Lefschetz fibration $f : X \to S^2$, the slope inequality $\lambda_f \geq 4-4/g$ holds.*

In this paper, we give a negative answer to Conjecture 1.1.

**Theorem 3.1.** *For each $g \geq 3$, there exists a genus-g Lefschetz fibration over $S^2$ with slope $\lambda_f = 4-4/g-1/3g$ whose total space is simply connected.*

Moreover, by fiber sum operations, we have the following results:

**Corollary 3.6.** *For each $g \geq 3$, $m \geq 0$ and $l \geq 0$, there exists a genus-g Lefschetz fibration $f_{m,l} : X_{m,l} \to S^2$ with slope $\lambda_{f_{m,l}} = 4-4/g-1/(m+3)g$ such that $\pi_1(X_{m,l}) = 1$. Moreover, if $(m,l) \neq (0,0)$, then $X_{m,l}$ is a minimal symplectic 4-manifold.*

**Corollary 3.7.** *For each $g \geq 3$, $m \geq 1$ and $l \geq 0$, there exists a genus-g Lefschetz fibration $f'_{m,l} : Y_{m,l} \to S^2$ with slope $\lambda_{f'_{m,l}} = 4-4/g-1/2g+1/(2\cdot 3^{m-1}g)$ such that $\pi_1(Y_{m,l}) = 1$. Moreover, if $l \geq 1$, then $Y_{m,l}$ is a minimal symplectic 4-manifold.*

As a consequence, we have the following results.

**Corollary 4.2.** *The Lefschetz fibrations $f_{m,l}$ $(m \geq 0)$ and $f'_{m,l}$ $(m \geq 2)$ are non-holomorphic.*

Let $f : X \to S^2$ be a relatively minimal genus-$g$ Lefschetz fibration with $n > 0$ singular fibers. From $e(X) = -4(g-1)+n$ and results from [Smith 1999; Stipsicz 1999; Ozbagci 2002], we have $\chi_f > 0$, $K_f^2 \geq 4g-4$ and $\lambda_f \leq 10$. Moreover, it is well known that any hyperelliptic Lefschetz fibration satisfies the slope inequality. This fact follows from the signature formula for genus-$g$ hyperelliptic Lefschetz fibrations obtained by Matsumoto [1983; 1996] for $g = 1, 2$ and Endo [2000] for $g \geq 3$. Therefore, genus-2 Lefschetz fibrations satisfy the slope inequality. In particular, if $f$ is a hyperelliptic Lefschetz fibration with only nonseparating vanishing cycles, then $\lambda_f$ is equal to $4-4/g$. For Lefschetz fibrations with $b_2^+ = 1$, we prove the following result.

**Theorem 5.1.** *Let $g \geq 2$ and let $f : X \to S^2$ be a genus-$g$ Lefschetz fibration with $b_2^+(X) = 1$.*

(1) *If $X$ is not diffeomorphic to the blow-up of a ruled surface, then*

(i) $$4 - 4/g \leq \lambda_f \leq 8 \quad \textit{for } b_1(X) = 0,$$

(ii) $$4 \leq \lambda_f \leq 8 \quad \textit{for } b_1(X) = 2.$$

(2) *If $X$ is diffeomorphic to the blow-up of an $S^2$-bundle over $\Sigma_k$, then*

$$4 + 4(k-1)/(g-k) \leq \lambda_f \leq 8,$$

*and the lower bound is sharp.*

The study of the slope of holomorphic fibrations was mainly motivated by Severi's inequality, which states that if $S$ is a minimal surface of general type of maximal Albanese dimension, then $K^2 \geq 4\chi_h$. Equivalently, if $K^2 < 4\chi_h$, then $S$ is a surface fibred over $C$ of genus $b_1(S)/2$. Severi [1932] claimed it, but his proof was not correct (see [Catanese 1983]). The inequality was independently posed as a conjecture by Reid [1979] and by Catanese [1983]. Xiao [1987] proved the conjecture when $S$ is a surface fibred over a curve of positive genus. He showed that if $S$ admits a holomorphic genus-$g$ fibration $f$ over $C$ of positive genus $k$ with $K^2 < 4\chi_h + 4(g-1)(k-1)$ (that is, $\lambda_f < 4$), then $k = b_1(S)/2$. Konno [1996] proved it in the case of *even* surfaces. The conjecture was solved by Manetti [2003] when $S$ has ample canonical bundle. Pardini [2005] proved the conjecture completely by using the slope inequality for holomorphic fibrations over $\mathbb{CP}^1$.

In Section 2, we review some standard facts on Lefschetz fibrations. Our main results are proved in Section 3. We give Lefschetz fibrations which violate the slope inequality. Consequently, we obtain examples of nonholomorphic Lefschetz fibrations in Section 4. In the last section, we investigate the slopes of Lefschetz fibrations with $b_2^+ = 1$.

**Remark 1.2.** The slope inequality of Conjecture 1.1 can be reformulated in terms of the Deligne–Mumford compactified moduli space of stable curves of genus $g$, denoted by $\overline{\mathcal{M}}_g$, as follows. For a relatively minimal genus-$g$ Lefschetz fibration $f : X \to S^2$ with $n$ singular fibers, we obtain a symplectic structure on $X$ such that for all $x \in S^2$, $f^{-1}(x)$ is a pseudoholomorphic curve. Since a 2-dimensional almost-complex structure is integrable, $f^{-1}(x)$ determines a point in $\overline{\mathcal{M}}_g$. Thus, we obtain the moduli map $\phi_f : S^2 \to \overline{\mathcal{M}}_g$ which is defined by $\phi_f(x) = [f^{-1}(x)] \in \overline{\mathcal{M}}_g$ for $x \in S^2$. We denote by $\mathcal{H}_g$ the Hodge bundle on $\overline{\mathcal{M}}_g$ with fiber the determinant line $\bigwedge^g H^0(C; K_C)$, where $C$ is the set of critical points of $f$. By Smith's signature formula [1999] and the slope inequality, we have the following inequality:

$$(8g+4)\langle c_1(\mathcal{H}_g), [\phi_f(S^2)]\rangle - g \cdot n \geq 0.$$

## 2. Preliminaries

In this section, we first recall the definition and basic properties of Lefschetz fibrations. More details can be found in [Gompf and Stipsicz 1999].

Let $\Sigma_g$ be a closed oriented surface of genus $g \geq 2$ and let $\Gamma_g$ be the *mapping class group* of $\Sigma_g$, which is the group of isotopy classes of orientation-preserving diffeomorphisms of $\Sigma_g$. We denote by $t_c$ the right-handed *Dehn twist* about a simple closed curve $c$ on an oriented surface. The notation $t_c t_d$ means that we first apply $t_d$ then $t_c$.

**Definition 2.1.** Let $X$ be a closed, oriented smooth 4-manifold. A smooth map $f : X \to S^2$ is a genus-$g$ *Lefschetz fibration* if it satisfies the following conditions:

(i) $f$ has finitely many critical values $b_1, \dots, b_n \in S^2$, and $f$ is a smooth $\Sigma_g$-bundle over $S^2 - \{b_1, \dots, b_n\}$.

(ii) For each $i$ $(i = 1, \dots, n)$, there exists a unique critical point $p_i$ in the *singular fiber* $f^{-1}(b_i)$ such that about each $p_i$ and $b_i$ there are local complex coordinate charts agreeing with the orientations of $X$ and $S^2$ on which $f$ is of the form $f(z_1, z_2) = z_1^2 + z_2^2$.

(iii) $f$ is relatively minimal (no fiber contains a $(-1)$-sphere).

Each singular fiber is obtained by collapsing a simple closed curve (the *vanishing cycle*) in the regular fiber. The monodromy of the fibration around a singular fiber is given by a right-handed Dehn twist along the corresponding vanishing cycle. A Lefschetz fibration $f : X \to S^2$ is *holomorphic* if there are complex structures on both $X$ and $S^2$ with holomorphic projection $f$.

Once we fix an identification of $\Sigma_g$ with the fiber over a base point of $S^2$, we can characterize the Lefschetz fibration $f : X \to S^2$ by its *monodromy representation* $\pi_1(S^2 - \{b_1, \dots, b_n\}) \to \Gamma_g$. This map is really an antihomomorphism, since elements of $\pi_1(S^2 - \{b_1, \dots, b_n\})$ are written left-to-right and elements of $\Gamma_g$ are written right-to-left. Let $\gamma_1, \dots, \gamma_n$ be an ordered system of generating loops for $\pi_1(S^2 - \{b_1, \dots, b_n\})$, such that each $\gamma_i$ encircles only $b_i$ and $\prod \gamma_i$ is homotopically trivial. Thus, the monodromy of $f$ comprises a factorization

$$t_{v_n} \dots t_{v_2} t_{v_1} = 1 \in \Gamma_g,$$

where $v_i$ are vanishing cycles of the singular fibers. This factorization is called the *positive relator*.

According to theorems of Kas [1980] and Matsumoto [1996], if $g \geq 2$, then the isomorphism class of a Lefschetz fibration is determined by a positive relator modulo simultaneous conjugations

$$t_{v_n} \dots t_{v_2} t_{v_1} \sim t_{\phi(v_n)} \dots t_{\phi(v_2)} t_{\phi(v_1)} \quad \text{for all } \phi \in \Gamma_g$$

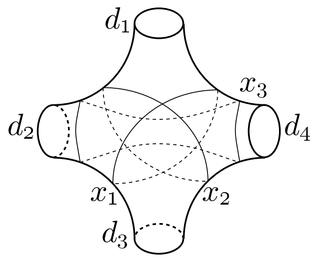



**Figure 1.** The curves $d_1, d_2, d_3, d_4, x_1, x_2, x_3$.

and elementary transformations

$$t_{v_n} \dots t_{v_{i+2}} t_{v_{i+1}} t_{v_i} t_{v_{i-1}} t_{v_{i-2}} \dots t_{v_1} \sim t_{v_n} \dots t_{v_{i+2}} t_{v_i} t_{t_{v_i}^{-1}(v_{i+1})} t_{v_{i-1}} t_{v_{i-2}} \dots t_{v_1},$$

$$t_{v_n} \dots t_{v_{i+2}} t_{v_{i+1}} t_{v_i} t_{v_{i-1}} t_{v_{i-2}} \dots t_{v_1} \sim t_{v_n} \dots t_{v_{i+2}} t_{v_{i+1}} t_{t_{v_i}(v_{i-1})} t_{v_i} t_{v_{i-2}} \dots t_{v_1}.$$

Note that $\phi t_{v_i} \phi^{-1} = t_{\phi(v_i)}$. For all $\phi \in \Gamma_g$, let $\phi(\varrho)$ be the positive relator which is obtained by applying simultaneous conjugations by $\phi$ to a positive relator $\varrho$. We denote a Lefschetz fibration associated to a positive relator $\varrho \in \Gamma_g$ by $f_\varrho : X_\varrho \to S^2$. Clearly, if $\varrho_1 \sim \varrho_2$ in $\Gamma_g$ (that is, $\varrho_2$ is obtained by applying elementary transformations or simultaneous conjugations to $\varrho_1$), then

$$\chi_{f_{\varrho_1}} = \chi_{f_{\varrho_2}} \quad \text{and} \quad K^2_{f_{\varrho_1}} = K^2_{f_{\varrho_2}}.$$

For positive relators $\varrho_1$ and $\varrho_2$ in $\Gamma_g$, the genus-$g$ Lefschetz fibration

$$f_{\varrho_1\varrho_2} : X_{\varrho_1\varrho_2} \to S^2$$

is the (trivial) fiber sum of $f_{\varrho_1}$ and $f_{\varrho_2}$. Since $\sigma(X_{\varrho_1\varrho_2}) = \sigma(X_{\varrho_1}) + \sigma(X_{\varrho_2})$ and $e(X_{\varrho_1\varrho_2}) = e(X_{\varrho_1}) + e(X_{\varrho_2}) + 4(g-1)$, we see that $\chi_{f_{\varrho_1\varrho_2}} = \chi_{f_{\varrho_1}} + \chi_{f_{\varrho_2}}$ and $K^2_{f_{\varrho_1\varrho_2}} = K^2_{f_{\varrho_1}} + K^2_{f_{\varrho_2}}$. In particular, if $\varrho_1 \sim \varrho_2$, then

$$\chi_{f_{\varrho_1\varrho_2}} = 2\chi_{f_{\varrho_1}} = 2\chi_{f_{\varrho_2}} \quad \text{and} \quad K^2_{f_{\varrho_1\varrho_2}} = 2K^2_{f_{\varrho_1}} = 2K^2_{f_{\varrho_2}}.$$

We next begin with a definition of the lantern relation (see [Dehn 1938; Johnson 1979]).

**Definition 2.2.** Let $\Sigma_0^4$ denote a sphere with 4 boundary components. Let $d_1$, $d_2$, $d_3$, $d_4$ be the 4 boundary curves of $\Sigma_0^4$ and let $x_1$, $x_2$, $x_3$ be the interior curves as shown in Figure 1. Then, we have the *lantern relation*

$$t_{d_1} t_{d_2} t_{d_3} t_{d_4} = t_{x_1} t_{x_2} t_{x_3}.$$

Let $\varrho$ be a positive relator of $\Gamma_g$. Let $d_1$, $d_2$, $d_3$, $d_4$, $x_1$, $x_2$, $x_3$ be curves as in Definition 2.2. Suppose that $\varrho$ includes $t_{d_1} t_{d_2} t_{d_3} t_{d_4}$ as a subword:

$$\varrho = U \cdot t_{d_1} t_{d_2} t_{d_3} t_{d_4} \cdot V,$$

where $U$ and $V$ are products of right-handed Dehn twists. Then, by the lantern relation, the product of right-handed Dehn twists

$$\varrho' = U \cdot t_{x_1} t_{x_2} t_{x_3} \cdot V$$

is also a positive relator of $\Gamma_g$.

This operation is one of substitution techniques introduced by Fuller.

**Definition 2.3.** We say that $\varrho'$ is obtained by applying an *$L$-substitution* to $\varrho$. Conversely, $\varrho$ is said to be obtained by applying an *$L^{-1}$-substitution* to $\varrho'$. We also call these two kinds of operations *lantern substitutions*.

**Proposition 2.4** [Endo and Nagami 2005, Theorem 4.3 and Proposition 3.12]. *Let $\varrho$, $\varrho'$ be positive relators of $\Gamma_g$ and let $X_\varrho$, $X_{\varrho'}$ be the corresponding Lefschetz fibrations over $S^2$, respectively. Suppose that $\varrho$ is obtained by applying an $L^{-1}$-substitution to $\varrho'$. Then, $\sigma(X_\varrho) = \sigma(X_{\varrho'}) - 1$ and $e(X_\varrho) = e(X_{\varrho'}) + 1$. Therefore,*

$$\chi_{f_\varrho} = \chi_{f_{\varrho'}} \quad \textit{and} \quad K^2_{f_\varrho} = K^2_{f_{\varrho'}} - 1.$$

**Remark 2.5.** Endo and Gurtas [2010] showed that $X_{\varrho'}$ is a rational blowdown of $X_\varrho$ introduced by Fintushel and Stern [1997]. Such relations were also generalized by Endo, Mark, and Van Horn-Morris [Endo et al. 2011].

## 3. Main results

In this section, we give a negative answer to Conjecture 1.1.

**Theorem 3.1.** *For each $g \geq 3$, there exists a genus-$g$ Lefschetz fibration over $S^2$ with slope $\lambda_f = 4 - 4/g - 1/3g$ whose total space is simply connected.*

In order to prove Theorem 3.1, we recall some standard facts on hyperelliptic Lefschetz fibrations. Let $\Delta_g$ be the *hyperelliptic mapping class group* of genus $g$, that is, the subgroup of $\Gamma_g$ which consists of all isotopy classes of orientation-preserving diffeomorphisms of $\Sigma_g$ commuting with the isotopy class of $\iota$, called the hyperelliptic involution. Note that $\Delta_g = \Gamma_g$ for $g = 1, 2$.

**Definition 3.2.** Let $\varrho = t_{a_1} \ldots t_{a_n}$ be a positive relator in $\Gamma_g$. A genus-$g$ Lefschetz fibration $f_\varrho : X_\varrho \to S^2$ is called *hyperelliptic* if for each $k \in \{1, \ldots, n\}$, $t_{a_k}$ is in $\Delta_g$. Equivalently, $\iota(a_k) = a_k$ for each $k$.

The following theorem was established in [Matsumoto 1983; 1996] for $g = 1, 2$ and in [Endo 2000] for $g \geq 3$.

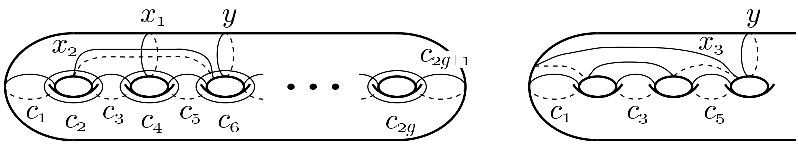



**Figure 2.** The curves $c_1, \ldots, c_{2g+1}, x_1, x_2, x_3, y$.

**Theorem 3.3** (Matsumoto, Endo). *Let $f_\varrho : X_\varrho \to S^2$ be a genus-g hyperelliptic Lefschetz fibration with m nonseparating and*

$$s = \sum_{h=1}^{[g/2]} s_h$$

*separating vanishing cycles, where $s_h$ denotes the number of separating vanishing cycles that separate $\Sigma_g$ into two surfaces, one of which has genus h. Then, we have*

$$\sigma(X_\varrho) = -\frac{g+1}{2g+1}m + \sum_{h=1}^{[g/2]}\left(\frac{4h(g-h)}{2g+1} - 1\right)s_h.$$

We need the following positive relator to prove Theorem 3.1. As shown in Figure 2, let $c_1, c_2, \ldots, c_{2g+1}$ be the curves in $\Sigma_g$. We denote by $h_g$ ($\in \Gamma_g$) the product of $8g+4$ right-handed Dehn twists

$$h_g := (t_{c_1} t_{c_2} \ldots t_{c_{2g+1}}^2 \ldots t_{c_2} t_{c_1})^2.$$

It is well known that $h_g$ is a positive relator in $\Delta_g$ and that $\sigma(X_{h_g}) = -4(g+1)$, by Theorem 3.3 and $e(X_{h_g}) = 4(g+2)$. This gives $\chi_{f_{h_g}} = g$, $K^2_{f_{h_g}} = 4g-4$ and $\lambda_{f_{h_g}} = 4 - 4/g$ (that is, $f_{h_g}$ is lying on the slope line).

*Proof of Theorem 3.1.* Suppose $g \geq 3$. Let $x_1, x_2, x_3, y$ be the curves as shown in Figure 2. Since $c_1, x_i$ are nonseparating curves, there exists a diffeomorphism $\phi_i$ such that $\phi_i(c_1) = x_i$. Hence, we have the following positive relator $r_i$ ($i = 1, 2, 3$):

$$\begin{aligned} r_i = \phi_i h_g \phi_i^{-1} &= \phi_i (t_{c_1} t_{c_2} \ldots t_{c_{2g+1}}^2 \ldots t_{c_2} t_{c_1})^2 \phi_i^{-1} \\ &= (t_{\phi_i(c_1)} t_{\phi_i(c_2)} \ldots t_{\phi_i(c_{2g+1})}^2 \ldots t_{\phi_i(c_2)} t_{\phi_i(c_1)})^2 \\ &= (t_{x_i} t_{\phi_i(c_2)} \ldots t_{\phi_i(c_{2g+1})}^2 \ldots t_{\phi_i(c_2)} t_{\phi_i(c_1)})^2 \\ &= 1 \in \Gamma_g. \end{aligned}$$

Let $r'_g = r_1 r_2 r_3 = (t_{x_1} \ldots t_{\phi_1(c_1)})^2 (t_{x_2} \ldots)^2 (t_{x_3} \ldots)^2$. Since $f_{r'_g}$ is the fiber sum of $f_{r_1}$, $f_{r_2}$ and $f_{r_3}$ which are obtained by applying simultaneous conjugations to $h_g$, we have

$$\chi_{f_{r'_g}} = 3\chi_{f_{h_g}} = 3g \quad \text{and} \quad K^2_{f_{r'_g}} = 3K^2_{f_{h_g}} = 3(4g-4).$$

We apply elementary transformations to $r'_g$ as follows:

$$
\begin{aligned}
r'_g &= r_1 r_2 r_3 \\
&= t_{x_1} t_{\phi_1(c_2)} \cdots t_{\phi_1(c_2)} \underline{t_{\phi_1(c_1)} \cdot t_{x_2}} t_{\phi_2(c_2)} \cdots t_{\phi_2(c_1)} \cdot t_{x_3} t_{\phi_3(c_2)} \cdots t_{\phi_3(c_1)} \\
&\sim t_{x_1} t_{\phi_1(c_2)} \cdots t_{\phi_1(c_2)} \underline{t_{x_2} t_{t_{x_2}^{-1}(\phi_1(c_1))}} t_{\phi_2(c_2)} \cdots t_{\phi_2(c_1)} \cdot t_{x_3} t_{\phi_3(c_2)} \cdots t_{\phi_3(c_1)} \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
&\sim t_{x_1} t_{x_2} t_{t_{x_2}^{-1}(\phi_1(c_2))} \cdots t_{t_{x_2}^{-1}(\phi_1(c_2))} t_{t_{x_2}^{-1}(\phi_1(c_1))} t_{\phi_2(c_2)} \cdots \underline{t_{\phi_2(c_1)} \cdot t_{x_3}} t_{\phi_3(c_2)} \cdots t_{\phi_3(c_1)} \\
&\sim t_{x_1} t_{x_2} t_{t_{x_2}^{-1}(\phi_1(c_2))} \cdots t_{t_{x_2}^{-1}(\phi_1(c_2))} t_{t_{x_2}^{-1}(\phi_1(c_1))} t_{\phi_2(c_2)} \cdots \underline{t_{x_3} t_{t_{x_3}^{-1}(\phi_2(c_1))}} t_{\phi_3(c_2)} \cdots t_{\phi_3(c_1)} \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
&\sim (t_{x_1} t_{x_2} t_{x_3}) W,
\end{aligned}
$$

where $W$ is a product of $24g+9$ right-handed Dehn twists. By the lantern relation, we get the following positive relator $r_g$:

$$r_g := (t_{c_1} t_{c_3} t_{c_5} t_y) W.$$

Since $r_g$ is obtained by applying an $L^{-1}$-substitution to $r'_g$, by Proposition 2.4

$$\chi_{f_{r_g}} = 3g \quad \text{and} \quad K^2_{f_{r_g}} = 3(4g-4) - 1.$$

Then, the slope of $f_{r_g}$ is equal to $4 - 4/g - 1/3g$.

Since it is easy to check that $r_g$ includes the Dehn twist about a curve $\phi_3(c_i)$ for $1 \le i \le 2g+1$, $\pi_1(X_{r_g}) = 1$. This follows from [Gompf and Stipsicz 1999] and the fact that $f_{r_g}$ has a section. This completes the proof of Theorem 3.1. □

**Remark 3.4.** Since $r_g$ is obtained by applying an $L^{-1}$-substitution to $r'_g$, $X_{r_g}$ is a *rational blow-up* of $X_{r'_g}$. By applying elementary transformations to a relator corresponding to a Lefschetz fibration which is obtained by taking a twisted fiber sum with sufficiently many Lefschetz fibrations, we obtain a positive relator such that we can apply a monodromy substitution, which corresponds to the operation of rational blowdown (resp. rational blow-up) in [Endo et al. 2011], to it.

**Remark 3.5.** Miyachi and Shiga [2011] produced genus-$g$ Lefschetz fibrations over $\Sigma_{2m}$ which do not satisfy the slope inequality.

Moreover, by fiber sum operations, we have the following results:

**Corollary 3.6.** *For each $g \ge 3$, $m \ge 0$ and $l \ge 0$, there exists a genus-$g$ Lefschetz fibration $f_{m,l} : X_{m,l} \to S^2$ with slope $\lambda_{f_{m,l}} = 4 - 4/g - 1/(m+3)g$ such that $\pi_1(X_{m,l}) = 1$. Moreover, if $(m,l) \ne (0,0)$, then $X_{m,l}$ is a minimal symplectic 4-manifold.*

*Proof of Corollary 3.6.* For any $m \geq 0$, we consider the Lefschetz fibration $f_{r_g h_g^m} : X_{r_g h_g^m} \to S^2$ which is the fiber sum of $f_{r_g}$ and $m$ copies of $f_{h_g}$. Then,

$$\chi_{f_{r_g h_g^m}} = \chi_{f_{r_g}} + m\chi_{f_{h_g}} = (3+m)g,$$
$$K^2_{f_{r_g h_g^m}} = K^2_{f_{r_g}} + mK^2_{f_{h_g}} = (3+m)(4g-4)-1.$$

Therefore, we obtain

$$\lambda_{f_{r_g h_g^m}} = 4 - 4/g - 1/(m+3)g. \tag{1}$$

Let $f_{m,l} : X_{m,l} \to S^2$ be the fiber sum of $l$ copies of $f_{r_g h_g^m}$ (that is, $f_{m,l} = f_{(r_g h_g^m)^l}$). By Using (1) and an argument similar to the proof of Theorem 3.1, we have $\lambda_{f_{m,l}} = 4 - 4/g - 1/(m+3)g$ and $\pi_1(X_{m,l}) = 1$. By the result of Usher [2006], $X_{m,l}$ is minimal for $(m,l) \neq (0,0)$. This completes the proof. □

**Corollary 3.7.** *For each $g \geq 3$, $m \geq 1$ and $l \geq 0$, there exists a genus-$g$ Lefschetz fibration $f'_{m,l} : Y_{m,l} \to S^2$ with slope $\lambda_{f'_{m,l}} = 4 - 4/g - 1/2g + 1/(2 \cdot 3^{m-1}g)$ such that $\pi_1(Y_{m,l}) = 1$. Moreover, if $l \geq 1$, then $Y_{m,l}$ is a minimal symplectic 4-manifold.*

*Proof of Corollary 3.7.* When we apply the argument of Theorem 3.1 again, with $\varrho_1 = h_g$ replaced by $\varrho_2 = r_g$, we obtain a genus-$g$ Lefschetz fibration $f_{\varrho_3} : X_{\varrho_3} \to S^2$ with

$$\chi_{f_{\varrho_3}} = 3\chi_{f_{\varrho_2}} = 3 \cdot 3\chi_{f_{\varrho_1}}$$
$$K^2_{f_{\varrho_3}} = 3K^2_{f_{\varrho_2}} - 1 = 3(3K^2_{f_{\varrho_1}} - 1) - 1.$$

By repeating this argument, we get a genus-$g$ Lefschetz fibration $f_{\varrho_m}$ $(m \geq 1)$ with

$$\begin{aligned}
\chi_{f_{\varrho_m}} &= 3^{m-1}\chi_{f_{\varrho_1}} = 3^{m-1}g, \\
K^2_{f_{\varrho_m}} &= 3\big(\ldots(3(3K^2_{f_{\varrho_1}} - 1) - 1)\ldots\big) - 1 = 3^{m-1}K^2_{f_{\varrho_1}} - 3^{m-2} - \cdots - 3 - 1 \\
&= 3^{m-1}(4g-4) - (3^{m-1}-1)/2.
\end{aligned}$$

Therefore, $\lambda_{f_{\varrho_m}} = 4 - 4/g - 1/2g + 1/(2 \cdot 3^{m-1}g)$.

Let $f'_{m,l} : Y_{m,l} \to S^2$ be the fiber sum of $l$ copies of $f_{\varrho_m}$, and so $\lambda_{f'_{m,l}} = 4 - 4/g - 1/2g + 1/(2 \cdot 3^{m-1}g)$. Similar to the proof of Corollary 3.6, we see that $\pi_1(Y_{m,l}) = 1$ and that $Y_{m,l}$ is minimal for $l \geq 1$. □

## 4. Nonholomorphic Lefschetz fibrations

There are various kinds of nonholomorphic Lefschetz fibrations. By fiber summing two copies of genus-2 Lefschetz fibration due to Matsumoto [1996], Ozbagci and Stipsicz [2000] constructed nonholomorphic genus-2 Lefschetz fibrations whose total space does not admit a complex structure. Korkmaz [2001] generalized their examples to $g \geq 3$. The above mentioned examples of Fintushel and Stern are also nonholomorphic Lefschetz fibrations. From the study of divisors in moduli space, Smith [2001] showed that a genus-3 Lefschetz fibration over $S^2$ which was produced by Fuller is nonholomorphic. Endo and Nagami [2005] constructed some examples of nonholomorphic Lefschetz fibrations which violate lower bounds of the slope for nonhyperelliptic fibrations of genus 3, 4 and 5 from the results of Konno [1991; 1993] and Chen [1993]. Hirose [2010] also gave some examples of $g = 3, 4$. In this section, we give new examples of nonholomorphic Lefschetz fibrations.

From the slope inequality for holomorphic fibrations, we have the following necessary condition for a Lefschetz fibration to be holomorphic:

**Proposition 4.1** [Xiao 1987]. *If a Lefschetz fibration $f$ is holomorphic, then the slope inequality $\lambda_f \geq 4 - 4/g$ holds.*

As a consequence, we have the following results.

**Corollary 4.2.** *The Lefschetz fibrations $f_{m,l}$ $(m \geq 0)$ and $f'_{m,l}$ $(m \geq 2)$ are nonholomorphic.*

**Remark 4.3.** The above mentioned examples of Fintushel and Stern satisfy the slope inequality but violate the Noether inequality. On the other hand, $f_{m,l}$ and $f'_{m,l}$ satisfy the Noether inequality but violate the slope inequality. Therefore, these two necessary conditions for a Lefschetz fibration to be holomorphic are independent in the sense that neither one implies the other.

## 5. The slopes of Lefschetz fibrations with $b_2^+ = 1$

We have the following natural question: Which Lefschetz fibrations satisfy the slope inequality? By Proposition 4.1, holomorphic Lefschetz fibrations satisfy the slope inequality. If a Lefschetz fibration is hyperelliptic, then $\lambda_f \geq 4 - 4/g$. This fact can proved as follows. In the notation of Theorem 3.3, we have $e(X_\varrho) = -4(g-1) + (m+s)$. Then, since $h \in \{1, \ldots, [g/2]\}$ and $g \geq 2$, we have

$$K_{f_\varrho}^2 - (4 - 4/g)\chi_{f_\varrho} = \sum_{h=1}^{[g/2]} \frac{4h(g-h)-g}{g} s_h \geq 0.$$

In particular, this means that for any hyperelliptic Lefschetz fibrations with only nonseparating vanishing cycles, $\lambda_f = 4 - 4/g$.

In this section, we investigate the slopes of Lefschetz fibrations with $b_2^+ = 1$. By combining the results of [Stipsicz 1999; 2002] and [Li 2000], we can show that Lefschetz fibrations with $b_2^+ = 1$ satisfy the slope inequality. Stipsicz showed that if $X \to S^2$ is a genus-$g$ Lefschetz fibration over $S^2$ with $b_2^+(X) = 1$ and $X$ is not diffeomorphic to the blow-up of a ruled surface (that is, diffeomorphic to an $S^2$-bundle over $\Sigma_k$), then $b_1(X) \in \{0, 2\}$ and $e \geq 0$ (see [Stipsicz 1999, Corollary 3.3

and 3.5]). In particular, if $X$ is the blow-up of an $S^2$-bundle over $\Sigma_k$, then $k \le g/2$ (see [Li 2000, Proposition 4.4]). Then, we obtain the following result.

**Theorem 5.1.** *Let $g \ge 2$ and let $f : X \to S^2$ be a genus-$g$ Lefschetz fibration with $b_2^+(X) = 1$.*

(1) *If $X$ is not diffeomorphic to the blow-up of a ruled surface, then*

(i) $$4 - 4/g \le \lambda_f \le 8 \quad \text{for } b_1(X) = 0,$$

(ii) $$4 \le \lambda_f \le 8 \quad \text{for } b_1(X) = 2.$$

(2) *If $X$ is diffeomorphic to the blow-up of an $S^2$-bundle over $\Sigma_k$, then*

$$4 + 4(k-1)/(g-k) \le \lambda_f \le 8,$$

*and the lower bound is sharp.*

An improvement of the previous result was suggested by the referee.

*Proof of Theorem 5.1.* For a genus-$g$ Lefschetz fibration, a regular fiber has zero self-intersection. Since the intersection form is nondegenerate, it follows that $b_2^{\pm} \ge 1$. Let $f : X \to S^2$ be a nontrivial genus-$g$ Lefschetz fibration with $b_2^+(X) = 1$. Note that $-4(g-1) \le K^2$, and so $4(g-1) \le K_f^2$ (see [Stipsicz 1999, Lemma 3.2]). Suppose that $X$ is not diffeomorphic to the blow-up of a ruled surface.

First, suppose that $b_1 = 0$. Since $b_2^+ = 1$ and $\chi_f = (\sigma + e)/4 + (g-1) = (b_2^+ - b_1 + 1)/2 + (g-1) = g$, we have $4(g-1)/g \le K_f^2/\chi_f = \lambda_f$. On the other hand, since $K^2 = 3\sigma + 2e = 5b_2^+ - b_2^- + 4 - 4b_1 = 9 - b_2^-$, by $b_2^- \ge 1$, we have $\lambda_f = K_f^2/\chi_f = \{9 - b_2^- + 8(g-1)\}/g \le 8$.

Next, suppose $b_1 = 2$. Then, $\chi_f = g - 1$. Therefore, by $4(g-1) \le K_f^2$, we have $4 \le \lambda_f$. Since $0 \le e = 2 - 2b_1 + b_2^+ + b_2^- = 2 - 4 + 1 + b_2^- = -1 + b_2^-$, we obtain $\lambda_f = \{1 - b_2^- + 8(g-1)\}/(g-1) \le 8$.

Finally, suppose that $X$ is the $m$-fold blow-up of an $S^2$-bundle over $\Sigma_k$. Let $Y$ be the $S^2$-bundle over $\Sigma_k$. Then, since $b_1(Y) = 2k$, $b_2^{\pm}(Y) = 1$ and $X = Y \# m\overline{\mathbb{CP}^2}$, we have $b_1(X) = 2k$, $b_2^+(X) = 1$, $b_2^-(X) = m+1$, $e(X) = 4 - 4k + m$ and $\sigma(X) = -m$. Hence, we have $\lambda_f = 8 - m/(g-k)$. From $m \ge 0$, $\lambda_f \le 8$. We will give lower bounds for $\lambda_f$. By Lemma 3.2 in [Stipsicz 2002], $4(2k-g) + m \le 4$. We have $\lambda_f \ge 4 + 4(k-1)/(g-k)$ from $\lambda_f = 8 - m/(g-k)$, $4(2k-g) + m \le 4$ and $0 \le k \le g/2$ [Li 2000, Proposition 4.4]. Fintushel and Stern [2004] showed that $(\Sigma_k \times S^2) \# 4m\overline{\mathbb{CP}^2}$ admits a genus-$(2k+m-1)$ Lefschetz fibration $f_{FS}$ over $S^2$. When $m = g - 2k + 1$, we find $b_2^+ = 1$ and that $\lambda_{f_{FS}} = 4 + 4(k-1)/(g-k)$. $\square$

**Remark 5.2.** If two Lefschetz fibrations $f_1$ and $f_2$ satisfy $\lambda_{f_1}, \lambda_{f_2} \ge 4 - 4/g$, then the (twisted) fiber sum $f_3$ of $f_1$ and $f_2$ satisfies $\lambda_{f_3} \ge 4 - 4/g$.

## Acknowledgments

The author would like to thank Hisaaki Endo for pointing out that the author's idea is effective in constructing a Lefschetz fibration which violates the slope inequality and for his encouragement. The author would like to express his gratitude to Kazuhiro Konno for his kind explanation of the history of the slope, and Mustafa Korkmaz and Kouichi Yasui for their comments on this paper. Finally, the referee's comments were especially helpful for improving the exposition.


## References


[Amorós et al. 2000] J. Amorós, F. Bogomolov, L. Katzarkov, and T. Pantev, "Symplectic Lefschetz fibrations with arbitrary fundamental groups", *J. Differential Geom.* **54**:3 (2000), 489–545. MR 2002g:57051 Zbl 1031.57021

[Arakelov 1971] S. J. Arakelov, "Families of algebraic curves with fixed degeneracies", *Izv. Akad. Nauk SSSR Ser. Mat.* **35** (1971), 1269–1293. In Russian; translated in *Math. USSR Izv.* **5** (1971), 1277–1302. MR 48 #298 Zbl 0238.14012

[Barth et al. 1984] W. Barth, C. Peters, and A. Van de Ven, *Compact complex surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **4**, Springer, Berlin, 1984. MR 86c:32026 Zbl 0718.14023

[Beauville 1979] A. Beauville, "L'application canonique pour les surfaces de type général", *Invent. Math.* **55**:2 (1979), 121–140. MR 81m:14025 Zbl 0403.14006

[Catanese 1983] F. Catanese, "Moduli of surfaces of general type", pp. 90–112 in *Algebraic geometry: open problems* (Ravello, 1982), edited by C. Ciliberto et al., Lecture Notes in Math. **997**, Springer, Berlin, 1983. MR 84k:14026 Zbl 0517.14011

[Chen 1993] Z. J. Chen, "On the lower bound of the slope of a nonhyperelliptic fibration of genus 4", *Internat. J. Math.* **4**:3 (1993), 367–378. MR 94g:14012 Zbl 0816.14006

[Dehn 1938] M. Dehn, "Die Gruppe der Abbildungsklassen", *Acta Math.* **69**:1 (1938), 135–206. MR 1555438 Zbl 0019.25301

[Donaldson 1999] S. K. Donaldson, "Lefschetz pencils on symplectic manifolds", *J. Differential Geom.* **53**:2 (1999), 205–236. MR 2002g:53154 Zbl 1040.53094

[Endo 2000] H. Endo, "Meyer's signature cocycle and hyperelliptic fibrations", *Math. Ann.* **316**:2 (2000), 237–257. MR 2001b:57047 Zbl 0948.57013

[Endo and Gurtas 2010] H. Endo and Y. Z. Gurtas, "Lantern relations and rational blowdowns", *Proc. Amer. Math. Soc.* **138**:3 (2010), 1131–1142. MR 2011a:57047 Zbl 1198.57017

[Endo and Nagami 2005] H. Endo and S. Nagami, "Signature of relations in mapping class groups and non-holomorphic Lefschetz fibrations", *Trans. Amer. Math. Soc.* **357**:8 (2005), 3179–3199. MR 2006g:57051 Zbl 1081.57023

[Endo et al. 2011] H. Endo, T. E. Mark, and J. Van Horn-Morris, "Monodromy substitutions and rational blowdowns", *J. Topol.* **4**:1 (2011), 227–253. MR 2012b:57051 Zbl 1234.57037

[Fintushel and Stern 1997] R. Fintushel and R. J. Stern, "Rational blowdowns of smooth 4-manifolds", *J. Differential Geom.* **46**:2 (1997), 181–235. MR 98j:57047 Zbl 0896.57022

[Fintushel and Stern 1998] R. Fintushel and R. J. Stern, "Constructions of smooth 4-manifolds", pp. 443–452 in *Proceedings of the International Congress of Mathematicians, II* (Berlin, 1998), edited by G. Fischer and U. Rehmann, 1998. MR 99g:57033 Zbl 0899.57012

[Fintushel and Stern 2004] R. Fintushel and R. J. Stern, "Families of simply connected 4-manifolds with the same Seiberg–Witten invariants", *Topology* **43**:6 (2004), 1449–1467. MR 2005d:57044 Zbl 1064.57036

[Gompf and Stipsicz 1999] R. E. Gompf and A. I. Stipsicz, 4*-manifolds and Kirby calculus*, Graduate Studies in Mathematics **20**, Amer. Math. Soc., Providence, RI, 1999. MR 2000h:57038 Zbl 0933.57020

[Hirose 2010] S. Hirose, "Presentations of periodic maps on oriented closed surfaces of genera up to 4", *Osaka J. Math.* **47**:2 (2010), 385–421. MR 2011h:57025 Zbl 1202.57022

[Johnson 1979] D. L. Johnson, "Homeomorphisms of a surface which act trivially on homology", *Proc. Amer. Math. Soc.* **75**:1 (1979), 119–125. MR 80h:57008 Zbl 0407.57003

[Kas 1980] A. Kas, "On the handlebody decomposition associated to a Lefschetz fibration", *Pacific J. Math.* **89**:1 (1980), 89–104. MR 82f:57012 Zbl 0457.14011

[Konno 1991] K. Konno, "A note on surfaces with pencils of nonhyperelliptic curves of genus 3", *Osaka J. Math.* **28**:3 (1991), 737–745. MR 93c:14035 Zbl 0766.14037

[Konno 1993] K. Konno, "Nonhyperelliptic fibrations of small genus and certain irregular canonical surfaces", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **20**:4 (1993), 575–595. MR 95b:14026 Zbl 0822.14009

[Konno 1996] K. Konno, "Even canonical surfaces with small $K^2$, III", *Nagoya Math. J.* **143** (1996), 1–11. MR 98a:14051 Zbl 0872.14033

[Korkmaz 2001] M. Korkmaz, "Noncomplex smooth 4-manifolds with Lefschetz fibrations", *Int. Math. Res. Not.* **2001**:3 (2001), 115–128. MR 2001m:57036 Zbl 0977.57020

[Li 2000] T.-J. Li, "Symplectic Parshin–Arakelov inequality", *Int. Math. Res. Not.* **2000**:18 (2000), 941–954. MR 2001i:57038 Zbl 0961.57022

[Manetti 2003] M. Manetti, "Surfaces of Albanese general type and the Severi conjecture", *Math. Nachr.* **261-262**:1 (2003), 105–122. MR 2004m:14081 Zbl 1044.14017

[Matsumoto 1983] Y. Matsumoto, "On 4-manifolds fibered by tori, II", *Proc. Japan Acad. Ser. A Math. Sci.* **59**:3 (1983), 100–103. MR 84j:57010b Zbl 0532.55020

[Matsumoto 1996] Y. Matsumoto, "Lefschetz fibrations of genus two: a topological approach", pp. 123–148 in *Topology and Teichmüller spaces* (Katinkulta, 1995), edited by S. Kojima et al., World Scientific, 1996. MR 2000h:14038 Zbl 0921.57006

[Miyachi and Shiga 2011] H. Miyachi and H. Shiga, "Holonomies and the slope inequality of Lefschetz fibrations", *Proc. Amer. Math. Soc.* **139**:4 (2011), 1299–1307. MR 2012a:57036 Zbl 1226.14015

[Ozbagci 2002] B. Ozbagci, "Signatures of Lefschetz fibrations", *Pacific J. Math.* **202**:1 (2002), 99–118. MR 2002k:57066 Zbl 1053.57013

[Ozbagci and Stipsicz 2000] B. Ozbagci and A. I. Stipsicz, "Noncomplex smooth 4-manifolds with genus-2 Lefschetz fibrations", *Proc. Amer. Math. Soc.* **128**:10 (2000), 3125–3128. MR 2000m:57036 Zbl 0951.57015

[Pardini 2005] R. Pardini, "The Severi inequality $K^2 \geq 4\chi$ for surfaces of maximal Albanese dimension", *Invent. Math.* **159**:3 (2005), 669–672. MR 2006a:14065 Zbl 1082.14041

[Reid 1979] M. Reid, "$\pi_1$ for surfaces with small $K^2$", pp. 534–544 in *Algebraic geometry* (Copenhagen, 1978), edited by K. Lønsted, Lecture Notes in Math. **732**, Springer, Berlin, 1979. MR 81j:14018 Zbl 0423.14021

[Severi 1932] F. Severi, "La serie canonica e la teoria delle serie principali di gruppi di punti sopra una superficie algebrica", *Comment. Math. Helv.* **4**:1 (1932), 268–326. MR 1509461 Zbl 0005.17602

[Smith 1999] I. Smith, "Lefschetz fibrations and the Hodge bundle", *Geom. Topol.* **3** (1999), 211–233. MR 2000j:57059 Zbl 0929.53047

[Smith 2001] I. Smith, "Lefschetz pencils and divisors in moduli space", *Geom. Topol.* **5** (2001), 579–608. MR 2002f:57056 Zbl 1066.57030

[Stipsicz 1999] A. I. Stipsicz, "On the number of vanishing cycles in Lefschetz fibrations", *Math. Res. Lett.* **6**:3-4 (1999), 449–456. MR 2000g:57046 Zbl 0955.57026

[Stipsicz 2002] A. I. Stipsicz, "Singular fibers in Lefschetz fibrations on manifolds with $b_2^+ = 1$", *Topology Appl.* **117**:1 (2002), 9–21. MR 2002j:57048 Zbl 1007.53059

[Usher 2006] M. Usher, "Minimality and symplectic sums", *Int. Math. Res. Not.* **2006**:16 (2006), Art. ID 49857. MR 2007h:53139 Zbl 1110.57017

[Xiao 1987] G. Xiao, "Fibered algebraic surfaces with low slope", *Math. Ann.* **276**:3 (1987), 449–466. MR 88a:14046 Zbl 0596.14028

Received November 1, 2012. Revised May 19, 2013.

NAOYUKI MONDEN
DEPARTMENT OF MATHEMATICS
KYOTO UNIVERSITY
GRADUATE SCHOOL OF SCIENCE
KYOTO
JAPAN
n-monden@math.kyoto-u.ac.jp

# Guidelines for Authors

Authors may submit manuscripts at msp.berkeley.edu/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

# PACIFIC JOURNAL OF MATHEMATICS

Volume 267 No. 1 January 2014

---